

A FORMAL GRAMMAR OF SPANISH
FOR PHRASE-LEVEL ANALYSIS
APPLIED TO INFORMATION RETRIEVAL

Lalia Collection
Series Maior
Nº 14

María Paula Santalla del Río

A FORMAL GRAMMAR OF SPANISH
FOR PHRASE-LEVEL ANALYSIS
APPLIED TO INFORMATION RETRIEVAL

To my parents

Contents

Contents	7
Acknowledgements	19
1 Introduction	21
1.1 The AGFL Formalism	24
1.1.1 First Level. Syntax Rules	25
1.1.2 Second Level. Affix Rules	28
1.1.3 <i>Grammar, Metagrammar and Lexicon</i>	29
1.2 A Model of Description	30
1.3 Coordination and Subordination	34
1.4 Levels of Formalization	38
1.4.1 First Level of Formalization. General Frames.	39
1.4.2 Second Level of Formalization. Coverage.	40
1.4.3 Third Level of Formalization. Resolution of Ambiguity. . .	41
1.5 Summary	42
2 VP	45
2.1 VP Description	45
2.1.1 Relevant Information for the Contextualization of VPs . . .	46
2.1.2 VPs not Containing Non-passive Semiauxiliaries	52
2.1.2.1 Elements of VPs that do not include functional clitic pronouns	53
2.1.2.2 Elements of VPs that include one functional clitic pronoun	55

2.1.2.3	Elements of VPs that include two functional clitic pronouns	58
2.1.3	Non-passive Semiauxiliaries: Periphrastic Groups	59
2.1.4	VPs	61
2.1.4.1	Elements of VPs that do not include functional clitic pronouns	61
2.1.4.2	Elements of VPs that include one functional clitic pronoun	64
2.1.4.3	Elements of VPs that include two functional clitic pronouns	67
2.1.5	Other Questions about VPs	67
2.2	AGFL Formalization of the VP	70
2.2.1	The Lexicon	71
2.2.2	Affix Rules, Terminals and Non-terminals.	74
2.2.2.1	Clitic Pronoun Position	74
2.2.2.2	Clitization	75
2.2.2.3	Mood and Tense	83
2.2.2.4	Person, Gender, Number and Case	87
2.2.2.5	Verb type, Voice and Preptype	88
2.2.2.6	Assignment of Verb Type and Voice Values to Verb Forms in the Lexicon. Verb Scheme Clusters Generation	97
2.2.3	The Grammar	109
2.2.3.1	Predicate Function: PREDICATE , NON FINITE PREDICATE and PARTICIPLE PREDICATE	113
2.2.3.2	Verb Phrase <i>Predicates</i>	121
2.2.3.2.1	case position in FCIPr	121
2.2.3.2.2	verb type priorities	122
2.2.3.2.3	compound tense	123
2.2.3.2.4	two persons values	123
2.2.3.2.5	three persons values	124
2.2.3.2.6	verb one CIPr	124

2.2.3.2.7	impersonal verb one ClPr	125
2.2.3.2.8	verb two ClPrs	125
2.2.3.2.9	negative imperative person number	126
2.2.3.3	Clitic Pronouns	126
2.2.3.4	Verb non-terminals	127
2.2.3.5	Non-Periphrastic VPs: Verb Phrase, Non Finite Verb Phrase and Participle Verb Phrase	
	Phrase	136
2.2.3.5.1	Verb Phrase/5	137
2.2.3.5.2	Verb Phrase/9	139
2.2.3.5.3	Verb Phrase/13	142
2.2.3.5.4	Non Finite Verb Phrase/3	143
2.2.3.5.5	Non Finite Verb Phrase/7	144
2.2.3.5.6	Non Finite Verb Phrase/11	147
2.2.3.5.7	Participle Verb Phrase/5	149
2.2.3.6	Periphrastic Groups: Periphrasis	149
2.2.3.6.1	Proper Voice Periphrastic Groups	149
2.2.3.6.1.1	ACT voice Periphrasis/3,5	149
2.2.3.6.1.2	ACT_IMP voice Periphrasis/3,5	151
2.2.3.6.1.3	MIDD voice Periphrasis/3,7,5	152
2.2.3.6.1.4	MIDD_IMP voice Periphrasis/3,5	155
2.2.3.6.2	<i>Improper_voice</i> Voice Periphrastic Groups with only Non-functional Clitic Pronouns	156
2.2.3.6.2.1	ip_MIDD voice Periphrasis/7,5	156
2.2.3.6.2.2	ip_MIDD_IMP voice Periphrasis/7,5	157
2.2.3.6.3	<i>Improper_voice</i> Voice Periphrastic Groups with One Functional Clitic Pronoun	158
2.2.3.6.3.1	ip_ACT voice Periphrasis/7,9	159
2.2.3.6.3.2	ip_ACT_IMP voice Periphrasis/7,9	160

2.2.3.6.3.3	ip_MIDD voice Periphrasis/11,9	163
2.2.3.6.3.4	ip_MIDD_IMP voice Periphrasis/11,9	164
2.2.3.6.4	<i>Improper_voice</i> Voice Periphrastic Groups with Two Functional Clitic Pronouns . .	166
2.2.3.6.4.1	ip_ACT voice Periphrasis/11,13	166
2.2.3.6.4.2	ip_ACT_IMP voice Periphrasis/11,13	167
2.2.3.7	Periphrastic VPs: Periphrastic Verb Phrase and Periphrastic Non Finite Verb Phrase . .	170
2.2.3.7.1	Periphrastic Verb Phrase/5	171
2.2.3.7.2	Periphrastic Verb Phrase/9	174
2.2.3.7.3	Periphrastic Verb Phrase/13	177
2.2.3.7.4	Periphrastic Non Finite Verb Phrase/3	179
2.2.3.7.5	Periphrastic Non Finite Verb Phrase/7	179
2.2.3.7.6	Periphrastic Non Finite Verb Phrase/11	180
3	NP, PrP, AvP and AjP	183
3.1	NP, PrP, AvP and AjP Description	183
3.1.1	Noun, Pronoun, Adverb and Adjective Phrase	186
3.1.2	Addition of Modifiers	189
3.1.2.1	Addition of Modifiers in Noun Phrase	190
3.1.2.2	Addition of Modifiers in Pronoun Phrase	198
3.1.2.3	Addition of Modifiers in Adverb Phrase	200
3.1.2.4	Addition of Modifiers in Adjective Phrase	202
3.1.3	Addition of Determiners	204
3.1.3.1	Addition of Determiners in Noun Phrase	205
3.1.3.2	Addition of Determiners in Pronoun Phrase	208

3.1.4	Modifiers	209
3.1.4.1	Modifiers Postponed to the Nucleus	209
3.1.4.1.1	Adjective Phrase	209
3.1.4.1.2	Prepositional Modifiers	209
3.1.4.1.3	Relative Clause	210
3.1.4.1.4	Apposition	212
3.1.4.1.5	Determiner	212
3.1.4.2	Modifiers Preceding the Nucleus	212
3.1.4.2.1	Adjective	212
3.1.4.2.2	Degree	212
3.1.4.2.3	Noun Phrase	216
3.1.5	Nominalization	216
3.1.6	Other constituents of the Clause	217
3.1.7	Prepositional Constituents	219
3.2	NP, PrP, AvP and AjP AGFL Formalization	220
3.2.1	The Lexicon	221
3.2.2	Affix Rules, Terminals and Non-terminals	224
3.2.2.1	Noun Type	224
3.2.2.2	Pronoun and Determiner Type and Subtype	224
3.2.2.3	Determiner Level	229
3.2.2.4	Definiteness	230
3.2.2.5	Adverb Type and Subtype	230
3.2.2.6	Adverb Position	232
3.2.2.7	Adjective Type	232
3.2.2.8	Position	232
3.2.2.9	Degree, Second Term Type and Type of (NoEqual) Comparisons	233
3.2.2.10	Person, Gender, Number and Case	235
3.2.2.11	Modifier Category and Lexical Category	236
3.2.2.12	Relative Clause Function	240
3.2.2.13	Conjunction and Preposition Type	241
3.2.2.14	Clause Type and Question Type	242
3.2.3	The Grammar	243

3.2.3.1	Noun, Pronoun, Adverb and Adjective	
	Non-terminals	244
3.2.3.2	Addition of Modifiers	248
3.2.3.2.1	From Non-terminals for Lexical Categories to Target Non-terminals for Phrase Categories	248
3.2.3.2.2	An AGFL General Frame for the Addition of Modifiers	249
3.2.3.2.3	More Details about the AGFL General Frame Dealing with Addition of Modifiers	252
3.2.3.2.4	Addition of Modifiers in Noun Phrase . .	255
3.2.3.2.4.1	HEAD Noun Phrase/9, HIDDEN RECURSIVE Noun Phrase/9	256
3.2.3.2.4.2	Noun Phrase/9	256
3.2.3.2.4.3	Noun Phrase NUCLEUS/9 . .	264
3.2.3.2.5	Addition of Modifiers in Pronoun Phrase	265
3.2.3.2.5.1	HEAD Pronoun Phrase/8,9, HIDDEN RECURSIVE Pronoun Phrase/8,9	265
3.2.3.2.5.2	Pronoun Phrase/8,9	268
3.2.3.2.5.3	Pronoun Phrase NUCLEUS/8,9	272
3.2.3.2.6	Addition of Modifiers in Adverb Phrase .	273
3.2.3.2.6.1	HEAD Adverb Phrase/6, Adverb Phrase/6,7, HIDDEN RECURSIVE Adverb Phrase/7 .	273
3.2.3.2.6.2	HIDDEN Adverb Phrase/7 . .	274
3.2.3.2.6.3	Adverb Phrase NUCLEUS/7 .	279
3.2.3.2.7	Addition of Modifiers in Adjective Phrase	281

3.2.3.2.7.1	HEAD Adjective Phrase/6, Adjective Phrase/6,7, HIDDEN RECURSIVE Adjective Phrase/7	281
3.2.3.2.7.2	HIDDEN Adjective Phrase/7	282
3.2.3.2.7.3	Adjective Phrase NUCLE- US/7	285
3.2.3.3	Addition of Determiners	286
3.2.3.3.1	From Non-terminals for Lexical Categor- ies to Target Non-terminals for Phrase Cat- egories	286
3.2.3.3.2	An AGFL General Frame for the Addition of Determiners	287
3.2.3.3.3	DETERMINER Function	291
3.2.3.3.3.1	DETERMINER/6	291
3.2.3.3.3.2	DETERMINER AND MODI- FIER/9,11	291
3.2.3.3.3.3	HIDDEN RECURSIVE PostDe- terminer and Modifier Cate- gory/9	292
3.2.3.3.3.4	PostDeterminer and Modifier Category/9	293
3.2.3.3.3.5	PostDeterminer and Modifier Category NUCLEUS/9	294
3.2.3.3.4	Addition of Determiners in Noun Phrase	294
3.2.3.3.4.1	Nominal Phrase/8	295
3.2.3.3.4.2	Nominal Phrase I/12	303
3.2.3.3.4.3	Nominal Phrase II/12	303
3.2.3.3.4.4	Nominal Phrase NOMINAL I/8	304
3.2.3.3.4.5	Nominal Phrase NOMINAL II/11	304

3.2.3.3.4.6	Nominal Phrase NUCLEUS	
	I/11	304
3.2.3.3.5	Addition of Determiners in Pronoun	
	Phrase	304
3.2.3.3.5.1	Pronominal Phrase/9	304
3.2.3.3.5.2	Pronominal Phrase I/9	309
3.2.3.3.5.3	Pronominal Phrase NOMINAL	
	I/7	309
3.2.3.3.5.4	Pronominal Phrase NOMINAL	
	II/9	309
3.2.3.3.5.5	Pronominal Phrase NUCLEUS	
	I/9	310
3.2.3.4	Modifiers	310
3.2.3.4.1	Modifiers Postponed to the Nucleus . . .	310
3.2.3.4.1.1	Adjective Phrase MODI-	
	FIER/5	310
3.2.3.4.1.2	Prepositional Phrase MODI-	
	FIER/1	310
3.2.3.4.1.3	Relative Clause MODI-	
	FIER/1,4	311
3.2.3.4.1.4	Apposition MODIFIER/4,5 .	314
3.2.3.4.1.5	Determiner MODIFIER/6 . .	314
3.2.3.4.2	Modifiers Preceding the Nucleus	315
3.2.3.4.2.1	Adjective MODIFIER/5 . . .	315
3.2.3.4.2.2	Degree MODIFIER/4, Compar-	
	ison Quantifier MODIFIER/4,5,	
	Factor MODIFIER/2, Second	
	Term MODIFIER/4,6,7	315
3.2.3.4.2.3	Noun MODIFIER	319
3.2.3.5	Nominalization	319
3.2.3.5.1	Nominalization/5	319
3.2.3.5.2	Nominalization NOMINAL I/4 . . .	320

3.2.3.5.3	HIDDEN FINAL Nominalized Phrase/6	320
3.2.3.5.4	Addition of Modifiers within Nominaliz- ation/5	321
3.2.3.5.4.1	HIDDEN RECURSIVE Nom- inalized Phrase/6	321
3.2.3.5.4.2	Nominalized Phrase/6	321
3.2.3.5.4.3	Nominalized Phrase NUCLE- US/6	322
3.2.3.5.5	Addition of Determiners within Nomin- alization/5	322
3.2.3.5.5.1	Nominalization I/5	322
3.2.3.5.5.2	Nominalization/5 , 1 st level of re- cursion	322
3.2.3.5.5.3	Nominalization NOMINAL II/5	323
3.2.3.5.6	Nominalization/6	323
3.2.3.5.7	Nominalization NOMINAL I/5	324
3.2.3.5.8	HIDDEN FINAL Nominalized Phrase/6	324
3.2.3.6	Other Constituents of the Clause. The Clause	324
3.2.3.6.1	<i>That</i> -clauses	325
3.2.3.6.2	Infinitive Clauses	328
3.2.3.6.3	Participle Clauses	328
3.2.3.6.4	Interrogative Clauses	329
3.2.3.6.4.1	<i>Yes/No</i> Questions	329
3.2.3.6.4.2	<i>Wh</i> -Questions	329
3.2.3.6.5	Relative Clauses	331
3.2.3.7	Prepositional Constituents	335
3.2.3.7.1	Prepositional Noun Phrase/10	335
3.2.3.7.2	Prepositional Pronoun Phrase/9,10	336
3.2.3.7.3	Prepositional Adverb Phrase/6,7,8	336

3.2.3.7.3.1	Prepositional Adverb Phrase/7	337
3.2.3.7.3.2	Prepositional Adverb Phrase/6	338
3.2.3.7.3.3	Prepositional Adverb Phrase/8	338
3.2.3.7.4	Prepositional Nominal Phrase/9	340
3.2.3.7.5	Prepositional Pronominal Phrase/10	341
3.2.3.7.6	Prepositional Nominalization/6,7	342
3.2.3.7.7	Prepositional Clause/3,4	344
3.2.3.7.8	Prepositional Interrogative Clause/4	344
3.2.3.7.9	Prepositional Relative Clause/5	345
4	Ambiguity	347
4.1	Full Sentence Analysis	348
4.1.1	Largest Constituents	350
4.1.1.1	verb type priorities	350
4.1.1.2	determiner better than pronoun	350
4.1.1.3	Relative Clauses	353
4.1.1.4	Prepositional MODIFIER	356
4.1.2	Ambiguity within Constituents	356
4.1.2.1	<i>el que</i>	356
4.1.2.2	(Quantifier) + (<i>más/menos</i>) + Adjective + Noun Quantifier + <i>más/menos</i>	357
4.1.2.3	Internal Structure of Modifiers	358
4.1.3	Lexical Ambiguity	359
4.2	Phrase Analysis	361
4.2.1	Largest Phrases	365
4.2.2	Optimization	365
4.2.2.1	Removed Ambiguities	366
4.2.2.1.1	VPs	366

4.2.2.1.2	<i>el que</i>	370
4.2.2.1.3	(Quantifier) + (<i>más/menos</i>) + Adjective + Noun. Quantifier + <i>más/menos</i>	370
4.2.2.1.4	Relative Clauses and Prepositional Relative Clauses	371
4.2.2.1.5	Noun/Pronoun Phrase person	372
4.2.2.2	Ambiguities Added Penalties	373
4.2.2.2.1	Internal Structure of Modifiers	373
4.2.2.2.2	Plural Proper Names	374
4.2.2.3	Lexical Ambiguity	374
5	Summary	375
A	Non-passive Semiauxiliaries	379
B	One Clause Example from BDS	381
C	Verb Scheme Cluster of <i>abandonar</i>	385
D	Pronouns	393
E	Adverbs	397
F	Determiners	401
G	Examples of Analysis	405
G.1	Examples from Section 3.1	405
G.2	Full Sentence Analysis	415
G.3	Phrase Analysis	418
	Bibliography	433

Acknowledgements

I would like to thank all those who have contributed to the work described in this book. Especial thanks are due to

Guillermo Rojo, for his guidance and encouragement, for having introduced me in the world of Corpus Linguistics and computers;

my collaborators in the DoRo project, Concepción Álvarez, Pilar Alvariño, Adelaida Gil and Susana Sotelo, for their dedication to the project which allowed me to concentrate on the development of the module described in this work;

all those who contributed to the development of the *Base de Datos Sintácticos*, professors and research assistants of the Department of *Lengua Española, Teoría de la Literatura y Lingüística General*, for their dedication to the project which allowed me to base my work on more and more improved data;

Susana Sotelo and Verónica Santalla, for their patience assisting me with computers;

Stefano Ninni, for his unfailing and continuous support;

the Foundation Caja de Madrid, which awarded me a grant between 1997 and 1999.

Chapter 1

Introduction

On the one hand, the work described in this book partially reports on the DoRo¹ research project, which gave the opportunity and set the main requirements for its concrete development. On the other hand, we will see that the manner in which verb government is dealt with in this work is directly related to the resource delivered by the Spanish research project *Estructura de la cláusula. Diccionario de Construcción y Régimen del Español Actual*², the so-called *Base de Datos Sintácticos*—henceforth BDS—. Moreover, even though the internal structure of phrases is not detailed in BDS, the grammar described in this book is, to a large extent, based on the descriptive model underlying syntactic analyses of clauses encoded in this database.

The first project mentioned above, the DoRo project, aims at improving traditional results of Information Retrieval applications, specifically Document Routing. Document Routing is the automatic routing of incoming text documents based on content analysis of text by a computer and knowledge of the characteristics of the candidate destinies. The innovative idea in the case of DoRo is

¹The DoRo project (1997–1999) was partly financed within the framework of the 6th Call of the Information Technologies Research and Technology Development (RTD) ESPRIT programme, managed by the European Commission Directorate General III (DGIII) and included in the Fourth Framework Programme (1994–1998) of European Union.

²Directed by Guillermo Rojo and supported by the Dirección Xeral de Educación e Ordenación Universitaria of the Consellería de Educación of the Xunta de Galicia (1989–1991) and the Dirección General de Investigación Científica y Técnica of the Ministerio de Educación (1991–1994). After, this project was continued and extended by new projects, *Hacia el Análisis Automático del Régimen Verbal en Español* and *El Análisis Automático del Régimen Verbal en Español*, directed also by Guillermo Rojo and supported by the Dirección Xeral de Universidades e Investigación of the Consellería de Educación of the Xunta de Galicia (1997–1999).

to achieve this improvement of results by performing a deep linguistic analysis, non-interpretative, of involved texts. This analysis should be based on large linguistic resources (formal grammars and lexicons) developed in order to be able to (i) precisely identify useful keywords and phrases that characterize the input document and are relevant for the classification, (ii) normalize them to avoid the dispersion of data caused by linguistic variation. That is, the main concern within the DoRo project is to make use of NLP techniques to prove the intuition that NLP is a possible solution for achieving retrieval effectiveness, a solution that might be worthwhile to take into consideration also for realistic and large-scale applications, such as document routing in a business environment that receives thousands of documents per day.

The role of the research group of the University of Santiago de Compostela involved in the DoRo project consisted of the development of the linguistic resources needed for Spanish: a formal grammar and, both, a domain specific lexicon and a general one, all of them written in the AGFL formalism (see Section 1.1), required by the AGFL system for parser generation. Within the context of the DoRo project, both the grammar and the lexicon should be strongly application-oriented, focusing in the processing of non-unrestricted text, text of the domain, for the robust and unambiguous syntactic identification of linguistic units with high probabilities of being relevant from the semantic point of view, that is, of being relevant for the classification of documents.

BDS³, on the other hand, is the result of the manual analysis —having in mind constitutive and functional analysis principles, see Section 1.2— of the syntactic context of 160.000 verb forms approximately, which belong to the contemporary part of the *Archivo de Textos Hispánicos de la Universidad de Santiago de Compostela* (ARTHUS). This corpus is constituted by million and a half words of texts taken from all the Hispanic countries, which include oral samples as well as novels, press and theater, published between 1980 and 1990. BDS is thus the result of a classical corpus-linguistics research project, that is to say, a research project that aims at creating a linguistic database which should contain a wealth of information (in this case, primarily about verb government) in the form of detailed analyses

³Section 2.2.2.6 and Appendix B respectively contain a more detailed description and an example of BDS. BDS is also described in the following works: García-Miguel [25] and Rojo [65], [66] and [67].

that might be used in the study of actual language use.

In BDS, the appearance of each verb from the corpus has been individually registered with detailed information about the syntactic configuration of the linguistic sequence governed by the verb. First of all, general information on the mentioned linguistic sequence: clause type, clause function, voice, modality, inflection of the verb form, number of arguments and order of them. Secondly, detailed information about each of the arguments is described for the sequence in question. Together with certain syntactic characteristic features of each of the arguments, the type of structural unit is specified by means of very detailed keys. Finally, certain semantic features that are syntactically relevant are also encoded—determination, animation and countability.

The availability of BDS, together with our research tradition, led since the very beginning the development of the linguistic resources that had to be delivered within the DoRo project. From this second perspective, the DoRo grammar—that we called AVALON—and lexicon were fed with BDS data and they were developed so as to generate a parser capable of producing automatically the type of analysis manually encoded in BDS. The AVALON parser constituted, and still constitutes, a very ambitious—and long-term—objective, both with respect to coverage and detail of the analysis. For the purposes of DoRo, nevertheless, somewhat partial goals had to be identified which could fulfill the requirements of the project and be achieved in the available period of time (two and a half years). This means that we had to find a balance between the classical corpus linguistics orientation towards the production of detailed and rich linguistic databases for language study, including where necessary manual intervention, and the robust, efficient and fully automatic orientation necessary to satisfy the requirements of a real environment application. This was done on the basis of the following principle: on the firm conviction that phrases already constitute main content units relevant for IR, robust phrase-level analysis and phrase normalization were given priority, although, from the point of view of analysis, identification of phrases to the extent that they match linguistic sequences filling syntactic functional slots of the clause was the real short-term goal of the project, required by the reusability of the delivered resource, the phrase-level parser, as the first step of analysis within the—future—full-sentence AVALON parser.

This book describes part of the work⁴ developed in USC to achieve this goal within the DoRo project: having always in mind the duplicity of objectives referred above, it describes both robust and unambiguous identification of phrases, and necessary adaptations required for integrating this description within the AVALON parser. The duplicity of the objectives of the work is formally accounted for in the book as well: Chapters 2 and 3 formally describe phrase-level analysis independently—to the extent that this is possible—of the use made after of the grammar, while Chapter 4 is dedicated to problems posed by the insertion of the grammar described in larger environments, a full-sentence grammar and, especially, a Document Routing application as the one designed in the DoRo project.

1.1 The AGFL Formalism

The AGFL⁵ system, which is a collection of software systems for NLP, constituted within the DoRo project the tool for generating the parsers for the three languages involved in the project (Dutch, Greek and Spanish). The AGFL system is a parser generator, i. e. a system for generating parsers from grammars and lexica written according to the notational conventions that constitute the AGFL formalism.

AGFL⁶ is a formalism in which non-terminals of a primary context-free grammar are extended with so-called affixes, defined in a secondary context-free grammar. Both grammars are rolled into one by the parser generator, and this gives the resulting AGFL parser minimally the power of a context-sensitive grammar.

⁴The modules for clause-level analysis and phrase normalization, as well as the process of collecting the lexicon of nouns and adjectives, are not described here. Information about these issues can be found in Álvarez *et al.* [6].

⁵The AGFL (*Affix Grammars over Finite Lattices*) system has been developed, and is still developed and maintained, by professor C.H.A. Koster and his collaborators in the Department of Software Engineering of the University of Nijmegen.

⁶The description of the formalism has been simplified to account exclusively for the resources used within the grammar described in Chapters 2, 3 and 4. Additional resources of the formalism that have not been necessary for the grammar described (commit operator, transduction facilities, for instance) have not been included. This simplification has the advantage of restricting the description to the more stable features of the AGFL formalism, which has been—it is still—evolving during the whole DoRo project (documentation about the system is, in fact, not fully updated at the moment). Our description, and the current parser, refer to 1.7.52 version of the AGFL system. Latest versions have introduced some new functionalities not relevant within the boundaries of our description. Also, certain old functionalities, as well as certain notational conventions, are being modified in these new, not yet documented, versions. None of them, however, will entail changes in the grammar that cannot be automatically done. Most of this Section, on the other hand, has been, even literally, taken from previous descriptions of the formalism delivered by its developers, especially from Dekkers *et al.* [14] and Derksen [15].

Therefore grammars written in AGFL are affix or two-level grammars: the first or basic level consists of context-free syntax rules, a number of rewrite rules for non-terminals, extended with affixes, which are meta-variables added to the non-terminals of the basic context-free grammar. The second level defines the domains of these affixes.

1.1.1 First Level. Syntax Rules

In the first or basic level, a rule has the following form:

```
left-hand side:
  right-hand side.
```

A colon is used to separate the two main parts of the rule, the one placed before the colon rewrites as the one placed after it. The rule is delimited by a period. In the left-hand side of the rule only one element is possible, which must be a non-terminal, see below. In the right-hand side various elements may be related by means of two different operators: a comma, which indicates concatenation of elements —called *members*—, and a semi-colon, which indicates disjunction of elements —called *alternatives*. From an internal point of view, these might also be regarded as constituted by one or more members.

Various types of members are possible:

Non-terminals A non-terminal is identified by its name and an *arity*, the number of affix parameters added to the non-terminal name. Non-terminals have the following form⁷:

```
arity 0: non-terminal A
arity 1: non-terminal A ( attribute a )
arity 2: non-terminal A ( attribute a, attribute b )
arity 3: non-terminal A ( attribute a, attribute b, attribute c )
...
```

Within formal grammars, non-terminals mostly refer to functions and categories whose names we usually want to show as labels for the nodes of the

⁷See Section 1.1.2 for a better understanding of the non-terminal notation and other features referred to in this Section. In this exposition, we use also the notation **non-terminal/n**, where *n* is the *arity* of the non-terminal. According to this notation, for instance, the non-terminals in the example might be also represented as **non-terminal A/0**, **non-terminal A/1**, **non-terminal A/2** and **non-terminal A/3**. This is, in fact, the notation used within the **DEFINES** module of the formalism introduced below.

parse trees that result from successful analyses. Not always so, however, because, as we will see in subsequent Sections and Chapters, we occasionally use a different kind of non-terminals, somewhat duplicating the former, which, as they are necessary not to describe the structure of the language but to produce an effect on certain behaviours of the formalism, we do not want to have in the output. The AGFL formalism provides a functionality which permits to hyde these non-terminals, which are then called *invisible non-terminals*. This functionality is activated by the **DEFINES** module, in which all and only the non-terminals that the grammar writer wants to show in the parse trees of the output have to be listed. This means that, according to the following **DEFINES** list, for instance, non-terminal **non-terminal A/1** of the example above could not be showed in the parse trees of the output:

```
DEFINES
non-terminal A,
non-terminal A/2,
non-terminal A/3.
```

Terminals Terminals are quoted sequences of characters that represent the words of the language:

```
"terminal a"
```

Guards Guards are structures that restrict affix variables to one or more literal or terminal affix values:

```
{ attribute a :: VALUE A }
```

Groups Groups are series of alternatives enclosed between brackets and placed in a member position. They might be seen as shorthands for calls of anonymous non-terminals which have a definition consisting of the given alternatives:

```
( non-terminal A;
  non-terminal B;
  "terminal a";
  { attribute a :: VALUE A } )
```

Wildcards Wildcards are possibly open classes of terminals, the structure of which is described by a regular expression.

```
$SKIP ( "regular expression" )
$MATCH ( "regular expression" )
```

All the members of concatenations —not simultaneously, however— can be characterized as optional by surrounding them with square brackets. In this respect, it is important to remark that the analyses in which the optional members are realized will be explored by the parser before the analyses in which such members are not realized.

Predicates, finally, constitute a especial type of rules. Used to impose restrictions on or to effect the generation or analysis of affix values, they are non-productive rules, so they do not correspond to any segment of input. *Predicates* have the form of rewrite rules whose right-hand side is constituted by alternatives that consist exclusively of guards. For instance, the following rule is a *predicate*:

```
non-terminal A ( attribute a ):
{ attribute a :: VALUE A };
{ attribute a :: VALUE B }.
```

According to this predicate, in all rules that call upon **non-terminal A**, *attribute a* will take either the value **VALUE A** or the value **VALUE B**. *Predicates* can alternatively be expressed by means of series of rules with empty right-hand sides, the predicate above, for instance, would then look as follows:

```
non-terminal A ( VALUE A ): .
non-terminal A ( VALUE B ): .
```

By default, the parsers generated from AGFL grammars and lexicons return for the input sequences all the possible analyses according to the grammar, in the order in which they were found. This behaviour, however, can be modified. On the one hand, we can force the parser to return only a certain number or parse trees and, on the other, the formalism allows for any alternative to be given an indeterminate number of penalties. Once, no matter where in the grammar, some alternative has been penalized, the behaviour of the parser changes in a way so that it starts to return all the analyses which have been found with the lowest number of penalties, in the order in which they were found. Penalties are included as additional members of the alternatives and they have the following form:

`$PENALTY (n)`

By default, the parser also uses *word analysis*, that is, lexical matching is only possible for sequences of characters between spaces. In spite of this, it is possible to use also so-called *prefix parsing*, so that lexical matching for sequences of characters not contained between spaces is allowed too.

1.1.2 Second Level. Affix Rules

Values of attributes are specified in the second level of the formalism, with the following type of rules:

```
attribute a :: VALUE A; VALUE B;... attribute b.  
attribute b :: VALUE C; VALUE D;... VALUE T.
```

Although recursion is not possible, attribute non-terminals can be hierarchically organized in a such way that affix non-terminals can be included as rewrite alternatives of other affix non-terminals different from themselves (this is the case for *attribute b* in previous rules). The collection of terminal values that, whether directly or indirectly, are a possible literal rewrite alternatives for an affix non-terminal constitutes the *domain* of such affix non-terminal. None of the domains defined in an AGFL formal grammar can contain more than 32 different values.

In addition to this, another important remark about affixes is that, on the basis of the so-called *consistent restriction*, within the same syntax rule, all attributes with the same representation have to take the same value. In spite of this, indexes can be added to attributes so as to overcome this restriction by indicating that the involved affixes can actually take different values. Attributes can also be restricted to one or more terminal values in the syntax rules. The vertical bar “|” is used to separate the values when affixes can take more than one literal value.

```
non-terminal B ( attribute a, attribute b ):  
  non-terminal A ( attribute a, VALUE D|VALUE T, attribute c ),  
  non-terminal A ( attribute a, attribute b, attribute c1 ).
```

In this example, **non-terminal B/2** is constituted by two consecutive non-terminals **non-terminal A/3**. The value for *attribute a* must be the same for all of them. The value for *attribute b* must be the same for **non-terminal B** and second **non-terminal A**. For first **non-terminal A**, affix *attribute b* must

take either one of the values **VALUE D** or **VALUE T**, or both. The values for *attribute c* in the two non-terminals **non-terminal A** are independent of each other.

1.1.3 Grammar, Metagrammar and Lexicon

Together with the distinction between the *grammar*, that is, the first or basic level of description of an AGFL grammar, and the *metagrammar*, that is, the definition of affixes or the second level of description of an AGFL grammar, we will introduce in this Section another useful distinction that concerns solely the *grammar* in this more restricted sense⁸.

This new distinction, which proves to be especially useful when writing formal grammars for natural languages, separates a further restricted notion of the *grammar* from what we call the *lexicon*. The *grammar*, although it might also contain, depending on the grammar-writer preferences, rules that rewrite non-terminals into terminals, will be considered to be mostly constituted by rules that rewrite non-terminals into more non-terminals, whereas the *lexicon* will be considered to be strictly constituted by rules that rewrite non-terminals into terminals.

However, as can be inferred from these definitions, the boundaries of the distinction between the *grammar* in this sense and the *lexicon* are not clearcut. As a matter of fact, we can see that, to a certain extent, they are delimited in each case by the grammar writer. In this respect, our practice so far has been the inclusion in the *grammar*, that is to say in the **.gra** file, of the rules that rewrite non-terminals into terminals which realize things such as clitic pronouns, which can be attached to verb forms, and grammatical prefixes, infixes or suffixes, these, obviously, only in case the description of morphology is comprised by the formal grammar.

The notions of *grammar*, *metagrammar* and *lexicon* introduced here are also recognized by the AGFL system itself. This, in fact, allows for the grammar writer to create three different files with extensions **.def**, **.gra** and **.dat**, respectively containing the *metagrammar*, the *grammar* and the *lexicon*. According to this, the structure of Sections 2.2 and 3.2, which respectively account for the AGFL

⁸So far we have been using the term *grammar* to refer to the whole set of syntax and affix rules that constitute together an AGFL formal description.

formalization of phrases described in Chapters 2 and 3, will be organized in three subsections: the first one will describe the *lexicon* —.dat file—, the second one the *metagrammar* —.def file— and the third one the *grammar* —.gra file—.

1.2 A Model of Description

The grammatical model used in the description of the structure of Spanish in this exposition is based on constitutive and functional principles. That is, on the one hand, from the point of view of the Grammar, linguistic sequences, words in context, are identified as linguistic structures hierarchically organized in such a way that higher level structures are constituted by progressively lower level structures. On the other hand, structures integrated in higher level structures are assigned a syntactic function within them. These syntactic functions, together with grammatical classes of words constituting the central element of the structure, are the base for the identification of the type of unity (*category*) constituted. Categories and functions are systematically alternated in syntactic descriptions. A syntactic category rewrites always as one function or a set of functions, and a function rewrites as a category, this is repeated until the word level is arrived at.

Apart from words, two main syntactic structures will be taken into consideration in this exposition: *clauses*, organized around verbs and constituted by a PREDICATE function and, optionally, functions like SUBJECT, DIRECT OBJECT, etc., and *phrases*, organized around nouns, pronouns, adverbs and adjectives —also around certain syntactic structures in the case of nominalizations. Recursion is possible, so that clauses might contain clauses, and phrases might contain phrases or clauses. Clauses in which recursion is not applied are constituted by phrases or lexical categories, phrases in which recursion is not applied are constituted by lexical categories.

Syntactic categories may be expanded by means of various general mechanisms of expansion: subordination and coordination have been traditionally identified as such. By means of coordination and subordination syntactic categories can be expanded without changing their functional capabilities. If a word is expanded, a phrase is obtained, and a phrase is nothing more than the —possibly recursive— expansion of a word category —a structural one in the case of nominalizations.

Coordination, by means of different connectors, repeats the expanded structure one or more times. After this process, the resulting structure is a coordinated one, whose constituents have not been traditionally assigned a syntactic function, having all of them the same relevance. Subordination, on the other hand, adds specifiers to the syntactic category expanded. While the structural category expanded is assigned the syntactic function nucleus, the specifiers have been traditionally assigned the syntactic function modifier. Both possibilities of expanding a structural category can be combined, and we may have a coordinated structure that is, as a whole, modified, or a modified structure that is also coordinated with another structure, this in turn modified or not.

For this grammatical model the best graphical representation are parse trees, showing i) the constitutive and functional relationships between progressively simpler segments of the analysed linguistic sequence, ii) the syntactic categories of these segments (see Figure 1.1).

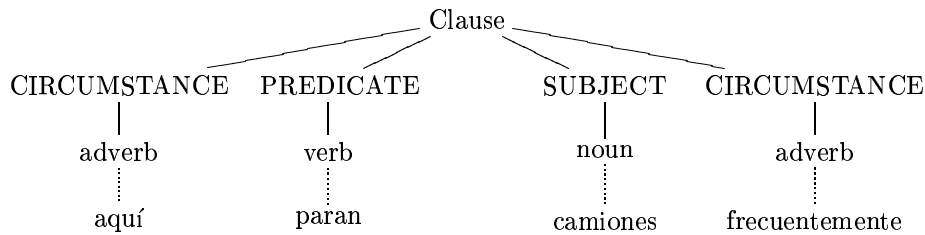


Figure 1.1: *Aquí paran camiones frecuentemente*

Looking at parse trees, the process of coordination creates two new levels of nodes: in one of them, the functions of the coordinated structure should be specified, in the other, the structural category name of the coordination —related to the structural category shared by all the coordinated structures— appears (see Figure 1.2).

The same holds for the process of subordination, two new levels of nodes are created also in this case: in one of them, nucleus and modifier functions are specified, in the other, the structural category name of the subordinated structure

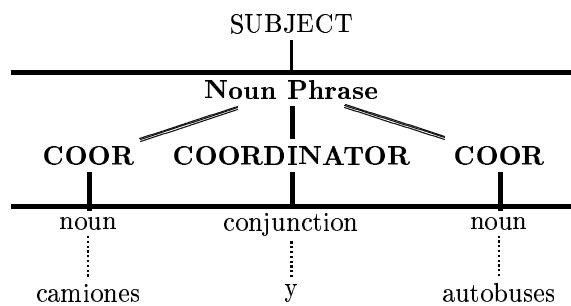


Figure 1.2: *camiones y autobuses*

—related to the structural category in the nucleus— appears (see Figure 1.3).

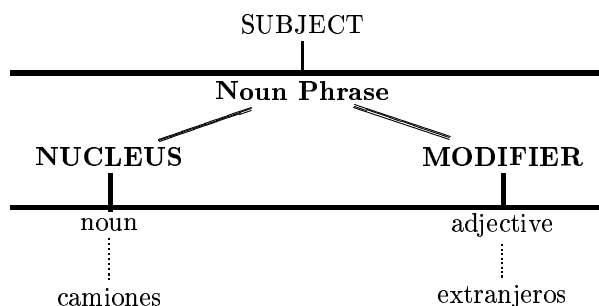


Figure 1.3: *camiones extranjeros*

Although coordination may add an **indeterminate** number of similar structures —probably because they have all the same syntactic relevance—, it seems that only **one** modifier can be added to any nucleus function of a structural category. That is, it seems that subordination, unlike coordination, is a binary process. This has consequences on the behaviour of these mechanisms with respect to recursion from an external point of view. Although coordination may be recursively applied and, at least theoretically, we may have different levels of coordinated structures, this is not frequent. On the contrary, recursive application of subordination is the most frequent case for subordinated structures, given the fact that the

addition of more than one specifier to a nucleus is very common (see Figure 1.4).

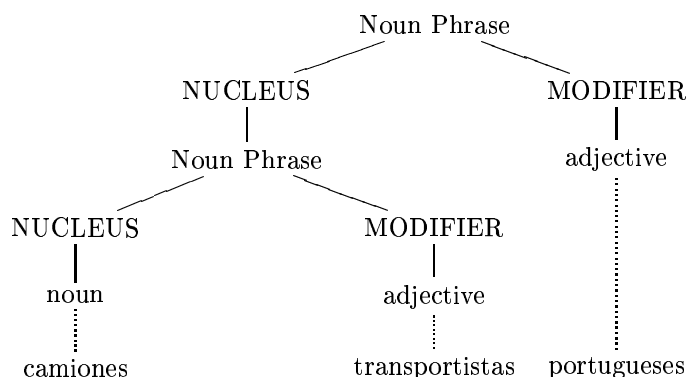


Figure 1.4: *camiones transportistas portugueses*

Verbs were not included above among the lexical categories that can be the base for the constitution of syntactic phrase categories. The combination of auxiliary and main verbs, and clitic pronouns, covered by what is considered here a verb phrase, behaves mostly as a morphological phenomenon, so, from an internal point of view, verb phrases are morphological, not syntactic, units. From an external point of view, however, verb phrases have an important syntactic relevance that lies on the fact that they constitute the PREDICATE function of clauses and therefore may have very important consequences on the syntactic information supplied by the main verb —subcategorization, in a very wide sense— for the organization of the clause. This must be carefully accounted for by any description of verb phrases that, as ours, aims at being integrated in a grammar for full-sentence analysis.

Within this framework, the purpose of this book is the formal description of segments of linguistic sequences dominated by nodes labelled as clause syntactic functions (PREDICATE, SUBJECT...) in syntactic parse trees, when no co-ordination has been applied. This is what we consider, from the point of view of coverage, the first layer for developing the AVALON parser. It concerns basically the description of various *phrases*: Verb Phrase, Noun Phrase, Pronoun Phrase, Adverb Phrase and Adjective Phrase, but it is also related to nominalization and other clause constituents that, on the basis of recursion, might occupy these syn-

tactic slots too. Chapter 2 describes Verb Phrases and Chapter 3 describes Noun, Pronoun, Adverb and Adjective Phrases, as well as nominalizations and clause constituents of clauses.

Even within these boundaries, certain phenomena could not be covered because of lack of the necessary data. This is the case for argumental behaviour of nouns and adjectives, whose integration in the grammar has been, however, foreseen and will be carried out in the near future. In addition to this, because of the very different nature of linguistic phenomena implied, the exhaustivity of the description has been a primary aim for Verb Phrase description, while in the case of Noun, Pronoun, Adverb and Adjective Phrases, the main concern has been the design of general AGFL formal frames where particular details or phenomena, perhaps left out of the description in first attempts that tried to account for larger phenomena, could be after easily integrated. This means that, in the case of non-verb Phrases, the structure and methodology of the grammar —based on subordination— itself have been given priority over coverage, which is, anyway, considerably wide.

From the point of view of linguistics, this is also congruent with the following situation: phrase level grammar and behaviour of closed categories of words have not been systematically described, either formally or not formally, for Spanish. Moreover, the available descriptions are based on general structures and definitions that can hardly be useful from the formal point of view, mainly when particular behaviours of closed categories of words have to be covered. This, together with certain requirements of the formalism —the definition of non-terminals—, leads us to give new names, not found in the linguistic tradition of Spanish, to certain, mainly functional, nodes of the trees (this is the case for various degree modifiers, see Section 3.1.4.2.2).

1.3 Coordination and Subordination: AGFL General Frames

The AGFL general frames for the formalization of coordination and subordination are the consequence of the opposite features that characterize each of these mechanisms:

- Addition of **equivalent** elements in the case of coordination, which means

that all constituents of a coordination are assigned the *same* syntactic function, as opposed to addition of one specifier to one central element, which means that one **MODIFIER** is added to a **NUCLEUS**.

- Addition of an **indeterminate** number of constituent functions in the case of coordination, that is, ***n*-ary** branching of the nodes that dominate coordinated structures, as opposed to addition of **one** modifier to a nucleus in the case of subordination, that is, **binary** branching of the nodes that dominate subordinated structures.

According to these facts, coordination must be formalized as an *internally* recursive process, so as to allow for the concatenation of *n* constituents. On the other hand, the recursion must be invisible, in order for all the coordinated constituents to appear at the same level, dominated by one and the same node, in the resulting parse tree. This can be achieved by means of the following AGFL structure:

```
Coordinated Structure:
  COOR Structure,
  [ HIDDEN COOR Structure ],
  COORDINATOR ( FINAL ),
  COOR Structure.
```

```
HIDDEN COOR Structure:
  COORDINATOR ( INTERMEDIATE ),
  COOR Structure,
  [ HIDDEN COOR Structure ].
```

```
COOR Structure:
  Structure.
```

```
COORDINATOR ( FINAL ):
  Conjunction ( Y ).
```

```
Conjunction ( Y ):
  "y".
```

```
COORDINATOR ( INTERMEDIATE ):
  Punctuation Mark ( COMMA ).
```

```
Punctuation Mark ( COMMA ):
  ",".
```

Recursion is ensured by the invisible non-terminal **HIDDEN COOR Structure**, which is the optional intermediate constituent of visible non-terminal **Coordinated Structure**. Non-terminal **HIDDEN COOR Structure** is not showed by the analysis, so that the node labelled **Coordinated Structure** in the tree **directly** dominates the nodes labelled **COOR Structure** (and, of course, **COORDINATOR**). It is thus obvious that an important part of this General AGFL Frame for Coordination is the following **DEFINES** module, where **HIDDEN COOR Structure** is not found:

```
DEFINES
Coordinated Structure,
COOR Structure,
COORDINATOR/1,
Structure,
Conjunction/1,
Punctuation Mark/1.
```

The AGFL analysis obtained is the following:

```
Coordinated Structure
  COOR Structure
    Structure
      "... "
    COORDINATOR(INTERMEDIATE)
      Punctuation Mark(COMMA)
        ","
    COOR Structure
      Structure
        "... "
    COORDINATOR(FINAL)
      Conjunction(Y)
        "y"
    COOR Structure
      Structure
        "... "
```

In order to account for various levels of coordination, that is, for recursive application of coordination itself, non-terminal **COOR Structure** must rewrite as follows:

```
COOR Structure:
  Structure;
  Coordinated Structure.
```

We may then note that, while internal recursion accounting for coordination is invisible and right recursive, recursion accounting for recursive application of coordination may be both right and left recursive, and it is visible.

Subordination, on the contrary, is not internally recursive. The following formalization frame accounts for subordination:

```
Subordinated Structure:
  Structure NUCLEUS,
    MODIFIER.
```

```
Structure NUCLEUS:
  Structure.
```

```
MODIFIER:
  Adjective Phrase;
  ... .
```

Subordinated Structure is constituted by **Structure NUCLEUS** and **MODIFIER**. The AGFL analysis obtained is the following:

```
Subordinated Structure
  Structure NUCLEUS
    Structure
      "... "
  MODIFIER
    Adjective Phrase
      "... "
```

Recursive application of subordination, that is, various levels of subordination, is enabled by rewriting **Structure NUCLEUS** as **Subordinated Structure**:

```
Structure NUCLEUS:
  Structure;
  Subordinated Structure.
```

The analysis obtained is in this case the following:

```
Subordinated Structure
  Structure NUCLEUS
    Subordinated Structure
      Structure NUCLEUS
        Structure
          "... "
      MODIFIER
        Adjective Phrase
          "... "
```

MODIFIER
Adjective Phrase
"..."

As can be observed, recursive application of subordination is visible, and, depending on the position of the modifier, it may be right or left recursive: if the modifier is positioned to the right of the nucleus, recursive application of subordination is left recursive, if the modifier is positioned to the left, recursive application of subordination is right recursive.

1.4 Levels of Formalization. The Design of the Grammar.

Formalization of language, or more exactly, the writing of a formal grammar, involves qualitatively different problems in different stages of development —which, from the point of view of the whole grammar, need not to be sequential in time. This is what is called here level of formalization, which must be understood in a wide sense, as a way of referring to the different nature of the decisions taken at different stages of development of the grammar. Three main levels of formalization, or three different kinds of decisions concerning the design of the grammar, will be illustrated in different points of this exposition. Each of these three levels of formalization deals primarily with one of the main issues involved by the design of a formal grammar: model of description, coverage and resolution of ambiguity. Chapters 2 and 3 describe the first and second level and formalization, that is, they deal with formal description and coverage, while problems related to the third level, that is, to resolution of ambiguity, are introduced in Chapter 4.

In writing the grammar, we opted for a modular structure. This makes easier for the grammarian to control the development of the grammar, because he can deal with only one structure, or a set of them, at a time, as well as test each module independently of the others. Modularity is also showed by the structure of the description of the grammar in this book: Chapter 2 deals with the module of the verb phrase, Chapter 3 deals methodically with the modules of the noun phrase, pronoun phrase, adverb phrase, adjective phrase, determined noun phrase, determined pronoun phrase, modifiers, nominalization, subordinated clauses and

prepositional phrases. Chapter 4, finally, must simultaneously deal with all these modules, since we describe there the ambiguity generated both by each module irrespective of the others and by the interaction between some or all of them.

1.4.1 First Level of Formalization. General Frames.

As outlined above, first level of formalization primarily concerns decisions, normally taken on the basis of tradition and familiarity, regarding the choice of the *descriptive framework* and the descriptive grammar that underlies the formal grammar.

Many of these decisions are actually taken before starting the writing of the grammar, but a large number of them is subject to revision later on, confirming the very well-known fact that the writing of a formal grammar is a *cyclic process* in which progressively improved versions of the grammar are elaborated on the basis of linguist's intuitions and knowledge of the structure of the corpus language, as well as test results obtained from previous versions of the grammar.

Moreover, since formalization forces them to be exhaustive in their descriptions, together with continuous revision, we have to refer too to one of the main problems that linguists writing formal grammar have to face, which is the fact that there are *lacunae* —that is, parts of the grammar forgotten in the description— in traditional, not formalized, or formalized but not automatically tested, descriptive grammars.

However, not only lacunae found in traditional descriptions create problems to linguists, but also inconsistencies —that is, similar problems treated in different ways— unforeseeably spread throughout them. First level of formalization deals too with this problem by including the definition of *general frames* of rules, *rule-schemata*, which account for processes considered equally active in various modules of the grammar. This implies that, within formal grammars, general frames capture linguistically relevant *generalizations*, as well as ensure throughout the grammar the *consistency* of the linguistic description, which, unlike the consistency of the analyses, ensured by the conversion of the grammar into a parser, remains a responsibility of the grammarian. Subordination and coordination, whose general frames of description were introduced in previous section, are both good examples of this —the main one— first level of formalization task.

Last but not least, concrete structures and requirements that are undoubtedly part of the language described, that is, the specification of concrete rules and affix values accounting for them, is also considered to be a task developed within the boundaries of the area of influence of the first level of formalization.

1.4.2 Second Level of Formalization. Coverage.

On a second level of formalization, the main issue is *coverage*, that is, where do we put the boundaries of the formalized description, both with respect to the detail of the analysis and to the number of structures which are described by the grammar. Although, strictly speaking, the possibilities are infinite, we can group them around three main approaches: according to these, and in all cases with different degrees of detail, a formal grammar might try to account for a) the competence of native speakers—including, in some way, their capability to understand ungrammatical sequences—, b) what is considered grammatical by the competence of native speakers, or c) what, being considered grammatical by the competence of native speakers, is also a frequent structure of the language. In terms of the utterances in the text material, decisions with respect to coverage will determine whether the parser is going to try to provide at least one analysis for all sequences in the text material or for only some—the grammatical or the grammatical and frequent ones— of them.

Obviously, decisions concerning coverage are very related to the use that we are going to make after of the grammar. While for the elaboration of a language database, especially if manual intervention is planned after the automatic analysis, one may decide to concentrate on detail and on a certain number of, exclusively, grammatical structures, for a real environment application one needs to pay attention to all structures without exception, grammatical or not, and to the selection of the adequate degree of detail, taking always into account both the needs of the application itself and the efficiency of the parser. Because of the duplicity of our objectives, we must, at least partially, satisfy both tendencies. On the one hand, although concentrating our main efforts on phrase structure, we have been quite demanding as to the detail of the analysis throughout the grammar. On the other, we have also provided the grammar with additional strategies adequate for the analysis, to the extent that it is possible, of ungrammatical utterances or

sequences of phrases or words.

From a practical point of view, however, particular problems related to coverage of certain structures have to be solved —ensuring coherency by means of previously defined general criteria— during the process of formalization itself, when particularly describing each structure. In this respect, we have found that decisions corresponding to this level of formalization are frequently connected with fuzzy requirements, often of lexical character, between syntactic constituents of relevant structures, decisions that concern whether phrases like, for instance, *se lo hay que decir*, with its sequence of clitic pronouns and auxiliary verb, or *todos los otros cuatro corredores*, with its sequence of determiners, will be analysed or not.

First and second level of formalization are strongly related to each other and, in fact, the best way to face the problems found in the second level of formalization is to ensure, in the first level, that all distinctions possibly required in the second level can be reached, so that rules and affix values can be handled —adjusted— if, depending on, for instance, the text that is going to be analysed by a concrete application, it is considered necessary. This somewhat complementary relationship between both levels of formalization is further proved by the fact that decisions typically concerning first level of formalization are apt to be expressed by means of syntactic rules, while decisions typically concerning second level of formalization are apt to be expressed by means of affixes or predicates contained in those rules.

1.4.3 Third Level of Formalization. Resolution of Ambiguity.

Although an ideal NLP system should aim at the full and detailed analysis of unrestricted text, including also aspects of a semantic and pragmatic nature, current NLP practice consists of the, almost exclusively, morpho-syntactic analysis of utterances. This is, no doubt, the case for formalized descriptions of language grammars, which do not include the level of semantics and/or pragmatics, nor do they comprise a description of text structure. Obviously, parsers generated from such grammars produce a certain amount of *spurious ambiguity* —in the sense that it is motivated by a difference in the amount of information handled by language as a communicative system and by parsers generated from current formal grammars and lexicons—. This ambiguity, or, strictly speaking, its resolution, is

the main issue in the so-called third level of formalization.

The resolution of ambiguity, on the other hand, is closely related to the environment in which the parser will be applied. The fact that the parser is run after a tagger, or the fact that human intervention is foreseen, for instance, are highly relevant from the point of view of ambiguity resolution within the parser itself. The more the external interventions solving ambiguities for the parser, the less the work that the writer of the grammar has to invest in order to handle ambiguity within the formal grammar itself.

In the case of the grammar described in this book, no external interventions are foreseen, so various additional heuristics have to be introduced in the grammar to handle ambiguity. Although this obviously satisfies the requirements of DoRo, as a project entailing a real environment application, it may cause undesirable erroneous analyses for a linguistic database for language study. For this reason, we tried to keep both in the grammar and in its description throughout this book, clear track of the strategies introduced to choose—on exclusively morpho-syntactic grounds—a certain analysis among all the possible ones for a certain sequence. Structures and resources for solving ambiguities are independent of each other within the grammar, so that the removal of these resources is very easy, allowing us to recover ambiguous results whenever we consider this more adequate for our scope.

1.5 Summary

In this study we try to give an exhaustive formal description of verb phrases from the point of view of their contextualization for constituting clauses. Besides, on the basis of subordination and the proposed general frame for its formalization, we try to give also as consistent and complete as possible a description of the syntactic structures that phrases—noun, pronoun, adverb and adjective phrases—have in Spanish. Our priority is, in this case, to show the structure of the grammar, more than to ensure full coverage, which is considered an issue of second level of formalization, that is, of writing the grammar in such a way that it can be easily added new specifications when and where necessary. It is also our intention to show how the grammar, by handling resources for the resolution of ambiguity, can

be adapted for different purposes, integration in a full-sentence grammar, on the one hand, and identification of relevant content unities within an application for information retrieval, on the other.

The formal grammar that we have developed with this aim, which is accounted for in Chapters 2 and 3, is certainly based on a wide range of grammars and monographs. As far as general principles of grammar are concerned we will not specifically refer during the exposition to the current handbooks and monographs listed in the bibliography. However, we want to point out here certain works which are especially illustrative for the way in which grammar is described in this book. The general model of description, constitutive and functionalist, and the set and character of syntactic unities handled in the grammar, are those described in Rojo [63] and [64], and in Rojo and Jiménez [69]. Voice is essentially described according to García-Miguel [24] and [26]. Subordination and coordination as general mechanisms of expansion of syntactic constituents are described in Jiménez [37], [38] and [39]. We want to point out also all previous works with AGFL (or EAG, *Extended Affix Grammars*). They have all contributed with practical solutions for various problems of formalization, we especially refer to Oostdijk [54], [56] for English, and Hallebeek [33], for Spanish.

Chapter 2

PREDICATE Constituent of the Clause: Verb Phrase

2.1 VP Description

In this Section we aim at describing a part of the linguistic facts that have a role in the constitution of the syntactic category **Verb Phrase**. Verb Phrases (henceforth VPs) are the syntactic categories underlying the syntactic function **PREDICATE** that appears at clause level analysis. From this point of view, VPs are expected to yield all the relevant information for the **PREDICATE** function to be combined with other syntactic functions at the level of the clause. In line with this, by means of VPs all the non-relevant information at this level should be handled earlier in the process of analysis, avoiding this way unwanted interferences of this information in the non-terminals for **PREDICATEs**.

While going through all the relevant and non-relevant linguistic facts having a role in the production of **PREDICATEs**, we will be making use in this work of certain common terms that have been previously used in traditional linguistic descriptions with partially different meanings. It might be worthwhile to introduce here these terms together with the meanings that we intend to give them in this work.

The term *inflected verb*, to begin with, is used here to refer to verb inflection caused by person, gender, number, tense and mood, formally supported by both:

- Suffixes in verb forms.

- The auxiliary verb *haber* inflected for person, number, tense and mood and followed by the past participle of the inflected verb.

The *main verb* of a VP is the verb form that gives its lexical meaning to the VP. The only *auxiliary verb* is *haber* when constituting the inflectional component of main or semiauxiliary verbs. *Semiauxiliaries* are *ser* for passive voice of main verbs, and periphrastic auxiliaries that modify main or semiauxiliary verbs inflected for infinitive or gerund. *Functional clitic pronouns* are unstressed personal pronouns assigned a syntactic function by case, DIRECT OBJECT, INDIRECT OBJECT or PREDICATIVE COMPLEMENT (*me, te, lo, la, le/se, nos, os, los, las, les/se*), while *verb clitic pronouns* are unstressed personal pronouns simply echoing person and number of the verb (*me, te, se, nos, os, se*). These are never assigned a syntactic function in our approach.

2.1.1 Relevant Information for the Contextualization of VPs

On the basis of the previous definitions we can now list the elements that supply the relevant information for the PREDICATE to be combined with other syntactic functions at the level of the clause, that is, for the verb to be *contextualized*. In our approach, this information is the following:

- **Configuration** of the VP that constitutes the PREDICATE.
- **Person, gender and number** of the VP.
- **Case, person, gender and number** of the one or two functional clitic pronouns possibly included in the VP.
- **Mood** of the VP, which is relevant for clauses embedded in other main clauses.

Except for the term **configuration**, all the remaining terms in this list are well known to the linguistic community. As in this work all these terms are used like in more widespread traditions in linguistics, they will not be specifically addressed here. We will instead explain the concept of **configuration**, introducing it along with the concept of **scheme**, to which it is closely related. As a matter of fact, both concepts —**scheme** as well as **configuration**—, refer to a conjunction of the

same four types of information: while schemes are concrete combinations of these four components as abstract or inherent possibilities of verbs, configurations are concrete combinations of the same four components as we can actually find them in linguistic discourse, that is, in real context. The process of **contextualization** thus can be —more technically— defined as the process by means of which, using verb schemes as a starting-point, we produce configurations of VPs acting in real context. By the same token, a description of VPs turns out to be primarily a description of all the mechanisms and formal procedures available in the language to express all the possibilities that the speakers of the language have for yielding configurations by means of the application of the process of contextualization to verb schemes.

Despite the close relation between them pointed out in previous paragraph, not necessarily all descriptions of schemes are also descriptions of configurations, and vice versa. Apart from language evidence about this, the number of different schemes and configurations and how they are related depends also on different linguistic theories. Our concrete approach to this problem, dependent on BDS, will be explained throughout this Section and in Sections 2.2.2.5 and 2.2.2.6. Here, we will only introduce the four components of schemes and configurations, giving at the same time some very preliminary ideas about how they behave in Spanish:

- **Verb type:** this is a representation of the syntactic functions inherently selected by the verb in question. In this approach, we consider the following relevant syntactic functions: SUBJECT, DIRECT OBJECT, INDIRECT OBJECT, PREPOSITIONAL COMPLEMENT, PREDICATIVE COMPLEMENT and AGENT. The possible combinations of verbs with syntactic functions in Spanish and how these combinations are extracted from BDS for each verb will be explained in Sections 2.2.2.5 and 2.2.2.6. An example of a possible combination, and so of a possible verb type, is that which describes the combination of a verb with a subject and a direct object. This is formally represented by **SDO**. Another example that we will be frequently handling in this exposition is that of the combination of a verb with a subject, a direct object and an indirect object, this one formally represented by **SDOIO**. Verb types in this approach are stable, which means that we have

the same verb types and representations throughout all the contextualization process undergone by a verb when, as users of the language, we select one of its possible schemes and produce an actual configuration.

- **Voice:** this is a representation of the various forms by means of which VPs express the organization of syntactic functions within the clause. We recognize the following voice values in Spanish:

Active The VP is constituted by:

- An inflected main verb—including participle verb forms—, or one or more inflected non-passive semiauxiliaries—not including participle verb forms— followed by the main verb inflected for Gerund or Infinitive.
- One or two functional clitic pronouns may also be included together with inflected verb forms different from participle verb forms.

Example: *Viene con nosotros. Ha venido desde París. Se lo ha tenido que decir para calmarlo. Habrías podido ir preparándolo mientras nos esperabas. Recién llegada de París, partió para Nueva York.*

Middle The VP is constituted by:

- An inflected main verb—including participle verb forms—, or one or more inflected non-passive semiauxiliaries—not including participle verb forms— followed by the main verb inflected for Gerund or Infinitive.
- One of the verb clitic pronouns *me, te, se, nos, os, se*. This clitic pronoun must be included together with any inflected verb form except for participle verb forms.
- One functional clitic pronoun may also be included together with any inflected verb form except for participle verb forms.

Example: *Se viene con nosotros. Se ha venido inmediatamente. Me lo he tenido que comprar para poder servirme del que había comprado antes. Habrías podido empezar a hacértelo tú mismo. Decidida a no perdonarlo, lo abandonó.*

Passive The VP is constituted by:

- i) An inflected form —different from participle— of the passive semiauxiliary verb *ser* followed by the main verb inflected for Past Participle, or ii) one or more inflected non-passive semiauxiliaries —not including participle verb forms— followed by the passive semiauxiliary *ser* inflected for Gerund or Infinitive and by the main verb inflected for Past Participle. A simple participle verb form may also constitute a passive VP.
- One functional clitic pronoun may also be included together with any passive voice VP not constituted by just one participle.

Example: *Al final de la carrera, el corredor fue alcanzado por sus contrincantes. Ha sido alcanzado a mitad de recorrido. Podría haber sido alcanzado por una bala perdida. Le ha sido ofrecido el puesto de director. Atraído por las ganancias, en ocasiones arriesga demasiado.*

In our approach, active and middle voices, unrestrictedly associated with different verb types, may be voice values for both schemes and configurations. Passive voice, on the contrary, is a value exclusively associated with configurations obtained from active schemes with a DIRECT OBJECT. In these configurations, the entities underlying the DIRECT OBJECT and the SUBJECT of the active scheme become respectively the SUBJECT and the AGENT of the passive configuration.

Example: the scheme *alcanzar*, *SDO*, *active voice* (*alguien*-SUBJECT *alcanza algo*-DIRECT OBJECT) can be contextualized in the passive configuration *SDO*, *passive voice* (*algo*-SUBJECT *es alcanzado por alguien*-AGENT).

Participle VPs have the same form in different voices. Only the context in which they are found, as well as original schemes of the involved verb, can help to decide on the appropriate voice value of participle verbs. While active voice schemes including a direct object constitute passive voice participle VPs and configurations, other active voice schemes constitute active voice participle VPs and configurations. Middle voice schemes, finally, constitute middle voice VPs and configurations.

- **Preposition:** this obviously stands for prepositions required by verbs that need to be combined with a prepositional complement.

- **Impersonality:** this accounts for the form of the reference to the entity underlying the SUBJECT in the original active or middle schemes for active and middle configurations and in acquired passive configuration for passive configurations. This form is mainly denoted by the VP feature person in the target configuration, which means that impersonality can only be considered with respect to indicative, subjunctive and, very rarely, imperative verb forms¹.

Personal Personal configurations are possible in active, middle or passive voice. For them, the VP feature person varies according to the person of the SUBJECT of the clause. In addition to this, the relationships between the SUBJECT and the entity that it covers in the source scheme or configuration are maintained for all target voice configurations. That is, personal configurations involve an explicit reference of the VP to the entity underlying the SUBJECT in the original active or middle schemes for active and middle configurations and in the acquired passive configuration for passive configurations.

Example: *como, me lavo, tenemos que ayudarlos, es entregado*

Impersonal Various possibilities must be taken into account in this case. All of them have in common the fact that they *explicitly* do not refer² to the entity underlying the SUBJECT in source schemes or configurations:

- Certain verbs with active and middle voice schemes whose verb type includes a SUBJECT inherently empty can only be contextualized by means of VPs blocked in third person singular. The clauses that contain PREDICATEs constituted by these VPs can never include a SUBJECT function. These verbs are considered to be inherently impersonal.

Example: *Hay buenas perspectivas. Se trata de su afición al alcohol.*

- Active or passive voice VPs containing a clitic pronoun *se* and

¹If we consider as such fixed forms as *Véase pág. . .*, which, being non-productive, will not be considered in this exposition.

²This apparent paradox will be explained in the following paragraphs.

whose feature person is blocked in third person singular, constitute PREDICATEs of clauses that block the specification of an explicit SUBJECT function, whose relation with the entity that it was covering in the original active scheme is not changed.

Example: *Se espera ansiosamente a Su Santidad. Se es liberado sólo si se merece serlo.*

- Active voice VPs containing a clitic pronoun *se* and whose feature person varies according to the SUBJECT, which is always third person singular or plural and covers the entity that was covered by the DIRECT OBJECT in the original active scheme, constitute PREDICATEs for clauses that, in this way, hide the entity covered by the SUBJECT in the original active scheme.

Example: *Se venden pisos de lujo.*

- Active, middle or passive voice VPs containing the impersonal semi-auxiliary *haber que*, which is always inflected for third person singular, constitute PREDICATEs for clauses that cannot specify an explicit SUBJECT syntactic function, whose relation with the entity that it was covering in the source scheme or configuration is maintained for all voice configurations.

Example: *Hay que vender todas las existencias. Hay que decidirse de una vez por todas. Habría que haber sido tratado de ese modo para comprenderlo.*

In our approach, personal and first type of impersonality, whose association with verb types has been restricted to S, SDO and SPC, are impersonality values for both schemes and configurations, the other three types of impersonality are only possible for configurations. The second type is obtained from active schemes with animate SUBJECT and, if there is one, animate DIRECT OBJECT. It can also be obtained from passive configurations with animate SUBJECT.

Example: the scheme *esperar*, SDO, active voice, personal (alguien–SUBJECT *espera a* alguien–DIRECT OBJECT) can be contextualized as the active impersonal configuration SDO, active voice, impersonal (*se espera a* alguien–DIRECT OBJECT). Similarly, the passive personal configuration SDO, passive voice, personal (alguien–SUBJECT *es liberado por* alguien–AGENT) can be further contextualized as the passive

impersonal configuration *SDO, passive voice, impersonal (se es liberado por alguien-AGENT)*.

The third type of impersonality is obtained from active schemes with animate SUBJECT and inanimate DIRECT OBJECT. The entity underlying the SUBJECT is not mentioned and the one underlying the DIRECT OBJECT is moved to the SUBJECT position. This syntactic movement of the DIRECT OBJECT is the same described above for passive voice.

Example: the scheme *vender, SDO, active voice, personal (alguien-SUBJECT vende algo-DIRECT OBJECT)* can be contextualized as the active impersonal configuration *SDO, active voice, impersonal (se vende algo-SUBJECT)*.

Finally, the fourth type of impersonality is obtained from active or middle voice schemes or from passive configurations, all of them with animate SUBJECT.

Example: the scheme *vender, SDO, active voice, personal (alguien-SUBJECT vende algo-DIRECT OBJECT)* can be contextualized as the active impersonal configuration *SDO, active voice, impersonal (hay que vender algo-DIRECT OBJECT)*.

In Table 2.1, for both voice and impersonality, we summarize all the possible contextualization processes which in Spanish might yield target configurations on the basis of verb schemes, specifying where necessary requirements on source original schemes, as well as including in all cases the formal procedures that support the final configurations.

2.1.2 VPs not Containing Non-passive Semiauxiliaries

The features identified as relevant elements of information for the contextualization of the VP —configuration, mood, person, gender and number of VP and case, person, gender and number of clitic pronouns— have to be extracted from all the linguistic facts which participate in the production of the VP. According to this, in this and the following Sections, we try to give as exhaustive as possible an enumeration of all the formal procedures that support the expression of VP configurations, specifying in each case the information that these VPs should yield for the PREDICATE functions that they constitute. Besides, since orthographical

<i>active, personal</i>		<i>active, personal</i>	<i>viene</i>
<i>active, personal</i>	direct object	<i>passive, personal</i>	<i>es alcanzado</i>
<i>active, personal</i>	animate direct object	<i>passive, impersonal</i>	<i>se es liberado</i> <i>hay que haber sido tratado</i>
<i>active, personal</i>	animate subject inanimate direct object	<i>active, impersonal</i>	<i>se venden</i>
<i>active, personal</i>	animate subject	<i>active, impersonal</i>	<i>se espera</i> <i>hay que vender</i>
<i>active, impersonal</i>		<i>active, impersonal</i>	<i>hay</i>
<i>middle, personal</i>		<i>middle, personal</i>	<i>se viene</i>
<i>middle, personal</i>	animate subject	<i>middle, impersonal</i>	<i>hay que decidirse</i>
<i>middle, impersonal</i>		<i>middle, impersonal</i>	<i>se trata</i>

Table 2.1: Schemes and final target configurations.

changes have to be carefully accounted for in the case of applications of language formalization to written language, we, even though this concerns a different and lower level of description, will be throughout all this Section specifically addressing the orthographical deviations experimented by verb forms when combined with clitic pronouns. On the other hand, although we are aware that certain structures of VPs will no doubt be found very infrequent in the following explanations, we do not refrain from including them, and leave the reader judge about it, since it is obviously a question concerning second level of formalization whether these formal procedures should be included or not in formal grammars developed for concrete applications.

2.1.2.1 Elements of VPs that do not include functional clitic pronouns

- An inflected main verb in an **active personal configuration**.

Example: *alcanzo, he alcanzado* should both yield the information *SDO, active voice, personal, indicative, first person, singular*; *alcanzar, haber alcanzado* should both yield the information *SDO, active voice, infinitive*; *llegada* should yield the information *S, active voice, participle, feminine, singular*.

- An inflected main verb in a **passive personal configuration**.

Example: *es alcanzado, ha sido alcanzado* should both yield the information *SDO, passive voice, personal, indicative, third person, masculine/neuter, singular*; *ser alcanzado, haber sido alcanzado* should both yield the information *SDO, passive voice, infinitive, masculine/neuter*,

singular; alcanzado should yield the information *SDO, passive voice, participle, masculine/neuter, singular*.

- An inflected main verb in a **middle personal configuration**.

Example: *me pongo, me he puesto* should both yield the information *SDO, middle voice, personal, indicative, first person, singular*; *ponerse, haberse puesto* should both yield the information *SDO, middle voice, infinitive, third person, singular/plural*; *decidida* should yield the information *SPC, middle voice, participle, feminine, singular*.

In middle voice, the combination with the necessary —non-functional— clitic pronoun involves certain changes in the orthographical form of the verb for imperative and gerund moods. For imperative, the following changes have to be considered:

- In second person singular, unaccented polysyllabic verb forms take an accent.

Example: *atrévete, espérate*.

- In second person singular, polysyllabic verb forms that have a stress mark in the last syllable, as well as those monosyllabic forms with diacritic stress mark, lose the accent.

Example: *componte, sete*.

- Unaccented polysyllabic subjunctive forms used for polite reference to second person singular and plural take an accent:

Example: *atrévase, atrévanse*.

- Third person singular monosyllabic subjunctive forms with diacritic stress mark, used for polite reference to second person singular imperative forms, lose the accent:

Example: *dese*.

- All first plural subjunctive forms lose the final *–s* and take an accent:

Example: *amémonos, cojámonos*.

- In second person plural, all verb forms lose the final *–d*. In third conjugation, unaccented verb forms also take an accent³:

³Verb *ir, id*, constitutes the only exception to this rule.

Example: *amaos, veníos.*

For gerund, the following changes have to be considered:

- All forms take an accent:

Example: *amándose.*

- An inflected main verb in an **active impersonal configuration**.

Example: *se alcanza, se ha alcanzado, hay, ha habido* should all yield the information *SDO, active voice, impersonal, indicative, third person, singular*; *se venden, se han vendido* should both yield the information *SDO, active voice, impersonal, indicative, third person, plural*.

- An inflected main verb in a **middle impersonal configuration**.

Example: *se trata, se ha tratado* should both yield the information *SPC, middle voice, impersonal, indicative, third person, singular*.

- An inflected main verb in a **passive impersonal configuration**.

Example: *se es liberado, se ha sido liberado* should both yield the information *SDO, passive voice, impersonal, indicative, third person, masculine, singular*.

2.1.2.2 Elements of VPs that include one functional clitic pronoun

- An inflected main verb in an **active personal configuration** with one functional clitic pronoun.

Example: *lo pongo, lo he puesto* should both yield the information *SDO, active voice, personal, indicative, first person, singular, accusative, third person, masculine/neuter, singular*; *ponerlo, haberlo puesto* should both yield the information *SDO, active voice, infinitive, accusative, third person, masculine/neuter, singular*.

The introduction of this clitic pronoun forces certain changes in the orthographical form of imperatives and gerunds. For imperative, changes are the following:

- In second person singular, unaccented polysyllabic verb forms take an accent.

Example: *cógelo, déjalo.*

- In second person singular, polysyllabic verb forms that have a stress mark in the last syllable, as well as those monosyllabic forms with diacritic stress mark, lose the accent.

Example: *componle, sele.*

- Unaccented polysyllabic subjunctive forms used for polite reference to second person singular and plural take an accent:

Example: *hágalo, háganlo.*

- Third person singular monosyllabic subjunctive forms with diacritic stress mark, used for polite reference to second person singular imperative forms, lose the accent:

Example: *dele.*

- All first person plural subjunctive forms take an accent:

Example: *hagámoslo, cojámoslo.*

For gerund, the following changes have to be considered:

- All forms take an accent:

Example: *haciéndolo, diciéndolo.*

- An inflected main verb in a **passive personal configuration** with one functional clitic pronoun (INDIRECT OBJECT):

Example: *le es entregado, le ha sido entregado* should both yield the information *SDO, passive voice, personal, indicative, third person, masculine/neuter, singular, dative, third person, masculine/feminine/neuter, singular*; *serle entregado, haberle sido entregado* should both yield the information *SDO, passive voice, infinitive, masculine/neuter, singular, dative, third person, masculine/feminine/neuter, singular*.

The addition of this clitic pronoun involves the appearance of an alternative form of the gerund for the passive semiauxiliary *ser*:

- Gerund form takes an accent:

Example: *siéndole entregado.*

- An inflected main verb in a **middle personal configuration** with one functional clitic pronoun:

Example: *me lo pongo, me lo he puesto* should both yield the information *SDO, middle voice, personal, indicative, first person, singular, accusative, third person, masculine/neuter, singular*; *ponérmelo, habérmelo puesto* should both yield the information *SDO, active voice, infinitive, first person, singular, accusative, third person, masculine/neuter, singular*.

With the addition of this functional clitic pronoun, new changes in some orthographical forms of verbs have to be added to those already mentioned for middle voice VPs without functional clitic pronouns. In imperative the following changes have to be considered:

- In second person singular, unaccented monosyllabic verb forms take an accent.

Example: *háztelo, póntelo*.

- In second person plural, for first and second conjugation verb forms, the final *-d-* is removed and an accent is added.

Example: *quedáoslo, imagináoslo*.

- Unaccented monosyllabic subjunctive forms used for polite reference to second person singular and plural take an accent.

Example: *dénsela*.

In infinitive:

- Unaccented verb forms take an accent:

Example: *entregársele, acercársenos*.

- An inflected main verb in an **active impersonal configuration** with one functional clitic pronoun.

Example: *se la presenta, se la ha presentado, la hay, la ha habido* should yield the information *SDOPR* —*SDO* in the case of *haber*—, *active voice, impersonal, third person, singular, accusative, third person, feminine, singular*; *se le entregan, se le han entregado* should both yield the information *SDOIO, active voice, impersonal, third person, plural, dative, third person, masculine/feminine/neuter, singular*.

- An inflected main verb in a **passive impersonal configuration** with one functional clitic pronoun.

Example: *se le es entregado, se le ha sido entregado* should both yield the information *SDO, passive voice, impersonal, third person, masculine, singular, dative, third person, masculine/feminine/neuter, singular*.

2.1.2.3 Elements of VPs that include two functional clitic pronouns

- An inflected main verb in an **active personal configuration** with two functional clitic pronouns.

Example: *se lo pongo, se lo he puesto* should both yield the information *SDOIO, active voice, personal, indicative, first person, singular, dative, third person, masculine/feminine/neuter, singular/plural, accusative, third person, masculine/neuter, singular; ponérselo, habérselo puesto* should both yield the information *SDOIO, active voice, infinitive, dative, third person, masculine/feminine/neuter, singular/plural, accusative, third person, masculine/neuter, singular*.

With the addition of the second functional clitic pronoun, new changes in some orthographical forms of verbs have to be added to those already mentioned for active voice VPs with one functional clitic pronoun. For imperative, the following changes have to be taken into account:

- In second person singular, unaccented monosyllabic verb forms take an accent.

Example: *dámelo, díselo*.

- In all first person plural subjunctive forms the final *-s-* is removed and an accent is added.

Example: *entreguémoselo*.

- In second person plural, unaccented verb forms take an accent.

Example: *entregádselo, devolvédsele*.

- Unaccented monosyllabic subjunctive forms used for polite reference to second person singular or plural take an accent:

Example: *dénselo*.

For infinitive:

- Unaccented verb forms take an accent.

Example: *decírselo, perdonárselo*.

2.1.3 Non-passive Semiauxiliaries: Periphrastic Groups

So far, we have been dealing with VPs that do not include non-passive semiauxiliaries. In this and the following Sections, we will instead be analysing VPs that contain them, proceeding as before from VPs that do not include functional clitic pronouns to those that include one and two of them.

VPs containing non-passive semiauxiliaries are constituted by a *periphrastic group*, the one or more non-passive semiauxiliaries, followed by a *main group*, an inflected main verb in active, passive or middle voice.

Example: *ha tenido que ponerse a trabajar* is a VP constituted by a periphrastic group, *ha tenido que ponerse a*, and a main group in active voice, *trabajar*. *Está empezando a vestirse* is a VP constituted by a periphrastic group, *está empezando a*, and a main group in middle voice, *vestirse*. *Puede haber sido expulsado* is a VP constituted by a periphrastic group, *puede*, and a main group in passive voice, *haber sido expulsado*.

Non-passive semiauxiliaries require its following verb to be an infinitive or gerund. In certain cases, a preposition or conjunction must be introduced between both verbs. Non-passive semiauxiliaries —as main verbs— can be inherently active or middle voice, personal or impersonal, this having the following consequences:

Active The inflected non-passive semiauxiliary can constitute a periphrastic group by itself.

Middle If this is the case, the inflected non-passive semiauxiliary needs a clitic pronoun from the set *me, te, se, nos, os, se* to constitute a periphrastic group. This pronoun echoes the person and number features of the verb and SUBJECT, and is never assigned a syntactic function.

Personal The periphrastic group constituted by the non-passive semiauxiliary enables the VP to refer to the entity underlying the SUBJECT in active and middle voice original schemes, or in acquired passive voice, of main verbs.

Impersonal If this is the case, the non-passive semiauxiliary is inherently blocked in third person singular. This fact blocks in the VP the reference to the entity underlying the SUBJECT in active and middle voice original schemes, or in acquired passive voice, of main verbs.

For the grammar described in this book, the set of non-passive semiauxiliaries is the same used for the analysis of BDS (see Section 2.2.2.6). In BDS, those displayed in Appendix A, Table A.1 are considered periphrastic auxiliaries.

Apart from the clitic pronoun required by middle voice semiauxiliaries, the VPs that contain non-passive semiauxiliaries may also include clitic pronouns selected by the main verb. The clitic pronoun required by middle voice semiauxiliaries is always placed in the periphrastic group within the VP. This clitic pronoun cannot be combined with any other clitic pronoun within the periphrastic group, if the main verb selects another clitic pronoun, this one must be placed within the main group. On the contrary, when the VP does not include a middle voice semiauxiliary in its periphrastic group, that is, for active voice periphrastic groups, clitic pronouns selected by the main verb may be placed either within the main group or within the periphrastic one. There are, nevertheless, some restrictions—with sometimes very unclear boundaries—related to voice and impersonality: these restrictions will be specified below, when particularly describing VPs that include non-passive semiauxiliaries.

Example: *puede hacerlo* and *lo puede hacer* are both VPs containing the active voice semiauxiliary *poder*. In both of them we have grammatically acceptable arrangements for the various elements that together supply the same information in these VPs. On the contrary, *se puso a hacerlo* is a VP containing the middle voice semiauxiliary *ponerse a*, but we cannot in this case move the accusative clitic pronoun to the periphrastic group and thus say something like: *se lo puso a hacer**. This restriction is a logical consequence of the fact that VPs with middle voice semiauxiliaries might include up to three clitic pronouns—one required by the middle voice semiauxiliary and two more clitic pronouns selected by the main verb, as in *me he puesto a hacérselo* (and we should not forget, on the other hand, that one of these two pronouns might also indicate middle voice of the main verb, as in *me he puesto a hacérmelo*).

Various non-passive semiauxiliaries may be combined within the same VP. Their sequence is not free, and therefore traditional approaches to linguistic description have tried to establish semantic classifications within them, attempting to account for how they are combined. However, since various classifications of this type have been produced which are at least partially different, our particular approach aims at avoiding this problem by considering each of the semiauxiliaries referred in Table A.1 as a class. Combinations between them are thus accepted only if they have been documented in corpora, in our case if they have been

documented in BDS. In spite of this, for the purposes of this work we have only accounted for one especially relevant fact of this type: middle voice semiauxiliaries, *echarse a*, *ponerse a*, must always be placed at the last position within periphrastic groups, that is, they always immediately precede the main verb.

When various non-passive semiauxiliaries are combined within the same VP, if clitic pronouns are included in the periphrastic group, they can be placed beside the first semiauxiliary, an intermediate one or the last one. There are only few restrictions to this rule which seem to be related with voice and impersonality of semiauxiliaries. These restrictions will be specifically addressed when particularly describing VPs containing non-passive semiauxiliaries.

Obviously, sequences of semiauxiliaries involve also combinations of different voice and impersonality features. In these cases, it is necessary to determine the voice and impersonality of the whole periphrastic group, because this will have to be combined after with a main verb for the production of the periphrastic VP. Two simple rules account for this fact:

- If a periphrastic group includes one middle voice semiauxiliary, it is a middle voice periphrastic group.

Example: *ha tenido que ponerse a*, *estuvimos a punto de echarnos a*, *puede haber tenido que ponerse a* are all middle voice periphrastic groups.

- If a periphrastic group includes one impersonal semiauxiliary, it is an impersonal periphrastic group.

Example: *hay que empezar a*, *puede haber que* are both active voice and impersonal periphrastic groups. *Hay que ponerse a*, *puede haber que ponerse a* are both middle voice and impersonal periphrastic groups.

2.1.4 VPs⁴ Containing Periphrastic Groups

2.1.4.1 Elements of VPs that do not include functional clitic pronouns

- An active or middle personal periphrastic group together with an active personal main verb results in an **active personal configuration**.

⁴Auxiliaries, when combined with clitic pronouns —one or two— undergo the same orthographical variations enumerated for verb forms of main verbs appearing in VPs not containing non-passive semiauxiliaries (see Section 2.1.2). We will not repeat the explanation of these variations in the following description.

Example: *pudo alcanzar, pudo ponerse a hacer/se pudo poner a hacer* should all yield the information *SDO, active voice, personal, indicative, third person, singular*; *poder alcanzar* should yield the information *SDO, active voice, infinitive*; *ponerse a hacer* should yield the information *SDO, active voice, infinitive, third person, singular/plural*.

Note that the clitic pronoun of the middle voice semiauxiliary can be placed anywhere within the periphrastic group. On the other hand, it can never be combined with any other clitic pronoun—which, if present, should be placed in the main group, see above—.

- An active or middle personal periphrastic group together with an active impersonal main verb results in an **active impersonal configuration**.

Example: *puede llover, pudo ponerse a llover/se pudo poner a llover* should all yield the information *S, active voice, impersonal, indicative, third person, singular*.

- An active or middle personal periphrastic group together with a middle personal main verb results in a **middle personal configuration**.

Example: *puede vestirse/se puede vestir, se puso a vestirse* should all yield the information *SDO, middle voice, personal, indicative, third person, singular*; *poder vestirse/poderse vestir, ponerse a vestirse* should all yield the information *SDO, middle voice, infinitive, third person, singular/plural*.

For active voice periphrastic groups, the middle voice clitic pronoun of the main verb can be placed anywhere in the VP, next to either the main verb or to any of the active voice semiauxiliaries constituting the periphrastic group.

- An active personal periphrastic group together with a middle impersonal main verb results in a **middle impersonal configuration**.

Example: *puede tratarse/se puede tratar* should both yield the information *SPC, middle voice, DE, impersonal, indicative, third person, singular*.

As before, the middle voice clitic pronoun of the main verb can be placed anywhere in the VP.

- An active personal periphrastic group together with a passive personal main verb results in a **passive personal configuration**.

Example: *puede haber sido alcanzado* should yield the information *SDO, passive voice, personal, indicative, third person, masculine/neuter, singular*; *poder ser tratado* should yield the information *SDO, passive voice, infinitive, masculine/neuter, singular*.

- An active or middle impersonal periphrastic group together with an active personal main verb results in an **active impersonal configuration**.

Example: *hay que hacer, hay que ponerse a hacer* should both yield the information *SDO, active voice, impersonal, indicative, third person, singular*.

Note that the middle voice clitic pronoun of the middle voice semiauxiliary cannot be placed beside or before the active voice impersonal semiauxiliary, that is, sequences like *Se hay que poner a** or *Se puede haber que poner a** are not possible.

- An active or middle impersonal periphrastic group together with a middle personal main verb results in a **middle impersonal configuration**.

Example: *hay que hacerse, hay que ponerse a hacerse* should both yield the information *SDO, middle voice, impersonal, indicative, third person, singular*.

For active voice periphrastic groups, the clitic pronoun of the main verb can be placed in the periphrastic group after the impersonal semiauxiliary: *hay que irse haciendo*.

- An active impersonal periphrastic group together with a passive personal main verb results in a **passive impersonal configuration**.

Example: *hay que haber sido maltratado* should yield the information *SDO, passive voice, impersonal, indicative, third person, masculine, singular*.

- An active personal periphrastic group in third person singular or plural together with the clitic pronoun *se* and an active personal main verb results in an **active impersonal configuration**:

Example: *puede recibirse/se puede recibir* should both yield the information *SDO, active voice, impersonal, indicative, third person, singular*; *se pueden vender/pueden venderse* should both yield the information *SDO, active voice, impersonal, indicative, third person, plural*.

The clitic pronoun *se* can be placed anywhere in the VP, either beside the main verb or beside any of the active voice semiauxiliaries constituting the periphrastic group.

- An active personal periphrastic group together with the clitic pronoun *se* and a passive personal main verb results in a **passive impersonal configuration**:

Example: *se puede ser encarcelado* should yield the information *SDO*, *passive voice, impersonal, indicative, third person, masculine, singular*.

The clitic pronoun *se* must in this case be placed somewhere in the periphrastic group, to avoid the cacophony that would be produced if the pronoun were attached to the infinitive of the passive semiauxiliary *ser* within the main group.

2.1.4.2 Elements of VPs that include one functional clitic pronoun

- An active or middle personal periphrastic group together with an active personal main verb and one functional clitic pronoun results in an **active personal configuration**.

Example: *puedo hacerlo/lo puedo hacer, se puso a hacerlo* should all yield the information *SDO*, *active voice, personal, indicative, first person, singular, accusative, third person, masculine/neuter, singular*; *poder hacerlo/poderlo hacer* should both yield the information *SDO*, *active voice, infinitive, accusative, third person, masculine/neuter, singular*; *ponerse a hacerlo* should yield the information *SDO*, *active voice, infinitive, third person, singular/plural, accusative, third person, masculine/neuter, singular*.

For active voice periphrastic groups, the functional clitic pronoun of the main verb can be placed anywhere in the VP, either beside the main verb or beside any of the active voice semiauxiliaries constituting the periphrastic group.

- An active or middle personal periphrastic group together with an active impersonal main verb and one functional clitic pronoun results in an **active impersonal configuration**.

Example: *puede haberlo/lo puede haber* should both yield the information *SDO*, *active voice, impersonal, indicative, third person, singular, accusative, third person, masculine/neuter, singular*.

For active voice periphrastic groups, the functional clitic pronoun of the main verb can be placed anywhere in the VP.

- An active or middle personal periphrastic group together with a middle personal main verb and one functional clitic pronoun results in a **middle personal configuration**.

Example: *puedo hacérmelo/me lo puedo hacer, me pongo a hacérmelo* should all yield the information *SDO, middle voice, personal, indicative, first person, singular, accusative, third person, masculine/neuter, singular*; *poder hacérmelo/podérmelo hacer, ponerme a hacérmelo* should all yield the information *SDO, middle voice, infinitive, first person, singular, accusative, third person, masculine/neuter, singular*.

For active voice periphrastic groups, clitic pronouns can be placed anywhere in the VP.

- An active personal periphrastic group together with a passive personal main verb and one functional clitic pronoun results in a **passive personal configuration**.

Example: *le puede ser entregado* should yield the information *SDO, passive voice, personal, indicative, third person, masculine/neuter, singular, dative, third person, masculine/feminine/neuter, singular*; *poderle ser entregado/poder serle entregado* should both yield the information *SDO, passive voice, infinitive, masculine/neuter, singular, accusative, third person, masculine/feminine/neuter, singular*.

The functional clitic pronoun can be placed anywhere in the VP.

- An active or middle impersonal periphrastic group together with an active personal main verb and one functional clitic pronoun results in an **active impersonal configuration**.

Example: *hay que hacerlo/lo hay que hacer, hay que ponerse a hacerlo* should all yield the information *SDO, active voice, impersonal, indicative, third person, singular, accusative, third person, masculine/neuter, singular*.

For active voice periphrastic groups, the functional clitic pronoun can be placed anywhere in the VP.

- An active or middle impersonal periphrastic group together with a middle personal main verb and one functional clitic pronoun results in a **middle impersonal configuration**.

Example: *hay que hacérselo, hay que ponerse a hacérselo* should both yield the information *SDO, middle voice, impersonal, indicative, third person, singular, accusative, third person, masculine/neuter, singular*.

For active voice periphrastic groups, clitic pronouns can be placed in the periphrastic group, but they must be placed after the impersonal semiauxiliary: *hay que írselo haciendo*.

- An active impersonal periphrastic group together with a passive personal main verb and one functional clitic pronoun results in a **passive impersonal configuration**.

Example: *hay que haberle sido presentado* should yield the information *SDO, passive voice, impersonal, indicative, third person, masculine, singular, dative, third person, masculine/feminine/neuter, singular*.

The clitic pronoun of the main verb cannot be placed next to or before the impersonal active voice semiauxiliary.

- An active personal periphrastic group in third person singular or plural together with the clitic pronoun *se*, an active personal main verb and one functional clitic pronoun results in an **active impersonal configuration**:

Example: *puede recibírsela/se la puede recibir* should yield the information *SDO, active voice, impersonal, indicative, third person, singular, accusative, third person, feminine, singular; se le pueden vender/pueden vendérsele* should both yield the information *SDO, active voice, impersonal, indicative, third person, plural, dative, third person, masculine/feminine/neuter, singular*.

The clitic pronouns can be placed anywhere in the VP.

- An active personal periphrastic group together with the clitic pronoun *se*, a passive personal main verb and one functional clitic pronoun results in a **passive impersonal configuration**:

Example: *se le tiene que haber sido presentado* should yield the information *SDO, passive voice, impersonal, indicative, third person, masculine, singular, dative, third person, masculine/feminine/neuter, singular*.

Clitic pronouns must be placed in this case within the periphrastic group, to avoid the cacophony that would be produced if the infinitive *ser* and the clitic pronoun *se* were combined.

2.1.4.3 Elements of VPs that include two functional clitic pronouns

- An active or middle personal periphrastic group together with an active personal main verb and two functional clitic pronouns results in an **active personal configuration**:

Example: *se lo puedes contar/puedes contárselo, se pone a contárselo* should all yield the information *SDOIO, active voice, personal, indicative, second person, singular, dative, third person, masculine/feminine/neuter, singular/plural, accusative, third person, masculine/neuter, singular*; *poder contárselo/podérselo contar* should both yield the information *SDOIO, active voice, personal, infinitive, dative, third person, masculine/feminine/neuter, singular/plural, accusative, third person, masculine/neuter, singular*.

Clitic pronouns can be placed anywhere in the VP.

- An active or middle impersonal periphrastic group together with an active personal main verb and two functional clitic pronouns results in an **active impersonal configuration**:

Example: *se lo hay que contar/hay que contárselo, hay que ponerse a hacérselo* should all yield the information *SDOIO, active voice, impersonal, indicative, third person, singular, dative, third person, masculine/feminine/neuter, singular/plural, accusative, third person, masculine/neuter, singular*.

Clitic pronouns can be placed anywhere in the VP.

2.1.5 Other Questions about VPs

Apart from configuration, the remaining features necessary for the contextualization of VPs —**person, gender, number** and **mood** of the verb component and **case, person, gender** and **number** of the clitic pronouns attached to it—pose also some minor questions that have to be dealt with when formalizing the production of VPs.

The fact that there is not a one-to-one correspondence between the functional combinations of the values for these features and verb forms has to be considered.

For regular Spanish verbs forms, this one-to-one correspondence fails in the following cases:

- Third person singular present indicative is equal to second person singular imperative, *ama* (except for irregular verbs which have experimented some type of apocope or contraction *sal* (*salir*), *pon* (*poner*), *ten* (*tener*), *ven* (*venir*), *haz* (*hacer*), *di* (*decir*), *ve* (*ir*)).
- Third person singular and first person singular are equal in imperfect (*amaba*), conditional (*amaría*), pluperfect (*había amado*) and past conditional (*habría amado*) of indicative mood, and in all tenses of subjunctive mood (*ame*, *amara*, *amase*, *amare*, *haya amado*, *hubiera amado*, *hubiese amado*, *hubiere amado*).
- In first and third conjugation, present indicative and past indicative are equal in first person plural, *amamos*, *partimos* (except for verbs with strong perfect, *estamos/estuvimos* (*estar*), *decimos/dijimos* (*decir*), *conducimos/condujimos* (*conducir*), *venimos/vinimos* (*venir*) and their derivatives).
- Third person singular and plural are always used as polite reference to second person.
- Polite reference to second person singular and plural together with first person plural of imperative affirmative, as well as all imperative negative forms are equal to subjunctive present forms, *no ames*, (*no*) *ame*, (*no*) *amemos*, *no améis*, (*no*) *amen*.

For clitic pronouns we assume the following —etymological— behaviour:

- Clitic pronouns *me*, *nos*, *te*, *os* indicate middle voice, direct or indirect object.
- Clitic pronouns *lo*, *la*, *los*, *las* indicate direct object.
- Clitic pronouns *le*, *les* indicate indirect object.
- Clitic pronoun *se* indicates middle voice, impersonality or indirect object preceding a direct object or predicative complement clitic pronoun.
- Clitic pronoun *lo* indicates predicative complement.

Clitic pronouns are placed in front of the verb in indicative, subjunctive or negative imperative forms, behind it in infinitive, gerund or affirmative imperative forms (*lo amo, lo ame, no lo ames, amarlo, amándolo, ámalo*).

From the point of view of **person**, clitic pronouns are combined with verbs as follows:

- Verb and one functional clitic pronoun have different person or both are third person.

Example: *te amo, me amas, lo amo, lo amas, lo ama.*

- Verb and two functional clitic pronouns have different person or two or three of them are third person.

Example: *te lo digo, me lo dices, se lo digo, se lo dices, se lo dice.*

Verb type determines the possibilities that a configuration has of including one or two functional clitic pronouns of a certain type. For personal and impersonal non-*se* configurations the situation is the following:

- Obviously, configurations including one functional clitic pronoun, a direct object, an indirect object or a predicative complement, are based on verb types that include, respectively, a direct object, an indirect object or a predicative complement argument. In passive voice configurations, since the obligatory direct object of the original scheme has been moved to the SUBJECT position, the only functional clitic pronoun allowed is the indirect object one.
- Active configurations including two functional clitic pronouns, an indirect object one in first position and a direct object or predicative complement one in second position, are based on verb types that include an indirect object as well as a direct object or predicative complement arguments.

For impersonal *se* configurations, the situation is the following:

- Configurations with a direct object clitic pronoun are based on verb types that include a direct object argument and that do not include an indirect object one (*se lo recibe*). The reason for this is that the clitic pronoun *se*

that expresses impersonality would otherwise be interpreted as a functional indirect object clitic pronoun (*se lo entrega*)⁵.

- Configurations with an indirect object clitic pronoun are based on verb types that include an indirect object clitic pronoun argument. This is the only functional clitic pronoun that can be combined with active or passive impersonal *se* configurations changing the relations between the SUBJECT and the entity that was covered by it in the original active scheme (*se le es presentado, se le venden*).
- Predicative complement clitic pronouns are not allowed.

2.2 AGFL Formalization of the VP

As was observed in Chapter 1, our main purpose in this work is the AGFL formal description of linguistic structures of various phrases, which should after be the base for an AGFL formal analysis of the clause in terms of syntactic functions. Restricting ourselves to VPs, this objective entails the reduction of all the linguistic facts that participate in the production of VPs, to obtain a simple and adequate formal set of values —AGFL terminal affixes— which together constitute the information necessary for the contextualization of the PREDICATE function at the level of the clause. So far we have described the linguistic facts that have a role in the production of VPs as we are going to formally account for them. In this Section the AGFL formalization of these linguistic facts and their progressive reduction into a relevant set of AGFL feature values at the level of the clause will be faced. In first place, we explain the structure of the lexical data that contain verb forms —verb lexicon non-terminals will be introduced and related to verb lexicon terminals—. In second place, features and feature values —affix non-terminals and terminals— will be considered in detail. Finally, rules for the production of the VP —syntax rules, grammar non-terminals and terminals— will be described.

⁵Obviously, this holds only for impersonal *-se* configurations that keep the original relations between SUBJECT, DIRECT OBJECT and the entities covered by them in the original schemes. For impersonal *-se* configurations changing these relations —obtained from a previous passive personal configuration (*se es presentado*) or from an active personal scheme (*se venden*)—, given the fact that the original DIRECT OBJECT is turned into the SUBJECT of the target configuration, a functional direct object clitic pronoun cannot be included in the VP.

2.2.1 The Lexicon

The lexicon associated with the grammar described in this exposition is a wordform lexicon. This means that, from both the morphological and the syntactic point of view —voice, impersonality, person, number, gender, tense, mood and functional enclitic variations are all accounted for—, the linguistic phenomena described in the VP grammar are those which occur beyond the level of the word, which is conceived as a sequence of non-space characters between two space characters, the only exception being enclitic pronouns. In other words, grammar describes compound tenses, grammatical voice and impersonality acquisition, periphrastic verb forms and clitic pronoun addition.

To be able to account for these phenomena, the lexicon must contain the following wordforms of auxiliary, semiauxiliary and main verb lemmas: present (*amo*), imperfect (*amaba*), past (*amé*), conditional (*amaría*) and future (*amaré*) of indicative mood, present (*ame*), imperfect (*amara/amase*) and future (*amare*) of subjunctive mood, present (*ama*) of imperative mood, present (*amar, amando*) of infinitive and gerund moods and participle (*amado*). Lexicon non-terminals for these verb forms are the following:

```
VerbSt ( vtype, voice, tense, mood )
Participle VerbSt ( vtype, voice, mood, gender, number )
VerbSt ( vtype, voice, person, number, tense, mood )
```

The first describes infinitives and gerunds, the second describes participles and the third finite forms. Infinitive and gerund forms are specified for tense and mood. Participles are specified for mood, gender and number. Finite forms are specified for **person, number**, tense and mood. All of them are specified for verb type and voice. Verbs whose verb type value includes some kind of prepositional complement add an affix to record information about the preposition introducing the prepositional complement. For these verbs, previous lexicon definitions are turned into the following ones:

```
VerbSt ( vtype, voice, linktype, tense, mood )
Participle VerbSt ( vtype, voice, linktype, mood, gender, number )
VerbSt ( vtype, voice, linktype, person, number, tense, mood )
```

Affix rules, affix non-terminals and terminal values for *vtype*, *voice*, *linktype*,

person, *gender*, *number*, *tense* and *mood* will be explained in detail in next Section 2.2.2.

Examples of lexicon definitions —lexicon rules— are the following:

```
VerbSt ( SDO, ACT, PRESENT, INFINITIVE ): "amar".
Participle VerbSt ( SDO, ACT, PARTICIPLE, MASC|NEUT, SING ): "amado".
VerbSt ( SDO, ACT, FIRST, SING, PRESENT, INDICATIVE ): "amo".
```

Without taking into account affix values for verb type, voice and prepositional requirements, we call *verb form cluster* the collection of definitions of this type needed to cover the inflection of each verb lemma within the lexicon: in principle, we would say, as many as combinations of single values for gender, person, number, tense and mood —*inflectional verb forms*— may be represented by the verb lemma.

Such a statement, however, is not exact because of syncretism, a linguistic phenomenon that is handled by the grammar as follows: for the *lexical verb forms* that can refer to more than one value for person, number, tense, mood or gender, these values are combined within the same lexicon definition, provided also that they can without restrictions be combined with all the other possibly more than one values of the other affixes involved in the definition. As a logical consequence of this, there is not a one-to-one relation between verb lexicon rules, inflectional verb forms and lexical verb forms. Unique lexicon definitions for *amaba* or *amamos*, for instance, are⁶:

```
VerbSt ( vtype, voice, FIRST|THIRD, SING, IMPERFECT, INDICATIVE ):
    "amaba".
VerbSt ( vtype, voice, FIRST, PLU, PRESENT|PAST, INDICATIVE ):
    "amamos".
```

Where, in each case, two inflectional verb forms are related to the same lexical verb form and simultaneously described in the same lexicon rule. Another example of this missadjustment is the lexical verb form *ama*, which must be described by two different lexicon definitions because two of the theoretical combinations of affix values —{THIRD, IMPERATIVE}, {SECOND, INDICATIVE}— do not correspond to real inflectional verb forms for *ama* in Spanish:

```
VerbSt ( vtype, voice, THIRD, SING, PRESENT, INDICATIVE ): "ama".
VerbSt ( vtype, voice, SECOND, SING, PRESENT, IMPERATIVE ): "ama".
```

⁶Henceforth, we will adopt the convention of not instantiating the attributes *vtype* and *voice* when referring to the verb form cluster.

The same holds for affix values referring to the scheme of the verb: verb type (*vtype*), voice (*voice*) and prepositional requirements (*preptype*). While one of them has more than one value and is freely variable with respect to all the other possibly more than one values of the other two affixes, these are combined in each verb form definition for all the verb form cluster involved. Each combination of one or more values for verb type, one or more values for voice and one or more values for prepositional requirements constitutes what we call *verb scheme cluster*, the production of which will be explained in detail in Section 2.2.2.6⁷.

In addition to those previously described, two more verb lexicon non-terminals are found in the lexicon. They have to be introduced to account for the orthographical variations that the attachment of clitic pronouns causes in base forms of verbs (these variations have been explained in Section 2.1.2). These new non-terminals are the following:

```
VerbSt_clitized ( vtype, voice, tense, mood, clitization )
VerbSt_clitized ( vtype, voice, person, number, tense, mood,
                  clitization )
```

The first describes infinitive and gerunds forms to which one or two clitic pronouns will be added by the grammar, the second describes imperative forms to which one or two clitic pronouns will be added by the grammar. Again, for verb scheme clusters with some type of prepositional complement, there are also two parallel lexicon non-terminals with the necessary affix *linktype*.

```
VerbSt_clitized ( vtype, voice, linktype, tense, mood, clitization )
VerbSt_clitized ( vtype, voice, linktype, person, number, tense,
                  mood, clitization )
```

Examples of this kind of verb lexicon definitions are the following:

```
VerbSt_clitized ( SDO, ACT, PRESENT, INFINITIVE, TWO ): "amár".
VerbSt_clitized ( SDO, ACT, PRESENT, GERUND, ONE ): "amándo".
VerbSt_clitized ( SDO, ACT, SECOND, SING, PRESENT, IMPERATIVE, ONE ):
"áma".
```

⁷It must be remarked that, at the moment, this is an internal technical concept, not linguistically motivated, see Section 2.2.2.6. An study about possible meaning of verb scheme clusters, or about what requirements should be imposed on their generation from BDS to make them meaningful from the linguistic point of view, goes far beyond the limits of this exposition and will be developed in further research.

For each verb form cluster of the same morphological class of *amar*, 67 forms of these types are contained in the lexicon for each verb scheme cluster. More or less the same number of verb forms are produced for the other morphological classes, little disparities being caused—in non-defective verbs—by the presence of different syncretic forms in different classes (the disparity, for instance, caused by second and first/third conjugations with respect to first person plural, present and past perfect indicative forms, *amamos*, *partimos*, *comemos/comimos*).

Some semiauxiliary verbs force the introduction of an intermediate element connecting the semiauxiliary and the main verb. For these intermediate connectors, specific non-terminals and lexicon definitions have been designed:

Periphrastic Connector (linktype)

Examples of lexicon definitions of this type are the following:

Periphrastic Connector (QUE): "que".
 Periphrastic Connector (A): "a".

2.2.2 Affix Rules, Terminals and Non-terminals.

For all features handled by the grammar and related with the description of VPs, that is, for all affixes associated either with clitic pronouns or with verb forms, we will now enumerate and explain in detail their values, as well as how they are assigned to wordforms in the lexicon or, in the case of clitic pronouns, to terminals in the grammar.

2.2.2.1 Clitic Pronoun Position

This affix is exclusive of functional clitic pronouns. It specifies their position with respect to other clitic pronouns possibly included in the same VP.

C1Pr_Pos :: ONE_OF_ONE; ONE_OF_TWO; TWO_OF_TWO.

Value **ONE_OF_ONE** is assigned to clitic pronouns which can be combined alone with verb forms (*me*, *te*, *lo*, *la*, *le*, *nos*, *os*, *los*, *las*, *les*). Value **ONE_OF_TWO** is assigned to clitic pronouns which can be followed by another clitic pronoun within the VP (*me*, *te*, *se*, *nos*, *os*, *se*). Value **TWO_OF_TWO** is assigned to clitic pronouns which can be preceded by another clitic pronoun within the VP (*lo*, *la*, *le*, *los*, *las*, *les*).

2.2.2.2 Clitization

Affix rule for *clitization* is:

`clitization :: ONE; TWO; JUST_ONE.`

Verb forms assigned value **ONE** must be added one or two clitic pronouns, verb forms assigned value **TWO** must be added two clitic pronouns, verb forms assigned value **JUST_ONE** must be added only one clitic pronoun. To account for all the orthographical variations mentioned in Section 2.1.2, these values must be assigned to the following forms obtained from the corresponding base forms:

ONE

- Gerund forms.

```
VerbSt_clitized ( SDOIO, ACT, PRESENT, GERUND, ONE ):  
    "entregándo".  
VerbSt_clitized ( SDO, MIDD, PRESENT, GERUND, ONE ):  
    "quedándo".
```

- Unaccented, polysyllabic, second person singular present imperative forms.

```
VerbSt_clitized ( SDOIO, ACT, SECOND, SING, PRESENT,  
    IMPERATIVE, ONE ): "entréga".  
VerbSt_clitized ( SDO, MIDD, SECOND, SING, PRESENT,  
    IMPERATIVE, ONE ): "quéda".
```

- First person plural present subjunctive forms used for imperative in middle voice.

```
VerbSt_clitized ( SDO, MIDD, FIRST, PLU, PRESENT,  
    IMPERATIVE, ONE ): "quedémo".
```

- Unaccented, third conjugation, second person plural present imperative forms for middle voice.

```
VerbSt_clitized ( SDO, MIDD, SECOND, PLU, PRESENT,  
    IMPERATIVE, ONE ): "serví".
```

- Unaccented, polysyllabic, third person present subjunctive forms used for polite reference to second person imperative forms.

VerbSt_clitized (SDOIO, ACT, THIRD, SING, PRESENT,
 IMPERATIVE, ONE): "entrégue".
 VerbSt_clitized (SDO, MIDD, THIRD, SING, PRESENT,
 IMPERATIVE, ONE): "quéde".
 VerbSt_clitized (SDOIO, ACT, THIRD, PLU, PRESENT,
 IMPERATIVE, ONE): "entréguen".
 VerbSt_clitized (SDO, MIDD, THIRD, PLU, PRESENT,
 IMPERATIVE, ONE): "quéden".

TWO

- Unaccented, infinitive forms.

VerbSt_clitized (SDOIO, ACT, PRESENT, INFINITIVE,
 TWO): "entregár".
 VerbSt_clitized (SDO, MIDD, PRESENT, INFINITIVE,
 TWO): "quedár".

- Unaccented, monosyllabic, second person singular present imperative forms.

VerbSt_clitized (SDOIO, ACT, SECOND, SING, PRESENT,
 IMPERATIVE, TWO): "dá".
 VerbSt_clitized (SDO, MIDD, SECOND, SING, PRESENT,
 IMPERATIVE, TWO): "pón".

- First person plural present subjunctive forms used for imperative in active voice.

VerbSt_clitized (SDOIO, ACT, FIRST, PLU, PRESENT,
 IMPERATIVE, TWO): "entreguém".

- First and second conjugation, second person plural present imperative forms for middle voice, all second person plural present imperative forms for active voice.

VerbSt_clitized (SDOIO, ACT, SECOND, PLU, PRESENT,
 IMPERATIVE, TWO): "entregád"; "recogéd"; "decíd".
 VerbSt_clitized (SDO, MIDD, SECOND, PLU, PRESENT,
 IMPERATIVE, TWO): "quedá"; "poné".

- Unaccented, monosyllabic, third person present plural subjunctive forms used for polite reference to second person plural imperative forms.

VerbSt_clitized (SDOIO, ACT, THIRD, PLU, PRESENT,
 IMPERATIVE, TWO): "dén".
 VerbSt_clitized (SDO, MIDD, THIRD, PLU, PRESENT,
 IMPERATIVE, TWO): "dén".

JUST_ONE

- Polysyllabic forms with stress mark in the last syllable (*compón*) and monosyllabic forms with diacritic stress mark (*sé*), second person singular, present imperative forms for active and middle voices.

```
VerbSt_clitized ( SDOIO, ACT, SECOND, SING, PRESENT,  
                  IMPERATIVE, JUST_ONE ): "compon".
```

```
VerbSt_clitized ( S, MIDD, SECOND, SING, PRESENT,  
                  IMPERATIVE, JUST_ONE ): "compon".
```

- Monosyllabic forms with diacritic stress mark (*dé*), third person singular, present subjunctive forms used for polite reference to second person singular present imperative forms in active and middle voices.

```
VerbSt_clitized ( SDOIO, ACT, THIRD, SING, PRESENT,  
                  IMPERATIVE, JUST_ONE ): "de".
```

```
VerbSt_clitized ( SDO, MIDD, THIRD, SING, PRESENT,  
                  IMPERATIVE, JUST_ONE ): "de".
```

- First person plural present subjunctive forms used for imperative in active voice.

```
VerbSt_clitized ( SDOIO, ACT, FIRST, PLU, PRESENT,  
                  IMPERATIVE, JUST_ONE ): "entreguemos".
```

- First and second conjugation, second person plural present imperative forms for middle voice.

```
VerbSt_clitized ( S, MIDD, SECOND, PLU, PRESENT,  
                  IMPERATIVE, JUST_ONE ): "queda"; "pone".
```

All verb form clusters are extended with these forms in the lexicon, taking into account the combinability of voice with only some of them, that is, the scheme *entregar*, *SDOIO*, *active voice*, *personal* will not produce forms like those only exemplified for the verb *quedar* in the previous list —*quedémo*, *queda*, *quedá*, respectively followed by at least the pronoun *nos*, by just the pronoun *os* and by a combination of the pronoun *os* with a second one. Conversely, the scheme *quedar*, *SDO*, *middle voice*, *personal* will not produce forms like those only exemplified for the verb *entregar* in the previous list —*entreguemos*, *entreguemo*, *entregád*, respectively followed by at least one pronoun distinct from *nos* or *se*, by pronoun *se* and by any combination of two pronouns, the first one distinct from *os*.

Middle voice schemes block also the production of personal imperative forms not combined with enclitic pronouns: *queda, quede, quedemos, quedad, queden*, which, in fact, do not exist in middle voice. On the contrary, non-finite gerund base forms, which appear to be in a similar situation, must be produced, because they can constitute periphrastic VPs in which the non-functional clitic pronoun of the middle voice main verb has been moved to the periphrastic group: *se está quedando*.

For inherently impersonal active or middle voice schemes, given the fact that they can only be associated with **S**, **SDO** and **SPC** verb types, the generation of alternative forms requiring the addition of two clitic pronouns is also blocked.

For auxiliary verbs, the relation between voice and alternative forms is different. Except for the passive semiauxiliary *ser*, auxiliaries, *haber* and active or middle voice semiauxiliaries, do not determine the voice of the VP in which they are integrated.

For the auxiliary verb *ser*, like for main verbs, all the alternative forms necessary to account for the combination of passive voice with enclitic pronouns must be generated. In passive configurations only one indirect object clitic pronoun can be found, so only the alternative forms which appear in combination with one functional enclitic pronoun must be generated. Considering also that there is not such thing as an imperative mood in passive voice (see Section 2.2.2.3), the only⁸ alternative form that we need to produce for the passive auxiliary *ser* is the following⁹:

```
VerbSt_clitized ( PARTICIPLE, PASS, PRESENT, GERUND, ONE ): "siendo".
```

The auxiliary *haber* is combined with all the remaining types of auxiliaries and with main verbs for the production of compound tenses. Functional and non-functional clitic pronouns are attached to the auxiliary *haber* in infinitive and gerund compound tenses. For these cases, the following alternative forms of the auxiliary *haber* are introduced in the lexicon:

```
VerbSt_clitized ( AUX, ACT|MIDD|ip_ACT_IMP|ip_no_IMP, PRESENT,
  INFINITIVE, TWO ): "habér".
```

⁸The attachment of clitic pronouns to passive semiauxiliaries within the quite rare impersonal configuration *Se le puede ser entregado*, for instance, is not allowed, *Puede sérsele entregado*.*

⁹The terminal values for verb type and voice in auxiliaries will be explained in Section 2.2.2.5.

```
VerbSt_clitized ( AUX, PASS|ip_ACT_IMP|act|mid|ip_no_IMP,
PRESENT, GERUND, ONE ): "habiendo".
```

Similarly, for personal active voice semiauxiliaries, which can also attract both functional and non-functional clitic pronouns required by the main verb of the VP, all the orthographically alternative forms must be produced:

```
VerbSt_clitized ( GERUND, ip_no_IMP, PRESENT, INFINITIVE,
TWO ): "seguír".
VerbSt_clitized ( GERUND, ip_no_IMP, PRESENT, GERUND,
ONE ): "siguiendo".
VerbSt_clitized ( GERUND, ip_no_IMP, SECOND, SING, PRESENT,
IMPERATIVE, ONE ): "sígue".
VerbSt_clitized ( GERUND, ip_no_IMP, THIRD, SING, PRESENT,
IMPERATIVE, ONE ): "siga".
VerbSt_clitized ( GERUND, ip_ACT, FIRST, PLU, PRESENT,
IMPERATIVE, JUST_ONE ): "sigámos".
VerbSt_clitized ( GERUND, ip_ACT, FIRST, PLU, PRESENT,
IMPERATIVE, TWO ): "sigámo".
VerbSt_clitized ( GERUND, ip_MIDD, FIRST, PLU, PRESENT,
IMPERATIVE, ONE ): "sigámo".
VerbSt_clitized ( GERUND, ip_ACT, SECOND, PLU, PRESENT,
IMPERATIVE, TWO ): "seguíd".
VerbSt_clitized ( GERUND, ip_MIDD, SECOND, PLU, PRESENT,
IMPERATIVE, ONE ): "seguí".
VerbSt_clitized ( GERUND, ip_no_IMP, THIRD, PLU, PRESENT,
IMPERATIVE, ONE ): "sigan".
```

The impersonal active voice semiauxiliary *haber que* might be attached functional clitic pronouns required by the main verb, the following alternative forms are produced for this case:

```
VerbSt_clitized ( INFINITIVE, ip_ACT_IMP, QUE, PRESENT,
INFINITIVE, TWO ): "habér".
VerbSt_clitized ( INFINITIVE, ip_ACT_IMP, QUE, PRESENT,
GERUND, ONE ): "habiendo".
```

Middle voice semiauxiliaries, finally, although they can modify active or middle voice main verbs, they cannot be attached the clitic pronouns required by those. For these verbs, only the alternative forms that appear in middle voice in combination with one enclitic pronoun —the non-functional clitic pronoun required by the middle voice semiauxiliary itself— are produced. This means, on the other hand, that middle voice semiauxiliaries block the generation of regular imperative verb

forms, because these cannot be combined with clitic pronouns (*pongase**). Regular present gerund forms, however, must be produced to account for the cases in which the clitic pronoun is attached to a previous active voice semiauxiliary (*me estaba poniendo a trabajar*).

```
VerbSt_clitized ( INFINITIVE, MIDD, A, PRESENT, GERUND,
ONE ): "poniéndo".
VerbSt_clitized ( INFINITIVE, MIDD, A, SECOND, SING, PRESENT,
IMPERATIVE, JUST_ONE ): "pon".
VerbSt_clitized ( INFINITIVE, MIDD, A, THIRD, SING, PRESENT,
IMPERATIVE, ONE ): "pónga".
VerbSt_clitized ( INFINITIVE, MIDD, A, FIRST, PLU, PRESENT,
IMPERATIVE, ONE ): "pongámo".
VerbSt__clitized ( INFINITIVE, MIDD, A, SECOND, PLU, PRESENT,
IMPERATIVE, JUST_ONE ): "pone".
VerbSt_clitized ( INFINITIVE, MIDD, A, THIRD, PLU, PRESENT,
IMPERATIVE, ONE ): "póngan".
```

The introduction of alternative forms in the lexicon for the verb forms that show orthographical variations forces the introduction of similar second descriptions for verbs that do not undergo such variations. For these verbs, the same lexical verb form is introduced by two different lexicon definitions with different lexicon non-terminals on the right-hand side of the rule. This is the situation in the following cases:

ONE

- Active and middle voice —we will include only examples for active voice—, second person singular present imperative forms: polysyllables having stress mark in the in the penultimate syllable —because of hiatus— (*guía*).

```
VerbSt_clitized ( SDO, ACT, SECOND, SING, PRESENT,
IMPERATIVE, ONE ): "guía".
```

- Active and middle voice, third person singular and plural present subjunctive forms used for polite reference to second person present imperative forms: polysyllables having stress mark in the penultimate syllable —because of hiatus— (*gué, guén*).

```
VerbSt_clitized ( SDO, ACT, THIRD, SING, PRESENT,
IMPERATIVE, ONE ): "gué".
```

VerbSt_clitized (SDO, ACT, THIRD, PLU, PRESENT,
IMPERATIVE, ONE): "guien".

TWO

- Active and middle voice, present infinitive forms: polysyllables having stress mark in the penultimate syllable —because of hiatus— (*oír*).

VerbSt_clitized (SDO, ACT, PRESENT, INFINITIVE, TWO):
"oír".

- Active and middle voice —we will include only examples for active voice—, second person singular present imperative forms: polysyllables having stress mark in the last syllable (*compón*) and monosyllables with diacritic stress mark (*sé*).

VerbSt_clitized (SDOIO, ACT, SECOND, SING, PRESENT,
IMPERATIVE, TWO): "compón".

- Active and middle voice —we will include only examples for active voice—, third person singular present subjunctive forms used for polite reference to second person singular present imperative forms: monosyllables with diacritic stress mark (*dé*).

VerbSt_clitized (SDOIO, ACT, THIRD, SING, PRESENT,
IMPERATIVE, TWO): "dé".

- Active voice, second person plural present imperative forms: polysyllables having stress mark in the penultimate syllable —because of hiatus— (*oíd*).

VerbSt_clitized (SDO, ACT, SECOND, PLU, PRESENT,
IMPERATIVE, TWO): "oíd".

JUST_ONE

- Active and middle voice, monosyllabic, second person singular present imperative forms: "da".

VerbSt_clitized (SDO, ACT, SECOND, SING, PRESENT,
IMPERATIVE, JUST_ONE): "da".

- Middle voice, second person plural present imperative of verb *ir*:

```
VerbSt_clitized ( S, MIDD, SECOND, PLU, PRESENT,  
IMPERATIVE, JUST_ONE ): "id".
```

- Active and middle voice, monosyllabic, third person plural present subjunctive forms used for polite reference to second person singular present imperative forms: "den".

```
VerbSt_clitized ( SDO, ACT, THIRD, PLU, PRESENT,  
IMPERATIVE, JUST_ONE ): "den".
```

The reason for these duplications is the following: having two different lexicon non-terminals for the description of two different lexical verb forms (*entrega/entrega*), both of them covering one and the same inflectional verb form (*second person, singular, present, imperative*) in two different contexts (with or without an enclitic personal pronoun), would allow the identification of forms in the wrong context when one of these non-terminals —normally, the basic one— is also used to describe one lexical verb form (*guía*), covering the same inflectional verb form (*second person, singular, present, imperative*), but in both contexts. That is, when describing in the grammar *guíalo, dale*, if *guía, da* and *entrega* are all described by the same non-terminal in the lexicon, we would also allow the identification of *entregalo**, *entregale**. Conversely, the grammar descriptions of *entregalo, entregale* would not identify *guíalo, dale*.

To solve this problem, the introduction of different non-terminals for describing the same inflectional verb form is not practical. First, from the point of view of the grammar, it contradicts the general principle of *one inflectional verb form in certain distribution/one lexicon non-terminal*, forcing the same inflectional verb form in the same distribution to be always called upon within the grammar by two different non-terminals. And second, the design of the lexicon would be more complicated with more different non-terminals or affixes associated to non-inflectional information. On the contrary, while the increase of the lexicon caused by these duplications is negligible, by having two different lexicon descriptions —those used for *entrega/entrega*— for the same lexical verb form *guía*, we assign only two different non-terminals —instead of three, one for *guía*, one for *entrega* and one for *entrega*— to refer, respectively, to *second person, singular, present, imperative* and to *second person, singular, present, imperative* when (en)clitic pronouns are attached to it.

2.2.2.3 Mood and Tense

Affix rules for *mood* and *tense* are the following:

MOOD

```
mood :: non_imperative; IMPERATIVE; non_finite_mood.  
non_imperative :: INDICATIVE; SUBJUNCTIVE.  
non_finite_mood :: PARTICIPLE; GERUND; INFINITIVE.
```

TENSE

```
tense :: simple; compound.  
simple :: PRESENT; PAST; IMPERFECT; FUTURE; CONDITIONAL.  
compound :: PRESENT_PERFECT; PAST_PERFECT; PLUPERFECT;  
            FUTURE_PERFECT; PAST_CONDITIONAL.
```

Terminal values contained in these rules are assigned to the following verb forms:

- ⇒ **PRESENT, INDICATIVE:** *corro*
- ⇒ **PAST, INDICATIVE:** *corrí*
- ⇒ **IMPERFECT, INDICATIVE:** *corría*
- ⇒ **FUTURE, INDICATIVE:** *correré*
- ⇒ **CONDITIONAL, INDICATIVE:** *correría*
- ⇒ **PRESENT_PERFECT, INDICATIVE:** *he corrido*
- ⇒ **PAST_PERFECT, INDICATIVE:** *hube corrido*
- ⇒ **PLUPERFECT, INDICATIVE:** *había corrido*
- ⇒ **FUTURE_PERFECT, INDICATIVE:** *habré corrido*
- ⇒ **PAST_CONDITIONAL, INDICATIVE:** *habría corrido*
- ⇒ **PRESENT, SUBJUNCTIVE:** *corra*
- ⇒ **IMPERFECT, SUBJUNCTIVE:** *corriera, corriese*
- ⇒ **FUTURE, SUBJUNCTIVE:** *corriere*
- ⇒ **PRESENT_PERFECT, SUBJUNCTIVE:** *haya corrido*
- ⇒ **PLUPERFECT, SUBJUNCTIVE:** *hubiera, hubiese corrido*
- ⇒ **FUTURE_PERFECT, INDICATIVE:** *hubiere corrido*
- ⇒ **PRESENT, IMPERATIVE:** *corre*

- ⇒ **PRESENT, INFINITIVE**: *correr*
- ⇒ **PRESENT_PERFECT, INFINITIVE**: *haber corrido*
- ⇒ **PRESENT, GERUND**: *corriendo*
- ⇒ **PRESENT_PERFECT, GERUND**: *habiendo corrido*
- ⇒ **PARTICIPLE**: *corrido*

Tense values in the domain of non-terminal *simple* are assigned to verb forms in the lexicon, those in the domain of non-terminal *compound* are always assigned by rules of the grammar.

The syncretism, for first and third conjugation, of inflectional verb forms first person plural, present and past indicative (*amamos, partimos*) produces one lexicon definition (see Section 2.2.1, page 72). The syncretism of inflectional verb forms third person singular, present indicative, and second person singular, present imperative (*ama*) produces two lexicon definitions (see Section 2.2.1, page 72). On the contrary, the use of third person present subjunctive forms for polite reference to second person present imperative (*ame*) produces two lexicon definitions, one of them with the value **SUBJUNCTIVE** and the other with the value **IMPERATIVE**. These two lexicon definitions were necessary in this case because in subjunctive mood, but not in imperative mood, the value for person can also be first:

```
VerbSt ( SDOIO, ACT, FIRST|THIRD, SING, PRESENT, SUBJUNCTIVE ):
    "entregue".
VerbSt ( SDOIO, ACT, THIRD, SING, PRESENT, IMPERATIVE ): "entregue".
```

In third and first person plural, we have instead only one lexicon definition:

```
VerbSt ( SDOIO, ACT, FIRST, PLU, PRESENT, SUBJUNCTIVE|IMPERATIVE ):
    "entreguemos".
VerbSt ( SDOIO, ACT, THIRD, PLU, PRESENT, SUBJUNCTIVE|IMPERATIVE ):
    "entreguen".
```

The assignment of **IMPERATIVE** mood to subjunctive forms used for affirmative imperative ensures, on the other hand, that one non-terminal in the grammar will be able to retrieve from the lexicon all imperative forms —both proper imperative forms in second person as well as subjunctive forms in third person and first person plural— of the verb paradigm. This seems also to be more

practical if we think of the fact that alternative forms produced to account for the attachment of clitic pronouns are associated solely with **IMPERATIVE** mood.

The same considerations, however, produced exactly the opposite effect for the use of subjunctive forms as negative imperative: no proper imperative forms are used in negative imperative—in second person, subjunctive forms cover also this function—and enclitic pronouns are not added to subjunctive verb forms in negative imperative VPs.

Summarizing all these explanations, full coverage of imperative mood for the verb scheme cluster *entregar*, *SDOIO*, *active voice*, *personal*, requires the following lexicon definitions:

- Affirmative commands:

```
VerbSt ( SDOIO, ACT, SECOND, SING, PRESENT, IMPERATIVE ):
    "entrega".
VerbSt ( SDOIO, ACT, THIRD, SING, PRESENT, IMPERATIVE ):
    "entregue".
VerbSt ( SDOIO, ACT, FIRST, PLU, PRESENT, SUBJUNCTIVE|
    IMPERATIVE ): "entreguemos".
VerbSt ( SDOIO, ACT, SECOND, PLU, PRESENT, IMPERATIVE ):
    "entregad".
VerbSt ( SDOIO, ACT, THIRD, PLU, PRESENT, SUBJUNCTIVE|
    IMPERATIVE ): "entreguen".
VerbSt_clitized ( SDOIO, ACT, SECOND, SING, PRESENT,
    IMPERATIVE, ONE ): "entréga".
VerbSt_clitized ( SDOIO, ACT, THIRD, SING, PRESENT,
    IMPERATIVE, ONE ): "entrégue".
VerbSt_clitized ( SDOIO, ACT, FIRST, PLU, PRESENT,
    IMPERATIVE, JUST_ONE ): "entreguemos".
VerbSt_clitized ( SDOIO, ACT, FIRST, PLU, PRESENT,
    IMPERATIVE, TWO ): "entreguemo".
VerbSt_clitized ( SDOIO, ACT, SECOND, PLU, PRESENT,
    IMPERATIVE, TWO ): "entregád".
VerbSt_clitized ( SDOIO, ACT, THIRD, PLU, PRESENT,
    IMPERATIVE, ONE ): "entréguen".
```

- Negative commands:

```
VerbSt ( SDOIO, ACT, SECOND, SING, PRESENT, SUBJUNCTIVE ):
    "entregues".
VerbSt ( SDOIO, ACT, FIRST|THIRD, SING, PRESENT,
    SUBJUNCTIVE ): "entregue".
VerbSt ( SDOIO, ACT, FIRST, PLU, PRESENT, SUBJUNCTIVE|
    IMPERATIVE ): "entreguemos".
```

```

VerbSt ( SDOIO, ACT, SECOND, PLU, PRESENT, SUBJUNCTIVE ):
    "entreguéis".
VerbSt ( SDOIO, ACT, THIRD, PLU, PRESENT, SUBJUNCTIVE |
    IMPERATIVE ): "entreguen".

```

The nested structure of the rule for affix non-terminal *tense* is self-explanatory. It reflects the different structure of simple and compound tenses, the latter are produced within the grammar by combining the forms of the auxiliary *haber* with the participle of the verb under consideration.

The nested structure of the rule for affix non-terminal *mood* is caused by the use of mood beyond the level of VP, at the level of the clause. At clause level analysis, the description of clauses was divided into different rules for different types of clauses: independent or *that* clauses, interrogative and relative finite clauses, interrogative and relative non-finite clauses, exhortative clauses, infinitive or gerund clauses and participle clauses. The structure of the rule for affix non-terminal *mood* optimizes the relation between mood and type of clause from the point of view of clause level formal description. Independent or *that* clauses and interrogative or relative finite clauses select values in the domain of *non_imperative*. Exhortative clauses select value **IMPERATIVE**. Infinitive or gerund clauses and interrogative or relative non-finite clauses select values in the domain of *non_finite_mood*. Participle clauses, finally, select value **PARTICIPLE**.

All tenses and moods are produced for main verbs. For participles, four inflected forms —for gender and number variations— are produced for each verb scheme cluster.

For auxiliaries, as expected, the situation is different, because, as these constitute a closed class of words, we can individually decide which tenses and moods must be produced for each auxiliary. For semiauxiliaries, forms for all moods and tenses are, in general, produced, only imperative mood is blocked in certain cases. With respect to participles, only forms inflected for masculine, singular —those which constitute compound tenses— are produced. Obviously, auxiliary *haber* has neither **PARTICIPLE** nor **IMPERATIVE** mood.

Impersonal verbs, both main and semiauxiliaries, have not **IMPERATIVE** mood.

2.2.2.4 Person, Gender, Number and Case

Affix rules for *person*, *gender*, *number* and *case* are the following:

PERSON

```
inherent person :: FOURTH; FIFTH; SIXTH; MVperson.  
MVperson :: person.  
person_one :: person_two.  
person_two :: person.  
person :: FIRST; SECOND; THIRD; NIL.
```

GENDER

```
gender_one :: gender_two.  
gender_two :: gender.  
gender :: NEUT; MASC; FEM; A_FEM; FEM_UN; NIL.
```

NUMBER

```
number_one :: number_two.  
number_two :: number.  
number :: SING; PLU; NIL.
```

CASE

```
case :: NOM; PREP; case_two.  
case_two :: case_one.  
case_one :: function_case; VERB.  
function_case :: ACC; DAT; PRTVE.
```

Terminal values for affixes *person*, *gender*, *number* and *case* are in general self-explanatory. Values **FOURTH**, **FIFTH** and **SIXTH** are used in possessive determiners and pronouns. Values **A_FEM** and **FEM_UN** are required by nouns beginning by [á-] and will be explained in Section 3.2.2.10. **VERB** is the case value assigned to clitic pronouns required by middle voice and impersonality. **NOM**, **ACC** and **DAT** refer to subject, direct object and indirect object functions. Traditional terminology does not provide a case name for prepositional complement and predicative complement functions, we use, respectively, **PREP** and **PRTVE** for these. **NOM** and **PREP** are restricted to non-clitic personal pronouns. **VERB**, **ACC**, **DAT** and **PRTVE**, which are the values directly related to VPs, are restricted to clitic personal pronouns.

Within VPs, *person*, *gender* and *number* values are assigned in the lexicon when associated with verb forms, they are assigned in the grammar when associated with

clitic pronouns. *Case* values are always assigned in the grammar, because they are always associated with clitic pronouns.

Syncretism of **FIRST** and **THIRD** person *SING* in **IMPERFECT** and **CONDITIONAL** tenses for **INDICATIVE** mood, and in **PRESENT**, **IMPERFECT** and **FUTURE** tenses for **SUBJUNCTIVE** mood, produces one lexicon definition (see Section 2.2.1, page 72). The use of third person singular to refer politely to second person, is assigned value **THIRD** for *person*. Syncretism phenomena related to *gender*, *number* or *case* in clitic pronouns will be explained in Section 2.2.2.1.

Nested structures of *person*, *number*, *gender* and *case* share one peculiarity: the progressive nesting of non-terminals with suffixes *_one*, *_two*. This structure is caused by the formalization of clitic pronouns as constituents of the VP. This will be described in Section 2.2.3.1.

Values immediately dominated by non-terminal *inherent person* are necessary for possessive pronouns (see Section 3.2.2.10). Non-terminal *MVperson* is useful beyond the level of the phrase, at the level of the clause, where this affix keeps track of the person of the verb of main clauses. This information is after used to disambiguate voice in subordinated infinitive or gerund clauses.

Within the subdomain of *case* that affects clitic pronouns, *function_case* separates the case values that refer to syntactic functions from those that indicate middle voice or impersonality.

All main verbs and auxiliaries are provided with all the verb forms necessary to cover all the different combinations of *person*, *gender* and *number*. Only inherently impersonal main and semiauxiliary verbs are restricted to third person singular inflected verb forms.

2.2.2.5 Verb type, Voice and Preptype

Affix rules for *vtype* and *voice* are the following:

VERB TYPE

```
vtype :: AUX; auxiliabile.  
auxiliabile :: semiaux; mvtype.  
semiaux :: non_finite_mood.  
mvtype :: S; SPC; one_clitic.  
one_clitic :: SDO; SI0; SPR; SDOPC; SDOPR; SIOPC; two_clitics.
```

two_clitics :: SDOIO; SIOPR; SDOIOPC.

VOICE

voice :: pass; no_pass; ACT_IMP+PASS.
pass :: PASS; PASS_IMP.
no_pass :: act; midd; improper_voice.
act :: ACT; ACT_IMP.
midd :: MIDD; MIDD_IMP.
improper_voice :: ip_no_IMP; ip_IMP.
ip_no_IMP :: ip_ACT; ip_MIDD.
ip_IMP :: ip_ACT_IMP; ip_MIDD_IMP.

Terminal values in the domain of *vtype* have the following meaning:

AUX Auxiliary verb *haber*

PARTICIPLE Semiauxiliary verb *ser* for passive voice, it requires the main verb to be a participle

INFINITIVE Active or middle voice semiauxiliaries that require the following verb to be an infinitive

GERUND Active or middle voice semiauxiliaries that require the following verb to be a gerund

S Main verb that subcategorizes SUBJECT

SPC Main verb that subcategorizes SUBJECT and PREPOSITIONAL COMPLEMENT

SDO Main verb that subcategorizes SUBJECT and DIRECT OBJECT

SIO Main verb that subcategorizes SUBJECT and INDIRECT OBJECT

SPR Main verb that subcategorizes SUBJECT and PREDICATIVE COMPLEMENT

SDOPC Main verb that subcategorizes SUBJECT, DIRECT OBJECT and PREPOSITIONAL COMPLEMENT

SDOPR Main verb that subcategorizes SUBJECT, DIRECT OBJECT and PREDICATIVE COMPLEMENT

SIOPC Main verb that subcategorizes SUBJECT, INDIRECT OBJECT and PREPOSITIONAL COMPLEMENT

SDOIO Main verb that subcategorizes SUBJECT, DIRECT OBJECT and INDIRECT OBJECT

SIOPR Main verb that subcategorizes SUBJECT, INDIRECT OBJECT and PREDICATIVE COMPLEMENT

SDOIOPC Main verb that subcategorizes SUBJECT, DIRECT OBJECT, INDIRECT OBJECT and PREPOSITIONAL COMPLEMENT

All these values are assigned to verb forms in the lexicon. The nested structure of affix *vtype* is self-explanatory. Immediately under the dominance of affix non-terminal *vtype* the **AUX** terminal value identifies the auxiliary verb *haber*, which can never be modified by any auxiliary or semiauxiliary. The affix non-terminal *auxiliable* subsequently refers to all the remaining verbs, which, on the contrary, can be modified by different types of auxiliary verbs. Immediately under the dominance of *auxiliable*, we find non-terminals *semiauxiliary* and *mvtype* (main verb type). The former refers to the passive semiauxiliary *ser* and active or middle voice periphrastic semiauxiliaries. The latter refers to the verbs that supply VPs with lexical meaning. All the verbs of these groups can in turn be modified either by verb *haber* or by other active or middle voice semiauxiliaries. Semiauxiliaries are subclassified taking into account their requirements with respect to the mood of the verb modified: **PARTICIPLE** (passive semiauxiliary *ser*), **INFINITIVE** or **GERUND** (active and middle voice semiauxiliaries). Main verbs are subclassified according to the number and type of arguments that they subcategorize. Immediately under the dominance of affix non-terminal *mvtype*, we find the terminal values associated with combinations of verbs and arguments not allowing the substitution of any argument by a clitic pronoun. Together with them, non-terminal *mvtype* dominates the non-terminal *one_clitic*, which, in turn, dominates non-terminal *two_clitics* and terminal values associated with verb-argument combinations allowing the substitution of only one argument by a clitic pronoun. Non-terminal *two_clitics*, finally, dominates terminal values associated with verb-argument combinations allowing the substitution of two arguments by clitic pronouns.

The affix *voice*, on the other hand, is designed to keep track simultaneously of all the information about both *voice* and *impersonality*. Besides, this affix must also provide an adequate coverage of voice for both non-periphrastic and periphrastic VPs. Terminal values in the domain of *voice* have the following meaning:

ACT_IMP+PASS Active voice, impersonal, passive meaning
PASS Passive voice, personal
PASS_IMP Passive voice, impersonal
ACT Active voice, personal
ACT_IMP Active voice, impersonal
MIDD Middle voice, personal
MIDD_IMP Middle voice, impersonal
ip_ACT Improper active voice, personal
ip_MIDD Improper middle voice, personal
ip_ACT_IMP Improper active voice, impersonal
ip_MIDD_IMP Improper middle voice, impersonal

The three values, **ACT_IMP+PASS**, **PASS_IMP** and **ip_MIDD_IMP** are exclusive of the grammar. In the lexicon, the value **PASS** is only assigned to the semiauxiliary *ser*. The remaining values are either assigned to verb forms in the lexicon or acquired in the grammar. Improper values are exclusive of semiauxiliaries in the lexicon, and of periphrastic groups in the grammar.

The nested structure of affix *voice* is motivated by the use of *voice* information beyond the level of the VP, at the level of the clause. *Voice* immediately dominates the terminal value **ACT_IMP+PASS**. According to the above definition of this terminal, this is the voice value assigned to active voice and *se* impersonal, with subject concordance, VPs (*se vende/venden*). Together with this terminal value, *voice* immediately dominates non-terminals *pass* and *no-pass*, the first obviously covering passive voice values, the second the remaining ones. With these three distinctions, the basic groupings of verb forms according to voice for the purposes of clause description are already identified, since, on the one hand, the same combinations of arguments are, at least theoretically, valid for active and middle voice VPs, while, on the other hand, passive voice requires a specific subset of them —as well as the introduction of a new argument, AGENT, which cannot be integrated in active or middle voice clauses.

Non-terminal *pass* subsequently dominates values **PASS** and **PASS_IMP**, respectively for passive voice personal and impersonal VPs. **PASS** is the voice value assigned to personal VPs that include the passive voice semiauxiliary *ser*

(*es entregado*). **PASS_IMP** is the voice value assigned to passive voice VPs that include the impersonal clitic pronoun *se* or the impersonal periphrasis *haber que* (*se es recibido, hay que haber sido tratado*).

Non-terminal *no_pass* dominates instead non-terminals *act*, *midd* and *improper_voice*. Values in the domains of *act* and *midd* non-terminals cover respectively active and middle voice VPs. In addition to this, within periphrastic VPs, these values are also assigned to periphrastic groups not containing clitic pronouns selected by the main verb, while, on the contrary, the values in the domain of *improper_voice* are those —*exclusively*— assigned to periphrastic groups when these contain clitic pronouns selected by the main verb. As a logical consequence of this, terminal values in the domain of *improper_voice* are never assigned to VPs: they are only used below this level of analysis to describe i) combinations of active voice semiauxiliaries and clitic pronouns required by the main verb, and ii) combinations of the resulting periphrastic groups and main verbs. In contrast to values in the domain of *improper_voice*, the term *proper* voice will be used to refer to the whole set of values out of the *improper_voice* subdomain, that is, *proper* voice will be used to refer to *voice* values which can be assigned to VPs and PREDICATE functions.

Non-terminal *act* dominates terminal values **ACT** and **ACT_IMP**, respectively for active voice personal and impersonal VPs and periphrastic groups. **ACT** is the value assigned to VPs with an inflected main verb and, possibly, one or two functional clitic pronouns (*escribo, puedo estar escribiendo, le escribo, le puedo estar escribiendo*), or to periphrastic groups with active voice semiauxiliaries and without clitic pronouns (*puedo estar escribiendo*). **ACT_IMP** is the value assigned to active voice VPs that include an inherently active voice impersonal main verb and the impersonal clitic pronoun *se* or the impersonal periphrasis *haber que* (*hay, se escribe, hay que escribir*), or to active voice periphrastic groups that include the impersonal semiauxiliary *haber que* and no clitic pronouns (*puede haber que escribir*).

Similarly, non-terminal *midd* dominates terminal values **MIDD** and **MIDD_IMP**, respectively for middle voice personal and impersonal VPs and periphrastic groups. **MIDD** is the value assigned to VPs with an inflected main verb and one non-functional clitic pronoun echoing *person* and *number* of the VP (*me atrevo,*

puedo atreverme, me lo como, puedo comérmelo), or to periphrastic groups with a middle voice semiauxiliary (*puedo ponerme a escribir*). **MIDD_IMP** is the value assigned to middle voice VPs that include an inherently middle voice impersonal main verb or the impersonal periphrasis *haber que* (*se trata, hay que vestirse*), or to middle voice periphrastic groups that include the impersonal semiauxiliary *haber que* (*hay que ponerse a trabajar*).

So far, the nested structure of affix voice shows that in our approach voice oppositions have taken priority over impersonality oppositions. This responds to the following two facts: on the one hand, we find that different behaviours with respect to clitic pronouns require voice based grouping of terminal values —for instance, the appearance of two functional clitic pronouns is only possible in active voice—, and, on the other, we also find that developments of clause description rules seem to be better focused on voice based grouping of alternatives —preliminary data about frequencies of middle and active voice for different argument combinations appear to be quite significant in this regard.

Finally, the last non-terminal immediately under the dominance of *no-pass*, *improper_voice*, dominates non-terminals *ip_no_IMP* and *ip_IMP*. The first refers to personal improper voice and the second to impersonal improper voice. Each in turn dominates terminal values for improper active and middle personal and impersonal voices, *ip_no_IMP* dominates **ip_ACT** and **ip_MIDD**, and *ip_IMP* dominates **ip_ACT_IMP** and **ip_MIDD_IMP**. **ip_ACT** is the voice value assigned to periphrastic groups that contain exclusively functional clitic pronouns (*me lo pudo haber dicho*). **ip_MIDD** is the voice value assigned to periphrastic groups that contain one non-functional clitic pronoun echoing *person* and *number* of the VP (*se puede vestir*). **ip_ACT_IMP** is the voice value assigned to improper active voice periphrastic groups that include the impersonal semiauxiliary *haber que* (*hay que írselo diciendo*). **ip_MIDD_IMP** is the voice value assigned to improper middle voice periphrastic groups that include the impersonal semiauxiliary *haber que* (*hay que irse preparando*).

Voice and verb types values of VPs maintain very important relationships. While voice values **ACT_IMP+PASS**, **PASS** and **PASS_IMP** can only be combined with verb types that include a DIRECT OBJECT, that is, **SDO**, **SDOIO**, **SDOPC**, **SDOPR** and **SDOIOPC**, voice values **ACT**, **MIDD**, **ACT_IMP**

and **MIDD_IMP** can be combined with any verb type. On the basis of, on the one hand, these relations between *voice* and *vtype* and, on the other, the relations between *voice* and clause level analysis, we can state the following principles regarding the formalization of clauses:

1. PREDICATE functions contained in clause alternatives that describe combinations of syntactic functions found in passive voice and that include an explicit SUBJECT function, are assigned value **PASS** for *voice*, so they call upon **PASS** VPs. Verb type of the VP should include a DIRECT OBJECT function.
2. PREDICATE functions contained in clause alternatives that describe combinations of syntactic functions found in passive voice and that do not include an explicit SUBJECT function, are assigned non-terminal *pass* for *voice*, so they call upon **PASS** and **PASS_IMP** VPs. Verb type of the VP should include a DIRECT OBJECT function.
3. PREDICATE functions contained in clause alternatives that describe combinations of syntactic functions found both in active and in middle voice and that include an explicit SUBJECT function, are assigned values **ACT|MIDD** for *voice*, so they call upon **ACT** and **MIDD** VPs.
4. PREDICATE functions contained in clause alternatives that describe combinations of syntactic functions found both in active and in middle voice and do not include an explicit SUBJECT function, are assigned non-terminal *no_pass*¹⁰ for *voice*, so they call upon **ACT**, **MIDD**, **ACT_IMP** and **MIDD_IMP** VPs.
5. PREDICATE functions contained in clause alternatives that describe combinations of syntactic functions found in active voice and that include an explicit SUBJECT function, are assigned value **ACT** for *voice*, so they call upon **ACT** VPs.

¹⁰We have already indicated, and the structure of the grammar, see Section 2.2.3, will further demonstrate this, that values in the domain of *improper_voice* cannot be assigned to VPs, so *no_pass* VPs will never call upon *improper_voice* VPs.

6. PREDICATE functions contained in clause alternatives that describe combinations of syntactic functions found in active voice and that do not include an explicit SUBJECT function, are assigned non-terminal *act* for *voice*, so they call upon **ACT** and **ACT_IMP** VPs.
7. PREDICATE functions contained in clause alternatives that describe combinations of syntactic functions found in middle voice and that include an explicit SUBJECT function, are assigned value **MIDD** for *voice*, so they call upon **MIDD** VPs.
8. PREDICATE functions contained in clause alternatives that describe combinations of syntactic functions found in middle voice and that do not include an explicit SUBJECT function, are assigned value *midd* for *voice*, so they call upon **MIDD** and **MIDD_IMP** VPs.
9. Every alternative that complies with points 4 and 6—that is, allowing an **ACT_IMP** voice value—and includes a DIRECT OBJECT syntactic function, should be preceded by an equal alternative, except for the fact that the DIRECT OBJECT is turned into the SUBJECT and the PREDICATE is assigned value **ACT_IMP+PASS** for *voice*. Obviously, verb type of the VP should include a DIRECT OBJECT function.

The affix rule that encodes the values that represent prepositional requirements is the following:

LINKTYPE

```

linktype :: preptype; cjtype.
preptype :: ADV; MODDE; NIL; preposition.
preposition :: DE; no_DE.
no_DE :: A; ANTE; BAJO; CON; CONTRA; DESDE; EN; ENTRE; HACIA;
        HASTA; PARA; POR; SIN; SOBRE; TRAS; OTHER.
cjtype :: COMO; QUE; SI; NIL.

```

Although we know that the exhaustive identification and classification for Natural Language Processing of prepositions and conjunctions is a very important and necessary task, this goes far beyond the limits of this work, and, as can be observed in the classification of linking elements included in our grammar, formally introduced above, we here restrict ourselves to prepositions and conjunctions that play

a role in the phrase level of analysis as accounted for by the proposed grammar. Description of prepositions is thus deliberately partial, covering only aspects of phrase level of analysis. The same holds for the description of conjunctions, which is even more remarkably reduced, since conjunctions are less relevant for phrase analysis than prepositions. In addition to this, we will in this Section restrict explanations about prepositions and conjunctions to just the elements concerning the VP. Further considerations about these categories of words will be introduced later, in Sections devoted to different linguistic structures.

Despite the fact that from the point of view of prepositional requirements of main verbs, values for affix *preptype* are self-explanatory, the case of semiauxiliaries would instead benefit from some additional explanations. As observed above, certain semiauxiliaries introduce the following verb by means of an intermediate connector which is in fact a particular property of the semiauxiliary verb in question. These intermediate connectors, though in their more frequent distributions are prepositions or conjunctions, as periphrastic connectors for semiauxiliaries, they have additional lexicon definitions (see Section 2.2.1). In these lexicon definitions, as well as in lexicon definitions of semiauxiliary verbs, values in the domain of non-terminal *linktype*, simultaneously covering values relevant for prepositions and conjunctions, are specified for their later unification in the grammar.

Possible values for affix *linktype* in semiauxiliaries and periphrastic connectors are obviously only those needed to cover the intermediate elements recognized as valid in Table A.1. The following six lexicon definitions of periphrastic connectors are sufficient to account for the requirements of all semiauxiliaries in this respect:

```
Periphrastic Connector ( A ): "a".
Periphrastic Connector ( DE ): "de".
Periphrastic Connector ( EN ): "en".
Periphrastic Connector ( POR ): "por".
Periphrastic Connector ( MODDE ): "a punto de"; "en un tris de";
    "a pique de"; "a un pelo de".
Periphrastic Connector ( QUE ): "que".
```

On the other hand, main verb prepositional requirements refer to the prepositions that more frequently introduce the prepositional complements possibly included in the scheme of the verb. The values specified by main verbs to account for prepositional requirements are those contained in the domain of non-terminal

preptype. The value **ADV** is assigned to verbs that require a prepositional complement documented only as an adverb phrase in BDS¹¹. The values in the domain of affix non-terminal *preposition* distinguish the prepositions that often introduce prepositional complements¹².

2.2.2.6 Assignment of Verb Type and Voice Values to Verb Forms in the Lexicon. Verb Scheme Clusters Generation

The current Section is divided in two parts. While in the first part, we explain how values for *vtype*, *voice* and *preptype* are obtained and assigned to the lexicon definitions of the main verbs collected in the lexicon, in the second part, the assignment of values for these affixes to the verb forms of auxiliary verbs is addressed.

To begin with the first issue it is necessary to bring to light again the concepts of *scheme* and *configuration*. As explained in Section 2.1.1, both concepts refer to a conjunction of verb type, voice, impersonality and prepositional requirements, *scheme* as an abstract, inherent possibility of a verb and *configuration* as a concrete fact, actually found in linguistic discourse. Obviously, in a formal lexicon what we need to find is all the information about the schemes corresponding to each of the main verbs collected in the lexicon. Since in our model, *mvtype* encodes the available information about the verb type of the main verbs, *voice* about both their voice and impersonality, and *preptype* about their prepositional requirements, in formal terms, what we are going to explain in the first part of this Section is how *mvtype*, *voice* and *preptype* obtain their values for each *scheme* of each main verb collected in the lexicon.

In the next paragraphs, we will see that in our approach these values are extracted from BDS real data about the syntactic behaviour of verbs. Briefly described in Chapter 1, BDS contains the syntactic analysis of 150.000 clauses found in the ARTHUS corpus of Spanish. BDS was specifically designed to offer detailed and corpus based information about the syntactic structure of the Spanish clause, giving main relevance to data about verb subcategorization. The analysis of clauses

¹¹This is a temporary value, it should be manually revised in order to add a concrete preposition in case this is found relevant.

¹²Value **NIL** is not relevant for VPs. This value is associated with relative clauses (see Sections 3.2.3.4.1.3, 3.2.3.7.6 and 3.2.3.6.5)

was manually performed and corrected in several phases to ensure coherency and consistency. To store the analysis of clauses, a hierarchy of numerical keys associated with different types of information was designed which allows the user to obtain the information encoded for different degrees of detail.

The information encoded refers in first place to general characteristics of the clause under consideration. This general information is followed by a detailed study of the arguments of the main verb of the clause. Possible arguments are subject, direct object, indirect object, prepositional complements —two of them are accepted—, agent and predicative complement. The fields of information encoded for each clause are the following ones —Appendix B shows the analysis of one real clause in BDS:

1. General information about the clause:

- ⇒ Verb lemma constituting the PREDICATE function of the clause.
- ⇒ Location in the corpus: page and line number in the paper format original text.
- ⇒ Clause type: 15 different types of clauses are distinguished (among them: independent clause, coordinated clause, *that*-clause, infinitive clause, relative clause, nominalized relative clause, adverbial clause, gerund clause, participle clause, etc.).
- ⇒ Clause syntactic function of embedded clauses: 26 different syntactic functions are distinguished (among them: subject, direct object, predicative complement, indirect object, prepositional complement, circumstance, prepositional modifier, apposition, members of conditional, causative, concessive, comparative, consecutive or contrastive sentences, etc.).
- ⇒ Voice: active, middle, passive.
- ⇒ Polarity: affirmative, negative.
- ⇒ Modality: declarative, interrogative, exhortative, exclamative.
- ⇒ Type of periphrastic auxiliary: *acabar de* + INFINITIVE, *estar* + GERUND, etc., those collected in Table A.1.
- ⇒ Tense and mood of the PREDICATE function.
- ⇒ Tense and mood of the PREDICATE of a possible main clause.
- ⇒ Number and person of the PREDICATE function.

- ⇒ Multiword character: multiword verbs, such as *darse cuenta*, *tener que ver*, etc.
- ⇒ Number of arguments in the clause.
- ⇒ Order of arguments in the clause.

2. Specific information about arguments:

⇒ **Subject**

- Type of subject: explicit or implicit, generic, absent because of grammatical impersonality, absent because of inherent impersonal character of the verb, etc.

⇒ **Direct Object**

- Clitic pronoun: functional clitic pronoun echoing **person**, **gender** and **number** of the direct object.
- Presence/absence of preposition *a* introducing the direct object.

⇒ **Indirect Object**

- Clitic pronoun 1: functional clitic pronoun echoing **person** and **number** of the indirect object.
- Clitic pronoun 2: clitic pronoun of the type of *te* in *Te lo ha comido todo*.

⇒ **Prepositional Complement 1 and 2**

- Type: type of prepositional complement: prepositional object, adverbial, modal or none of these.
- Prepositions introducing them.

⇒ **Agent**

- Preposition introducing the agent.

⇒ **Predicative Complement**

- Preposition or conjunction introducing the predicative complement.
- Referent: subject, direct object, indirect object, prepositional object or none of these.

3. Information included for all arguments:

- ⇒ Type of the argument: present or absent, reported discourse, doubtful.
- ⇒ Type of syntactic category: 55 syntactic structures are distinguished (among them: noun phrase, different types of pronouns, different types of clauses, etc.).
- ⇒ Animation and countability of the entities underlying the syntactic functions.
- ⇒ Determination of the involved syntactic category: defined, undefined.
- ⇒ Number.

Crude BDS, therefore, offers deep and rich analyses of examples of verbs used in real context: directly, however, it does not illustrate the verb schemes associated with these verbs, but the final configurations in which they are contextualized. For this reason, the schemes associated with the verbs must indeed be derived from the configurations, since it is precisely this association between schemes and verb lemmas which constitutes the information that must be supplied to the lexicon. This information, on the other hand, should be extracted from BDS taking into account the necessities of the grammar designed, which means, in our case, that we had to reduce the amount of the information contained in BDS to conform to the possible values of affixes *voice* and *mvtype* as described in previous Section¹³.

This reduction was performed—or, in other words, information about schemes was derived from information about configurations—applying the following rules¹⁴:

1. All the arguments encoded in BDS analyses will be included in the schemes, except for second prepositional complements. SUBJECT and, sometimes, DIRECT OBJECT, are specifically addressed in the following rules. All the different types of prepositional complements in BDS are considered the same.

¹³Obviously, different theories about schemes, configurations and relationships between them would perform different reductions on BDS data. Our particular approach is primarily based on linguistic criteria, but it is also conditioned by the fact that we desired the process of extraction to be fully automatic. Other approaches, which perform further reductions or profit from certain relationships between schemes resulting from our approach, needed manual intervention to ensure the level of correction reached by this approach. Future refinements, however, might be easily integrated in the current design.

¹⁴Reductions are explained both in general and formal terms, that is, by means of both the BDS keys assigned to the fields involved in the reduction and the values that encode the data within the formal grammar. The programming tasks required by the extraction of the scheme clusters from BDS were carried out by Guillermo Rojo. I want to especially acknowledge his work on this matter.

In formal terms: the target scheme will include **DIRECT OBJECT**, **INDIRECT OBJECT**, **PREPOSITIONAL COMPLEMENT** and **PREDICATIVE COMPLEMENT** if fields 21 (type of direct object), 28 (type of indirect object), 35 (type of prepositional complement) and 55 (type of predicative complement) of BDS have keys 1 (explicit), 2 (reported statement) or * (somewhat doubtful). The key in field 36 (type of prepositional complement) is not taken into account.

2. Prepositions documented in BDS for prepositional complements are discarded when different from the traditional and more frequent ones. In this case, the scheme is assigned a general value for *preptype*. In formal terms: if one of the traditional or more frequent prepositions appears in field 37 (preposition required by prepositional complement), this value is stored in *preptype*, otherwise the affix is assigned one of the values **OTHER** or **ADV**.
3. Examples illustrating active voice personal configurations are reduced to active voice personal schemes whose verb type includes all the arguments specified in the analysis. All verb types are accepted. In formal terms: all the examples of BDS with key 1 (active voice) in field 6 (voice) and keys 1 (explicit), 2 (implicit), 30 (generic subjects of infinitives and gerunds), 31 (1st person plural with impersonal meaning), 32 (2nd person singular with impersonal meaning), 33 (3rd person plural with impersonal meaning), 34 (pronoun *uno* with impersonal meaning) or 5 (reported statement) in field 16 (type of the subject), are reduced to scheme *S, ACT* —where *vtype* is assigned a different verb type value if fields 21 (type of direct object), 28 (type of indirect object), 35 (type of prepositional complement) and 55 (type of predicative complement) of BDS have keys 1, 2 or * explained above.
4. Examples illustrating middle voice personal configurations are reduced to middle voice personal schemes whose verb type includes all the arguments specified in the analysis. All verb types are accepted. In formal terms: all the examples of BDS with key 2 (middle voice) in field 6 (voice) and keys 1 (explicit), 2 (implicit), 30 (generic subjects of infinitives and gerunds), 31 (1st person plural with impersonal meaning), 32 (2nd person singular with impersonal meaning), 33 (3rd person plural with impersonal meaning), 34

(pronoun *uno* with impersonal meaning) or 5 (reported statement) in field 16 (type of the subject), are reduced to scheme *S, MIDD* —where *vtype* is assigned a different verb type value if fields 21 (type of direct object), 28 (type of indirect object), 35 (type of prepositional complement) and 55 (type of predicative complement) of BDS have keys 1, 2 or *.

5. Examples illustrating passive voice personal or impersonal configurations are reduced to active voice personal schemes whose verb type includes SUBJECT, DIRECT OBJECT and all the arguments specified in the analysis except, obviously, for subject and agent of the passive configuration. All verb types containing DIRECT OBJECTs are accepted. In formal terms: all the examples with key 3 (passive voice) in field 6 (voice) of BDS are reduced to scheme *SDO, ACT* —where *vtype* is SDOIO, SDOPR, SDOPC or SDOIOPC if fields 28 (type of indirect object), 35 (type of prepositional complement) and 55 (type of predicative complement) have keys 1, 2 or *.
6. Examples illustrating active voice *se* impersonal configurations with no SUBJECT argument are reduced to active voice personal schemes whose verb type includes SUBJECT and the rest of the arguments specified in the analysis. All verb types are accepted. In formal terms: all the examples of BDS with key 2 (middle voice¹⁵) in field 6 (voice) and with key 35 (*se* impersonal) in field 16 (type of the subject) are reduced to scheme *S, ACT* —where *vtype* is any verb type value if fields 21 (type of direct object), 28 (type of indirect object), 35 (type of prepositional complement) and 55 (type of predicative complement) of BDS have keys 1, 2 or *.
7. Examples illustrating active voice impersonal configurations due to inherently active voice impersonal verbs are reduced to active voice impersonal schemes whose verb type includes all the arguments specified in the analysis. Only verb types including SUBJECT, SUBJECT and DIRECT OBJECT or SUBJECT and PREPOSITIONAL COMPLEMENT are accepted. In formal terms: all examples of BDS with key 1 (active voice) in field 6 (voice), key 4 (impersonal) in field 16 (type of the subject) and key different from 22 (*haber*

¹⁵*Se* impersonal configurations are classified as middle voice in BDS.

que) in field 9 (periphrasis), are reduced to scheme *S, ACT_IMP* —where *vtype* is SDO or SPC if fields 21 (type of DIRECT OBJECT) or 35 (type of PREPOSITIONAL COMPLEMENT) have keys 1, 2 or *.

8. Examples illustrating middle voice impersonal configurations due to inherently middle voice impersonal verbs are reduced to middle voice impersonal schemes whose verb type includes all the arguments specified in the analysis. Only verb types that include SUBJECT, SUBJECT and DIRECT OBJECT or SUBJECT and PREPOSITIONAL COMPLEMENT are accepted. In formal terms: all the examples of BDS with key 2 (middle voice) in field 6 (voice), key 4 (impersonal) in field 16 (type of the subject) and key different from 22 (*haber que*) in field 9 (periphrasis), are reduced to scheme *S, MIDD_IMP* —where *vtype* is SDO or SPC if fields 21 (type of direct object) or 35 (type of prepositional complement) have keys 1, 2 or *.
9. Examples illustrating active voice impersonal configurations due to periphrasis *haber que* are reduced to active voice personal schemes whose verb type includes SUBJECT and all the arguments specified in the analysis. All verb types are accepted. In formal terms: all the examples of BDS with key 1 (active voice) in field 6 (voice), key 4 (impersonal) in field 16 (type of the subject) and key 22 (*haber que*) in field 9 (periphrasis) are reduced to scheme *S, ACT* —where *vtype* is a different verb type value if fields 21 (type of direct object), 28 (type of indirect object), 35 (type of prepositional complement) and 55 (type of predicative complement) of BDS have keys 1, 2 or *.
10. Examples illustrating middle voice impersonal configurations due to periphrasis *haber que* are reduced to middle voice personal schemes whose verb type includes SUBJECT and all the arguments specified in the analysis. All verb types are accepted. In formal terms: all the examples of BDS with key 2 (middle voice) in field 6 (voice), key 4 (impersonal) in field 16 (type of the subject) and key 22 (*haber que*) in field 9 (periphrasis) are reduced to scheme *S, MIDD* —where *vtype* is a different verb type value if fields 21 (type of direct object), 28 (type of indirect object), 35 (type of prepositional

complement) and 55 (type of predicative complement) of BDS have keys 1, 2 or *.

11. Examples of active voice *se impersonal* configurations that include a SUBJECT argument were initially encoded in BDS as middle voice personal configurations in third person singular or plural. The ambiguity generated by this encoding is being solved by the addition of new fields to BDS during the last process of correction, currently in course. After this disambiguation, these examples will be reduced to active voice personal schemes whose verb type includes SUBJECT, DIRECT OBJECT and all of the arguments specified in the analysis except, obviously, for the subject of the source active voice impersonal configuration. All verb types that include a DIRECT OBJECT will be accepted.

By means of the application of these rules to all the clauses that document each verb in BDS, we obtain the set of possible combinations of *vtype*, *voice* and *preptype* in which the verb has been documented. Each combination constitutes a scheme for the verb under consideration and thus will be used for the assignment of values to *vtype*, *voice* and *preptype* affixes in the AGFL lexicon definitions of verb forms. However, before this assignment takes place, the set of schemes derived from BDS still has to undergo a process of clustering. This means that, trying to keep the lexicon as short as possible, partially different schemes are merged for the same lemma. The process of merging is very simple: for each verb lemma, we sort the set of the scheme descriptions, giving priority to *voice* values in first place, *vtype* values in second place, and *preptype* values in third place. The process is then carried out in three steps:

- Taking the value of affix *voice* as a key, while values for *vtype* are equal, values for *preptype* are merged.
- Taking the value of affix *voice* as a key, while values for *preptype* are identical, values for *vtype* are merged.
- Taking the value of affix *vtype* as a key, while values for *preptype* are identical, values for *voice* are merged.

At the end of the process, for the same verb lemma all schemes sharing at least two of the fields *voice*, *vtype* or *preptype* —giving priority to *voice* variations— are merged into the same *verb scheme cluster*. An example of the whole process for the verb lemma *abandonar* is included in Appendix C.

In contrast to this method, the assignment of *voice*, *vtype* and *preptype* values to auxiliary or semiauxiliary verb forms is done manually, since these verbs constitute a closed class of words whose values for *vtype*, *voice* and *preptype* can —and must— be predefined according to their function in the grammar, which is primarily the constitution of compound tenses and periphrastic VPs. In this regard, while only four values of *vtype* are relevant for auxiliaries, almost all *voice* values are somewhat pertinent for them. As to *vtype*, **AUX**, **PARTICIPLE**, **INFINITIVE** and **GERUND** values are respectively associated with *haber*, *ser* (+ participle), and the active or middle voice semiauxiliaries collected in Table A.1. As to *voice*, value **PASS** is associated with semiauxiliary verb *ser*, while values **ACT**, **ACT_IMP**, **MIDD**, **MIDD_IMP**, **ip_ACT**, **ip_MIDD**, **ip_ACT_IMP** and **ip_MIDD_IMP** are all associated with periphrastic groups (see Section 2.2.2.5) in which non-passive semiauxiliaries take part.

For regular verb forms —not including orthographical variations—, values for affix *voice* are the following:

- Middle voice personal semiauxiliaries, which cannot be combined within the periphrastic group with clitic pronouns selected by the main verb of the VP, are assigned value **MIDD**.

```
VerbSt ( INFINITIVE, MIDD, FIRST, SING, PRESENT, INDICATIVE ):
    "pongo".
```

MIDD voice semiauxiliaries produce **MIDD** voice periphrastic groups (se puso a *trabajar*) and, combined with the active voice impersonal semiauxiliary *haber que*, they produce also **MIDD_IMP** voice periphrastic groups (ha habido que ponerse a *trabajar*).

- Active voice personal semiauxiliaries, which can be combined within the periphrastic group with clitic pronouns selected by the main verb of the VP, are assigned value **ACT**.

VerbSt (INFINITIVE, ACT, FIRST, SING, PRESENT,
INDICATIVE): "puedo".

When they are not combined with other voice semiauxiliaries, i) **ACT** voice semiauxiliaries constitute **ACT** voice periphrastic groups if they alone constitute the periphrastic group (puede haber tenido que *salir*); ii) they constitute **ip_ACT** voice periphrastic groups if there are functional clitic pronouns in the periphrastic group (me lo ha tenido que *entregar*); iii) they constitute **ip_MIDD** periphrastic groups if there is one clitic pronoun that echoes *person* and *number* of the VP in the periphrastic group (me lo he tenido que *quedar*). **ACT** voice semiauxiliaries, on the other hand, may participate in periphrastic groups of all voice values (**ACT**: puede empezar a *tocar*, **ACT_IMP**: hay que empezar a *tocar*, **MIDD**: puede ponerse a *gritar*, **MIDD_IMP**: puede haber que ponerse a *pensar*, **ip_ACT**: puede habérselo tenido que *decir*, **ip_ACT_IMP**: hay que irlo *teniendo en cuenta*, **ip_MIDD**: se tiene que ir *habituando*, **ip_MIDD_IMP**: hay que irse *adaptando*).

- Active voice impersonal semiauxiliary *haber que*, which can be combined within the periphrastic group with clitic pronouns required by the main verb of the VP, is assigned value **ACT_IMP**.

VerbSt (INFINITIVE, ACT_IMP, THIRD, QUE, SING, PRESENT,
INDICATIVE): "hay".

ACT_IMP voice semiauxiliary *haber que* constitutes **ACT_IMP** voice periphrastic groups (hay que *decirle*). Combined with middle voice semiauxiliaries, it produces **MIDD_IMP** voice periphrastic groups (hay que ponerse a *trabajar*). Combined with functional clitic pronouns, it produces **ip_ACT_IMP** voice periphrastic groups (hay que irlo *acostumbrando*). Combined with one non-functional clitic pronoun echoing *person* and *number* of the VP, it produces **ip_MIDD_IMP** voice periphrastic groups (hay que irse *acostumbrando*).

- Passive voice semiauxiliary verb *ser* is assigned value **PASS**.

VerbSt (PARTICIPLE, PASS, FIRST, SING, PRESENT, INDICATIVE):
"soy".

It constitutes passive voice VPs (*es repartido*).

- Auxiliary *haber* has no inherent voice. Value for affix *voice* of semiauxiliary *haber* is determined within the grammar by means of unification with *voice* values of the participles combined with *haber* for the production of compound tenses. It means, obviously, that lexicon definitions of *haber* must be assigned *open* values for affix *voice*.

To be able to unify with *voice* values of all verbs in the lexicon, second and first person regular verb forms are assigned values **ACT|MIDD|PASS**, third person and infinitive regular verb forms add **ACT_IMP|MIDD_IMP** to the previous ones. Because of syncretism of third and first person singular, for third person singular, *voice* values involving impersonality are sometimes encoded in different lexicon definitions:

```
VerbSt ( AUX, ACT|MIDD|PASS, FIRST, SING, PRESENT,  
INDICATIVE ): "he".  
VerbSt ( AUX, PASS|act|midd, THIRD, SING, PRESENT,  
INDICATIVE ): "ha".  
VerbSt ( AUX, ACT|MIDD|PASS, FIRST|THIRD, SING, PRESENT,  
SUBJUNCTIVE ): "haya".  
VerbSt ( AUX, ACT_IMP|MIDD_IMP, THIRD, SING, PRESENT,  
SUBJUNCTIVE ): "haya".
```

Regular infinitive and gerund verb forms should be simultaneously assigned all values listed above. However, gerund is not assigned middle voice values. This is due to the fact that present perfect gerund of non-periphrastic middle voice VPs uses the alternative gerund form of the auxiliary, because it must be attached the non-functional clitic pronoun required by middle voice (*habiéndose propuesto*)¹⁶.

```
VerbSt ( AUX, PASS|act|midd, PRESENT, INFINITIVE ): "haber".  
VerbSt ( AUX, PASS|act, PRESENT, GERUND ): "habiendo".
```

¹⁶There is only one context in which regular gerund forms are needed in middle voice: within periphrastic VPs when clitic pronouns are moved to the periphrastic group (*se está vistiendo*). Nevertheless, present perfect gerund is not accepted in this distribution, that is, *habiendo vestido* will never be found in the position of *vistiendo* in our example, which means that *habiendo* will never need to unify with middle voice participles.

Participles of active voice personal and impersonal semiauxiliaries are respectively assigned **ACT|ip_no_IMP** and **ACT_IMP|ip_ACT_IMP** voice values.

```
Participle VerbSt ( INFINITIVE, ACT|ip_no_IMP, PARTICIPLE,
  MASC, SING ): "podido".
Participle VerbSt ( INFINITIVE, ACT_IMP|ip_ACT_IMP, QUE,
  PARTICIPLE, MASC, SING ): "habido".
```

The cooccurrence of both values in each case ensures that semiauxiliary participles can be combined with both regular and alternative forms of the auxiliary verb *haber*, constituting, respectively, proper and improper voice periphrastic groups (haber podido *decírselo*, habérselo podido *decir*, haber podido *ponérselo*, habérselo podido *poner*). Unlike participles of personal semiauxiliaries, those of impersonal semiauxiliaries do not need to cover improper middle voice, since non-functional clitic pronouns —either those selected by main verbs or those corresponding to middle voice semiauxiliaries— cannot be attached to impersonal active voice semiauxiliaries (that is, the production of an improper middle voice impersonal periphrastic group such as this found in *habiéndose habido que vestir**, as well as the production of a proper middle voice impersonal periphrastic group such as this found in *habiéndose habido que poner a trabajar**, are blocked by the grammar).

Lexicon definitions of non-regular verb forms of auxiliaries —alternative forms that account for orthographical variations caused by clitic pronoun attachment—, were introduced in Section 2.2.2.2.

All the alternative forms of active voice semiauxiliaries are assigned voice values in the domain of *improper_voice*, since these are forms that must be attached clitic pronouns selected by the main verb of the VP. These verb forms constitute **ip_ACT** (tenérselo que *contar*), **ip_MIDD** (tenérselo que *tragar*), **ip_ACT_IMP** (hay que írselo *diciendo*) or **ip_MIDD_IMP** (hay que írselo *creyendo*) voice periphrastic groups. All alternative forms need to be produced for active voice semiauxiliaries:

- Alternative forms only compatible with functional clitic pronouns (*seguídselo contando*), which are assigned values **ip_ACT** or **ip_ACT_IMP**.
- Alternative forms only compatible with non-functional clitic pronouns (*seguíoslo poniendo*), which are assigned values **ip_MIDD** or **ip_MIDD_IMP**.

- Alternative forms compatible with both types of clitic pronouns (*seguírsele contando*, *seguírsele poniendo*), which are assigned non-terminals **ip_no_IMP** or **ip_IMP**. Alternative forms *habér* and *habiéndo* of active impersonal semiauxiliary *haber que*, which should be included here, are only assigned **ip_ACT_IMP** value for voice, and not **ip_IMP**, which would cover both (improper) active and middle impersonal voices. This is due to the fact that the grammar blocks the production of middle voice impersonal periphrastic groups in which the non-functional clitic pronoun selected either by a middle voice main verb or by a following semiauxiliary verb is attached to the active voice impersonal semiauxiliary (*habiéndose que vestir**, *habiéndose que poner a trabajar**).

Passive voice semiauxiliary *ser* and middle voice semiauxiliaries —as main verbs— have the same voice value for regular and alternative verb forms. Of these, only the verb forms necessary to cover the distributions in which these semiauxiliaries can be found are produced. Both passive and middle voice semiauxiliaries can be combined with not more than one clitic pronoun, so only alternative forms with values **ONE** or **JUST_ONE** for affix cliticization are produced (that is, *siéndo* or *poniéndo* are generated, while *sér** or *ponér** are not).

The alternative forms of auxiliary *haber* must include all the values in the domain of *improper_voice*, except for value **ip_MIDD_IMP**, for the same reason that causes the absence of this value for participles and alternative forms of semiauxiliary *haber que* (see above in this Section). The alternative form of the infinitive lacks also value **PASS**, because this voice value cannot be combined with two clitic pronouns, and **ACT_IMP** and **MIDD_IMP** values, because these values must be always associated with verb types that do not include more than one argument replaceable by a clitic pronoun.

2.2.3 The Grammar

This Section is devoted to the description of the rules of the grammar that join the elements that constitute VPs. The grammar of VPs, as remarked above (see Chapter 1 and Section 2.2.1), describes linguistic phenomena that occur beyond the level of the word, this being considered as the sequence of characters between

two space characters, with the only exception of enclitic pronoun attachment to infinitive, gerund and imperative forms of verbs. Although the nature of such linguistic phenomena has, in principle, scarce interest from the syntactic point of view, these linguistic phenomena have to be carefully accounted for here, because they are ultimately responsible for the contextualization capabilities of the **PREDICATE** function to organize syntactic arguments at the level of the clause. A side-effect of this fact, however, is that the intermediate constituents arising between nodes labelled as lexical items and nodes labelled as **PREDICATE**s, as they are, on the one hand, necessary from the formal point of view, but, on the other, syntactically not relevant by themselves, they are all veiled in the output analyses.

From the point of view of words, the elements joined by the grammar of VPs are clitic pronouns (*me, te, se, lo, la, le, nos, os, se, los, las, les*) and verb forms. From the formal point of view, the elements joined by rules of VPs are the non-terminals which in the end rewrite as clitic pronouns or verb forms. To them, we have to add the non-terminals of *predicates* that impose restrictions on the combination and generation of values of the non-terminal affixes involved.

Now, if we, just for the time being, disregard the portions of names enclosed between square brackets—except for the one in component 2 with respect to the connection between 1 and 2—Figure 2.1 shows the relations between the various productive non-terminals that take part in the description of personal VPs.

The **PREDICATE** function, which is the main component of main non-finite clause descriptions, rewrites either as non-terminal **Periphrastic Verb Phrase** or as non-terminal **Verb Phrase**. Both types of VPs supply the **PREDICATE** function with the same amount of information, but their components are different. On the one hand, non-terminal **Verb Phrase** is constituted by one or two optional **Clitic Pronoun** non-terminals and one main **Verb** non-terminal in **INDICATIVE**, **SUBJUNCTIVE** or **IMPERATIVE** mood. On the other, non-terminal **Periphrastic Verb Phrase** is constituted by non-terminal **Periphrasis** and non-terminal **Non Finite Verb Phrase**. Of these, whereas non-terminal **Non Finite Verb Phrase** will be specifically addressed below, we can already state that non-terminal **Periphrasis** is constituted by one or two optional **Clitic Pronoun** non-terminals and one semiauxiliary **Verb** non-terminal in **INDICATIVE**,

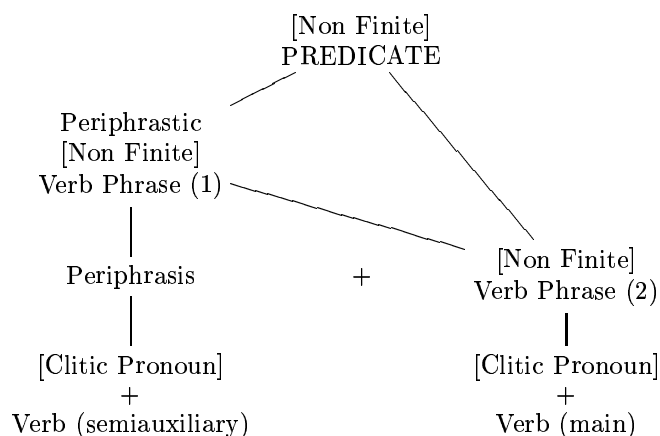


Figure 2.1: VP productive non-terminals.

SUBJUNCTIVE or IMPERATIVE mood.

Simultaneously, but taking now into account the portions of names enclosed between square brackets, Figure 2.1 shows the relations between the productive non-terminals that take part in the description of non-finite VPs. The **NON FINITE PREDICATE** function, which will be the main component of finite clause descriptions, rewrites either as non-terminal **Periphrastic Non Finite Verb Phrase** or as non-terminal **Non Finite Verb Phrase**. Both of them supply the **NON FINITE PREDICATE** function with the same information, but their components are different. Non-terminal **Periphrastic Non Finite Verb Phrase** is constituted by non-terminal **Periphrasis** and non-terminal **Non Finite Verb Phrase** again. Of these, whereas non-terminal **Periphrasis** is constituted by one or two optional **Clitic Pronoun** non-terminals and one semiauxiliary **Verb** non-terminal in **INFINITIVE** or **GERUND** mood, non-terminal **Non Finite Verb Phrase**¹⁷ is constituted by one or two optional **Clitic Pronoun** non-terminals and one main **Verb** non-terminal in **INFINITIVE** or **GERUND** mood.

Although this is not accounted for in Figure 2.1, together with these PREDIC-

¹⁷Note that, with the non-terminals under consideration in this paragraph, **Non Finite Verb Phrase** is by now an immediate constituent of i) **Periphrastic Verb Phrase**, ii) **NON FINITE PREDICATE** and iii) **Periphrastic Non Finite Verb Phrase** non-terminals.

ATE non-terminals, there is also a non-terminal **PARTICIPLE PREDICATE**. This is the main component of participle non-finite clauses, and is constituted by a **Participle Verb Phrase** non-terminal, which in turn rewrites as a **Participle Verb** non-terminal.

What might, however, be more surprising about Figure 2.1 is the absence of lexicon non-terminals —or grammar terminals— **VerbSt**, **VerbSt_clitized** and **Participle VerbSt**, described in previous Section 2.2.1 and rewritten as verb forms. Although one would have expected these lexicon non-terminals to be the constituents of VP non-terminals **Verb Phrase**, **Periphrasis** and **Non Finite Verb Phrase**, we find that these rewrite as certain **Verb** non-terminals instead. These **Verb** non-terminals are in fact intermediate constituents between VP non-terminals and **VerbSt** lexicon non-terminals (see Section 2.2.3.4), introduced in the grammar for the following reasons. First, this permits the introduction of verb morphology in the grammar, in which case **Verb** non-terminals would rewrite as verb endings and stems contained in the lexicon instead of as wordforms. Currently, however, as morphology is not accounted for by the grammar, the node created by these intermediate non-terminals is used to rewrite word categories as regular expressions identified solely by means of their endings, trying to assign a category tag, in this case **Verb**, to unknown words in the input (that is, **Verb**, which first rewrites as **VerbSt** non-terminal from the lexicon, rewrites also as any sequence of characters matching a certain pattern, one ending in *-ríamos*, for example, which can be identified as a conditional indicative, first person plural verb form).

In addition to this, by means of these intermediate non-terminals we can handle simple and compound tenses in a similar way within the grammar, since **Verb** non-terminals called upon by VP rules can rewrite, for example, both as *amo* and as *he amado*. In other words, this means that the auxiliary *haber* is, within the grammar, paired with the morphological endings that play the same inflectional role within the lexicon. As a result, no differences are appreciated in the grammar between inflectional compound and simple verb forms.

To conclude, if we present a summarization of what we have presented so far in this Section, we will also get an idea of the organization of the rest of this exposition about the grammar of VPs. As observed above, the target non-terminals of the VP

module of the grammar are **PREDICATE** and **NON FINITE PREDICATE** functions, which are constituted by periphrastic or simple finite or non-finite VPs, **Periphrastic Verb Phrase**, **Periphrastic Non Finite Verb Phrase**, **Verb Phrase** and **Non Finite Verb Phrase** (**PREDICATE** and **NON FINITE PREDICATE** non-terminals will be described in Section 2.2.3.1). On the other hand, non-terminals for VP *predicates*, non-terminals which ultimately rewrite as clitic pronouns (**Clitic Pronoun**) and non-terminals which ultimately rewrite as verb forms (**Verb**) constitute **Verb Phrase**, **Periphrasis** and **Non Finite Verb Phrase** non-terminals for VPs (*predicate* non-terminals will be described in Section 2.2.3.2, **Clitic Pronoun** non-terminals in Section 2.2.3.3, **Verb** non-terminals in Section 2.2.3.4, **Verb Phrase** and **Non Finite Verb Phrase** non-terminals in Section 2.2.3.5 and **Periphrasis** non-terminals in Section 2.2.3.6). Finally, periphrastic VPs **Periphrastic Verb Phrase** and **Periphrastic Non Finite Verb Phrase** rewrite as periphrastic groups —finite or non-finite **Periphrasis**— followed by non-finite VPs —**Non Finite Verb Phrase**— (**Periphrastic Verb Phrase** and **Periphrastic Non Finite Verb Phrase** non-terminals will be described in Section 2.2.3.7).

A central role is played in the grammar of VPs by affix *voice*. The reason for this lies on the fact that *voice*, which in our formalization encodes both voice and impersonality, is the element primarily concerned in the process of contextualization of verb schemes to yield configurations. If the VP grammar is fundamentally conceived of as a formalization of the procedures whereby verb schemes yield configurations, *voice* naturally becomes the core of the VP grammar.

2.2.3.1 Predicate Function: **PREDICATE**, **NON FINITE PREDICATE** and **PARTICIPLE PREDICATE**

In the approach described in Section 2.1 *configuration*, *person*, *gender* and *number* of the VP, as well as *case*, *person*, *gender* and *number* of possible clitic pronouns included in it, were identified as the relevant information for carrying out the process of contextualization of VPs to constitute the predicate syntactic functions of clauses. This means that, from a formal point of view, the non-terminal that represents the predicate function should specify exclusively this information, given that this is the only information actually used in the occurrences of the predicate

function in clause description rules.

Different predicate function non-terminals had to be designed for VPs based on different verb forms, not all of them able to supply the predicate function with information for all and the same attributes. We have indeed different predicate non-terminals for VPs constituted either by finite or by non-finite verb forms. Not only this, however, because prepositional requirements and clitic pronouns determine the appearance of other predicate non-terminals too.

All predicate function non-terminals directly rewrite either as periphrastic or as non-periphrastic VPs. Affixes whose names present the suffixes *_one* or *_two* record information about clitic pronouns attached to VPs. These affixes dominate the set of values —more exactly the subset of these values that can be associated with clitic pronouns— in the domain of the respective non-*_one* (or *_two*) non-terminal affixes. This strategy of nesting affix values serves within the grammar to refer in the same non-terminal to independent values for *person*, *gender*, *number* and *case*, associated either with the verb form itself or with each of the possibly up to two clitic pronouns present in each predicate function.

Predicate function non-terminals —and their rewrite rules— constituted by VPs based on finite verb forms are the following¹⁸:

```
PREDICATE ( mvtype, voice, person, number, mood ):  
  Periphrastic Verb Phrase ( mvtype, voice, person, number,  
    mood );  
  Verb Phrase ( mvtype, voice, person, number, mood ).  
  
PREDICATE ( mvtype, voice, preptype, person, number, mood ):  
  Periphrastic Verb Phrase ( mvtype, voice, preptype, person,  
    number, mood );  
  Verb Phrase ( mvtype, voice, preptype, person, number,  
    mood ).  
  
PREDICATE ( mvtype, voice, person, number, mood, case_one,  
  person_one, gender_one, number_one ):  
  Periphrastic Verb Phrase ( mvtype, voice, person, number,  
    mood, case_one, person_one, gender_one, number_one );  
  Verb Phrase ( mvtype, voice, person, number, mood,  
    case_one, person_one, gender_one, number_one ).
```

¹⁸Almost every VP rule is duplicated in the grammar to account for verb schemes specifying prepositional requirements, the only difference between first rules and duplications being the introduction of affix *preptype*. Unless there are additional differences, we will not henceforth display the duplicated rules that include affix *preptype*.

```

PREDICATE ( mvtype, voice, preptype, person, number, mood,
            case_one, person_one, gender_one, number_one ):
  Periphrastic Verb Phrase ( mvtype, voice, preptype, person,
                             number, mood, case_one, person_one, gender_one,
                             number_one );
  Verb Phrase ( mvtype, voice, preptype, person, number, mood,
               case_one, person_one, gender_one, number_one ).

PREDICATE ( mvtype, voice, person, number, mood, case_one,
            person_one, gender_one, number_one, case_two, person_two,
            gender_two, number_two ):
  Periphrastic Verb Phrase ( mvtype, voice, person, number,
                             mood, case_one, person_one, gender_one, number_one,
                             case_two, person_two, gender_two, number_two );
  Verb Phrase ( mvtype, voice, person, number, mood, case_one,
               person_one, gender_one, number_one, case_two, person_two,
               gender_two, number_two ).

PREDICATE ( mvtype, voice, preptype, person, number, mood,
            case_one, person_one, gender_one, number_one, case_two,
            person_two, gender_two, number_two ):
  Periphrastic Verb Phrase ( mvtype, voice, preptype, person,
                             number, mood, case_one, person_one, gender_one, number_one,
                             case_two, person_two, gender_two, number_two );
  Verb Phrase ( mvtype, voice, preptype, person, number, mood,
               case_one, person_one, gender_one, number_one, case_two,
               person_two, gender_two, number_two ).

```

Although these rewrite rules are mostly self-explanatory, we want to remark the fact that the clitic pronouns that yield information for **PREDICATE** non-terminals, information recorded by *_one* and *_two* affixes, are exclusively **functional** clitic pronouns, since these, contrary to what happens with non-functional clitic pronouns, which indicate middle voice or impersonality, supply the **PREDICATE** non-terminal with new information not present in the verb form itself.

After this remark, however, it might be of help to digress slightly from the strict explanation of these rules, in order to describe how the target **PREDICATE** non-terminals on the right-hand side of these rules are, in higher levels of the analysis, integrated within the whole grammar, since it will facilitate the understanding of the general structure of it —divided, as specified in Chapter 1, in two wide areas: clause and phrase—, and of the rest of this Section, specifically devoted to VP structure.

As regards the integration of **PREDICATE** non-terminals in clause rewrite

rules, to begin with, we can observe that every clause rewrite rule in our grammar consists of a large series of alternatives, each of which is constituted by a concrete combination and arrangement of a set of syntactic functions, one of these being, precisely, a **PREDICATE** function. The combination of syntactic functions and the values taken by affixes *mvtype* and *voice* —in other words, the *configuration*— of the **PREDICATE** non-terminal, are mutually dependent for each particular alternative. If one concrete combination and arrangement of arguments is considered possible for a certain combination of *mvtype* and *voice* values —a concrete configuration—, there must be an alternative that describes the clause as a sequence of the arguments in that order, and where the values taken by affixes *mvtype* and *voice* of the **PREDICATE** function are those that correspond to the involved configuration.

In addition to this, the integration of the **PREDICATE** function in clause description rules complies with the principle that all the arguments —except for the **SUBJECT**— symbolized in the value of the affix *mvtype* in the **PREDICATE** function must be somewhat present within the alternative in question: they can be either explicit arguments of the alternative or implicit arguments substituted by clitic pronouns found in the **PREDICATE** non-terminal. As a result, coverage of the alternatives of clause rewrite rules with respect to the arguments symbolized by *mvtype* of their **PREDICATE** non-terminal is always **exhaustive**, which means that there are not optional syntactic functions symbolized neither in the *mvtype* values nor in the alternatives. This is a logical consequence of the way values for *mvtype*, *voice* and *preptype* —*schemes*— contained in the lexicon are encoded in BDS and extracted from there (see Section 2.2.2.6).

To make an example, if a **PREDICATE** non-terminal is part of an alternative with both explicit **SUBJECT** and **PREPOSITIONAL COMPLEMENT** functions, affix *mvtype* of the **PREDICATE** non-terminal must take one of the values **SPC**, **SDOPC** or **SDOIOPC**, with, respectively, one and two clitic pronouns in the latter two cases. *Voice*, on the other hand, takes the values **ACT**|**MIDD**, if the combination and arrangement of arguments in the alternative are considered relevant for both voice values.

Since the set of all the alternatives for clause rewrite rules requires that affixes *mvtype* and *voice* of their **PREDICATE** non-terminals take a large number of

different combinations of literal values, a correspondingly large variety of **PREDICATEs**, with different values for *mvtype* and *voice*, must be produced by VP rewrite rules. Each of the rules presented above for **PREDICATE** non-terminals gathers the generation of a certain group of the **PREDICATE** configurations required by the alternatives of clause rewrite rules. The configurations covered by the above rewrite rules and their corresponding **PREDICATE** non-terminals are only outlined here, more details about them will arise from later explanations about the VP grammar.

- **PREDICATE/5** covers personal and impersonal active, middle and passive voice configurations. Any verb type not containing a PREPOSITIONAL COMPLEMENT may appear here. In formal terms: for VPs not containing functional clitic pronouns, any possible combination of non-**PC**- verb type and proper voice (**ACT**, **MIDD**, **ACT_IMP**, **MIDD_IMP**, **PASS**, **PASS_IMP**, **ACT_MIDD+PASS**) is handled by **PREDICATE/5**.
- **PREDICATE/6** covers personal and impersonal active, middle and passive voice configurations. Any verb type containing a PREPOSITIONAL COMPLEMENT may appear here. In formal terms: for VPs not containing functional clitic pronouns, any possible combination of **-PC**- verb type and proper voice (**ACT**, **MIDD**, **ACT_IMP**, **MIDD_IMP**, **PASS**, **PASS_IMP**, **ACT_MIDD+PASS**) is handled by **PREDICATE/6**.
- **PREDICATE/9** covers personal and impersonal active, middle and passive voice configurations. Any verb type not containing a PREPOSITIONAL COMPLEMENT and including at least one complement that it is possible to substitute by a clitic pronoun may appear here. In formal terms: for VPs containing one functional clitic pronoun, any possible combination of non-**PC**- verb type in the domain of non-terminal affix *one_clitic*, and proper voice (**ACT**, **MIDD**, **ACT_IMP**, **MIDD_IMP**, **PASS**, **PASS_IMP**, **ACT_MIDD+PASS**) is handled by **PREDICATE/9**.
- **PREDICATE/10** covers personal and impersonal active, middle and passive voice configurations. Any verb type containing a PREPOSITIONAL COMPLEMENT and including at least one complement that it is possible

to substitute by a clitic pronoun may appear here. In formal terms: for VPs containing one functional clitic pronoun, any possible combination of **-PC-** verb type in the domain of non-terminal affix *one_clitic*, and proper voice (**ACT**, **MIDD**, **ACT_IMP**, **MIDD_IMP**, **PASS**, **PASS_IMP**, **ACT_-MIDD+PASS**) is handled by **PREDICATE/10**.

- **PREDICATE/13** covers personal and (periphrastic) impersonal active voice configurations. Any verb type not containing a **PREPOSITIONAL COMPLEMENT** and including two complements that it is possible to substitute by clitic pronouns may appear here. In formal terms: for VPs containing two functional clitic pronouns, any possible combination of non-**-PC-** verb type in the domain of non-terminal affix *two_clitics*, and active voice (**ACT**, **ACT_IMP**) is handled by **PREDICATE/13**.
- **PREDICATE/14** covers personal and (periphrastic) impersonal active voice configurations. Any verb type containing a **PREPOSITIONAL COMPLEMENT** and including two complements that it is possible to substitute by clitic pronouns may appear here. In formal terms: for VPs containing two functional clitic pronouns, any possible combination of **-PC-** verb type in the domain of non-terminal affix *two_clitics*, and active voice (**ACT**, **ACT_IMP**) is handled by **PREDICATE/14**.

The predicate function non-terminals —and their rewrite rules— constituted by VPs based on non-finite verb forms are the following:

```
NON FINITE PREDICATE ( mvtype, voice, mood ):
  Periphrastic Non Finite Verb Phrase ( mvtype, voice, mood );
  Non Finite Verb Phrase ( mvtype, voice, mood ).

NON FINITE PREDICATE ( mvtype, voice, mood, case_one, person_one,
  gender_one, number_one ):
  Periphrastic Non Finite Verb Phrase ( mvtype, voice, mood,
    case_one, person_one, gender_one, number_one );
  Non Finite Verb Phrase ( mvtype, voice, mood, case_one,
    person_one, gender_one, number_one ).

NON FINITE PREDICATE ( mvtype, voice, mood, case_one, person_one,
  gender_one, number_one, case_two, person_two, gender_two,
  number_two ):
  Periphrastic Non Finite Verb Phrase ( mvtype, voice, mood,
```

```

    case_one, person_one, gender_one, number_one, case_two,
    person_two, gender_two, number_two );
Non Finite Verb Phrase ( mvtype, voice, mood, case_one,
    person_one, gender_one, number_one, case_two, person_two,
    gender_two, number_two ).

```

As the previous ones, these rewrite rules are self-explanatory. Above them also, only one remark has to be done, which is the fact that, unlike finite predicates, non-finite ones, apart from the information about functional clitic pronouns, record information that concerns non-functional clitic pronouns too, since, for middle voice, these supply the **NON FINITE PREDICATE** non-terminal with values for *person* and *number* of the VP, which, as we all know, is information that cannot be expressed by the non-finite verb form alone (*ponerme* in *Quiero ponerme a la cola cuanto antes* has value **FIRST** for *person* and value **SING** for *number*).

Configurations covered by **NON FINITE PREDICATE** non-terminals of these rules are the following:

- **NON FINITE PREDICATE/3** covers active and passive voice configurations. Any verb type not containing a **PREPOSITIONAL COMPLEMENT** may appear here. In formal terms: for VPs not containing clitic pronouns, any possible combination of non--**PC**- verb type and active or passive voice (**ACT**, **PASS**) is handled by **NON FINITE PREDICATE/3**.
- **NON FINITE PREDICATE/4** covers active and passive voice configurations. Any verb type containing a **PREPOSITIONAL COMPLEMENT** may appear here. In formal terms: for VPs not containing clitic pronouns, any possible combination of **-PC**- verb type and active or passive voice (**ACT**, **PASS**) is handled by **NON FINITE PREDICATE/4**.
- **NON FINITE PREDICATE/7** covers active, middle and passive voice configurations. For active and passive voice any verb type not containing a **PREPOSITIONAL COMPLEMENT** and including at least one complement that it is possible to substitute by a clitic pronoun may appear here, for middle voice any verb type not containing a **PREPOSITIONAL COMPLEMENT** might be specified. In formal terms: for VPs containing one clitic pronoun, any possible combination of non--**PC**- verb type in the domain of

non-terminal affix *one_clitic*, and active or passive voice (**ACT**, **PASS**), and any possible combination of non-**PC**- verb type and middle voice (**MIDD**), are handled by **NON FINITE PREDICATE/7**.

- **NON FINITE PREDICATE/8** covers active, middle and passive voice configurations. For active and passive voice any verb type containing a **PREPOSITIONAL COMPLEMENT** and including at least one complement that it is possible to substitute by a clitic pronoun may appear here, for middle voice any verb type containing a **PREPOSITIONAL COMPLEMENT** might be specified. In formal terms: for VPs containing one clitic pronoun, any possible combination of **-PC**- verb type in the domain of non-terminal affix *one_clitic*, and active or passive voice (**ACT**, **PASS**), and any possible combination of **-PC**- verb type and middle voice (**MIDD**), are handled by **NON FINITE PREDICATE/8**.
- **NON FINITE PREDICATE/11** covers active and middle voice configurations. For active voice, any verb type not containing a **PREPOSITIONAL COMPLEMENT** and including two complements that it is possible to substitute by clitic pronouns may appear here, for middle voice any verb type not containing a **PREPOSITIONAL COMPLEMENT** and including a complement that it is possible to substitute by a clitic pronoun might be specified. In formal terms: for VPs containing two clitic pronouns, any possible combination of non-**PC**- verb type in the domain of non-terminal affix *two_clitics*, and active voice (**ACT**), and any possible combination of non-**PC**- verb type in the domain of non-terminal affix *one_clitic* and middle voice (**MIDD**), are handled by **NON FINITE PREDICATE/11**.
- **NON FINITE PREDICATE/12** covers active and middle voice configurations. For active voice, any verb type containing a **PREPOSITIONAL COMPLEMENT** and including two complements that it is possible to substitute by clitic pronouns may appear here, for middle voice any verb type containing a **PREPOSITIONAL COMPLEMENT** and including a complement that it is possible to substitute by a clitic pronoun might be specified. In formal terms: for VPs containing two clitic pronouns, any possible combination of **-PC**- verb type in the domain of non-terminal affix *two_clitics*,

and active voice (**ACT**), and any possible combination of **-PC-** verb type in the domain of non-terminal affix *one_clitic* and middle voice (**MIDD**), are handled by **NON FINITE PREDICATE/12**.

The rewrite rules of non-terminals **PARTICIPLE PREDICATE**, constituted by VPs based on participles, are the following:

```
PARTICIPLE PREDICATE ( mvtype, voice, mood, gender, number ):
  Participle Verb Phrase ( mvtype, voice, mood, gender, number ).
```

Configurations covered by these non-terminals and rewrite rules are the following:

- **PARTICIPLE PREDICATE/5** covers active, middle and passive voice configurations. Any verb type not containing a **PREPOSITIONAL COMPLEMENT** may appear here. In formal terms: any possible combination of non-**-PC-** verb type and active, middle or passive voice (**ACT**, **MIDD**, **PASS**) is handled by **PARTICIPLE PREDICATE/5**.
- **PARTICIPLE PREDICATE/6** covers active, middle and passive voice configurations. Any verb type containing a **PREPOSITIONAL COMPLEMENT** may appear here. In formal terms: any possible combination of **-PC-** verb type and active, middle or passive voice (**ACT**, **MIDD**, **PASS**) is handled by **PARTICIPLE PREDICATE/6**.

2.2.3.2 Verb Phrase *Predicates*

The rewrite rules of VPs are in part constituted by *predicates* that are repeatedly called upon by many of them. These *predicates* are the following:

2.2.3.2.1 case position in FC1Pr This predicate is only used in rewrite rules of functional clitic pronouns (Section 2.2.3.3). It imposes restrictions on the possible values of affix *case*, taking into account the position of the clitic pronoun.

```
case position in FC1Pr ( case, ClPr_Pos ):
  { case :: DAT|ACC }, { ClPr_Pos :: ONE_OF_ONE };
  { case :: DAT }, { ClPr_Pos :: ONE_OF_TWO }.
```

All the functional clitic pronouns that include this *predicate* in their rewrite rule are restricted to case dative when another functional clitic pronoun follows them. This is the case for first and second person functional clitic pronouns *me*, *te*, *nos*, *os*, which can be direct or indirect objects when they are alone attached to the verb form, but which are undoubtedly indirect objects when they are followed by *lo*, *la*, *los*, *las*.

2.2.3.2.2 verb type priorities This predicate is used in all the rules that rewrite **Verb** non-terminals into **VerbSt** non-terminals from the lexicon (Section 2.2.3.4). It avoids the identification of a verb as a main one if it is possible to identify it as an auxiliary or semiauxiliary one.

```
verb type priorities ( AUX ):.  
verb type priorities ( semiaux ):.  
verb type priorities ( mvtype ): $PENALTY ( 2 ).
```

In cases of structural ambiguity caused by the amount of information currently used within the grammar, this predicate ensures the selection of the more frequently correct analyses. For instance, on the basis of this predicate the ambiguity generated whenever verb *deber* happens to be followed by an infinitive¹⁹ is solved in favour of the semiauxiliary reading. This is, however, a momentaneous side-effect of the use of the predicate within the grammar in its current stage of development, because the interpretation of *deber* as a main verb in suchlike cases, should rather be overcome on the basis of syntactic and semantic subcategorization information about the arguments of the verb.

By contrast, it is in cases of real structural ambiguity where *verb type priorities* proves to be really useful. In these cases, it ensures the selection of the analyses considered more probable, while the others are downgraded by their lower priority. It is on the basis of this predicate, for example, that the ambiguity generated by verb *ir* in the context of a following preposition *a* and an infinitive²⁰ is solved in favour of the, more frequent, semiauxiliary reading.

¹⁹This ambiguity arises from the fact that, while infinitive is a possible realization of direct object, on the one hand, we have that main verb *deber* has a scheme *SDO, ACT* and, on the other hand, we also have that semiauxiliary verb *deber* requires the following verb to be an infinitive.

²⁰The same as in the above case of *deber*, this ambiguity arises from the fact that, while infinitive is a possible realization of prepositional complement, we also have that, on the one hand, main verb *ir* has a scheme *SPC, ACT, A* and, on the other hand, semiauxiliary verb *ir* requires the following verb to be an infinitive introduced by preposition *a*.

2.2.3.2.3 compound tense This predicate is used in all the rules that rewrite **Verb** non-terminals into **VerbSt** non-terminals for the generation of compound tense verb forms. It assigns the appropriate compound tense value, taking into account tense of the verb form of the auxiliary *haber*.

```
compound tense ( simple, compound ):
  { simple :: PRESENT }, { compound :: PRESENT_PERFECT };
  { simple :: PAST }, { compound :: PAST_PERFECT };
  { simple :: IMPERFECT }, { compound :: PLUPERFECT };
  { simple :: FUTURE }, { compound :: FUTURE_PERFECT };
  { simple :: CONDITIONAL }, { compound :: PAST_CONDITIONAL }.
```

2.2.3.2.4 two persons values This predicate is used in all the rules that combine **Verb** non-terminals with one functional clitic pronoun. It checks their respective *person* values to determine whether they can be combined or not.

```
two persons values ( mood, person, person1 ):
  { mood :: INDICATIVE|SUBJUNCTIVE|INFINITIVE|GERUND },
    { person :: FIRST|THIRD }, { person1 :: SECOND };
  { person :: FIRST|THIRD }, { person1 :: THIRD };
  { person :: THIRD }, { person1 :: FIRST };
  { person :: SECOND }, { person1 :: FIRST|THIRD }.
```

The predicate allows the combination of **FIRST** person verbs with **SECOND** and **THIRD** person clitic pronouns (*te/os espero, lo/los espero, te/os esperamos, lo/los esperamos*), of **SECOND** person verbs with **FIRST** and **THIRD** person clitic pronouns (*me/nos esperas, lo/los esperas, me/nos esperáis, lo/los esperáis*), and of **THIRD** person verbs with **FIRST**, **SECOND** and **THIRD** person clitic pronouns (*me/nos espera, te/os espera, lo/los espera, me/nos esperan, te/os esperan, lo/los esperan*). In **IMPERATIVE** mood the combination of **FIRST** and **THIRD** person with **SECOND** person clitic pronouns is blocked —because, on the one hand, **THIRD** person is in this context a second polite person and, on the other, **FIRST** person comprises the whole audience in **IMPERATIVE** mood. The predicate is organized in order to avoid that the pair **FIRST|THIRD** person, which is the person value assigned to a large series of syncretic verb forms in the lexicon (*amaba, amaría, ame, amara, amase, amare*), instead of producing only one analysis that takes an ambiguous value for the affix *person*, results in the appearance of two parse trees in the output, different only with respect to this affix value.

2.2.3.2.5 three persons values This predicate is used in all the rules that combine **Verb** non-terminals with two functional clitic pronouns. It checks their respective *person* values to determine whether they can be combined or not.

```
three persons values ( mood, person, person1, person2 ):
  { mood :: INDICATIVE|SUBJUNCTIVE|INFINITIVE|GERUND },
  { person :: FIRST|THIRD }, { person1 :: SECOND },
  { person2 :: THIRD };
  { person :: SECOND|THIRD }, { person1 :: FIRST },
  { person2 :: THIRD };
  { person1 :: THIRD }, { person2 :: THIRD }.
```

The predicate allows the combination of **FIRST** person verbs with **SECOND** person followed by **THIRD** person clitic pronouns (*te/os lo/los digo, te/os lo/los decimos*), of **SECOND** person verbs with **FIRST** person followed by **THIRD** person clitic pronouns (*me/nos lo/los dices, me/nos lo/los decís*), of **THIRD** person verbs with **FIRST** or **SECOND** person followed by **THIRD** person clitic pronouns (*me/nos lo/los dice, te/os lo/los dice, me/nos lo/los dicen, te/os lo/los dicen*), and of **FIRST**, **SECOND** and **THIRD** person verbs with two **THIRD** person clitic pronouns (*se lo/los digo, se lo/los decimos, se lo/los dices, se lo/los decís, se lo/los dice, se lo/los dicen*). In **IMPERATIVE** mood, as in previous predicate and for the same reasons, the combination of **FIRST** and **THIRD** person verbs with **SECOND** person followed by **THIRD** person clitic pronouns is blocked.

2.2.3.2.6 verb one ClPr This predicate is used in all the rules that combine main **Verb** non-terminals with one functional clitic pronoun to produce personal or impersonal non-*se* VPs. It checks their respective *vtype* and *case* values to determine whether they can be combined or not.

```
verb one ClPr ( one_clitic, case ):
  { one_clitic :: SDO|SDOIO|SDOPC|SDOPR|SDOIOPC }, { case :: ACC };
  { one_clitic :: SIO|SDOIO|SIOPC|SIOPR|SDOIOPC }, { case :: DAT };
  { one_clitic :: SPR|SIOPR }, { case :: PRTVE }.
```

Any **-DO-** verb type value can be combined with **ACC** case value. Any **-IO-** verb type value can be combined with **DAT** case value. Any **-PR-** verb type value except **SDOPR** can be combined with **PRTVE** case value. For all verb type values except **SDOIO** and **SDOIOPC**, this predicate disambiguates the case value

of the ambiguous clitic pronouns (*me, te, nos, os* for **ACC** and **DAT**, *lo* for **ACC** and **PRTVE**). For **SDOIO** and **SDOIOPC** values, on the contrary, the predicate permits the assignment of both case values in different alternatives. Despite this, given the general structure of the grammar, it does not cause the appearance of ambiguous parse trees in the output —because the exhaustive identification of arguments ensures that if only one of the arguments is explicit in the clause rewrite rule, the clitic pronoun should refer to the other one, if both of them are explicit instead, the reference to each complement should be described in different clause rules, in order to enable agreement to be used for the resolution of the ambiguity.

2.2.3.2.7 impersonal verb one ClPr This predicate is used in all the rules that combine main **Verb** non-terminals and one functional clitic pronoun to produce impersonal *se* VPs. It checks their respective *vtype* and *case* values to determine whether they can be combined or not.

```
impersonal verb one ClPr ( one_clitic, case ):
  { one_clitic :: SDO|SDOPC|SDOPR }, { case :: ACC };
  { one_clitic :: SIO|SDOIO|SIOPC|SIOPR|SDOIOPC }, { case :: DAT }.
```

There are two differences with respect to the previous predicate: neither case **PRTVE** is allowed by this predicate, because **PREDICATIVE COMPLEMENT** clitic pronouns cannot be combined with impersonal *se* VPs, nor **-DOIO-** verb types can be combined with **ACC** case value for clitic pronoun, because, for **-DOIO-** verb types, if a clitic pronoun *se* is present, it is always interpreted as a **DAT** case clitic pronoun, rather than as one indicating impersonality.

2.2.3.2.8 verb two ClPrs This predicate is used in all the rules that combine main **Verb** non-terminals and two functional clitic pronouns. It checks *vtype* of the main verb and *case* value of the second functional clitic pronoun to determine whether they can be combined or not.

```
verb two ClPrs ( two_clitics, case ):
  { two_clitics :: SDOIO|SDOIOPC }, { case :: ACC };
  { two_clitics :: SIOPR }, { case :: PRTVE }.
```

For verb types that can be combined with two functional clitic pronouns, those that contain a direct object are combined with an **ACC** case second functional

clitic pronoun, those that contain a predicative complement are combined with a **PRTVE** case second functional clitic pronoun. This predicate is useful to disambiguate clitic pronoun *lo* preceded by an indirect object clitic pronoun.

2.2.3.2.9 negative imperative person number This predicate is used in all the rules that rewrite negative imperative VPs into subjunctive verb forms. It ensures that the combination of values **FIRST** for affix *person* and **SING** for affix *number* will not be allowed in imperative mood.

```
negative imperative person number ( IMPERATIVE, person, number ):
  { person :: FIRST }, { number :: PLU };
  { person :: SECOND|THIRD }, { number :: SING|PLU }.
```

2.2.3.3 Clitic Pronouns

The lexical definitions of clitic pronouns are included in the grammar, not in the lexicon. There are different definitions for functional and non-functional clitic pronouns.

Functional clitic pronouns are rewritten as follows²¹:

```
Clitic Pronoun ( FIRST, MASC|FEM, SING, case, ClPr_pos ):
  case position in FC1Pr ( case, ClPr_pos ), "me".
```

```
Clitic Pronoun ( SECOND, MASC|FEM, SING, case, ClPr_pos ):
  case position in FC1Pr ( case, ClPr_pos ), "te".
```

```
Clitic Pronoun ( FIRST, MASC|FEM, PLU, case, ClPr_pos ):
  case position in FC1Pr ( case, ClPr_pos ), "nos".
```

```
Clitic Pronoun ( SECOND, MASC|FEM, PLU, case, ClPr_pos ):
  case position in FC1Pr ( case, ClPr_pos ), "os".
```

```
Clitic Pronoun ( THIRD, MASC|NEUT, SING, ACC,
  ONE_OF_ONE|TWO_OF_TWO ): "lo".
```

```
Clitic Pronoun ( THIRD, NEUT, SING, PRTVE,
  ONE_OF_ONE|TWO_OF_TWO ): "lo".
```

```
Clitic Pronoun ( THIRD, FEM, SING, ACC,
  ONE_OF_ONE|TWO_OF_TWO ): "la".
```

²¹Functional clitic pronouns have been defined according to their etymological behaviour. Extensions to account for so-called *leísmo* —or *laísmo*— can be easily added as penalized rewrite alternatives within the relevant rules.

Clitic Pronoun (THIRD, MASC, PLU, ACC,
ONE_OF_ONE|TWO_OF_TWO): "los".

Clitic Pronoun (THIRD, FEM, PLU, ACC,
ONE_OF_ONE|TWO_OF_TWO): "las".

Clitic Pronoun (THIRD, MASC|FEM|NEUT, SING|PLU, DAT,
ONE_OF_TWO): "se".

Clitic Pronoun (THIRD, MASC|FEM|NEUT, SING, DAT,
ONE_OF_ONE|TWO_OF_TWO): "le".

Clitic Pronoun (THIRD, MASC|FEM, PLU, DAT,
ONE_OF_ONE|TWO_OF_TWO): "les".

Non-functional clitic pronouns are rewritten as follows:

Clitic Pronoun (FIRST, MASC|FEM, SING, VERB): "me".

Clitic Pronoun (SECOND, MASC|FEM, SING, VERB): "te".

Clitic Pronoun (FIRST, MASC|FEM, PLU, VERB): "nos".

Clitic Pronoun (SECOND, MASC|FEM, PLU, VERB): "os".

Clitic Pronoun (THIRD, MASC|FEM|NEUT, SING|PLU, VERB): "se".

2.2.3.4 Verb non-terminals

Verb, **Verb_clitized**, **Participle Verb** and **Compound Verb_clitized** non-terminals are intermediate non-terminals (see above, Section 2.2.3) between non-periphrastic VP non-terminals and non-terminals rewritten as verb forms in the lexicon. Thus, on the one hand, they are members of the alternatives of the rewrite rules of non-terminals **Periphrasis**, **Verb Phrase**, **Non Finite Verb Phrase** and **Participle Verb Phrase**. On the other, **Verb**, **Verb_clitized**, **Participle Verb** and **Compound Verb_clitized** non-terminals rewrite as **VerbSt**, **VerbSt_clitized** and **Participle VerbSt** non-terminals of the lexicon: for each of the **VerbSt**-lexicon non-terminals that rewrite as finite, gerund or infinitive, or participle verb forms there is a parallel **Verb**-non-terminal working in the grammar.

Verb and **Verb_clitized** non-terminals —and their rewrite rules— for regular and orthographically modified, finite and non-finite (gerund and infinitive) verb

forms are the following ²²:

```
Verb ( vtype, voice, person, number, simple, mood ):
# he, soy, hay (que + infinitive)
# sigo (+ gerund), (me) pongo (a + infinitive)
# salgo, llueve, (me) visto, (se) trata
  VerbSt ( vtype, voice, person, number, simple, mood ),
  verbttype priorities ( vtype ).

Verb ( vtype, voice, simple, mood ):
# haber, ser, haber (que + infinitive)
# seguir (+ gerund), poner(se a + infinitive)
# salir, llover, vestir(se), tratar(se)
  VerbSt ( vtype, voice, simple, mood ),
  verbttype priorities ( vtype ).

Verb_clitized ( auxiliabile, voice, person, number, simple, mood,
  clitization ):
# sigue(lo) (+ gerund), écha(te a + infinitive)
# entréga(lo), viste(te)
  VerbSt_clitized ( auxiliabile, voice, person, number, simple,
mood, clitization ),
  verbttype priorities ( auxiliabile ).
```

²²To facilitate the understanding of the grammar, we will throughout the discussion illustrate each alternative with examples —introduced as AGFL comments preceding the alternatives— of verb phrases or fragments of them that should be parsed by the alternative in question.

Obviously, different verb types and voices of verbs are dealt with in different alternatives and rules. For the purposes of examples, we distinguish the following main groups of different behaviours of verbs within the VP. These groups of verb types are very helpful in the understanding of the structure of the formal description of the VP. We list here the example verbs used to represent each group, its definition and the schemes corresponding to the selected example verbs.

haber Auxiliary *haber*, (AUX, ACT|MIDD|PASS).

ser Passive voice semiauxiliary *ser*, (PARTICIPLE, PASS).

haber (que + infinitive) Impersonal active voice semiauxiliary *haber que + infinitive*, (INFINITIVE, ACT+IMP).

seguir (+ gerund), dejar (de + infinitive) Active voice semiauxiliaries, respectively (GERUND, ACT) and (INFINITIVE, ACT).

echar(se a + infinitive), poner(se a + infinitive) Middle voice semiauxiliaries (INFINITIVE, MIDD).

salir, llegar, vender, querer, entregar, presentar, considerar, obligar, poner, hacer Active voice main verbs, respectively (S, ACT), (SPC, ACT, DE|A), (SDO, ACT) —inanimate direct object—, (SDO, ACT) —animate direct object—, (SDOIO, ACT) —inanimate direct object—, (SDOIO, ACT) —animate direct object—, (SDOPR, ACT), (SDOPC, ACT, A), (SDOIOPC, ACT, EN) and (SDO, ACT) or (SDOIO, ACT).

llover, haber Impersonal active voice main verbs, respectively (S, ACT_IMP), (SDO, ACT_IMP).

vestir(se), creer(se), hacer(se) Middle voice main verbs, respectively (S, MIDD), (SDO, MIDD) and (SDO, MIDD).

tratar(se) Impersonal middle voice main verbs (SPC, MIDD_IMP).

```

Verb_clitized ( vtype, voice, simple, mood, clitization ):
# haber(selo), siendo(le), haber(selo) (que + infinitive)
# seguir(selo) (+ gerund), poniendo(se a + infinitive)
# entregár(selo), habiéndolo), creér(selo), tratádo(se)
  VerbSt_clitized ( vtype, voice, simple, mood, clitization ),
  verbtype priorities ( vtype ).

```

Obviously, the first two rules correspond to regular finite and non-finite verb forms, the last two to orthographically modified verb forms. Note also the use of affix non-terminal *vtype* in all these rules except for the third; as we already know, this affix non-terminal dominates all the possible verb types of the verb forms in the lexicon, included that of the auxiliary verb *haber*. Such an affix non-terminal, however, will no longer be used in the grammar, given that none of the non-terminals of the grammar, apart from **Verb**, can present value **AUX** in the position of affix *vtype*. In the entire grammar, indeed, value **AUX** is only specified in the **Verb** non-terminals that constitute the first member of the alternatives on the right-hand side of the rewrite rules of compound tenses, see below. In consequence, the corresponding rewrite rules for verbs that specify some type of prepositional requirement, as well as the third rule above, which accounts for orthographically modified imperative verb forms, which do not exist in the case of the auxiliary *haber*, might restrict the values of affix verb type to those values in the domain of affix non-terminal *auxiliable*:

```

Verb ( auxiliable, voice, linktype, person, number, simple, mood ):
  VerbSt ( auxiliable, voice, linktype, person, number, simple,
    mood ),
  verbtype priorities ( auxiliable ).

```

```

Verb ( auxiliable, voice, linktype, simple, mood ):
  VerbSt ( auxiliable, voice, linktype, simple, mood ),
  verbtype priorities ( auxiliable ).

```

```

Verb_clitized ( auxiliable, voice, linktype, person, number, simple,
  mood, clitization ):
  VerbSt_clitized ( auxiliable, voice, linktype, person, number,
    simple, mood, clitization ),
  verbtype priorities ( auxiliable ).

```

```

Verb_clitized ( auxiliable, voice, linktype, simple, mood,
  clitization ):
  VerbSt_clitized ( auxiliable, voice, linktype, simple, mood,
    clitization ),

```

verbtype priorities (auxiliable).

The rewrite rule of **Participle Verb** non-terminal is the following:

```
Participle Verb ( auxiliable, voice, mood, gender, number ):  
# sido, habido (que + infinitive)  
# seguido (+ gerund), puesto (a + infinitive)  
# salido, llovido, vestido, tratado  
  Participle VerbSt ( auxiliable, voice, mood, gender, number ),  
  verbtype priorities ( auxiliable ).
```

We need specific rules for the production of passive participles, which must be obtained from participle verb forms with **-DO-** verb type and **ACT** voice²³:

```
Participle Verb ( mvtype, PASS, mood, gender, number ):  
# vendido, entregado, considerado  
  { mvtype :: SDO|SDOIO|SDOPR },  
  Participle VerbSt ( mvtype, ACT, mood, gender, number ),  
  verbtype priorities ( mvtype ).  
  
Participle Verb ( mvtype, PASS, linktype, mood, gender, number ):  
# obligado, puesto  
  { mvtype :: SDOPC|SDOIOPC },  
  Participle VerbSt ( mvtype, ACT, linktype, mood, gender,  
    number ),  
  verbtype priorities ( mvtype ).
```

To account for compound tense verb forms, **Verb** non-terminals rewrite as **Verb** non-terminals of auxiliary verb *haber* followed by **Participle Verb** non-terminals of main verbs. Because they are considered a single unity, the inflected form of the auxiliary *haber* and the participle are assigned the same voice value.

```
Verb ( auxiliable, voice, person, number, compound, mood ):  
# he sido, ha habido (que + infinitive)  
# he seguido (+ gerund), (me) he puesto (a + infinitive)  
# he salido, ha llovido, (me) he vestido, (se) ha tratado  
  Verb ( AUX, voice, person, number, simple, mood ),  
  Participle Verb ( auxiliable, voice, mood1, MASC, SING ),  
  compound tense ( simple, compound ).
```

```
Verb ( auxiliable, voice, compound, mood ):  
# haber sido, haber habido (que + infinitive)  
# haber seguido (+ gerund), haber puesto (a + infinitive)  
# haber salido, haber llovido, haber vestido, haber tratado
```

²³Henceforth, if, as in this case, the only difference between rewrite rules specifying and not specifying prepositional requirements lies on verb type restrictions, we will only include the first.

```

Verb ( AUX, voice, simple, mood ),
  Participle Verb ( auxiliabile, voice, mood1, MASC, SING ),
  compound tense ( simple, compound ).

```

By means of these rewrite rules compound tenses are paired with simple tenses, so that **Verb** non-terminals found in **Periphrasis**, **Verb Phrase** and **Non Finite Verb Phrase** rewrite rules can simultaneously call upon both simple and compound tense verb forms.

Another remarkable point with respect to these rules is the production of middle voice non-finite compound verb forms in which the middle voice clitic pronoun is not present between the auxiliary and the participle. These, apparently wrong, middle voice verb forms that lack the clitic pronoun are necessary to account for periphrastic VPs in which the clitic pronoun has been attached to a previous semiauxiliary: *se puede haber puesto a*, *se puede haber vestido*.

Apart from the previous ones, the combination of non-finite compound verb forms with clitic pronouns requires the introduction of specific rules to account for the fact that the clitic pronoun is in these cases attached to the infinitive or gerund auxiliary, inserted between this one and the participle of the inflected verb. For these forms we have a **Compound Verb_clitized** non-terminal that includes the necessary information about the clitic pronouns possibly attached to the non-finite verb form.

To present the rewrite rules of **Compound Verb_clitized** non-terminals we proceed from **Compound Verb_clitized** non-terminals to which one clitic pronoun has been attached to **Compound Verb_clitized** non-terminals to which two clitic pronouns have been attached.

When it includes just one clitic pronoun and has a participle that is main verb, the rewrite rule of non-terminal **Compound Verb_clitized** is the following:

```

Compound Verb_clitized ( mvtype, voice, compound, mood, case_one,
  person_one, gender_one, number_one ):
# haberlo/habiéndolo vendido, haberlo/habiéndolo habido
  { voice :: ACT|ACT_IMP },
  ( { mood :: INFINITIVE },
    Verb ( AUX, voice, simple, mood );
    Verb_clitized ( AUX, voice, simple, mood, ONE ) ),
    Clitic Pronoun ( person_one, gender_one, number_one,
      case_one, ONE_OF_ONE ),
    Participle Verb ( mvtype, voice, PARTICIPLE, MASC, SING ),

```

```

        compound tense ( simple, compound );
# haberme/habiéndome vestido, haberse/habiéndose tratado
  { voice :: MIDD|MIDD_IMP },
    ( { mood :: INFINITIVE },
      Verb ( AUX, voice, simple, mood );
      Verb_clitized ( AUX, voice, simple, mood, ONE ) ),
      Clitic Pronoun ( person_one, gender_one, number_one,
        case_one ),
      Participle Verb ( mvtype, voice, PARTICIPLE, MASC, SING ),
      compound tense ( simple, compound );
# haberse/habiéndose salido, haberse/habiéndose vendido
  { voice :: ACT_IMP|ACT_IMP+PASS },
    ( { mood :: INFINITIVE },
      Verb ( AUX, ACT, simple, mood );
      Verb_clitized ( AUX, ACT, simple, mood, ONE ) ),
      Clitic Pronoun ( person_one, gender_one, number_one,
        case_one ),
      Participle Verb ( mvtype, ACT, PARTICIPLE, MASC, SING ),
      compound tense ( simple, compound ).

```

The first alternative combines **ACT** or **ACT_IMP** main verbs with one functional clitic pronoun: this results in a **Compound Verb_clitized** non-terminal with the same *voice* value. The second alternative combines **MIDD** or **MIDD_IMP** main verbs with one non-functional clitic pronoun: this results in a **Compound Verb_clitized** non-terminal with the same *voice* value. The third alternative combines **ACT** main verbs with one non-functional clitic pronoun: this results in a **Compound Verb_clitized** non-terminal with **ACT_IMP** or **ACT_IMP+PASS** voice value. Both the third and the first alternatives produce impersonal active voice (**ACT_IMP**) **Compound Verb_clitized** non-terminals, the values for affix case are used by the rewrite rules of **Verb Phrase** to distinguish them: **VERB** is the *case* value of the clitic pronoun in impersonal *se* VPs, and values in the domain of affix non-terminal *function_case* are the relevant *case* values for impersonal VPs that include an inherently impersonal verb.

When it includes just one clitic pronoun and has a participle that is an auxiliary verb which does not require an intermediate connector, the rewrite rule of non-terminal **Compound Verb_clitized** is the following:

```

Compound Verb_clitized ( semiaux, voice, compound, mood, case_one,
  person_one, gender_one, number_one ):
# haberlo seguido (+ gerund)
  { voice :: ip_ACT }, { mood :: INFINITIVE }, { voice1 :: ACT },
  Verb ( AUX, voice1, simple, mood ),

```

```

        Clitic Pronoun ( person_one, gender_one, number_one, case_one,
            ONE_OF_ONE ),
        Participle Verb ( semiaux, voice1, PARTICIPLE, MASC, SING ),
        compound tense ( simple, compound );
# haberle sido
    { voice :: PASS }, { mood :: INFINITIVE },
    Verb ( AUX, voice, simple, mood ),
    Clitic Pronoun ( person_one, gender_one, number_one, case_one,
        ONE_OF_ONE ),
    Participle Verb ( semiaux, voice, PARTICIPLE, MASC, SING ),
    compound tense ( simple, compound );
# habiéndolo seguido (+ gerund), habiéndole sido
    { voice :: ip_ACT|PASS },
    Verb_clitized ( AUX, voice, simple, mood, ONE ),
    Clitic Pronoun ( person_one, gender_one, number_one, case_one,
        ONE_OF_ONE ),
    Participle Verb ( semiaux, voice, PARTICIPLE, MASC, SING ),
    compound tense ( simple, compound );
# haberme seguido (+ gerund)
    { voice :: ip_MIDD }, { mood :: INFINITIVE }, { voice1 :: ACT },
    Verb ( AUX, voice1, simple, mood ),
    Clitic Pronoun ( person_one, gender_one, number_one,
        case_one ),
    Participle Verb ( semiaux, voice1, PARTICIPLE, MASC, SING ),
    compound tense ( simple, compound );
# habiéndome seguido (+ gerund)
    { voice :: ip_MIDD },
    Verb_clitized ( AUX, voice, simple, mood, ONE ),
    Clitic Pronoun ( person_one, gender_one, number_one,
        case_one ),
    Participle Verb ( semiaux, voice, PARTICIPLE, MASC, SING ),
    compound tense ( simple, compound ).

```

When it includes just one clitic pronoun and has a participle that is an auxiliary verb which requires an intermediate connector, the rewrite rule of non-terminal **Compound Verb_clitized** is the following:

```

Compound Verb_clitized ( semiaux, voice, linktype, compound, mood,
    case_one, person_one, gender_one, number_one ):
# haberlo dejado (de + infinitive)
# haberlo habido (que + infinitive)
    ( { voice :: ip_ACT }, { voice1 :: ACT };
    { voice :: ip_ACT_IMP }, { voice1 :: ACT_IMP } ),
    { mood :: INFINITIVE },
    Verb ( AUX, voice1, simple, mood ),
    Clitic Pronoun ( person_one, gender_one, number_one, case_one,
        ONE_OF_ONE ),
    Participle Verb ( semiaux, voice1, linktype, PARTICIPLE, MASC,
        SING ),

```

```

        compound tense ( simple, compound );
# habiéndolo dejado (de + infinitive)
# habiéndolo habido (que + infinitive)
  { voice :: ip_ACT|ip_ACT_IMP },
    Verb_clitized ( AUX, voice, simple, mood, ONE ),
    Clitic Pronoun ( person_one, gender_one, number_one, case_one,
                     ONE_OF_ONE ),
    Participle Verb ( semiaux, voice, linktype, PARTICIPLE, MASC,
                     SING ),
    compound tense ( simple, compound );
# haberme dejado (de + infinitive)
  { voice :: ip_MIDD }, { mood :: INFINITIVE }, { voice1 :: ACT },
    Verb ( AUX, voice1, simple, mood ),
    Clitic Pronoun ( person_one, gender_one, number_one,
                     case_one ),
    Participle Verb ( semiaux, voice1, linktype, PARTICIPLE, MASC,
                     SING ),
    compound tense ( simple, compound );
# haberme puesto (a + infinitive)
  { voice :: MIDD }, { mood :: INFINITIVE },
    Verb ( AUX, voice, simple, mood ),
    Clitic Pronoun ( person_one, gender_one, number_one,
                     case_one ),
    Participle Verb ( semiaux, voice, linktype, PARTICIPLE, MASC,
                     SING ),
    compound tense ( simple, compound );
# habiéndome dejado (de + infinitive)
# habiéndome puesto (a + infinitive)
  { voice :: ip_MIDD|MIDD },
    Verb_clitized ( AUX, voice, simple, mood, ONE ),
    Clitic Pronoun ( person_one, gender_one, number_one,
                     case_one ),
    Participle Verb ( semiaux, voice, linktype, PARTICIPLE, MASC,
                     SING ),
    compound tense ( simple, compound ).

```

When it includes two clitic pronouns and has a participle that is a main verb, the rewrite rule of non-terminal **Compound Verb_clitized** is the following:

```

Compound Verb_clitized ( mvtype, voice, compound, mood, case_one,
                          person_one, gender_one, number_one, case_two, person_two,
                          gender_two, number_two ):
# habérselo entregado, habiéndoselo entregado
  { voice :: ACT },
    Verb_clitized ( AUX, voice, simple, mood, ONE|TWO ),
    Clitic Pronoun ( person_one, gender_one, number_one, case_one,
                     ONE_OF_TWO ),
    Clitic Pronoun ( person_two, gender_two, number_two, case_two,
                     TWO_OF_TWO ),

```

```

    Participle Verb ( mvtype, voice, PARTICIPLE, MASC, SING ),
    compound tense ( simple, compound );
# habérmelo creído, habiéndomelo creído
{ voice :: MIDD },
    Verb_clitized ( AUX, voice, simple, mood, ONE|TWO ),
    Clitic Pronoun ( person_one, gender_one, number_one,
                    case_one ),
    Clitic Pronoun ( person_two, gender_two, number_two, case_two,
                    ONE_OF_ONE ),
    Participle Verb ( mvtype, voice, PARTICIPLE, MASC, SING ),
    compound tense ( simple, compound );
# habérselo querido, habersele entregado
{ voice :: ACT_IMP|ACT_IMP+PASS },
    Verb_clitized ( AUX, ACT, simple, mood, ONE|TWO ),
    Clitic Pronoun ( person_one, gender_one, number_one,
                    case_one ),
    Clitic Pronoun ( person_two, gender_two, number_two, case_two,
                    ONE_OF_ONE ),
    Participle Verb ( mvtype, ACT, PARTICIPLE, MASC, SING ),
    compound tense ( simple, compound ).

```

The first alternative combines **ACT** main verbs with two functional clitic pronouns: this results in a **Compound Verb_clitized** non-terminal with the same *voice* value. The second alternative combines **MIDD** main verbs with one non-functional clitic pronoun and one functional clitic pronoun: this results in a **Compound Verb_clitized** non-terminal with the same *voice* value. The third alternative combines **ACT** voice main verbs with one non-functional clitic pronoun and one functional clitic pronoun: this results in a **Compound Verb_clitized** non-terminal with **ACT_IMP** or **ACT_IMP+PASS** voice values.

When it includes two clitic pronouns and has a participle that is an auxiliary verb which does not require an intermediate connector, the rewrite rule of non-terminal **Compound Verb_clitized** is the following:

```

Compound Verb_clitized ( semiaux, voice, compound, mood, case_one,
    person_one, gender_one, number_one, case_two, person_two,
    gender_two, number_two ):
# habérselo/habiéndoselo seguido (+ gerund)
# habérmelo/habiéndomelo seguido (+ gerund)
    Verb_clitized ( AUX, voice, simple, mood, ONE|TWO ),
    ( { voice :: ip_ACT },
        Clitic Pronoun ( person_one, gender_one, number_one,
                        case_one, ONE_OF_TWO ),
        Clitic Pronoun ( person_two, gender_two, number_two,
                        case_two, TWO_OF_TWO );
    { voice :: ip_MIDD },

```

```

        Clitic Pronoun ( person_one, gender_one, number_one,
                        case_one ),
        Clitic Pronoun ( person_two, gender_two, number_two,
                        case_two, ONE_OF_ONE ),
        Participle Verb ( semiaux, voice, PARTICIPLE, MASC, SING ),
        compound tense ( simple, compound ).

```

When it includes two clitic pronouns and has a participle that is an auxiliary verb which requires an intermediate connector, the rewrite rule of non-terminal **Compound Verb_clitized** is the following:

```

Compound Verb_clitized ( semiaux, voice, linktype, compound, mood,
                        case_one, person_one, gender_one, number_one, case_two,
                        person_two, gender_two, number_two ):
# habérselo/habiéndoselo dejado (de + infinitive)
# habérselo/habiéndoselo habido (que + infinitive)
# habérmelo/habiéndomelo dejado (de + infinitive)
  Verb_clitized ( AUX, voice, simple, mood, ONE|TWO ),
    ( { voice :: ip_ACT|ip_ACT_IMP },
      Clitic Pronoun ( person_one, gender_one, number_one,
                      case_one, ONE_OF_TWO ),
      Clitic Pronoun ( person_two, gender_two, number_two,
                      case_two, TWO_OF_TWO );
    { voice :: ip_MIDD },
      Clitic Pronoun ( person_one, gender_one, number_one,
                      case_one ),
      Clitic Pronoun ( person_two, gender_two, number_two,
                      case_two, ONE_OF_ONE ) ),
  Participle Verb ( semiaux, voice, linktype, PARTICIPLE, MASC,
                  SING ),
  compound tense ( simple, compound ).

```

2.2.3.5 Non-Periphrastic VPs: Verb Phrase, Non Finite Verb Phrase and Participle Verb Phrase

First in this Section, **Verb Phrase** non-terminals and their rewrite rules are explained, proceeding from rules for **Verb Phrase** non-terminals that do not include functional clitic pronouns, to those that include one and two of them. Then, production of **Non Finite Verb Phrase** non-terminal is addressed, proceeding also from rules for **Non Finite Verb Phrase** non-terminals that do not include clitic pronouns to those that include one and two of them²⁴. In each of these groups of rules, the rewrite rules for different *voice* values of **Verb Phrase** and **Non Finite**

²⁴Note that we are now taking into account both functional and non-functional clitic pronouns.

Verb Phrase non-terminals will be separated. **Participle Verb Phrase**, finally, is described at the end of the Section.

2.2.3.5.1 Verb Phrase/5 Verb Phrase non-terminals without functional clitic pronouns account for a number of active, middle or passive personal or impersonal configurations, the whole collection of possible PREDICATE *voice* values: **ACT**, **ACT_IMP**, **MIDD**, **MIDD_IMP**, **PASS**, **PASS_IMP**, **ACT_-IMP+PASS**.

Without non-functional clitic pronouns, non-terminal **Verb Phrase/5** accounts for active personal or impersonal configurations (**ACT**, **ACT_IMP**), based on, respectively, active personal (**ACT**) and impersonal (**ACT_IMP**) schemes.

```
Verb Phrase ( mvtype, act, person, number, mood ):
# (no) salgo, (no) salga, sal
# (no) llueve, (no) llueva
  [ { mood :: INDICATIVE|SUBJUNCTIVE },
    "no" ],
    Verb ( mvtype, act, person, number, tense, mood );
# no salgas
  { act :: ACT },
    "no",
    Verb ( mvtype, act, person, number, PRESENT, SUBJUNCTIVE ),
    negative imperative person number ( mood, person, number ).
```

The first alternative identifies non-periphrastic VPs constituted by one verb form in **INDICATIVE**, **SUBJUNCTIVE** or **IMPERATIVE** mood for affirmative polarity, only in **INDICATIVE** or **SUBJUNCTIVE** mood for negative polarity. The second alternative identifies non-periphrastic **IMPERATIVE** VPs for negative polarity, constituted by one verb form in **SUBJUNCTIVE** mood. *Mood* of the target non-terminal is in this case determined by the predicate **negative imperative person number**, whose primary function, as stated in Section 2.2.3.2, is to prevent the identification of first person singular verb forms in imperative. The structure of this rule shows the strategy used throughout the grammar to deal formally with the problem posed by polarity of imperative mood. Because this is a recurrent structure, with only slight variations, in the rewrite rules of non-periphrastic VPs and periphrastic groups, in the rest of this exposition we will no longer show the alternatives accounting for negative imperative mood. Note, finally, that compound tenses are also possible realizations of

VPs described by these rules, since **Verb** non-terminals cover both simple and compound tenses (see Section 2.2.3.4). The same is valid for all VP rules with **Verb** non-terminals in their right-hand sides.

Without non-functional clitic pronouns, non-terminal **Verb Phrase/5** accounts also for passive personal configurations (**PASS**)²⁵.

```
Verb Phrase ( mvtype, PASS, person, number, mood ):
# (no) es vendido
  [ "no" ],
  Verb ( PARTICIPLE, PASS, person, number, tense, mood ),
  Participle Verb ( mvtype, PASS, mood1, gender, number ).
```

With one non-functional clitic pronoun, non-terminal **Verb Phrase/5** accounts for middle personal and middle, active or passive impersonal configurations (**MIDD**, **MIDD_IMP**, **ACT_IMP+PASS**, **ACT_IMP**, **PASS_IMP**).

```
Verb Phrase ( mvtype, voice, person, number, mood ):
# (no) me visto, (no) me vista
# (no) se trata, (no) se trate
  { voice :: MIDD|MIDD_IMP }, { mood :: INDICATIVE|SUBJUNCTIVE },
  [ "no" ],
  Clitic Pronoun ( person, gender, number, case ),
  Verb ( mvtype, voice, person, number, tense, mood );
# vistete
  { voice :: MIDD },
  Verb_clitized ( mvtype, voice, person, number, tense, mood,
    ONE|JUST_ONE ),
  Clitic Pronoun ( person, gender, number, case );
# no te vistas
(...);
# (no) se vende, (no) se venda
  { mvtype :: SDO|SDOIO|SDOPR }, { voice :: ACT_IMP+PASS },
  { person :: THIRD }, { mood :: INDICATIVE|SUBJUNCTIVE },
  [ "no" ],
  Clitic Pronoun ( person, gender, number, case ),
  Verb ( mvtype, ACT, person, number, tense, mood );
# (no) se sale, (no) se salga
  { voice :: ACT_IMP }, { person :: THIRD }, { number :: SING },
  { mood :: INDICATIVE|SUBJUNCTIVE },
  [ "no" ],
  Clitic Pronoun ( person, gender, number, case ),
  Verb ( mvtype, ACT, person, number, tense, mood );
# (no) se es querido
  { voice :: PASS_IMP }, { person :: THIRD }, { number :: SING },
```

²⁵Although this is veiled here, remind that passive configurations are always based on **-DO-** active personal (**ACT**) schemes (see rewrite rule of **Participle Verb** in Section 2.2.3.4).

```

[ "no" ],
Clitic Pronoun ( person, gender_one, number, case_one ),
Verb ( PARTICIPLE, PASS, person, number, tense, mood ),
Participle Verb ( mvtype, PASS, mood1, gender, number ).

```

All the alternatives, except for the second and the third ones, identify non-periphrastic VPs in **INDICATIVE** or **SUBJUNCTIVE** mood, for both affirmative and negative polarity. The first alternative identifies middle personal or impersonal configurations (**MIDD**, **MIDD_IMP**) based on, respectively, middle personal (**MIDD**) or impersonal (**MIDD_IMP**) schemes. The fourth alternative identifies active impersonal configurations with subject concord (**ACT_IMP+PASS**) based on **-DO-** active personal (**ACT**) schemes. *Mvtype* values are conveniently restricted by means of a guard. Another guard takes care of the fact that only value **THIRD** is allowed for affix *person* in this context. The fifth alternative identifies active impersonal configurations (**ACT_IMP**) based on active personal (**ACT**) schemes. A guard takes care of the fact that only values **THIRD** and **SING** are allowed for, respectively, affixes *person* and *number*. The sixth alternative identifies passive impersonal configurations (**PASS_IMP**) based on **-DO-** active personal (**ACT**) schemes. A guard restricts the values for affixes *person* and *number* to, respectively, **THIRD** and **SING**. The second and the third alternatives, finally, identify non-periphrastic VPs in **IMPERATIVE** mood, for, respectively, affirmative and negative polarity. Since clitic pronouns occur in different positions in indicative or subjunctive moods and in imperative mood, every alternative of the type of the first one in this rule must be supplemented by a parallel alternative that accounts for basically the same verb type, voice and combination of clitic pronouns, but for imperative mood. In the rest of this exposition, we will no longer display these supplementary alternatives that account for affirmative imperative VPs.

2.2.3.5.2 Verb Phrase/9 Verb Phrase non-terminals with one functional clitic pronoun account for a number of active, middle or passive personal or impersonal configurations, the whole collection of PREDICATE voice values: **ACT**, **ACT_IMP**, **MIDD**, **MIDD_IMP**, **PASS**, **PASS_IMP**, **ACT_IMP+PASS**. All alternatives of these rewrite rules include the predicate **two persons values** to check the compatibility of the values for affix *person* of the functional clitic

pronoun and the verb form. Almost all of the alternatives —except those that account for passive voice— include also the predicate **verb one ClPr**, or its alternative formulation for impersonal *se* configurations, **impersonal verb one ClPr**, to regulate the combination between the values for affixes *mvtype* and *case* of, respectively, the verb form and the functional clitic pronoun.

Without non-functional clitic pronouns, non-terminal **Verb Phrase/9** accounts for active personal or impersonal configurations (**ACT**, **ACT_IMP**), based on, respectively, active personal (**ACT**) and impersonal (**ACT_IMP**) schemes.

```
Verb Phrase ( one_clitic, act, person, number, mood, case_one,
             person_one, gender_one, number_one ):
# (no) lo vendo, (no) lo venda
# (no) lo hay, (no) lo haya
  { mood :: INDICATIVE|SUBJUNCTIVE },
  [ "no" ],
  Clitic Pronoun ( person_one, gender_one, number_one, case_one,
                  ONE_OF_ONE ),
  Verb ( one_clitic, act, person, number, tense, mood ),
  two persons values ( mood, person, person_one ),
  verb one ClPr ( one_clitic, case_one );
# véndelo
(...);
# no lo vendas
(...).
```

The first alternative identifies indicative or subjunctive active personal or impersonal configurations (**ACT**, **ACT_IMP**) based on, respectively, active personal (**ACT**) or impersonal (**ACT_IMP**) schemes. The second and the third alternatives identify imperative active personal configurations (**ACT**) based on active personal (**ACT**) schemes.

Without non-functional clitic pronouns, non-terminal **Verb Phrase/9** accounts also for passive personal configurations (**PASS**).

```
Verb Phrase ( one_clitic, PASS, person, number, mood, case_one,
             person_one, gender_one, number_one ):
# (no) le es entregado
  { one_clitic :: SD0IO }, { case_one :: DAT },
  [ "no" ],
  Clitic Pronoun ( person_one, gender_one, number_one, case_one,
                  ONE_OF_ONE ),
  Verb ( PARTICIPLE, PASS, person, number, tense, mood ),
  Participle Verb ( one_clitic, PASS, PARTICIPLE, gender,
                   number ),
  two persons values ( mood, person, person_one ).
```

A guard must restrict the value of affix *vtype* to **SDOIO**, since this is the only **-DO-** scheme that permits the introduction of a clitic pronoun with value **DAT** for affix case.

With one non-functional clitic pronoun, non-terminal **Verb Phrase/9** accounts for middle personal and middle, active or passive impersonal configurations (**MIDD**, **MIDD_IMP**, **ACT_IMP+PASS**, **ACT_IMP**, **PASS_IMP**).

```
Verb Phrase ( one_clitic, voice, person, number, mood, case_one,
             person_one, gender_one, number_one ):
# (no) me lo creo, (no) me lo crea
  { voice :: MIDD|MIDD_IMP }, { mood :: INDICATIVE|SUBJUNCTIVE },
  [ "no" ],
  Clitic Pronoun ( person, gender, number, case ),
  Clitic Pronoun ( person_one, gender_one, number_one, case_one,
                  ONE_OF_ONE ),
  Verb ( one_clitic, voice, person, number, tense, mood ),
  two persons values ( mood, person, person_one ),
  verb one ClPr ( one_clitic, case_one );
# créetelo
(...);
# no te lo creas
(...);
# (no) se le entrega, (no) se le entregue
  { one_clitic :: SDOIO }, { voice :: ACT_IMP+PASS },
  { person :: THIRD }, { mood :: INDICATIVE|SUBJUNCTIVE },
  { case_one :: DAT },
  [ "no" ],
  Clitic Pronoun ( person, gender, number, case ),
  Clitic Pronoun ( person_one, gender_one, number_one, case_one,
                  ONE_OF_ONE ),
  Verb ( one_clitic, ACT, person, number, tense, mood );
# (no) se lo quiere, (no) se lo quiera
  { voice :: ACT_IMP }, { person :: THIRD }, { number :: SING },
  { mood :: INDICATIVE|SUBJUNCTIVE },
  [ "no" ],
  Clitic Pronoun ( person, gender, number, case ),
  Clitic Pronoun ( person_one, gender_one, number_one, case_one,
                  ONE_OF_ONE ),
  Verb ( one_clitic, ACT, person, number, tense, mood ),
  impersonal verb one ClPr ( one_clitic, case_one );
# (no) se le es presentado
  { one_clitic :: SDOIO }, { voice :: PASS_IMP },
  { person :: THIRD }, { number :: SING }, { case_one :: DAT },
  [ "no" ],
  Clitic Pronoun ( person, gender, number, case ),
  Clitic Pronoun ( person_one, gender_one, number_one, case_one,
                  ONE_OF_ONE ),
  Verb ( PARTICIPLE, PASS, person, number, tense, mood ),
```

```
Participle Verb ( one_clitic, PASS, PARTICIPLE, gender,
                 number ).
```

Except for the introduction of the functional clitic pronoun and the predicates **two persons values** and **verb one ClPr** (or **impersonal verb one ClPr**), this set of alternatives identifies the same set of configurations, based on the same set of schemes, as the alternatives that constituted the rewrite rule, displayed above, of non-terminal **Verb Phrase/5** with only one non-functional clitic pronoun.

2.2.3.5.3 Verb Phrase/13 Verb Phrase non-terminals with two functional clitic pronouns account for a number of active personal configurations, **ACT** configurations. All alternatives include the predicate **three persons values** to check the compatibility of the values for affix *person* of the functional clitic pronouns and the verb form. All of them include also the predicate **verb two ClPrs** to regulate the combination of the values for *mvtype* and *case* of, respectively, the verb form and the second functional clitic pronoun.

```
Verb Phrase ( two_clitics, ACT, person, number, mood, case_one,
              person_one, gender_one, number_one, case_two, person_two,
              gender_two, number_two ):
# (no) se lo entrego, (no) se lo entregue
  { mood :: INDICATIVE|SUBJUNCTIVE },
  [ "no" ],
  Clitic Pronoun ( person_one, gender_one, number_one, case_one,
                  ONE_OF_TWO ),
  Clitic Pronoun ( person_two, gender_two, number_two, case_two,
                  TWO_OF_TWO ),
  Verb ( two_clitics, ACT, person, number, tense, mood ),
  three persons values ( mood, person, person_one, person_two ),
  verb two ClPrs ( two_clitics, case_two );
# entrégaselo
(...);
# no se lo entregues
(...).
```

The first alternative identifies indicative or subjunctive active personal configurations (**ACT**) based on active personal (**ACT**) schemes. The second and the third alternatives identify imperative active personal configurations (**ACT**) based on active personal (**ACT**) schemes.

2.2.3.5.4 Non Finite Verb Phrase/3 Non-terminal **Non Finite Verb Phrase** without clitic pronouns accounts for active configurations, **ACT** configurations.

```
Non Finite Verb Phrase ( mvtype, act, mood ):
# (no) salir, (no) llover
  [ "no" ],
  Verb ( mvtype, act, tense, mood ).
```

This rule identifies infinitive and gerund active personal configurations (**ACT**), based on active personal (**ACT**) schemes. It appears to identify also impersonal active voice (**ACT_IMP**) configurations based on impersonal active voice schemes. In this case, however, we have to consider that impersonal active voice **Non Finite Verb Phrase** non-terminals cannot be contextualized by themselves as PREDICATE functions of clause description rules, so we should not use for these examples the term configurations. In spite of the fact that they do not constitute configurations, such non-terminals must be produced, to make possible their integration in impersonal finite periphrastic VPs (*puede llover*). All the alternatives that account for apparently impersonal **Non Finite Verb Phrase** non-terminals are present in their rewrite rules for this reason.

A similar situation can be observed in the case of **MIDD** schemes that are the base for **Non Finite Verb Phrase/3** non-terminals lacking the obligatory non-functional clitic pronoun. Although neither do these **Non Finite Verb Phrase** non-terminals, as those above and for the same reason, describe configurations, they must be also accounted for because they are necessary for the production of the main groups of periphrastic VPs in which the clitic pronoun has been moved to the periphrastic group (*se puede vestir*). **MIDD_IMP** schemes obviously combine both problems (*se puede tratar de una trampa*).

```
Non Finite Verb Phrase ( mvtype, midd, mood ):
# (no) vestir, (no) tratar
  [ "no" ],
  Verb ( mvtype, midd, tense, mood ).
```

Non-terminal **Non Finite Verb Phrase/3** accounts also for passive configurations, **PASS** configurations.

```
Non Finite Verb Phrase ( mvtype, PASS, mood ):
# (no) ser vendido
```

```
[ "no" ],
  Verb ( PARTICIPLE, PASS, tense, mood ),
  Participle Verb ( mvtype, PASS, mood1, gender, number ).
```

2.2.3.5.5 Non Finite Verb Phrase/7 Non Finite Verb Phrase non-terminals with one clitic pronoun account for active, middle and passive configurations, **ACT**, **MIDD** and **PASS** configurations. Alternatives for simple and compound tenses are separated in the rules for non-finite VPs that include clitic pronouns, due to the fact that clitic pronouns are placed in compound tenses between the infinitive or gerund auxiliary and the participle main verb (see Section 2.2.3.4). For active and passive voice, the clitic pronoun is functional. In active voice, the predicate **verb one ClPr** is introduced to check the compatibility between the values for affixes *mvtype* and *case* of, respectively, the verb form and the clitic pronoun. For impersonal and middle voices, the clitic pronoun is non-functional.

The rewrite rule for active or middle personal or impersonal voices is the following:

```
Non Finite Verb Phrase ( mvtype, voice, mood, case_one,
  person_one, gender_one, number_one ):
# (no) venderlo, (no) haberlo
  { voice :: ACT|ACT_IMP },
  [ "no" ],
  ( { mood :: INFINITIVE },
    Verb ( mvtype, voice, simple, mood );
    Verb_clitized ( mvtype, voice, simple, mood, ONE ) ),
  Clitic Pronoun ( person_one, gender_one, number_one, case_one,
    ONE_OF_ONE ),
  verb one ClPr ( mvtype, case_one );
# (no) vestirme
  { voice :: MIDD },
  [ "no" ],
  ( { mood :: INFINITIVE },
    Verb ( mvtype, voice, simple, mood );
    Verb_clitized ( mvtype, voice, simple, mood, ONE ) ),
  Clitic Pronoun ( person_one, gender_one, number_one,
    case_one );
# (no) venderse
  { mvtype :: SDO|SDOIO|SDOPR }, { voice :: ACT_IMP+PASS },
  { person_one :: THIRD },
  [ "no" ],
  ( { mood :: INFINITIVE },
    Verb ( mvtype, ACT, simple, mood );
```

```

        Verb_clitized ( mvtype, ACT, simple, mood, ONE ) ),
        Clitic Pronoun ( person_one, gender_one, number_one,
            case_one );
# (no) salirse
{ voice :: ACT_IMP }, { person_one :: THIRD },
{ number_one :: SING },
[ "no" ],
( { mood :: INFINITIVE },
    Verb ( mvtype, ACT, simple, mood );
    Verb_clitized ( mvtype, ACT, simple, mood, ONE ) ),
    Clitic Pronoun ( person_one, gender_one, number_one,
        case_one );
# (no) tratarse
{ voice :: MIDD_IMP }, { person_one :: THIRD },
{ number_one :: SING },
[ "no" ],
( { mood :: INFINITIVE },
    Verb ( mvtype, voice, simple, mood );
    Verb_clitized ( mvtype, voice, simple, mood, ONE ) ),
    Clitic Pronoun ( person_one, gender_one, number_one,
        case_one );
# (no) haberlo vendido, (no) haberlo habido
{ voice :: ACT|ACT_IMP },
[ "no" ],
Compound Verb_clitized ( mvtype, voice, compound, mood,
    case_one, person_one, gender_one, number_one ),
verb one ClPr ( mvtype, case_one );
# (no) haberme vestido
{ voice :: MIDD },
[ "no" ],
Compound Verb_clitized ( mvtype, voice, compound, mood,
    case_one, person_one, gender_one, number_one );
# (no) haberse vendido
{ mvtype :: SDO|SDOIO|SDOPR }, { voice :: ACT_IMP+PASS },
{ person_one :: THIRD },
[ "no" ],
Compound Verb_clitized ( mvtype, voice, compound, mood,
    case_one, person_one, gender_one, number_one );
# (no) haberse salido
{ voice :: ACT_IMP }, { case_one :: VERB },
{ person_one :: THIRD }, { number_one :: SING },
[ "no" ],
Compound Verb_clitized ( mvtype, voice, compound, mood,
    case_one, person_one, gender_one, number_one );
# (no) haberse tratado
{ voice :: MIDD_IMP }, { person_one :: THIRD },
{ number_one :: SING },
[ "no" ],
Compound Verb_clitized ( mvtype, voice, compound, mood,
    case_one, person_one, gender_one, number_one ).

```

The first alternative identifies active personal configurations (**ACT**), and active impersonal main groups for periphrastic VPs (**ACT_IMP**), based on, respectively, active personal (**ACT**) and impersonal (**ACT_IMP**) schemes. The second alternative identifies middle personal configurations (**MIDD**), based on middle personal (**MIDD**) schemes. The third alternative identifies active impersonal main groups with subject concord for periphrastic VPs (**ACT_IMP+PASS**), based on **-DO-** active personal (**ACT**) schemes. A guard takes care of the fact that third person clitic pronouns alone are allowed in this context. The fourth alternative identifies active impersonal main groups for periphrastic VPs (**ACT_IMP**), based on active personal (**ACT**) schemes. As in the previous alternative, a guard takes care of the fact that third person clitic pronouns alone are allowed in this context. Although both the first and the fourth alternatives produce impersonal active voice (**ACT_IMP**) **Non Finite Verb Phrase/7** non-terminals, this ambiguity is solved at the level of periphrastic VPs, see Sections 2.2.3.7.1 and 2.2.3.7.2, by means of case of the clitic pronoun: **VERB** for impersonal *se* VPs, and values in the domain of *function_case* for impersonal VPs based on inherently impersonal active voice verb schemes. The fifth alternative identifies middle impersonal main groups for periphrastic VPs (**MIDD_IMP**), based on middle impersonal (**MIDD_IMP**) schemes. The remaining sixth to tenth alternatives correspond to the previous ones, but they account for compound tenses.

The rewrite rule for passive voice is the following:

```

Non Finite Verb Phrase ( one_clitic, PASS, mood, case_one,
    person_one, gender_one, number_one ):
# (no) serle presentado, (no) haberle sido presentado
{ one_clitic :: SD0IO }, { case_one :: DAT },
[ "no" ],
( ( { mood :: INFINITIVE },
    Verb ( PARTICIPLE, PASS, simple, mood );
    Verb_clitized ( PARTICIPLE, PASS, simple, mood, ONE ) ),
    Clitic Pronoun ( person_one, gender_one, number_one,
        case_one, ONE_OF_ONE );
    Compound Verb_clitized ( PARTICIPLE, PASS, compound, mood,
        case_one, person_one, gender_one, number_one ) ),
    Participle Verb ( one_clitic, PASS, PARTICIPLE, gender,
        number ) ).

```

This rule identifies passive voice configurations (**PASS**), based on **-DO-** active personal (**ACT**) schemes.

2.2.3.5.6 Non Finite Verb Phrase/11 Non Finite Verb Phrase non-terminals with two clitic pronouns account for active and middle configurations, **ACT** and **MIDD** configurations. All alternatives include the predicate **two person values** to check the compatibility of the affix *person* of the clitic pronouns. For active voice, both clitic pronouns are functional. In active voice, the predicate **verb two ClPrs** regulates the combination between the values for affixes *mvtype* and *case of*, respectively, the verb form and the second functional clitic pronoun. For impersonal or middle voices, one clitic pronoun is non-functional and the other is functional. In middle voice, the predicates **verb one ClPr** and **impersonal verb one ClPr** regulate the combination between the values for affixes *mvtype* and *case of*, respectively, the verb form and the functional clitic pronoun.

The rewrite rule for active, personal or impersonal, and for middle voices is the following:

```

Non Finite Verb Phrase ( mvtype, voice, mood, case_one,
    person_one, gender_one, number_one, case_two, person_two,
    gender_two, number_two ):
# (no) entregárselo
  { voice :: ACT },
  [ "no" ],
  Verb_clitized ( mvtype, voice, simple, mood, ONE|TWO ),
  Clitic Pronoun ( person_one, gender_one, number_one, case_one,
    ONE_OF_TWO ),
  Clitic Pronoun ( person_two, gender_two, number_two, case_two,
    TWO_OF_TWO ),
  two persons values ( mood, person_one, person_two ),
  verb two ClPrs ( mvtype, case_two );
# (no) creérmelo
  { voice :: MIDD },
  [ "no" ],
  Verb_clitized ( mvtype, voice, simple, mood, ONE|TWO ),
  Clitic Pronoun ( person_one, gender_one, number_one,
    case_one ),
  Clitic Pronoun ( person_two, gender_two, number_two, case_two,
    ONE_OF_ONE ),
  two persons values ( mood, person_one, person_two),
  verb one ClPr ( mvtype, case_two );
# (no) entregársele
  { mvtype :: SDOIO }, { voice :: ACT_IMP+PASS },
  { person_one :: THIRD }, { case_two :: DAT },
  [ "no" ],
  Verb_clitized ( mvtype, ACT, simple, mood, ONE|TWO ),
  Clitic Pronoun ( person_one, gender_one, number_one,
    case_one ),

```

```

        Clitic Pronoun ( person_two, gender_two, number_two, case_two,
            ONE_OF_ONE ),
        two persons values ( mood, person_one, person_two);
# (no) querérselo
{ voice :: ACT_IMP }, { person_one :: THIRD },
  { number_one :: SING },
  [ "no" ],
  Verb_clitized ( mvtype, ACT, simple, mood, ONE|TWO ),
  Clitic Pronoun ( person_one, gender_one, number_one,
    case_one ),
  Clitic Pronoun ( person_two, gender_two, number_two, case_two,
    ONE_OF_ONE ),
  two persons values ( mood, person_one, person_two),
  impersonal verb one ClPr ( mvtype, case_two );
# (no) habérselo entregado
{ voice :: ACT },
  [ "no" ],
  Compound Verb_clitized ( mvtype, voice, compound, mood,
    case_one, person_one, gender_one, number_one, case_two,
    person_two, gender_two, number_two ),
  two persons values ( mood, person_one, person_two ),
  verb two ClPrs ( mvtype, case_two );
# (no) habérmelo creído
{ voice :: MIDD },
  [ "no" ],
  Compound Verb_clitized ( mvtype, voice, compound, mood,
    case_one, person_one, gender_one, number_one, case_two,
    person_two, gender_two, number_two ),
  two persons values ( mood, person_one, person_two ),
  verb one ClPr ( mvtype, case_two );
# (no) habersele entregado
{ mvtype :: SDOIO }, { voice :: ACT_IMP+PASS },
  { person_one :: THIRD }, { case_two :: DAT },
  [ "no" ],
  Compound Verb_clitized ( mvtype, voice, compound, mood,
    case_one, person_one, gender_one, number_one, case_two,
    person_two, gender_two, number_two ),
  two persons values ( mood, person_one, person_two );
# (no) habérselo querido
{ voice :: ACT_IMP }, { person_one :: THIRD },
  { number_one :: SING },
  [ "no" ],
  Compound Verb_clitized ( mvtype, voice, compound, mood,
    case_one, person_one, gender_one, number_one, case_two,
    person_two, gender_two, number_two ),
  two persons values ( mood, person_one, person_two ),
  impersonal verb one ClPr ( mvtype, case_two ).

```

Except for the absence of the alternatives that account for **ACT_IMP** and

MIDD_IMP main groups based on, respectively, **ACT_IMP** and **MIDD_IMP** schemes, this set of alternatives identifies the same set of configurations, based on the same set of schemes, as the alternatives that constituted the rewrite rule of **Non Finite Verb Phrase/7** with one clitic pronoun.

2.2.3.5.7 Participle Verb Phrase/5 Non-terminal **Participle Verb Phrase** accounts for active, middle and passive configurations, **ACT**, **MIDD** and **PASS** configurations.

```
Participle Verb Phrase ( mvtype, voice, mood, gender, number ):
# llegado, vestido, vendido
  { voice :: ACT|MIDD|PASS },
  Participle Verb ( mvtype, voice, mood, gender, number ).
```

2.2.3.6 Periphrastic Groups: Periphrasis

First in this Section, we deal with **Periphrasis** non-terminals that constitute periphrastic groups that do not include clitic pronouns selected by the main group of the periphrastic VP, that is, periphrastic groups that have proper voice values. Then we deal with **Periphrasis** non-terminals that include clitic pronouns selected by the main group of the periphrastic VP, that is periphrastic groups that have instead voice values in the domain of *improper_voice*. In both cases, we will proceed from the periphrastic groups that do not include functional clitic pronouns to those that include one and two of them. Within each of the resulting groups of rules, we will separate rewrite rules that describe **Periphrasis** non-terminals with different voice values, both for finite and for non-finite verb forms.

2.2.3.6.1 Proper Voice Periphrastic Groups **Periphrasis** non-terminals that do not include clitic pronouns selected by the main groups of the target periphrastic VPs have **ACT**, **ACT_IMP**, **MIDD** or **MIDD_IMP** voice values.

2.2.3.6.1.1 ACT voice Periphrasis/3,5 Non-terminals **Periphrasis/3,5** with active voice (**ACT**) are constituted by one active voice (**ACT**) semiauxiliary verb form, or by a series of them.

```
Periphrasis ( semiaux, ACT, mood ):
# (no) seguir
  [ "no" ],
```

```

    ( Verb ( semiaux, ACT, tense, mood );
      Verb ( non_finite_mood, ACT, tense, mood ),
        Periphrasis ( semiaux, ACT, non_finite_mood ) );
# (no) dejar de
[ "no" ],
  ( Verb ( semiaux, ACT, linktype, tense, mood ),
    Periphrastic Connector ( linktype );
    Verb ( non_finite_mood, ACT, linktype, tense, mood ),
      Periphrastic Connector ( linktype ),
        Periphrasis ( semiaux, ACT, non_finite_mood ) ).

```

Both alternatives identify active voice (**ACT**) **Periphrasis** non-terminals, constituted exclusively by active voice (**ACT**) semiauxiliary verb forms. While the first alternative accounts for periphrastic groups whose first auxiliary verb form does not require an intermediate connector to introduce the following verb form, the second instead accounts for periphrastic groups whose first auxiliary verb form does require such an intermediate connector. All the alternatives that account for periphrastic groups non-terminals that contain active voice semiauxiliaries are duplicated in this way. In the rest of this exposition, unless there are some additional differences, we will only include the alternatives that account for semiauxiliaries not requiring intermediate connectors.

In addition, we can observe too that in both alternatives the group that constitutes their second member (after the optional negative particle) is there to account for the fact that either just one verb form or a series of them might constitute the periphrastic group. Whereas, in the first case, the value for attribute *vtype* (restricted to the domain of *semiaux*) of the target **Periphrasis** non-terminal will be that of the verb form that constitutes alone the periphrastic group, in the second case, this value will be that of the last semiauxiliary verb form found in the series of verb forms that constitutes together the periphrastic group. In this second case, the value in question corresponds to that of the **Periphrasis** non-terminal found in the last position of the second alternative of the group. This, in fact, is a recursive call of non-terminal **Periphrasis**, and thus enables the identification of a series of active voice semiauxiliaries, the last of which provides its value for *vtype* to the **Periphrasis** non-terminal in the group. This structure is recurrent in almost all the alternatives that constitute **Periphrasis** non-terminals. In the rest of this exposition, we follow the convention of not fully developing groups of alternatives of this type. Instead, we will substitute the second member of the

group by the (non-AGFL) notation <-, **Periphrasis**, meaning that this second member can be modeled from the first member by means of the following simple rules: i) substitution of the affix non-terminal *semiaux* of **Verb** (or **Verb_clitized**) for *non_finite_mood* and ii) addition of the **Periphrasis** non-terminal, with parameters (*semiaux*, **ACT**, *non_finite_mood*).

```
Periphrasis ( semiaux, ACT, person, number, mood ):
# (no) sigo, (no) siga, sigue
  [ { mood :: INDICATIVE|SUBJUNCTIVE },
    "no" ],
  ( Verb ( semiaux, ACT, person, number, tense, mood );
    <-, Periphrasis );
# no sigas
(...).
```

The first alternative identifies indicative, subjunctive and affirmative imperative periphrastic groups, the second identifies negative imperative periphrastic groups.

2.2.3.6.1.2 ACT_IMP voice Periphrasis/3,5 Non-terminals **Periphrasis/3,5** with impersonal active voice (**ACT_IMP**) are constituted by i) one impersonal active voice (**ACT_IMP**) semiauxiliary verb form, or ii) a series of active voice (**ACT**) semiauxiliary verb forms with in between one impersonal active voice (**ACT_IMP**) semiauxiliary verb form.

```
Periphrasis ( semiaux, ACT_IMP, mood ):
# (no) haber que
  [ "no" ],
  ( Verb ( semiaux, ACT_IMP, linktype, tense, mood ),
    Periphrastic Connector ( linktype );
    <-, Periphrasis );
# (no) seguir habiendo que
  [ "no" ],
  Verb ( non_finite_mood, ACT, tense, mood ),
  Periphrasis ( semiaux, ACT_IMP, non_finite_mood ).
```

While the first alternative accounts for **Periphrasis** structures with an initial impersonal active voice (**ACT_IMP**) semiauxiliary verb form, the second accounts for **Periphrasis** structures with an initial active voice (**ACT**) semiauxiliary verb form.

As a logical consequence of the fact that impersonal non-finite verb phrases do not constitute configurations by themselves (see Section 2.2.3.5.4), neither do

impersonal non-finite **Periphrasis** non-terminals constitute periphrastic groups on their own. In spite of this, impersonal non-finite **Periphrasis** non-terminals must be covered by the grammar, since they might appear as the final component of an impersonal active voice (**ACT_IMP**) periphrastic group with an initial finite active voice (**ACT**) verb form (puede haber que *ayudarlo*, see below in this Section).

```
Periphrasis ( semiaux, ACT_IMP, person, number, mood ):
# (no) hay que, (no) haya que
  [ "no" ],
  ( Verb ( semiaux, ACT_IMP, linktype, person, number, tense,
          mood ),
    Periphrastic Connector ( linktype );
    <-, Periphrasis );
# sigue habiendo que, (no) siga habiendo que
  { person :: THIRD }, { number :: SING },
  { mood :: INDICATIVE|SUBJUNCTIVE },
  [ "no" ],
  Verb ( non_finite_mood, ACT, person, number, tense, mood ),
  Periphrasis ( semiaux, ACT_IMP, non_finite_mood ).
```

The first alternative accounts for **Periphrasis** non-terminals with an initial impersonal active voice (**ACT_IMP**) semiauxiliary verb form. The second alternative accounts for **Periphrasis** non-terminals with an initial active voice (**ACT**) semiauxiliary verb form.

2.2.3.6.1.3 MIDD voice Periphrasis/3,7,5 Non-terminals **Periphrasis/3,7,5** with middle voice (**MIDD**) are constituted by a non-functional clitic pronoun and i) one middle voice (**MIDD**) semiauxiliary verb form, or ii) a series of active voice (**ACT**) semiauxiliary verb forms followed by one middle voice (**MIDD**) semiauxiliary verb form.

```
Periphrasis ( semiaux, MIDD, mood ):
# (no) poner a
  [ "no" ],
  Verb ( semiaux, MIDD, linktype, tense, mood ),
  Periphrastic Connector ( linktype );
# (no) seguir poniendo a
  [ "no" ],
  Verb ( non_finite_mood, ACT, tense, mood ),
  Periphrasis ( semiaux, MIDD, non_finite_mood ).
```

The first alternative accounts for **Periphrasis** non-terminals constituted by one middle voice (**MIDD**) semiauxiliary verb form. The second accounts for **Periphrasis** non-terminals constituted by one middle voice (**MIDD**) semiauxiliary verb form preceded by one or more active voice (**ACT**) semiauxiliaries.

Neither of the alternatives contains in this rule the obligatory clitic pronoun required by the middle voice semiauxiliary, which means that non-finite **Periphrasis/3** non-terminals cannot constitute periphrastic groups on their own. In spite of this, they must be accounted for in the grammar because they can constitute the final part of a middle voice periphrastic group with an initial active voice semiauxiliary. In this case, the pronoun appears in the periphrastic group before the series of verb forms identified by the non-terminal **Periphrasis/3**, attached to a previous semiauxiliary and making possible the analysis of sequences as, for example, *te puedes poner a, puedes seguirte poniendo a* (see below **Periphrasis/7,5**).

```
Periphrasis ( semiaux, MIDD, mood, case_one, person_one, gender_one,
              number_one ):
# (no) ponerse a
  [ "no" ],
  ( { mood :: INFINITIVE },
    Verb ( semiaux, MIDD, linktype, simple, mood );
    Verb_clitized ( semiaux, MIDD, linktype, simple, mood,
                    ONE ) ),
    Clitic Pronoun ( person_one, gender_one, number_one,
                    case_one ),
    Periphrastic Connector ( linktype );
# (no) haberse puesto a
  [ "no" ],
  Compound Verb_clitized ( semiaux, MIDD, linktype, compound,
                           mood, case_one, person_one, gender_one, number_one ),
  Periphrastic Connector ( linktype );
# (no) seguir poniéndose a
  [ "no" ],
  Verb ( non_finite_mood, ACT, tense, mood ),
  Periphrasis ( semiaux, MIDD, non_finite_mood, case_one,
               person_one, gender_one, number_one );
# (no) seguirse poniendo a
  [ "no" ],
  ( { mood :: INFINITIVE },
    Verb ( non_finite_mood, ACT, simple, mood );
    Verb_clitized ( non_finite_mood, ip_MIDD, simple, mood,
                    ONE ) ),
    Clitic Pronoun ( person_one, gender_one, number_one,
                    case_one ),
```

```

    Periphrasis ( semiaux, MIDD, non_finite_mood );
# (no) haberse seguido poniendo a
  [ "no" ],
  Compound Verb_clitized ( non_finite_mood, ip_MIDD, compound,
    mood, case_one, person_one, gender_one, number_one ),
  Periphrasis ( semiaux, MIDD, non_finite_mood ).

```

The first and the second alternatives account for **Periphrasis** non-terminals constituted by one middle voice (**MIDD**) semiauxiliary verb form. While the third alternative accounts for **Periphrasis** non-terminals constituted by an initial active voice (**ACT**) semiauxiliary verb form to which the non-functional clitic pronoun has not been attached, the fourth and the fifth alternatives account for **Periphrasis** non-terminals constituted by an initial active voice (**ACT**) semiauxiliary verb form to which the non-functional clitic pronoun has been attached.

```

Periphrasis ( semiaux, MIDD, person, number, mood ):
# (no) me pongo a, (no) me ponga a
  { mood :: INDICATIVE|SUBJUNCTIVE },
  [ "no" ],
  Clitic Pronoun ( person, gender, number, case ),
  Verb ( semiaux, MIDD, linktype, person, number, tense, mood ),
  Periphrastic Connector ( linktype );
# ponte a
(...);
# no te pongas a
(...);
# (no) me siga poniendo a, (no) me siga poniendo a
  { mood :: INDICATIVE|SUBJUNCTIVE },
  [ "no" ],
  Clitic Pronoun ( person, gender, number, case ),
  Verb ( non_finite_mood, ACT, person, number, tense, mood ),
  Periphrasis ( semiaux, MIDD, non_finite_mood );
# síguete poniendo a
(...);
# no te sigas poniendo a
(...);
# (no) siga poniéndome a, (no) siga poniéndome a, sigue poniéndote a
  [ { mood :: INDICATIVE|SUBJUNCTIVE },
    "no" ],
  Verb ( non_finite_mood, ACT, person, number, tense, mood ),
  Periphrasis ( semiaux, MIDD, non_finite_mood, case, person,
    gender, number );
# no sigas poniéndote a
(...).

```

The first, the second and the third alternatives account for **Periphrasis** non-terminals constituted by one middle voice (**MIDD**) semiauxiliary verb form.

While the fourth, the fifth and the sixth alternatives account for **Periphrasis** non-terminals in which the non-functional clitic pronoun has been attached to an initial finite active voice (**ACT**) semiauxiliary, the seventh and the eighth alternatives account for **Periphrasis** non-terminals in which the non-functional clitic pronoun has been attached to a middle voice (**MIDD**) semiauxiliary verb form preceded by one or more active voice (**ACT**) semiauxiliary verb forms.

2.2.3.6.1.4 MIDD_IMP voice Periphrasis/3,5 Non-terminals **Periphrasis/3,5** with impersonal middle voice (**MIDD_IMP**) are constituted by one non-functional clitic pronoun, one impersonal active voice (**ACT_IMP**) semiauxiliary verb form in infinitive or gerund and one final middle voice (**MIDD**) semiauxiliary verb form. Active voice (**ACT**) semiauxiliary verb forms can be introduced before the impersonal active voice (**ACT_IMP**) semiauxiliary verb form and between this and the middle voice (**MIDD**) semiauxiliary verb form. The clitic pronoun can be attached to any verb form placed after the impersonal active voice (**ACT_IMP**) semiauxiliary verb form.

```
Periphrasis ( semiaux, MIDD_IMP, mood ):
# (no) haber que ponerse a
  [ "no" ],
  Verb ( non_finite_mood, ACT_IMP, linktype, tense, mood ),
  Periphrastic Connector ( linktype ),
  Periphrasis ( semiaux, MIDD, non_finite_mood, VERB, THIRD,
    gender, SING );
# (no) seguir habiendo que ponerse a
  [ "no" ],
  Verb ( non_finite_mood, ACT, tense, mood ),
  Periphrasis ( semiaux, MIDD_IMP, non_finite_mood ).
```

While the first alternative accounts for **Periphrasis** non-terminals with an initial impersonal active voice (**ACT_IMP**) semiauxiliary verb forms, the second accounts for **Periphrasis** non-terminals with an initial active voice (**ACT**) semiauxiliary verb forms.

```
Periphrasis ( semiaux, MIDD_IMP, person, number, mood ):
# (no) hay que ponerse a, (no) haya que ponerse a
  [ "no" ],
  Verb ( non_finite_mood, ACT_IMP, linktype, person, number,
    tense, mood ),
  Periphrastic Connector ( linktype ),
  Periphrasis ( semiaux, MIDD, non_finite_mood, VERB, person,
```

```

        gender, number );
# sigue habiendo que ponerse a, (no) siga habiendo que ponerse a
  { person :: THIRD }, { number :: SING },
  { mood :: INDICATIVE|SUBJUNCTIVE },
  [ "no" ],
  Verb ( non_finite_mood, ACT, person, number, tense, mood ),
  Periphrasis ( semiaux, MIDD_IMP, non_finite_mood ).

```

As before, while the first alternative accounts for **Periphrasis** non-terminals with an initial impersonal active voice (**ACT_IMP**) semiauxiliary verb forms, the second accounts for **Periphrasis** non-terminals with an initial active voice (**ACT**) semiauxiliary verb forms.

2.2.3.6.2 Improper_voice Voice Periphrastic Groups with only Non-functional Clitic Pronouns **Periphrasis** non-terminals that include only non-functional clitic pronouns selected by the main group of the target periphrastic VPs have improper middle (**ip_MIDD**) or improper impersonal middle (**ip_MIDD_-IMP**) voices.

2.2.3.6.2.1 ip_MIDD voice Periphrasis/7,5 Non-terminals **Periphrasis/7,5** with improper middle voice (**ip_MIDD**) are constituted by one non-functional clitic pronoun and one or more active voice (**ACT**) semiauxiliary verb forms. The clitic pronoun can be attached to any verb form in the periphrastic group.

```

Periphrasis ( semiaux, ip_MIDD, mood, case, person, gender, number ):
# (no) seguirme
  [ "no" ],
  ( ( { mood :: INFINITIVE },
      Verb ( semiaux, ACT, simple, mood );
      Verb_clitized ( semiaux, ip_MIDD, simple, mood, ONE ) ),
      Clitic Pronoun ( person, gender, number, case );
      <- , Periphrasis );
# (no) haberme seguido
  [ "no" ],
  ( Compound Verb_clitized ( semiaux, ip_MIDD, compound, mood,
      case, person, gender, number );
      <- , Periphrasis );
# (no) seguir dejándome de
  [ "no" ],
  Verb ( non_finite_mood, ACT, tense, mood ),
  Periphrasis ( semiaux, ip_MIDD, non_finite_mood, case,
      person, gender, number ).

```

While the first and the second alternatives account for **Periphrasis** non-terminals in which the clitic pronoun is attached to the first semiauxiliary verb form, the third alternative accounts for **Periphrasis** non-terminals in which the clitic pronoun is attached to one of the semiauxiliaries that follow the first one of a series of them.

```
Periphrasis ( semiaux, ip_MIDD, person, number, mood ):
# (no) me sigo, (no) me siga
  { mood :: INDICATIVE|SUBJUNCTIVE },
  [ "no" ],
  Clitic Pronoun ( person, gender, number, case ),
  ( Verb ( semiaux, ACT, person, number, tense, mood );
    <- , Periphrasis );
# siguete
(...);
# no te sigas
(...);
# (no) sigo dejándome de, (no) siga dejándome de
  # sigue dejándote de
  [ { mood :: INDICATIVE|SUBJUNCTIVE },
    "no" ],
    Verb ( non_finite_mood, ACT, person, number, tense, mood ),
    Periphrasis ( semiaux, ip_MIDD, non_finite_mood, case,
      person, gender, number );
# no sigas dejándote de
(...).
```

While the first, the second and the third alternatives account for **Periphrasis** non-terminals in which the clitic pronoun is attached to the finite semiauxiliary verb form, the fourth and the fifth alternatives account for **Periphrasis** non-terminals in which the clitic pronoun is attached to one of the non-finite verb forms that follow the first finite semiauxiliary verb form.

2.2.3.6.2.2 ip_MIDD_IMP voice Periphrasis/7,5 Non-terminals **Periphrasis/7,5** with improper impersonal middle voice (**ip_MIDD_IMP**) are constituted by one non-functional clitic pronoun, one impersonal active voice (**ACT_-IMP**) semiauxiliary verb form and one or more active voice (**ACT**) semiauxiliary verb forms. The clitic pronoun can be attached to any active voice (**ACT**) verb form following the impersonal active voice (**ACT_IMP**) semiauxiliary verb form in the periphrastic group.

```
Periphrasis ( semiaux, ip_MIDD_IMP, mood, case, person, gender,
```

```

    number ):
# (no) haber que dejarse de
  { person :: THIRD }, { number :: SING },
  [ "no" ],
  Verb ( non_finite_mood, ACT_IMP, linktype, tense, mood ),
  Periphrastic Connector ( linktype ),
  Periphrasis ( semiaux, ip_MIDD, non_finite_mood, case,
    person, gender, number );
# (no) seguir habiendo que dejarse de
  [ "no" ],
  Verb ( non_finite_mood, ACT, tense, mood ),
  Periphrasis ( semiaux, ip_MIDD_IMP, non_finite_mood, case,
    person, gender, number ).

```

While the first alternative accounts for **Periphrasis** non-terminals in which the impersonal active voice (**ACT_IMP**) semiauxiliary verb form appears in the first position, the second accounts for **Periphrasis** non-terminals in which the impersonal active voice (**ACT_IMP**) semiauxiliary is instead preceded by one or more active voice (**ACT**) semiauxiliaries.

```

Periphrasis ( semiaux, ip_MIDD_IMP, person, number, mood ):
# (no) hay que dejarse de, (no) haya que dejarse de
  [ "no" ],
  Verb ( non_finite_mood, ACT_IMP, linktype, person, number,
    tense, mood ),
  Periphrastic Connector ( linktype ),
  Periphrasis ( semiaux, ip_MIDD, non_finite_mood, case,
    person, gender, number );
# sigue habiendo que dejarse de, (no) siga habiendo que dejarse de
  { mood :: INDICATIVE|SUBJUNCTIVE },
  [ "no" ],
  Verb ( non_finite_mood, ACT, person, number, tense, mood ),
  Periphrasis ( semiaux, ip_MIDD_IMP, non_finite_mood, case,
    person, gender, number ).

```

While the first alternative accounts for **Periphrasis** non-terminals with an initial impersonal active voice (**ACT_IMP**) semiauxiliary verb form, the second alternative accounts for **Periphrasis** non-terminals with an initial active voice (**ACT**) semiauxiliary verb form.

2.2.3.6.3 Improper_voice Voice Periphrastic Groups with One Functional Clitic Pronoun **Periphrasis** non-terminals that include one functional clitic pronoun selected by the main group of the target periphrastic VPs have **ip_ACT**, **ip_ACT_IMP**, **ip_MIDD** or **ip_MIDD_IMP** voices.

2.2.3.6.3.1 ip_ACT voice Periphrasis/7,9 Non-terminals **Periphrasis-7,9** with improper active voice (**ip_ACT**) are constituted by one functional clitic pronoun and one or more active voice (**ACT**) semiauxiliary verb forms. The clitic pronoun can be attached to any verb form in the periphrastic group.

```
Periphrasis ( semiaux, ip_ACT, mood, case_one, person_one,
              gender_one, number_one ):
# (no) seguirlo
  [ "no" ],
  ( ( { mood :: INFINITIVE },
      Verb ( semiaux, ACT, simple, mood );
      Verb_clitized ( semiaux, ip_ACT, simple, mood, ONE ) ),
      Clitic Pronoun ( person_one, gender_one, number_one,
                      case_one, ONE_OF_ONE );
    <-, Periphrasis );
# (no) haberlo seguido
  [ "no" ],
  ( Compound Verb_clitized ( semiaux, ip_ACT, compound, mood,
                             case_one, person_one, gender_one, number_one );
    <-, Periphrasis );
# (no) seguir dejándolo de
  [ "no" ],
  Verb ( non_finite_mood, ACT, tense, mood ),
  Periphrasis ( semiaux, ip_ACT, non_finite_mood, case_one,
                person_one, gender_one, number_one ).
```

While the first and the second alternatives account for **Periphrasis** non-terminals in which there is only one semiauxiliary or in which the clitic pronoun is attached to the first semiauxiliary of a series of them, the third alternative accounts for **Periphrasis** non-terminals in which the clitic pronoun is attached to one of the semiauxiliary verb forms placed after the first one of a series of them.

```
Periphrasis ( semiaux, ip_ACT, person, number, mood, case_one,
              person_one, gender_one, number_one ):
# (no) lo sigo, (no) lo siga
  { mood :: INDICATIVE|SUBJUNCTIVE },
  [ "no" ],
  Clitic Pronoun ( person_one, gender_one, number_one, case_one,
                  ONE_OF_ONE ),
  ( Verb ( semiaux, ACT, person, number, tense, mood );
    <-, Periphrasis ),
  two persons values ( mood, person, person_one );
# siguelo
  (...);
# no lo sigas
  (...);
```

```

# (no) siga dejándolo de, (no) siga dejándolo de, sigue dejándolo de
  [ { mood :: INDICATIVE|SUBJUNCTIVE },
    "no" ],
    Verb ( non_finite_mood, ACT, person, number, tense, mood ),
    Periphrasis ( semiaux, ip_ACT, non_finite_mood, case_one,
      person_one, gender_one, number_one ),
    two persons values ( mood, person, person_one );
# no sigas dejándolo de
(...).

```

All the alternatives of this rewrite rule include the predicate **two person values** to regulate the compatibility of the values for affix *person* of the functional clitic pronoun and the finite verb form. The first, the second and the third alternatives account for **Periphrasis** non-terminals in which the clitic pronoun is attached to the finite semiauxiliary verb form. The fourth and the fifth alternatives, on the contrary, account for **Periphrasis** non-terminals in which the clitic pronoun is attached to one of the non-finite verb forms that follow the first finite semiauxiliary verb form.

2.2.3.6.3.2 ip_ACT_IMP voice Periphrasis/7,9 Non-terminals **Periphrasis/7,9** with improper impersonal active voice (**ip_ACT_IMP**) are constituted by one functional clitic pronoun and one impersonal active voice (**ACT_-IMP**) semiauxiliary verb form. Active voice (**ACT**) semiauxiliaries can precede or follow the impersonal active voice (**ACT_IMP**) semiauxiliary verb form. The clitic pronoun can be attached to any verb form in the periphrastic group.

```

Periphrasis ( semiaux, ip_ACT_IMP, mood, case_one, person_one,
  gender_one, number_one ):
# (no) haber que dejarlo de
  [ "no" ],
    Verb ( non_finite_mood, ACT_IMP, linktype, tense, mood ),
    Periphrastic Connector ( linktype ),
    Periphrasis ( semiaux, ip_ACT, non_finite_mood, case_one,
      person_one, gender_one, number_one );
# (no) haberlo que
  [ "no" ],
    ( ( { mood :: INFINITIVE },
      Verb ( semiaux, ACT_IMP, linktype, simple, mood );
      Verb_clitized ( semiaux, ip_ACT_IMP, linktype, simple,
        mood, ONE ) ),
      Clitic Pronoun ( person_one, gender_one, number_one,
        case_one, ONE_OF_ONE ),
      Periphrastic Connector ( linktype );

```

```

        <-, Periphrasis );
# (no) haberlo habido que
  [ "no" ],
  ( Compound Verb_clitized ( semiaux, ip_ACT_IMP, linktype,
    compound, mood, case_one, person_one, gender_one,
    number_one ),
    Periphrastic Connector ( linktype );
  <-, Periphrasis );
# (no) seguir habiéndolo que
  [ "no" ],
  Verb ( non_finite_mood, ACT, tense, mood ),
  Periphrasis ( semiaux, ip_ACT_IMP, non_finite_mood,
    case_one, person_one, gender_one, number_one );
# (no) seguirlo habiendo que
  [ "no" ],
  ( { mood :: INFINITIVE },
    Verb ( non_finite_mood, ACT, simple, mood );
    Verb_clitized ( non_finite_mood, ip_ACT, simple, mood,
      ONE ) ),
  Clitic Pronoun ( person_one, gender_one, number_one,
    case_one, ONE_OF_ONE ),
  Periphrasis ( semiaux, ACT_IMP, non_finite_mood );
# (no) haberlo seguido habiendo que
  [ "no" ],
  Compound Verb_clitized ( non_finite_mood, ip_ACT, compound,
    mood, case_one, person_one, gender_one, number_one ),
  Periphrasis ( semiaux, ACT_IMP, non_finite_mood ).

```

The first alternative accounts for **Periphrasis** non-terminals in which the clitic pronoun has been attached to a semiauxiliary verb form placed after the impersonal active voice (**ACT_IMP**) semiauxiliary verb form. The second, the third and the fourth alternatives account for **Periphrasis** non-terminals in which the clitic pronoun has been attached to the impersonal active voice (**ACT_IMP**) semiauxiliary verb form. In the second and the third alternatives, the impersonal active voice (**ACT_IMP**) semiauxiliary is in first position in the periphrastic group, in the fourth alternative instead it is preceded by one or more active voice (**ACT**) semiauxiliary verb forms. The fifth and the sixth alternatives account for **Periphrasis** non-terminals in which the clitic pronoun has been attached to the first active voice (**ACT**) semiauxiliary verb form of the series of them possibly preceding the impersonal active voice (**ACT_IMP**) semiauxiliary verb form.

```

Periphrasis ( semiaux, ip_ACT_IMP, person, number, mood, case_one,
  person_one, gender_one, number_one ):
# (no) hay que dejarlo de

```

```

[ "no" ],
  Verb ( non_finite_mood, ACT_IMP, linktype, person, number,
        tense, mood ),
  Periphrastic Connector ( linktype ),
  Periphrasis ( semiaux, ip_ACT, non_finite_mood, case_one,
               person_one, gender_one, number_one );
# (no) lo hay que
[ "no" ],
  Clitic Pronoun ( person_one, gender_one, number_one, case_one,
                  ONE_OF_ONE ),
  ( Verb ( semiaux, ACT_IMP, linktype, person, number, tense,
          mood ),
    Periphrastic Connector ( linktype );
  <- , Periphrasis );
# (no) sigue habiéndolo que, (no) siga habiéndolo que
{ person :: THIRD }, { number :: SING },
{ mood :: INDICATIVE|SUBJUNCTIVE },
[ "no" ],
  Verb ( non_finite_mood, ACT, person, number, tense, mood ),
  Periphrasis ( semiaux, ip_ACT_IMP, non_finite_mood,
               case_one, person_one, gender_one, number_one );
# (no) lo sigue habiendo que, (no) lo siga habiendo que
{ person :: THIRD }, { number :: SING },
{ mood :: INDICATIVE|SUBJUNCTIVE },
[ "no" ],
  Clitic Pronoun ( person_one, gender_one, number_one,
                  case_one, ONE_OF_ONE ),
  Verb ( non_finite_mood, ACT, person, number, tense, mood ),
  Periphrasis ( semiaux, ACT_IMP, non_finite_mood ).

```

The first alternative accounts for **Periphrasis** non-terminals in which the clitic pronoun has been attached to a semiauxiliary verb form placed after the impersonal active voice (**ACT_IMP**) semiauxiliary verb form. The second and the third alternatives account for **Periphrasis** non-terminals in which the clitic pronoun has been attached to the impersonal active voice (**ACT_IMP**) semiauxiliary verb form. In the second alternative, the impersonal active voice (**ACT_IMP**) semiauxiliary is in indicative or subjunctive mood, in the third alternative the impersonal active (**ACT_IMP**) voice semiauxiliary is in gerund of infinitive mood. The fifth alternative accounts for **Periphrasis** non-terminals in which the clitic pronoun has been attached to an active voice (**ACT**) semiauxiliary verb form placed in first position in the periphrastic group.

2.2.3.6.3.3 ip_MIDD voice Periphrasis/11,9 Non-terminals **Periphrasis/11,9** with improper middle voice (**ip_MIDD**) are constituted by one non-functional clitic pronoun, one functional clitic pronoun and one or more active voice (**ACT**) semiauxiliary verb forms. The clitic pronouns can be attached to any verb form in the periphrastic group.

```
Periphrasis ( semiaux, ip_MIDD, mood, case_one, person_one,
              gender_one, number_one, case_two, person_two, gender_two,
              number_two ):
# (no) seguírmelo
[ "no" ],
  ( Verb_clitized ( semiaux, ip_MIDD, simple, mood, ONE|TWO ),
    Clitic Pronoun ( person_one, gender_one, number_one,
                    case_one ),
    Clitic Pronoun ( person_two, gender_two, number_two,
                    case_two, ONE_OF_ONE );
  <-, Periphrasis ),
  two persons values ( mood, person_one, person_two);
# (no) habérmelo seguido
[ "no" ],
  ( Compound Verb_clitized ( semiaux, ip_MIDD, compound, mood,
                             case_one, person_one, gender_one, number_one, case_two,
                             person_two, gender_two, number_two );
  <-, Periphrasis ),
  two persons values ( mood, person_one, person_two);
# (no) seguir dejándomelo de
[ "no" ],
  Verb ( non_finite_mood, ACT, tense, mood ),
  Periphrasis ( semiaux, ip_MIDD, non_finite_mood, case_one,
               person_one, gender_one, number_one, case_two, person_two,
               gender_two, number_two ).
```

All alternatives of this rewrite rule include the predicate **two person values** to regulate the compatibility between the values for affix *person* of the non-functional clitic pronoun and the functional one. The first and the second alternatives account for **Periphrasis** non-terminals in which there is only one semiauxiliary verb form or in which the clitic pronouns are attached to the first semiauxiliary verb form of a series of them. The third alternative accounts for **Periphrasis** non-terminals in which the clitic pronouns are attached to one of the semiauxiliary verb forms placed after the first one of a series of them.

```
Periphrasis ( semiaux, ip_MIDD, person, number, mood, case_one,
              person_one, gender_one, number_one ):
# (no) me lo sigo, (no) me lo siga
```

```

{ mood :: INDICATIVE|SUBJUNCTIVE },
  [ "no" ],
  Clitic Pronoun ( person, gender, number, case ),
  Clitic Pronoun ( person_one, gender_one, number_one, case_one,
    ONE_OF_ONE ),
  ( Verb ( semiaux, ACT, person, number, tense, mood );
    <- , Periphrasis ),
  two persons values ( mood, person, person_one );
# siguetelo
(...);
# no te lo sigas
(...);
# (no) sigo dejándomelo de, (no) siga dejándotelo de
# sigue dejándotelo de
[ { mood :: INDICATIVE|SUBJUNCTIVE },
  "no" ],
  Verb ( non_finite_mood, ACT, person, number, tense, mood ),
  Periphrasis ( semiaux, ip_MIDD, non_finite_mood, case,
    person, gender, number, case_one, person_one, gender_one,
    number_one ),
  two persons values ( mood, person, person_one );
# no sigas dejándotelo de
(...).

```

All alternatives of this rewrite rule include the predicate **two person values** to regulate the compatibility between the values for affix *person* of the functional clitic pronoun and the finite verb form. The first, the second and the third alternatives account for **Periphrasis** non-terminals in which the clitic pronouns are attached to the finite semiauxiliary verb form. The fourth and the fifth alternatives, on the contrary, account for **Periphrasis** non-terminals in which the clitic pronouns are attached to one of the non-finite verb forms that follow the first finite semiauxiliary verb form.

2.2.3.6.3.4 ip_MIDD_IMP voice Periphrasis/11,9 Non-terminals **Periphrasis/11,9** with improper impersonal middle voice (**ip_MIDD_IMP**) are constituted by one non-functional clitic pronoun, one functional clitic pronoun and one impersonal active voice (**ACT_IMP**) semiauxiliary verb form. Active voice (**ACT**) semiauxiliary verb forms can precede or follow the impersonal active voice (**ACT_IMP**) semiauxiliary verb form. The clitic pronouns can be attached to any active voice (**ACT**) verb form following the impersonal active voice (**ACT_IMP**) semiauxiliary verb form in the periphrastic group.

```

Periphrasis ( semiaux, ip_MIDD_IMP, mood, case, person, gender,
              number, case_two, person_two, gender_two, number_two ):
# (no) haber que dejárselo de
  { person :: THIRD }, { number :: SING },
  [ "no" ],
  Verb ( non_finite_mood, ACT_IMP, linktype, tense, mood ),
  Periphrastic Connector ( linktype ),
  Periphrasis ( semiaux, ip_MIDD, non_finite_mood, case,
                person, gender, number, case_two, person_two,
                gender_two, number_two );
# (no) seguir habiendo que dejárselo de
  [ "no" ],
  Verb ( non_finite_mood, ACT, tense, mood ),
  Periphrasis ( semiaux, ip_MIDD_IMP, non_finite_mood, case,
                person, gender, number, case_two, person_two, gender_two,
                number_two ).

```

The first alternative accounts for **Periphrasis** non-terminals in which the impersonal active voice (**ACT_IMP**) semiauxiliary verb form is placed in first position in the periphrastic group. The second alternative accounts for **Periphrasis** non-terminals in which the impersonal active voice (**ACT_IMP**) semiauxiliary verb form is preceded by one or more active voice (**ACT**) semiauxiliary verb forms.

```

Periphrasis ( semiaux, ip_MIDD_IMP, person, number, mood, case_two,
              person_two, gender_two, number_two ):
# (no) hay que dejárselo de, (no) haya que dejárselo de
  [ { mood :: INDICATIVE|SUBJUNCTIVE },
    "no" ],
  Verb ( non_finite_mood, ACT_IMP, linktype, person, number,
         tense, mood ),
  Periphrastic Connector ( linktype ),
  Periphrasis ( semiaux, ip_MIDD, non_finite_mood, case,
                person, gender, number, case_two, person_two, gender_two,
                number_two );
# (no) sigue habiendo que dejárselo de
  # (no) siga habiendo que dejárselo de
  { mood :: INDICATIVE|SUBJUNCTIVE },
  [ "no" ],
  Verb ( non_finite_mood, ACT, person, number, tense, mood ),
  Periphrasis ( semiaux, ip_MIDD_IMP, non_finite_mood, case,
                person, gender, number, case_two, person_two, gender_two,
                number_two ).

```

The first alternative accounts for **Periphrasis** non-terminals with an initial impersonal active voice (**ACT_IMP**) semiauxiliary verb form. The second altern-

ative accounts for **Periphrasis** non-terminals with an initial active voice (**ACT**) semiauxiliary verb form.

2.2.3.6.4 Improper-voice Voice Periphrastic Groups with Two Functional Clitic Pronouns **Periphrasis** non-terminals that include two functional clitic pronouns selected by the main group of the target periphrastic VPs have **ip_ACT** or **ip_ACT_IMP** voices.

2.2.3.6.4.1 ip_ACT voice Periphrasis/11,13 Non-terminals **Periphrasis/11,13** with improper active voice (**ip_ACT**) are constituted by two functional clitic pronouns and one or more active voice (**ACT**) semiauxiliary verb forms. The clitic pronouns can be attached to any verb form in the periphrastic group.

```
Periphrasis ( semiaux, ip_ACT, mood, case_one, person_one,
              gender_one, number_one, case_two, person_two, gender_two,
              number_two ):
# (no) seguirselo
  [ "no" ],
  ( Verb_clitized ( semiaux, ip_ACT, simple, mood, ONE|TWO ),
    Clitic Pronoun ( person_one, gender_one, number_one,
                    case_one, ONE_OF_TWO ),
    Clitic Pronoun ( person_two, gender_two, number_two,
                    case_two, TWO_OF_TWO );
    <-, Periphrasis ) ),
  two persons values ( mood, person_one, person_two );
# (no) habérselo seguido
  [ "no" ],
  ( Compound Verb_clitized ( semiaux, ip_ACT, compound, mood,
                             case_one, person_one, gender_one, number_one, case_two,
                             person_two, gender_two, number_two );
    <-, Periphrasis ),
  two persons values ( mood, person_one, person_two );
# (no) seguir dejándoselo de
  [ "no" ],
  Verb ( non_finite_mood, ACT, tense, mood ),
  Periphrasis ( semiaux, ip_ACT, non_finite_mood, case_one,
               person_one, gender_one, number_one, case_two, person_two,
               gender_two, number_two ).
```

All alternatives of this rewrite rule include the predicate **two person values** to regulate the compatibility between the values for affix *person* of the two functional clitic pronouns. The first and second alternatives account for **Periphrasis** non-terminals in which there is only one semiauxiliary verb form or in which the clitic

pronouns are attached to the first semiauxiliary verb form of a series of them. The third alternative accounts for **Periphrasis** non-terminals in which the clitic pronouns are attached to any semiauxiliary verb form placed after the first one of a series of them.

```

Periphrasis ( semiaux, ip_ACT, person, number, mood, case_one,
              person_one, gender_one, number_one, case_two, person_two,
              gender_two, number_two ):
# (no) se lo sigo, (no) se lo siga
  { mood :: INDICATIVE|SUBJUNCTIVE },
  [ "no" ],
  Clitic Pronoun ( person_one, gender_one, number_one, case_one,
                  ONE_OF_TWO ),
  Clitic Pronoun ( person_two, gender_two, number_two, case_two,
                  TWO_OF_TWO ),
  ( Verb ( semiaux, ACT, person, number, tense, mood );
    <- , Periphrasis ),
  three persons values ( mood, person, person_one, person_two );
# sigueselo
(...);
# no se lo sigas
(...);
# (no) sigo dejándoselo de, (no) siga dejándoselo de
  # sigue dejándoselo de
  [ { mood :: INDICATIVE|SUBJUNCTIVE },
    "no" ],
  Verb ( non_finite_mood, ACT, person, number, tense, mood ),
  Periphrasis ( semiaux, ip_ACT, non_finite_mood, case_one,
               person_one, gender_one, number_one, case_two, person_two,
               gender_two, number_two ),
  three persons values ( mood, person, person_one, person_two );
# no sigas dejándoselo de
(...).

```

All alternatives of this rewrite rule include the predicate **three person values** to regulate the compatibility between the values for affix *person* of the two functional clitic pronouns and the finite verb form. The first, the second and the third alternatives account for **Periphrasis** non-terminals in which the clitic pronouns are attached to the finite semiauxiliary. The fourth and the fifth alternatives account for **Periphrasis** non-terminals in which the clitic pronouns are attached to one of the non-finite verb forms that follow the first finite semiauxiliary verb form.

2.2.3.6.4.2 ip_ACT_IMP voice Periphrasis/11,13 Non-terminals **Periphrasis/11,13** non-terminals with improper impersonal active voice (**ip_ACT_-**

IMP) are constituted by two functional clitic pronouns and one impersonal active voice (**ACT_IMP**) semiauxiliary verb form. Other active voice (**ACT**) semiauxiliary verb forms can precede or follow the impersonal active voice (**ACT_IMP**) semiauxiliary verb form. The clitic pronouns can be attached to any verb form in the periphrastic group.

```

Periphrasis ( semiaux, ip_ACT_IMP, mood, case_one, person_one,
              gender_one, number_one, case_two, person_two, gender_two,
              number_two ):
# (no) haber que dejárselo de
  [ "no" ],
  Verb ( non_finite_mood, ACT_IMP, linktype, tense, mood ),
  Periphrastic Connector ( linktype ),
  Periphrasis ( semiaux, ip_ACT, non_finite_mood, case_one,
                person_one, gender_one, number_one, case_two, person_two,
                gender_two, number_two );
# (no) habérselo que
  [ "no" ],
  ( Verb_clitized ( semiaux, ip_ACT_IMP, linktype, simple, mood,
                    ONE|TWO ),
    Clitic Pronoun ( person_one, gender_one, number_one,
                    case_one, ONE_OF_TWO ),
    Clitic Pronoun ( person_two, gender_two, number_two,
                    case_two, TWO_OF_TWO ),
    Periphrastic Connector ( linktype );
  <-, Periphrasis ),
  two persons values ( mood, person_one, person_two );
# (no) habérselo habido que
  [ "no" ],
  ( Compound Verb_clitized ( semiaux, ip_ACT_IMP, linktype,
                             compound, mood, case_one, person_one, gender_one,
                             number_one, case_two, person_two, gender_two,
                             number_two ),
    Periphrastic Connector ( linktype );
  <-, Periphrasis ),
  two persons values ( mood, person_one, person_two );
# (no) seguir habiéndoselo que
  [ "no" ],
  Verb ( non_finite_mood, ACT, tense, mood ),
  Periphrasis ( semiaux, ip_ACT_IMP, non_finite_mood, case_one,
                person_one, gender_one, number_one, case_two, person_two,
                gender_two, number_two );
# (no) seguirselo habiendo que
  [ "no" ],
  Verb_clitized ( non_finite_mood, ip_ACT, simple, mood,
                  ONE|TWO ),
  Clitic Pronoun ( person_one, gender_one, number_one, case_one,
                  ONE_OF_TWO ),

```

```

    Clitic Pronoun ( person_two, gender_two, number_two, case_two,
                    TWO_OF_TWO ),
    Periphrasis ( semiaux, ACT_IMP, non_finite_mood ),
    two persons values ( mood, person_one, person_two );
# (no) habérselo seguido habiendo que
[ "no" ],
    Compound Verb_clitized ( non_finite_mood, ip_ACT, compound,
                             mood, case_one, person_one, gender_one, number_one,
                             case_two, person_two, gender_two, number_two ),
    Periphrasis ( semiaux, ACT_IMP, non_finite_mood ),
    two persons values ( mood, person_one, person_two ).

```

The first alternative accounts for **Periphrasis** non-terminals in which the clitic pronouns have been attached to a semiauxiliary verb form placed after the impersonal active voice (**ACT_IMP**) semiauxiliary verb form. The second, the third and the fourth alternatives account for **Periphrasis** non-terminals in which the clitic pronouns have been attached to the impersonal active voice (**ACT_IMP**) semiauxiliary verb form. While in the second and the third alternatives, the impersonal active voice (**ACT_IMP**) semiauxiliary is in first position in the periphrastic group, in the fourth it is preceded by one or more active voice (**ACT**) semiauxiliaries. The fifth and the sixth alternatives account for **Periphrasis** non-terminals in which the clitic pronouns have been attached to the first active voice (**ACT**) semiauxiliary verb form of the series of them possibly preceding the impersonal active voice (**ACT_IMP**) semiauxiliary verb form.

```

Periphrasis ( semiaux, ip_ACT_IMP, person, number, mood, case_one,
             person_one, gender_one, number_one, case_two, person_two,
             gender_two, number_two ):
# (no) hay que dejárselo de, (no) haya que dejárselo de
[ "no" ],
    Verb ( non_finite_mood, ACT_IMP, linktype, person, number,
          tense, mood ),
    Periphrastic Connector ( linktype ),
    Periphrasis ( semiaux, ip_ACT, non_finite_mood, case_one,
                 person_one, gender_one, number_one, case_two, person_two,
                 gender_two, number_two );
# (no) se lo hay que
[ "no" ],
    Clitic Pronoun ( person_one, gender_one, number_one, case_one,
                    ONE_OF_TWO ),
    Clitic Pronoun ( person_two, gender_two, number_two, case_two,
                    TWO_OF_TWO ),
    ( Verb ( semiaux, ACT_IMP, linktype, person, number, tense,
            mood ),

```

```

        Periphrastic Connector ( linktype );
        <-, Periphrasis );
# (no) sigue habiéndoselo que, (no) siga habiéndoselo que
  { person :: THIRD }, { number :: SING },
  { mood :: INDICATIVE|SUBJUNCTIVE },
  [ "no" ],
  Verb ( non_finite_mood, ACT, person, number, tense, mood ),
  Periphrasis ( semiaux, ip_ACT_IMP, non_finite_mood, case_one,
    person_one, gender_one, number_one, case_two, person_two,
    gender_two, number_two );
# (no) se lo sigue habiendo que, (no) se lo sigue habiendo que
  { person :: THIRD }, { number :: SING },
  { mood :: INDICATIVE|SUBJUNCTIVE },
  [ "no" ],
  Clitic Pronoun ( person_one, gender_one, number_one, case_one,
    ONE_OF_TWO ),
  Clitic Pronoun ( person_two, gender_two, number_two, case_two,
    TWO_OF_TWO ),
  Verb ( non_finite_mood, ACT, person, number, tense, mood ),
  Periphrasis ( semiaux, ACT_IMP, non_finite_mood ).

```

The first alternative accounts for **Periphrasis** non-terminals in which the clitic pronouns have been attached to a semiauxiliary placed after the impersonal active voice (**ACT_IMP**) semiauxiliary verb form. The second and the third alternatives account for **Periphrasis** non-terminals in which the clitic pronouns have been attached to the impersonal active voice (**ACT_IMP**) semiauxiliary verb form. While in the second alternative, the impersonal active voice (**ACT_IMP**) semiauxiliary is in indicative or subjunctive, in the third one the impersonal active (**ACT_IMP**) voice semiauxiliary is in gerund or infinitive. The fifth alternative accounts for **Periphrasis** non-terminals in which the clitic pronouns have been attached to a finite active voice (**ACT**) semiauxiliary verb form in first position in the periphrastic group.

2.2.3.7 Periphrastic VPs: Periphrastic Verb Phrase and Periphrastic Non Finite Verb Phrase

This Section is organized the same as Section 2.2.3.5: first, **Periphrastic Verb Phrase** production is explained, proceeding from rules for **Periphrastic Verb Phrase** non-terminals that do not include functional clitic pronouns to those that include one and two of them²⁶. Then, production of **Periphrastic Non Finite**

²⁶Note that, in this Section, when talking about clitic pronouns and about their position in periphrastic VPs, we refer exclusively to clitic pronouns, functional or non-functional, selected

Verb Phrase non-terminal will be introduced, proceeding also from rules for **Periphrastic Non Finite Verb Phrase** non-terminals that do not include clitic pronouns to those that include one and two of them. Within each of these two group of rules, we will separate rules for the production of different *voice* values for **Periphrastic Verb Phrase** and **Periphrastic Non Finite Verb Phrase** non-terminals.

2.2.3.7.1 Periphrastic Verb Phrase/5 **Periphrastic Verb Phrase** non-terminals without functional clitic pronouns account for a number of personal or impersonal active, middle or passive voice configurations, the whole collection of possible **PREDICATE** *voice* values: **ACT**, **ACT_IMP**, **MIDD**, **MIDD_IMP**, **PASS**, **PASS_IMP**, **ACT_IMP+PASS**.

Without non-functional clitic pronouns, the non-terminal **Periphrastic Verb Phrase/5** accounts for personal or impersonal active voice configurations (**ACT**, **ACT_IMP**).

```
Periphrastic Verb Phrase ( mvtype, act, person, number, mood ):
# sigo saliendo, me pongo a hacer
  { act :: ACT },
  Periphrasis ( non_finite_mood, ACT|MIDD, person, number,
    mood ),
  Non Finite Verb Phrase ( mvtype, act, non_finite_mood );
# sigue lloviendo, se pone a llover
  { act :: ACT_IMP }, { person :: THIRD }, { number :: SING },
  { mood :: INDICATIVE|SUBJUNCTIVE },
  Periphrasis ( non_finite_mood, ACT|MIDD, person, number,
    mood ),
  Non Finite Verb Phrase ( mvtype, act, non_finite_mood );
# hay que salir, hay que ponerse a hacer
  { act :: ACT_IMP },
  Periphrasis ( non_finite_mood, ACT_IMP|MIDD_IMP, person,
    number, mood ),
  Non Finite Verb Phrase ( mvtype, ACT, non_finite_mood ).
```

As can be observed in this rewrite rule, the number and character of the alternatives of the rewrite rules of **Periphrastic Verb Phrase** non-terminals is determined by the number of possible combinations of voice values of their constituents **Periphrasis** and **Non Finite Verb Phrase**. In this case, for instance, while the

by the main verb contained in the main group of the periphrastic VP (clitic pronouns required by middle voice semiauxiliaries must be included in the periphrastic group and cannot be combined with any other clitic pronoun within it, see Section 2.1.3).

first alternative accounts for active voice (**ACT**) configurations constituted by active or middle voice (**ACT, MIDD**) **Periphrasis** non-terminals followed by active voice (**ACT**) **Non Finite Verb Phrase** non-terminals, the second one accounts for impersonal active voice (**ACT_IMP**) configurations constituted by active or middle voice (**ACT, MIDD**) **Periphrasis** non-terminals followed by impersonal active voice (**ACT_IMP**) **Non Finite Verb Phrase** non-terminals, and the third one, finally, accounts for impersonal active voice (**ACT_IMP**) configurations constituted by impersonal active or middle voice (**ACT_IMP, MIDD_IMP**) **Periphrasis** non-terminals followed by active voice (**ACT**) **Non Finite Verb Phrase** non-terminals.

Without non-functional clitic pronouns, **Periphrastic Verb Phrase/5** non-terminal accounts also for personal and impersonal passive voice configurations (**PASS, PASS_IMP**).

```
Periphrastic Verb Phrase ( mvtype, pass, person, number, mood ):
# sigue siendo vendido
  { pass :: PASS }, { mood :: INDICATIVE|SUBJUNCTIVE },
    Periphrasis ( non_finite_mood, ACT, person, number, mood ),
    Non Finite Verb Phrase ( mvtype, pass, non_finite_mood );
# hay que ser querido
  { pass :: PASS_IMP },
    Periphrasis ( non_finite_mood, ACT_IMP, person, number,
      mood ),
    Non Finite Verb Phrase ( mvtype, PASS, non_finite_mood ).
```

With one non-functional clitic pronoun, **Periphrastic Verb Phrase/5** non-terminal accounts for personal middle voice and impersonal middle, active or passive voice configurations (**MIDD, MIDD_IMP, ACT_IMP+PASS, ACT_IMP, PASS_IMP**).

```
Periphrastic Verb Phrase ( mvtype, voice, person, number, mood ):
# sigo vistiéndome, me pongo a hacerme
  { voice :: MIDD },
    Periphrasis ( non_finite_mood, ACT|MIDD, person, number,
      mood ),
    Non Finite Verb Phrase ( mvtype, voice, non_finite_mood,
      case, person, gender, number );
# sigue saliéndose, sigue vendiéndose
  { voice :: ACT_IMP|ACT_IMP+PASS },
    { mood :: INDICATIVE|SUBJUNCTIVE },
    Periphrasis ( non_finite_mood, ACT, person, number, mood ),
    Non Finite Verb Phrase ( mvtype, voice, non_finite_mood,
```

```

        VERB, person, gender, number );
# sigue tratándose
  { voice :: MIDD_IMP }, { mood :: INDICATIVE|SUBJUNCTIVE },
    Periphrasis ( non_finite_mood, ACT|MIDD, person, number,
      mood ),
    Non Finite Verb Phrase ( mvtype, voice, non_finite_mood,
      case, person, gender, number );
# hay que vestirse, hay que ponerse a hacerse
  { voice :: MIDD_IMP },
    Periphrasis ( non_finite_mood, ACT_IMP|MIDD_IMP, person,
      number, mood ),
    Non Finite Verb Phrase ( mvtype, MIDD, non_finite_mood,
      case, person, gender, number );
# me sigo vistiendo
  { voice :: MIDD },
    Periphrasis ( non_finite_mood, ip_MIDD, person, number,
      mood ),
    Non Finite Verb Phrase ( mvtype, voice,
      non_finite_mood );
# se sigue vendiendo
  { mvtype :: SDO|SDOIO|SDOPR }, { voice :: ACT_IMP+PASS },
    { person :: THIRD }, { mood :: INDICATIVE|SUBJUNCTIVE },
    Periphrasis ( non_finite_mood, ip_MIDD, person, number,
      mood ),
    Non Finite Verb Phrase ( mvtype, ACT, non_finite_mood );
# se sigue saliendo
  { voice :: ACT_IMP }, { person :: THIRD }, { number :: SING },
    { mood :: INDICATIVE|SUBJUNCTIVE },
    Periphrasis ( non_finite_mood, ip_MIDD, person, number,
      mood ),
    Non Finite Verb Phrase ( mvtype, ACT, non_finite_mood );
# se sigue tratando
  { voice :: MIDD_IMP }, { person :: THIRD }, { number :: SING },
    { mood :: INDICATIVE|SUBJUNCTIVE },
    Periphrasis ( non_finite_mood, ip_MIDD, person, number,
      mood ),
    Non Finite Verb Phrase ( mvtype, voice,
      non_finite_mood );
# hay que seguirse vistiendo
  { voice :: MIDD_IMP },
    Periphrasis ( non_finite_mood, ip_MIDD_IMP, person, number,
      mood ),
    Non Finite Verb Phrase ( mvtype, MIDD, non_finite_mood );
# se sigue siendo querido
  { voice :: PASS_IMP }, { person :: THIRD }, { number :: SING },
    { mood :: INDICATIVE|SUBJUNCTIVE },
    Periphrasis ( non_finite_mood, ip_MIDD, person, number,
      mood ),
    Non Finite Verb Phrase ( mvtype, PASS, non_finite_mood ).

```

As can be observed in the rewrite rule of **Periphrastic Verb Phrase/5**, the rewrite rules of periphrastic VPs that include clitic pronouns must contain different alternatives for the placement of the clitic pronouns either in the periphrastic group or in the main one—in the first case, the periphrastic groups have voice values in the domain of the affix non-terminal *improper_voice*. While in the previous rule the first four alternatives account for the attachment of the non-functional clitic pronoun to the **Non Finite Verb Phrase** non-terminal, the remaining six of them account for the attachment of the non-functional clitic pronoun to the **Periphrasis** non-terminal.

2.2.3.7.2 Periphrastic Verb Phrase/9 **Periphrastic Verb Phrase** non-terminals with one functional clitic pronoun account for active, middle or passive personal or impersonal configurations, the whole collection of **PREDICATE voice** values: **ACT**, **ACT_IMP**, **MIDD**, **MIDD_IMP**, **PASS**, **PASS_IMP**, **ACT_IMP+PASS**.

In the alternatives that include the functional clitic pronoun in the main group, predicate **two person values** regulates the compatibility of the values for affix *person* of, respectively, the verb form contained in the periphrastic group and the functional clitic pronoun supplied by the main group. In the alternatives that include the functional clitic pronoun in the periphrastic group, predicates **verb one CIPr** or **impersonal verb one CIPr** evaluate the compatibility of the clitic pronoun supplied by the periphrastic group and the main verb type in the main group. In general, in the rewrite rules of periphrastic finite VPs that include functional clitic pronouns, the predicates that regulate the compatibility of the values for affix **person** (**two person values**, **three person values**) appear in the alternatives in which the clitic pronouns are attached to the **Non Finite Verb Phrase** non-terminals. Conversely, the guards or predicates that regulate the compatibility of the values for affixes *case* and *vtype* (**verb one CIPr**, **verb two CIPrs**, **impersonal verb one CIPr**) appear in the alternatives in which the clitic pronouns are attached to the **Periphrasis** non-terminal.

Without non-functional clitic pronouns, **Periphrastic Verb Phrase/9** non-terminals account for active personal or impersonal configurations (**ACT**, **ACT_IMP**).

```

Periphrastic Verb Phrase ( one_clitic, act, person, number, mood,
    function_case, person_one, gender_one, number_one ):
# sigo vendiéndolo, me pongo a hacerlo
{ act :: ACT },
    Periphrasis ( non_finite_mood, ACT|MIDD, person, number,
        mood ),
    Non Finite Verb Phrase ( one_clitic, act,
        non_finite_mood, function_case, person_one, gender_one,
        number_one ),
    two persons values ( mood, person, person_one );
# sigue habiéndolo
{ act :: ACT_IMP }, { person :: THIRD }, { number :: SING },
    { mood :: INDICATIVE|SUBJUNCTIVE },
    Periphrasis ( non_finite_mood, ACT|MIDD, person, number,
        mood ),
    Non Finite Verb Phrase ( one_clitic, act,
        non_finite_mood, function_case, person_one, gender_one,
        number_one );
# hay que venderlo, hay que ponerse a hacerlo
{ act :: ACT_IMP },
    Periphrasis ( non_finite_mood, ACT_IMP|MIDD_IMP, person,
        number, mood ),
    Non Finite Verb Phrase ( one_clitic, ACT,
        non_finite_mood, function_case, person_one, gender_one,
        number_one );
# lo sigo vendiendo
{ act :: ACT },
    Periphrasis ( non_finite_mood, ip_ACT, person, number, mood,
        function_case, person_one, gender_one, number_one ),
    Non Finite Verb Phrase ( one_clitic, act,
        non_finite_mood ),
    verb one ClPr ( one_clitic, function_case );
# lo sigue habiendo
{ act :: ACT_IMP }, { person :: THIRD }, { number :: SING },
    { mood :: INDICATIVE|SUBJUNCTIVE },
    Periphrasis ( non_finite_mood, ip_ACT, person, number, mood,
        function_case, person_one, gender_one, number_one ),
    Non Finite Verb Phrase ( one_clitic, act,
        non_finite_mood ),
    verb one ClPr ( one_clitic, function_case );
# lo hay que vender
{ act :: ACT_IMP },
    Periphrasis ( non_finite_mood, ip_ACT_IMP, person, number,
        mood, function_case, person_one, gender_one, number_one ),
    Non Finite Verb Phrase ( one_clitic, ACT,
        non_finite_mood ),
    verb one ClPr ( one_clitic, function_case ).

```

Without non-functional clitic pronouns, **Periphrastic Verb Phrase/9 non-**

terminal accounts also for passive personal or impersonal configurations (**PASS**, **PASS_IMP**).

```
Periphrastic Verb Phrase ( one_clitic, pass, person, number, mood,
    case_one, person_one, gender_one, number_one ):
# sigue siéndole entregado
  { pass :: PASS }, { mood :: INDICATIVE|SUBJUNCTIVE },
  Periphrasis ( non_finite_mood, ACT, person, number, mood ),
  Non Finite Verb Phrase ( one_clitic, pass,
    non_finite_mood, case_one, person_one, gender_one,
    number_one ),
  two persons values ( mood, person, person_one );
# hay que serle presentado
  { pass :: PASS_IMP },
  Periphrasis ( non_finite_mood, ACT_IMP, person, number,
    mood ),
  Non Finite Verb Phrase ( one_clitic, PASS,
    non_finite_mood, case_one, person_one, gender_one,
    number_one );
# le sigue siendo entregado
  { one_clitic :: SDOIO }, { pass :: PASS },
  { mood :: INDICATIVE|SUBJUNCTIVE }, { case_one :: DAT },
  Periphrasis ( non_finite_mood, ip_ACT, person, number, mood,
    case_one, person_one, gender_one, number_one ),
  Non Finite Verb Phrase ( one_clitic, pass,
    non_finite_mood ).
```

With one non-functional clitic pronoun, **Periphrastic Verb Phrase/9** non-terminals account for personal middle voice and impersonal middle, active or passive voice configurations (**MIDD**, **MIDD_IMP**, **ACT_IMP+PASS**, **ACT_IMP**, **PASS_IMP**).

```
Periphrastic Verb Phrase ( one_clitic, voice, person, number, mood,
    case_one, person_one, gender_one, number_one ):
# sigo creyéndomelo, me pongo a hacérmelo
  { voice :: MIDD },
  Periphrasis ( non_finite_mood, ACT|MIDD, person, number,
    mood ),
  Non Finite Verb Phrase ( one_clitic, voice,
    non_finite_mood, case, person, gender, number, case_one,
    person_one, gender_one, number_one );
# sigue queriéndoselo, sigue entregándosele
  { voice :: ACT_IMP|ACT_IMP+PASS },
  Periphrasis ( non_finite_mood, ACT, person, number, mood ),
  Non Finite Verb Phrase ( one_clitic, voice,
    non_finite_mood, case, person, gender, number, case_one,
    person_one, gender_one, number_one );
# hay que creérselo, hay que ponerse a hacérselo
```

```

{ voice :: MIDD_IMP },
  Periphrasis ( non_finite_mood, ACT_IMP|MIDD_IMP, person,
    number, mood ),
  Non Finite Verb Phrase ( one_clitic, MIDD,
    non_finite_mood, case, person, gender, number, case_one,
    person_one, gender_one, number_one );
# me lo sigo creyendo
{ voice :: MIDD },
  Periphrasis ( non_finite_mood, ip_MIDD, person, number, mood,
    case_one, person_one, gender_one, number_one ),
  Non Finite Verb Phrase ( one_clitic, voice,
    non_finite_mood ),
  verb one ClPr ( one_clitic, case_one );
# se le sigue entregando
{ one_clitic :: SDOIO }, { voice :: ACT_IMP+PASS },
  { mood :: INDICATIVE|SUBJUNCTIVE }, { person :: THIRD },
  { case_one :: DAT },
  Periphrasis ( non_finite_mood, ip_MIDD, person, number, mood,
    case_one, person_one, gender_one, number_one ),
  Non Finite Verb Phrase ( one_clitic, ACT,
    non_finite_mood ),
  impersonal verb one ClPr ( one_clitic, case_one );
# se lo sigue queriendo
{ voice :: ACT_IMP }, { person :: THIRD }, { number :: SING },
  { mood :: INDICATIVE|SUBJUNCTIVE },
  Periphrasis ( non_finite_mood, ip_MIDD, person, number, mood,
    case_one, person_one, gender_one, number_one ),
  Non Finite Verb Phrase ( one_clitic, ACT,
    non_finite_mood ),
  impersonal verb one ClPr ( one_clitic, case_one );
# hay que seguirselo creyendo
{ voice :: MIDD_IMP },
  Periphrasis ( non_finite_mood, ip_MIDD_IMP, person, number,
    mood, case_one, person_one, gender_one, number_one ),
  Non Finite Verb Phrase ( one_clitic, MIDD, tense,
    non_finite_mood ),
  verb one ClPr ( one_clitic, case_one );
# se le sigue siendo presentado
{ one_clitic :: SDOIO }, { voice :: PASS_IMP }, { person :: THIRD },
  { number :: SING }, { mood :: INDICATIVE|SUBJUNCTIVE },
  { case_one :: DAT },
  Periphrasis ( non_finite_mood, ip_MIDD, person, number, mood,
    case_one, person_one, gender_one, number_one ),
  Non Finite Verb Phrase ( one_clitic, PASS,
    non_finite_mood ).

```

2.2.3.7.3 Periphrastic Verb Phrase/13 Periphrastic Verb Phrase non-terminals with two functional clitic pronouns account for personal or impersonal

active voice configurations, **ACT** or **ACT_IMP** voice values. In the alternatives that include the clitic pronouns in the main group, the predicate **three person values** regulates the compatibility of the values for affix *person* of, respectively, the verb form contained in the periphrastic group and the functional clitic pronouns supplied by the main group. In the alternatives that include the clitic pronouns in the periphrastic group, the predicate **verb two CIPrs** evaluates the compatibility of the second clitic pronoun supplied by the periphrastic group and the main verb type in the main group.

```

Periphrastic Verb Phrase ( two_clitics, act, person, number, mood,
    case_one, person_one, gender_one, number_one, case_two,
    person_two, gender_two, number_two ):
# sigo entregándoselo, me pongo a hacérselo
{ act :: ACT },
    Periphrasis ( non_finite_mood, ACT|MIDD, person, number,
        mood ),
    Non Finite Verb Phrase ( two_clitics, act,
        non_finite_mood, case_one, person_one, gender_one,
        number_one, case_two, person_two, gender_two,
        number_two ),
    three persons values ( mood, person, person_one, person_two );
# hay que entregárselo, hay que ponerse a hacérselo
{ act :: ACT_IMP },
    Periphrasis ( non_finite_mood, ACT_IMP|MIDD_IMP, person,
        number, mood ),
    Non Finite Verb Phrase ( two_clitics, ACT,
        non_finite_mood, case_one, person_one, gender_one,
        number_one, case_two, person_two, gender_two,
        number_two );
# se lo sigo entregando
{ act :: ACT },
    Periphrasis ( non_finite_mood, ip_ACT, person, number, mood,
        case_one, person_one, gender_one, number_one, case_two,
        person_two, gender_two, number_two ),
    Non Finite Verb Phrase ( two_clitics, act,
        non_finite_mood ),
    verb two CIPrs ( two_clitics, case_two );
# se lo hay que entregar
{ act :: ACT_IMP },
    Periphrasis ( semiaux, ip_ACT_IMP, person, number, mood,
        case_one, person_one, gender_one, number_one, case_two,
        person_two, gender_two, number_two ),
    Non Finite Verb Phrase ( two_clitics, ACT,
        non_finite_mood ),
    verb two CIPrs ( two_clitics, case_two ).

```

2.2.3.7.4 Periphrastic Non Finite Verb Phrase/3 Periphrastic Non Finite Verb Phrase non-terminals without clitic pronouns account for active configurations (**ACT**).

```
Periphrastic Non Finite Verb Phrase ( mvtype, ACT, mood ):
# seguir saliendo, ponerse a hacer
  ( Periphrasis ( non_finite_mood, ACT, mood );
    Periphrasis ( non_finite_mood, MIDD, mood, case, person,
                  gender, number ) ),
Non Finite Verb Phrase ( mvtype, ACT, non_finite_mood ).
```

Periphrastic Non Finite Verb Phrase/3 non-terminals account also for passive configurations (**PASS**).

```
Periphrastic Non Finite Verb Phrase ( mvtype, PASS, mood ):
# seguir siendo vendido
  Periphrasis ( non_finite_mood, ACT, mood ),
  Non Finite Verb Phrase ( mvtype, PASS, non_finite_mood ).
```

2.2.3.7.5 Periphrastic Non Finite Verb Phrase/7 Periphrastic Non Finite Verb Phrase non-terminals with one clitic pronoun account for active and middle configurations, **ACT** and **MIDD** configurations. In active voice the clitic pronoun is functional. If in this voice the clitic pronoun is included in the periphrastic group, predicate **verb one ClPr** regulates the compatibility between the values for affixes *mvtype* and *case* of, respectively, the verb form and the clitic pronoun. If, on the contrary, the clitic pronoun is included in the main group, provided that the periphrastic group is in middle voice, the predicate **two persons values** regulates the compatibility of the values for affix *person* of the VP —*person* of the VP is that of the clitic pronoun of the periphrastic group— and the functional clitic pronoun. On the other hand, in middle voice periphrastic VPs the clitic pronoun is non-functional. In this case, when the periphrastic group is in middle voice, the values for affixes *person*, *gender* and *number* of the clitic pronouns of the periphrastic group and of the main group are the same.

```
Periphrastic Non Finite Verb Phrase ( mvtype, voice, mood,
  case_one, person_one, gender_one, number_one ):
# seguir vendiéndolo, seguir vistiéndome
  { voice :: ACT|MIDD },
  Periphrasis ( non_finite_mood, ACT, mood ),
  Non Finite Verb Phrase ( mvtype, voice, non_finite_mood,
    case_one, person_one, gender_one, number_one );
```

```

# ponerme a hacerlo
{ voice :: ACT },
  Periphrasis ( non_finite_mood, MIDD, mood, case, person,
    gender, number ),
  Non Finite Verb Phrase ( mvtype, voice, non_finite_mood,
    case_one, person_one, gender_one, number_one ),
  two persons values ( mood, person, person_one );
# ponerme a hacerme
{ voice :: MIDD },
  Periphrasis ( non_finite_mood, MIDD, mood, case_one,
    person_one, gender_one, number_one ),
  Non Finite Verb Phrase ( mvtype, voice, non_finite_mood,
    case_one, person_one, gender_one, number_one );
# seguirlo vendiendo
{ voice :: ACT },
  Periphrasis ( non_finite_mood, ip_ACT, mood, case_one,
    person_one, gender_one, number_one ),
  Non Finite Verb Phrase ( mvtype, voice,
    non_finite_mood ),
  verb one ClPr ( mvtype, case_one );
# seguirme vistiendo
{ voice :: MIDD },
  Periphrasis ( non_finite_mood, ip_MIDD, mood, case_one,
    person_one, gender_one, number_one ),
  Non Finite Verb Phrase ( mvtype, voice,
    non_finite_mood ).

```

Periphrastic Non Finite Verb Phrase/7 non-terminals account for passive configurations (**PASS**).

```

Periphrastic Non Finite Verb Phrase ( one_clitic, PASS, mood,
  case_one, person_one, gender_one, number_one ):
# seguir siéndole entregado
  Periphrasis ( non_finite_mood, ACT, mood ),
  Non Finite Verb Phrase ( one_clitic, PASS,
    non_finite_mood, case_one, person_one, gender_one,
    number_one );
# seguirle siendo entregado
{ case_one :: DAT },
  Periphrasis ( non_finite_mood, ip_ACT, mood, case_one,
    person_one, gender_one, number_one ),
  Non Finite Verb Phrase ( one_clitic, PASS,
    non_finite_mood ).

```

2.2.3.7.6 Periphrastic Non Finite Verb Phrase/11 **Periphrastic Non Finite Verb Phrase** non-terminals with two clitic pronouns account for active and middle configurations, **ACT** and **MIDD** configurations. In active voice both

clitic pronouns are functional. If in this voice the clitic pronouns are included in the periphrastic group, predicate **verb two CIPrs** regulates the compatibility between the values for affixes *mvtype* and *case* of, respectively, the verb form and the second functional clitic pronoun. If, on the contrary, the clitic pronouns are included in the main group, provided that the periphrastic group is in middle voice, the predicate **three persons values** regulates the compatibility of the values for affix *person* of the VP and the clitic pronouns. In middle voice periphrastic VPs, on the other hand, one clitic pronoun is non-functional and the other is functional. In this voice, if the clitic pronouns are included in the periphrastic group, **verb one CIPr** regulates the compatibility between the values for affixes *mvtype* and *case* of, respectively, the verb form and the functional clitic pronoun. If, on the opposite, the clitic pronouns are included in the main group, provided that the periphrastic group is in middle voice, the values for affixes *person*, *gender* and *number* of the clitic pronouns of the periphrastic group and of the main group are the same.

```

Periphrastic Non Finite Verb Phrase ( mvtype, voice, mood,
    case_one, person_one, gender_one, number_one, case_two,
    person_two, gender_two, number_two ):
# seguir entregándoselo, seguir creyéndoselo
  { voice :: ACT|MIDD },
    Periphrasis ( non_finite_mood, ACT, mood ),
    Non Finite Verb Phrase ( mvtype, voice, non_finite_mood,
        case_one, person_one, gender_one, number_one, case_two,
        person_two, gender_two, number_two );
# ponerme a hacérselo
  { voice :: ACT },
    Periphrasis ( non_finite_mood, MIDD, mood, case, person,
        gender, number ),
    Non Finite Verb Phrase ( mvtype, voice, non_finite_mood,
        case_one, person_one, gender_one, number_one, case_two,
        person_two, gender_two, number_two ),
    three persons values ( mood, person, person_one, person_two );
# ponerme a hacérmelo
  { voice :: MIDD },
    Periphrasis ( non_finite_mood, MIDD, mood, case_one,
        person_one, gender_one, number_one ),
    Non Finite Verb Phrase ( mvtype, voice, non_finite_mood,
        case_one, person_one, gender_one, number_one, case_two,
        person_two, gender_two, number_two );
# seguirselo entregando
  { voice :: ACT },
    Periphrasis ( non_finite_mood, ip_ACT, mood, case_one,

```

```

        person_one, gender_one, number_one, case_two, person_two,
        gender_two, number_two ),
    Non Finite Verb Phrase ( mvtype, voice,
        non_finite_mood ),
    verb two ClPrs ( mvtype, case_two );
# seguirselo creyendo
{ voice :: MIDD },
    Periphrasis ( non_finite_mood, ip_MIDD, mood, case_one,
        person_one, gender_one, number_one, case_two, person_two,
        gender_two, number_two ),
    Non Finite Verb Phrase ( mvtype, voice,
        non_finite_mood ),
    verb one ClPr ( mvtype, case_two ).

```

Chapter 3

Argument and Circumstance Constituents of the Clause: Noun, Pronoun, Adverb and Adjective Phrases

3.1 Description of Noun, Pronoun, Adverb and Adjective Phrases

From the point of view of formalization of clause analysis, the description of —non-verb— phrase-level analysis needs to cover the linguistic resources that play a role in the conversion of nouns, pronouns, adverbs and adjectives into syntagms labelled as functional nodes at clause-level. Or, in other words, non-verb phrase-level analysis should describe how noun, pronoun, adverb and adjective lexical categories are contextualized. Various entirely syntactic mechanisms were identified in Chapter 1 which are recursively used in the process of contextualization of lexical categories to form the phrases that constitute clauses. There, both a formal description and, according to this, a schematic AGFL formalization for these mechanisms was introduced. Of these mechanisms, we deal in the current Chapter with subordination, trying to show practically how the abstract schemes proposed in Chapter 1 are applied in a concrete AGFL formal grammar for Spanish.

Taking into account these considerations with respect to the whole Chapter, the aim of the current Section is to illustrate the application of subordination mech-

anisms for the production of linguistic unities labelled as the syntactic categories **Noun Phrase**, **Pronoun Phrase**, **Adjective Phrase** and **Adverb Phrase**—next Section 3.2 will show how this production can be expressed in AGFL formal terms. Even so, rather than developing an exhaustive and strictly proper description of the linguistic structures under consideration, we want to remark that our claim in this Chapter is to show how the proposed AGFL formalization of subordination can be similarly expressed for all of these structures, creating so-called *frames* which permit regular extensions of the grammar if necessary.

That is, as was observed in Chapter 1—but it must be especially remarked here—the objective of this exposition with respect to non-verb phrases is the effective formalization of certain linguistic mechanisms, showing the structure and methodology of the proposed formal grammar—and ensuring, of course, an acceptable coverage—, rather than the grammar of structures itself, which is in fact conceived of as a never-ending task. Apart from language structure itself, three main—more practical—reasons for this were already pointed out in Chapter 1, so we will refer to them here very briefly: the lack of descriptions, both non-formal and formal, about the phrase-level grammar and about the behaviour of closed categories of words in Spanish,—closely related to this—the lack of linguistic structures and definitions—or their rigidity—in which the particular behaviours of, mainly, closed categories of words can be properly described, and adaptation to the current state of the lexicon, which, among other less relevant deficiencies, lacks all kind of semantic information for nouns and adjectives. It must be clear, in consequence, that we will not develop a detailed description of those linguistic facts only partially or not at all covered by the grammar. As already pointed out in Chapter 1, the main example of this is precisely the argumental behaviour of nouns and adjectives. For them, the absence of data made impossible to consider, given the time available, the effective formalization of such phenomena, although their integration in the grammar has been foreseen and will be carried out in the future.

In line with this, it is necessary to insist also (see Chapter 1) on the fact that completeness of the description, due to the very different nature of linguistic phenomena involved, has been pursued and evaluated with very different criteria in the case of Verb Phrases and in the case of Noun, Pronoun, Adverb and Adject-

ive Phrases. While for VPs, completion of the description, even beyond corpus evidence, was found feasible, for Noun, Pronoun, Adjective and Adverb Phrases completion must be estimated on the basis of how simple it is to extend the grammar for covering structures and lexical combinations previously left out on the grounds of corpus evidence. Being the current specifications strictly derived from random searches performed on the ARTHUS corpus (non-oral section), we consider them circumstantial, and we specify them in the AGFL formalization, see next Section 3.2, of the more general linguistic facts that we describe in the current one.

Coordination apart, from an external point of view Noun Phrase, Pronoun Phrase, Adverb Phrase and Adjective Phrase are the syntactic categories that constitute the syntagms that underly the syntactic functions **SUBJECT**, **DIRECT OBJECT**, **INDIRECT OBJECT**, **PREPOSITIONAL COMPLEMENT**, **PREDICATIVE COMPLEMENT**, **AGENT** and all kind of **CIRCUMSTANCES**. From an internal point of view, Noun, Pronoun, Adverb and Adjective Phrases obviously result from the contextualization, by means of the application of subordination —as defined in Chapter 1—, of, respectively, noun, pronoun, adverb and adjective lexical categories. During this process of contextualization, two different kinds of subordinating operations expanding lexical categories may be applied, have been traditionally identified and will be accounted for here: **addition of modifiers** and **addition of determiners**.

The intersection between these two subordinating operations —addition of modifiers and addition of determiners— and the four lexical and phrase categories —noun, pronoun, adverb and adjective— will be leading our description in this Section, turned into an AGFL formal description, in the next one. After a brief definition of phrasal and lexical categories (Section 3.1.1), we will start from the description of addition of modifiers (Section 3.1.2), as a subordinating mechanism, to all the lexical categories to which it can be applied —nouns, pronouns, adverbs and adjectives—, thus yielding the corresponding phrasal categories. After, we will describe, again as a subordinating mechanism, addition of determiners (Section 3.1.3) to, again, all the lexical categories to which it can be applied —nouns and pronouns—, thus yielding the corresponding phrasal categories too. Finally, we will conclude this Section with brief descriptions of modifiers by themselves

(Section 3.1.4), of other constituents of the clause (Sections 3.1.5 and 3.1.6) and of prepositional contextualizations of previously described structures (Section 3.1.7).

Apart from subordination, our particular formalization of Noun, Pronoun, Adverb and Adjective Phrases is conditioned by the purpose of discriminating the phrases that are capable of introducing particular types of clauses —interrogative, exclamative or relative clauses—, whose formal description, on the basis of the phrases that introduce them, will be dealt with at clause-level by means of specific rules in each case.

3.1.1 Noun, Pronoun, Adverb and Adjective Phrase

Noun Phrase (henceforth **NP**) is the syntactic category based on a noun lexical category. Nouns, as lexical categories, can be proper (*Santiago Nasar, La Voz de Galicia, L^AT_EX 2_ε*) or common nouns (*hombre, periódico, procesador de textos*). Certain nouns establish with prepositional phrases relationships similar to those established by verbs with their various complements (*el comprador de la casa, la venta de la casa, la entrega de las llaves de la ciudad a Su Santidad, la tendencia de los iraníes a magnificar sus acciones militares*).

Nouns are inflected for gender and number. Values for gender and number, together with noun type (common or proper), will be the only information available for analysis in our formal grammar.

In addition to noun type, gender and number, NPs refer also to person. Traditionally, NPs have been considered to be always third person singular or plural items. However, this is only partially true: singular NPs always agree for third person singular, but plural NPs, even though more frequently agree for third person plural, can also agree for first and second person plural (*Las mujeres ya no aceptamos que los hombres pretendáis que no os corresponde una parte igual del trabajo doméstico*).

Nouns to which determiners or modifiers can be added are common nouns. Sometimes also proper nouns are modified or determined, but in these cases they are somewhat used as common nouns (*la María que yo conocía*).

Pronoun Phrase (henceforth **PrP**) is the syntactic category based on a pronoun lexical category. Pronouns, as lexical categories, constitute a closed category of words. This means, on the one hand, that they can be enumerated, and, on

the other hand, that the items in the class show a very large variation in their particular behaviours.

Pronoun items have been traditionally classified into different types and subtypes, but different linguistic descriptions have identified different groups of pronouns, taking into account different particularities of their behaviour. To account for all these particularities, almost each pronoun lemma should be considered a different type. To account for those considered especially relevant in this book, we have distinguished the types and subtypes presented in Section 3.2.2.2, assigning them to pronoun items as showed in Appendix D. The collection of subtypes is very detailed to distinguish each minimal set of pronouns with exactly the same syntactic behaviour, for each set of this type, a pronoun lemma was arbitrary selected as its representative. Combined with the collection of types, this collection of subtypes can account for the most particular behaviours of pronouns, remarkably for how they can be combined with determiners in the PrP.

Pronoun types permit the distinction of relative, interrogative, exclamative and general —declarative— PrPs (*La mano que mece la cuna es la mano que domina el mundo*, *¿Quién teme a Virginia Wolf?*, *¡Qué me dices!*, *Alguien voló sobre el nido del cuco*). They ensure also that the especial behaviours of comparative and personal pronouns, consisting of the facts that comparative pronouns might introduce a second term of comparison or a consequence (*Aquí hay más que allí*, *Había tantos que no se podían contar*), and that personal pronouns are inflected for case and person (*De mí no se ríe nadie*, *Yo no tengo rival en esta competición*), can be accounted for.

All pronouns are inherently marked for determination, but different types of pronouns are inflected for other categories different from determination. Personal pronouns are inflected for gender, number, person and case. Possessive pronouns are inflected for gender, number and person. The rest of pronoun types are only inflected for gender and number. Although most non-personal pronouns have with respect to person the same behaviour registered for nouns, their capability to refer to first and second person depends to a certain extent on their type. Most of them can refer to first and second person only in plural (*Muchos hemos decidido quedarnos*, *¿Quiénes hemos de hacerlo?*, *Unos os habéis portado peor que otros*), but, relative pronouns, for instance, can also refer to them in singular (*Yo que*

te hubiese querido hasta el fin, sé que te arrepentirás, Tú que nunca dices nada, hoy has dicho demasiado), while possessive pronouns, on the contrary, neither in singular nor in plural number can refer to first or second person.

Adverb Phrase (henceforth **AvP**) is the syntactic category based on an adverb lexical category. Adverbs, as lexical categories, include items derived from adjectives —mainly by means of suffix *-mente*, but also by fixing the adjective stem in neuter gender—, and non-derivative adverbs, which, as pronouns, are a closed category of words that shows quite large variations in their behaviour.

Adverb items are also classified into different types and subtypes. Adverbs derived from adjectives are generally classified as manner adverbs. There is, nevertheless, less general agreement with respect to the classification of non-derivative adverbs, both with respect to the classification of each item and to the collection of types needed to classify all the items. Trying to account for the relevant particular behaviours of each item in the collection of the non-derivative adverbs, we have distinguished here the types and subtypes described in Section 3.2.2.5, assigning them to particular items as showed by Appendix E.

Not every adverb can be the nucleus of an AvP, those which cannot —modifying adverbs— will be introduced below (Sections 3.1.4.2.2, 3.2.2.9 and 3.2.2.6). For those which can, on the contrary, adverb types permit the distinction of relative, interrogative, exclamative and general AvPs (*La casa donde nació ya no existe, ¿Dónde está?, ¡Adónde hemos llegado!, Ayer se fue*).

Certain adverbs are inherently marked for degree either as positive adverbs (*tarde*), which can be graded according to the values described in Sections 3.1.4.2.2 and 3.2.2.9 (*muy tarde, demasiado tarde, más tarde que los demás*), or as inherently graded adverbs (*mejor que los demás, mucho, demasiado*). Possible values for degree are enumerated in Section 3.2.2.9, and their association with lexical items can be found in Appendix E.

Finally, **Adjective Phrase** (henceforth **AjP**) is the syntactic category based on an adjective lexical category. Adjectives, as lexical categories, are classified into different types. Those considered in this work can be found in Section 3.2.2.7. A number of adjectives, as nouns do, are specified by prepositional complements (*fiel a sus reaccionarios principios éticos, difícil de llevar a cabo en tan poco tiempo, indiferente al sufrimiento de su propio pueblo*). Most adjectives can be placed

both before or after the nucleus that they modify, but some of them, for different reasons, can only be placed in one of these positions, either before or after the nucleus.

Adjectives, as adverbs, can be inherently marked for degree, as positive adjectives (*simpático*), which can be graded according to the values described in Sections 3.1.4.2.2 and 3.2.2.9 (*muy simpático*, *poco inteligente*, *tan extrovertido como su hermano menor*), or as inherently graded adjectives (*mejor de lo que esperábamos*, *óptimo*).

Adjectives are inflected for gender and number. Values for gender and number, together with adjective type, adjective position and degree, will be the only information available for analysis in our formal grammar.

NPs, **PrPs** and **AvPs**, very rarely **AjPs**, might be introduced by prepositions, thus constituting prepositional NPs, PrPs and AvPs (*por una buena causa*, *por ella*, *por debajo de la mesa*). This means that in our approach prepositional phrases are simply conceived of as particular organizations of phrasal categories, required by their contextualization.

3.1.2 Addition of Modifiers

Noun, pronoun, adverb and adjective meanings are specified by means of the addition of modifiers. A large number of different categories can be used as modifiers. Not all of them, nevertheless, in all contexts, that is, modifying any phrasal category. Certain types of modifiers are only compatible with certain phrasal categories, or, at least, they have some particularities when modifying some phrasal categories. In spite of these remarks, however, the claim that the mechanism of adding modifiers is in all cases the same cannot be refuted, since little differences regard almost exclusively the productivity of the mechanism and are caused by the different nature of the modified categories.

Modifiers can be added before or after the modified item. We deal with premodifiers and postmodifiers in different paragraphs. Besides, because the most productive modifiers show their highest degree of productivity within the context of NPs, we will start by describing addition of modifiers within NPs, proceeding after through PrPs and AvPs to AjPs, showing for each of these the specific particularities that they present with respect to the addition of modifiers within the

context of NPs.

3.1.2.1 Addition of Modifiers in Noun Phrase

Different categories can be found as noun modifiers, and at different levels. A noun can be modified by the addition of an adjective postponed to the noun (*una propuesta interesante*). The adjective in this position can in turn be modified, and so we can find an AjP in this position (*la propuesta menos interesante de todas*) (see Section 3.1.2.4). More than one adjective can be added behind the noun. As this addition is recursive, each adjective modifies the NPs constituted by the noun modified by all the previous adjectives found in the linguistic sequence (*la crisis económica mundial*). Semantic properties govern the compatibility of adjectives and the order in which they can be added to the noun. The more restrictive ones are placed the nearer to the noun. Only the last one can in turn be modified (*la crisis económica mundial más aguda de la historia*).

We have already mentioned the fact that some nouns can maintain with prepositional complements syntactic relationships similar to those entailed by verbs with their complements. We also remarked that the kind of nouns —and prepositional phrases required by them— that maintain this type of syntactic functional relations could not be accounted for here because of lack of data. Even so, we make here an attempt to find out how these prepositional complements relate to other modifiers in the whole NP, outlining at the same time how they should be integrated in the grammar in the future. Thus, our description here of prepositional phrase noun complements will be very brief, although it will refer to both the prepositional complements entering argumental structures, which mainly modify nouns morphologically derived from verbs, and the general ones, which modify all kind of nouns. Various patterns of relations between verbs and nouns morphologically derived from them can be distinguished:

- Nouns that refer to the subject of the original verb scheme from which they are obtained: *productor* from *producir*, *amante* from *amar*. These nominalizations are usually referred to as nominalizations of the agent. Quite regularly, they keep —except, obviously, for the subject— the same elements of the scheme of the verb. While the direct object is turned into a

prepositional phrase introduced by preposition *de*, the rest of the arguments maintain the same structures that appear in the verb schemes. The nominalization of the subject argument of the scheme *amar*, *SDO*, *ACT*, for example, produces *amante*, *DO*: *el amante de lady Chatterly*. The nominalization of the subject argument of the scheme *introducir*, *SDOPC*, *ACT*, *EN* produces *introducción*, *DOPC*, *EN*: *el principal introducción de la mecanización en la agricultura*.

- Nouns that refer in general to the direct object of the original verb scheme from which they are obtained: *producto* from *producir*, *composición* from *componer*. These nominalizations are usually considered as nominalizations of the result of the activity of the original verb. Although this is correct, it must be further specified and we should, more exactly, say that these are nominalizations of the result of the activity referred to by the verb when the direct object of the verb in question is created by means of the activity referred to by the verb—if, on the contrary, it is not, the resulting nominalization corresponds to the following group. While, quite obviously, nominalizations of this type do not keep the direct object of the verb as an argument, the subject can be considered to be preserved—instead of considering it just as a possessive prepositional complement, see below in this Section—, although turned into a prepositional phrase introduced by preposition *de*, as well as the rest of the arguments, which maintain the same structures found in the verb schemes. The nominalization of the direct object argument of the scheme *componer*, *SDO*, *ACT*, for example, produces *composición*, *S*: *la importante composición de Mozart*. The nominalization of the direct object argument of the scheme *indicar*, *SDOIO*, *ACT*, for example, produces *indicación*, *SIO*: *ciertas indicaciones del maestro a sus colaboradores*.
- Nouns that refer to the result of the activity of the original verb, when this result is not itself an argument of the original verb: *análisis* from *analizar*, *copia* from *copiar*. These nominalizations have been also considered as nominalizations of the result of the activity of the original verb, but an important difference with the previous type of nominalizations arises from the character of the direct object of the original verb, because in this case

the direct object of the verb is not itself the result of the activity of the verb. On the contrary, the direct object of these verbs exists before the activity referred to by the verb takes place, it is in fact the entity to which the activity, which produces something else, is applied. Nominalizations of this type maintain the direct object of the verb. Both the subject, if we consider that it is preserved, and the direct object, are turned into prepositional phrases introduced by preposition *de*. The rest of the arguments keep the same structures found in the verb schemes. The nominalization of the scheme *analizar*, *SDO*, *ACT*, for example, produces *análisis*, *SDO*: *el análisis de la cuestión de esta empresa encuestadora*.

- Nouns that refer to the action named by the original verb scheme from which they are obtained: *llegada* from *llegar*, *absorción* from *absorber*. In this case, transitive verbs preserve the reference to the direct object, which is turned into a prepositional phrase introduced by preposition *de*. The subject is also maintained. When the direct object is not specified, the subject can be introduced by preposition *de*, when combined with the direct object it is usually introduced by prepositions *por*, *por parte de* and the like. For intransitive verbs, the subject, if we consider that it is preserved, is always introduced by preposition *de*. Other complements keep the structures found in the verb schemes. Nominalization of the scheme *llegar*, *SPC*, *A*, produces *llegada*, *SPC*, *A*: *la llegada de la presidenta a la reunión*. Nominalization of the scheme *absorber*, *SDO*, produces *absorción*, *SDO*: *la absorción de la empresa por la banca*.

However, these are but general tendencies for nouns to behave, nouns can be more or less close to one of these groups, or to more than one of them if they have different meanings. Quite frequently, for example, we find nouns which can be considered as nominalizations either of the result or of the action of the verb in question (*el análisis de sangre*, *el análisis de los resultados por los evaluadores*). Not only this, but we find also that the status of the complement that introduces the subject of the original verb has caused too considerable controversy. Occasionally it has been in fact interpreted simply as a possessive complement, an interpretation which appears to be especially more appropriate in the case of nominalizations of

the result of the verb, because these mostly account for concrete objects of the world that are usually owned by a possessor. On the other hand, there are also nouns which can hardly be connected to any of the possible schemes of the verb from which they are obtained, developing instead an argument structure that is particular to them (*la introducción de Coseriu a la semántica estructural*).

In addition to this, nouns derived from verbs can also be modified by prepositional phrases specifying the kind of peripheral information introduced by circumstances, adjuncts at the level of the clause (*la llegada de S.M. Don Juan Carlos a las 14:30, la ejecución del reo en la cárcel*). Intuitively, the nature of these adjuncts is similar to the nature of clause adjuncts. This intuition is supported by the fact that other linguistic categories, apart from prepositional phrases, can be found in both contexts: gerund clauses, for instance (*la extracción de piezas dentales utilizando las tenazas*).

Nouns not derived from verbs, on the other hand, are also modified by prepositional complements. These complements can be more or less required by the noun, depending on its meaning. Prepositional complements of nouns derived from adjectives (*la palidez de los vampiros, la brevedad de la vida*), of nouns indicating relationships between human beings (*la madre de Dios, el marido de la peluquera, los amigos de Peter*), and of iconic nouns, both those referring to the objects and those referring to the authors (*la fotografía del Che Guevara, el retrato de Dorian Gray, el libro de Gabriel García Márquez*), have been traditionally considered as complements required by the noun¹. Prepositional complements indicating possessor (*el arca de Noé, la caja de Pandora*), and all kind of prepositional complements indicating place, time, manner. . . , mainly, but not exclusively, introduced by preposition *de* (*el faro del fin del mundo, el expreso de medianoche, la muchacha de las bragas de oro, el hombre del saco*), have been traditionally considered not required by nouns. Especial groups, finally, are constituted by nouns indicating portions or amounts of their complements (*un millón de muertos, por un puñado de dólares*), by prepositional complements indicating the kind of ma-

¹Moreover, prepositional complements required by nouns derived from adjectives have been considered argumental in the same sense as prepositional complements of nouns derived from verbs, a claim that is supported by the fact that some of them, together with the subject of the predication *to be + adjective* (*los vampiros son pálidos/la palidez de los vampiros*), can also refer to an indirect object complement (*Penélope es fiel a Ulises/la fidelidad de Penélope a Ulises*).

terial that something is made of (*vellocino de oro, magnolias de acero, potaje de garbanzos*), and by prepositional complements lexically linked to their nucleus to indicate an especial type of the entity referred to by the nucleus (*coche de caballos, coche de carreras, coche de punto, reloj de pulsera, reloj de pared, reloj de cuco, libro de caballerías, libro de cuentos, libro de aventuras*).

Relative clauses, see Section 3.1.4.1.3, act also as NP modifiers. Although the modified NP mostly agrees on person, gender and number with the relative clause modifier (*el rayo que no cesa, las personas a quienes nada les interesa, el caso en el que te ocupas, la música de la cual sufres la influencia*), there are also certain relative phrases for which the modified NP and the relative clause do not agree (*la empresa cuyos trabajadores se encuentran a gusto, el lugar donde mejor me siento, el lugar de donde vengo, el modo como lo presentas*). With regard to relative clause modifiers, it is doubtful also whether more than one relative clause can be added to the modified NP. The fact that a possible NP introduced after the verb in the innermost relative clause could create confusion about the antecedent of the outermost relative clause, seems to indicate that recursion of relative clause modifiers is not allowed (if possible, this should rarely occur and only in very particular utterances, *las personas a las que vimos ayer que luego estuviste criticando durante todo el día*).

Yet another possible postmodifier of NPs, the last one that we are going to present here, is constituted by appositions. With respect to them, we have to say that in this formalization we only account for non-commas appositions that introduce proper names and specify which concrete entity of the world the common nouns that they modify are applied to (*el tío Pepe, la virgen María, el niño Jesús, el piso A, la calle Zamora*).

So far, then, we have been describing noun modifiers postponed to the noun. Preceding it, on the contrary, we can only find adjective modifiers. Only one adjective can precede the noun (*la buena tierra, grandes esperanzas*). There are also certain adjectives that have especial forms when they precede the noun (*el buen ladrón, el gran Gatsby, un mal enemigo*).

Under certain contextual conditions, either the adjective preceding the noun can be inherently marked for degree (*mejor profesor, peor coche*) or the noun preceded by a positive adjective can still be modified by an adverb that indicates

degree (*muy buena persona, tan difícil tarea*). In both cases, a second term of comparison or a consequence can be introduced after the noun (*más buena persona que tú, tan buena persona que lo toman por tonto*)². Section 3.1.4.2.2 below will be devoted to degree and degree modifiers. The addition of the exclamative degree modifier must be, however, especially remarked here (*¡qué buena persona!*), because it enables the NP to introduce an especial type of clauses, the exclamative ones (*¡Qué buen tiempo hace hoy!*).

In addition to this, comparative degree adverbs *más, menos* allow also the introduction of another adverb that quantifies the comparison (*mucho más buena persona de lo que creíamos*). Apart from other types of adverbs, a relative adverb can be placed in this position, thus enabling the NP to introduce relative clauses with especial functions beyond the level of the clause, at sentence-level (*Cuanto más buena persona seas, más te querremos*). When the degree or comparison quantifier is adverb *poco*, it is still possible to add a second degree adverb (*muy poco buena persona, muy poco más buena persona que tú*)³.

An especial type of modifying NP that indicates also quantification may also appear in similar contexts, although its syntactic position is hierarchically more external, as demonstrated by the fact that it can even modify the quantifier that modifies the comparison (*mil veces más buena persona que tú, mil veces mucho más buena persona que tú*). This NP is mainly used in informal style, it quantifies the comparison using as unit of measurement the difference set by the comparison itself, that is, indicating “how many times more (or less)...”. As a matter of fact, this modifier is always associated to plural of lemma *vez* in Spanish.

Proper nouns, finally, mainly human proper nouns, can also be preceded by an

²As regards the question of the boundaries of the syntactic influence of the degree modifier, we refrain from discussing it here. Obviously, it implies different hierarchical organizations of the linguistic sequence: either the adverb modifies only the adjective, and the AjP constituted by both of them modifies the noun, or the adverb modifies the NP constituted by the adjective and the noun. We opted for the second structure because second terms of comparison allowed by the adverb and placed after the noun, can be formally better dealt with at the same hierarchical level of the adverb that enables them to occur (*la más buena persona del mundo entero*).

³As was observed above with respect to degree adverb modifiers, the boundaries of the syntactic influence of these quantifiers or second degree adverbs can also be discussed. Our choice is again motivated by the fact that the second terms of comparison that these adverbs introduce are placed after the noun together with all its modifiers (*tanto más buen consejero que tú como el mismo demonio*), and formalization is simpler if second terms of comparison are analysed each on a different level, that of the adverb which enables them to occur (see Sections 3.1.4.2.2 and 3.2.3.2.4.2).

especial class of words—that we have considered an especial type of adjectives—, which represent titles given to the person (*san Andrés, sor Citroen*).

The addition of all these modifiers to the noun in the nucleus of the NP is regulated by certain rules that govern the cooccurrence of different types of modifiers, as well as the order in which they appear in the linguistic sequence. With respect to postmodifiers, the following general rules should be considered:

- The addition of each type of postmodifiers (AjP, prepositional complement or relative clause) is independent of the addition of other types of modifiers. That is, each type of modifiers may appear alone or together with another one or two different types of modifiers.
- In case two or three of them appear together, they should appear in the following order: AjP, prepositional complement, relative clause (*un incremento abusivo de los precios que el gobierno quiso frenar por decreto*).
- Prepositional complements lexically linked to the noun—which indicate an especial type of the entity denoted by the nucleus— may be placed immediately after the noun, thus preceding other possible adjectives that could also be modifying the whole NP (*reloj de pared antiguo*).

With respect to premodifiers, the following general rules should be considered:

- Degree modifiers require the previous addition of an adjective premodifier.
- Comparison quantifier modifiers require the previous addition of a comparative degree premodifier.
- Degree modifiers precede adjective modifiers and comparison quantifier modifiers precede degree modifiers in the linguistic sequence.
- Titles are never combined with other modifiers (to a certain extent, they constitute a lexical compound with the proper names that they modify).

Hierarchically, the addition of modifiers is recursive in the following way:

- Premodifiers are more external modifiers than postmodifiers.

- In both cases the items placed further from the nucleus are more external modifiers than those placed nearer.

The graphical representation of *una mucho más vertiginosa subida de los precios que en Brasil*, for example, is showed by Figure 3.1⁴.

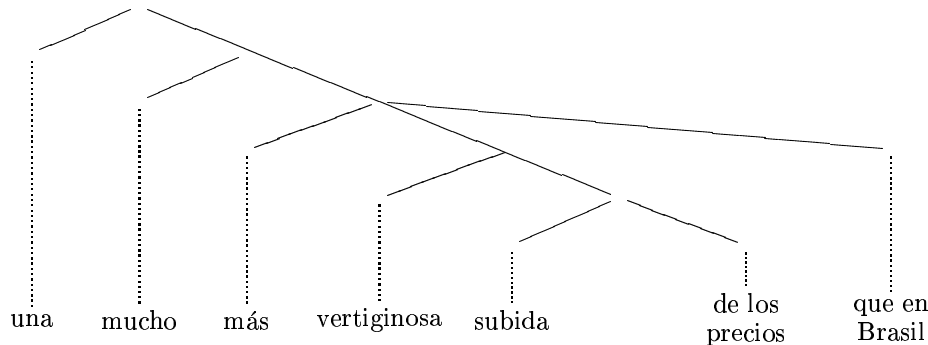


Figure 3.1: *una mucho más vertiginosa subida de los precios que en Brasil*

The structure of *una manifestación evidente de esquizofrenia a la que nunca se le ha prestado la atención adecuada* is represented in Figure 3.2.

We claimed above that this general structure of modifiers will be found, with more or less productivity, in descriptions of phrases different from NPs: PrPs, AvPs and AjPs. As a matter of fact, the claim is, on the one hand, that modifiers are the actual traces of the application of the available mechanisms for expanding linguistic unities, and, on the other, that these mechanisms are, except for the differences of their productivity in different contexts, independent from the categories of the extended linguistic unities. The first example that proves this claim is found within the module of the NP description itself, because some of the premodifiers described above can be also used to expand the noun preceded by certain types of determiners.

The comparison set by comparative determiners *más* and *menos* can be quantified as that set by the corresponding comparative degree adverbs (see above in

⁴The AGFL analysis of this and the following examples can be found in Appendix G.1.

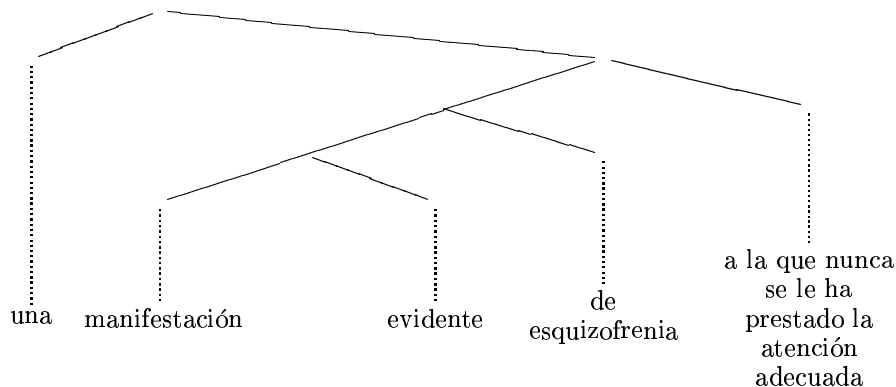


Figure 3.2: *una manifestación evidente de esquizofrenia a la que nunca se le ha prestado la atención adecuada*

this Section). Apart from quantifier determiners (*muchos más libros*), NPs with nucleus *veces* (*mil veces más tiempo que cualquier otro*), can play this role in this context too. The quantifier determiner *poco*, on the other hand, may also be specified by the addition of degree adverbs (*muy pocos libros*, *tan pocos libros de caballerías como en la biblioteca de don Quijote*), including the exclamative degree adverb *qué* (*¡qué poca gracia!*), which enables the NP to introduce exclamative clauses (*¡Qué poca gracia ha tenido eso!*).

Modifiers enabled to occur by comparative determiners *más* and *menos* and quantifier *poco* embrace the NP constituted by the corresponding determiners and noun (including, of course, other possible modifiers), so the structure of, for instance, *muy pocos días de expansión* is that represented in Figure 3.3.

3.1.2.2 Addition of Modifiers in Pronoun Phrase

In PrPs the possibilities are considerably reduced with respect to the addition of modifiers in NPs. Modifiers in PrPs can only be postponed to the nucleus, preceding modifiers are only allowed if the pronoun type in the nucleus can be graded or quantified (see below in this Section).

Adjective modifiers can be postponed to the nucleus of PrPs (*pocos compe-*

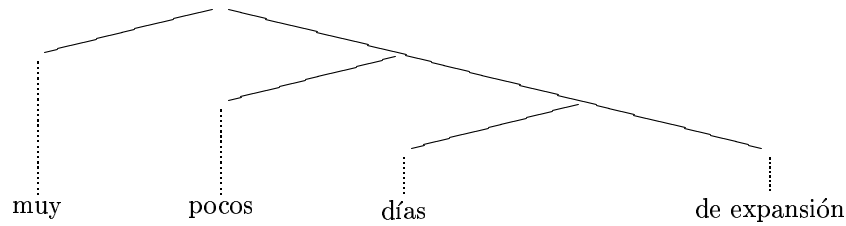


Figure 3.3: *muy pocos días de expansión*

tentes). Only one adjective is accepted in this position, but it can in turn be modified (*pocos verdaderamente competentes*). Such a structure, however, is not frequent, and very often it is debatable whether or not the adjective and the pronoun belong to the same constituent (*Hay pocos competentes*). On the other hand, not all pronoun types can be modified by adjectives, personal pronouns, for instance, cannot.

Partitive prepositional complements can also be postponed to the nucleus of PrPs (*algunos de ellos, dos de los convocados*). Neither in this case can all pronoun types be modified by this kind of complements, personal pronouns, again, cannot.

Relative clauses too can constitute modifiers for PrPs (*otros que habrían podido ayudarnos, nosotros que siempre los servimos fielmente*). As in the case of adjective modifiers, in the case of relative clauses it is also often debatable whether both elements —pronoun and relative clause— belong to the same constituent or to different ones.

Comparative pronouns *más, menos* can also be added a modifying NP with lemma *vez* (*mil veces más que allí*).

Quantifier pronoun *poco* can be specified by degree modifiers (*muy pocos, tan pocos como en cualquier otro sitio*), including the exclamative degree adverb *qué* (*¡qué pocos!*).

The rules that govern the relations between the modifiers and of these with the pronoun nucleus are those described in the previous Section for NPs, with the particularity that concatenation of postmodifiers, if possible, is very rare. The

structure of *muy pocos que puedan resolverlo*, for instance, is that represented in Figure 3.4.

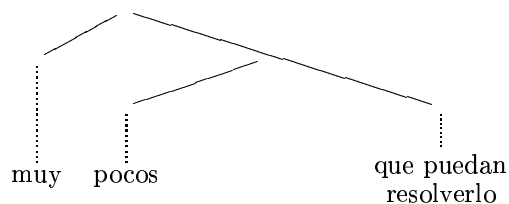


Figure 3.4: *muy pocos que puedan resolverlo*

3.1.2.3 Addition of Modifiers in Adverb Phrase

Nouns and adverbs share also some type of modifiers, but adverbs, more even than pronouns, behave with respect to modifiers as closed categories of words, establishing lexical —or more exactly lexicalized— relations between certain adverbs and certain types of modifiers, especially postmodifiers.

Only pairs of adverbs and modifiers considered more representative will be mentioned here. Even more than in previous descriptions of nouns and pronouns, we want to emphasize that our intention is not to be fully comprehensive, but to show how this heterogeneous type of modifier addition can also be dealt with in a similar way by an AGFL formal description. Adverb types are used to determine which modifiers can be added to which adverbs, constituting an example of highly lexically motivated syntax. As in the case of pronouns and determiners, we have tried to account for the more particular behaviours of adverbs by means of a rich collection of types and, mainly, subtypes.

With respect to postmodifiers, we are not able to identify one of them which can modify all adverbs without exception. Certain adverbs can be modified by the apposition of other adverbs, this is the case of the adverb of time *nunca*, which can be modified by the adverb of time *jamás*, and for demonstrative adverbs (*aquí*, *ahí*, *allí*) which can be modified by other adverbs of place (*aquí arriba*, *ahí dentro*). These accept also the apposition of prepositional phrases (*aquí en España*).

Demonstrative adverbs, as well as the adverb of time *ahora*, can also be mod-

ified by the determiner *mismo* (*aquí mismo, ahora mismo, debajo mismo de su ventana*). In this respect, the pair constituted by the adverb of time *nunca* and the postdeterminer *más* deserves a particular mention, because of its similarity with the postdeterminer *más* when it modifies NPs to give an indication of how many more times something happened or will happen (*Después de la desgracia, ella aún fue por allí tres veces más. Él, en cambio, no regresó nunca más*).

Non-directional adverbs of place accept a prepositional modifier introduced by preposition *de* (*cerca de la gloria, lejos del pecado, dentro de lo razonable, alrededor de la ciudad*).

Adverbs of time *ahora, siempre* can be modified by a relative clause introduced by relative adverb *que* (*Ahora que estás aquí, todo será distinto, Siempre que vengo me trata bien*)⁵. Demonstrative adverbs can also be modified by relative clauses. Lexicalization is very strong in this case too, so that the lexical components of the resulting structure can hardly be substituted by anything else (*Allá donde vayas, te perseguirá su recuerdo, Ahí donde la ves, acaba de licenciarse*).

Regarding premodifiers, adverbs of place signifying movement, *abajo, arriba, adentro, afuera, adelante, atrás*, can be modified by nouns indicating where the movement occurs (*río abajo, escaleras arriba, mar adentro, puerta afuera, carretera adelante, calle atrás*).

Apart from this, addition of premodifiers is more homogeneous than addition of postmodifiers. Degree adverbs can be added to positive adverbs (*muy levemente, más tarde, más cerca del tiovivo*), including exclamative adverb *qué*, which permits the introduction of an exclamative clause (*¡Qué cerca del desastre hemos estado esta vez!, ¡Qué bien lo has hecho!*).

In premodifier position, comparative degree adverbs *más, menos* can in turn be modified by quantifier adverbs, including the relative *cuanto*, which allows for the introduction of relative clauses (*mucho más tarde, un poco más cerca de la solución, cuanto más lejos*). Of course, adverbs inherently marked for comparative degree—including *más, menos* when they are the nucleus of the AvP—can also be modified in this manner (*mucho mejor, poco después, mucho más*). If the quantifier adverb is *poco*, the AvP can still be added a second degree adverb (*muy*

⁵Note that these pairs have been both lexicalized also as conjunctions (*Ahora que, yo no pienso ayudarte, ¿eh?, Os ayudaré siempre que primero lleguéis a un acuerdo*).

poco amablemente, muy poco después, muy poco más tarde)⁶. NPs may also act as comparison quantifiers in AvPs. Semantically, these NPs are strongly dependent on the adverb, because their nucleus must always be the noun of a possible unit of measurement for the dimension indicated by the modified adverb (*tres calles más abajo, varios años antes*).

AvPs already modified by comparative degree adverbs, even when the comparison is further quantified as described in previous paragraph, might still be added quantifying NPs with lemma *vez* (*mil veces mucho más tarde que tú, mil veces más, mil veces mejor*).

The relations between the modifiers and of these with the adverb nucleus are those described in previous Sections for NPs and PrPs. Concatenation of post-modifiers is, as in the case of PrPs, very rare. The structure of *ahora que nos hemos reconciliado*, for instance, is that represented in Figure 3.5.

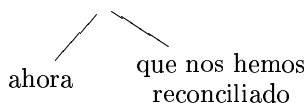


Figure 3.5: *ahora que nos hemos reconciliado*

The structure of *tanto más cerca de la pantalla que nosotros como ellos* is that represented in Figure 3.6.

3.1.2.4 Addition of Modifiers in Adjective Phrase

Adjectives, as nouns, may require prepositional complements as modifiers. The behaviour of adjectives with respect to these complements is very heterogeneous, more even than that of nouns, so it is much more difficult for them to give regular patterns which could be used to encode and automatically process this information. In the current situation of the lexical data available for this work this information was absent, so, as in the case of nouns and prepositional complements, we will

⁶The boundaries of the syntactic influence of this degree or comparison quantifier adverbs are difficult to determine. As in Section 3.1.2.1, our decision here is motivated by formalization, and implies that the first adverb on the left modifies the whole AvP constituted by the adverb in the nucleus, the comparison and comparison quantifier adverbs if present, and the second degree adverb after *poco* if present.

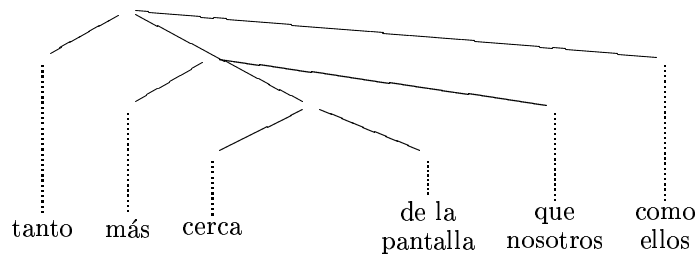


Figure 3.6: *tanto más cerca de la pantalla que nosotros como ellos*

only indicate where in the grammar the availability of these additional resources should be integrated in the future.

The main modifiers of adjectives are degree adverbs. Degree adverbs —the exclamative adverb *qué*, which introduces exclamative clauses, included— precede adjectives that can be graded (*muy guapo*, *tan agradable*, *nada inteligente*, ¡*Qué verde era mi valle!*). In this position, the degree adverb *poco* might still introduce a second level of degree (*muy poco atractivo*, *qué poco delicado*), and the comparative degree adverbs *más*, *menos* can be added quantifier adverbs that measure the comparison (*mucho más guapo*, *un poco más interesante*). Among these, the relative quantifier *cuanto* allows for the introduction of a relative clause (*cuanto más interesante*). The AjP with the quantifier adverb *poco* might be graded in this secondary position too (*muy poco más colaboradora*).

The relations between premodifiers and of these with the adjective nucleus are those described for previous Sections about NPs and AvPs. The structure of *qué poco delicado*, for instance, is that represented in Figure 3.7.

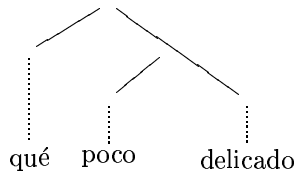


Figure 3.7: *qué poco delicado*

3.1.3 Addition of Determiners

Among other means of identifying the reference of nouns, addition of determiners is the main one. Determiners constitute a closed category of words, closely related to pronouns. Like these, determiner items are also classified according to different types and subtypes, many of which are shared by both determiners and pronouns. In line with this, the exact definition of types and subtypes of determiners is as difficult as that of pronouns, and we find that different linguistic descriptions identify also in this case different sets of determiner types and subtypes. Those considered relevant in this grammar can be found in Section 3.2.2.2, their assignment to lexical items in Appendix F. Interrogative, exclamative and relative determiners enable the phrases in which they participate to introduce the corresponding particular types of clauses (*¿Qué clase de gente asistirá?*, *¡Qué porquería de televisión nos tenemos que tragar día tras día!*, *El caso de cuya instrucción más se ha hablado es sin duda el del GAL*). Comparative determiners (*más*, *menos*, *tanto*) deserve also a especial mention, because they may introduce a second term of comparison or a consequence (*más días que en Roma*, *tanta agua que se inundó hasta la buhardilla*).

Determiners are, obviously, inherently marked for determination. Most of them can only be placed before the item determined, but some of them allow or even need to be postponed (*empresa alguna*, *otro cualquiera*), interrupting in this case the natural sequential relation between the noun or pronoun nucleus and its modifiers (*empresa alguna de cosméticos*). Determiners are inflected for gender and number. Although possessive determiners are also inflected for person, no agreement relationship of the language accounts for this (Nuestros hijos *nunca serán tratados de ese modo*).

In spite of the fact that determiners can be added to both nouns and pronouns, there are not significant differences as to the way in which the addition is performed in each case. Variations are strictly caused by the different nature of the categories to which determiners are added. As pronouns are already marked for determination, determiners preceding them play the role of “over-determiners”, more or less as second level determiners added to nouns—think of pairs like, for instance, *otros dos chicos/otros dos*, *los dos primeros corredores/los dos primeros*.

Because they can be combined with pronouns, which might have neuter gender, determiners, if they have a form for masculine singular, they need to cover also value neuter for attribute gender.

We will first describe the addition of determiners to nouns, and after to pronouns, in this case referring only to some specific particularities with respect to the addition of determiners to nouns.

3.1.3.1 Addition of Determiners in Noun Phrase

Determiners are added to common nouns. Proper nouns are added determiners only rarely, when they are being used as common nouns (Las Carmelas *suelen ser muy mandonas*). The addition of determiners to nouns complies with the following rules:

- One or more determiners can be placed before the noun: *el acuerdo, todos los acuerdos, los demás colaboradores, los otros cuatro candidatos*. Even four determiners appear to be possible in this position: *todos esos otros seis participantes*.
- The order in which determiners are placed before the noun is not free. The possibility that each determiner has of appearing in an NP depends on all the other determiners previously added to the noun in the nucleus. That is, if determiner *dos* has appeared in first position, determiner *los* can appear in second position, *los dos participantes*, but determiner *algunos*, for instance, cannot. If determiners *otros dos* have been previously added to the noun, determiner *los* can still be added in third position, but if the previous sequence of determiners is *esos dos*, it cannot. The concrete sequences of determiners that can be considered acceptable are, however, very difficult to enumerate. This is, in fact, one of the problems in which corpus evidence becomes the best criterion for determining what should be actually allowed by our formal grammar and what should not.
- Certain combinations of determiners occur only either in singular or in plural number (*tantos otros participantes*).

- Not all determiners can close the sequence of determiners added to the noun. *Mismos*, for instance, or *primeros*, cannot close a determiner sequence.
- Certain determiners can close a sequence of determiners only under the condition of having a certain value for number. *Todo hombre*, for instance, is correct, while *todos hombres** is not.
- Some determiners are postponed to the noun, sometimes separating the noun from its modifiers (*libro alguno de Camilo José Cela*, *libro verde alguno*, *dos litros más de agua que de vino*).
- Determiner *el* is required by a noun modified by an adjective and graded by adverbs *más* or *menos* if we want to introduce a general second term of comparison (*el personaje más sorprendente de toda la historia*, *el mejor libro que se ha escrito nunca*).
- Feminine singular nouns beginning by [á-], select determiners *el*, *un*, instead of *la*, *una* (*el área*, *un hada*).

The syntactic structure corresponding to the combination of determiners and nouns is binary and right recursive. That is, hierarchically, in a sequence of determiners, the outermost determiner is the first one, which is placed furthest from the noun. The structure of *todos los demás participantes*, for example, is that represented in Figure 3.8.

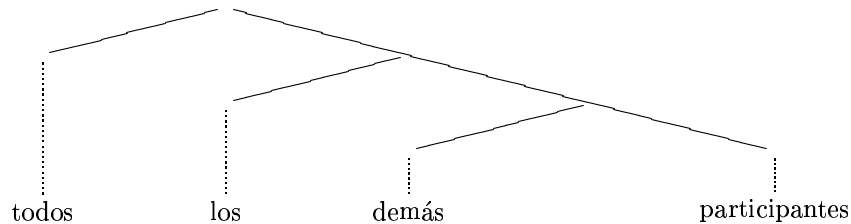


Figure 3.8: *todos los demás participantes*

When comparative determiners introduce a second term of comparison the structure of the NP is, only *apparently*, trinary, since the second term of com-

parison is fully dependent on the determiner, and must in fact disappear if the determiner is removed. The structure of *mucha más importancia de la que pensaba*, for instance, is that represented in Figure 3.9.

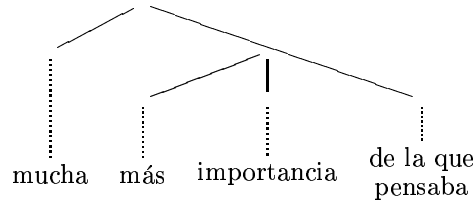


Figure 3.9: *mucha más importancia de la que pensaba*

Postponed determiners are more external determiners than preceding ones, but, contrary to our expectations, those supplying the relevant characteristics not shared by all determiners (type, subtype, determination) to the whole NP are preceding determiners. For example, even if *más* is more external than *cuántos* in *cuántos días más*, the whole NP behaves as a determined interrogative NP.

On the other hand, the interposition of postponed determiners between the nucleus and the modifier causes a problem for graphical representations. We will not establish a best graphical representation for the fact that, in this case too, the direct constitutive relationship between the nucleus and the modifier must be displayed. From a strictly formal point of view, our concern will be simply to account for the whole structure ensuring to the extent that it is possible that the same restrictions between the nucleus and the modifiers will hold in this case and in the case of the rewrite rules for the same structure when nucleus and modifiers are consecutive in the linguistic sequence.

Determiners are the main elements responsible for the identification of NPs as relative, exclamative, interrogative or general NPs, thus enabling these to introduce the corresponding types of clauses (*cuyas buenas intenciones, ¡qué suerte!, ¡qué día?, todo el clan*).

3.1.3.2 Addition of Determiners in Pronoun Phrase

With respect to the addition of determiners to NPs, the addition of determiners to PrPs presents the particularity that the types of the pronouns and determiners involved must be checked for their compatibility.

For pronoun-determiner —or determiner-pronoun— sequences, when both items can be either pronouns or determiners, it is often difficult to determine which item is acting as determiner and which one as pronoun, our decision is taken on the basis of the structures of the corresponding NPs when both items are used as determiners: in *ese otro*, for example, we consider *ese* a determiner and *otro* a pronoun because in the corresponding NP the determiner *otro* is closer to the noun, *ese otro problema*. In *dos más*, on the opposite, we consider *dos* a pronoun and *más* a determiner because in the corresponding NP the determiner *dos* is closer to the noun, *dos días más*.

Pronouns may have inflectional variation capabilities which are not fully realized in all contexts: *lo otro*, *eso otro*, for instance, have value neuter for attribute gender, but no other combination of *otro* with a determiner —not even *otro* alone— is found with value neuter for attribute gender.

All combinations of the items listed in Appendix D and Appendix F are considered PrPs: *lo mío*, for instance, is a PrP and not a nominalization, which is the title given to the structures described in Section 3.1.5.

The syntactic structure that corresponds to the combination of determiners and pronouns is that which corresponds to the combination of determiners and nouns. The structure of *los otros cuatro*, for instance, is that represented in Figure 3.10.

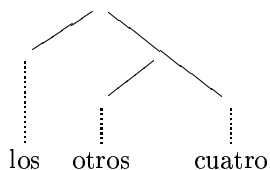


Figure 3.10: *los otros cuatro*

3.1.4 Modifiers

In Section 3.1.1 modifiers were described by their relation to the nucleus of the phrase that they modify. In this Section, on the contrary, modifiers are described by themselves, giving, where necessary, more details about their structures and about the closed categories of items that play the role of modifiers, as well as summarizing all the information concerning the types of nucleus to which different modifiers can be added.

As usual, we will proceed from postmodifiers to premodifiers, from the more internal ones to the more external ones, and from those which can be added to nouns—and other lexical categories—to those which are exclusive of nouns or other lexical categories.

3.1.4.1 Modifiers Postponed to the Nucleus

3.1.4.1.1 Adjective Phrase The structure of AjPs was explained in Section 3.1.2.4. See Section 3.1.4.2.2 for details about degree premodifiers. AjPs are modifiers of both NPs and PrPs.

3.1.4.1.2 Prepositional Modifiers As prepositional phrases are in this proposal considered as particular configurations of NPs, PrPs and AvPs (very rarely AjPs), we can say here that prepositional phrases have indeed already been described in previous Sections 3.1.1, 3.1.3 and 3.1.2. However, not only phrases, but practically all types of categories—although not with the same frequency—can act as prepositional modifiers, see Section 3.1.7.

On the other hand, as data were not available about prepositional requirements of nouns and adjectives, it was impossible to account for these prepositional complements in the proposed formal grammar. Nevertheless, their future integration in the system was foreseen and, so, a preliminary methodology for obtaining and encoding these data⁷ was actually outlined in Section 3.1.2.1, if we combine the piece of information given there with that provided in Chapter 2. On the contrary, within the boundaries of the closed class of adverbs, prepositional modifiers could be accounted for without problems, since the nature of these prepositional

⁷With regard to nouns, since role character of the adjective prepositional complements is less clear.

modifiers is different: first, they are not role based in the same way as nouns or verbs do; second, no more than one prepositional complement can be required by an adverb and, third, it seems possible to handle it quite efficiently if we know simply which preposition is required by the adverb.

Prepositional modifiers can be added to NPs, PrPs, AjPs and AvPs. Prepositional modifiers of PrPs, however, are almost exclusively partitive prepositional modifiers; other types of prepositional modifiers, when they appear in this context, must be related to the entity underlying the pronoun.

3.1.4.1.3 Relative Clause Relative clauses are embedded clauses introduced by relative NPs, PrPs, AvPs or AjPs. The type of relative phrase that introduces the relative clause enables this to act as a modifier, as a functional constituent of a clause (SUBJECT, DIRECT OBJECT, INDIRECT OBJECT, PREPOSITIONAL COMPLEMENT, PREDICATIVE, AGENT, CIRCUMSTANCE), or as the nucleus of a nominalization. In this Section we will restrict ourselves to relative clauses that act as modifiers of phrases —nominalized and argument relative clauses will be respectively described in Sections 3.1.5 and 3.1.6. In the cases in question here, the modified NP —hierarchically precedent— is the antecedent of the relative phrase that introduces the modifying relative clause. Depending on the concrete relative phrase, the antecedent will or will not determine person, gender and number of the relative phrase that introduces the relative clause —which means, of course, that, when the relative phrase is the SUBJECT of the relative clause, the antecedent determines also person and number of the PREDICATE of the relative clause. Not only this, however, because the relative phrases that introduce relative clause modifiers determine also the kind of categories that can be modified by the relative clause, we will specify here which relative phrases can modify NPs (or PrPs), which can modify AvPs, and which can modify both. The relative phrases that introduce relative clause modifiers are the following:

- Noun phrases with determiner *cuyo*. The antecedent of the relative NP and the relative NP itself agree on none of the attributes person, gender or number. Prepositional relative NPs with determiner *cuyo* can also introduce relative clause modifiers with the same characteristics (*la persona cuyo trabajo ya haya sido premiado en alguna ocasión, la persona de cuyo trabajo*

mejor nos han hablado). The relative clauses introduced by these types of relative phrases modify NPs.

- Relative pronoun *que*. The antecedent of the relative pronoun determines person, gender and number of the relative pronoun (*la espía que me amó, el hombre que vino del frío, el hombre que mató a Liberty Valance*). The relative clauses introduced by this type of relative phrases modify NPs.
- Relative pronoun *quien* preceded by preposition. The antecedent of the relative phrase determines person, gender and number of the relative pronoun (*las personas a quienes hayamos ofendido*). The relative clauses introduced by this type of relative phrases modify NPs.
- Relative PrP *el que* preceded by preposition. The antecedent of the relative phrase determines person, gender and number of the relative pronoun (*las personas a las que hayamos ofendido, el asunto del que hablas, las personas a las que os parecía que no era justo*). The relative clauses introduced by this type of relative phrases modify NPs.
- Relative PrP *el cual* preceded by preposition. The antecedent of the relative phrase determines person, gender and number of the relative pronoun (*las personas a las cuales todo les parece negativo, las personas a las cuales nos estamos refiriendo en este caso*). The relative clauses introduced by this type of relative phrases modify NPs.
- Relative adverb *donde*. The antecedent of the relative AvP and the relative AvP itself agree on none of the attributes person, gender or number. The relative adverb *donde* preceded by preposition can also introduce relative clause modifiers with the same characteristics (*el lugar donde nací, el lugar de donde vengo, allá donde te encuentres*). The relative clauses introduced by these types of relative phrases modify both NPs and AvPs.
- Relative adverb *como*. The antecedent of the relative AvP and the relative AvP itself do not agree on either person, gender or number (*el modo como me miras*). The relative clauses introduced by this type of relative phrases modify NPs.

- Relative adverb *que*. The antecedent of the relative AvP can only be one of a limited subset of adverbs (see Section 3.2.3.2.6.2, epigraph **Relative Clause MODIFIER**) that indicate place, manner or time, so the relative clause cannot inflectionally agree with them (*dondequiera que estés, comoquiera que lo haya sabido, ahora que ya somos amigos, siempre que nos vemos*). The relative clauses introduced by this type of relative phrases modify AvPs.

3.1.4.1.4 Apposition Current coverage of appositions has been limited to non-commas appositions in NPs and AvPs. Appositions of NPs cover proper names that specify common nouns, as in *el día D, mi tía Mercedes, el caso GAL*. Appositions of AvPs cover adverbs that specify other adverbs (*nunca jamás, aquí abajo*), or prepositional phrases that specify adverbs (*allí en la playa*, see Sections 3.1.2.1 and 3.1.2.3).

3.1.4.1.5 Determiner Postponed determiners are considered modifiers in AvPs (*aquí mismo, nunca más*, see Section 3.1.2.3).

3.1.4.2 Modifiers Preceding the Nucleus

3.1.4.2.1 Adjective Adjectives have been described in Section 3.1.2.1. They can only modify NPs.

An especial type of these modifiers are words that indicate titles, such as *san, fray, don, etc.*, a closed class of words that have particular requirements with respect to their nucleus (proper names).

3.1.4.2.2 Degree NPs, PrPs, AvPs and AjPs can all be specified for degree. Although comparative and quantifier determiners and pronouns are means for grading NPs and PrPs, we will restrict here the term degree modifier to adverbial premodifiers added to NPs, PrPs, AvPs and AjPs. Even with this restriction, different levels have to be distinguished, because more than one level of adverbs can be present in front of the nucleus of the phrase, each of which involving a different contribution to degree in a general sense.

The possibilities that adverbs have of indicating degree are considerably wide and might reach high levels of detail. There are, in first place, a number of adverbs, derived from adjectives, that can indicate degree (*ligeramente retrasado*, *verdaderamente idiota*, *increíblemente astuto*), or at least modify in precedent position certain types of nucleus. This information, however, is not encoded in the lexicon in its current stage, so at the moment these adverbs can only be analysed as independent constituents. Non-derivative adverbs more frequently indicating degree, on the other hand, are *nada*, *poco*, *algo*, *un poco*, *bastante*, *demasiado*, *harto*, *muy/mucho*, *tan/tanto*, *más*, *menos*, which have all different values for degree (see Section 3.2.2.9 and Appendix E). In addition to these, there is the exclamative adverb *qué* too, which, because it has the capability of introducing exclamative clauses, deserves an especial mention here.

As can be observed in the inventory above, some of the non-derivative adverbs have two different forms, each of which must be used in a different distribution: either in stressed or in unstressed positions in the linguistic discourse. Unstressed degree adverbs *muy*, *tan*, and the remaining degree adverbs with only one form, are used as degree modifiers for NPs with an adjective premodifier (*muy buen estudiante*), for AjPs (*muy independiente*) and for AvPs (*muy tarde*). In these positions, degree adverbs *más*, *menos* allow for the introduction of a second degree adverb that quantifies the comparison introduced by *más*, *menos* by themselves. Stressed degree adverbs *mucho*, *tanto*, and a number of the remaining degree adverbs, perform this function for NPs with an adjective premodifier (*mucho más buen estudiante*), for AjPs (*mucho más independiente*) and for AvPs (*mucho más tarde*). Adjectives and adverbs inherently marked for comparative degree are, of course, directly modified by the second level of degree modification (*mucho mejores*, *mucho más*, *mucho antes*)⁸.

Poco, no matter what lexical category we take into consideration, can be graded. NPs with determiner *poco* and PrPs with pronoun *poco* can be added degree modifiers, including the exclamative *qué* (*qué poca gente*, *muy poca gente*, *qué pocos*, *muy pocos*). AvPs and AjPs including adverb *poco* can still be added

⁸We will henceforth use the term **unstressed degree position** to refer to distributions in which the unstressed forms *muy*, *tan* of adverb lemmas *mucho*, *tanto* are required. Conversely, we will use the term **stressed degree position** for distributions in which the stressed forms *mucho*, *tanto* of adverb lemmas *mucho*, *tanto* appear.

other degree adverbs in a second unstressed degree position (*muy poco detallista, qué poco detallista, muy poco amablemente, qué poco amablemente*). Even when it is quantifying a comparison, in a stressed degree position, the adverb *poco* can still be added another degree adverb in an unstressed degree position (*muy poco más amable, muy poco más tarde*). It is, however, quite difficult to determine the degree values that can be found in this latter —quite infrequent— distribution. As in all decisions concerning second level of formalization, the point is here to create formal structures that can be easily extended if this should be required by a concrete application (Section 3.2.3.2 illustrates our current coverage with respect to this issue).

Comparative degree items *tan/tanto, más, menos*, whether they are determiners, pronouns or adverbs, may introduce a new constituent in the phrase under consideration. This constituent expresses either the second term of the comparison set by the degree item or a consequence of its value. In addition to this, if the comparisons introduced by adverbs *más, menos* are found in the contexts of NPs with determiner *el* or of nominalized AjPs, these comparisons may have either **general** or **particular** character. In the first case, the second term of comparison identifies the absolute situation of the first term of comparison with respect to the whole class of entities having the feature in question. In the second case, it identifies the relative situation of the first term of comparison with respect to a different member of the explicit or implicit class under consideration.

Second terms of comparison have different forms depending both on the item that introduces the comparison and on the type of the comparison itself. *Tan/tanto* introduce either consecutive or comparative —particular— second terms:

- On the one hand, consecutive second terms are constituted by *that*-clauses (*Acoge a sus invitados con tanta solicitud que acaba por causarles incomodidad, Sabe tanto de casi todo que acompleja, Se cuida tanto que dudo mucho que le quede tiempo para ninguna otra cosa, Lo hizo tan bien que nadie sospechó de ella, Actúa de manera tan cauta que me infunde sospechas*).
- On the other, comparative particular second terms are constituted by *como* followed by NPs, PrPs, AvPs, AjPs or nominalizations (*Ella tiene tantos achaques como tú, De eso no tengo tanto como él, Ella se cuida tanto como*

tú, *Ella puede hacerlo tan bien como su hermano, Es tan guapo como el otro*).

Más, menos introduce comparative particular second terms with different structures depending on the type of the phrase that contains the comparative item.

- All kind of phrases might introduce a second term of comparison constituted by *que* followed by NPs, PrPs, AvPs, AjPs or nominalizations (*Rivaldo marca más goles que Raúl, Rivaldo marcó más que Raúl, Rivaldo me gusta más que Raúl, Rivaldo actuó más deprisa que Raúl, Rivaldo es más rápido que Raúl*).
- All kind of phrases might introduce second terms of comparison constituted by nominalizations of, mainly —some participle clauses rarely appear in this context too—, relative clauses (see Section 3.1.5) introduced by preposition *de*. For AjPs and AvPs, gender of these nominalizations is neuter (*Se cuidaba más de lo que pensaba, Llegó más tarde de lo que esperaba, Es más largo de lo que pensaba*), for NPs and PrPs, these nominalizations agree in gender and number with the nucleus of the phrase (*Había más personas de las que esperábamos, Había más de los que creíamos*) —it is difficult to determine whether NPs accept also neuter nominalizations (*más personas de lo que esperábamos*).
- Noun phrases introduced by preposition *de* constitute also a possibility for particular second terms of comparison associated with a limited set of AvPs (*más allá de la cúpula del trueno, después de su reaparición*).

Más, menos also introduce comparative general second terms which might present different structures.

- Phrases introduced by preposition *de* (*el personaje más siniestro de la historia, el mejor futbolista de Europa, el más atractivo de todos*).
- Nominalizations, mainly of relative clauses —although participles clauses rarely appear in this context too—, introduced by preposition *de* (*el mejor futbolista de los que he conocido, la mejor de las que he visto*). These nominalizations agree for gender with the nucleus.

- Relative clauses introduced by relative PrPs that agree for gender, number and third person singular or plural with the nucleus of the nominalization (*el mejor profesor que he tenido, el mejor que he visto*). It is difficult, however, to determine whether relative clauses introduced by prepositional PrPs *el que* are also accepted in this position (*el mejor del que he tenido noticia*).

3.1.4.2.3 Noun Phrase Various NPs can be used as premodifiers for different phrase categories and performing different functions.

The more external premodifier is the NP with lemma *vez*, which quantifies how many times something is more or a lot more than something else. This is quite familiar and usually contains a determiner that expresses the number of times that something is more than something else (*mil veces más tiempo que yo, mil veces más buena persona que tú, De ese tipo de armamento, los americanos tienen mil veces más que nosotros, mil veces más detallista que él, mil veces más tarde*). This type of NP can be added to almost any structure containing a comparison or quantification, whatever the phrase category under consideration.

Certain comparative AvPs can be quantified by NPs whose nucleus is a noun lemma that can be used as a unit of measurement for the dimension indicated by the adverb. These NPs should either be introduced by a quantifier determiner or have plural number (*dos meses después, meses después*).

Singular undetermined and unmodified NPs, nouns, are also used as premodifiers when specifying where the movement implied by the meaning of some adverbs occurs (*río abajo*).

3.1.5 Nominalization

Nominalizations are constituted by determiner *el* (sometimes preceded by determiner *todo*) followed by AjPs, prepositional phrases or relative clauses introduced by either pronoun *que* or prepositional PrPs with *el que* (see below in this Section).

Nominalized AjPs can be modified by *más, menos* comparative degree adverbs—it is difficult to determine whether other degree values are also acceptable in this context—, yielding general comparisons—it is also difficult to determine whether particular comparisons too are acceptable here: in this respect, in fact, we find that when the second term of comparison is not present, the interpretation of the

type of the comparison appears to be always the general one— (*el más travieso de todos, los más antiguos que he visto, la más linda*).

Prepositional NPs, PrPs and nominalizations preceded by determiner *el* constitute nominalizations as those found in the examples *el de su madre, el de ella, el de la de la esquina*. Prepositional clauses, whether infinitive or *that*-clauses, constitute also nominalizations, always preceded by definite articles in neuter gender (*lo, lo de que no te guste, lo de vestirse de largo*).

With respect to nominalizations of relative clauses:

- *Que* relative clauses preceded by determiner *el* constitute nominalizations frequently used in linguistic discourse (*el que vino ayer, los que no estén seguros*).
- It is, however, more difficult to determine whether relative clauses introduced by prepositional PrP *el que* can also be nominalized, that is, whether examples such as *A los que se lo hemos propuesto han dicho ya que asistirán*, are acceptable or not (see also Sections 3.1.6 and 3.1.7).

Addition of modifiers in nominalizations is problematic, examples as *Los que vimos ayer que no nos gustaron*, with a relative clause modifier, seem to be the only ones possible, although they are quite rare.

3.1.6 Other constituents of the Clause

Phrases are not the only syntactic structures that can play syntactic roles at the level of the clause. Recursion is indeed a property of language, and clauses themselves, conveniently contextualized, can play these roles too. In this Section we give a preliminary description of how clauses are contextualized in order to be used as clause constituents that perform syntactic functions such as SUBJECT, DIRECT OBJECT, etc.

The resources⁹ used to contextualize clauses as clause constituents have been traditionally considered as those determining the appearance of the so-called clause

⁹We do not intend at all to be exhaustive with respect to these resources here. On the contrary, our intention is simply to account for the more frequent and generally accepted structures and connectors. However, although this work exceeds the aims of this book and thus has not been covered by this research, it is obvious that an individual and extensive evaluation of the items (single or multi-word) used as clause connectors would be necessary to define clearly the type of the connections —and at which levels— that can be established between clauses in each language.

types, which in Spanish are the following:

- *That*-clauses: clauses introduced by conjunction *que* (*A nadie le gusta que le pongan defectos*), sometimes preceded by determiner *el* (*El que no quieras venir me deja absolutamente indiferente*).
- Infinitive clauses: clauses in which an infinitive verb form plays the role of **PREDICATE** (*A todos nos tranquiliza descubrir defectos en los demás*), sometimes preceded by determiner *el* (*El no haber sido invitado me deja absolutamente indiferente*).
- Participle clauses: clauses in which a participle verb form plays the role of **PREDICATE** (*La perseguía sin duda cegado por el odio que sentía por ella*).
- Indirect yes/no questions: clauses introduced by conjunction *si* followed by either finite or non-finite verb forms (*No sé si quiere hacerlo, No sé si hacerlo*).
- Indirect Wh-questions: clauses introduced by interrogative NPs, PrPs or AvPs followed by either finite or non-finite verb forms (*No sé quién lo hizo, Nunca sé qué hacer en esos casos*).
- Argument relative clauses: clauses introduced by relative NPs, PrPs, AvPs or AjPs¹⁰, only rarely combined with infinitive verb forms. These relative phrases are not all of them equally capable of introducing relative clauses that can play any syntactic role in the main clause. This is, however, something very difficult to determine and decisions should be taken only on the basis of a representative corpus. Relative connectors are the following:

- NPs with relative determiner *cuanto* (*Corren peligro cuantas personas sepan la verdad*). In this case we have the same problem pointed out

¹⁰ We do not include in this group, because their syntactic influence extends beyond the level of the clause, relative clauses introduced by determiner or adverb *cuanto* when these precede, respectively, one of the determiners or degree adverbs *más*, *menos*. This means that, in the current stage of development of the grammar, both the relative constituent and the relative clause introduced by *cuanto* in this context can be analysed, what still has to be solved—even simply from the descriptive point of view—is the further contextualization of these examples in the linguistic discourse.

in previous Section with respect to prepositional phrases with pronoun *que* when these introduce nominalizations: it is difficult to determine whether examples as *A cuantos se lo hemos propuesto han asentido* are acceptable, that is, whether relative clauses that start with prepositions strictly required within the limits of the relative clause itself, can perform in the main clause syntactic functions which usually reject prepositional syntagms. This problem exists also in the cases of the relative connectors referred to in the next two paragraphs.

- Relative pronoun *cuanto*, sometimes preceded by determiner *todo*, (*Cuantos están presentes estarían dispuestos a testificar*).
- Relative pronoun *quien* (*Quien mal anda, mal acaba*).
- Relative adverb *donde* (*Siéntate donde quieras*).
- Relative adverb *cuando* (*Ven cuando se haya puesto el sol*).
- Relative adverb *como* (*Hágalo como se le ha ordenado*).

3.1.7 Prepositional Constituents

Structures of prepositional NPs, PrPs and nominalizations of AjPs or prepositional phrases do not need additional explanations. The only particularity that has to be taken into account with respect to these is the appearance of the forms *al*, *del*, which merge prepositions and determiners. Prepositional AjPs are very rare, found only in a short number of certain restricted contexts, such as *Los toman por tontos*.

Neither prepositional AvPs need additional explanations. They are not the only phrases that can be added certain modifiers, but this is for them more frequent and natural (*mucho más hacia allá, muy hacia dentro*).

Except for participle clauses, clause constituents of main clauses might also be introduced by prepositions (*Se olvidó de que se lo había pedido, Me obligó a seguirlo, No me acuerdo de si lo he hecho yo o no, No consigo acordarme de quién me lo dijo*). As for prepositional nominalizations of relative clauses and prepositional argument relative clauses we want to remark the following:

- In first place, it is worthwhile insisting on the fact that both nominalizations of relative clauses and argument relative clauses might be introduced by

prepositions strictly required within the limits of the relative clause itself, and that, in this case, the whole syntagm does not constitute a prepositional nominalization or a prepositional argument relative clause (*A los que se lo propusimos asintieron, A cuantos se les propuso lo aceptaron*).

- Prepositional nominalizations of relative clauses or prepositional argument relative clauses are instead constituted by prepositions required within the level of the main clause followed by nominalizations of *que* relative clauses or argument relative clauses (*Nos dirigimos a los que/quienes puedan estar interesados*). Sometimes, these nominalizations or argument relative clauses might also include a different preposition required within the boundaries of their direct syntactic influence (*Nunca nos olvidamos de a los que/quienes les debemos algo*).
- It is also possible that only one preposition, required both at the level of the main clause and within the relative one, is present and performs syntactic functions simultaneously at both levels, the main and the embedded clause (*Amamos a los que/quienes podemos amar*).

3.2 AGFL Formalization of Noun, Pronoun, Adverb and Adjective Phrases

We undertake in this Section the specific description of the AGFL formalization of all the linguistic facts that were described in the previous one. In Section 3.2.1 we describe the structure of the lexical data that account for nouns, pronouns, adverbs, adjectives, determiners and other lexical items related to NPs, PrPs, AvPs and AjPs—in formal terms, lexicon non-terminals corresponding to these items are introduced and associated with the adequate lexicon terminals. Subsequently, the features and feature values that play a role as parameters for these lexicon non-terminals, as well as those that play a role for the non-terminals of the grammar described in the following Section,—affix non-terminals and terminals—are specifically addressed in Section 3.2.2. Finally, the AGFL syntax rules that describe the syntactic structures of NPs, PrPs, AvPs, AjPs, nominalizations and other clause constituents—syntax rules, grammar non-terminals and terminals—

are accounted for in Section 3.2.3.

3.2.1 The Lexicon

The description of the lexicon non-terminals for nouns, pronouns, adverbs, adjectives, determiners, prepositions and conjunctions is quite simple. To begin with, there is only one lexicon non-terminal that accounts for all noun wordforms:

```
NounSt ( ntype, gender, number )
```

Noun wordforms are specified for noun type, gender and number. Values for *ntype* will be specified in Section 3.2.2.1. Values for *gender* and *number* were already specified in Section 2.2.2.4, but they will be further dealt with in Section 3.2.2.10. Examples of lexicon definitions are the following:

```
NounSt ( COMMON, FEM, SING ): "alumna".
NounSt ( COMMON, FEM, PLU ): "alumnas".
NounSt ( COMMON, MASC, SING ): "alumno".
NounSt ( COMMON, MASC, PLU ): "alumnos".
```

Lexicon non-terminals for pronoun wordforms are the following:

```
PronounSt ( prtype, prsubtype, definiteness, gender, number )
PronounSt ( prtype, prsubtype, definiteness, inherent person,
           gender, number )
PronounSt ( prtype, prsubtype, definiteness, person, gender,
           number, case )
```

The first describes all pronoun types except possessive and personal pronouns, the second describes possessive pronouns (*mío, etc.*), and the third non-clitic personal pronouns (*yo, etc.*). Some personal pronouns are merged with preposition *con*, yielding forms *conmigo, contigo, consigo*. For them an additional pronoun lexicon definition has been designed:

```
Prepositional PronounSt ( preposition, prtype, prsubtype,
                        definiteness, person, gender, number, case )
```

All pronouns are specified for pronoun type, pronoun subtype, definiteness, gender and number. Possessive pronouns are also specified for person—inherent person—and personal pronouns are specified for person and case. Prepositional items add a specification for preposition type, which obviously contains the value

of the preposition merged with the item. Values for *prtype* and *prsubtype* will be described in Section 3.2.2.2. Values for *definiteness* will be described in Section 3.2.2.4. Values for *person*, *gender*, *number* and *case* were already specified in Section 2.2.2.4, but they will be further explained in Section 3.2.2.10. Values for *preposition* were described in Section 2.2.2.5. A catalogue of pronoun lexicon definitions can be found in Appendix D.

Lexicon non-terminals for adverb wordforms are the following:

```
AdvSt ( avtype, avsubtype, preptype, degree, AVposition )
AdvSt ( degree )
```

The first describes the closed collection of non-derivative adverbs, the second describes adverbs derived from adjectives. Some adverbs are merged with prepositions, yielding prepositional forms as *adonde*. For them an additional lexicon definition has also been designed:

```
Prepositional AdvSt ( preposition, avtype, avsubtype, preptype,
                    degree, AVposition )
```

Non-derivative adverbs are all specified for adverb type and adverb subtype, preposition required for a possible prepositional modifier, degree and adverb position. Adverbs derived from adjectives are only specified for degree. Values for *avtype* and *avsubtype* will be described in Section 3.2.2.5. Values for *degree* will be described in Section 3.2.2.9. Values for *AVposition* will be described in Section 3.2.2.6. Values for *preposition* were described in Section 2.2.2.5. A catalogue of lexicon definitions for non-derivative adverbs can be found in Appendix E. Some examples of lexicon definitions for adverbs derived from adjectives are the following:

```
AdvSt ( POS ): "abruptamente".
AdvSt ( POS ): "absurdamente".
```

Lexicon non-terminals for adjective wordforms are the following:

```
AdjSt ( adjtype, position, degree, gender, number )
```

All adjectives are specified for adjective type, position with respect to the nucleus, degree, gender and number. Values for *adjtype* will be explained in Section 3.2.2.7. Values for *degree* will be explained in Section 3.2.2.9. Values for

gender and *number* were already introduced in Section 2.2.2.4, but they will be further dealt with in Section 3.2.2.10. Examples of adjective lexicon definitions are the following:

```
AdjSt ( QUAL, position, POS, FEM, SING ): "abrupta".
AdjSt ( QUAL, position, POS, FEM, PLU ): "abruptas".
AdjSt ( QUAL, position, POS, MASC|NEUT, SING ): "abrupto".
AdjSt ( QUAL, position, POS, MASC, PLU ): "abruptos".
```

Lexicon non-terminals for determiner wordforms are the following:

```
DeterminerSt ( dttype, dtsubtype, position, definiteness,
  inherent person, gender, number )
DeterminerSt ( dttype, dtsubtype, position, definiteness,
  gender, number )
```

The first describes possessive determiners (*mi*, etc.), the second describes all determiners except possessive determiners. As article *el* can be merged with prepositions *a*, *de*, yielding forms *al*, *del*, for these forms an additional lexicon definition had to be designed:

```
Prepositional DeterminerSt ( preptype, dttype, dtsubtype, position,
  definiteness, gender, number )
```

All determiners are specified for determiner type and determiner subtype, position with respect to the nucleus, definiteness, gender and number. Possessive determiners are also specified for person—inherent person. Values for *dttype* and *dtsubtype* will be explained in Section 3.2.2.2. Values for *definiteness* and *position* will be explained in Sections 3.2.2.4 and 3.2.2.8. Values for *person*, *gender* and *number* were already specified in Section 2.2.2.4, but they will be further explained in Section 3.2.2.10. Values for *preposition* were described in Section 2.2.2.5. A catalogue of determiner lexicon definitions can be found in Appendix F.

Lexicon non-terminals for preposition and conjunction wordforms are the following:

```
PrepositionSt ( preptype )
ConjunctionSt ( cjtype )
```

Prepositions are specified for preposition type and conjunctions are specified for conjunction type. Currently accepted values for *preptype* and *cjtype* were explained in Section 2.2.2.5.

3.2.2 Affix Rules, Terminals and Non-terminals

For each feature considered in the grammar and related to NPs, PrPs, AvPs, AjPs, nominalizations or other constituents of the clause, we will in this Section deal in detail with their possible values, structure and main functionality. Some features were also relevant for VPs, and therefore they have already been introduced in Chapter 2, for these we will only include additional explanations in case additional particularities require them.

3.2.2.1 Noun Type

Affix rule for *ntype* is the following:

```
## NOUN TYPE

ntype :: COMMON; proper.
proper :: Alphabetic; Numeric; Alphanumeric.
```

Value **COMMON** is assigned to common nouns. Proper nouns are assigned values in the domain of non-terminal *proper* according to their formal characteristics. Currently, proper nouns are not contained in the lexicon, so they are exclusively identified by means of regular expression rules included in the grammar: these rules assign value **Alphabetic** to proper nouns constituted by letters, value **Numeric** to proper nouns constituted by numbers, and value **Alphanumeric** to proper nouns constituted by both letters and numbers. Assuming certain risks, this information is after used in higher levels of the grammar to restrict the functional possibilities of each type of proper noun.

3.2.2.2 Pronoun and Determiner Type and Subtype

Affix rules for *prtype* and *dttype* are the following:

```
## PRONOUN TYPE

prtype :: prgeneral; prno_general.
prno_general :: no_general.
prgeneral :: COMPTVE; prno_COMPTVE.
prno_COMPTVE :: PERS; no_comptve.

## DETERMINER TYPE

I_dttype :: dttype.
```

```

dttype :: dtgeneral; dtno_general.
dtno_general :: no_general.
dtgeneral :: COMPTVE; dtno_COMPTVE.
dtno_COMPTVE :: ART; no_comptve.

no_general :: no_declarative; REL.
no_declarative :: INTG; EXCL.
no_comptve :: GLOBAL; DEM; POSS; QUANT; DIS; INDEF; REFL; CARD;
            ORD; PART.

```

Since there are strong relations between the two categories of pronouns and determiners, most of the types identified for each of them are shared by both. The names assigned to the affix terminals for the identification of types of pronouns and determiners are in general self-explanatory, most of them have in fact been traditionally used in language descriptions. Besides these, however, we introduce certain names which will probably be unknown to the linguistic community, of these we will give an example together with a preliminary proposal for a corresponding term: **REL**, *relative*, **INTG**, *interrogative*, **EXCL**, *exclamative*, **COMPTVE**, *comparative*, **GLOBAL**, *global (todo)*, **DEM**, *demonstrative*, **POSS**, *possessive*, **QUANT**, *quantifier*, **DIS**, *distributive (cada)*, **INDEF**, *indefinite*, **REFL**, *reflexive (mismo)*, **CARD**, *cardinal*, **ORD**, *ordinal*, **PART**, *partitive*. There are only two types in this catalogue which are exclusive of, respectively, pronouns and determiners: **PERS**, *personal*, and **ART**, *article*. The association of types with pronoun and determiner items —we have tried to be exhaustive with respect to both of them— can be found in Appendixes D and F.

As for the nested structure of affix non-terminals *prtype* and *dttype*, it is motivated by the following two reasons:

- Having always the possibility of referring to types either of pronouns or of determiners, on the one hand, or to both of them, on the other, gives us maximum of flexibility and of economy of terminal values. For this reason, while non-terminals beginning by *pr-* or *dt-* have the power to refer, respectively, to types of only pronouns and to types of only determiners, non-terminals not beginning by *pr-* or *dt-* —found in the domain of *pr-* and *dt-* non-terminals— have, on the contrary, the power to refer to types of pronouns and determiners simultaneously, ensuring as well that we do not have to define terminals for shared types twice in the grammar —one in the domain of *prtype* and

another in that of *dttype*¹¹.

- Now, if, just for a while, we do not pay attention to the prefixes *pr-* and *dt-*, we can see also that type non-terminals are nested so as to separate groups of terminal values which are somewhat relevant for the grammar. Non-terminals *-general/-no_general*, to begin with, separate terminal values for pronoun and determiner types that introduce especial types of clauses —interrogative, exclamative and relative ones— from terminal values for types that cannot do that¹². Non-terminal *-no_comptve* subsequently separates **COMPTVE** pronouns and determiners, which can introduce a second term of comparison, from other general types of pronouns, which cannot. Non-terminal *no_declarative*, finally, separates **INTG** and **EXCL** affix terminals, which identify phrases that introduce clauses in interrogative and exclamative moods, from **REL** affix terminal, which identifies phrases that introduce relative clauses in declarative mood.

Affix non-terminal *Ldttype*, which dominates *dttype* affix non-terminal, is used by certain rules of the grammar to distinguish two different levels of determiners (see Sections 3.2.2.3, 3.2.3.3.4 and 3.2.3.3.5¹³).

Affix rules for *prsubtype* and *dtsubtype* are the following:

```
## PRONOUN SUBTYPE
```

```
prsubtype :: prno_declarative; prrel; ESTE; prquant; prnum;
           prcomptve; prindef; NIL.
prno_declarative :: QUE; QUIEN; CUAL; CUANTO.
prrel :: QUE; QUIEN; CUAL; CUANTO.
prquant :: quant.
prnum :: num.
prcomptve :: comp.
```

¹¹There is only one exception to this rule. When, while distinguishing groups of types, a group includes just one shared type, the terminal value for this type appears repeated in both the *prtype* and the *dttype* domains. This is indeed the case for the affix terminal **COMPTVE**: if we would nest a *comptve* non-terminal with just one shared type **COMPTVE**, we would see that in all the analyses the non-terminal symbol would appear instead of the terminal value. Thus, just for *aesthetic* reasons —and this will be always our choice in similar situations—, we prefer to have in this case duplicate values for the same type.

¹²Note that non-terminal *no_general*, shared by *prno_general* and *dtno_general*, will be found in the domain of other non-terminals, that we describe in next sections, as well.

¹³The function in the grammar of *L-* affix non-terminals can be compared to that of *-one* and *-two* affix non-terminals, found in the domains of *gender*, *number*, *person* and *case* affix non-terminals, see Sections 2.2.2.4 and 2.2.3.1.

```

prindef :: UNO; QUIENQUIERA; ALGUIEN; indef.

## DETERMINER SUBTYPE

I_dtsubtype :: dtsubtype.
dtsubtype :: dstart; dtno_declarative; dtrel; dttem; dtquant; dtnum;
            dtcomptve; dtindef; NIL.
dstart :: EL; UNO.
dtno_declarative :: QUE; CUANTO.
dtrel :: CUYO; CUAL; CUANTO.
dttem :: TAL; ESTE.
dtquant :: SENDOS; quant.
dtnum :: num.
dtcomptve :: comp.
dtindef :: CIERTO; indef.

quant :: AMBOS; VARIOS; POCO; MUCHO; BASTANTE; DEMASIADO.
num :: card; no_card.
card :: UNO; no_UNO.
no_UNO :: NoUNIT; nounit+uno.
nounit+uno :: NoUNIT+UNO; NoUNIT_0-A+UNO.
no_card :: MEDIO; CUARTO.
indef :: OTRO; ALGUNO; CUALQUIERA; DEMAS.

```

The identification of subtypes of pronouns and determiners is very detailed. This detail is a requirement of the function attributed to the subtype feature in the grammar, since this feature is used to distinguish each minimum group of pronouns and determiners with exactly the same contextualization capabilities¹⁴. For most of the groups we have selected one of the items —or the only one— as the representative of the group, the name of the subtype terminal value assigned to the group is the capitalized lemma of the selected representative. Almost all names are, then, self-explanatory, their association with different items and, consequently, the groups of items identified by the same subtype, can be found in Appendixes D and F. The following explanations, however, can be of benefit:

NoUNIT Value assigned to simple cardinal pronouns and determiners different from *uno/un*, and to compound cardinal pronouns and determiners whose last component is different from *un/uno*: *dos*, *treinta y dos*.

NoUNIT+UNO Value assigned to compound cardinal pronouns and determiners whose last component is *un/uno*. They do not include any other

¹⁴As a side-effect of this goal, a number of groups are constituted by just one pronoun or determiner item.

component that can indicate *gender* of the whole pronoun or determiner: *veintiún/veintiuno*.

NoUNIT_O-A+UNO Value assigned to compound cardinal pronouns and determiners whose last component is *un/uno*. They include at least one more element that indicates *gender* of the whole pronoun or determiner: *doscientos un/doscientos uno*.

NIL Value assigned to pronouns and determiners whose value for *prtype* or *dttype* is sufficient to identify a group of pronouns or determiners with similar contextualization capabilities. That is, this is the value assigned when *subtype* does not exist or is the same as *type*.

As to the nested structure of affix non-terminals *prsubtype* and *dtsubtype*, this is due to the following reasons:

- Internal control and readability. Each type containing different groups of contextualization preferences is represented by a different subtype non-terminal whose name, if we eliminate the *dt-/pr-* prefixes, results from decapitalization of the terminal value assigned to the type. This correspondence between type affix terminals and subtype affix non-terminals is obviously useful for internal control of the grammar, but it must be clear that only the grammar and the lexicon as designed by the linguist ensure that type **QUANT**, for instance, will always be associated with subtypes in the domain of *prquant* or *dtquant* (which both immediately dominate *quant* affix non-terminal). In other words, there is nothing whatsoever in the affix structure as such—not even the correspondence between type affix terminals and subtype affix non-terminals—which by itself forces a necessary relation between certain types and certain subtypes.
- Economy of terminal values. Terminal values which represent the same contextualization capabilities of the NPs and PrPs containing the corresponding determiners and pronouns constitute the domain of one and the same affix non-terminal, which is in turn dominated by the corresponding specific pronoun and determiner affix non-terminals: this is the case, for instance, for

non-terminals *prquant*, *dtquant* and *quant*¹⁵.

- Relevance for the grammar. This is the case for the affix structure that accounts for **NoUNIT**, **NoUNIT+UNO** and **NoUNIT-O-A+UNO** affix terminals, which is necessary because these affix terminals have to be combined with nouns beginning by [á] under different conditions, see Section 3.2.3.3.4.

As regards the affix non-terminal *comp*, as it is also found in the domain of affix non-terminal *degree*, it will be explained in detail in Section 3.2.2.9 dealing with degree. At this point it is sufficient if we note that, by means of the values found in the domain of *comp*, which are shared by pronouns, determiners and adverbs, we can ensure in the grammar that comparative pronouns and determiners will require the same type of second terms of comparison required by comparative adverbs, even if pronouns and determiners, unlike adverbs and second terms of comparison, are not strictly specified for affix non-terminal *degree* (see Sections 3.2.2.9, 3.2.3.2.5 and 3.2.3.3.4).

Affix non-terminal *Ldsubtype*, as affix non-terminal *Ldttype* explained above, is used by certain rules of the grammar to distinguish different levels of determiners (see details in Sections 3.2.2.3, 3.2.3.3.4 and 3.2.3.3.5).

3.2.2.3 Determiner Level

Affix rule for *level* is the following:

```
## DETERMINER LEVEL
```

```
level :: ONE; TWO; THREE; FOUR.
```

Values are obvious. These values are used in the grammar to keep track of the recursive addition of determiners in order to restrict their combinations at

¹⁵On the contrary, terminal values which represent different contextualization capabilities of the NPs and PrPs containing the corresponding determiners and pronouns are exclusively found under the domain either of *dt-* or of *pr-* subtype affix non-terminals. These terminal values might even have the same name if they correspond to the same lexical item. An example of this situation is affix terminal **QUE**, which in the domain of *prrel* is the value associated with the neuter pronoun *qué*, and in the domain of *dtrel* is the value associated with the determiner *qué*. The point is that pronoun *qué* is never a substitute for an NP introduced by determiner *qué*, and so, as PrPs and NPs based on the item *qué* have not the same contextualization capabilities, the corresponding pronouns and determiners must be identified by means of different, although with the same form, subtype affix non-terminals.

each level of recursion. The syntax rules that account for the recursive addition of determiners are showed in Section 3.2.3.3.

3.2.2.4 Definiteness

Affix rule for *definiteness* is the following:

```
## DEFINITENESS  
  
definiteness :: DEFINITE; UNDEFINITE.
```

Values are self-explanatory. Their association with *pronoun* and *determiner* items in the lexicon can be found in Appendixes D and F. Common and proper nouns are assigned values for *definiteness* within the grammar.

3.2.2.5 Adverb Type and Subtype

Affix rule for *avtype* is the following:

```
## ADVERB TYPE  
  
avtype :: avgeneral; avno_general.  
avno_general :: no_general.  
avgeneral :: POL; avno_discoursive.  
avno_discoursive :: COMPTVE; avno_comptve.  
avno_comptve :: avdimensions.  
avdimensions :: PLACE; MANNER; TIME; QUANT.
```

Traditional types of adverbs are distinguished, together with the type **POL**—which stands for *polarity*—, assigned to affirmative or negative adverbs required at the level of discourse analysis. The rest of terminal values are self-explanatory. Their association with the closed collection of adverb lexicon items can be found in Appendix E. The nested structure of affix non-terminal *avtype* is similar to that of pronouns and determiners. The affix terminals for adverb types show also certain similarities with the affix terminals for pronoun and determiner types. The fact that all of them share the types found in the domain of non-terminal *no_general*, as well as the affix terminal **COMPTVE**, is particularly relevant for the grammar, since these affix terminals are ultimately responsible both for the identification of especial types of clauses, such as interrogative, exclamative, etc., as well as for the introduction of second terms of comparison.

Affix rule for *avsubtype* is the following:

```

## ADVERB SUBTYPE

avsubtype :: avrel; avexcl; avcomptve; avquant; avplace; avtime;
    avmanner; NIL.
avrel :: DONDE; COMO; CUANDO; CUANTO; MIENTRAS.
avno_declarative :: DONDE; COMO; CUANDO; CUANTO.
avexcl :: avno_declarative; QUE.
avcomptve :: comp.
avquant :: POCO; BASTANTE; DEMASIADO; HARTO; MUCHO; ALGO; CASI; NADA.
avplace :: AQUI; ACA; DETRAS; ATRAS; DONDEQUIERA.
avtime :: AHORA; SIEMPRE; MIENTRAS; NUNCA; JAMAS.
avmanner :: ASI; COMOQUIERA.

```

Subtypes of adverbs, as pronoun and determiner subtypes do, try to group together adverbs with the same contextualization capabilities. Likewise, the names assigned to the non-terminals that represent groups of adverbs with a similar syntactic behaviour are in this case too obtained from capitalized lemmas of the adverb items arbitrary selected as representatives of the groups identified. All terminal values are, then, self-explanatory, their association with adverb items in the lexicon, as well as the groups of items that constitute each subtype can be found in Appendix E.

With respect to value **NIL**, however, we have to note that it is assigned to adverb items under slightly different conditions with respect to the assignment of the same value to pronoun and determiner items. Within the domain of *avsubtype*, in fact, value **NIL** is assigned to all the remaining items which should be assigned the same subtype, once all the other particular behaviours within the whole type group have been conveniently identified by means of subtypes. That is, once, for instance, time adverb items with particular behaviours have been assigned subtypes **AHORA**, **SIEMPRE**, **MIENTRAS**, **NUNCA** or **JAMAS**, the remaining ones, for which no particular contextualization capabilities have been identified, are assigned value **NIL**¹⁶.

As regards the nested structure of *avsubtype*, finally, this responds to the same principles that were adduced to justify the nested structures of affix non-terminals *prsubtype* and *dtsubtype*.

¹⁶Of course, both in the case of pronouns and determiners and in the case of adverbs, value **NIL** is also a means to keep lower the number of terminal values that can be distinguished within the large domain of subtypes (remember that the AGFL maximum for the number of affix terminals under the domain of one non-terminal is 32, see Chapter 1).

3.2.2.6 Adverb Position

Affix rule for *AVposition* is the following:

```
## ADVERB POSITION
```

```
AVposition :: +STRESS; -STRESS.
```

Value **+STRESS** is assigned to all adverbs that can be placed in high stress positions (usually, although not exclusively, nucleus positions, see Section 3.2.3.4). Value **-STRESS** is assigned to all adverbs that can be placed in low stress positions (always modifier positions). We identify as high stress positions those in which *mucho*, *tanto* may appear. We identify as low stress positions those in which *muy*, *tan* may appear.

3.2.2.7 Adjective Type

Affix rule for *adjtype* is the following:

```
## ADJECTIVE TYPE
```

```
adjtype :: TITLE; QUAL; PPLE.
```

All adjectives except those classified as **TITLE**, are classified as qualifiers. Value **TITLE** is assigned to certain items that precede human proper names and indicate titles given to human beings (*don*, *doña*, *fray*, *frey*, *sor*, *san*, *santo*, *santa*, *monseñor*). Value **PPLE** is assigned within the grammar to participle verb forms used as adjective modifiers.

3.2.2.8 Position

Affix rule for *position* is the following:

```
## POSITION
```

```
position :: PRE; POST.
```

Values are self-explanatory: **PRE** is assigned to determiners or adjectives that precede the nucleus of the phrase, and **POST** is assigned to determiners or adjectives that follow it.

At this moment, we only identify as exclusively preceding adjectives those formally identified as such by apocope (*buen*, *mal*, *gran*), as well as the adjectives that indicate titles (*san*, *sor*, *etc.*).

3.2.2.9 Degree, Second Term Type and Type of (NoEqual) Comparisons

Affix rule for *degree* is the following:

```
## DEGREE

degree :: degree-; degree+.
degree- :: NIL; POS.
degree+ :: neg.
neg :: ZERO; pos+++ .
pos+++ :: POS+++; pos++.
pos++ :: POS++; pos+.
pos+ :: POS+; pos-.
pos- :: POS-; comp; SUP.
comp :: EQUAL; NoEQUAL.
```

As can be observed, values distinguished for *degree* are quite detailed. Value **NIL** is assigned to items that cannot be graded. Value **POS**, *positive*, is assigned to items that can be graded. The remaining values are assigned to items inherently graded. Various degrees have been distinguished from **ZERO** to **SUP**, or, in other words, by means of *degree* it is possible to state that something is not the case, that something is the case to a certain extent, or that something is no doubt the case. Some of these values are well known in the linguistic tradition and, in general, they are self-explanatory. All of them can be better understood, however, on the basis of the items to which they are assigned in the lexicon, see Appendix E —especial attention deserve **QUANT** and **COMPTVE** adverbs, which will be after responsible for the assignment of *degree* within the grammar, see Section 3.2.3.4.2.2.

Despite the heterogeneous behaviour of adverbs with respect to degree, which often makes that all relevant values for a rule have to be explicitly specified in it, we can state that the nested structure of *degree* is primarily important for the grammar structure. The best example of this is, precisely, the distinction between affix non-terminals *degree+* and *degree-*, which separates values that indicate grammatical capability to accept degree from grammatical capability to assign it. Apart from this, the nested structure of *degree* obviously responds to the desire of accounting in a very precise way for the rich detail and gradation of the scale of degree in Spanish. And, last but not least, the presence of affix non-terminal *comp* deserves a particular mention in this regard too. The reason

for this is that values **EQUAL** and **NoEQUAL** found in the domain of this affix are both valid affix terminal values for all *prsubtype*, *dtsubtype* and *degree*, which ensures that we can formalize for all of them the same associations with second terms of comparison.

Affix rule for *second_term_type* is the following:

```
## SECOND TERM TYPE

second_term_type :: COMPTVE; CONSTVE.
```

Adverbs assigned **COMPTVE** degree and pronouns or determiners assigned **COMPTVE** type might introduce a constituent semantically associated with the **COMPTVE** item. This second term is assigned value **CONSTVE** for *second_term_type* if it is introduced by conjunction *que* and the **COMPTVE** item has value **EQUAL** for *degree* (*tan tarde que no pudo ver nada*). It is assigned value **COMPTVE** for *second_term_type* if it is introduced by conjunction *como* and the **COMPTVE** item has value **EQUAL** for *degree* (*tan tarde como siempre*). If, on the contrary, the **COMPTVE** item has value **NoEQUAL** for *degree*, affix non-terminal *second_term_type* is always assigned **COMPTVE** value (*más tarde que nunca*).

Affix rule for *CH_comp* is the following:

```
## TYPE OF (NoEQUAL) COMPARISON

CH_comp :: NIL; CH_comp+.
CH_comp+ :: particular; general.
particular :: PARTICULAR; PARTICULAR+.
general :: GENERAL; GENERAL+.
```

Types of comparison, *CH_comp* values, are assigned within the grammar taking into account values for *degree*, second terms of comparison and determiners found in the whole structure assigned degree. Every non-terminal of the grammar that is assigned a value for *degree* must be also assigned a value for *CH_comp*, this assignment takes place according to the following rules:

- Value **NIL** is assigned in case value for *degree* is different from **NoEQUAL** (*tanta gente, bastante aceptable, muy pronto, pronto*).

- Value **PARTICULAR** is assigned in case value for *degree* is **NoEQUAL** and no second term of comparison could be identified. For NPs containing a comparative adjective or adjective phrase, determiner type must be different from **ART**. Nominalizations are never assigned this value for *CH.comp* (*más libros, unos libros mejores, más bueno, más cerca*).
- Value **GENERAL** is assigned to nominalizations in case value for *degree* is **NoEQUAL** and no second term of comparison could be identified. Under the same conditions, it is also assigned to NPs containing a comparative adjective or adjective phrase if determiner type is **ART** (*el más rápido, la mejor profesora*).
- Values **PARTICULAR+** or **GENERAL+** are assigned in case value for *degree* is **NoEQUAL** and, respectively, **PARTICULAR+** or **GENERAL+** second terms of comparison could be identified (second terms of comparison are formalized in Section 3.2.3.4.2.2). **GENERAL+** requires the whole comparative structure to be a nominalization or an NP with determiner type **ART** (*mejor idea que la tuya, más que nosotros, más bonito que el otro, más cerca que ellos, el más espabilado de todos, el mejor corredor que he conocido*).

Nested structure of *CH.comp* is self-explanatory: in a first step, positive values for *CH.comp* are separated from negative ones. In a second step, *particular* values are separated from *general* values.

3.2.2.10 Person, Gender, Number and Case

Affix rules for *person*, *gender*, *number* and *case* have already been introduced in Section 2.2.2.4.

As regards affix non-terminal *gender*, only values **A_FEM** and **FEM_UN** need additional explanations. **A_FEM** is the value assigned to feminine nouns beginning by [á-] (*hada*), which have to be combined with determiners *el*, *un* (see Section 3.1.3.1). **FEM_UN** is the value assigned to compound cardinal determiners when its last component is *un* and one of the preceding components indicates feminine gender (*veintiuna mil un, doscientas un*). Both values are complement-

ary: nouns with **A.FEM** gender require **FEM_UN** value from determiners that have this possibility.

With respect to affix non-terminal *person*, on the other hand, values **FOURTH**, **FIFTH** and **SIXTH**, immediately dominated by non-terminal *inherent person*, are exclusively used in possessive determiners and pronouns. These values, as well as those found in the domain of *person*, simultaneously encode for possessive determiners and pronouns what for verbs is encoded by the two affix non-terminals *number* and *person*¹⁷.

3.2.2.11 Modifier Category and Lexical Category

Affix rule for *MDcategory* is the following:

```
## MODIFIER CATEGORY

MDcategory :: postmodifier; premodifier.
postmodifier :: apposition; DETERMINER; recursive_post.
apposition :: HOMO_APP; HETERO_APP.
recursive_post :: RELCLAUSE; no_RELCLAUSE.
no_RELCLAUSE :: PRPPHRASE; no_PRPPHRASE.
no_PRPPHRASE :: AJPHRASE; NOTHING.
premodifier :: prmd_general; prmdno_general.
prmd_general :: ADJECTIVE; NOUN; TITLE; DEGREE; QUANT; FACTOR;
                NOTHING.
prmdno_general :: no_general.
```

Affix non-terminal *MDcategory* contains information about the character or the structural category of the modifiers added to the various possible categories of nucleus. Consequently, this affix is exclusively assigned within the grammar, where it is primarily used to control the structure of the addition of modifiers to different nucleus categories. This will be explained in detail in Section 3.2.3.2.2.

Although most terminal values in the domain of affix non-terminal *MDcategory* are self-explanatory, additional explanations might be of benefit in certain cases:

DETERMINER This value is assigned to nucleus structures extended with a determiner postmodifier. AvPs can be extended with this type of modifier: *ahora mismo*.

¹⁷Affix non-terminal *number* of possessive pronouns or determiners encodes a *number* value that indicates the number of things owned by the possessor. This value, contrary to what happens with the values for *number* and *person* simultaneously encoded by affix non-terminal *inherent person*, is taken into account for agreement.

HOMO_APP This value is assigned to nucleus structures extended with apposition postmodifiers with the same structural category found in the nucleus. NPs and AvPs can be extended with this type of apposition modifiers: *la casa Santalla, aquí abajo*¹⁸.

HETERO_APP This value is assigned to nucleus structures extended with apposition postmodifiers with structural categories different from that of the nucleus. AvPs can be extended with prepositional phrases as apposition modifiers: *aquí en Italia*.

RELCLAUSE This value is assigned to nucleus structures extended with relative clause postmodifiers. NPs, PrPs and certain AvPs can be extended with this type of modifiers: *el hombre que susurraba a los caballos, aquéllos a los que más se ha ayudado, ahora que vamos despacio*.

PRPPRHASE This value is assigned to nucleus structures extended with prepositional phrase postmodifiers. NPs, PrPs and certain AvPs can be extended with this type of modifiers: *la casa de los espíritus, algunos de los que vinieron, cerca del límite*.

AJPHRASE This value is assigned to nucleus structures extended with adjective phrase postmodifiers. NPs and PrPs (although this is very rare) can be extended with this type of modifiers: *el ladrón más avisado, la casa verde*.

ADJECTIVE This value is assigned to nucleus structures extended with adjective premodifiers. NPs can be extended with this type of modifiers: *el buen ladrón*.

NOUN This value is assigned to nucleus structures extended with noun premodifiers. AvPs can be extended with this type of modifiers: *calle abajo*.

TITLE This value is assigned to nucleus structures extended with adjective premodifiers that indicate titles. NPs can be extended with this type of modifiers: *santo Tomás de Aquino*.

¹⁸This possibility of modification forced the introduction of affix **lexical_category** —see below in this Section— to record the necessary information about the structural category found in the nucleus.

DEGREE This value is assigned to nucleus structures extended with degree modifiers. NPs, PrPs, AvPs and AjPs can be extended with this type of modifiers: *muy buen carpintero, muy pocos hombres, muy pocos, muy temprano, muy hacia arriba, muy atractivo.*

EXCL This value is found in the domain of non-terminal *no_general*, see Section 3.2.2.2. It is assigned to nucleus structures extended with exclamative degree adverb *qué*. NPs, PrPs, AvPs and AjPs can be extended with this type of modifiers: *qué poco tiempo, qué gran acontecimiento, qué pocos, qué tarde, qué bonitas.*

QUANT This value is assigned to nucleus structures extended with quantification of comparative degree. NPs, AvPs and AjPs can be extended with this type of modifiers: *mucho más buenas razones que las tuyas, mucho más deprisa, dos años antes, mucho más hacia arriba, mucho mejores.*

REL This value is also found in the domain of non-terminal *no_general*, see Section 3.2.2.2. It is assigned to nucleus structures extended with relative quantification of comparative degree. NPs, AvPs and AjPs can be extended with this type of modifiers: *cuanto más buenas soluciones, cuanto más cerca, cuanto más hacia dentro, cuanto más listas.*

FACTOR This value is assigned to nucleus structures extended with quantification by means of quantified NPs with nucleus *veces*. NPs, PrPs, AvPs and AjPs can be extended with this type of modifiers: *mil veces más tiempo que la mayoría, mil veces más mala persona, mil veces más, mil veces más deprisa, mil veces mejores.*

NOTHING This value is assigned in case nucleus structures are not extended with modifiers.

The nested structure of affix *MDcategory* —see Section 3.2.3.2.2— is of great importance for the grammar. In a first step, by means of non-terminals *premodifier* and *postmodifier*, the values that correspond to structures of modifiers that precede the nucleus are separated from the values that correspond to structures of

modifiers that follow it. In a second step, within the domain of non-terminal *post-modifier*, non-terminal *recursive_post* separates the values that represent modifiers which can be combined with other modifiers from those that represent modifiers which cannot —**DETERMINER** and values dominated by affix non-terminal *apposition*.

As regards the internal structure of the domain of *recursive_post* non-terminal, it is due to the necessity of controlling the addition of relative clause, prepositional phrase and adjective phrase modifiers postponed to the nucleus. In order to control the recursive addition of these modifiers, affix non-terminal *recursive_post* must be hierarchically organized in a way so that, together with each terminal value that represents one of the modifier structures in question, there is always only one affix non-terminal that dominates all the remaining terminal values that represent those modifiers which might precede the modifier structure denoted by the terminal value in question. For instance, immediately in the domain of non-terminal *recursive_post*, we find terminal value **REL_CLAUSE**, as well as non-terminal *no_REL_CLAUSE*, which dominates all **PREPPHRASE**, **ADJPHRASE** and **NOTHING** terminal values. Obviously, the affix structure described in this paragraph needs to be after conveniently used within the grammar if we want to ensure that the order in which modifiers will be recursively recognized will be correct —see Section 3.2.3.2.4.2 for an explanation about how this affix structure is actually integrated in the rules of the grammar.

Non-terminal *premodifier*, finally, dominates non-terminals *prmd_general* and *prmdno_general*. These non-terminals, as those found in the domains of *prtype* and *dttype*, separate premodifiers which are able to introduce especial types of clauses —exclamative and relative ones— from premodifiers which are not.

Affix rule for *lexical_category* is the following:

```
## LEXICAL CATEGORY
lexical_category :: N; Av.
```

Values, if we still bear the above **HOMO_APP** appositions in mind, are self-explanatory. Only values required to solve the problem of homogeneous appositions are currently accounted for. We can expect, however, that future extensions of the grammar will highly benefit from the introduction of new values in the

domain of this affix non-terminal.

3.2.2.12 Relative Clause Function

Affix rule for *RCfunction* is the following:

```
## RELATIVE CLAUSE FUNCTION

RCfunction :: RCgeneral; CUANTO_MAS.
RCgeneral  :: modifier; argument; LINKED; NOMINAL.
modifier   :: AGR_MOD; no_AGR_MOD.
no_AGR_MOD :: no_AdvP; AdvP.
argument   :: S; DO; IO; PR; PC; AG; C.
```

Feature *RCfunction* encodes information about the syntactic function of relative clauses. The values of affix *RCfunction*, which will restrict the syntactic functional possibilities of the whole relative clause, are always assigned within the grammar, where different relative constituents (NPs, PrPs, AvPs or AjPs that include relative pronouns, determiners or adverbs) determine different values for the affix. This point, which is only briefly outlined here, will be further clarified by the explanation of the relevant grammar rules, see Sections 3.2.3.4.1.3, 3.2.3.5 and 3.2.3.6.

Most terminal values in the domain of *RCfunction* are self-explanatory. The following specifications, however, can be of help:

CUANTO_MAS This value indicates that the relative clause has a syntactic function that is useful beyond the clause level of analysis (see Footnote 10):
Cuantos más nos ayuden, antes acabaremos.

LINKED This value indicates that the relative clause is a prepositional modifier:
las vidas de quienes no aman la vida.

NOMINAL This value indicates that the relative clause is the nucleus of a nominalization: *lo que el viento se llevó.*

AGR_MOD This value indicates that the relative clause is a modifier and that the antecedent of the relative NP or PrP and the relative NP or PrP itself agree on *person*, *gender* and *number*: *las personas a las que te referes.*

no_AdvP This value indicates that the relative clause is the modifier of a category structure different from AvP, and that the antecedent of the relative NP, PrP or AvP and the relative NP, PrP or AvP itself do not agree on none of the features *person*, *gender* or *number*: *el caso de cuyas consecuencias estamos hablando aquí*.

AdvP This value indicates that the relative clause is the modifier of an AvP, and that the antecedent of the relative AvP and the relative AvP itself do not agree on none of the features *person*, *gender* or *number*: *siempre que pasa por aquí*.

S, DO, IO, PR, PC, AG, C These values are self-explanatory. They indicate that the relative clause is, respectively, the subject, direct object, indirect object, predicative complement, prepositional complement, agent or one of the circumstances of the main clause.

The association between the relative constituents of relative clauses and the syntactic functions of these relative clauses has been described in the previous Sections 3.1.4.1.3, 3.1.5 and 3.1.6, the formal rules that account for this association will be displayed in Sections 3.2.3.4.1.3, 3.2.3.5 and 3.2.3.6.

The nested structure of feature *RCfunction* is self-explanatory. It is required by the grammar, which benefits from the main division between, on the one hand, non-terminal *argument*, which dominates all syntactic functions found at the level of the clause and, on the other hand, non-terminal *modifier*, which dominates the terminal value **AGR_MOD** as well as the affix non-terminal *no_AGR_MOD*—which in turn dominates all terminal values relevant for the identification of not agreeing modifiers.

3.2.2.13 Conjunction and Preposition Type

Affix rules for affix non-terminals *cjtype* and *preptype*, both in the domain of *linktype*, were introduced in Section 2.2.2.5.

Although terminal values found in these domains are mostly self-explanatory, it might be worthwhile—in spite of their apparent transparency and because of its relevance for the analysis of relative clauses—to remark that there exist

certain differences between the meanings of value **NIL** for each one of the two affix non-terminals *preptype* and *cjtype*. In the first case, **NIL** is assigned within the grammar, and means that no preposition has been found occupying the syntactic position that corresponds to the syntax non-terminal containing the affix non-terminal *preptype* in question. In the second case, **NIL** is assigned in the lexicon, and is the value associated with any conjunction different from those already assigned a value in the domain of *cjtype*¹⁹.

3.2.2.14 Clause Type and Question Type

Affix rule for *clause_type* is the following:

```
## CLAUSE TYPE
```

```
clause_type :: MAIN; SUBORDINATED; EL_SUBORDINATED; QUE; EL_QUE; SI.
```

Value **MAIN** is assigned to clauses which are not dominated by a higher level *clause* node (*Has comido lo necesario*, *¿Has comido lo necesario?*, *¿Quién come lo necesario?*, *¡Qué necesario es comer!*, *Come lo necesario*). Value **SUBORDINATED** is assigned to clauses which, on the contrary, are dominated by a higher level *clause* node (*No creo que comas lo necesario*, *No creo comer lo necesario*, *No sé si comes lo necesario*, *No sé quién come lo necesario*, *Las personas citadas por la Audiencia deberán presentarse en un plazo más corto*). Value **EL_SUBORDINATED** is assigned to clauses which are dominated by a higher level *clause* node and preceded by determiner *el* (*El decir la verdad sólo ayuda a veces*). Value **QUE** is assigned to subordinated clauses introduced by conjunction *que* (*Espero que esta vez lo hayas entendido mejor*). Value **EL_QUE** is assigned to subordinated clauses introduced by conjunction *que* and preceded by determiner *el* (*Nunca me importó el que no tuviéramos más hermanos*).

Affix rule for *question_type* is the following:

```
## QUESTION TYPE
```

```
question_type :: Yes_No; Wh.
```

Values are self-explanatory. Value **Yes_No** is assigned to main or subordinated interrogative clauses that wait for an affirmative or negative answer: *¿Ha llegado*

¹⁹This difference is primarily caused by the strong provisional character of the lexicon definitions of conjunctions, see Section 2.2.2.5.

ya?, *No sé* si ha llegado aún. Value **Wh** is assigned to main or subordinated interrogative clauses introduced by interrogative phrases: *¿Quién ha llegado?*, *No sé* quién ha llegado.

3.2.3 The Grammar

This Section describes the AGFL syntax rules that join the elements that constitute NPs, PrPs, AvPs and AjPs. This means that the grammar of these phrases describes the contextualization of the corresponding lexical categories, or, in other words, how the elements, determiners or modifiers, that expand the lexical categories in question, are added to these, thus yielding the syntagms that realize clause syntactic functions (see Section 3.1). In formal terms, the grammar of phrases describes how the non-terminals rewritten as determiners or as categories that can function as modifiers are combined with the non-terminals rewritten as lexical categories, thus constituting the various phrase non-terminals into which clause syntactic function non-terminals are rewritten. To account for this, this Section will be organized following the same structure seen in Section 3.1: the rules that rewrite non-terminals for lexical categories of the grammar into lexicon non-terminals described in Section 3.2.1 will be introduced in Section 3.2.3.1, the rules that join nucleus and modifiers will be described in Section 3.2.3.2, the rules that account for the addition of determiners will be described in Section 3.2.3.3, and the rules that rewrite modifiers and other constituents of the clause, as well as prepositional constituents, will be respectively introduced in Sections 3.2.3.4, 3.2.3.6 and 3.2.3.7.

Besides, as we have, on the one hand, that the target non-terminals obtained by means of the rules described in these Sections are the non-terminals into which syntactic functions are rewritten, and, on the other, that these syntactic functions, together with the **PREDICATE** function that organizes them, constitute the syntactic category *clause*, we can conclude that syntactic functions, along with the non-terminals described in this Section, are the nodes that support the effective integration of *clause* and *phrase* grammars. Certain particularities associated with this integration, as well as illustrative examples of the relevant rewrite rules, will be showed in Section 3.2.3.6 too.

Finally, it might be useful to close the introduction to this Section by insisting

on the fact that, as was observed above (see Chapter 1), the linguistic phenomena which have a role in the constitution of NPs, PrPs, AvPs and AjPs, contrary to those which take part in the constitution of VPs, have a plain syntactic nature. For this reason, within our framework these structures —unlike VPs— must be formally accounted for in accordance with a strict category-function model of description. Although which non-terminals are the names of the categories and which of the functions is always quite obvious, we have adopted the convention of using capital letters to name functions and small letters to name categories. Although category and function non-terminals are both mostly showed by the output of the parser, there are certain category non-terminals which are not. These invisible non-terminals, which are used for different reasons that we will explain where necessary throughout the exposition, might always be recognized because their names start with words **HEAD** or **HIDDEN**.

3.2.3.1 Noun, Pronoun, Adverb and Adjective Non-terminals

The non-terminals described in this Section rewrite as the lexicon non-terminals described in Section 3.2.1. The former, and so their relation with the corresponding **-St** lexicon non-terminals, can be compared to **Verb** non-terminals and their relation with **VerbSt** non-terminals within the module of VPs. The main reason adduced in Section 2.2.3 for the introduction of these non-terminals in the grammar was that this step of the analysis can be used for the insertion of morphological rules within the grammar, as well as of regular expressions for the identification of unknown or partially incorrect input. There are, however, yet other benefits of using these intermediate non-terminals within the grammar. These will be pointed out as we go through all the rules accounted for here.

With regard to **-St** non-terminals specifically, it is worthwhile to remark the fact that these, as well as all non-terminals found in levels of analysis lower than that of non-terminals **Noun**, **Pronoun**, **Adverb**, **Adjective** or **Determiner** in the parse tree, are never showed by the output of the parser, that is to say, they cannot be found in any of the analyses returned for the input sequences.

The rules that rewrite noun and other lexical categories into non-terminals of the lexicon are the following:

Noun (*n*type, *g*ender, *n*umber):

```

NounSt ( ntype, gender, number);
{ gender :: MASC|FEM },
  ( { number :: SING };
    { number :: PLU } ),
  Proper Name ( ntype ).

```

This rule already shows the first of the above mentioned additional benefits of using nodes **X/XSt** in the grammar: as can be observed here, alternative rewrite non-terminals for lexical categories can be introduced under the **X** nodes together with the **XSt** non-terminals, which means that, formally, the former can be treated the same as the latter in higher levels of analysis. A good example of this are proper names, which, as they are not included in the lexicon, must be identified in the grammar by means of a regular expression rule that has non-terminal **Proper Name** on its left-hand-side —this rule is not to be showed here. As a result, therefore, of the inclusion of the non-terminal **Proper Name** as a second alternative, along with non-terminal **NounSt**, in the above rule for non-terminal **Noun**, the input sequences identified as proper names under the dominance of a node **Proper Name** will be formally paired, from the point of view of further analysis, with the input sequences identified under the dominance of a node **NounSt**. For proper names, on the other hand, values for *gender* and *number* affix non-terminals are assigned to the corresponding affixes of the non-terminal **Noun** by means of guards, value **SING** is in this case preferred to value **PLU**.

```

Pronoun ( prtype, prsubtype, definiteness, gender, number ):
  PronounSt ( prtype, prsubtype, definiteness, gender, number );
  { prtype :: CARD }, { definiteness :: UNDEFINITE },
    ( { prsubtype :: NoUNIT },
      Written Cardinal Number ( numbering, gd, last, POST,
        gender, number );
      ( { prsubtype :: UNO }, { number :: SING };
        { prsubtype :: NoUNIT }, { number :: PLU } ),
        { gender :: MASC|FEM },
        Arabic Numbers ( number ) );
  { prtype :: ORD }, { prsubtype :: CUARTO },
    { definiteness :: UNDEFINITE },
    Written Ordinal Number ( numbering, POST, gender, number );
  { prtype :: PART }, { prsubtype :: CUARTO },
    { definiteness :: UNDEFINITE }, { gender :: MASC },
    Written Partitive Number ( numbering, POST, gender, number ).

```

```
Pronoun ( prtype, prsubtype, definiteness, person, gender, number ):  
  PronounSt ( prtype, prsubtype, definiteness, person, gender,  
    number ).
```

```
Pronoun ( prtype, prsubtype, definiteness, person, gender, number,  
  case ):  
  PronounSt ( prtype, prsubtype, definiteness, person, gender,  
    number, case ).
```

```
Prepositional Pronoun ( preposition, prtype, prsubtype, definiteness,  
  person, gender, number, case ):  
  Prepositional PronounSt ( preposition, prtype, prsubtype,  
    definiteness, person, gender, number, case ).
```

The first rule accounts for all pronouns except possessive and personal pronouns. In this rule, the last three alternatives, which account both for arabic numbers and for compound cardinal, ordinal or partitive written numbers²⁰ used as pronouns, form another example of how the treatment given to the non-terminals for items contained in the lexicon and to alternative rewrite non-terminals can be paired by means of the nodes **X/XSt** of the analysis. The second, third and fourth rules respectively deal, the former two, with possessive and personal pronouns and, the latter, with personal pronouns merged with prepositions.

```
Adverb ( avtype, avsubtype, preptype, degree, AVposition ):  
  AdvSt ( avtype, avsubtype, preptype, degree, AVposition ).
```

```
Adverb ( MANNER, NIL, NIL, degree, +STRESS ):  
  AdvSt ( degree ).
```

```
Prepositional Adverb ( preposition, avtype, avsubtype, preptype,  
  degree, AVposition ):  
  Prepositional AdvSt ( preposition, avtype, avsubtype, preptype,  
    degree, AVposition ).
```

The first rule accounts for the items that constitute the closed collection of spanish non-derivative adverbs. The second deals with derivative adverbs obtained from adjectives. By way of these rules the treatment given to both types of adverbs is actually paired from the point of view of higher levels of the analysis. The third rule accounts for adverbs merged with prepositions.

²⁰The rewrite rules of non-terminals **Written Cardinal Number**, **Written Ordinal Number**, **Written Partitive Number** and **Arabic Numbers** non-terminals will not be showed here.

```

Adjective ( ajtype, position, degree, gender, number ):
  AdjSt ( ajtype, position, degree, gender, number ).

Determiner ( dtttype, dtsubtype, position, definiteness, gender,
  number ):
  DeterminerSt ( dtttype, dtsubtype, position, definiteness, gender,
    number );
  { dtttype :: CARD }, { definiteness :: UNDEFINITE },
  ( ( { dtsubtype :: NoUNIT }, { last :: +MORE };
    { dtsubtype :: NoUNIT+UNO }, { gd :: BOTH },
    { last :: +UNO };
    { dtsubtype :: NoUNIT_0-A+UNO }, { gd :: MASC_OR_FEM },
    { last :: +UNO } ),
    Written Cardinal Number ( numbering, gd, last, position,
      gender, number );
    ( { dtsubtype :: UNO }, { number :: SING };
    { dtsubtype :: NoUNIT }, { number :: PLU } ),
    { gender :: MASC|FEM },
    Arabic Numbers ( number ) );
  { dtttype :: ORD }, { dtsubtype :: CUARTO },
  { definiteness :: UNDEFINITE },
  ( Written Ordinal Number ( numbering, position, gender,
    number );
    Ordinal numbers ( gender, number ) );
  { dtttype :: PART }, { dtsubtype :: CUARTO },
  { definiteness :: UNDEFINITE }, { gender :: FEM },
  Written Partitive Number ( numbering, position, gender,
    number ).

Determiner ( dtttype, dtsubtype, position, definiteness,
  inherent person, gender, number ):
  DeterminerSt ( dtttype, dtsubtype, position, definiteness,
    inherent person, gender, number ).

Prepositional Determiner ( preposition, dtttype, dtsubtype, position,
  definiteness, gender, number ):
  Prepositional DeterminerSt ( preposition, dtttype, dtsubtype,
    position, definiteness, gender, number ).

```

The first rule of determiners accounts for all determiners except possessive determiners. The second, third and fourth alternatives of this rule, as in the case of pronouns, account for arabic numbers and compound written cardinal, ordinal and partitive numbers used as determiners. The second rule accounts for possessive determiners, and the third for determiners merged with prepositions.

```

Preposition ( preptype ):
  PrepositionSt ( preptype ).

```

```
Conjunction ( cjtype ):
    ConjunctionSt ( cjtype ).
```

3.2.3.2 Addition of Modifiers

3.2.3.2.1 From Non-terminals for Lexical Categories to Target Non-terminals for Phrase Categories Addition of modifiers expands lexical categories and yields phrase categories. From the formal point of view, it expands non-terminals for lexical categories, described in the previous Section, and yields non-terminals for phrase categories. In this Section the non-terminals for lexical categories are connected with the corresponding target non-terminals for phrase categories. Because we think that this will be of benefit for the general understanding of the remaining explanations introduced along Section 3.2.3.2, although very simply, we present here these non-terminals in order to define properly what is done, from the formal point of view, by the process of addition of modifiers—how this is done will be explained in Section 3.2.3.2.2.

- By means of the addition of modifiers to non-terminals for lexical category noun, **Noun/3**, the following non-terminal for NPs must be obtained:

```
Noun Phrase ( ntype, premodifier, postmodifier, CH_comp, degree,
    definiteness, person, gender, number )
```

- By means of the addition of modifiers to non-terminals for lexical category non-personal pronoun, **Pronoun/5,6**, the following non-terminal for PrPs must be obtained:

```
Pronoun Phrase ( premodifier, postmodifier, prtype, prsubtype,
    definiteness, person, gender, number )
```

- By means of the addition of modifiers to non-terminals for lexical category personal pronoun, **Pronoun/7**, the following non-terminal for PrPs must be obtained:

```
Pronoun Phrase ( premodifier, postmodifier, prtype, prsubtype,
    definiteness, person, gender, number, case )
```

- By means of the addition of modifiers to non-terminals for lexical category adverb, **Adverb/5**, the following non-terminal for AvPs must be obtained:

Adverb Phrase (*premodifier*, *postmodifier*, *avtype*, *avsubtype*,
CH_comp, degree)

- By means of the addition of modifiers to non-terminals for lexical category adjective, **Adjective/5**, the following non-terminal for AjPs must be obtained:

Adjective Phrase (*premodifier*, *ajtype*, CH_comp, degree,
gender, number)

All of them are visible non-terminals, but, for reasons that will be clarified in the following explanations, invisible non-terminals closely related to them must be introduced in different points of the grammar; these too will be described in the following sections.

3.2.3.2.2 An AGFL General Frame for the Addition of Modifiers Two main factors determine the formalization of the addition of modifiers:

- Modifiers can be both left and right recursively subordinated to the nucleus. That is to say, addition of modifiers is carried out by means of both left and right recursive application of subordination.
- The recursive addition of modifiers is not free. It must comply with certain general rules to which different phrase categories add different specifications.

According to these statements, addition of modifiers must be formalized as a recursive subordinating process, within the general scheme proposed in Chapter 1 for the AGFL formalization of subordination. Besides, as addition of modifiers must be carefully controlled in order to ensure the correct concatenation of modifiers, the two affixes *premodifier* and *postmodifier* must be introduced in the grammar and used to keep track of the last modifier added to the nucleus at each level of recursion. Moreover, the structure of the affix rules for the definition of these affix non-terminals must be designed so as to ensure maximum exploitation of our knowledge about the way different modifiers behave with respect to each other.

Given the fact that, in the context of addition of modifiers, we decided to assign function name **NUCLEUS** to the non-terminals that play the role of nucleus

and function name **MODIFIER** to the non-terminals that play the roles of modifiers, our AGFL general frame for the formalization of addition of modifiers, both premodifiers and postmodifiers, can be represented as follows²¹:

```

structure ( DEGREE, postmodifier ):
  Degree MODIFIER ( avtype, avsubtype, CH_comp, degree ),
  NUCLEUS ( DEGREE, postmodifier ).

NUCLEUS ( DEGREE, postmodifier ):
  structure ( ADJECTIVE, postmodifier ).

structure ( ADJECTIVE, postmodifier ):
  Adjective MODIFIER ( ajtype, CH_comp, degree, gender, number ),
  NUCLEUS ( ADJECTIVE, postmodifier ).

NUCLEUS ( ADJECTIVE, postmodifier ):
  structure ( NOTHING, postmodifier ).

```

According to these rules, any **structure** in which a **Degree MODIFIER** has been identified as a premodifier, should contain a non-terminal **structure** with an **Adjective MODIFIER** premodifier in the following node. This is ensured by the fact that the **NUCLEUS** of **structure** when a **Degree MODIFIER** has been identified is assigned value **DEGREE** for *premodifier*, and the only rewrite rule of **NUCLEUS** with such a value for *premodifier* is that which rewrites it into **structure** with value **ADJECTIVE** for *premodifier*.

```

structure ( NOTHING, RELCLAUSE ):
  NUCLEUS ( NOTHING, RELCLAUSE ),
  ( Relative Clause MODIFIER ( AGR_MOD, person, gender, number );
    Relative Clause MODIFIER ( no_AdvP ) ).

NUCLEUS ( NOTHING, RELCLAUSE ):
  structure ( NOTHING, no_RELCLAUSE ).

structure ( NOTHING, PRPPHRASE ):
  NUCLEUS ( NOTHING, PRPPHRASE ),
  Prepositional Phrase MODIFIER ( preposition ).

NUCLEUS ( NOTHING, PRPPHRASE ):
  structure ( NOTHING, no_PRPPHRASE ).

structure ( NOTHING, AJPHRASE ):

```

²¹The rewrite rules that account for the **MODIFIER** non-terminals found in these rules and in those included in the next Sections are displayed in Section 3.2.3.4. For the purposes of this Section and the next ones, however, their names are self-explanatory.

```

NUCLEUS ( NOTHING, AJPHRASE ),
  Adjective Phrase MODIFIER ( ajtype, CH_comp, degree, gender,
    number ).

NUCLEUS ( NOTHING, AJPHRASE ):
  structure ( NOTHING, NOTHING ).

structure ( NOTHING, NOTHING ):
  lexical category.

```

According to these rules, any **structure** in which a **Relative Clause MODIFIER** has been identified as a postmodifier should contain a subsequent node of **structure** with a postponed **Prepositional Phrase MODIFIER**, **Adjective Phrase MODIFIER** or no postponed modifier at all. This is ensured by the fact that the **NUCLEUS** of **structure** when a **Relative Clause MODIFIER** has been identified is assigned value **RELCLAUSE** for *postmodifier*, and the rewrite rule of **NUCLEUS** with such a value for *postmodifier* states that the following **structure** node should obtain for *postmodifier* any terminal value in the domain of non-terminal *no_RELCLAUSE*, that is, **PRPPHRASE**, **AJPHRASE** or **NOTHING**. Similarly, any **structure** containing a **Prepositional Phrase MODIFIER**, as the **NUCLEUS** of *structure* when a **Prepositional Phrase MODIFIER** has been identified is assigned value **PRPPHRASE**, and the rewrite rule of **NUCLEUS** with such a value for *postmodifier* states that the following **structure** node should obtain for *postmodifier* any terminal value in the domain of non-terminal *no_PRPPHRASE*, can only select terminal values within the domain of this non-terminal, that is **AJPHRASE** or **NOTHING**. A further reduction of the possible values of *postmodifier* is finally carried out in case an **Adjective Phrase MODIFIER** has been identified; in this context, the subsequent **structure** node—the one in the right-hand side of the rewrite rule of **NUCLEUS** with value **AJPHRASE** for affix *postmodifier*—can only be assigned value **NOTHING** for affix non-terminal *postmodifier*.

This *rule-schemata* permits that **structure** constituents with more external modifiers embrace **structure** constituents with progressively more internal modifiers, ensuring as well that, on the one hand, **Relative Clause MODIFIER** will always and only be recognized if it is a more external modifier than **Prepositional Phrase MODIFIER**, and, on the other, that this in turn will also be

recognized only if more external than **Adjective Phrase MODIFIER**. None of these modifiers, however, is obligatory.

According to this formal description, there is a one-to-one relation between premodifiers that require other premodifiers —once certain type of premodifier has been introduced, it is clear which type can be introduced in the next level of analysis—, whereas relations between postmodifiers are more flexible and, in consequence, less predictable —a relative clause modifier, for instance, might be preceded by a prepositional phrase, an adjective phrase or none of them. These differences between, on the one hand, the behaviour of various categories of premodifiers with respect to each other and, on the other, the behaviour of various categories of postmodifiers with respect to each other, result in the different affix structures of affix non-terminals *premodifier* and *postmodifier*. The assignment of values for affix *premodifier*, in addition, is completely independent from the assignment of values for affix *postmodifier*. We can state, however, that in case premodifiers and postmodifiers appear combined within the same structure, values for affix *premodifier* will always be assigned in higher levels of analysis than values for affix *postmodifier* (given the fact that premodifiers are more external modifiers than postmodifiers).

Besides, the structure of the described general frame for the addition of modifiers, has been designed so as to be open and flexible enough to allow for the introduction or removal of affix values for both *premodifier* and *postmodifier*, as well as to account for the concrete relationships —more or less restricted— that these values can maintain with each other. To account for these concrete relations, restrictions typical of the second level of formalization must also be imposed on the proposed general structure by means of concrete assignments of values to the affix non-terminals *premodifier* and *postmodifier*, as will be showed by Sections 3.2.3.2.4.2, 3.2.3.2.5.2, 3.2.3.2.6.2 and 3.2.3.2.7.2.

3.2.3.2.3 More Details about the AGFL General Frame Dealing with Addition of Modifiers If we look more carefully at the *rule-schemata* described in the previous Section, we discover an inconsistency: the last rule rewrites the category **structure** —when this is not added modifiers at all— into a lexical category, which should never be the case in a strict category-function model of

description.

From a descriptive point of view, this problem has been traditionally overcome by means of a **NUCLEUS** function introduced in the last rewrite rule of **structure**, even though this is not added any modifier. We will not enter here the discussion about the adequacy of this solution, given the fact that, within an AGFL formal description, the corresponding formal strategy cannot be applied unless a different **NUCLEUS** non-terminal —with different arity— is created ²².

Instead of this, we preferred to extend the previous scheme with the following rules, making use of the AGFL capability to hide non-terminals:

```
HEAD structure ( premodifier, postmodifier ):
  structure ( premodifier, postmodifier );
  HIDDEN FINAL structure ( premodifier, postmodifier ).

...
structure ( NOTHING, PRPPHRASE ):
  NUCLEUS ( NOTHING, PRPPHRASE ),
  Prepositional Phrase MODIFIER ( preposition ).

NUCLEUS ( NOTHING, PRPPHRASE ):
  HIDDEN RECURSIVE structure ( NOTHING, no_PRPPHRASE ).
...

HIDDEN FINAL structure ( NOTHING, NOTHING ):
  lexical category.

HIDDEN RECURSIVE structure ( premodifier, postmodifier ):
  structure ( premodifier, postmodifier );
  HIDDEN FINAL structure ( premodifier, postmodifier ).
```

Conceptually, this solution is congruent with the idea that any function that can be associated with a category structure, conceived of as a modified lexical category, can also be associated with the lexical category that supports that structure, that is, this is congruent with the idea that, in certain circumstances, lexical categories are contextualized as such. From the formal point of view, this means, that any function, the **NUCLEUS** of **structure**, for instance, should be rewritten into both **structure** —which includes at least one modifier— and the **lexical**

²²The same **NUCLEUS** non-terminal cannot be introduced in the rule in question because the system rejects cycles of rules such as that which would be created by **structure/2** if rewritten into **NUCLEUS/2**, since this in turn rewrites as **structure/2** by means of the higher level rules that account for recursive addition of modifiers.

category on the basis of **structure**.

Since this duplication for every structural category found in the rewrite rule of every syntactic function was not desirable, we decided to introduce a new invisible non-terminal for each structural category. This invisible non-terminal should be rewritten into either the structural category or the lexical category on the basis of the structural category in question. For reasons that will be clarified in the following paragraphs, Sections 3.2.3.2.5.1, 3.2.3.2.6.1, 3.2.3.2.7.1 and Chapter 4, we preferred to separate **structure** non-terminals which would be used by **NUCLEUS** in further **structure** recursive calls, from **structure** non-terminals used in other points of the grammar. In formal terms, it means that we created two invisible non-terminals, **HEAD structure** and **HIDDEN RECURSIVE structure**, for each structural category, and that one of them, **HEAD structure**, will be called upon from modules of the grammar different from the **structure** module itself, while the other, **HIDDEN RECURSIVE structure**, will be called upon only within the boundaries of the **structure** module itself.

Besides, we still included a third invisible non-terminal for **structure** with zero modifiers, **HIDDEN FINAL structure**, instead of having **lexical category** as a direct rewriting or **HEAD** and **HIDDEN RECURSIVE structure**. The reason for this is that we wanted to avoid the repeated call of all the guards that would be necessary in these two rules in order to assign values to the affixes that, being present in **structure**, are absent from **lexical category**.

Last but not least, we should also note that this formal scheme must still be further extended for certain categories when it appears convenient to reduce the arity of the non-terminal **structure** with respect to attributes that will not be useful in levels of the analysis beyond that of the **structure** module itself. The result of this extension is that represented in the following *rule-schemata*. Imagine, for instance, that we do not consider affixes *premodifier* and *postmodifier* of **structure** relevant for levels of analysis beyond that of **structure**²³:

```
HEAD structure:
  structure;
  HIDDEN FINAL structure ( premodifier, postmodifier ).
```

²³Of course, it must be clear that we want to make this evident within the **structure** module of the grammar, at the highest level of analysis of **structure** in the parse tree.

```

structure:
  HIDDEN structure ( premodifier, postmodifier ).

...
HIDDEN structure ( NOTHING, PRPPHRASE ):
  NUCLEUS ( NOTHING, PRPPHRASE ),
  Prepositional Phrase MODIFIER ( preposition ).

NUCLEUS ( NOTHING, PRPPHRASE ):
  HIDDEN RECURSIVE structure ( NOTHING, no_PRPPHRASE ).
...

HIDDEN FINAL structure ( NOTHING, NOTHING ):
  lexical category.

HIDDEN RECURSIVE structure ( premodifier, postmodifier ):
  structure ( premodifier, postmodifier );
  HIDDEN FINAL structure ( premodifier, postmodifier ).

structure ( premodifier, postmodifier ):
  HIDDEN structure ( premodifier, postmodifier ).

```

HEAD structure/0 rewrites as **structure/0**. Non-terminal **structure/0** rewrites as **HIDDEN structure/2**. **HIDDEN structure/2** starts recursion. By way of **NUCLEUS/2** it calls upon **HIDDEN RECURSIVE structure/2**, which in turn rewrites as **structure/2**. This non-terminal, finally, rewrites again as **HIDDEN structure/2**.

3.2.3.2.4 Addition of Modifiers in Noun Phrase This Section and the following ones will show how the AGFL General Frame for Addition of Modifiers, introduced in previous Sections 3.2.3.2.2 and 3.2.3.2.3, is concretely applied for the formalization of addition of modifiers within various phrase categories. With respect to this, we have already remarked the fact that NPs are the structural categories in which the productivity of the linguistic phenomena that expand lexical categories is larger. For this reason, the exposition of these linguistic phenomena within the context of NPs will be more exhaustive than that of other structural categories, which will be described taking into account what will have already been explained about NPs.

3.2.3.2.4.1 HEAD Noun Phrase/9, HIDDEN RECURSIVE Noun Phrase/9 All the information provided by the attributes handled within the boundaries of the NP module is considered relevant somewhere else in the grammar, so the number of attributes of NP non-terminals is not reduced for their insertion as rewrite alternatives for syntactic functions other than the **NUCLEUS** of NPs. Rules for non-terminals **HEAD Noun Phrase/9** and **HIDDEN RECURSIVE Noun Phrase/9** are, consequently, the following:

```

HEAD Noun Phrase ( ntype, premodifier, postmodifier, CH_comp, degree,
  definiteness, person, gender, number );
Noun Phrase ( ntype, premodifier, postmodifier, CH_comp, degree,
  definiteness, person, gender, number );
HIDDEN FINAL Noun Phrase ( ntype, premodifier, postmodifier,
  CH_comp, degree, definiteness, person, gender, number ).

HIDDEN RECURSIVE Noun Phrase ( ntype, premodifier, postmodifier,
  CH_comp, degree, definiteness, person, gender, number ):
Noun Phrase ( ntype, premodifier, postmodifier, CH_comp, degree,
  definiteness, person, gender, number );
HIDDEN FINAL Noun Phrase ( ntype, premodifier, postmodifier,
  CH_comp, degree, definiteness, person, gender, number ).

```

3.2.3.2.4.2 Noun Phrase/9 The rewrite rules for **Noun Phrase/9** are the following:

- **No MODIFIER** Example: *ajetreo*

```

HIDDEN FINAL Noun Phrase ( ntype, NOTHING, NOTHING, NIL, NIL,
  definiteness, person, gender, number ):
Noun ( ntype, gender, number),
  ntype definiteness ( ntype, definiteness ),
  Noun Phrase person ( person, number ).

```

By means of their insertion in this rule, noun lexical categories can be contextualized. That is to say, they can be added modifiers or assigned syntactic functions. Values initially assigned to various affixes in **Noun Phrase** non-terminals are, for *ntype*, *gender* and *number*, those provided by **Noun**. For *premodifier* and *postmodifier*, values assigned at this level are both **NOTHING**. For *CH_comp* and *degree*, values are both **NIL**—the addition of adjective modifiers might after change these values. Values for *definiteness* and *person*, finally, are assigned according to

the rules formalized in predicates **ntype definiteness** and **Noun Phrase person**. These predicates are the following:

```
ntype definiteness ( COMMON, UNDEFINITE ):.
ntype definiteness ( proper, DEFINITE ):.

Noun Phrase person ( person, number ):
  { person :: THIRD };
  { person :: FIRST|SECOND }, { number :: PLU }.
```

The first is self-explanatory. The second tries first value **THIRD** for *person*. If no analysis can be found with this value, and value for *number* is **PLU**, it tries also values **FIRST** or **SECOND**.

- **Adjective Phrase MODIFIER Example:** *ajetreo continuo*

```
Noun Phrase ( ntype, NOTHING, AJPHRASE, CH_comp, degree,
  definiteness, person, gender, number ):
  Noun Phrase NUCLEUS ( ntype, NOTHING, AJPHRASE, NIL, POS|NIL,
    definiteness, person, gender, number ),
    Adjective Phrase MODIFIER ( QUAL|PPLE, CH_comp, degree,
      gender, number );
  { gender :: A_FEM },
  Noun Phrase NUCLEUS ( ntype, NOTHING, AJPHRASE, NIL,
    POS|NIL, definiteness, person, gender, number ),
    Adjective Phrase MODIFIER ( QUAL|PPLE, CH_comp, degree,
      FEM, number ).
```

The values **QUAL|PPLE** for affix *ajtype* indicate that both adjectives and participles are accepted in this position (see Section 3.2.3.2.7.2). Values **POS|NIL** for *degree* of **Noun Phrase NUCLEUS** ensure that recursive addition of **Adjective Phrase MODIFIER** is possible. **NIL** is the value assigned if no previous AjP has been added, that is, if the **Noun Phrase NUCLEUS** is rewritten by means of the previous rule. **POS**, on the contrary, is the value acquired by **Noun Phrase** when added an AjP modifier, that is, this value is assigned by recursive calls of this rule. Values for *degree* and *CH_comp* assigned to the **Noun Phrase** are those of the last **Adjective Phrase MODIFIER** added, the only one which may have for *degree* a value in the domain of *degree+*.

Nouns beginning by [á-] motivate the introduction of the second alternative in this rule. A complete explanation of the subsystem centred on this problem will result

from various explanations distributed along Section 3.2.3 where necessary (see below in this Section, epigraph **Adjective MODIFIER**, and Section 3.2.3.3.4). In all rules in which postmodifiers are added which agree for *gender* with the nucleus, specific rewrite alternatives similar to this must be included for these nouns—we will omit them in the future. In these alternatives, the **Noun Phrase NUCLEUS**, whose value for *gender* is **A_FEM**, passes on this value to the **Noun Phrase** non-terminal, while the value for *gender* assigned to the modifiers is always **FEM**. The point here is to maintain the specificity of the *gender* value **A_FEM**, in order to ensure correct matching with determiners *un* and *el* in higher levels of the analysis, while series of feminine modifiers are added to the nucleus.

- **Prepositional Phrase MODIFIER Example:** *el día* de la fiesta

```
Noun Phrase ( ntype, NOTHING, PRPPHRASE, CH_comp, degree,
              definiteness, person, gender, number ):
  Noun Phrase NUCLEUS ( ntype, NOTHING, PRPPHRASE, CH_comp,
                       degree, definiteness, person, gender, number ),
  Prepositional Phrase MODIFIER ( preposition ).
```

According to this rule (certain restrictions will be specified in Chapter 4), **Prepositional Phrase MODIFIER** can be introduced by any preposition. Therefore, it is at this point of the grammar where the data about the argumental behaviour of nouns should be considered and play their role for unification with **Prepositional Phrase MODIFIER**. Nevertheless, as we already pointed out in Chapter 1 and Section 3.1, the best way of encoding this information for nouns, as well as the formalization of its matching for unification with **Prepositional Phrase MODIFIER**, are still problems that have to be solved on the basis of the—not yet—available data about nouns and their requirements with respect to noun complements.

- **Relative Clause MODIFIER Example:** *dinero* que llueve del cielo

```
Noun Phrase ( ntype, NOTHING, RELCLAUSE, CH_comp, degree,
              definiteness, person, gender, number ):
  Noun Phrase NUCLEUS ( ntype, NOTHING, RELCLAUSE, CH_comp,
                       degree, definiteness, person, gender, number ),
  ( Relative Clause MODIFIER ( AGR_MOD, person, gender,
                              number );
    Relative Clause MODIFIER ( no_AdvP ) ).
```

- **Apposition Example:** *la madre* Teresa de Calcuta

```

Noun Phrase ( COMMON, NOTHING, HOMO_APP, CH_comp, degree,
              definiteness, person, gender, number ):
  Noun Phrase NUCLEUS ( COMMON, NOTHING, HOMO_APP, CH_comp,
                       degree, definiteness, person, gender, number ),
  Apposition MODIFIER ( HOMO_APP, N, gender, number ).

```

- **Adjective MODIFIER Examples:** *buen vino*, *gran área*

```

Noun Phrase ( ntype, ADJECTIVE, postmodifier, CH_comp, degree-,
              definiteness, person, gender, number ):
  Adjective MODIFIER ( QUAL|PPLE, CH_comp, degree-, gender,
                      number ),
  Noun Phrase NUCLEUS ( ntype, ADJECTIVE, postmodifier,
                       CH_comp1, degree, definiteness, person, gender,
                       number );
  { gender :: FEM },
  Adjective MODIFIER ( QUAL|PPLE, CH_comp, degree-, gender,
                      number ),
  Noun Phrase NUCLEUS ( ntype, ADJECTIVE, postmodifier,
                       CH_comp1, degree, definiteness, person, A_FEM,
                       number ).

```

The cooccurrence of values **QUAL|PPLE** in the parameter position of affix *ajtype* indicates that both adjectives and participles are accepted in this case too (see Section 3.2.3.4.2.1). The fact that inherently graded adjectives are not allowed here is ensured by affix non-terminal *degree-*.

The second alternative in this rule is, again (see above in this Section, epigraph **Adjective Phrase MODIFIER**), motivated by nouns beginning by [á-]. The point here is that, once a previous modifier is added to the NP with nucleus [á-], the fact that the noun begins by [á-] is not relevant anymore. This is formalized in the following way: **Noun Phrase NUCLEUS** with **A_FEM** value for *gender* is, by means of the alternative in question, combined with an **Adjective MODIFIER** whose value for *gender* is **FEM**. From this moment on, all the structure is assigned **FEM** value for *gender*, so higher levels of the analysis will not appreciate any difference between this **Noun Phrase** non-terminal and **Noun Phrase** non-terminals constituted by **Noun Phrase NUCLEUS** with regular **FEM** value for *gender*. The same situation is found in the next rule: the corresponding rewrite alternative will not be specified there.

Examples: mejor *persona*, mejor *alma*

```
Noun Phrase ( ntype, ADJECTIVE, postmodifier, CH_comp, degree+,
  definiteness, person, gender, number ):
  Adjective MODIFIER ( QUAL, CH_comp, degree+, gender, number ),
  Noun Phrase NUCLEUS ( ntype, ADJECTIVE, postmodifier, NIL,
    POS|NIL, definiteness, person, gender, number ),
  ( { CH_comp :: NIL|PARTICULAR|GENERAL };
    ( Second Term MODIFIER ( second_term_type, CH_comp,
      degree+, QUE|COMO );
      { CH_comp :: GENERAL+ },
      Second Term MODIFIER ( second_term_type, CH_comp,
        degree+, DE );
      { CH_comp :: GENERAL+ },
      Second Term MODIFIER ( COMPTVE, CH_comp, degree+,
        DE, person, gender, PLU );
      { CH_comp :: PARTICULAR+ },
      Second Term MODIFIER ( COMPTVE, CH_comp, degree+,
        DE, THIRD, NEUT, SING );
      Second Term MODIFIER ( COMPTVE, CH_comp, degree+,
        person, gender, number ) ) ).
```

This rule accounts for modifiers in which **Degree MODIFIER** and **Adjective MODIFIER** functions are merged, due to the presence of inherently graded adjectives. Non-terminal *degree+* ensures that only inherently graded adjectives will be accepted by this rule. Apart from this, we refrain here from giving more detailed explanations about the subsystem of degree in the grammar, since these appear to be more appropriately placed under the next epigraph **Degree MODIFIER**.

• **Degree MODIFIER Example:** más *buena persona*

```
Noun Phrase ( ntype, DEGREE, postmodifier, CH_comp, degree+,
  definiteness, person, gender, number ):
  Degree MODIFIER ( avtype, avsubtype, CH_comp, degree+ ),
  Noun Phrase NUCLEUS ( ntype, DEGREE, postmodifier, NIL,
    POS, definiteness, person, gender, number ),
  ( ... ).
```

This rule accounts for **Noun Phrase** when this is constituted by **Degree MODIFIER** and **Noun Phrase NUCLEUS**²⁴. The value **POS** for affix *degree* of **Noun Phrase NUCLEUS** ensures that nouns followed or preceded by already graded adjectives will not be added another **Degree MODIFIER**.

²⁴Dots in this alternative stand for the same group specified in the third member of the previous rule.

Degree modifiers are functional in a large number of different points of the grammar. They are mainly responsible for the values obtained by affixes *degree* and *CH_comp*. The explanations that account for this connection between degree modifiers and affixes *degree* and *CH_comp* are distributed, where they are relevant and can contribute to a general understanding of the subsystem of degree, throughout the current Section 3.2.3.

At the point illustrated by this rule, we can verify the following:

- **Degree MODIFIER** might introduce a **Second Term MODIFIER**.
- Both **Degree MODIFIER** and **Second Term MODIFIER** are provided with values for *CH_comp* and *degree*.
- The value for *degree* of **Noun Phrase** is that supplied by **Degree MODIFIER** and matched by **Second Term MODIFIER**.
- According to the rules described in Section 3.2.3.4.2.2 the value for *degree* in this rule is a single value, already fully specified, while the value for *CH_comp*, on the contrary, might still be open for further specification.
- If **Second Term MODIFIER** is realized, the value for *CH_comp*, which must be fully determined, will have to be matched by the addition of determiners or the assignment of syntactic function to **Noun Phrase** in higher levels of the analysis. Otherwise, at the level represented by this rule, the value might still remain open.
 - It will be fully determined if **Degree MODIFIER** has value **NIL** for *CH_comp*, that is, if it has a value for *degree* different from **NoEQUAL**.
 - It will be restricted to **PARTICULAR|GENERAL** if **Degree MODIFIER** has value **NoEQUAL** for *degree*. In this case, the value for *CH_comp* will not be fully specified until **Noun Phrase** is extended with determiners or assigned syntactic function (see Sections 3.2.3.3.4 and 3.2.3.6).

The rewrite rules of non-terminal **Second Term MODIFIER** are described in Section 3.2.3.4.2.2. We have considered as acceptable **Second Term MODIFIERS** within the context of **Noun Phrase** those exemplified by the following NPs—we include one example for each of the alternatives found in the group that

introduces **Second Term MODIFIER**, in the same order in which these alternatives appear in the group: *mejores sedas* que en Pekín, *la más buena persona* del mundo, *el mejor consejero* de los que he tenido, *más buen tiempo* de lo que creía, *el mejor profesor* que he tenido.

We will include similar examples of various possibilities for the realization of **Second Term MODIFIER** in future explanations of different phrase categories. It must be clear, however, that we consider this an adjustable detail of the type identified in Chapter 1 as corresponding to *second level of formalization*, that is, these are linguistic specifications associated with vague requirements, frequently lexically motivated, between syntactic components. What is important when dealing with them (see Chapter 1), is the fact that the grammar, in this case the AGFL general frame developed to account for the addition of modifiers, is designed so as to be flexible enough to allow for a quick introduction or removal of alternatives depending on, for instance, the kind of text that has to be parsed.

Examples: *muy poco buena persona*, *muy poco mejor persona*

```
Noun Phrase ( ntype, DEGREE, postmodifier, CH_comp, pos++,
  definiteness, person, gender, number ):
  { pos++ :: POS++|EQUAL|SUP },
  Degree MODIFIER ( avtype, avsubtype, CH_comp, pos++ ),
  Noun Phrase NUCLEUS ( ntype, DEGREE, postmodifier, NIL,
    POS-, definiteness, person, gender, number ),
  [ Second Term MODIFIER ( second_term_type, CH_comp, pos++,
    QUE|COMO ) ].
```

This rule accounts for secondary degree positions added after *poco*, which can be either a **Degree MODIFIER** or a **Comparison Quantifier MODIFIER**.

- **(Exclamative) Degree MODIFIER Example:** *qué buena persona*

```
Noun Phrase ( ntype, EXCL, recursive_post, CH_comp, degree,
  definiteness, person, gender, number ):
  Degree MODIFIER ( EXCL, NIL, CH_comp, degree ),
  Noun Phrase NUCLEUS ( ntype, EXCL, recursive_post,
    CH_comp, POS|POS-, definiteness, person, gender,
    number ).
```

- **Comparison Quantifier MODIFIER Example:** mucho *más buena persona*.

```

Noun Phrase ( ntype, QUANT, recursive_post, CH_comp, pos++,
  definiteness, person, gender, number ):
{ pos++ :: POS++|POS+|POS-|EQUAL|SUP },
  Comparison Quantifier MODIFIER ( avtype, avsubtype,
    CH_comp, pos++ ),
  Noun Phrase NUCLEUS ( ntype, QUANT, recursive_post,
    particular, NoEQUAL, definiteness, person, gender,
    number ),
  [ Second Term MODIFIER ( second_term_type, CH_comp,
    pos++, QUE|COMO ) ].

```

A guard determines in this rule the *degree* values accepted for **Comparison Quantifier MODIFIER**. Specifications of this type correspond to what has been identified as second level of formalization, where our main concern, see above, is to guarantee flexibility and ease of enlargement of the coverage of the rules in question.

- **(Relative) Comparison Quantifier MODIFIER Example:** cuanto *más buena persona*.

```

Noun Phrase ( ntype, REL, recursive_post, CH_comp, degree,
  definiteness, person, gender, number ):
  Comparison Quantifier MODIFIER ( REL, CUANTO, CH_comp,
    degree ),
  Noun Phrase NUCLEUS ( ntype, REL, recursive_post,
    particular, NoEQUAL, definiteness, person, gender,
    number ).

```

- **Factor MODIFIER Example:** mil veces *más buena persona*.

```

Noun Phrase ( ntype, FACTOR, recursive_post, CH_comp, degree,
  definiteness, person, gender, number ):
  Factor MODIFIER ( CH_comp, degree ),
  Noun Phrase NUCLEUS ( ntype, FACTOR, recursive_post,
    CH_comp1, degree1, definiteness, person, gender,
    number ),
  [ Second Term MODIFIER ( second_term_type, CH_comp,
    degree, QUE|COMO ) ].

```

- (Title) Adjective MODIFIER Example: *san Juan Bosco*.

```

Noun Phrase ( ntype, TITLE, NOTHING, CH_comp, degree,
              definiteness, person, gender, number ):
  Adjective MODIFIER ( TITLE, CH_comp, degree, gender, number ),
    Noun Phrase NUCLEUS ( ntype, TITLE, NOTHING, NIL, NIL,
                          definiteness, person, gender, number ).

```

3.2.3.2.4.3 Noun Phrase NUCLEUS/9 The rewrite rules for **Noun Phrase NUCLEUS/9** are the following:

```

Noun Phrase NUCLEUS ( ntype, NOTHING, AJPHRASE, CH_comp, degree,
                      definiteness, person, gender, number ):
  HIDDEN RECURSIVE Noun Phrase ( ntype, NOTHING, AJPHRASE|NOTHING,
                                  CH_comp, degree, definiteness, person, gender, number ).

```

```

Noun Phrase NUCLEUS ( ntype, NOTHING, PRPPHRASE, CH_comp, degree,
                      definiteness, person, gender, number ):
  HIDDEN RECURSIVE Noun Phrase ( ntype, NOTHING, PRPPHRASE, CH_comp,
                                  degree, definiteness, person, gender, number );
  HIDDEN RECURSIVE Noun Phrase ( ntype, NOTHING, no_PRPPHRASE,
                                  CH_comp, degree, definiteness, person, gender, number ).

```

The value **AJPHRASE** of affix *postmodifier* in the first rule as well as the value **PRPPHRASE** in the first alternative of the second rule allow for the recursive additions of, respectively, **Adjective Phrase MODIFIER** and **Prepositional Phrase MODIFIER**.

```

Noun Phrase NUCLEUS ( ntype, NOTHING, RELCLAUSE, CH_comp, degree,
                      definiteness, person, gender, number ):
  HIDDEN RECURSIVE Noun Phrase ( ntype, NOTHING, no_RELCLAUSE,
                                  CH_comp, degree, definiteness, person, gender, number ).

```

```

Noun Phrase NUCLEUS ( ntype, NOTHING, HOMO_APP, CH_comp, degree,
                      definiteness, person, gender, number ):
  HIDDEN RECURSIVE Noun Phrase ( ntype, NOTHING, NOTHING, CH_comp,
                                  degree, definiteness, person, gender, number ).

```

```

Noun Phrase NUCLEUS ( ntype, ADJECTIVE, postmodifier, CH_comp,
                      degree, definiteness, person, gender, number ):
  HIDDEN RECURSIVE Noun Phrase ( ntype, NOTHING, postmodifier,
                                  CH_comp, degree, definiteness, person, gender, number ).

```

```

Noun Phrase NUCLEUS ( ntype, DEGREE, recursive_post, CH_comp, degree,
                      definiteness, person, gender, number ):
  HIDDEN RECURSIVE Noun Phrase ( ntype, ADJECTIVE, recursive_post,
                                  CH_comp, degree, definiteness, person, gender, number );

```

```

{ degree :: POS- },
  HIDDEN RECURSIVE Noun Phrase ( ntype, DEGREE|QUANT,
    recursive_post, CH_comp, degree, definiteness, person,
    gender, number ).

Noun Phrase NUCLEUS ( ntype, EXCL, recursive_post, CH_comp, degree,
  definiteness, person, gender, number ):
  HIDDEN RECURSIVE Noun Phrase ( ntype, ADJECTIVE, recursive_post,
    CH_comp, degree, definiteness, person, gender, number );
{ degree :: POS- },
  HIDDEN RECURSIVE Noun Phrase ( ntype, DEGREE, recursive_post,
    CH_comp, degree, definiteness, person, gender, number ).

```

The second alternatives in the last two rules allow for the addition of secondary degree positions after **Degree MODIFIER** *poco*.

```

Noun Phrase NUCLEUS ( ntype, QUANT|REL, recursive_post, CH_comp,
  degree, definiteness, person, gender, number ):
  HIDDEN RECURSIVE Noun Phrase ( ntype, DEGREE|ADJECTIVE,
    recursive_post, CH_comp, degree, definiteness, person,
    gender, number ).

Noun Phrase NUCLEUS ( ntype, FACTOR, recursive_post, CH_comp, degree,
  definiteness, person, gender, number ):
  HIDDEN RECURSIVE Noun Phrase ( ntype, QUANT, recursive_post,
    CH_comp, degree, definiteness, person, gender, number );
{ degree :: NoEQUAL },
  HIDDEN RECURSIVE Noun Phrase ( ntype, DEGREE|ADJECTIVE,
    recursive_post, CH_comp, degree, definiteness, person,
    gender, number ).

```

3.2.3.2.5 Addition of Modifiers in Pronoun Phrase

3.2.3.2.5.1 HEAD Pronoun Phrase/8,9, HIDDEN RECURSIVE Pronoun Phrase/8,9 All the information provided by the attributes handled within the PrP module is, as in the case of NPs, considered relevant somewhere else in the grammar, so the number of attributes of PrPs is not reduced for further integration of PrPs beyond the PrP module itself. On the other hand, we can observe that the whole module of PrPs is somewhat duplicated, due to the fact that personal pronouns are always dealt with by means of specific rules and non-terminals that include affix case, which is not relevant for the rest of pronouns. We will include specific rules dealing with personal pronouns only in case they incorporate additional particularities with respect to the corresponding rules for non-personal

pronouns. It must be also remarked that the capability of personal pronouns to accept the addition of modifiers is considerably reduced with respect to that of non-personal pronouns. To begin with, we have prevented (see Section 3.1.2.2) their combination with adjective and prepositional modifiers. Nevertheless, these are the kind of adjustable decisions that concern second level of formalization, and therefore their formalization has been designed so as to be easily modified if necessary.

Rules for non-terminals **HEAD Pronoun Phrase/8** and **HIDDEN RECURSIVE Pronoun Phrase/8** are the following:

```

HEAD Pronoun Phrase ( premodifier, postmodifier, prtype, prsubtype,
    definiteness, person, gender, number ):
  Pronoun Phrase ( premodifier, postmodifier, prtype, prsubtype,
    definiteness, person, gender, number ),
  Top Pronoun ( gender, number, prtype, prsubtype );
HIDDEN FINAL Pronoun Phrase ( premodifier, NOTHING, prtype,
    prsubtype, definiteness, person, gender, number ),
  Top Pronoun ( gender, number, prtype, prsubtype ).

```

Not all pronouns can be contextualized by themselves, that is, without being added determiners (partitive pronouns different from *medio*, or relative pronoun *cual*, for instance, cannot). The purpose of predicate non-terminal **Top Pronoun** included in the rewrite rule of **HEAD Pronoun Phrase/8** is to account for this fact. Besides, it also accounts for the fact that sometimes pronouns that can be contextualized by themselves, cannot equally be contextualized for all their inflectional forms; that is, certain inflectional forms of certain pronouns appear only in the context of certain determiners (*otro*, for instance, is not assigned **NEUT** value for *gender* unless it is preceded by neuter determiners, *eso otro*, *lo otro*, see also Section 3.2.3.3.5). The current rewrite rule of this predicate is the following:

```

Top Pronoun ( gender, number, prtype, prsubtype ):
  { prtype :: INTG|COMPTVE|GLOBAL|DEM|POSS|QUANT|DIS };
  { prtype :: REL }, { prsubtype :: QUE|QUIEN|CUANTO };
  { gender :: MASC|FEM }, { prtype :: INDEF|CARD },
    { prsubtype :: UNO|OTRO|NoUNIT };
  { gender :: MASC|FEM }, { number :: PLU },
    { prtype :: INDEF|CARD }, { prsubtype :: DEMAS };
  { prtype :: INDEF }, { prsubtype :: QUIENQUIERA|ALGUNO
    |ALGUIEN|CUALQUIERA };
  { prtype :: PART }, { prsubtype :: MEDIO }.

```

The problem now is obvious: non-terminal **HEAD Pronoun Phrase/8** cannot be the basis for the addition of determiners, because some pronouns (or their association with certain inflectional values) would never be identified. The solution, however, is quite straightforward and it illustrates yet another advantage of the use of hidden **HEAD-** non-terminals on the top level of the description of the modules of the grammar that account for the addition of modifiers to different phrase categories. This invisible node allows, in fact, for the writing of as much rules for different **HEAD-** non-terminals as needed by further contextualizations of the phrase category in question. In the case of PrPs, for instance, we have a second rule that accounts for these PrPs that will be added determiners. This rule is the following:

```
HEAD Pronoun Phrase for NOMINAL ( premodifier, postmodifier, prtype,
    prsubtype, definiteness, person, gender, number ):
Pronoun Phrase ( premodifier, postmodifier, prtype, prsubtype,
    definiteness, person, gender, number );
HIDDEN FINAL Pronoun Phrase ( premodifier, NOTHING, prtype,
    prsubtype, definiteness, person, gender, number ).
```

Rule for **HIDDEN RECURSIVE Pronoun Phrase/8** is the following:

```
HIDDEN RECURSIVE Pronoun Phrase ( premodifier, postmodifier, prtype,
    prsubtype, definiteness, person, gender, number ):
Pronoun Phrase ( premodifier, postmodifier, prtype, prsubtype,
    definiteness, person, gender, number );
HIDDEN FINAL Pronoun Phrase ( premodifier, NOTHING, prtype,
    prsubtype, definiteness, person, gender, number ).
```

Obviously, the corresponding **HEAD-** non-terminal for personal pronouns, since these are never specified by determiners, need not to be broken down into two different rules. The rewrite rules of **HEAD Pronoun Phrase/9** and **HIDDEN RECURSIVE Pronoun Phrase/9** are the following:

```
HEAD Pronoun Phrase ( premodifier, postmodifier, prtype,
    prsubtype, definiteness, person, gender, number, case ):
Pronoun Phrase ( premodifier, postmodifier, prtype, prsubtype,
    definiteness, person, gender, number, case );
HIDDEN FINAL Pronoun Phrase ( premodifier, postmodifier, prtype,
    prsubtype, definiteness, person, gender, number, case ).

HIDDEN RECURSIVE Pronoun Phrase ( premodifier, postmodifier,
    prtype, prsubtype, definiteness, person, gender, number,
    case ):
```

```

Pronoun Phrase ( premodifier, postmodifier, prtype, prsubtype,
                 definiteness, person, gender, number, case );
HIDDEN FINAL Pronoun Phrase ( premodifier, postmodifier, prtype,
                               prsubtype, definiteness, person, gender, number, case ).

```

3.2.3.2.5.2 Pronoun Phrase/8,9 The rewrite rules for **Pronoun Phrase/8,9** are the following:

- **No MODIFIER** Example: *alguien*

```

HIDDEN FINAL Pronoun Phrase ( NOTHING, NOTHING, prtype,
                              prsubtype, definiteness, person, gender, number ):
( Pronoun ( prtype, prsubtype, definiteness, gender, number );
  Pronoun ( prtype, prsubtype, definiteness, person1, gender,
            number ) ),
  Pronoun Phrase person ( prtype, prsubtype, person,
                          number ).

```

By means of this rule, pronoun lexical categories are contextualized. At the level of the analysis represented by this rule, the values of affixes *premodifier* and *postmodifier* are **NOTHING**; the values for *prtype*, *prsubtype*, *definiteness*, *gender* and *number* are provided by **Pronoun**; the value for *person* is acquired according to the rules formalized in the predicate **Pronoun Phrase person**. This predicate assigns values for *person* depending on *prtype*, *prsubtype* and *number*.

```

Pronoun Phrase person ( prtype, prsubtype, person, number ):
{ prtype :: INTG|EXCL|COMPTVE|GLOBAL|DEM|DIS|QUANT|INDEF|
  REFL|CARD|ORD|PART },
  ( { person :: THIRD };
    { number :: PLU }, { person :: FIRST|SECOND } );
{ prtype :: REL }, { prsubtype :: QUE|QUIEN },
  ( { person :: THIRD };
    { person :: FIRST|SECOND } );
{ prtype :: REL }, { prsubtype :: CUAL|CUANTO },
  ( { person :: THIRD };
    { number :: PLU }, { person :: FIRST|SECOND } );
{ prtype :: POSS }, { person :: THIRD }.

```

For pronouns and determiners, degree is contained in the meaning of some pronoun or determiner types and subtypes. This has consequences, and certain rules have to be broken down into different rules to allow for the optional addition of

Second Term MODIFIERS only if we have **COMPTVE** pronouns or determiners. When no modifier is identified, as in the previous rule, there is no need to break down the rule in question, but an additional rule is also necessary to account for the fact that **COMPTVE** pronouns might introduce a second term of comparison. This rule, though it has value **NOTHING** for both *premodifier* and *postmodifier*, must have a visible left-hand-side non-terminal, since this non-terminal gives the name of the node that joins the comparative pronoun and the second term modifier:

Example: *tantos como nosotros*

```
Pronoun Phrase ( NOTHING, NOTHING, COMPTVE, comp, definiteness,
  person, gender, number ):
  Pronoun Phrase NUCLEUS ( NOTHING, NOTHING, COMPTVE, comp,
    definiteness, person, gender, number ),
    ( Second Term MODIFIER ( second_term_type, CH_comp, comp,
      QUE|COMO );
      Second Term MODIFIER ( COMPTVE, PARTICULAR+, degree+,
        DE, person, gender, number ) ).
```

Obviously, the corresponding rule for personal pronouns does not include predicate

Pronoun Phrase person:

Example: *nosotros*

```
HIDDEN FINAL Pronoun Phrase ( NOTHING, NOTHING, prtype,
  prsubtype, definiteness, person, gender, number, case ):
  Pronoun ( prtype, prsubtype, definiteness, person, gender,
    number, case ).
```

- **Adjective Phrase MODIFIER** The rules that add postmodifiers to PrPs must be broken down into different rules to account for the addition of **Second Term MODIFIERS** to, exclusively, **COMPTVE** pronouns, see above in this Section. Apart from the rules for **PERS** pronouns, three more rules are necessary if we want to guarantee that only **COMPTVE** pronouns will be allowed to introduce a **Second Term MODIFIER**: of these three rules, one should deal with **COMPTVE** pronouns themselves, another one with **no_COMPTVE** pronouns, and the other with **prno_general** pronouns, which are not contained in the **no_COMPTVE** domain. The addition of **Adjective Phrase MODIFIERS** to PrPs, however, is very rare, and at present only the rule dealing with

no_COMPTVE pronouns is included in the grammar. A similar situation is found in the case of **Prepositional Phrase MODIFIERS**, although their productivity is somewhat larger. Only when dealing with the addition of **Relative Clause MODIFIERS** to PrPs, did we decide to include in the grammar the four rules necessary to cover all pronoun types in Spanish, but, even in this case, some examples still appear to be very rare.

Example: *pocos* inteligentes

```
Pronoun Phrase ( NOTHING, AJPHRASE, prno_COMPTVE, prsubtype,
  definiteness, person, gender, number ):
  Pronoun Phrase NUCLEUS ( NOTHING, AJPHRASE, prno_COMPTVE,
    prsubtype, definiteness, person, gender, number ),
  Adjective Phrase MODIFIER ( QUAL|PPLE, CH_comp, degree,
    gender, number ).
```

- **Prepositional Phrase MODIFIER Example:** *algunos* de los asistentes

```
Pronoun Phrase ( NOTHING, PRPPHRASE, prno_COMPTVE, prsubtype,
  definiteness, person, gender, number ):
  Pronoun Phrase NUCLEUS ( NOTHING, PRPPHRASE, prno_COMPTVE,
    prsubtype, definiteness, person, gender, number ),
  Prepositional Phrase MODIFIER ( preptype ).
```

- **Relative Clause MODIFIER Example:** *aquellos* que puedan ayudar

```
Pronoun Phrase ( NOTHING, RELCLAUSE, prno_COMPTVE, prsubtype,
  definiteness, person, gender, number ):
  Pronoun Phrase NUCLEUS ( NOTHING, RELCLAUSE, prno_COMPTVE,
    prsubtype, definiteness, person, gender, number ),
  ( Relative Clause MODIFIER ( AGR_MOD, person, gender,
    number );
  Relative Clause MODIFIER ( no_AdvP ) ).
```

Example: *¿Quién* que lo supiera *podría* callarlo?

```
Pronoun Phrase ( NOTHING, RELCLAUSE, prno_general, prsubtype,
  definiteness, person, gender, number ):
  Pronoun Phrase NUCLEUS ( NOTHING, RELCLAUSE, prno_general,
    prsubtype, definiteness, person, gender, number ),
  ( Relative Clause MODIFIER ( AGR_MOD, person, gender,
    number );
  Relative Clause MODIFIER ( no_AdvP ) ).
```

Example: *más* que se van antes de tiempo

```
Pronoun Phrase ( NOTHING, RELCLAUSE, COMPTVE, comp, definiteness,
  person, gender, number ):
  Pronoun Phrase NUCLEUS ( NOTHING, RELCLAUSE, COMPTVE, comp,
    definiteness, person, gender, number ),
    ( Relative Clause MODIFIER ( AGR_MOD, person, gender,
      number );
      Relative Clause MODIFIER ( no_AdvP ) ),
    [ Second Term MODIFIER ( second_term_type, CH_comp, comp,
      QUE|COMO );
      Second Term MODIFIER ( COMPTVE, PARTICULAR+, degree+, DE,
        person, gender, number ) ] .
```

As we can observe here, it is the **Pronoun Phrase NUCLEUS** itself which, being the comparative element, causes the introduction of the optional **Second Term MODIFIER** in this rule. Formally, this means that more than one **Second Term MODIFIER** could be identified at different levels of recursion for just one comparative element. This problem, which, in fact, appears wherever the introduction of an optional term in a subordinating recursive structure depends on the nucleus, can be easily solved by the addition of an affix, functional within the recursive module, to keep track of the realization of the optional term in question. Although in the case of PrPs this role could be played by the affix *CH_comp* itself, taking into account that modified structures of PrPs are rarely —perhaps never— nested, we preferred, for the time being, to maintain these rules in this, simpler, form.

Unlike for NPs and AjPs, value of *CH_comp* in PrPs is fully specified at this level of analysis, due to the fact that it only depends on **Pronoun Phrase NUCLEUS** and **Second Term MODIFIER**²⁵. Examples of **Second Term MODIFIERS** covered by this rule are the following: *tantos que nos quieren ayudar que es abrumador, más que quieren participar de los que esperaba*.

Example: *nosotros* que le teníamos tanto respeto

```
Pronoun Phrase ( NOTHING, RELCLAUSE, prtype, prsubtype,
  definiteness, person, gender, number, case ):
  Pronoun Phrase NUCLEUS ( NOTHING, RELCLAUSE, prtype,
    prsubtype, definiteness, person, gender, number, case ),
```

²⁵Values for *CH_comp* found in the domain of *general* are not relevant for PrPs.

```
( Relative Clause MODIFIER ( AGR_MOD, person, gender,
    number );
  Relative Clause MODIFIER ( no_AdvP ) ).
```

- Degree MODIFIER Example: *muy pocas*

```
Pronoun Phrase ( DEGREE, recursive_post, QUANT, POCO,
  definiteness, person, gender, number ):
{ pos+++ :: POS+++|POS++|EQUAL|SUP },
  Degree MODIFIER ( avtype, avsubtype, NIL, pos+++ ),
  Pronoun Phrase NUCLEUS ( DEGREE, recursive_post, QUANT,
    POCO, definiteness, person, gender, number ),
  [ Second Term MODIFIER ( second_term_type, NIL, pos+++,
    QUE|COMO ) ].
```

- (Exclamative) Degree MODIFIER Example: *qué pocas*

```
Pronoun Phrase ( EXCL, recursive_post, QUANT, POCO,
  definiteness, person, gender, number ):
  Degree MODIFIER ( EXCL, NIL, CH_comp, NIL ),
  Pronoun Phrase NUCLEUS ( EXCL, recursive_post, QUANT,
    POCO, definiteness, person, gender, number ).
```

- Factor MODIFIER Example: *mil veces más*

```
Pronoun Phrase ( FACTOR, recursive_post, COMPTVE, comp,
  definiteness, person, gender, number ):
  Factor MODIFIER ( CH_comp, degree ),
  Pronoun Phrase NUCLEUS ( NOTHING, recursive_post, COMPTVE,
    NoEQUAL, definiteness, person, gender, number ),
  [ Second Term MODIFIER ( second_term_type, CH_comp, degree,
    QUE|COMO ) ].
```

3.2.3.2.5.3 Pronoun Phrase NUCLEUS/8,9 The rewrite rules for Pronoun Phrase NUCLEUS/8,9 are the following:

```
Pronoun Phrase NUCLEUS ( NOTHING, NOTHING, prtype, prsubtype,
  definiteness, person, gender, number ):
  Pronoun ( prtype, prsubtype, definiteness, gender, number ),
  Pronoun Phrase person ( prtype, prsubtype, person,
    number ).
```

```
Pronoun Phrase NUCLEUS ( NOTHING, AJPHRASE, prtype, prsubtype,
  definiteness, person, gender, number ):
  HIDDEN RECURSIVE Pronoun Phrase ( NOTHING, NOTHING, prtype,
    prsubtype, definiteness, person, gender, number ).
```

Pronoun Phrase NUCLEUS (NOTHING, PREPPHRASE, prtype, prsubtype, definiteness, person, gender, number):
 HIDDEN RECURSIVE Pronoun Phrase (NOTHING, no_PREPPHRASE, prtype, prsubtype, definiteness, person, gender, number).

Pronoun Phrase NUCLEUS (NOTHING, RELCLAUSE, prtype, prsubtype, definiteness, person, gender, number):
 HIDDEN RECURSIVE Pronoun Phrase (NOTHING, no_RELCLAUSE, prtype, prsubtype, definiteness, person, gender, number).

Pronoun Phrase NUCLEUS (NOTHING, RELCLAUSE, prtype, prsubtype, definiteness, person, gender, number, case):
 HIDDEN RECURSIVE Pronoun Phrase (NOTHING, no_RELCLAUSE, prtype, prsubtype, definiteness, person, gender, number, case).

Pronoun Phrase NUCLEUS (DEGREE|EXCL, recursive_post, prtype, prsubtype, definiteness, person, gender, number):
 HIDDEN RECURSIVE Pronoun Phrase (NOTHING, recursive_post, prtype, prsubtype, definiteness, person, gender, number).

Pronoun Phrase NUCLEUS (FACTOR, recursive_post, prtype, prsubtype, definiteness, person, gender, number):
 HIDDEN RECURSIVE Pronoun Phrase (NOTHING, recursive_post, prtype, prsubtype, definiteness, person, gender, number).

3.2.3.2.6 Addition of Modifiers in Adverb Phrase

3.2.3.2.6.1 HEAD Adverb Phrase/6, Adverb Phrase/6,7, HIDDEN RECURSIVE Adverb Phrase/7 The piece of information that refers to the prepositional requirements of the adverbs that constitute **Adverb Phrases** is only relevant within the AvP module of analysis itself. This explains the reduction of the affix *preptype* that contains this information for further integration of AvPs beyond the limits of the AvP module.

The rewrite rules of non-terminals **HEAD Adverb Phrase** and **Adverb Phrase/6** are the following:

HEAD Adverb Phrase (premodifier, postmodifier, avtype, avsubtype, CH_comp, degree):
 Adverb Phrase (premodifier, postmodifier, avtype, avsubtype, CH_comp, degree);
 HIDDEN FINAL Adverb Phrase (premodifier, postmodifier, avtype, avsubtype, preptype, CH_comp, degree).

Adverb Phrase (premodifier, postmodifier, avtype, avsubtype,

```

    CH_comp, degree ):
HIDDEN Adverb Phrase ( premodifier, postmodifier, avtype,
    avsubtype, preptype, CH_comp, degree ).

```

The rewrite rules of non-terminals **HIDDEN RECURSIVE Adverb Phrase/7** and **Adverb Phrase/7** are the following:

```

HIDDEN RECURSIVE Adverb Phrase ( premodifier, postmodifier,
    avtype, avsubtype, preptype, CH_comp, degree ):
    Adverb Phrase ( premodifier, postmodifier, avtype, avsubtype,
        preptype, CH_comp, degree );
HIDDEN FINAL Adverb Phrase ( premodifier, postmodifier,
    avtype, avsubtype, preptype, CH_comp, degree ).

Adverb Phrase ( premodifier, postmodifier, avtype, avsubtype,
    preptype, CH_comp, degree ):
    HIDDEN Adverb Phrase ( premodifier, postmodifier, avtype,
        avsubtype, preptype, CH_comp, degree ).

```

3.2.3.2.6.2 HIDDEN Adverb Phrase/7

- **No MODIFIER** Example: *aquí*

```

HIDDEN FINAL Adverb Phrase ( NOTHING, NOTHING, avtype,
    avsubtype, preptype, CH_comp, degree ):
    ( { CH_comp :: PARTICULAR }, { degree :: NoEQUAL };
    { CH_comp :: NIL }, { degree :: NIL|POS|ZERO|POS+++
        |POS++|POS+|EQUAL|POS-|SUP } ),
    Adverb ( avtype, avsubtype, preptype, degree, +STRESS ).

```

By means of this rule, adverb lexical categories are contextualized. At this level of the analysis *premodifier* and *postmodifier* affixes are both assigned value **NOTHING**. Values for *avtype*, *avsubtype*, *preptype* and *degree* are provided by **Adverb**. Value for *CH_comp* is assigned according to both the value for *degree* and the fact that no **Second Term MODIFIER** can be identified in this case. Only adverbs assigned value **+STRESS** for attribute *AVposition* can be placed in this syntactic position.

As was the case with respect to PrPs, an additional rule is necessary to account for the fact that inherently graded adverbs might introduce a **SECOND Term MODIFIER**:

Example: *más que amor*

```
HIDDEN Adverb Phrase ( NOTHING, NOTHING, avtype, avsubtype,
  preptype, CH_comp, degree+ ):
  Adverb Phrase NUCLEUS ( NOTHING, NOTHING, avtype, avsubtype,
    preptype, CH_comp, degree+ ),
    ( Second Term MODIFIER ( COMPTVE, CH_comp, degree+, QUE );
      { avtype :: TIME }, { avsubtype :: NIL },
        { CH_comp :: PARTICULAR+ },
          Second Term MODIFIER ( second_term_type, CH_comp,
            degree+, DE );
        { CH_comp :: PARTICULAR+ },
          Second Term MODIFIER ( COMPTVE, CH_comp, degree+, DE,
            THIRD, NEUT, SING ) ).
```

Non-terminal *degree+* ensures that only inherently graded adverbs will be accepted by this rule.

We already pointed out in Section 3.1.2.3 the fact that addition of modifiers in AvPs shows the highest degree of lexical interference with syntax. Examples of this fact will be recurrent in formal rules that describe AvPs. We will only remark the most relevant ones.

In general, guards determining the values of affixes *avtype* and *avsubtype* are used in AvP rules to restrict the domain of the application of these rules to certain items. All the specifications expressed by these guards are, obviously, typical decisions of the second level of formalization. In the current rule, for instance, guards establish that one of the possible realizations of **Second Term MODIFIER**, the second one —some phrase introduced by preposition *de*— is only allowed for inherently graded adverbs with value **TIME** for affix *avtype* and value **NIL** for affix *avsubtype*, (*antes, después*, see Appendix E, *después del recreo*).

• **Noun MODIFIER Example:** *río arriba*

```
HIDDEN Adverb Phrase ( NOUN, NOTHING, avtype, avsubtype,
  preptype, CH_comp, degree ):
  { avtype :: PLACE }, { avsubtype :: ATRAS },
    Noun MODIFIER,
      Adverb Phrase NUCLEUS ( NOUN, NOTHING, avtype, avsubtype,
        preptype, CH_comp, degree ).
```

In order to account properly for this kind of modifier-adverb combinations, the semantic type or, more exactly, the lemma of the noun underlying the **Noun**

MODIFIER, should be compared with the semantic type or the lemma of the adverb item occupying the position of **Adverb Phrase NUCLEUS**. To check this compatibility, however, was not possible at the moment, because the 1.7.52 version of the AGFL system —unlike later versions— does not permit the definition of affix non-terminals that contain an unlimited number of values, as would be necessary for storing and handling information about lemmas of nouns and adverbs.

- **Determiner MODIFIER Example:** *nunca* más

```
HIDDEN Adverb Phrase ( NOTHING, DETERMINER, avtype,
  avsubtype, preptype, CH_comp, degree ):
( { avtype :: TIME }, { avsubtype :: NUNCA },
  { dttype :: COMPTVE }, { dtsubtype :: NoEQUAL };
  { avtype :: TIME }, { avsubtype :: AHORA },
  { dttype :: REFL }, { dtsubtype :: NIL };
  { avtype :: PLACE }, { avsubtype :: AQUI },
  { dttype :: REFL }, { dtsubtype :: NIL } ),
  Adverb Phrase NUCLEUS ( NOTHING, DETERMINER, avtype,
    avsubtype, preptype, CH_comp, degree ),
  Determiner MODIFIER ( dttype, dtsubtype, POST,
    definiteness, NEUT, SING ).
```

- **Apposition MODIFIER Example:** *aquí* delante

```
HIDDEN Adverb Phrase ( NOTHING, HOMO_APP, avtype, avsubtype,
  preptype, CH_comp, degree ):
( { avtype :: TIME }, { avsubtype :: NUNCA },
  { avsubtype1 :: JAMAS };
  { avtype :: PLACE }, { avsubtype :: AQUI },
  { avsubtype1 :: DETRAS|ATRAS } ),
  Adverb Phrase NUCLEUS ( NOTHING, HOMO_APP, avtype,
    avsubtype, preptype, CH_comp, degree ),
  Apposition MODIFIER ( HOMO_APP, Av, avtype, avsubtype1,
    degree1 ).
```

Example: *aquí* en la playa

```
HIDDEN Adverb Phrase ( NOTHING, HETERO_APP, avtype, avsubtype,
  preptype, CH_comp, degree ):
{ avtype :: PLACE }, { avsubtype :: AQUI },
  { preposition :: EN },
  Adverb Phrase NUCLEUS ( NOTHING, HETERO_APP, avtype,
    avsubtype, preptype, CH_comp, degree ),
  Apposition MODIFIER ( HETERO_APP, Av, preposition ).
```

- **Prepositional Phrase MODIFIER Example:** *cerca* del cielo

```
HIDDEN Adverb Phrase ( NOTHING, PRPPHRASE, avtype, avsubtype,
  preposition, CH_comp, degree ):
  Adverb Phrase NUCLEUS ( NOTHING, PRPPHRASE, avtype,
    avsubtype, preposition, CH_comp, degree ),
  Prepositional Phrase MODIFIER ( preposition ).
```

- **Relative Clause MODIFIER Example:** *ahora* que lo dices

```
HIDDEN Adverb Phrase ( NOTHING, RELCLAUSE, avtype, avsubtype,
  preptype, CH_comp, degree ):
  ( { avtype :: TIME }, { avsubtype :: AHORA|SIEMPRE|MIENTRAS } );
  { avtype :: MANNER }, { avsubtype :: COMOQUIERA };
  { avtype :: PLACE }, { avsubtype :: AQUI|DONDEQUIERA
    |COMOQUIERA } ),
  Adverb Phrase NUCLEUS ( NOTHING, RELCLAUSE, avtype,
    avsubtype, preptype, CH_comp, degree ),
  Relative Clause MODIFIER ( AdvP ).
```

- **Degree MODIFIER Example:** *muy bien*

```
HIDDEN Adverb Phrase ( DEGREE, recursive_post, avtype,
  avsubtype, preptype, CH_comp, degree+ ):
  Degree MODIFIER ( avtype1, avsubtypel, CH_comp, degree+ ),
  Adverb Phrase NUCLEUS ( DEGREE, recursive_post, avtype,
    avsubtype, preptype, NIL, POS ),
  ( { CH_comp :: NIL|PARTICULAR };
    ( Second Term MODIFIER ( second_term_type, CH_comp,
      degree+, QUE|COMO );
      { avtype :: PLACE }, { avsubtype :: ACA },
      { CH_comp :: PARTICULAR+ },
      Second Term MODIFIER ( second_term_type, CH_comp,
        degree+, DE );
      { CH_comp :: PARTICULAR+ },
      Second Term MODIFIER ( COMPTVE, CH_comp, degree+,
        DE, THIRD, NEUT, SING ) ) ).
```

Unlike NPs (see Section 3.2.3.2.4.2, epigraph **Degree MODIFIER**), AvPs have already fully specified values for affix *CH_comp*. This is possible because AvPs are not concerned by *CH_comp* values in the domain of affix non-terminal *general*. That is, for AvPs the value assigned for *CH_comp* depends exclusively on the non-terminals **Degree MODIFIER** and **Second Term MODIFIER** present in this rule.

Examples of **Second Term MODIFIERS** covered by this rule are the following:
más que cualquiera de los otros, más allá del bien y del mal, más de lo que cabría esperar.

Example: *muy poco amablemente*

```
HIDDEN Adverb Phrase ( DEGREE, recursive_post, avtype, avsubtype,
  preptype, CH_comp, pos++ ):
{ pos++ :: POS++|EQUAL|SUP },
  Degree MODIFIER ( avtype1, avsubtype1, CH_comp, pos++ ),
  Adverb Phrase NUCLEUS ( DEGREE, recursive_post, avtype,
    avsubtype, preptype, NIL, POS- ),
  [ Second Term MODIFIER ( second_term_type, CH_comp, pos++,
    QUE|COMO ) ].
```

This rule accounts for secondary degree positions added after a previous **Degree MODIFIER** *poco*.

- **(Exclamative) Degree MODIFIER Example:** *qué cerca*

```
HIDDEN Adverb Phrase ( EXCL, recursive_post, avtype, avsubtype,
  preptype, CH_comp, degree ):
  Degree MODIFIER ( EXCL, NIL, CH_comp, degree ),
  Adverb Phrase NUCLEUS ( EXCL, recursive_post, avtype,
    avsubtype, preptype, NIL, POS|POS- ).
```

- **Comparison Quantifier MODIFIER Example:** *un poco más cerca*

```
HIDDEN Adverb Phrase ( QUANT, recursive_post, avtype, avsubtype,
  preptype, CH_comp, pos++ ):
{ pos++ :: POS++|POS+|POS-|EQUAL|SUP },
  Comparison Quantifier MODIFIER ( avtype1, avsubtype1,
    CH_comp, pos++ ),
  Adverb Phrase NUCLEUS ( QUANT, recursive_post, avtype,
    avsubtype, preptype, particular, NoEQUAL ),
  [ Second Term MODIFIER ( second_term_type, CH_comp, pos++,
    QUE|COMO ) ].
```

Example: *doscientos metros más lejos*

```
HIDDEN Adverb Phrase ( QUANT, recursive_post, avtype, avsubtype,
  preptype, CH_comp, degree ):
{ avtype :: PLACE|TIME },
  Comparison Quantifier MODIFIER ( CH_comp, degree, THIRD,
    gender, number ),
```

```

Adverb Phrase NUCLEUS ( QUANT, recursive_post, avtype,
    avsubtype, preptype, particular, NoEQUAL ),
[ Second Term MODIFIER ( second_term_type, CH_comp, degree,
    QUE|COMO ) ].

```

As observed above with respect to **Noun MODIFIERs**, the semantic types of the NPs underlying **Comparison Quantifier MODIFIERs** should be checked to prove that they can be used as possible units of measurement for the dimensions indicated by the adverb item occupying the syntactic position of **Adverb Phrase NUCLEUS**. For the same reasons adduced above, however, this could not be done in this work.

- **(Relative) Comparison Quantifier MODIFIER Example:** *cuanto más cerca*

```

HIDDEN Adverb Phrase ( REL, recursive_post, avtype, avsubtype,
    preptype, CH_comp, degree ):
Comparison Quantifier MODIFIER ( REL, CUANTO, CH_comp,
    degree ),
    Adverb Phrase NUCLEUS ( REL, recursive_post, avtype,
        avsubtype, preptype, particular, NoEQUAL ).

```

- **Factor MODIFIER Example:** *mil veces más rápido*

```

HIDDEN Adverb Phrase ( FACTOR, recursive_post, avtype, avsubtype,
    preptype, CH_comp, degree ):
Factor MODIFIER ( CH_comp, degree ),
    Adverb Phrase NUCLEUS ( FACTOR, recursive_post, avtype,
        avsubtype, preptype, particular, NoEQUAL ),
    [ Second Term MODIFIER ( second_term_type, CH_comp, degree,
        QUE|COMO ) ].

```

3.2.3.2.6.3 Adverb Phrase NUCLEUS/7

```

Adverb Phrase NUCLEUS ( NOTHING, NOTHING, avtype, avsubtype,
    preptype, CH_comp, degree ):
Adverb ( avtype, avsubtype, preptype, degree, +STRESS ).

```

```

Adverb Phrase NUCLEUS ( NOUN, NOTHING, avtype, avsubtype,
    preptype, CH_comp, degree ):
HIDDEN RECURSIVE Adverb Phrase ( NOTHING, NOTHING, avtype,
    avsubtype, preptype, CH_comp, degree ).

```

```

Adverb Phrase NUCLEUS ( NOTHING, DETERMINER, avtype, avsubtype,
    preptype, CH_comp, degree ):
HIDDEN RECURSIVE Adverb Phrase ( NOTHING, NOTHING, avtype,

```

```

        avsubtype, preptype, CH_comp, degree ).

Adverb Phrase NUCLEUS ( NOTHING, HOMO_APP, avtype, avsubtype,
    preptype, CH_comp, degree ):
    HIDDEN RECURSIVE Adverb Phrase ( NOTHING, NOTHING, avtype,
        avsubtype, preptype, CH_comp, degree ).

Adverb Phrase NUCLEUS ( NOTHING, HETERO_APP, avtype, avsubtype,
    preptype, CH_comp, degree ):
    HIDDEN RECURSIVE Adverb Phrase ( NOTHING, DETERMINER|NOTHING,
        avtype, avsubtype, preptype, CH_comp, degree ).

Adverb Phrase NUCLEUS ( NOTHING, PRPPHRASE, avtype, avsubtype,
    preptype, CH_comp, degree ):
    HIDDEN RECURSIVE Adverb Phrase ( NOTHING, NOTHING, avtype,
        avsubtype, preptype, CH_comp, degree );
    HIDDEN RECURSIVE Adverb Phrase ( NOTHING, HOMO_APP, avtype,
        avsubtype, preptype, CH_comp, degree ).

Adverb Phrase NUCLEUS ( NOTHING, RELCLAUSE, avtype, avsubtype,
    preptype, CH_comp, degree ):
    HIDDEN RECURSIVE Adverb Phrase ( NOTHING, NOTHING, avtype,
        avsubtype, preptype, CH_comp, degree ).

```

The fact that **Adverb Phrase NUCLEUS** never rewrites as a **HIDDEN RECURSIVE Adverb Phrase** non-terminal with an affix non-terminal in the position of affix *postmodifier* is due to the rigidity of the process of addition of modifiers within the context of AvPs. Obviously, the specifications expressed by the values assigned to affix *postmodifier* in these rules are decisions that concern the second level of formalization, and our experience is that they are indeed frequently revised depending on concrete applications.

```

Adverb Phrase NUCLEUS ( DEGREE, recursive_post, avtype, avsubtype,
    preptype, CH_comp, degree ):
    HIDDEN RECURSIVE Adverb Phrase ( NOTHING, recursive_post,
        avtype, avsubtype, preptype, CH_comp, degree );
    { degree :: POS- },
    HIDDEN RECURSIVE Adverb Phrase ( DEGREE|QUANT,
        recursive_post, avtype, avsubtype, preptype, CH_comp,
        degree ).

Adverb Phrase NUCLEUS ( EXCL, recursive_post, avtype, avsubtype,
    preptype, CH_comp, degree ):
    HIDDEN RECURSIVE Adverb Phrase ( NOTHING, recursive_post,
        avtype, avsubtype, preptype, CH_comp, degree );
    { degree :: POS- },
    HIDDEN RECURSIVE Adverb Phrase ( DEGREE, recursive_post,

```

```
avtype, avsubtype, preptype, CH_comp, degree ).
```

The second alternative in the last two rules accounts for secondary degree positions introduced after a previous **Degree MODIFIER** *poco*.

```
Adverb Phrase NUCLEUS ( QUANT|REL, recursive_post, avtype,
  avsubtype, preptype, CH_comp, degree ):
  HIDDEN RECURSIVE Adverb Phrase ( DEGREE|NOTHING,
    recursive_post, avtype, avsubtype, preptype, CH_comp,
    degree ).
```

```
Adverb Phrase NUCLEUS ( FACTOR, recursive_post, avtype,
  avsubtype, preptype, CH_comp, degree ):
  HIDDEN RECURSIVE Adverb Phrase ( QUANT, recursive_post,
    avtype, avsubtype, preptype, CH_comp, degree );
  { degree :: NoEQUAL },
  HIDDEN RECURSIVE Adverb Phrase ( DEGREE|NOTHING,
    recursive_post, avtype, avsubtype, preptype, CH_comp,
    degree ).
```

3.2.3.2.7 Addition of Modifiers in Adjective Phrase

3.2.3.2.7.1 HEAD Adjective Phrase/6, Adjective Phrase/6,7, HIDDEN RECURSIVE Adjective Phrase/7 As regards affix *postmodifier* of AjPs, it seems to be currently useless for rewrite rules of AjPs, given the fact that it is always assigned value **NOTHING**. Despite the apparent uselessness of the affix, however, it has been introduced because it is in this affix position in which in the near future we plan to specify the information that concerns prepositional complements required by the adjective item occupying the position of the nucleus of the adjective phrase. The affix *postmodifier*, on the other hand, is at the moment, and so will be in the future, only functional within the AjP module itself, for this reason it is removed from the **Adjective Phrase** non-terminals that play a role beyond the AjP module.

The rewrite rules of non-terminals **HEAD Adjective Phrase/6** and **Adjective Phrase/6** are the following:

```
HEAD Adjective Phrase ( premodifier, ajtype, CH_comp, degree, gender,
  number ):
  Adjective Phrase ( premodifier, ajtype, CH_comp, degree, gender,
    number );
  HIDDEN FINAL Adjective Phrase ( premodifier, postmodifier, ajtype,
```

```
CH_comp, degree, gender, number ).
```

```
Adjective Phrase ( premodifier, ajtype, CH_comp, degree, gender,  
number ):  
HIDDEN Adjective Phrase ( premodifier, NOTHING, ajtype, CH_comp,  
degree, gender, number ).
```

The rewrite rules of non-terminals **HIDDEN RECURSIVE Adjective Phrase/7** and **Adjective Phrase/7** are the following:

```
HIDDEN RECURSIVE Adjective Phrase ( premodifier, postmodifier,  
ajtype, CH_comp, degree, gender, number ):  
Adjective Phrase ( premodifier, postmodifier, ajtype, CH_comp,  
degree, gender, number );  
HIDDEN FINAL Adjective Phrase ( premodifier, postmodifier,  
ajtype, CH_comp, degree, gender, number ).  
  
Adjective Phrase ( premodifier, postmodifier, ajtype, CH_comp,  
degree, gender, number ):  
HIDDEN Adjective Phrase ( premodifier, postmodifier, ajtype,  
CH_comp, degree, gender, number ).
```

3.2.3.2.7.2 HIDDEN Adjective Phrase/7

- **No MODIFIER Example:** *agotador*

```
HIDDEN FINAL Adjective Phrase ( NOTHING, NOTHING, ajtype, CH_comp,  
degree, gender, number ):  
( { degree :: NoEQUAL }, { CH_comp :: PARTICULAR|GENERAL };  
{ degree :: POS|SUP }, { CH_comp :: NIL } ),  
Adjective ( ajtype, POST, degree, gender, number );  
{ ajtype :: PPLE }, { CH_comp :: NIL }, { degree :: POS },  
Participle Verb ( mvtype, ACT|MIDD, mood, gender, number ).
```

Adjective lexical categories are contextualized by means of their insertion in this rule. As can be observed here, **Adjective Phrases** can be constituted by either the lexicon non-terminal **Adjective** or the lexicon non-terminal **Participle Verb**. In both cases, however, at this level of analysis the value assigned to affix *premodifier*, as well as that assigned to affix *postmodifier*, is **NOTHING**. As regards the rest of the affixes, the lexicon non-terminal **Adjective** provides values for affixes *ajtype*, *degree*, *gender* and *number*, while the lexicon non-terminal **Participle Verb** only for affixes *gender* and *number* (affix *ajtype* is in this case always assigned value **PPLE**, and affix *degree* value **NIL**). The value for *CH_comp*, finally,

is, no matter which category we find in the nucleus, assigned according to the value of *degree*.

An additional rule is necessary to account for inherently graded adverbs that might introduce a **SECOND Term MODIFIER**:

Example: *mejor que la otra vez*

```
HIDDEN Adjective Phrase ( NOTHING, NOTHING, ajtype, CH_comp, degree+,
    gender, number ):
  Adjective Phrase NUCLEUS ( NOTHING, NOTHING, ajtype, CH_comp,
    degree+, gender, number ),
  ( Second Term MODIFIER ( COMPTVE, CH_comp, degree+, QUE );
    { CH_comp :: GENERAL+ },
    Second Term MODIFIER ( COMPTVE, CH_comp, degree+, DE );
    { CH_comp :: GENERAL+ },
    Second Term MODIFIER ( COMPTVE, CH_comp, degree+, DE,
      person, gender, number1 );
    { CH_comp :: PARTICULAR+ },
    Second Term MODIFIER ( COMPTVE, CH_comp, degree+, DE,
      THIRD, NEUT, SING );
    Second Term MODIFIER ( COMPTVE, CH_comp, degree+, person,
      gender, number ) ).
```

• **Degree MODIFIER Example:** *muy apetecible*

```
HIDDEN Adjective Phrase ( DEGREE, recursive_post, ajtype, CH_comp,
    degree+, gender, number ):
  Degree MODIFIER ( avtype, avsubtype, CH_comp, degree+ ),
  Adjective Phrase NUCLEUS ( DEGREE, recursive_post, ajtype, NIL,
    POS, gender, number ),
  ( ... ).
```

AjP values for *CH_comp* and *degree* are the same found in NPs. This means that the values of *CH_comp* in the domain of non-terminal *general* are relevant for AjPs and, moreover, that the values of affixes *CH_comp* of AjPs might still not be fully determined at this level of analysis. These values, like for NPs, can only be fully determined when **Degree MODIFIER**, **Second Term MODIFIER** and determiners —think of nominalizations of AjPs, see Section 3.2.3.5— found in higher levels of analysis have all been taken into account.

Examples²⁶ of **Second Term MODIFIER** covered by this rule are the following:

²⁶Dots in this alternative stand for the same group specified in the previous rule.

tan divertido como el primero, la más cansada de todas, el más interesante de los que me han ofrecido, más malo de lo que creía, lo más simple que pude encontrar.

Example: *muy poco apetecible*

```
HIDDEN Adjective Phrase ( DEGREE, NOTHING, ajtype, CH_comp, pos++,
  gender, number ):
  { pos++ :: POS++|EQUAL|SUP },
  Degree MODIFIER ( avtype, avsubtype, CH_comp, pos++ ),
  Adjective Phrase NUCLEUS ( DEGREE, NOTHING, ajtype, NIL, POS-,
    gender, number ),
  [ Second Term MODIFIER ( second_term_type, CH_comp, pos++,
    QUE|COMO ) ].
```

This rule accounts for secondary degree positions added beside a **Degree MODIFIER** *poco*.

- **(Exclamative) Degree MODIFIER Example:** *qué interesante*

```
HIDDEN Adjective Phrase ( EXCL, NOTHING, ajtype, CH_comp, degree,
  gender, number ):
  Degree MODIFIER ( EXCL, avsubtype, CH_comp, degree ),
  Adjective Phrase NUCLEUS ( EXCL, NOTHING, ajtype, NIL,
    POS|POS-, gender, number ).
```

- **Comparison Quantifier MODIFIER Example:** *mucho más atractivo*

```
HIDDEN Adjective Phrase ( QUANT, NOTHING, ajtype, CH_comp, pos++,
  gender, number ):
  { pos++ :: POS++|POS+|POS-|EQUAL|SUP },
  Comparison Quantifier MODIFIER ( avtype, avsubtype, CH_comp,
    pos++ ),
  Adjective Phrase NUCLEUS ( QUANT, NOTHING, ajtype, particular,
    NoEQUAL, gender, number ),
  [ Second Term MODIFIER ( second_term_type, CH_comp, pos++,
    QUE|COMO ) ].
```

- **(Relative) Comparison Quantifier MODIFIER Example:** *cuanto más atractivo*

```
HIDDEN Adjective Phrase ( REL, NOTHING, ajtype, CH_comp, degree,
  gender, number ):
  Comparison Quantifier MODIFIER ( REL, CUANTO, CH_comp, degree ),
  Adjective Phrase NUCLEUS ( REL, NOTHING, ajtype, particular,
    NoEQUAL, gender, number ).
```

- **Factor MODIFIER Example:** mil veces *más divertido*

```
HIDDEN Adjective Phrase ( FACTOR, NOTHING, ajtype, CH_comp, degree,
gender, number ):
  Factor MODIFIER ( CH_comp, degree ),
  Adjective Phrase NUCLEUS ( FACTOR, NOTHING, ajtype, particular,
    NoEQUAL, gender, number ),
  [ Second Term MODIFIER ( second_term_type, CH_comp, degree,
    QUE|COMO ) ].
```

3.2.3.2.7.3 Adjective Phrase NUCLEUS/7

```
Adjective Phrase NUCLEUS ( NOTHING, NOTHING, ajtype, CH_comp, degree,
gender, number ):
  Adjective ( ajtype, POST, degree, gender, number ).
```

```
Adjective Phrase NUCLEUS ( DEGREE, NOTHING, ajtype, CH_comp, degree,
gender, number ):
  HIDDEN RECURSIVE Adjective Phrase ( NOTHING, NOTHING, ajtype,
    CH_comp, degree, gender, number );
  { degree :: POS- },
  HIDDEN RECURSIVE Adjective Phrase ( QUANT|DEGREE, NOTHING,
    ajtype, CH_comp, degree, gender, number ).
```

The second alternative accounts for secondary degree positions introduced after a **Degree MODIFIER** *poco*.

```
Adjective Phrase NUCLEUS ( EXCL, NOTHING, ajtype, CH_comp, degree,
gender, number ):
  HIDDEN RECURSIVE Adjective Phrase ( NOTHING, NOTHING, ajtype,
    CH_comp, degree, gender, number );
  { degree :: POS- },
  HIDDEN RECURSIVE Adjective Phrase ( DEGREE, NOTHING, ajtype,
    CH_comp, degree, gender, number ).
```

```
Adjective Phrase NUCLEUS ( QUANT|REL, NOTHING, ajtype, CH_comp,
degree, gender, number ):
  HIDDEN RECURSIVE Adjective Phrase ( DEGREE|NOTHING, NOTHING,
    ajtype, CH_comp, degree, gender, number ).
```

```
Adjective Phrase NUCLEUS ( FACTOR, NOTHING, ajtype, CH_comp, degree,
gender, number ):
  HIDDEN RECURSIVE Adjective Phrase ( QUANT, NOTHING, ajtype,
    CH_comp, degree, gender, number );
  { degree :: NoEQUAL },
  HIDDEN RECURSIVE Adjective Phrase ( DEGREE|NOTHING, NOTHING,
    ajtype, CH_comp, degree, gender, number ).
```

3.2.3.3 Addition of Determiners

3.2.3.3.1 From Non-terminals for Lexical Categories to Target Non-terminals for Phrase Categories Together with addition of modifiers, addition of determiners is also an important operation for the contextualization of lexical categories to yield phrase categories.

Addition of determiners might be combined with addition of modifiers. In this case, determiners are always more external than modifiers, and we can say that, after the addition of modifiers, determiners **further** expand lexical categories.

Although two different operations are applied to lexical categories, from the point of view of contextualization there is only one target linguistic unit which, to a large extent, has a homogeneous syntactic behaviour at the level of the clause. That is, from the formal point of view, **Noun Phrase** non-terminals should result from the addition of modifiers to **Noun** non-terminals, which is described in previous Section 3.2.3.2, but they should also result from the addition of determiners to **Noun** —or **Noun Phrase**— non-terminals.

However, our AGFL grammar gives different names to the non-terminals that result from the addition of modifiers and to the non-terminals that result from the addition of determiners. There is merely a technical reason for this, consisting of the fact that current AGFL system does not support polymorphism, that is, it is not possible to have non-terminals which, having the same name and arity, are different only because they have different affixes (see Chapter 1). It must be clear, in consequence, that the different target non-terminal names are not at all intended to reflect different natures of addition of modifiers and addition of determiners, whose different behaviours are, in our opinion, sufficiently represented by the different functions and general AGFL frames that participate in their constitution (see Sections 3.2.3.2.2 and 3.2.3.3.2).

In this Section we simply connect non-terminals for lexical categories with target non-terminals for phrase categories expanded by means of addition of determiners. Our intention at the moment, as in previous Section 3.2.3.2.1 for addition of modifiers, is just to introduce these non-terminals to properly define from the formal point of view what is done by the process of addition of determiners —how it is done will be explained in Section 3.2.3.3.2.

- By means of the addition of determiners to non-terminals for lexical category noun, **Noun/3**, the following non-terminal for NPs must be obtained:

Nominal Phrase (ntype, premodifier, dttype, dtsubtype,
definiteness, person, gender, number)

- By means of the addition of determiners to non-terminals for lexical category non-personal pronoun, **Pronoun/5,6**, the following non-terminal for PrPs must be obtained:

Pronominal Phrase (premodifier, dttype, dtsubtype, prtype,
prsubtype, definiteness, person, gender, number)

- By means of the addition of determiners to non-terminals for lexical category personal pronoun, **Pronoun/7**, the following non-terminal for PrPs must be obtained:

Pronominal Phrase (premodifier, dttype, dtsubtype, prtype,
prsubtype, definiteness, person, gender, number, case)

3.2.3.3.2 An AGFL General Frame for the Addition of Determiners

Two main factors determine the way addition of determiners is formalized:

- Addition of determiners is carried out by means of right recursive application of subordination.
- The recursive addition of determiners is not free. Determiners must be organized according to very heterogeneous and vague rules that it is difficult to describe as a system. From a formal point of view, these rules are considered lexically motivated.

These factors, except for the much higher heterogeneous character of the rules that restrict concatenation of determiners, are not different from those which modulate addition of modifiers (see Section 3.2.3.2.2). As a matter of fact, according to these factors, addition of determiners must also be formalized as a recursive subordinating process, within the general scheme proposed in Chapter 1 for the AGFL formalization of subordination. Besides, addition of determiners must also

be controlled in order to guarantee the correct concatenation of determiners. Nevertheless, because of the much higher heterogeneous character of the rules that restrict concatenation of determiners, the way this is formalized is quite different from the way the same problem was formalized in the case of addition of modifiers. While affixes *premodifier* and *postmodifier*, conveniently handled within the recursive structure that accounts for the addition of modifiers was enough to control the concatenation of modifiers, various predicates are necessary, together with affixes *dttype*, *dsubtype*, *Ldttype*, *Ldsubtype*, to control the concatenation of determiners.

Given the fact that, in the case of addition of determiners, we decided to assign function name **NOMINAL** to the non-terminals that play the role of nucleus and function name **DETERMINER** to the non-terminals that play the role of modifiers, our AGFL general frame for the formalization of addition of determiners can be represented as follows:

```

structure ( dttype, dsubtype ):
  DETERMINER ( dttype, dsubtype, PRE ),
  NOMINAL I,
  Top Determiner Level ( dttype, dsubtype, ONE ).

structure I ( dttype, dsubtype ):
  DETERMINER ( dttype, dsubtype, PRE ),
  NOMINAL I.

NOMINAL I:
  lexical category;
  modified structure.

```

According to these rules, **structure** is constituted by one **DETERMINER** followed by one **NOMINAL I**. Predicate **Top Determiner Level** checks *dttype* and *dsubtype* to control which determiners may appear within **structure** when there is only one determiner. Non-terminal **structure I** is also constituted by one **DETERMINER** followed by one **NOMINAL I**, but no predicate restricts the possible values of *dttype* and *dsubtype*. **NOMINAL I** is in turn constituted by **lexical category** or **modified structure**. There is not recursion anywhere.

```

structure ( dttype, dsubtype ):
  DETERMINER ( dttype, dsubtype, PRE ),
  NOMINAL II ( I_dttype, I_dsubtype ),
  Top Determiner Level ( dttype, dsubtype, TWO ),

```

```

Determiner Transition Level ( I_ddtype, dttype,
                             I_dsubtype, dsubtype, ONE ).

structure II ( dttype, dsubtype ):
  DETERMINER ( dttype, dsubtype, PRE ),
  NOMINAL II ( I_ddtype, I_dsubtype ),
  Determiner Transition Level ( I_ddtype, dttype,
                               I_dsubtype, dsubtype, ONE ).

NOMINAL II ( dttype, dsubtype ):
  structure I ( dttype, dsubtype ).

```

According to these rules, **structure** is constituted by one **DETERMINER** followed by one **NOMINAL II**. Predicate **Top Determiner Level** checks *dttype* and *dsubtype* to control which determiners may close **structure** when there are two determiners. Non-terminal **structure II** is also constituted by one **DETERMINER** followed by one **NOMINAL II**, but predicate **Top Determiner Level** is not used to check values of *dttype* and *dsubtype*. In both rules, predicate **Determiner Transition Level** checks *dttype*, *dsubtype*, *Ldtype* and *Ldsubtype* to control the concatenation of two determiners. **NOMINAL II** is constituted by **structure I** described above, so there is not recursion anywhere, but two determiners need to be concatenated if **structure** or **structure II** are rewritten by means of these rules.

This scheme can be extended for as much levels as necessary, to allow for the concatenation of three, four or more determiners. From the formal point of view, strict recursion disappears, because each level of recursion is assigned a different non-terminal name (**structure I**, **NOMINAL I**, **structure II**, **NOMINAL II**...). This way we always know in which level of recursion a new determiner is added, and we can control whether this addition is possible or not, taking into account both the level of recursion and the determiner previously added. This is done by means of predicate **Determiner Transition Level** (or similar predicates with different names).

Yet another advantage of this *rule schemata* is that it draws a distinction between the non-terminals **I**, **II**, **III**..., which account for recursion, and the non-terminals **structure**, which are the only ones that play a role beyond the recursive addition of determiners itself. This distinction permits to control, in the case of non-terminals **structure**, whether the last determiner found at a certain

level of recursion can close the concatenation at this level or not, while, in the case of non-terminals **I, II, III...**, only the combination of determiners at the level in question is checked. An additional predicate, **Top Determiner Level**, is used in the rewrite rules of non-terminals *structure* for the first task, while for the second, as in the rewrite rules of non-terminals **I, II, III...**, predicate **Determiner Transition Level** is used again.

Obviously, the number and type of concrete combinations of determiners considered acceptable for each level of recursion—that is, the rewrite rules of predicate **Determiner Transition Level**—, as well as the type and subtype of determiners that are allowed to close the sequence of determiners at the level in question—that is, the rewrite rules of predicate **Top Determiner Level**—, are the kind of decisions that typically concern second level of formalization.

Within such a scheme of formalization, the information that is not relevant beyond the level of **structure** itself is directly removed from **structure**, while carried on by all non-terminals **structure I, II...** This means that no additional extensions—unlike for addition of modifiers—are needed to remove the non-(further-)relevant affixes.

On the contrary, an extension is necessary to account for the possibility that determined NPs have of being modified by certain types of premodifiers. This extension consists of a sort of insertion of the AGFL General Frame for Addition of Modifiers, described in Section 3.2.3.2.2, into the AGFL General Frame for Addition of Determiners, described above in this Section.

By way of this insertion, the affix *premodifier* is introduced in **structure** non-terminals, and **structure I, II...** non-terminals are enabled to constitute, together with **NOMINAL** functions, **NUCLEUS** functions. The resulting extended AGFL General Frame for Addition of Determiners is as follows:

```
structure ( NOTHING, dttype, dtsubtype ):
  DETERMINER ( dttype, dtsubtype, PRE ),
  NOMINAL I,
  Top Determiner Level ( dttype, dtsubtype, ONE ).
```

```
structure I ( dttype, dtsubtype ):
  DETERMINER ( dttype, dtsubtype, PRE ),
  NOMINAL I.
```

```
NOMINAL I:
```

```
lexical category;  
modified structure.
```

```
structure ( DEGREE, dttype, dtsubtype ):  
  Degree MODIFIER ( avtype, avsubtype, CH_comp, degree),  
  NUCLEUS I.
```

```
NUCLEUS I:  
  structure I ( dttype, dtsubtype ).
```

Non-terminal **structure** is assigned value **NOTHING** for premodifier when constituted by **DETERMINER** and **NOMINAL I**. It is assigned value **DEGREE** when constituted by **Degree MODIFIER** and **NUCLEUS I**. **NUCLEUS I** is itself constituted by **structure I**. This extension can be carried out for each level of recursion.

3.2.3.3.3 DETERMINER Function

3.2.3.3.3.1 DETERMINER/6 The rewrite rule for **DETERMINER/6** is the following:

```
DETERMINER ( dttype, dtsubtype, position, definiteness, gender,  
  number ):  
  Determiner ( dttype, dtsubtype, position, definiteness, gender,  
    number );  
  Determiner ( dttype, dtsubtype, position, definiteness,  
    inherent person, gender, number ).
```

The first alternative accounts for all determiners except possessive determiners, the second accounts for possessive determiners.

3.2.3.3.3.2 DETERMINER AND MODIFIER/9,11 As was observed in Section 3.1.3.1, certain determiners might be postponed to the noun, thus interrupting the natural sequence of noun and modifiers, given that possible modifiers of the noun are in such cases placed after the postponed determiner. To account for this phenomenon, we had to introduce non-terminals for what we consider *false functions* —**DETERMINER AND MODIFIER**, **PostDeterminer and Modifier Category NUCLEUS**— and *false categories* —**PostDeterminer and Modifier Category**—. These non-terminals, which identify sequences of postponed determiners followed by modifiers, are after combined with the noun

and other possible previous determiners exactly the same as unmodified postponed determiners do (see Sections 3.2.3.3.4.1 and 3.2.3.3.5.1).

The identification of sequences of determiners and modifiers is formalized according to the AGFL General Frame for Addition of Modifiers described in Sections 3.2.3.2.2 and 3.2.3.2.3.

The rewrite rules of **DETERMINER AND MODIFIER/9,11** are the following:

```
DETERMINER AND MODIFIER ( ddtype, dtsubtype, definiteness,
    recursive_post, CH_comp, degree, person, gender, number ):
  Postdeterminer and Modifier Category ( ddtype, dtsubtype,
    definiteness, recursive_post, CH_comp, degree, person,
    gender, number ).
```

```
DETERMINER AND MODIFIER ( prtype, prsubtype, ddtype, dtsubtype,
    definiteness, recursive_post, CH_comp, degree, person, gender,
    number ):
  Postdeterminer and Modifier Category ( prtype, prsubtype, ddtype,
    dtsubtype, definiteness, recursive_post, CH_comp, degree,
    person, gender, number ).
```

The first rule is used within the NP module, that is, **DETERMINER AND MODIFIER/9** non-terminals are always postponed to nouns. The second rule is used within the PrP module, that is, **DETERMINER AND MODIFIER/11** are always postponed to pronouns. Except for the introduction of affixes *prtype* and *prsubtype*, needed to enable, in PrPs with postponed determiners and modifiers, the specification of the same requirements between pronoun types and recursive modifiers specified in regularly modified PrPs (see Section 3.2.3.2.5), the two rules are equal. The two entire submodules, in fact, respectively devoted to **Postdeterminer and Modifier Category/9** and to **Postdeterminer and Modifier Category/11** are equal, so we will only include here the rules that account for the former.

3.2.3.3.3.3 HIDDEN RECURSIVE PostDeterminer and Modifier Category/9 The rewrite rule of non-terminal **HIDDEN RECURSIVE PostDeterminer and Modifier Category/9** is the following:

```
HIDDEN RECURSIVE Postdeterminer and Modifier Category ( ddtype,
    dtsubtype, definiteness, recursive_post, CH_comp, degree,
    person, gender, number ):
```

```

Postdeterminer and Modifier Category ( dttype, dtsubtype,
    definiteness, recursive_post, CH_comp, degree, person,
    gender, number );
HIDDEN FINAL Postdeterminer and Modifier Category ( dttype,
    dtsubtype, definiteness, recursive_post, CH_comp, degree,
    person, gender, number ).

```

3.2.3.3.3.4 PostDeterminer and Modifier Category/9 The rewrite rules of non-terminal **PostDeterminer and Modifier Category/9** are the following:

- **No MODIFIER** This non-terminal is never called upon alone—that is, constituting an alternative by itself—from outside of the recursive module of analysis of **Postdeterminer and Modifier Category**. For this reason, no examples can be, strictly speaking, supplied for this rule, given that any NP followed by just a postdeterminer, *persona* alguna, for instance, should be analysed by means of the rules described in Section 3.2.3.3.4 below alone, that is, not involving at all the submodule in question here, which accounts exclusively for *postdeterminers when these are followed by modifiers*.

```

HIDDEN FINAL Postdeterminer and Modifier Category ( dttype,
    dtsubtype, definiteness, NOTHING, NIL, NIL, person, gender,
    number ):
Determiner ( dttype, I_dtsubtype, POST, definiteness, gender,
    number ).

```

- **Adjective Phrase MODIFIER Example:** *persona* alguna más egoísta

```

Postdeterminer and Modifier Category ( dttype, dtsubtype,
    definiteness, AJPHRASE, CH_comp, degree+, person, gender,
    number ):
Postdeterminer and Modifier Category NUCLEUS ( dttype, dtsubtype,
    definiteness, AJPHRASE, NIL, NIL, person, gender, number ),
Adjective Phrase MODIFIER ( QUAL|PPLE, CH_comp, degree+,
    gender, number ).

```

- **Prepositional Phrase MODIFIER Example:** *un día* más de vacaciones

```

Postdeterminer and Modifier Category ( dttype, dtsubtype,
    definiteness, PRPPHRASE, CH_comp, degree, person, gender,
    number ):
Postdeterminer and Modifier Category NUCLEUS ( dttype, dtsubtype,

```

definiteness, PRPPHRASE, CH_comp, degree, person, gender,
number),
Prepositional Phrase MODIFIER (pretype).

• **Relative Clause MODIFIER Example:** *una persona más que nos deja*

Postdeterminer and Modifier Category (dttype, dtsubtype,
definiteness, RELCLAUSE, CH_comp, degree, person, gender,
number):
Postdeterminer and Modifier Category NUCLEUS (dttype, dtsubtype,
definiteness, RELCLAUSE, CH_comp, degree, person, gender,
number),
(Relative Clause MODIFIER (AGR_MOD, person, gender, number);
Relative Clause MODIFIER (no_AdvP)).

3.2.3.3.3.5 PostDeterminer and Modifier Category NUCLEUS/9

The rewrite rules of non-terminal **PostDeterminer and Modifier Category NUCLEUS/9** are the following:

Postdeterminer and Modifier Category NUCLEUS (dttype, dtsubtype,
definiteness, RELCLAUSE, CH_comp, degree, person, gender,
number):
HIDDEN RECURSIVE Postdeterminer and Modifier Category (dttype,
dtsubtype, definiteness, no_RELCLAUSE, CH_comp, degree,
person, gender, number).

Postdeterminer and Modifier Category NUCLEUS (dttype, dtsubtype,
definiteness, PRPPHRASE, CH_comp, degree, person, gender,
number):
HIDDEN RECURSIVE Postdeterminer and Modifier Category (dttype,
dtsubtype, definiteness, no_PRPPHRASE, CH_comp, degree,
person, gender, number).

Postdeterminer and Modifier Category NUCLEUS (dttype, dtsubtype,
definiteness, AJPHRASE, CH_comp, degree, person, gender,
number):
HIDDEN RECURSIVE Postdeterminer and Modifier Category (dttype,
dtsubtype, definiteness, NOTHING, CH_comp, degree, person,
gender, number).

3.2.3.3.4 Addition of Determiners in Noun Phrase This Section and the following one will show how the AGFL General Frame for Addition of Determiners, introduced in previous Section 3.2.3.3.2, is concretely applied for the formalization of addition of determiners in both the contexts of NPs and of PrPs. As in the case of addition of modifiers, taking into account that productivity of the operation

is larger in NPs, these will be more exhaustively described than PrPs. Both the descriptions of NPs and of PrPs will be, however, restricted to only first and second levels of recursion²⁷.

3.2.3.3.4.1 Nominal Phrase/8 The rewrite rules of non-terminal **Nominal Phrase/8** are the following:

- **Level ONE** The rules that add determiners must be broken down into different rules if we want to ensure that only NPs introduced by **COMPTVE** determiners will be allowed to specify a **Second Term MODIFIER**. Three rules are thus necessary: one for **COMPTVE** determiners themselves, another for *no_COMPTVE* determiners and the other for *dtno_general* determiners, which are not contained in the *no_COMPTVE* domain.

Example: *el mejor tiempo*

```
Nominal Phrase ( ntype, NOTHING, dtno_COMPTVE, dtsubtype,
    definiteness, person, gender, number ):
    DETERMINER ( dtno_COMPTVE, dtsubtype, PRE, definiteness, gender,
        number ),
    Nominal Phrase NOMINAL I ( ntype, premodifier, postmodifier,
        CH_comp, degree, person, gender, number ),
    eval character degree ( dtsubtype, CH_comp, degree ),
    Top Determiner Level ( number, dtno_COMPTVE, dtsubtype, ONE ).
```

As regards the predicate **eval character degree** in this rule, this is the element of the grammar ultimately responsible for the assignment of a value for affix *CH_comp*, taking into account the value of affix *dtsubtype*. The rewrite rules of this predicate are the following:

```
eval character degree ( EL, GENERAL|GENERAL+|PARTICULAR+, NoEQUAL ): .
eval character degree ( EL, NIL, NIL|ZERO|POS|POS+|POS++|POS+++|
    EQUAL|SUP ): .
eval character degree ( UNO, NIL|PARTICULAR|PARTICULAR+, degree ): .
eval character degree ( no_UNO, NIL|PARTICULAR|PARTICULAR+,
    degree ): .
eval character degree ( no_card, NIL|PARTICULAR|PARTICULAR+,
    degree ): .
eval character degree ( dtdem, NIL|PARTICULAR|PARTICULAR+, degree ): .
```

²⁷The grammar currently accepts four levels for NPs, but the formalization of the third and fourth levels can be easily inferred from the description of the first and second levels.

```

eval character degree ( dtquant, NIL|PARTICULAR|PARTICULAR+,
    degree ):.
eval character degree ( dtindef, NIL|PARTICULAR|PARTICULAR+,
    degree ):.
eval character degree ( NIL, NIL|PARTICULAR|PARTICULAR+, degree ):.

```

According to this predicate, values in the domain of affix non-terminal *general* are only matched if affix *dsubtype* is assigned value **EL** (value for *degree* must be **NoEQUAL**). Otherwise the value of affix *CH_comp* must either be **NIL** or it must be found in the domain of affix non-terminal *particular*. In case *CH_comp* is still an open value (see Section 3.2.3.2.4.2, epigraph **Degree MODIFIER**), predicate **eval character degree** assigns value **GENERAL** if affix *dsubtype* has value **EL**, otherwise it assigns value **PARTICULAR**.

On the other hand, as was observed in Section 3.1.3.1, not any determiner alone can introduce an NP. Moreover, certain determiners are allowed to introduce it only when they have a certain value for *number*. Predicate **Top Determiner Level** accounts for these facts. The rewrite rule of predicate **Top Determiner Level** is the following:

```

Top Determiner Level ( number, dttype, dsubtype, ONE ):
  { dttype :: INTG|REL|COMPTVE|ART|POSS|DEM|QUANT|DIS|INDEF|CARD };
  { dttype :: PART }, { dsubtype :: MEDIO };
  { number :: SING }, { dttype :: GLOBAL }.

```

Obviously, similar definitions of this predicate exist for each level of recursion, we will no longer include them in future explanations.

Example: el *área*

```

Nominal Phrase ( ntype, NOTHING, dtno_COMPTVE, dsubtype,
    definiteness, person, FEM, number ):
  DETERMINER ( dtno_COMPTVE, dsubtype, PRE, definiteness, gender,
    number ),
  Nominal Phrase NOMINAL I ( ntype, premodifier, postmodifier,
    CH_comp, degree, person, A_FEM, number ),
  eval false gender agreement ( dsubtype, gender, number ),
  eval character degree ( dsubtype, CH_comp, degree ),
  Top Determiner Level ( number, dtno_COMPTVE, dsubtype, ONE ).

```

This rule, which can also be included as a second alternative of the previous one, accounts for nouns beginning by [á-]. While **Nominal Phrase** non-terminal is

always assigned value **FEM** for *gender*, predicate **eval false gender agreement** ensures here that nouns beginning by [á-] will be combined with the adequate inflectional forms of determiners. The rewrite rules of this predicate are the following:

```
eval false gender agreement ( EL|UNO, MASC, SING ):.
eval false gender agreement ( EL|UNO, FEM, PLU ):.
eval false gender agreement ( NoUNIT+UNO, MASC, PLU ):.
eval false gender agreement ( NoUNIT_0-A+UNO, FEM_UN, PLU ):.
eval false gender agreement ( dtDEM, FEM, SING|PLU ):.
eval false gender agreement ( dtQUANT, FEM, SING|PLU ):.
eval false gender agreement ( no_CARD, FEM, SING|PLU ):.
eval false gender agreement ( dtINDEF, FEM, SING|PLU ):.
eval false gender agreement ( NoUNIT|NIL, FEM, SING|PLU ):.
```

According to this predicate, the following decisions are formalized:

- If *dtsubtype* is **EL** or **UNO** and *number* is **SING**, *gender* must be **MASC**, so *el/un área* are accepted.
- If *dtsubtype* is **EL** or **UNO** and *number* is **PLU**, *gender* must be **FEM**, so *las/unas áreas* are accepted.
- If *dtsubtype* is **NoUNIT+UNO**, *number* is always **PLU** and *gender* must be **MASC**, so *ciento un áreas* is accepted.
- If *dtsubtype* is **NoUNIT_0-A+UNO**, *number* is always **PLU** and *gender* must be **FEM_UN**, so *doscientas un áreas* is accepted.
- Otherwise, determiner must have value **FEM** for *gender*, while *number* can be either **SING** or **PLU**.

Examples: *qué día, qué alma cándida*

```
Nominal Phrase ( ntype, NOTHING, dtno_general, dtsubtype,
  definiteness, person, gender, number ):
  DETERMINER ( dtno_general, dtsubtype, PRE, definiteness, gender,
    number ),
  Nominal Phrase NOMINAL I ( ntype, premodifier, postmodifier,
    NIL|PARTICULAR|PARTICULAR+, degree, person, gender,
    number ),
  Top Determiner Level ( number, dtno_general, dtsubtype, ONE );
{ gender :: FEM },
  DETERMINER ( dtno_general, dtsubtype, PRE, definiteness, gender,
```

```

    number ),
Nominal Phrase NOMINAL I ( ntype, NOTHING, postmodifier,
    NIL|PARTICULAR|PARTICULAR+, degree, person, A_FEM,
    number ),
Top Determiner Level ( number, dtno_general, dtsubtype, ONE ).

```

Value **EL** for *dtsubtype* is not associated with any determiner type in the domain of *dtno_general*, so predicate **eval character degree** is not necessary in this case and, therefore, we can simply assign values **NIL|PARTICULAR|PARTICULAR+** to affix *CH.comp*. The second alternative accounts for nouns beginning by [á-]. This alternative is necessary to enable these nouns to combine with feminine determiners, thus constituting regular feminine NPs (see Section 3.2.3.2.4.2, epigraph **Adjective MODIFIER**). A similar alternative exists for all the remaining rules in the first level of recursion —and for their counterparts for **Nominal Phrase I**—. We will omit these alternatives in the following rules.

Example: *más tiempo*

```

Nominal Phrase ( ntype, NOTHING, COMPTVE, comp, definiteness, person,
    gender, number ):
  DETERMINER ( COMPTVE, comp, PRE, definiteness, gender, number ),
  Nominal Phrase NOMINAL I ( ntype, premodifier, postmodifier,
    NIL|PARTICULAR|PARTICULAR+, degree, person, gender,
    number ),
  Top Determiner Level ( number, COMPTVE, comp, ONE ),
  [ Second Term MODIFIER ( second_term_type, NIL|PARTICULAR+,
    comp, QUE|COMO );
    Second Term MODIFIER ( COMPTVE, PARTICULAR+, comp, DE,
    person, gender, number ) ].

```

CH.comp values in the domain of affix non-terminal *general* are not relevant for NPs introduced by **COMPTVE** determiners, so *CH.comp* is here fully specified. Examples of **Second Term MODIFIER** covered by this rule are the following: *tantos libros como en la biblioteca, más gente de la que cabía esperar*.

Examples: *esperanza alguna, esperanza alguna de triunfar*

```

Nominal Phrase ( ntype, NOTHING, dttype, dtsubtype, definiteness,
    person, gender, number ):
  Nominal Phrase NOMINAL I ( ntype, NOTHING|ADJECTIVE,
    NOTHING|AJPHRASE, NIL, NIL|POS, person, gender, number ),
  DETERMINER ( dttype, dtsubtype, POST, definiteness, gender,

```

```

        number ),
    Top PostDeterminer Level ( number, dttype, dtsubtype, ONE );
Nominal Phrase NOMINAL I ( ntype, NOTHING|ADJECTIVE,
    NOTHING|AJPHRASE, NIL, NIL|POS, person, gender, number ),
    DETERMINER AND MODIFIER ( dttype, dtsubtype, definiteness,
        recursive_post, NIL|PARTICULAR|PARTICULAR+, degree,
        person, gender, number ),
    Top PostDeterminer Level ( number, dttype, dtsubtype, ONE ).

```

Values **NOTHING|ADJECTIVE** for affix *premodifier*, values **NOTHING|AJPHRASE** for affix *postmodifier*, **NIL** for affix *CH_comp* and **NIL|POS** for affix *degree*, found in the non-terminal **Nominal Phrase NOMINAL I**, ensure that only adjective premodifiers or AjP postmodifiers, not graded, are accepted beside the noun when a determiner is postponed. Other modifiers are only accepted if postponed to **DETERMINER**, the second alternative in this rule accounts for this fact. These are, however, decisions that concern second level of formalization and can be easily changed without affecting the general structure designed to account for postdeterminers. Predicate **Top PostDeterminer Level** is the same as **Top Determiner Level**, but, obviously, it is associated with determiners postponed to the noun.

Example: *muy pocos días*

```

Nominal Phrase ( ntype, DEGREE, QUANT, POCO, definiteness, person,
    gender, number ):
    Degree MODIFIER ( avtype, avsubtype, NIL, pos+++ ),
        Nominal Phrase NUCLEUS I ( QUANT, POCO, definiteness, ntype,
            premodifier, postmodifier, CH_comp, degree, person,
            gender, number ),
        [ Second Term MODIFIER ( second_term_type, NIL, pos+++,
            QUE|COMO ) ].

```

Example: *qué pocos días*

```

Nominal Phrase ( ntype, EXCL, QUANT, POCO, definiteness, person,
    gender, number ):
    Degree MODIFIER ( EXCL, NIL, CH_comp, NIL ),
        Nominal Phrase NUCLEUS I ( QUANT, POCO, definiteness, ntype,
            premodifier, postmodifier, CH_comp1, degree, person,
            gender, number ).

```

Example: mil veces *más tiempo*

```
Nominal Phrase ( ntype, FACTOR, COMPTVE, comp, definiteness, person,
gender, number ):
  Factor MODIFIER ( CH_comp, degree ),
  Nominal Phrase NUCLEUS I ( COMPTVE, comp, definiteness, ntype,
premodifier, postmodifier, CH_comp1, degree1, person,
gender, number ),
  [ Second Term MODIFIER ( second_term_type, CH_comp, degree,
QUE|COMO ) ].
```

The three last rules account for predetermined and modified NPs. Only few modifiers are accepted in this context, that we have formalized for first level of recursion alone (but, again, these are the kind of adjustable decisions that concern second level of formalization and can be easily modified).

- **Level TWO** As **COMPTVE** determiners cannot be placed in the syntactic position of non-first-level determiners, it is no longer necessary to break down the rules that account for the addition of determiners at any level of analysis higher than the first.

Example: los *dos ladrones*

```
Nominal Phrase ( ntype, NOTHING, dttype, dtsubtype, definiteness,
person, gender, number ):
  DETERMINER ( dttype, dtsubtype, PRE, definiteness, gender,
number ),
  Nominal Phrase NOMINAL II ( I_dttype, I_dtsubtype, ntype,
premodifier, postmodifier, CH_comp, degree,
definiteness1, person, gender, number ),
  Top Determiner Level ( number, dttype, dtsubtype, TWO ),
  Determiner Transition Level ( number, I_dttype, dttype,
I_dtsubtype, dtsubtype, ONE ).
```

As regards the predicate **Determiner Transition Level**, it is here used to impose restrictions on the possible combinations of two determiners, taking also into account the value of affix *number*. The rewrite rule of this predicate is the following²⁸:

```
Determiner Transition Level ( number, I_dttype, dttype, I_dtsubtype,
dtsubtype, ONE ):
```

²⁸We include here only two alternatives, but this is, obviously, a much longer predicate.

```

(...);
{ I_dtttype :: CARD }, { dttype :: ART },
  { I_dsubtype :: NoUNIT|NoUNIT+UNO|NoUNIT_0-A+UNO },
  { dsubtype :: EL };
{ I_dtttype :: INDEF }, { dttype :: ART },
  { I_dsubtype :: CIERTO }, { dsubtype :: UNO };
(...).

```

According to these alternatives of this predicate, NPs such as *las tres Marías* or *una cierta persona*, for instance, are accepted.

Conversely, **COMPTVE** determiners might be postponed to nouns preceded by certain determiners, and therefore the rules for postponed determiners at second level of recursion must be broken down into two different rules.

Examples: *la casa esta, el niño ese de los tirantes*

```

Nominal Phrase ( ntype, NOTHING, I_dtttype, I_dsubtype,
  definiteness, person, gender, number ):
  Nominal Phrase NOMINAL II ( I_dtttype, I_dsubtype, ntype,
    NOTHING|ADJECTIVE, NOTHING|AJPHRASE, NIL, NIL|POS,
    definiteness, person, gender, number ),
  DETERMINER ( dtno_COMPTVE, dsubtype, POST, definiteness1,
    gender, number ),
  Top PostDeterminer Level ( number, dttype, dsubtype, TWO ),
  PostDeterminer Transition Level ( number, I_dtttype,
    dtno_COMPTVE, I_dsubtype, dsubtype, ONE );
Nominal Phrase NOMINAL II ( I_dtttype, I_dsubtype, ntype,
  NOTHING|ADJECTIVE, NOTHING|AJPHRASE, NIL, NIL|POS,
  definiteness, person, gender, number ),
  DETERMINER AND MODIFIER ( dtno_COMPTVE, dsubtype,
    definiteness1, recursive_post, NIL|PARTICULAR|PARTICULAR+,
    degree, person, gender, number ),
  Top PostDeterminer Level ( number, dttype, dsubtype, TWO ),
  PostDeterminer Transition Level ( number, I_dtttype,
    dtno_COMPTVE, I_dsubtype, dsubtype, ONE ).

```

Just note here —the corresponding explanation will be supplied with the following rule— that the values for *dttype*, *dsubtype* and *definiteness* assigned to **Nominal Phrase** are those provided by **Nominal Phrase NOMINAL II**, that is, those provided by the last predeterminer added by the recursive application of the operation of addition of determiners —even if postdeterminers are always analysed as more external determiners than the most external predeterminer found in the input phrase.

The predicate **PostDeterminer Transition Level** is similar to predicate **Determiner Transition Level**, but while the latter checks the combination of predeterminers, the former checks the combination of the highest level predeterminer placed before the noun with a possible postdeterminer following it.

Examples: *dos días más, dos litros más de vino que de agua*

```

Nominal Phrase ( ntype, NOTHING, I_ddtype, I_dtsubtype,
  definiteness, person, gender, number ):
  Nominal Phrase NOMINAL II ( I_ddtype, I_dtsubtype, ntype,
    NOTHING|ADJECTIVE, NOTHING|AJPHRASE, NIL, NIL|POS,
    definiteness, person, gender, number ),
  DETERMINER ( COMPTVE, dtsubtype, POST, definiteness1, gender,
    number ),
  [ Second Term MODIFIER ( second_term_type, PARTICULAR+, comp,
    QUE );
    Second Term MODIFIER ( COMPTVE, PARTICULAR+, comp, DE, person,
    gender, number ) ],
  Top PostDeterminer Level ( number, dtype, dtsubtype, TWO ),
  PostDeterminer Transition Level ( number, I_ddtype, COMPTVE,
    I_dtsubtype, dtsubtype, ONE );
Nominal Phrase NOMINAL II ( I_ddtype, I_dtsubtype, ntype,
  NOTHING|ADJECTIVE, NOTHING|AJPHRASE, NIL, NIL|POS,
  definiteness, person, gender, number ),
  DETERMINER AND MODIFIER ( COMPTVE, dtsubtype, definiteness1,
    recursive_post, NIL|PARTICULAR|PARTICULAR+, degree,
    person, gender, number ),
  [ Second Term MODIFIER ( second_term_type, PARTICULAR+, comp,
    QUE );
    Second Term MODIFIER ( COMPTVE, PARTICULAR+, comp, DE,
    person, gender, number ) ],
  Top PostDeterminer Level ( number, dtype, dtsubtype, TWO ),
  PostDeterminer Transition Level ( number, I_ddtype, COMPTVE,
    I_dtsubtype, dtsubtype, ONE ).

```

By way of the preceding rule a phrase like *cuántos días más* should be analysed as an NP containing an interrogative predeterminer and a comparative, more external, postdeterminer. Because the affix values for *Ldtype*, *Ldtsubtype* and *definiteness* passed on to non-terminal **Nominal Phrase** are those of non-terminal **Nominal Phrase NOMINAL II**, that is, those of the highest level predeterminer, the whole NP can be properly characterized, from the point of view of its further contextualization, as an interrogative definite NP. That is, even if postdeterminers are more external determiners than predeterminers, these are those which supply the NPs with the adequate characterization for their inclusion in

higher level units of analysis; therefore, it is their values for the relevant affixes that we have to conserve for further integration of the NP in question.

3.2.3.3.4.2 Nominal Phrase I/12 Apart from the rules that account for the addition of postdeterminers, all the rules that account for **Nominal Phrases** and first level of recursion have a counterpart for non-terminal **Nominal Phrase I**. Except for the absence of predicate **Top Determiner Level**, these rules are equal to those described above. For this reason, we will only include one example rule, the remaining ones can be easily deduced from this:

```
Nominal Phrase I ( PRE, dtno_COMPTVE, dtsubtype, definiteness, ntype,
  premodifier, postmodifier, CH_comp, degree, person, gender,
  number ):
  DETERMINER ( dtno_COMPTVE, dtsubtype, PRE, definiteness, gender,
    number ),
  Nominal Phrase NOMINAL I ( ntype, premodifier, postmodifier,
    CH_comp, degree, person, gender, number ),
  eval character degree ( dtsubtype, CH_comp, degree ).
```

3.2.3.3.4.3 Nominal Phrase II/12 The same observation holds for **Nominal Phrase II** and second level of recursion of addition of determiners in **Nominal Phrase**: the rule for the latter has also a similar counterpart for the former, with only the difference that predicate **Top Determiner Level** has been removed.

Affix *definiteness*, on the other hand, has also disappeared from the non-terminal **Nominal Phrase II**. The reason for this is that, as no more postponed determiners might be added after this level of analysis (that is, at third level of recursion or beyond), we can be sure that the final value for *definiteness* will not be that of the determiner identified by means of this rule, but that of the last determiner found in the recursion (at least one more level of recursion must be present if a non-terminal of the type **structure I, II,...** is recognized, see Section 3.2.3.3.2).

Non-terminal **Nominal Phrase II** is called upon from **Nominal Phrase NOMINAL III** (see Section 3.2.3.3.2), but this rule will not be displayed here.

```
Nominal Phrase II ( PRE, dttype, dtsubtype, ntype, premodifier,
  postmodifier, CH_comp, degree, person, gender, number ):
  DETERMINER ( dttype, dtsubtype, PRE, definiteness, gender,
    number ),
  Nominal Phrase NOMINAL II ( I_dttype, I_dtsubtype, ntype,
    premodifier, postmodifier, CH_comp, degree,
```

definiteness1, person, gender, number),
Determiner Transition Level (number, I_ddtype, dtype,
I_dtsubtype, dsubtype, ONE).

3.2.3.3.4.4 Nominal Phrase NOMINAL I/8 The rewrite rule of non-terminal **Nominal Phrase NOMINAL I** is the following:

Nominal Phrase NOMINAL I (ntype, premodifier, postmodifier, CH_comp,
degree, person, gender, number):
HEAD Noun Phrase (ntype, premodifier, postmodifier, CH_comp,
degree, definiteness, person, gender, number).

3.2.3.3.4.5 Nominal Phrase NOMINAL II/11 The rewrite rule of non-terminal **Nominal Phrase NOMINAL II** is the following:

Nominal Phrase NOMINAL II (dtype, dsubtype, ntype, premodifier,
postmodifier, CH_comp, degree, definiteness, person, gender,
number):
Nominal Phrase I (dtype, dsubtype, definiteness, ntype,
premodifier, postmodifier, CH_comp, degree, person,
gender, number).

3.2.3.3.4.6 Nominal Phrase NUCLEUS I/11 The rewrite rule of non-terminal **Nominal Phrase NUCLEUS I** is the following:

Nominal Phrase NUCLEUS I (dtype, dsubtype, definiteness, ntype,
premodifier, postmodifier, CH_comp, degree, person, gender,
number):
Nominal Phrase I (dtype, dsubtype, definiteness, ntype,
premodifier, postmodifier, CH_comp, degree, person, gender,
number).

3.2.3.3.5 Addition of Determiners in Pronoun Phrase Except for the introduction of affix case, the rules that account for personal pronouns are equal to those which account for non-personal pronouns, the latter are the only ones displayed here.

3.2.3.3.5.1 Pronominal Phrase/9 The rewrite rules of non-terminal **Pronominal Phrase/9** are the following:

- **Level ONE** These rules —like those for **Nominal Phrase** or **Pronoun Phrase** and for the same reason— must be broken down into four different rules associated with different types of determiners.

Example: los *dos*

```
Pronominal Phrase ( NOTHING, dtno_COMPTE, dtsubtype, prtype,
  prsubtype, definiteness, person, gender, number ):
  DETERMINER ( dtno_COMPTE, dtsubtype, PRE, definiteness, gender,
    number ),
  Pronominal Phrase NOMINAL I ( prtype, prsubtype, premodifier,
    postmodifier, definiteness1, person, gender, number ),
  Pronominal Phrase Top Determiner Level ( gender, number,
    dtno_COMPTE, dtsubtype, ONE ),
  Pronoun Determiner Transition ( gender, number, prtype,
    prsubtype, dtno_COMPTE, dtsubtype ).
```

As regards predicate **Pronominal Phrase Top Determiner Level**, it is here used to control the types of determiners that can be alone when they precede pronouns. Apart from *dttype* and *dtsubtype*, this predicate must also check the values of affixes *gender* and *number* to account for the fact that some determiners can appear in this context only in certain inflectional forms. Currently, all determiners introduced by sequences of determiners and pronouns accepted by predicate **Pronoun Determiner Transition** (see below) can precede, alone, the corresponding pronoun, so the rewrite rule of predicate **Pronominal Phrase Top Determiner Level** is:

```
Pronominal Phrase Top Determiner Level ( gender, number, dttype,
  dtsubtype, ONE):.
```

In addition to predicate **Pronominal Phrase Top Determiner Level**, the predicate **Pronoun Determiner Transition** is used in this rule to control which sequences of determiners and pronouns are allowed. We only include here some example alternatives, but the predicate is much longer:

```
Pronoun Determiner Transition ( gender, number, prtype, prsubtype,
  dttype, dtsubtype ):
  (...);
  { prtype :: CARD }, { prsubtype :: NoUNIT },
    { dttype :: ART }, { dtsubtype :: EL };
  { prtype :: CARD }, { prsubtype :: NoUNIT },
    { dttype :: INDEF }, { dtsubtype :: OTRO };
  (...).
```

According to these alternatives, for instance, instances like *los tres, otros dos* are accepted.

Example: *cuantos más*

```
Pronominal Phrase ( NOTHING, dtno_general, dtsubtype, prtype,
  prsubtype, definiteness, person, gender, number ):
  DETERMINER ( dtno_general, dtsubtype, PRE, definiteness, gender,
    number ),
  Pronominal Phrase NOMINAL I ( prtype, prsubtype, premodifier,
    postmodifier, definiteness1, person, gender, number ),
  Pronominal Phrase Top Determiner Level ( gender, number,
    dtno_general, dtsubtype, ONE ),
  Pronoun Determiner Transition ( gender, number, prtype,
    prsubtype, dtno_general, dtsubtype ).
```

Example: *tantos otros*

```
Pronominal Phrase ( NOTHING, COMPTVE, comp, prtype, prsubtype,
  definiteness, person, gender, number ):
  DETERMINER ( COMPTVE, comp, PRE, definiteness, gender, number ),
  Pronominal Phrase NOMINAL I ( prtype, prsubtype, premodifier,
    postmodifier, definiteness1, person, gender, number ),
  Pronominal Phrase Top Determiner Level ( gender, number,
    COMPTVE, comp, ONE ),
  Pronoun Determiner Transition ( gender, number, prtype,
    prsubtype, COMPTVE, comp ),
  [ Second Term MODIFIER ( second_term_type, CH_comp, comp,
    QUE|COMO );
    { CH_comp :: PARTICULAR+ },
    Second Term MODIFIER ( COMPTVE, CH_comp, comp, DE,
      person, gender, number ) ].
```

Examples: *uno cualquiera, uno cualquiera de los otros*

```
Pronominal Phrase ( NOTHING, dtno_COMPTVE, dtsubtype, prtype,
  prsubtype, definiteness, person, gender, number ):
  Pronominal Phrase NOMINAL I ( prtype, prsubtype, premodifier,
    postmodifier, definiteness, person, gender, number ),
  DETERMINER ( dtno_COMPTVE, dtsubtype, POST, definiteness1,
    gender, number ),
  Pronominal Phrase Top PostDeterminer Level ( gender, number,
    dtno_COMPTVE, dtsubtype, ONE ),
  Pronoun PostDeterminer Transition ( gender, number, prtype,
    prsubtype, dtno_COMPTVE, dtsubtype );
  Pronominal Phrase NOMINAL I ( prtype, prsubtype, premodifier,
    postmodifier, definiteness, person, gender, number ),
```

```

DETERMINER AND MODIFIER ( prtype, prsubtype, dtno_COMPTVE,
    dtsubtype, definiteness1, recursive_post, CH_comp, degree,
    person, gender, number ),
Pronominal Phrase Top PostDeterminer Level ( gender, number,
    dtno_COMPTVE, dtsubtype, ONE ),
Pronoun PostDeterminer Transition ( gender, number, prtype,
    prsubtype, dtno_COMPTVE, dtsubtype ).

```

When determiners are postponed, the values for *definiteness* passed on to non-terminal **Pronominal Phrase** must be those supplied by **Pronominal Phrase NOMINAL I** if we want to be able to characterize instances like *quién más* as interrogative definite PrPs (see Section 3.2.3.3.4.1). A rule similar to this one exists also for personal pronouns (*vosotros mismos*, *nosotros dos*).

Examples: *dos más*, *dos más de los verdes*

```

Pronominal Phrase ( NOTHING, COMPTVE, comp, prtype, prsubtype,
    definiteness, person, gender, number ):
  Pronominal Phrase NOMINAL I ( prtype, prsubtype, premodifier,
    postmodifier, definiteness, person, gender, number ),
  DETERMINER ( COMPTVE, comp, POST, definiteness1, gender,
    number ),
  [ Second Term MODIFIER ( second_term_type, CH_comp, comp,
    QUE|COMO );
    { CH_comp :: PARTICULAR+ },
    Second Term MODIFIER ( COMPTVE, CH_comp, comp, DE, person,
    gender, number ) ],
  Pronominal Phrase Top PostDeterminer Level ( gender, number,
    COMPTVE, comp, ONE ),
  Pronoun PostDeterminer Transition ( gender, number, prtype,
    prsubtype, COMPTVE, comp );
Pronominal Phrase NOMINAL I ( prtype, prsubtype, premodifier,
    postmodifier, definiteness, person, gender, number ),
  DETERMINER AND MODIFIER ( prtype, prsubtype, COMPTVE, comp,
    definiteness1, recursive_post, CH_comp, degree, person,
    gender, number ),
  [ Second Term MODIFIER ( second_term_type, CH_comp, comp,
    QUE|COMO );
    { CH_comp :: PARTICULAR+ },
    Second Term MODIFIER ( COMPTVE, CH_comp, comp, DE, person,
    gender, number ) ],
  Pronominal Phrase Top PostDeterminer Level ( gender, number,
    COMPTVE, comp, ONE ),
  Pronoun PostDeterminer Transition ( gender, number, prtype,
    prsubtype, COMPTVE, comp ).

```

CH_comp values in the domain of *general* are not relevant for PrPs, so affix

CH_comp can be fully specified here. Examples of **Second Term MODIFIERs** recognized by this rule are the following: *dos más que la otra vez*, *dos más de los que habríamos deseado*.

Example: *muy pocos más*

```
Pronominal Phrase ( DEGREE, QUANT, POCO, prtype, prsubtype,
  definiteness, person, gender, number ):
{ pos++ :: POS++|EQUAL|SUP },
  Degree MODIFIER ( avtype, avsubtype, NIL, pos++ ),
  Pronominal Phrase NUCLEUS I ( QUANT, POCO, person, gender,
    number, prtype, prsubtype, premodifier, postmodifier ),
  [ Second Term MODIFIER ( second_term_type, NIL, pos++,
    QUE|COMO ) ].
```

Example: *qué pocos más*

```
Pronominal Phrase ( EXCL, QUANT, POCO, prtype, prsubtype,
  definiteness, person, gender, number ):
  Degree MODIFIER ( EXCL, NIL, CH_comp, NIL ),
  Pronominal Phrase NUCLEUS I ( QUANT, POCO, PRE, person, gender,
    number, prtype, prsubtype, premodifier, postmodifier ).
```

• **Level TWO Example:** *los otros dos*

```
Pronominal Phrase ( NOTHING, dttype, dtsubtype, prtype, prsubtype,
  definiteness, person, gender, number ):
  DETERMINER ( dttype, dtsubtype, PRE, definiteness, gender,
    number ),
  Pronominal Phrase NOMINAL II ( I_dttype, I_dtsubtype, person,
    gender, number, prtype, prsubtype, premodifier,
    postmodifier ),
  Pronominal Phrase Top Determiner Level ( gender, number,
    dttype, dtsubtype, TWO ),
  Pronominal Phrase Determiner Transition Level ( I_dttype,
    I_dtsubtype, dttype, dtsubtype, ONE ).
```

The predicate **Pronominal Phrase Determiner Transition Level** regulates the combination of determiners when more than one of these precede together a pronoun. There are few possibilities of combinations at the moment, but these are, anyway, the kind of specifications that concern second level of formalization and that can be enlarged without difficulty. The rewrite rule of this predicate is the following:

```

Pronominal Phrase Determiner Transition Level ( I_ddtype,
  I_dtsubtype, ddtype, dtsubtype, ONE ):
{ I_ddtype :: ART }, { I_dtsubtype :: EL },
  { ddtype :: GLOBAL }, { dtsubtype :: NIL };
{ I_ddtype :: INDEF }, { I_dtsubtype :: OTRO },
  { ddtype :: ART }, { dtsubtype :: EL }.

```

3.2.3.3.5.2 Pronominal Phrase I/9 The rewrite rule of non-terminal **Pronominal Phrase I/9** is the following:

```

Pronominal Phrase I ( ddtype, dtsubtype, person, gender, number,
  prtype, prsubtype, premodifier, postmodifier ):
  DETERMINER ( ddtype, dtsubtype, PRE, definiteness, gender,
    number ),
  Pronominal Phrase NOMINAL I ( prtype, prsubtype, premodifier,
    postmodifier, definiteness1, person, gender, number ),
  Pronoun Determiner Transition ( gender, number, prtype,
    prsubtype, ddtype, dtsubtype ).

```

Higher levels of recursion might only add predeterminers; as these must supply their value for affix *definiteness* to non-terminal **Pronominal Phrase**, affix *definiteness* can be removed from non-terminal **Pronominal Phrase I** or higher level non-terminals of the type **structure I, II...** (see Section 3.2.3.3.2) exclusively concerning recursion.

3.2.3.3.5.3 Pronominal Phrase NOMINAL I/7 The rewrite rule of non-terminal **Pronominal Phrase I/7** is the following:

```

Pronominal Phrase NOMINAL I ( prtype, prsubtype, premodifier,
  postmodifier, definiteness, person, gender, number ):
  HEAD Pronoun Phrase for NOMINAL ( premodifier, postmodifier,
    prtype, prsubtype, definiteness, person, gender, number ).

```

3.2.3.3.5.4 Pronominal Phrase NOMINAL II/9 The rewrite rule of non-terminal **Pronominal Phrase II/9** is the following:

```

Pronominal Phrase NOMINAL II ( ddtype, dtsubtype, person, gender,
  number, prtype, prsubtype, premodifier, postmodifier ):
  Pronominal Phrase I ( ddtype, dtsubtype, person, gender, number,
    prtype, prsubtype, premodifier, postmodifier ).

```

3.2.3.3.5.5 Pronominal Phrase NUCLEUS I/9 The rewrite rule of non-terminal **Pronominal Phrase NUCLEUS I/9** is the following:

Pronominal Phrase NUCLEUS I (dttype, dtsubtype, person, gender, number, prtype, prsubtype, premodifier, postmodifier):
Pronominal Phrase I (dttype, dtsubtype, person, gender, number, prtype, prsubtype, premodifier, postmodifier).

3.2.3.4 Modifiers

3.2.3.4.1 Modifiers Postponed to the Nucleus

3.2.3.4.1.1 Adjective Phrase MODIFIER/5 The rewrite rule of non-terminal **Adjective Phrase MODIFIER/5** is the following²⁹:

Adjective Phrase MODIFIER (ajtype, CH_comp, degree, gender, number):
HEAD Adjective Phrase (premodifier, ajtype, CH_comp, degree, gender, number).

3.2.3.4.1.2 Prepositional Phrase MODIFIER/1 The rewrite rule of non-terminal **Prepositional Phrase MODIFIER/1** is the following:

Prepositional Phrase MODIFIER (preposition):
HEAD Prepositional Adverb Phrase (premodifier, preposition, premodifier1, postmodifier, PLACE, avsubtype, CH_comp, degree);
Prepositional Noun Phrase (preposition, ntype, prmd_general, postmodifier, CH_comp, degree, definiteness, person, gender, number);
Prepositional Pronoun Phrase (preposition, premodifier, postmodifier, prgeneral, prsubtype, definiteness, person, gender, number);
HEAD Prepositional Pronoun Phrase (preposition, premodifier, postmodifier, prgeneral, prsubtype, definiteness, person, gender, number, case);
Prepositional Nominal Phrase (preposition, ntype, premodifier, dtgeneral, dtsubtype, definiteness, person, gender, number);
Prepositional Pronominal Phrase (preposition, premodifier, dtgeneral, dtsubtype, prgeneral, prsubtype, definiteness, person, gender, number);
Prepositional Pronominal Phrase (preposition, premodifier, dtgeneral, dtsubtype, prgeneral, prsubtype, definiteness, person, gender, number, case);
Prepositional Nominalization (preposition, MDcategory, preptype,

²⁹ **Adjective Phrase/6** was described in Section 3.2.3.2.7.

```

    definiteness, person, gender, number );
Prepositional Nominalization ( preposition, MDcategory,
    definiteness, person, gender, number );
Prepositional Relative Clause ( preposition, LINKED, person,
    gender, number );
Prepositional Clause ( preposition, QUE, non_imperative );
Prepositional Clause ( preposition, clause_type, INFINITIVE,
    person ).

```

At present, **Prepositional Phrase MODIFIER** keeps only information about the preposition that introduces the modifier. This non-terminal, however, will soon be enriched with additional affixes to integrate data about prepositional requirements of nouns and adjectives.

Prepositional versions of almost all linguistic structures might be used as prepositional modifiers. The rewrite rules for these prepositional versions of linguistic structures such as NPs or PrPs are described in Section 3.2.3.7.

3.2.3.4.1.3 Relative Clause MODIFIER/1,4 The rewrite rules of non-terminals **Relative Clause MODIFIER/1,4** are the following:

```

Relative Clause MODIFIER ( no_AGR_MOD ):
    Relative Clause ( no_AGR_MOD, preptype, person, gender, number );
    Relative Clause ( no_AGR_MOD, preptype, person, gender, number,
        MVperson ).

Relative Clause MODIFIER ( AGR_MOD, person, gender, number ):
    Relative Clause ( AGR_MOD, preptype, person, gender, number );
    Relative Clause ( AGR_MOD, preptype, person, gender, number,
        MVperson ).

```

Relative Clause MODIFIER/1 accounts for relative clauses introduced by relative NPs, PrPs or AvPs which do not agree for *person*, *gender* or *number* with their antecedent (*los padres cuyos hijos, la casa donde nació*). Relative clauses of this type can be distinguished from other relative clauses because they select values in the domain of affix non-terminal *no_AGR_MOD* for affix *RCfunction*.

Relative Clause MODIFIER/4, on the other hand, accounts for relative clauses introduced by relative NPs, PrPs or AvPs which do agree for *person*, *gender* and *number* with their antecedent (*el hermano que tengo, la mujer a quien amo*). Relative clauses of this type are distinguished from other relative clauses because they take value **AGR_MOD** for affix *RCfunction*.

The rewrite rules of the various types of clauses —of relative clauses too—, possible in Spanish will be described in Section 3.2.3.6. For the purposes of the current explanations, however, we can say here that relative clauses have rewrite rules with right-hand sides introduced by relative NPs, PrPs or AvPs that have a syntactic function within the relative clause. Or, in other words, we can say that relative clauses are introduced by **Relative SUBJECT**, **Relative DIRECT OBJECT**, **Relative INDIRECT OBJECT**, etc. or, yet making a further generalisation, we can³⁰ say that relative clauses are introduced by **Relative CONSTITUENTS**.

Now, as in this Section we are only considering relative clauses used as modifiers within syntactic constituents of main clauses —that is, having values in the domain of non-terminal *modifier* for affix *RCfunction*—, a rewrite rule for **Relative CONSTITUENT**, which should theoretically merge the rewrite rules of the real non-terminals **Relative SUBJECT**, **Relative DIRECT OBJECT**, etc., could be the following:

```
Relative CONSTITUENT ( RCfunction, preptype, person, gender,
    number ):
    Propositional Nominal Phrase ( preptype, ntype, prmd_general,
        REL, CUYO, definiteness, person, gender, number ),
        { RCfunction :: no_AdvP };
## cuyo trabajo
    Nominal Phrase ( ntype, prmd_general, REL, CUYO, definiteness,
        person, gender, number ),
        { RCfunction :: no_AdvP }, { preptype :: NIL };
## del que
    Propositional Pronominal Phrase ( preptype, prmd_general,
        GLOBAL|ART, NIL|EL, REL, QUE, definiteness, person, gender,
        number ),
        { RCfunction :: AGR_MOD };
    Propositional Pronominal Phrase ( preptype, prmd_general,
        GLOBAL|ART, NIL|EL, REL, CUAL, definiteness, person, gender,
        number ),
        { RCfunction :: AGR_MOD };
    HEAD Pronoun Phrase ( prmd_general, postmodifier, REL, QUE,
        definiteness, person, gender, number ),
        { RCfunction :: AGR_MOD }, { preptype :: NIL };
    Propositional Pronoun Phrase ( preptype, prmd_general,
        postmodifier, REL, QUIEN, definiteness, person, gender,
```

³⁰Just for the purposes of these explanations, since **Relative CONSTITUENT** is not a real non-terminal of the grammar, but an abstraction that in this exposition is used to refer simultaneously to all the real non-terminals of the grammar **Relative SUBJECT**, **Relative DIRECT OBJECT**, etc.

```

    number ),
  { RCfunction :: AGR_MOD };
HEAD Prepositional Adverb Phrase ( prmd_general, preptype,
  prmd_general1, postmodifier, REL, DONDE, NIL, NIL ),
  { RCfunction :: no_AdvP|AdvP },
  { person :: NIL }, { gender :: NIL }, { number :: NIL };
HEAD Adverb Phrase ( prmd_general, postmodifier, REL, QUE,
  CH_comp, degree ),
  { RCfunction :: AdvP }, { preptype :: NIL },
  { person :: NIL }, { gender :: NIL }, { number :: NIL };
HEAD Adverb Phrase ( prmd_general, postmodifier, REL, DONDE, NIL,
  NIL ),
  { RCfunction :: no_AdvP|AdvP }, { preptype :: NIL },
  { person :: NIL }, { gender :: NIL }, { number :: NIL };
HEAD Adverb Phrase ( prmd_general, postmodifier, REL, COMO, NIL,
  NIL ),
  { RCfunction :: no_AdvP|AdvP }, { preptype :: NIL },
  { person :: NIL }, { gender :: NIL }, { number :: NIL }.

```

As can be observed in this rule, by handling types and subtypes of determiners, pronouns and adverbs, the distinction between different relative phrases can be very accurate. On the basis of this detailed identification of the relative phrases that introduce relative clauses, one or more syntactic functions can be already proposed at this level of analysis for the whole relative clauses. This is done by means of guards which assign a value to affix *RCfunction* in each rewrite alternative in this rule, that is to say, by means of guards which assign a different value to affix *RCfunction* for each different relative phrase that can introduce a **Relative Clause MODIFIER**.

To give some examples, by means of second alternative, relative phrase *cuyo trabajo*—which introduces, for instance, the relative clause modifying the noun *persona* within the clause *La persona cuyo trabajo sobre la contaminación sea mejor optará a la beca*— is identified. The guard in the alternative assigns value **no_AdvP** to affix *RCfunction*, so the relative clause *cuyo trabajo sobre la contaminación sea mejor* will have to be dominated by a node **Relative Clause MODIFIER/1**. By means of third alternative, on the other hand, relative phrase *del que*—which introduces, for instance, the relative clause modifying the noun *hombre* within the clause *El hombre del que se ha enamorado no merece su abnegado amor*— is identified. The guard in the alternative assigns in this case value **AGR_MOD** to affix *RCfunction*, so the relative clause *del que se ha enamorado*

will have to be dominated by a node **Relative Clause MODIFIER/4**.

The collection of accepted relative phrases and their association with syntactic functions are adjustable decisions that concern second level of formalization. Affix *preptype* holds information about the preposition that introduces prepositional relative phrases that start a relative clause. This is necessary because this preposition might simultaneously act as the preposition that introduces the whole relative clause (thus constituting a prepositional relative clause). We will explain this in detail in Section 3.2.3.7.

3.2.3.4.1.4 Apposition MODIFIER/4,5 The rewrite rules of non-terminal **Apposition MODIFIER/4,5** are the following:

```
Apposition MODIFIER ( HOMO_APP, N, gender, number ) :  
  Noun ( proper, gender, number ).
```

```
Apposition MODIFIER ( HOMO_APP, Av, avtype, avsubtype, degree ) :  
  Adverb ( avtype, avsubtype, preptype, degree, +STRESS ).
```

```
Apposition MODIFIER ( HETERO_APP, Av, preposition ) :  
  Prepositional Nominal Phrase ( preposition, ntype, NOTHING,  
    dtgeneral, dtsubtype, definiteness, person, gender, number );  
  Prepositional Noun Phrase ( preposition, ntype, NOTHING,  
    postmodifier, CH_comp, degree, definiteness, person, gender,  
    number ).
```

The first rule accounts for appositions that modify common nouns (*la tía Paula*), the second and third ones for appositions that modify adverbs (*aquí atrás*, *aquí en mi tierra*).

3.2.3.4.1.5 Determiner MODIFIER/6 The rewrite rule of non-terminal **Determiner MODIFIER/6** is the following:

```
Determiner MODIFIER ( dttype, dtsubtype, position, definiteness,  
  gender, number ) :  
  Determiner ( dttype, dtsubtype, position, definiteness, gender,  
    number ).
```

3.2.3.4.2 Modifiers Preceding the Nucleus

3.2.3.4.2.1 Adjective MODIFIER/5 The rewrite rule of non-terminal **Adjective MODIFIER/5** is the following:

```
Adjective MODIFIER ( ajtype, CH_comp, degree, gender, number ) :
  ( { CH_comp :: NIL }, { degree :: POS|SUP };
    { CH_comp :: PARTICULAR|PARTICULAR+|GENERAL|GENERAL+ },
      { degree :: NoEQUAL } ),
  Adjective ( ajtype, PRE, degree, gender, number );
{ CH_comp :: NIL }, { degree :: POS }, { ajtype :: PPLE },
  Participle Verb ( mvtype, ACT|MIDD, mood, gender, number ).
```

Both adjectives and participles are accepted as **Adjective MODIFIERS**. The latter are assigned value **PPLE** for affix *ajtype* and value **POS** for affix *degree*. *CH_comp* values are always assigned according to the values of affix *degree*. **NoEQUAL** is associated with all the values in the domain of affix non-terminal *CH_comp+*, the rest of the values of *degree* are associated with value **NIL**.

3.2.3.4.2.2 Degree MODIFIER/4, Comparison Quantifier MODIFIER/4,5, Factor MODIFIER/2, Second Term MODIFIER/4,6,7 The rewrite rule of non-terminal **Degree MODIFIER/4** is the following:

Examples: *muy buena persona*, *muy pocos libros*, *muy pocos*, *muy tarde*, *muy bueno*

```
Degree MODIFIER ( avtype, avsubtype, CH_comp, degree ) :
  ( { avtype :: QUANT|EXCL }, { CH_comp :: NIL };
    { avtype :: COMPTVE }, { avsubtype :: EQUAL},
      { CH_comp :: NIL };
    { avtype :: COMPTVE }, { avsubtype :: NoEQUAL },
      { CH_comp :: PARTICULAR|PARTICULAR+|GENERAL|GENERAL+ } ),
  Adverb ( avtype, avsubtype, preptype, degree, -STRESS ).
```

Only **QUANT**, **COMPTVE** and **EXCL** adverbs with value **-STRESS** for affix *AVposition* are accepted as **Degree MODIFIERS**. *CH_comp* values are assigned according to the value of affix *degree*: value **NoEQUAL** is associated with all the values in the domain of affix non-terminal *CH_comp+*, whilst value **NIL** is associated with the rest of possible values of *degree*.

The rewrite rules of non-terminals **Comparison Quantifier MODIFIER/4,5** are the following:

Examples: mucho *más buena persona*, mucho *más tarde*, mucho *más interesante*

```
Comparison Quantifier MODIFIER ( avtype, avsubtype, NIL, degree ):
  { avtype :: QUANT|COMPTVE|REL },
  Adverb ( avtype, avsubtype, preptype, degree, +STRESS ).
```

Example: muchos años *más tarde*

```
Comparison Quantifier MODIFIER ( NIL, degree, person, gender, number ):
  { degree :: NIL },
  ( Nominal Phrase ( ntype, NOTHING, QUANT,
    VARIOS|POCO|MUCHO|BASTANTE, definiteness, person,
    gender, number );
    Nominal Phrase ( ntype, NOTHING, CARD, dtsubtype,
    definiteness, person, gender, number );
    Nominal Phrase ( ntype, NOTHING, PART, MEDIO, definiteness,
    person, gender, number );
    Noun Phrase ( ntype, NOTHING, definiteness, person, gender,
    PLU ) );
  { degree :: EQUAL },
  Nominal Phrase ( ntype, NOTHING, COMPTVE, EQUAL, definiteness,
    person, gender, number ).
```

The first rule accounts for **Comparison Quantifier MODIFIERs** constituted by **QUANT**, **COMPTVE** or **REL** adverbs with value **+STRESS** for *AVposition*³¹. The second rule accounts for NPs that modify graded AvPs. Affix *CH.comp* is always assigned value **NIL**, due the fact that value **NoEQUAL** for *degree* is not accepted for **Comparison Quantifier MODIFIER**.

The rewrite rule of non-terminal **Factor MODIFIER/2** is the following:

Examples: mil veces *más buena persona*, mil veces *más*, mil veces *más tarde*, mil veces *más interesante*

```
Factor MODIFIER ( NIL, degree ):
  ( { degree :: NIL },
    Determiner ( CARD, no_UNO, PRE, definiteness, FEM, PLU ) );
```

³¹Of course, not any relative adverb, for instance, might appear in this position. However, the restrictions specified in higher levels of the analysis make unnecessary to specify the adequate subtypes here.

```

{ degree :: EQUAL },
Determiner ( COMPTVE, EQUAL, PRE, definiteness, FEM, PLU ) ),
"veces".

```

The three preceding types of modifiers presented in this Section —as well as **Adjective MODIFIERS** when constituted by inherently graded adjectives— can introduce a **Second Term MODIFIER** when they have for *degree* a value in the domain of affix non-terminal *comptve*: **EQUAL** or **NoEQUAL**.

The rewrite rules of non-terminals **Second Term MODIFIER/4,5** are the following:

```

Second Term MODIFIER ( second_term_type, CH_comp, comp, linktype ):
{ second_term_type :: CONSTVE }, { CH_comp :: NIL },
  { comp :: EQUAL }, { linktype :: QUE },
  Clause ( QUE, mood );
{ second_term_type :: COMPTVE },
  Comparison Second Term ( second_term_type, CH_comp, comp,
  linktype ).

```

Second Term MODIFIER might take for affix *CH_comp* values **NIL**, **PARTICULAR+** or **GENERAL+**. These values must be matched by values for affix *CH_comp* supplied by **Adjective**, **Degree**, **Comparison Quantifier** or **Factor MODIFIERS** in higher levels of the analysis. The first alternative in the previous rule accounts for **Second Term MODIFIER** with consecutive meaning, associated with value **EQUAL** for *degree* and constituted by *that*-clauses. According to value **EQUAL** of affix *degree*, the affix *CH_comp* in this alternative is assigned value **NIL** (*tan lento que no llegará nunca*).

The second alternative, on the other hand, accounts for **Second Term MODIFIERS** with comparative meaning, associated with values **EQUAL** or **NoEQUAL** for affix *degree*, and introduced by *como* for value **EQUAL** and by *que* or *de* for value **NoEQUAL**. The value for affix *CH_comp* is **NIL** for value **EQUAL** of affix *degree* (*tan guapa como la que más*). For value **NoEQUAL** of affix *degree*, when **Second Term MODIFIER** is introduced by *que*, affix *CH_comp* takes value **PARTICULAR+** (*más listo que el hambre*). When **Second Term MODIFIER** is introduced by *de*, except for certain AvPs which select value **PARTICULAR+** (*más allá de la muerte*), the value for affix *CH_comp* is always **GENERAL+** (*la más inteligente de todas*). The rewrite rules of non-terminal **Comparison Second Term/4** are the following:

```

Comparison Second Term ( COMPTVE, CH_comp, comp, linktype ):
  Comparison LINKER ( CH_comp, comp, linktype ),
    LINKED Second Term ( linktype );
  { CH_comp :: GENERAL+|PARTICULAR+ }, { comp :: NoEQUAL },
    { linktype :: DE },
    LINKED Second Term ( linktype ).

```

```

Comparison LINKER ( PARTICULAR+, NoEQUAL, QUE ):
  Conjunction ( QUE ).

```

```

Comparison LINKER ( NIL, EQUAL, COMO ):
  Conjunction ( COMO ).

```

```

LINKED Second Term ( linktype ):
  Second Term ( linktype ).

```

```

Second Term ( QUE|COMO ):
  (...).

```

```

Second Term ( DE ):
  (...).

```

We omit the rewrite alternatives of non-terminal **Second Term**. Almost every structure might appear in the right-hand side of the first rule for **Second Term**. Prepositional versions of NPs, although prepositional versions of other linguistic structures are also possible, are the more frequent ones in the right-hand side of the second rule.

```

Second Term MODIFIER ( COMPTVE, GENERAL+, NoEQUAL, person, gender,
  number ):
  Relative Clause ( AGR_MOD, preptype, person, gender, number ).

```

This rule accounts for **Second Term MODIFIERs** constituted by relative clauses. When it is constituted by relative clauses, **Second Term MODIFIER** takes value **GENERAL+** for affix *CH_comp* (*lo más alucinante que he visto*). Although this is not very frequent, relative clauses in this position might also be introduced by prepositional phrases (*la mujer más interesante a la que jamás me atreví a acercarme*).

```

Second Term MODIFIER ( COMPTVE, CH_comp, NoEQUAL, DE, person, gender,
  number ):
  ( { CH_comp :: PARTICULAR+ };
    { CH_comp :: GENERAL+ },
    ( { gender :: MASC|FEM }, { number :: PLU };
      { gender :: NEUT } ) ),

```

Prepositional Nominalization (DE, RELCLAUSE, definiteness,
person, gender, number).

This rule accounts for **Second Term MODIFIERs** constituted by prepositional relative clauses introduced by preposition *de*. Depending on the context, **Second Term MODIFIER** will have either value **GENERAL+** or value **PARTICULAR+** for affix *CH.comp*. Within **Nominal Phrase**, **Pronominal Phrase**, **Pronoun Phrase** and **Adverb Phrase** non-terminals, affix *CH.comp* takes value **PARTICULAR+** (*más historias de las que puedo contar, más de lo que se puede desear*). Within **Noun Phrase** and **Adjective Phrase** non-terminals, affix *CH.comp* might take both **GENERAL+** and **PARTICULAR+** values: it takes value **GENERAL+** when it agrees for *gender* with the nucleus—for **MASC** or **FEM** values for *gender*, value for *number* will always be **PLU** (*el mejor músico de los que hay aquí, la mejor de las que conozco*). It might take value **PARTICULAR+** when it is associated with value **NEUT** of affix *gender* (*mejor músico de lo que parece a primera vista, mejor de lo que parece*).

3.2.3.4.2.3 Noun MODIFIER The rewrite rule of non-terminal **Noun MODIFIER/0** is the following:

Noun MODIFIER:
Noun (ntype, gender, number).

3.2.3.5 Nominalization

Two different non-terminals of the grammar are associated with structures of nominalizations. **Nominalization/5** accounts for nominalizations of adjective and prepositional phrases, while **Nominalization/6** accounts for nominalizations of relative clauses.

3.2.3.5.1 Nominalization/5 The rewrite rule of non-terminal **Nominalization/5** is the following:

Examples: *la más bonita, la de la esquina*

Nominalization (MDcategory, definiteness, person, gender, number):
DETERMINER (ART, EL, position, definiteness, gender, number),
Nominalization NOMINAL I (MDcategory, person, gender, number).

Affix *MDcategory* in this rule contains information about the structural category nominalized, either AjPs or prepositional phrases.

3.2.3.5.2 Nominalization NOMINAL I/4

```
Nominalization NOMINAL I ( MDcategory, person, gender, number ):
  HIDDEN FINAL Nominalized Phrase ( premodifier, postmodifier,
    MDcategory, person, gender, number );
  Nominalized Phrase ( premodifier, postmodifier, MDcategory,
    person, gender, number ).
```

As the adjective or prepositional phrases that constitute nominalizations might be modified, the AGFL General Frame for Addition of Modifiers described in Section 3.2.3.2.2 applies also in the context of the rewrite rules of non-terminal **Nominalization**. The first alternative in the previous rule, in fact, accounts for nominalizations of unmodified adjective or prepositional phrases, while the second accounts for nominalization of the modified ones.

3.2.3.5.3 HIDDEN FINAL Nominalized Phrase/6

Examples: Lo de que siempre tenga hambre *resulta un poco fastidioso*, Lo de salir siempre en grupo *es síntoma de inseguridad*

```
HIDDEN FINAL Nominalized Phrase ( NOTHING, NOTHING, MDcategory,
  THIRD, gender, number ):
  { MDcategory :: PRPPHRASE }, { gender :: NEUT },
  ( Prepositional Clause ( preposition, QUE, non_imperative );
    Prepositional Clause ( preposition, SUBORDINATED, INFINITIVE,
      person ) ).
```

Examples: *la verde, el repetido, el del sombrero, el de casa*

```
HIDDEN FINAL Nominalized Phrase ( NOTHING, NOTHING, MDcategory,
  person, gender, number ):
  ( { MDcategory :: AJPHRASE },
    HEAD Adjective Phrase ( premodifier, QUAL,
      NIL|GENERAL|GENERAL+|PARTICULAR+, degree, gender,
      number );
    { MDcategory :: AJPHRASE },
    HEAD Participle Clause ( premodifier,
      NIL|GENERAL|GENERAL+|PARTICULAR+, degree, gender,
      number );
    { MDcategory :: PRPPHRASE },
    Prepositional Nominal Phrase ( preposition, ntype,
      premodifier, dttype, dtsubtype, definiteness, person1,
```

```

        gender1, number1 );
{ MDcategory :: PRPPHRASE },
  Prepositional Noun Phrase ( preposition, ntype, premodifier,
    postmodifier, CH_comp, degree, definiteness, person1,
    gender1, number1 );
(... ) ,
  Noun Phrase person ( person, number ).

```

AjPs might be nominalized, as well as prepositional versions of almost every structural category. Nominalizations of adjective and prepositional phrases are assigned values for affix *person* the same as **Noun Phrase** does, that is, by means of predicate **Noun Phrase person**, described in Section 3.2.3.2.4.2, epigraph **No MODIFIER** (*Vamos a ganar los mejores*). Nominalizations of prepositional clauses, on the contrary, are always assigned values **THIRD** for affix *person* and **NEUT** for affix *gender* (*Lo de ir a verla no me motiva en absoluto*).

All *CH_comp* values except **PARTICULAR** are allowed for **Adjective Phrase**, which means that, if the nominalized **Adjective Phrase** includes a **Second Term MODIFIER**, it will determine the value of *CH_comp* within the nominalization (*el más listo de todos*). Otherwise, affix *CH_comp* will be assigned either value **NIL**, if the value for affix *degree* is different from **NoEQUAL** (*los buenos*), or value **GENERAL**, if the value for affix *degree* is **NoEQUAL** (*los más buenos*)³².

3.2.3.5.4 Addition of Modifiers within Nominalization/5

3.2.3.5.4.1 HIDDEN RECURSIVE Nominalized Phrase/6

```

HIDDEN RECURSIVE Nominalized Phrase ( premodifier, postmodifier,
  MDcategory, person, gender, number ):
  Nominalized Phrase ( premodifier, postmodifier, MDcategory,
    person, gender, number );
HIDDEN FINAL Nominalized Phrase ( NOTHING, NOTHING, MDcategory,
  person, gender, number ).

```

3.2.3.5.4.2 Nominalized Phrase/6

Example: *los rojos que resulten más llamativos*

³²Obviously, it is a question for second level of formalization which *degree* and *CH_comp* values are allowed within nominalizations of AjPs. The values in the domain of affix *degree+*, apart from value **NoEQUAL**, appear to be very rare.

Nominalized Phrase (NOTHING, RELCLAUSE, MDcategory, person, gender, number):
 Nominalized Phrase NUCLEUS (NOTHING, RELCLAUSE, MDcategory, person, gender, number),
 (Relative Clause MODIFIER (AGR_MOD, person, gender, number);
 Relative Clause MODIFIER (no_AdvP)).

According to this rule, nominalizations of adjective and prepositional phrases might only include one **Relative Clause MODIFIER**. This, however, is a decision that concerns second level of formalization and that, in consequence, can be revised without difficulty.

3.2.3.5.4.3 Nominalized Phrase NUCLEUS/6

Nominalized Phrase NUCLEUS (NOTHING, RELCLAUSE, MDcategory, person, gender, number):
 HIDDEN RECURSIVE Nominalized Phrase (NOTHING, NOTHING, MDcategory, person, gender, number).

3.2.3.5.5 Addition of Determiners within Nominalization/5 On the other hand, as **Nominalizations**, like PrPs, might be added determiners—*over-determiners*—, the AGFL General Frame for Addition of Determiners, described in Section 3.2.3.3.2, applies also in the context of **Nominalizations**.

3.2.3.5.5.1 Nominalization I/5

Nominalization I (MDcategory, definiteness, person, gender, number):
 DETERMINER (ART, EL, position, definiteness, gender, number),
 Nominalization NOMINAL I (MDcategory, person, gender, number).

3.2.3.5.5.2 Nominalization/5, 1st level of recursion

Example³³: *todo lo más caro*

Nominalization (MDcategory, definiteness, person, gender, number):
 DETERMINER (GLOBAL, NIL, position, definiteness, gender, number),
 Nominalization NOMINAL II (MDcategory, definiteness, person, gender, number).

³³Only **GLOBAL** determiners may be added to **Nominalization**.

3.2.3.5.5.3 Nominalization NOMINAL II/5

```
Nominalization NOMINAL II ( MDcategory, definiteness, person, gender,
number ):
  Nominalization I ( MDcategory, definiteness, person, gender,
number ).
```

3.2.3.5.6 Nominalization/6 The rewrite rules of non-terminal **Nominalization/6** are the following:

Examples: Lo que he visto siempre *seguiré viéndolo hasta el final*, A los que se lo hayan pedido *deberían acudir*

```
Nominalization ( MDcategory, NIL, definiteness, person, gender,
number ):
  DETERMINER ( ART, EL, position, definiteness, gender, number ),
  Nominalization NOMINAL I ( MDcategory, NIL, person, gender,
number ).
```

```
Nominalization ( MDcategory, preposition, definiteness, person,
gender, number ):
{ definiteness :: DEFINITE },
  Nominalization NOMINAL I ( MDcategory, preposition, person,
gender, number ).
```

Affix *preptype* of non-terminal **Nominalization/6** stores the value of the preposition that introduces the prepositional phrase that in turn introduces the nominalized relative clause.

When the relative phrase that introduces the relative clause is not a prepositional phrase, we consider that the relative clause starts with pronoun *que*, and that the whole relative clause is after combined with determiner *el* in order to form the nominalization. That is, in our example, the relative clause *que he visto siempre* is combined with determiner *lo* within the nominalization *lo que he visto siempre*. The first rule above accounts for this type of nominalizations of relative clauses. Affix *preptype* takes here value **NIL**.

On the contrary, when the phrase that introduces the relative clause is a prepositional phrase, the relative clause starts with a prepositional PrP with determiner *el* and pronoun *que*. This relative clause is nominalized as such. That is, in our example, relative clause *a los que se lo hayan pedido* is ready to be nominalized. The second rule above accounts for this type of nominalizations of relative clauses.

Affix non-terminal *preposition* in the position of affix *preptype* ensures that relative clauses not introduced by prepositional phrases will never be nominalized as such.

3.2.3.5.7 Nominalization NOMINAL I/5

```
Nominalization NOMINAL I ( MDcategory, preptype, person, gender,
    number ):
  HIDDEN FINAL Nominalized Phrase ( premodifier, postmodifier,
    MDcategory, preptype, person, gender, number );
  Nominalized Phrase ( premodifier, postmodifier, MDcategory,
    preptype, person, gender, number ).
```

Nominalizations of relative clauses might also contain a relative clause modifier, as well as be added the determiner *todo*. These possibilities are formalized the same as described for adjective and prepositional phrase nominalizations, so we will omit here the corresponding rules.

3.2.3.5.8 HIDDEN FINAL Nominalized Phrase/6

```
HIDDEN FINAL Nominalized Phrase ( NOTHING, NOTHING, MDcategory,
    preptype, person, gender, number ):
  { MDcategory :: RELCLAUSE },
  Relative Clause ( NOMINAL, preptype, person, gender, number ).
```

The relative clauses that can be nominalized are assigned value **NOMINAL** for affix *RCfunction* (see also Sections 3.2.3.4.1.3 and 3.2.3.6.5).

```
Relative CONSTITUENT ( RCfunction, preptype, person, gender,
    number ):
  HEAD Pronoun Phrase ( prmd_general, postmodifier, REL, QUE,
    definiteness, person, gender, number ),
  { RCfunction :: NOMINAL }, { preptype :: NIL };
  Prepositional Pronominal Phrase ( preptype, prmd_general,
    GLOBAL|ART, NIL|EL, REL, QUE, definiteness, person, gender,
    number ),
  { RCfunction :: NOMINAL }.
```

The relative phrases associated with value **NOMINAL** for affix *RCfunction* of **Relative CONSTITUENT** are *que* or *al que*, *a la que*, etc., and the like.

3.2.3.6 Other Constituents of the Clause. The Clause

In this Section an outline is given of the formalization of clauses that function as constituents of main clauses. It is obvious that, in order to account for these

constituents, we have to include a discussion of the formalization of clause itself. Nevertheless, this issue exceeds the argument of this exposition, and will be developed only to the extent that the clause is the context in which all described structures for NPs, PrPs, AvPs and AjPs should be integrated.

3.2.3.6.1 That-clauses The rewrite rules that account for *that*-clauses are the following:

Examples: *Aseguraba que les contaría el plan a todos, El que les haya contado el plan a todos demuestra que no se puede confiar en él*

```
Clause ( QUE, non_imperative ):
  Conjunctive LINKER ( QUE ),
  LINKED Clause ( SUBORDINATED, non_imperative ).
```

```
Clause ( EL_QUE, non_imperative ):
  DETERMINER ( ART, NIL, PRE, DEFINITE, MASC, SING ),
  Conjunctive LINKER ( QUE ),
  LINKED Clause ( SUBORDINATED, non_imperative ).
```

```
Conjunctive LINKER ( cjtype ):
  Conjunction ( cjtype ).
```

```
LINKED Clause ( clause_type, mood ):
  Clause ( clause_type, mood ).
```

That-clauses are constituted by conjunction *que* followed by subordinated clauses, which are accounted for by the rewrite rules of non-terminal **Clause/2**. Non-terminal **Clause/2** is rewritten into the syntactic functions **SUBJECT**, **DIRECT OBJECT**, **INDIRECT OBJECT**, **PREPOSITIONAL COMPLEMENT**, **PREDICATIVE COMPLEMENT**, **AGENT**, **CIRCUMSTANCE** and **PREDICATE**. In principle, as many alternatives as possible dispositions of combinations of one or more syntactic functions for all possible combinations of *verb scheme* and *voice* within the **PREDICATE** function should constitute the rewrite rule of **Clause**. To illustrate the integration of clause and phrase modules within the full sentence grammar, we include here one example alternative of this rule:

```
Clause ( MAIN|SUBORDINATED, mood ):
  (...);
```

```

SUBJECT ( const_category1, person, gender, number ),
  ( PREDICATE ( SDOIO, ACT|MIDD, person, number, non_imperative,
    DAT, person_one, gender_one, number_one );
    PREDICATE ( SDOIO, ACT|MIDD, person, number,
    non_imperative ) ),
  DIRECT OBJECT ( const_category2, person1, gender1, number1,
    person ),
  INDIRECT OBJECT ( const_category3, person_one, gender_one,
    number_one, person );
(...).

```

As can be observed in the rule above, main and subordinated clauses share the same internal structure, being instead the hierarchical position of the clause within its linguistic context which determines whether it is a main or a subordinated clause. For this reason precisely, values **MAIN** and **SUBORDINATED** for affix *clause_type* cooccur in the example rule: depending on the non-terminal that calls upon this rule in the analysis, the affix in question will take either the **MAIN** or the **SUBORDINATED** value. If non-terminal **Clause** is called upon within the rewrite rule of non-terminal **UTTERANCE**, affix *clause_type* will take the value **MAIN**; if, on the other hand, it is called upon within the rewrite rule of non-terminal **LINKED Clause**, the affix *clause_type* will take the value **SUBORDINATED**.

From what we have stated in the previous paragraph, one more important inference is straightforward and must be pointed out: this is the fact that **SUBORDINATED** clauses can be dominated by phrase or clause nodes, whereas **MAIN** clauses, on the contrary, cannot. This ultimately means that **MAIN** clauses are the top non-terminals of the clause module of the grammar and that, with their rewrite rule, both phrase and clause modules, as well as their integration, can be considered complete.

With respect, on the other hand, to syntactic functions, the rewrite rules of non-terminal **PREDICATE** were introduced in Chapter 2, whilst the rewrite rules of the remaining syntactic functions call upon the non-terminals described in previous Sections of this Chapter, which, as these are rewritten into NPs, PrPs, AvPs and AjPs, means that the rewrite rules of syntactic functions act indeed as a bridge between the clause and the phrase modules of the grammar. The rewrite rule of the syntactic function **SUBJECT**, for instance, is the following:

```

SUBJECT ( const_category, person, gender, number ):

```

```

HEAD Noun Phrase ( ntype, prmd_general, postmodifier,
NIL|PARTICULAR|PARTICULAR+, degree, definiteness, person,
gender, number );
Nominal Phrase ( ntype, prmd_general, dtgeneral, dtsubtype,
definiteness, person, gender, number );
Nominalization ( MDcategory, definiteness, person, gender,
number );
Nominalization ( MDcategory, NIL, definiteness, person, gender,
number );
HEAD Pronoun Phrase ( prmd_general, postmodifier, prgeneral,
prsubtype, definiteness, person, gender, number );
HEAD Pronoun Phrase ( prmd_general, postmodifier, prgeneral,
prsubtype, definiteness, person, gender, number, NOM );
Pronominal Phrase ( prmd_general, dtgeneral, dtsubtype, prgeneral,
prsubtype, definiteness, person, gender, number );
Pronominal Phrase ( prmd_general, dtgeneral, dtsubtype, prgeneral,
prsubtype, definiteness, person, gender, number, NOM );
Relative Clause ( S, NIL, person, gender, number );
{ person :: THIRD }, { gender :: NEUT }, { number :: SING },
( Clause ( QUE|EL_QUE, mood );
Clause ( SUBORDINATED|EL_SUBORDINATED, INFINITIVE,
MVperson ) ).

```

What we want to point out here —it will not be the case for other types of clauses (see below)—, is the fact that all the phrases identified as a possible **SUBJECT** by this rule, if they include one of the affixes *premodifier*, *dttype* or *prtype*, these always take values within the domains of, respectively, *prmd_general*, *dtgeneral* and *prgeneral*. The same holds for the rewrite rules of all the other syntactic functions enumerated above, and this means that interrogative, exclamative or relative phrases are excluded from the syntactic functions that play a role within **MAIN** or **SUBORDINATED Clauses**.

Yet another important focus of attention in this rule is constituted by the fact that affix *CH_comp* of the non-terminals whose value for *CH_comp* might remain open longer during the analysis, **Noun Phrase** and **Adjective Phrase**, takes a value here (see Sections 3.2.3.2.4.2 and 3.2.3.2.7.2, epigraph **Degree MODIFIER**): this is because, if a syntactic function rewrites as one of the non-terminals **Noun Phrase** or **Adjective Phrase**, we know that no determiner *el* can be added to further expand the NP or to form a nominalization on the basis of the AjP, so we can state that values in the domain of *general* cannot be selected for affix *CH_comp*.

3.2.3.6.2 Infinitive Clauses

Examples: *No me gustaría* tener que darle la enhorabuena a ella, El no haber podido darle el puesto a su candidato *ha acabado por exasperarlo*

```
Clause ( SUBORDINATED, mood, person ) :
  (...);
  ( NON FINITE PREDICATE ( SDOIO, ACT, mood, DAT, person_one,
    gender_one, number_one );
    NON FINITE PREDICATE ( SDOIO, ACT, mood, ACC, person_one1,
    gender_one1, number_one1 );
    NON FINITE PREDICATE ( SDOIO, ACT, mood ) ),
    DIRECT OBJECT ( const_category, person_one1, gender_one1,
    number_one1, person ),
    INDIRECT OBJECT ( const_category1, person_one, gender_one,
    number_one, person );
  (...).
```

```
Clause ( EL_SUBORDINATED, INFINITIVE, MVperson ) :
  DETERMINER ( ART, NIL, PRE, DEFINITE, MASC, SING ),
  LINKED Clause ( SUBORDINATED, INFINITIVE, MVperson ).
```

```
LINKED Clause ( clause_type, mood, MVperson ) :
  Clause ( clause_type, mood, MVperson ).
```

These rules are self-explanatory. **Clause/3**, infinitive clause, is rewritten into the syntactic functions **SUBJECT**, **DIRECT OBJECT**, **INDIRECT OBJECT**, **PREPOSITIONAL COMPLEMENT**, **PREDICATIVE COMPLEMENT**, **AGENT**, **CIRCUMSTANCE** and **NON PERSONAL PREDICATE**. The rewrite rules of **NON FINITE PREDICATE** were introduced in Chapter 2, the remaining syntactic functions are the same found in **Clause/2**.

Infinitive clauses are assigned a unique value **SUBORDINATED** for affix *clause_type*. This is no doubt their main function, but it is a question that concerns second level of formalization whether we assign also value **MAIN** for affix *clause_type* within **Clause/3**, thus including this non-terminal as another possible rewrite alternative of non-terminal **UTTERANCE**.

3.2.3.6.3 Participle Clauses

Example: *cifrada por el Informe de la Cámara en varios millones de pesetas*

```
Participle Clause ( SUBORDINATED, gender, number ) :
  (...);
  PARTICIPLE PREDICATE ( SDOPC, PASS, preptype, PARTICIPLE,
```

```

    gender, number ),
AGENT ( const_category1 ),
PREPOSITIONAL COMPLEMENT ( const_category2, preptype, person );
(...).

```

These rules are self-explanatory. **Participle Clause** is rewritten into syntactic functions **SUBJECT**, **DIRECT OBJECT**, **INDIRECT OBJECT**, **PREPOSITIONAL COMPLEMENT**, **PREDICATIVE COMPLEMENT**, **AGENT**, **CIRCUMSTANCE** and **PARTICIPLE PREDICATE**. The rewrite rules of **PARTICIPLE PREDICATE** were introduced in Chapter 2, the remaining syntactic functions are the same found in **Clause/2,3**.

Participle clauses may be modified (*muy apreciada por los expertos*). We account for this possibility, introducing a **HEAD Participle Clause/5** non-terminal, as we account for the addition of modifiers to AjPs, so we will not display here the corresponding rules—which can be easily deduced from Section 3.2.3.2.7.

3.2.3.6.4 Interrogative Clauses

3.2.3.6.4.1 Yes/No Questions

Examples: *No sé si le dijo las mismas barbaridades al otro candidato*, *No sé si concederle el beneficio de la duda a una persona que lo merece tan poco*

```

Interrogative Clause ( SI, Yes_No, non_imperative ):
  Conjunctive LINKER ( SI ),
  LINKED Clause ( SUBORDINATED, non_imperative ).

```

```

Interrogative Clause ( SI, Yes_No, INFINITIVE, MVperson ):
  Conjunctive LINKER ( SI ),
  LINKED Clause ( SUBORDINATED, INFINITIVE, MVperson ).

```

These rules are self-explanatory. *Yes/no*-questions are constituted by conjunction *si* followed by subordinated clauses.

3.2.3.6.4.2 Wh-Questions

Examples: *¿Se puede saber de qué me acusa usted?*, *Nunca sé dónde ponerme el distintivo*

```

Interrogative Clause ( MAIN|SUBORDINATED, Wh, non_imperative ):
  (...);

```

```

Interrogative PREPOSITIONAL COMPLEMENT ( const_category2,
  preptype ),
( PREDICATE ( SDOPC, ACT|MIDD, preptype, person, number,
  INDICATIVE|SUBJUNCTIVE, ACC, person_one, gender_one,
  number_one );
  PREDICATE ( SIOPC, ACT|MIDD, preptype, person, number,
  INDICATIVE|SUBJUNCTIVE, DAT, person_one, gender_one,
  number_one );
  PREDICATE ( SPC, ACT|MIDD, preptype, person, number,
  INDICATIVE|SUBJUNCTIVE ) ),
SUBJECT ( const_category1, person, gender, number );
(...).

```

```

Interrogative Clause ( SUBORDINATED, Wh, INFINITIVE, person ):
(...);
Interrogative PREPOSITIONAL COMPLEMENT ( const_category2,
  preptype ),
( NON FINITE PREDICATE ( SDOPC, MIDD, preptype, INFINITIVE,
  VERB, person, gender_one, number_one );
  NON FINITE PREDICATE ( SDOIOPC, ACT, preptype, INFINITIVE,
  DAT, person_one, gender_one, number_one );
  NON FINITE PREDICATE ( SDOPC, ACT, preptype, INFINITIVE ) ),
DIRECT OBJECT ( const_category1, person1, gender, number,
  person );
(...).

```

Interrogative Clause/3,4 are introduced by an interrogative constituent, which is in turn followed by one or more syntactic functions as those found in **Clause/2,3**. All syntactic functions have interrogative versions. The rewrite rule of non-terminal **Interrogative PREPOSITIONAL COMPLEMENT**, for instance, is the following:

```

Interrogative PREPOSITIONAL COMPLEMENT ( const_category, preptype ):
  HEAD Prepositional Adverb Phrase ( prmd_general, preptype,
  prmd_general1, postmodifier, INTG, DONDE, CH_comp, degree );
  HEAD Adverb Phrase ( prmd_general, postmodifier, INTG, DONDE,
  CH_comp, degree );
  Prepositional Nominal Phrase ( preptype, ntype, prmd_general,
  INTG, dtno_declarative, definiteness, person, gender,
  number );
  Prepositional Pronoun Phrase ( preptype, prmd_general,
  postmodifier, INTG, prno_declarative, definiteness, person,
  gender, number );
  Prepositional Pronominal Phrase ( preptype, prmd_general,
  dtgeneral, dtsubtype, INTG, prno_declarative, definiteness,
  person, gender, number );
  Prepositional Pronominal Phrase ( preptype, prmd_general, INTG,
  dtno_declarative, prgeneral, prsubtype, definiteness, person,

```

gender, number).

What we want to point out here is the fact that all phrases in this rule are assigned for affixes *avtype*, *dttype* or *prtype* value **INTG**, which means that interrogative clauses are identified by means of interrogative phrases. If we consider that also relative and exclamative clauses are similarly identified by means of relative and exclamative phrases, this explains why, together with subordination in the context of the addition of modifiers and determiners, discrimination of interrogative, relative and exclamative phrases is the next working-principle which modulates the structure of the formalization of phrase-level analysis as described here.

As can be appreciated in the above rules, the behaviour of **Interrogative Clause** with respect to affix *clause_type* is the same as that of **Clause**.

3.2.3.6.5 Relative Clauses

Examples: *El rey Alfonso hizo entonces pregonar que quien diera posada al Cid o a su gente sufriría su misma suerte, Hoy no tengo nada que decir a mis feligreses*

```
Relative Clause ( RCfunction, preptype, person, gender, number ):
(...);
Relative SUBJECT ( rel_const_category, RCfunction, person, gender,
number ),
( PREDICATE ( SDOIO, ACT|MIDD, person, number,
non_imperative );
PREDICATE ( SDOIO, ACT|MIDD, person, number, non_imperative,
DAT, person_one, gender_one, number_one ) ),
DIRECT OBJECT ( const_category1, person1, gender1, number1,
person ),
INDIRECT OBJECT ( const_category2, person_one, gender_one,
number_one, person ),
{ preptype :: NIL };
(...).
```

```
Relative Clause ( RCfunction, preptype, person, gender, number,
MVperson ):
(...);
Relative DIRECT OBJECT ( rel_const_category, RCfunction, preptype,
person, gender, number ),
( NON FINITE PREDICATE ( SDOIO, ACT, INFINITIVE );
NON FINITE PREDICATE ( SDOIO, ACT, INFINITIVE, DAT,
person_one, gender_one, number_one ) ),
INDIRECT OBJECT ( const_category, person_one, gender_one,
number_one, MVperson );
```

(...).

Relative Clause/5,6 is introduced by a relative constituent, which is in turn followed by other syntactic functions as those found in **Clause/2,3** and **Interrogative Clause/3,4**. All syntactic functions have also relative versions. The rewrite rule of non-terminal **Relative SUBJECT**, for instance, is the following—obviously, all phrases in this rule are assigned for *dttype* or *prtype* value **REL**:

```
Relative SUBJECT ( rel_const_category, RCfunction, person, gender,
number ):
  Nominal Phrase ( ntype, prmd_general, REL, CUYO, definiteness,
person, gender, number ),
  { RCfunction :: no_AdvP };
  Nominal Phrase ( ntype, prmd_general, REL, CUANTO, definiteness,
person, gender, number ),
  { RCfunction :: S|DO|IO|PC|PR|AG|C|LINKED|CUANTO_MAS };
  Pronominal Phrase ( prmd_general, REL, CUANTO, COMPTVE, NoEQUAL,
definiteness, person, gender, number ),
  { RCfunction :: CUANTO_MAS };
  Pronominal Phrase ( prmd_general, TOT, NIL, REL, CUANTO,
definiteness, person, gender, number ),
  { RCfunction :: S|DO|IO|PC|PR|AG|C|LINKED };
  HEAD Pronoun Phrase ( prmd_general, postmodifier, REL, QUE,
definiteness, person, gender, number ),
  { RCfunction :: AGR_MOD };
  HEAD Pronoun Phrase ( prmd_general, postmodifier, REL, QUE,
definiteness, person, gender, number ),
  { RCfunction :: NOMINAL };
  HEAD Pronoun Phrase ( prmd_general, postmodifier, REL, CUANTO,
definiteness, person, gender, number ),
  { RCfunction :: S|DO|IO|PC|PR|AG|C|LINKED };
  HEAD Pronoun Phrase ( prmd_general, postmodifier, REL, QUIEN,
definiteness, person, gender, number ),
  { RCfunction :: S|DO|IO|PC|PR|AG|LINKED }.
```

At this point of this Section, as we have already accounted for the various structures of phrases and clauses in Spanish, it appears worthwhile to summarize what we have been trying to illustrate about the integration of clause and phrase modules within the grammar: clauses are rewritten into syntactic functions, syntactic functions are rewritten into phrases, and both phrases and clauses are classified according to the same categories relevant for the analysis of utterances: they might have main—general, when talking about phrases—, interrogative, relative or exclamative (not showed here) character.

As regards their integration within the whole grammar, **Relative Clauses**, however, present certain particularities that it might be of benefit to emphasize here:

- They are always subordinated clauses, so affix *clause_type* might be omitted in this case.
- They have more functional capabilities than **Clause**, **Participle Clause** or **Interrogative Clause**. These functional capabilities are closely related to the relative phrase that introduces the relative clause (see Section 3.2.3.4.1.3). This relation is formalized by means of affix *RCfunction*, which is assigned value according to the concrete relative phrase into which in each case the **Relative SUBJECT**, **Relative DIRECT OBJECT**, etc. in question is rewritten. Conversely, the rewrite alternatives of syntactic functions (**SUBJECT**, **DIRECT OBJECT**, etc.) constituted by **Relative Clauses** specify in their turn the value of affix *RCfunction* convenient in each case. For instance, the alternative that rewrites **SUBJECT** —see Section 3.2.3.6.1— as **Relative Clause** specifies for affix *RCfunction* value **S**, while the alternatives that rewrite **Relative Clause MODIFIER** in Section 3.2.3.4.1.3 specify for affix *RCfunction* values in the domain of *modifier*.

In Sections 3.2.3.4.1.3 and 3.2.3.5, we included as rewrite alternatives of the generic term **Relative CONSTITUENT**, the collection of phrases that can introduce relative clauses which must after either function as **Relative Clause MODIFIERS** or be nominalized. We include here the rewrite alternatives of phrases that enable relative clauses to constitute **SUBJECT**, **DIRECT OBJECT**, etc. syntactic functions, or **Prepositional Relative Clauses** (see Section 3.2.3.7)³⁴. This is a very long rule, but it is very illustrative both with respect to phrase-clause modules integration and to the highly detailed classification of pronouns, determiners and adverbs proposed in this work:

Relative CONSTITUENT (*RCfunction*, *preptype*, *person*, *gender*,
number):

³⁴The alternatives that enable relative clauses to have a role beyond the level of the clause, associated with value **CUANTO_MAS** for affix *RCfunction*, see Section 3.1.6, are included too.

```

Prepositional Nominal Phrase ( preptype, ntype, prmd_general, REL,
    CUANTO, definiteness, person, gender, number ),
    { RCfunction :: DO|IO|PC|PR|AG|C|LINKED|CUANTO_MAS };
Nominal Phrase ( ntype, prmd_general, REL, CUANTO, definiteness,
    person, gender, number ),
    { RCfunction :: S|DO|IO|PC|PR|AG|C|LINKED|CUANTO_MAS },
    { preptype :: NIL };
HEAD Noun Phrase ( ntype, REL, postmodifier,
    NIL|PARTICULAR|PARTICULAR+, degree, definiteness, person,
    gender, number ),
    { RCfunction :: CUANTO_MAS }, { preptype :: NIL };
Pronominal Phrase ( prmd_general, REL, CUANTO, COMPTVE, NoEQUAL,
    definiteness, person, gender, number ),
    { RCfunction :: CUANTO_MAS }, { preptype :: NIL };
Prepositional Pronominal Phrase ( preptype, prmd_general, REL,
    CUANTO, COMPTVE, NoEQUAL, definiteness, person, gender,
    number ),
    { RCfunction :: CUANTO_MAS };
Prepositional Pronominal Phrase ( preptype, prmd_general, GLOBAL,
    NIL, REL, CUANTO, definiteness, person, gender, number ),
    { RCfunction :: DO|IO|PC|PR|AG|C|LINKED };
Pronominal Phrase ( prmd_general, GLOBAL, NIL, REL, CUANTO,
    definiteness, person, gender, number ),
    { RCfunction :: S|DO|IO|PC|PR|AG|C|LINKED },
    { preptype :: NIL };
Prepositional Pronoun Phrase ( preptype, prmd_general,
    postmodifier, REL, CUANTO, definiteness, person, gender,
    number ),
    { RCfunction :: DO|IO|PC|PR|AG|C|LINKED };
HEAD Pronoun Phrase ( prmd_general, postmodifier, REL, CUANTO,
    definiteness, person, gender, number ),
    { RCfunction :: S|DO|IO|PC|PR|AG|C|LINKED },
    { preptype :: NIL };
Prepositional Pronoun Phrase ( preptype, prmd_general,
    postmodifier, REL, QUIEN, definiteness, person, gender,
    number ),
    { RCfunction :: DO|IO|PC|PR|AG|LINKED };
HEAD Pronoun Phrase ( NOTHING, postmodifier, REL, QUIEN,
    definiteness, person, gender, number ),
    { RCfunction :: S|DO|IO|PC|PR|AG|LINKED },
    { preptype :: NIL };
HEAD Adjective Phrase ( REL, ajtype, NIL|PARTICULAR|PARTICULAR+,
    degree, gender, number ),
    { RCfunction :: CUANTO_MAS }, { preptype :: NIL };
HEAD Participle Clause ( REL, NIL|PARTICULAR|PARTICULAR+, degree,
    gender, number ),
    { RCfunction :: CUANTO_MAS }, { preptype :: NIL };
HEAD Prepositional Adverb Phrase ( REL, preptype, prmd_general1,
    postmodifier, avgeneral, avsubtype, CH_comp, degree ),
    { RCfunction :: CUANTO_MAS },

```

```

    { person :: NIL }, { gender :: NIL }, { number :: NIL };
HEAD Prepositional Adverb Phrase ( prmd_general, preptype,
    prmd_general1, postmodifier, REL, DONDE, NIL, NIL ),
    { RCfunction :: PC|C },
    { person :: NIL }, { gender :: NIL }, { number :: NIL };
HEAD Adverb Phrase ( REL, postmodifier, avgeneral, avsubtype,
    CH_comp, degree ),
    { RCfunction :: CUANTO_MAS }, { preptype :: NIL },
    { person :: NIL }, { gender :: NIL }, { number :: NIL };
HEAD Adverb Phrase ( prmd_general, postmodifier, REL, avrel, NIL,
    NIL ),
    ( { RCfunction :: PC|C|LINKED }, { avrel :: DONDE };
      { RCfunction :: C|LINKED }, { avrel :: CUANDO };
      { RCfunction :: C }, { avrel :: MIENTRAS };
      { RCfunction :: PC|C|PR|LINKED }, { avrel :: COMO } ),
    { preptype :: NIL },
    { person :: NIL }, { gender :: NIL }, { number :: NIL }.

```

The concrete relationships between relative phrases and syntactic functions of the whole relative clauses might of, course, yet be refined, but these are, again, details that concern second level of formalization.

3.2.3.7 Prepositional Constituents

Phrases and subordinated clauses can be contextualized by means of prepositions too. The prepositional versions of phrases and subordinated clauses are all formalized in the same way.

3.2.3.7.1 Prepositional Noun Phrase/10

Example: *de madera*

```

Prepositional Noun Phrase ( preptype, ntype, premodifier,
    postmodifier, CH_comp, degree, definiteness, person, gender,
    number ):
    Prepositional LINKER ( preptype ),
        LINKED Noun Phrase ( ntype, premodifier, postmodifier, CH_comp,
            degree, definiteness, person, gender, number ).

Prepositional LINKER ( preptype ):
    Preposition ( preptype ).

LINKED Noun Phrase ( ntype, premodifier, postmodifier, CH_comp,
    degree, definiteness, person, gender, number ):
    { CH_comp :: NIL|PARTICULAR|PARTICULAR+ },
        HEAD Noun Phrase ( ntype, premodifier, postmodifier, CH_comp,
            degree, definiteness, person, gender, number ).

```

3.2.3.7.2 Prepositional Pronoun Phrase/9,10

Examples: *de alguien, de ellos, conmigo*

```
Prepositional Pronoun Phrase ( preptype, premodifier, postmodifier,
    prtype, prsubtype, definiteness, person, gender, number ):
  Prepositional LINKER ( preptype ),
    LINKED Pronoun Phrase ( premodifier, postmodifier, prtype,
        prsubtype, definiteness, person, gender, number ).

HEAD Prepositional Pronoun Phrase ( preptype, premodifier,
    postmodifier, prtype, prsubtype, definiteness, person, gender,
    number, case ):
  Prepositional Pronoun Phrase ( preptype, premodifier,
    postmodifier, prtype, prsubtype, definiteness, person,
    gender, number, case );
  { premodifier :: NOTHING }, { postmodifier :: NOTHING },
    Prepositional Pronoun ( preptype, prtype, prsubtype,
    definiteness, person, gender, number, case ).
```

In this rule, the first alternative accounts for personal pronouns preceded by prepositions, and the second for personal pronouns merged with prepositions. The invisible non-terminal **HEAD Prepositional Pronoun Phrase** is used to combine both alternatives for further integration within the grammar.

```
Prepositional Pronoun Phrase ( preptype, premodifier, postmodifier,
    prtype, prsubtype, definiteness, person, gender, number,
    case ):
  Prepositional LINKER ( preptype ),
    LINKED Pronoun Phrase ( premodifier, postmodifier, prtype,
        prsubtype, definiteness, person, gender, number, case ).

LINKED Pronoun Phrase ( premodifier, postmodifier, prtype,
    prsubtype, definiteness, person, gender, number ):
  HEAD Pronoun Phrase ( premodifier, postmodifier, prtype,
    prsubtype, definiteness, person, gender, number ).

LINKED Pronoun Phrase ( premodifier, postmodifier, prtype,
    prsubtype, definiteness, person, gender, number, PREP ):
  HEAD Pronoun Phrase ( premodifier, postmodifier, prtype,
    prsubtype, definiteness, person, gender, number, PREP ).
```

3.2.3.7.3 Prepositional Adverb Phrase/6,7,8

Examples: *muy hacia dentro, hacia fuera, adondequiera que vayas, afuera*

```
HEAD Prepositional Adverb Phrase ( premodifier, preptype,
    premodifier1, postmodifier, avtype, avsubtype, CH_comp,
    degree ):
```

```

Prepositional Adverb Phrase ( premodifier, preptype,
    premodifier1, postmodifier, avtype, avsubtype, CH_comp,
    degree );
{ premodifier :: NOTHING },
    Prepositional Adverb Phrase ( preptype, premodifier1,
        postmodifier, avtype, avsubtype, CH_comp, degree );
{ premodifier :: NOTHING }, { premodifier1 :: NOTHING },
    { postmodifier :: NOTHING }, { CH_comp :: NIL },
    Prepositional Adverb ( preptype, avtype, avsubtype,
        preptype1, degree, +STRESS ).

```

Prepositional phrases might be added modifiers. At the moment, we have only formalized this possibility for the prepositional versions of AvPs, but the same formalization scheme could be easily applied to prepositional versions of other phrases.

The first alternative in the rule above, which is further developed in next paragraph 3.2.3.7.3.3, accounts for premodified prepositional AvPs, the second, which is further developed in paragraph 3.2.3.7.3.1, for unmodified prepositional AvPs, and the third, finally, for prepositions merged with adverbs. The invisible non-terminal **HEAD Prepositional Adverb Phrase** is used to combine all these possibilities for further integration within the grammar.

3.2.3.7.3.1 Prepositional Adverb Phrase/7

```

Prepositional Adverb Phrase ( preptype, premodifier, postmodifier,
    avtype, avsubtype, CH_comp, degree ):
Prepositional LINKER ( preptype ),
    LINKED Adverb Phrase ( premodifier, postmodifier, avtype,
        avsubtype, CH_comp, degree );
{ premodifier :: NOTHING },
    Prepositional Adverb Phrase ( preptype, postmodifier, avtype,
        avsubtype, CH_comp, degree ).

```

The first alternative of this rule is self-explanatory, the second accounts for prepositional adverbs which can introduce a postmodifier, *adondequiera que vayas*. The non-terminal **Prepositional Adverb Phrase/6**, described in the next paragraph 3.2.3.7.3.2, is used to formalize this possibility. At this point, however, it is worthwhile to remark that this structure is not formalized as an addition of postmodifiers to prepositional AvPs, but as a prepositional version of an AvP that includes a postmodifier.

```

LINKED Adverb Phrase ( premodifier, postmodifier, avtype, avsubtype,

```

```
CH_comp, degree ):
HEAD Adverb Phrase ( premodifier, postmodifier, avtype,
    avsubtype, CH_comp, degree ).
```

3.2.3.7.3.2 Prepositional Adverb Phrase/6 The addition of modifiers to **Prepositional Adverb Phrase/6** is formalized according to the AGFL General Frame for Addition of Modifiers presented in Section 3.2.3.2.2. The rewrite rules constituting the corresponding subset of rules are the following:

```
HIDDEN RECURSIVE Prepositional Adverb Phrase ( preptype,
    postmodifier, avtype, avsubtype, CH_comp, degree ):
Prepositional Adverb Phrase ( preptype, postmodifier, avtype,
    avsubtype, CH_comp, degree );
HIDDEN FINAL Prepositional Adverb Phrase ( preptype,
    postmodifier, avtype, avsubtype, CH_comp, degree ).
```

```
HIDDEN FINAL Prepositional Adverb Phrase ( preptype, NOTHING,
    avtype, avsubtype, CH_comp, degree ):
{ CH_comp :: NIL },
Prepositional Adverb ( preptype, avtype, avsubtype,
    preptype1, degree, +STRESS ).
```

```
Prepositional Adverb Phrase ( preptype, RELCLAUSE, avtype,
    avsubtype, CH_comp, degree ):
Prepositional Adverb Phrase NUCLEUS ( preptype, RELCLAUSE,
    avtype, avsubtype, CH_comp, degree ),
Relative Clause MODIFIER ( AdvP ).
```

```
Prepositional Adverb Phrase NUCLEUS ( preptype, RELCLAUSE, avtype,
    avsubtype, CH_comp, degree ):
HIDDEN RECURSIVE Prepositional Adverb Phrase ( preptype,
    no_RELCLAUSE, avtype, avsubtype, CH_comp, degree ).
```

3.2.3.7.3.3 Prepositional Adverb Phrase/8 The addition of modifiers to **Prepositional Adverb Phrase/7** yields **Prepositional Adverb Phrase/8**, and it is also formalized according to the AGFL General Frame for Addition of Modifiers presented in Section 3.2.3.2.2.

```
HIDDEN RECURSIVE Prepositional Adverb Phrase ( premodifier, preptype,
    premodifier1, postmodifier, avtype, avsubtype, CH_comp,
    degree ):
```

Prepositional Adverb Phrase (premodifier, preptype, premodifier1,
 postmodifier, avtype, avsubtype, CH_comp, degree);
 HIDDEN FINAL Prepositional Adverb Phrase (premodifier, preptype,
 premodifier1, postmodifier, avtype, avsubtype, CH_comp,
 degree).

HIDDEN FINAL Prepositional Adverb Phrase (NOTHING, preptype,
 premodifier, postmodifier, avtype, avsubtype, CH_comp,
 degree):
 Prepositional Adverb Phrase (preptype, premodifier,
 postmodifier, avtype, avsubtype, CH_comp, degree).

• Degree MODIFIER

Example: *muy hacia fuera*

Prepositional Adverb Phrase (DEGREE, preptype, premodifier,
 postmodifier, avtype, avsubtype, CH_comp, degree+):
 Degree MODIFIER (avtype1, avsubtype1, CH_comp, degree+),
 Prepositional Adverb Phrase NUCLEUS (DEGREE, preptype,
 premodifier, postmodifier, avtype, avsubtype, NIL,
 POS|NIL),
 ({ CH_comp :: NIL|PARTICULAR };
 (Second Term MODIFIER (second_term_type, CH_comp, degree+,
 QUE|COMO);
 { CH_comp :: PARTICULAR+ },
 Second Term MODIFIER (COMPTVE, CH_comp, degree+, DE,
 THIRD, NEUT, SING))).

• Comparison Quantifier MODIFIER

Example: *mucho más hacia dentro*

Prepositional Adverb Phrase (QUANT, preptype, premodifier,
 postmodifier, avtype, avsubtype, CH_comp, pos++):
 { pos++ :: POS++|POS+|POS-|EQUAL|SUP },
 Comparison Quantifier MODIFIER (avtype1, avsubtype1, CH_comp,
 pos++),
 Prepositional Adverb Phrase NUCLEUS (QUANT, preptype,
 premodifier, postmodifier, avtype, avsubtype, particular,
 NoEQUAL),
 [Second Term MODIFIER (second_term_type, CH_comp, pos++,
 QUE|COMO)].

• (Relative) Comparison Quantifier MODIFIER

Example: *cuanto más hacia abajo*

Prepositional Adverb Phrase (REL, preptype, premodifier,

postmodifier, avtype, avsubtype, CH_comp, degree):
 Comparison Quantifier MODIFIER (REL, CUANTO, CH_comp, degree),
 Prepositional Adverb Phrase NUCLEUS (REL, preptype,
 premodifier, postmodifier, avtype, avsubtype, particular,
 NoEQUAL).

Prepositional Adverb Phrase NUCLEUS (QUANT|REL, preptype,
 premodifier, postmodifier, avtype, avsubtype, CH_comp,
 degree):
 HIDDEN RECURSIVE Prepositional Adverb Phrase (DEGREE, preptype,
 premodifier, postmodifier, avtype, avsubtype, CH_comp,
 degree).

Prepositional Adverb Phrase NUCLEUS (DEGREE, preptype, premodifier,
 postmodifier, avtype, avsubtype, CH_comp, degree):
 HIDDEN RECURSIVE Prepositional Adverb Phrase (NOTHING, preptype,
 premodifier, postmodifier, avtype, avsubtype, CH_comp,
 degree).

3.2.3.7.4 Prepositional Nominal Phrase/9

Example: *de la gran huida*

Prepositional Nominal Phrase (preptype, ntype, premodifier, dttype,
 dtsubtype, definiteness, person, gender, number):
 Prepositional LINKER (preptype),
 LINKED Nominal Phrase (ntype, premodifier, dttype, dtsubtype,
 definiteness, person, gender, number).

LINKED Nominal Phrase (ntype, premodifier, dttype, dtsubtype,
 definiteness, person, gender, number):
 Nominal Phrase (ntype, premodifier, dttype, dtsubtype,
 definiteness, person, gender, number).

Examples: *del otro coche, del coche, del coche ese, del coche ese de la ventanilla rota*

Prepositional Nominal Phrase (preptype, ntype, NOTHING, dttype,
 dtsubtype, definiteness, person, gender, number):
 Prepositional LINKER (preptype, dttype, dtsubtype, PRE,
 definiteness, gender, number),
 (Nominal Phrase NOMINAL II (I_dttype, I_dtsubtype, ntype,
 premodifier, postmodifier, CH_comp, degree,
 definiteness1, person, gender, number),
 Top Determiner Level (number, dttype, dtsubtype, TWO),
 Determiner Transition Level (number, I_dttype, dttype,
 I_dtsubtype, dtsubtype, ONE);
 Nominal Phrase NOMINAL I (ntype, premodifier, postmodifier,

```

        CH_comp, degree, person, gender, number ),
        eval character degree ( dsubtype, CH_comp, degree ),
        Top Determiner Level ( number, dttype, dsubtype, ONE ) );
Prepositional LINKER ( preptype, dttype, dsubtype, PRE,
        definiteness, gender, number ),
Nominal Phrase NOMINAL I ( ntype, NOTHING|ADJECTIVE,
        NOTHING|AJPHRASE, NIL, NIL|POS, person, gender, number ),
DETERMINER ( I_dttype, I_dsubtype, POST, definiteness1,
        gender, number ),
Top PostDeterminer Level ( number, dttype, dsubtype, TWO ),
PostDeterminer Transition Level ( number, dttype, I_dttype,
        dsubtype, I_dsubtype, ONE );
Prepositional LINKER ( preptype, dttype, dsubtype, PRE,
        definiteness, gender, number ),
Nominal Phrase NOMINAL I ( ntype, NOTHING|ADJECTIVE,
        NOTHING|AJPHRASE, NIL, NIL|POS, person, gender, number ),
DETERMINER AND MODIFIER ( I_dttype, I_dsubtype, definiteness1,
        recursive_post, NIL|PARTICULAR|PARTICULAR+, degree,
        person, gender, number ),
Top PostDeterminer Level ( number, dttype, dsubtype, TWO ),
PostDeterminer Transition Level ( number, dttype, I_dttype,
        dsubtype, I_dsubtype, ONE ).

```

The second rule accounts for **Prepositional Nominal Phrases** introduced by prepositions merged with determiners (*al, del*). In this rule non terminal **Prepositional LINKER/7** is combined with all the non-terminals that can be combined with the determiner *el* for constituting **Nominal Phrases**. A similar rule is necessary for nouns beginning by [á-], but we will not display here this rule, which can be inferred both from the previous one and from other rules, showed in previous Sections of this exposition, that account for this kind of nouns in different contexts.

3.2.3.7.5 Prepositional Pronominal Phrase/10

Example: *de los otros*

```

Prepositional Pronominal Phrase ( preptype, premodifier, dttype,
        dsubtype, prtype, prsubtype, definiteness, person, gender,
        number ):
Prepositional LINKER ( preptype ),
LINKED Pronominal Phrase ( premodifier, dttype, dsubtype,
        prtype, prsubtype, definiteness, person, gender,
        number ).
LINKED Pronominal Phrase ( premodifier, dttype, dsubtype, prtype,

```

prsubtype, definiteness, person, gender, number):
Pronominal Phrase (premodifier, dttype, dtsubtype, prtype,
prsubtype, definiteness, person, gender, number).

Example: *del otro*

Prepositional Pronominal Phrase (preptype, NOTHING, dttype,
dtsubtype, prtype, prsubtype, definiteness, person, gender,
number):
Prepositional LINKER (preptype, dttype, dtsubtype, position,
definiteness, gender, number),
(Pronominal Phrase NOMINAL II (I_dttype, I_dtsubtype,
position, person, gender, number, prtype, prsubtype,
premodifier, postmodifier),
Pronominal Phrase Top Determiner Level (gender, number,
dttype, dtsubtype, TWO),
Pronominal Phrase Determiner Transition Level (gender,
I_dttype, I_dtsubtype, dttype, dtsubtype, ONE);
Pronominal Phrase NOMINAL I (prtype, prsubtype,
premodifier, postmodifier, definiteness1, person,
gender, number),
Pronominal Phrase Top Determiner Level (gender, number,
dttype, dtsubtype, ONE),
Pronoun Determiner Transition (gender, number, prtype,
prsubtype, dttype, dtsubtype)).

3.2.3.7.6 Prepositional Nominalization/6,7 The first two rules account for prepositional nominalizations of adjective and prepositional phrases. The second two rules, instead, account for prepositional nominalizations of relative clauses.

Examples: *de la verde, de la de la esquina*

Prepositional Nominalization (preposition, MDcategory, definiteness,
person, gender, number):
Prepositional LINKER (preposition),
LINKED Nominalization (MDcategory, definiteness, person,
gender, number);
Prepositional LINKER (preposition, dttype, dtsubtype, position,
definiteness, gender, number),
Nominalization NOMINAL I (MDcategory, person, gender,
number).
LINKED Nominalization (MDcategory, definiteness, person, gender,
number):
Nominalization (MDcategory, definiteness, person, gender,
number).

Examples: *Se lo daremos* al que podamos dárselo, *Estamos hablando* de a los que les pasó lo del derrumbe, *Desde hace mucho tiempo no hemos tenido noticias* de los que salieron primero, *Nadie sabe nada* del que salió solo

```

Prepositional Nominalization ( preposition, MDcategory, definiteness,
    person, gender, number ):
    LINKED Nominalization ( MDcategory, preposition, definiteness,
        person, gender, number );
Prepositional LINKER ( preposition ),
    LINKED Nominalization ( MDcategory, pretype, definiteness,
        person, gender, number );
Prepositional LINKER ( pretype, dttype, dtsubtype, position,
    definiteness, gender, number ),
    Nominalization NOMINAL I ( MDcategory, NIL, person, gender,
        number ).

LINKED Nominalization ( MDcategory, pretype, definiteness, person,
    gender, number ):
    Nominalization ( MDcategory, pretype, definiteness, person,
        gender, number ).

```

The first alternative deals with nominalizations of relative clauses introduced by prepositional relative phrases (*a lo que, al que, etc.*, see Section 3.2.3.5). If the preposition required at the level of the prepositional nominalization —main clause— is the same required at the level of the prepositional phrase —within the boundaries of the relative clause—, the preposition must be shared by both levels of analysis. That is, the first alternative identifies **Prepositional Nominalization** non-terminals introduced by prepositions which have a syntactic function also within the limits of the nominalized **Relative Clause** non-terminal. Formally, the preposition is analysed at the level of the relative phrase, while at the level of the prepositional nominalization —the one represented by this rule—, no **Prepositional LINKER** is found, and affixes *preposition* of both **Prepositional Nominalization** and **LINKED Nominalization** are unified.

On the contrary, the second alternative deals with prepositional nominalizations of relative clauses introduced by prepositions that have a syntactic function only at the level of the prepositional nominalization itself (that is, with respect to the main clause). The nominalized relative clause, on the other hand, might or might not be introduced by a prepositional relative phrase. Formally, non-terminals **Prepositional LINKER** and **Prepositional Nominalization** have the same value for affix *preposition*, whereas affix *pretype* of non-terminal

LINKED Nominalization is not restricted at all —it will have value **NIL** if the nominalized relative clause is not introduced by a prepositional relative phrase and, otherwise, it will have one of the values in the domain of affix non-terminal *preposition*.

The third alternative, finally, deals with prepositional nominalizations of relative clauses not introduced by prepositional phrases, more specifically it deals with these prepositional nominalizations when the prepositions, which belong to the main clause, and the determiners, which mark the nominalization itself, are merged, *del que, al que*.

3.2.3.7.7 Prepositional Clause/3,4

Example: *Me acuerdo perfectamente* de que me traicionaste

```
Prepositional Clause ( preposition, clause_type, non_imperative ):
  { clause_type :: QUE },
  Prepositional LINKER ( preposition ),
  LINKED Clause ( clause_type, non_imperative ).
```

Examples: *No me acuerdo nunca* de preguntarle por ella, *Al llegar ella se calló todo el mundo*

```
Prepositional Clause ( preposition, clause_type, INFINITIVE,
  person ):
  Prepositional LINKER ( preposition ),
  LINKED Clause ( clause_type, INFINITIVE, person );
  { clause_type :: EL_SUBORDINATED },
  Prepositional LINKER ( preposition, dttype, dtsubtype,
  position, definiteness, gender, number ),
  LINKED Clause ( SUBORDINATED, INFINITIVE, person ).
```

3.2.3.7.8 Prepositional Interrogative Clause/4

Examples: *No me acuerdo* de si vale más o menos de lo que dices, *No consigo acordarme* de quién me lo dijo

```
Prepositional Interrogative Clause ( preposition, clause_type,
  question_type, non_imperative ):
  Prepositional LINKER ( preposition ),
  LINKED Interrogative Clause ( clause_type, question_type,
  non_imperative ).
```

```
LINKED Interrogative Clause ( clause_type, question_type,
  non_imperative ):
```

Interrogative Clause (clause_type, question_type,
non_imperative).

3.2.3.7.9 Prepositional Relative Clause/5

Examples: *Se lo daremos a quien podamos dárselo, No se obligará a nada a quien no lo desee*

Prepositional Relative Clause (preposition, argument, person,
gender, number):
LINKED Relative Clause (argument, preposition, person, gender,
number);
Prepositional LINKER (preposition),
LINKED Relative Clause (argument, pretype, person, gender,
number).

Example: *Ésas son las ideas de quienes nunca han sido solidarios con los demás*

Prepositional Relative Clause (preposition, LINKED, person, gender,
number):
Prepositional LINKER (preposition),
LINKED Relative Clause (LINKED, pretype, person, gender,
number).

LINKED Relative Clause (RCfunction, pretype, person, gender,
number):
Relative Clause (RCfunction, pretype, person, gender, number).

The first rule accounts for prepositional relative clauses directly dominated by clause-level syntactic functions. The second rule, instead, accounts for prepositional relative clauses directly dominated by phrase-level (**Prepositional MODIFIER**) syntactic functions.

In the first rule, the first alternative deals with relative clauses introduced by prepositional relative phrases (*a quien, a cuanto, etc.*, see Section 3.2.3.6). As in the case of prepositional nominalizations of relative clauses, if the preposition required at the level of the prepositional clause —main clause— is the same required at the level of the prepositional phrase —within the boundaries of the relative clause—, the preposition must be shared by both levels of analysis. Formally, the preposition is analysed at the level of the relative phrase, and the value for affix non-terminal *preposition* is shared by the relative phrase itself, the **Relative Clause** and the **Prepositional Relative Clause**.

The second alternative deals with prepositional relative clauses introduced by prepositions that play a role only at the level of the prepositional relative clause (the relative clause itself might or might not be in turn introduced by a prepositional phrase with a different preposition).

Chapter 4

Ambiguity, Scope and Environment of the Grammar

In Chapters 2 and 3, an AGFL formal grammar was described which deals with certain linguistic processes that expand verbs, nouns, pronouns, adverbs and adjectives to yield VPs, NPs, PrPs, AvPs and AjPs. According to this description, the automatic identification of these structures and their combinations within clauses must be based on a certain, limited, amount of information primarily provided by inflectional characteristics of wordforms —enriched in the case of verbs with verb schemes as described in Chapter 2. Ambiguity occurs when more than one structure or combination of structures are identified for one linguistic sequence.

Ambiguity arises from the fact that inflectional information, even with the addition of verb schemes, is not sufficient to identify a linguistic sequence as the adequate, from the communicative point of view, combination —or sequence, depending on the use made of the grammar, see below— of VPs, NPs, PrPs, AvPs and AjPs structures. That is, for the proposed AGFL grammar, as for formal grammars and parsing in general, the computational problem of ambiguity arises from the fact that language, as a communicative system, makes use of much more information —sometimes of a different nature— than that encoded in current computational lexicons.

While computational lexicons and formal grammars grow and progressively

approach the level of information handled in communication —whatever the real limit of this scope—, this difference of information must be balanced somehow within current applications. Two main resources are available within the AGFL formalism that can be used for this purpose: order of alternatives, given the fact that the parser produces the analyses in the order in which they are found, and *penalties*, which give low priority to the analyses that call upon rules that contain them. Applications themselves, as well as the whole environment in which the parser is going to be applied (i.e. other software-tools involved), must be taken into account when determining the best way of using the grammar and dealing with ambiguity. This, in fact, can be considered a third level of formalization.

In this Chapter, a more detailed account is given of how order and *penalties* are handled to deal with ambiguity for two different applications made of the proposed grammar within two different environments. The first is the grammar for full-sentence analysis itself, in its current state, a work in progress constantly subject to what for the moment is only a laboratory test performed on preselected groups of sentences. The second one is the grammar for phrase analysis being used within the DoRo project for an IR application. This is currently being tested on a user environment, although results are not yet available. A large scale test is, anyway, needed in both cases. However, this task exceeds the purpose of this Chapter, which aims at illustrating how the two different situations determine the formalization itself.

All this Chapter is based on the assumption that, even risking the lost of the correct analysis in certain cases, the grammar will be used to generate a parser taking decisions to solve ambiguities. Of course, if, for instance, for a certain application the main interest is to ensure the collection of correct analyses by means of later manual intervention, the proposed solutions should be revised whenever they introduce penalties somewhere in the rules described by the previous Chapters.

4.1 Full Sentence Analysis

In line with the general design of verb schemes described in Chapter 2, clause analysis, taking into account that most verbs might have more than one verb scheme, tries to identify as much arguments as possible. That is, it tries to identify

a contextualization for the scheme with the highest number of arguments encoded for the verb in question. For this reason, alternatives with more arguments are placed before in rewrite rules of different clauses (that is, the analysis of *alguien le dijo algo a alguien* is found before the analysis of *alguien dijo algo*). Besides, for the same number of arguments, frequency determines the order of alternatives taking into account the more frequent dispositions of the arguments within the clauses found in BDS (that is, the analysis of *alguien dijo algo* is found before the analysis of *algo lo dijo alguien*).

Nevertheless, the order of alternatives cannot ensure by itself that preference will be given to the alternatives with the highest number of arguments in the case of subordinated clauses. In this context, a corrective must be introduced by means of penalties in order to ensure that the highest number of arguments within subordinated clauses will be given priority over the analysis of arguments at the level of the main clause. With this scope, progressively higher numbers of penalties are added to the alternatives that contain lower numbers of arguments within the rewrite rules of subordinated clauses (i.e. the identification of *alguien* as the **SUBJECT** of *dijo* within *dijo que lo había dicho alguien* is found before its identification as the **SUBJECT** of *había dicho* within the subordinated clause. However, the second option is preferred because the analysis of the subordinated clause *que lo había dicho alguien* is less penalized than that of the subordinated clause *que lo había dicho*).

The identification of constituents as **CIRCUMSTANCES**, on the other hand, sometimes interferes with their identification as arguments. Arguments are, nevertheless, preferred, due to the fact that the identification of **CIRCUMSTANCES** is always found later during the analysis (i.e. the analysis of *alguien dijo ayer algo* is found later than the analysis of *alguien dijo algo*).

So far in this Section we have given an outline of how clause rewrite rules are organized in our sentence grammar. Therefore at this point we are prepared to start the main argument here, which is the description of how the attempt to handle ambiguity determines a number of aspects of phrase-level formalization for its integration in a sentence grammar for full-sentence analysis.

4.1.1 Largest Constituents

The preference for the highest number of arguments needs to be applied a corrective at the level of phrase analysis. As a consequence of this, the identification of arguments will be the result of the tension between clause preference for the highest number of arguments and phrase preference for the largest constituents. As explained above, the highest number of arguments is given priority by means of the order of alternatives in the case of main clauses, by means of penalties in the case of the subordinated ones. Hence, to be given priority over analyses with a higher number of arguments, analyses with larger constituents but with a lower number of arguments, need to be less penalized. In order to produce this effect, we add penalties to all syntagms that can be placed in both nucleus and modifier positions whenever they are placed in nucleus positions. The number of penalties added must be obviously higher than that added to progressively shorter subordinated clauses, i.e. higher than the difference of penalties between, for instance, subordinated clauses with two arguments and subordinated clauses with three arguments. All this might be better understood looking at the concrete points of the grammar where the introduction of penalties prevents the interpretation of something as a constituent if it can be integrated, as a modifier or determiner, with an adjacent item or syntagm¹.

4.1.1.1 verb type priorities

Predicate **verb type priorities**, see Section 2.2.3.2.2, ensures that auxiliary and semiauxiliary verbs will always be preferred to main ones. Therefore, if possible, adjacent verb forms will always be combined for the constitution of a unique **PREDICATE** constituent.

4.1.1.2 determiner better than pronoun

This predicate ensures that all the items that can be interpreted either as determiners or as pronouns will be, if possible, interpreted as determiners.

¹Of course, the preference for the largest proper names is also a kind of preference for the largest constituents. The regular expressions that identify proper names, combined with the action of penalties ensure that the largest proper name possible will always be preferred. However, see Section 3.2.3.1, these rules are not showed in this exposition.

For instance, by means of this predicate, the analysis preferred for sequence *Le dijo que cogiera eso otro* is that which identifies *eso otro* as the **DIRECT OBJECT** of the verb *cogiera*, and not that which identifies *eso* and *otro* as, respectively, the **DIRECT OBJECT** and the **SUBJECT** of the same verb. The analysis that identifies *otro* as the **SUBJECT** of *dijo*, on the other hand, is rejected both because of its lower number of arguments within the subordinated clause and because of its interpretation of *eso* as a pronoun and not as a determiner.

Formally, the analyses are found in the following order:

- (1) *Le dijo que cogiera eso otro*: *otro* as **SUBJECT** of *dijo*².
- (2) *Le dijo que cogiera eso otro*: *otro* as **SUBJECT** of *cogiera*.
- (3) *Le dijo que cogiera eso otro*: *eso otro* as **DIRECT OBJECT** of *cogiera*.

Because of their lower number of arguments within the subordinated clause, analyses (1) and (3) are given +1 penalty with respect to analysis (2). Because of their interpretation of *eso* as a pronoun, analyses (1) and (2) are given +2 penalties with respect to analysis (3), thus the preferred analysis is (3)³.

Predicate **determiner better than pronoun** is called upon by the rules that rewrite **Pronoun** non-terminals into **PronounSt** non-terminals from the lexicon. These rules were already showed in Section 3.2.3.1, after the introduction of the predicate **determiner better than pronoun**, they look as follows:

```
Pronoun ( prtype, prsubtype, definiteness, gender, number ):
  ( PronounSt ( prtype, prsubtype, definiteness, gender, number );
    { prtype :: CARD }, { definiteness :: UNDEFINITE },
      ( { prsubtype :: NoUNIT },
        Written Cardinal Number ( numbering, gd, last, POST,
          gender, number );
        ( { prsubtype :: UNO }, { number :: SING };
          { prsubtype :: NoUNIT }, { number :: PLU } ),
          { gender :: MASC|FEM },
          Arabic Numbers ( number ) );
```

² In this and the following example, Section 4.1.1.3, the ambiguity is, in fact, higher than that which we present here, because, for instance, *que cogiera eso* and *otro*, on purely syntactic grounds, might also interchange their functions, being either **SUBJECT** or **DIRECT OBJECT** in different parse trees. However, these ambiguities, which are easily solved by pure order of alternatives, are not taken into consideration here, to facilitate the understanding of the mechanism that regulates the preference for the largest constituents.

³The AGFL analysis is showed in Appendix G.2.

```

{ prtype :: ORD }, { prsubtype :: CUARTO },
  { definiteness :: UNDEFINITE },
  Written Ordinal Number ( numbering, POST, gender, number );
{ prtype :: PART }, { prsubtype :: CUARTO },
  { definiteness :: UNDEFINITE }, { gender :: MASC },
  Written Partitive Number ( numbering, POST, gender, number ) ),
determiner better than pronoun ( prtype, prsubtype ).

```

```

Pronoun ( prtype, prsubtype, definiteness, person, gender, number ):
PronounSt ( prtype, prsubtype, definiteness, person, gender,
  number ),
determiner better than pronoun ( prtype, prsubtype ).

```

The rewrite rule of the predicate adds 2 penalties each time that a pronoun that can also be a determiner is actually interpreted as a pronoun:

```

determiner better than pronoun ( prtype, prsubtype ):
{ prtype :: INTG }, { prsubtype :: QUIEN|CUAL };
$PENALTY ( 2 ),
  { prtype :: INTG }, { prsubtype :: QUE|CUANTO };
{ prtype :: REL }, { prsubtype :: CUAL|QUIEN|QUE };
$PENALTY ( 2 ),
  { prtype :: REL }, { prsubtype :: CUANTO };
$PENALTY ( 2 ),
  { prtype :: GLOBAL|DEM|POSS|REFL|QUANT|DIS|COMPTVE };
$PENALTY ( 2 ),
  { prtype :: INDEF|CARD }, { prsubtype :: UNO|NoUNIT|
    NoUNIT+UNO };
{ prtype :: INDEF }, { prsubtype :: ALGUIEN|QUIENQUIERA };
$PENALTY ( 2 ),
  { prtype :: INDEF }, { prsubtype :: ALGUNO|CUALQUIERA|
    OTRO|DEMÁS };
$PENALTY ( 2 ),
  { prtype :: PART|ORD }, { prsubtype :: CUARTO|MEDIO }.

```

There exists, however, one exception to this norm, because proper names preceded by pronoun-determiner ambiguous wordforms should not be combined with them for the constitution of one NP. Predicate **ntype NOMINAL** is used to balance the effect of predicate **determiner better than pronoun** in this case. This predicate has the following form:

```

ntype NOMINAL ( COMMON ): .
ntype NOMINAL ( proper ): $PENALTY ( 2 ).

```

This predicate is included in the rules that account for the addition of determiners to NPs or nouns. If the noun in question is a proper name, two penalties

are added when a determiner is found which expands the noun. As the analysis of this determiner as a pronoun, which is also added two penalties, is found before its analysis as a determiner, the analysis preferred is that one. For instance, the first rewrite rule of non-terminal **Nominal Phrase** described in Section 3.2.3.3.4.1 remains as follows:

```
Nominal Phrase ( ntype, NOTHING, dtno_COMPTVE, dtsubtype,
  definiteness, person, gender, number ):
  DETERMINER ( dtno_COMPTVE, dtsubtype, PRE, definiteness, gender,
    number ),
  Nominal Phrase NOMINAL I ( ntype, premodifier, postmodifier,
    CH_comp, degree, person, gender, number ),
  eval character degree ( dtsubtype, CH_comp, degree ),
  ntype NOMINAL { ntype },
  Top Determiner Level ( number, dtno_COMPTVE, dtsubtype, ONE ).
```

Obviously, both **verb type priorities** and **determiner better than pronoun** become unnecessary if the parser is run over previously disambiguated text, that is, within an environment that uses some kind of *tagger*. In this case, the annotation supplied by the tagger should be considered in detail, because if, for instance, verb forms were not classified into auxiliary, semiauxiliary or main verb forms, predicate **verb type priorities** could not be removed whatsoever.

4.1.1.3 Relative Clauses

A similar situation is found in the case of relative clauses, because, when introduced by certain relative phrases, the same internal structures of relative clauses may be used either as modifiers within NPs, PrPs or AvPs, or as arguments within main clauses.

For instance, the analysis that identifies the relative clause *a las que no habían podido responder* as a modifier of *algunas preguntas* in *Dijo que les había hecho algunas preguntas a las que no habían podido responder* is preferred to the analyses that identifies the relative clause as an **INDIRECT OBJECT** of one of the verbs *dijo* or *había hecho*. The example is similar to that introduced in the previous Section. Formally, the analyses are found in the same —relevant— order:

- (1) *Dijo que les había hecho algunas preguntas a las que no habían podido responder*: *a las que no habían podido responder* analysed as **INDIRECT OBJECT** of *dijo*.

- (2) *Dijo* que les había hecho algunas preguntas a las que no habían podido responder: *a las que no habían podido responder* analysed as **INDIRECT OBJECT** of *había hecho*.
- (3) *Dijo* que les había hecho algunas preguntas a las que no habían podido responder: *a las que no habían podido responder* analysed as **MODIFIER** of *algunas preguntas*.

Because of their lower number of arguments within the subordinated clause, analyses (1) and (3) are given +1 penalty with respect to analysis (2). Because of their interpretation of the —nominalized— relative clause as an argument, analyses (1) and (2) are given +2 penalties with respect to analysis (3), thus the preferred analysis is (3)⁴.

In these examples, it is only the function of the relative clause what is changed from one interpretation to the other, whereas the category of the item that supports the relative phrase that introduces the relative clause is always the same. Therefore the penalties must be associated with the assignment of function to the relative clause, or, in other words, with the rules that rewrite **Relative SUBJECT**, **DIRECT OBJECT**, etc. into relative phrases, given the fact that the functional capabilities of relative clauses are determined by the relative phrases that introduce them. These rules were explained in Sections 3.2.3.4.1.3, 3.2.3.5 and 3.2.3.6.

With respect to these rules as they were presented in the mentioned Sections, we can observe here that all the alternatives that rewrite **Relative CONSTITUENT** into a relative phrase that allows for the relative clause to function as an argument of the main clause will be added penalties whenever there is another alternative for the same relative phrase that allows for the relative clause to function as a modifier within a constituent of the main clause. For instance, there are two alternatives that rewrite **Relative CONSTITUENT** into relative phrase *a quien*, and we can observe that, while the one which assigns values in the domain of *argument* for affix **RCfunction** is penalized, the one which assigns value **AGR_MOD** is not. The same happens with relative phrases *al que*, *a la que*, etc., except for the fact that the penalized value for affix *RCfunction* in this case

⁴The AGFL analysis is showed in Appendix G.2.

is **NOMINAL**. The alternatives that rewrite **Relative CONSTITUENT** into these phrases remain then as follows:

```
Relative CONSTITUENT ( RCfunction, preptype, person, gender, number ):
  (...);
  Prepositional Pronoun Phrase ( preptype, prmd_general, postmodifier,
    REL, QUIEN, definiteness, person, gender, number ),
    { RCfunction :: AGR_MOD };
  $PENALTY ( 2 ),
  Prepositional Pronoun Phrase ( preptype, prmd_general,
    postmodifier, REL, QUIEN, definiteness, person, gender,
    number ),
    { RCfunction :: DO|IO|PC|PR|AG|LINKED };
  Prepositional Pronominal Phrase ( preptype, prmd_general,
    GLOBAL|ART, NIL|EL, REL, QUE, definiteness, person, gender,
    number ),
    { RCfunction :: AGR_MOD };
  $PENALTY ( 2 ),
  Prepositional Pronominal Phrase ( preptype, prmd_general,
    GLOBAL|ART, NIL|EL, REL, QUE, definiteness, person, gender,
    number ),
    { RCfunction :: NOMINAL };
  (...).
```

The ambiguity generated by semiauxiliary/main verbs, determiner/pronoun items and relative clauses is very important. The examples showed here, which can be considered as communicatively ambiguous—at least, within the communicative context that can be showed here—are the less frequent ones. The computational problem, on the contrary, actually affects much more examples than from the communicative point of view could never be considered ambiguous whatsoever. That is, for the parser, sequences as *Dijo que cogieran una entrada* and *Dijo que les había hecho algunas preguntas a las que tendrían que dar muchas vueltas antes de responder* are, from the point of view of ambiguity, exactly the same instances as, respectively, *Dijo que cogiera eso otro* and *Dijo que les había hecho algunas preguntas a las que no habían podido responder*. Obviously, more information than the purely inflectional one, needs to be included in the grammar to block the wrong analyses in these cases. The nature of this information—semantic, probabilistic, more accurate grammatical information, or, probably, a combination of all of them—is still a matter for study.

4.1.1.4 Prepositional MODIFIER

In contrast to relative clauses in similar contexts, the interpretation of prepositional phrases as **Prepositional MODIFIERS** is not forced. On the one hand, the information about the prepositional requirements of verbs, together with our grammatical knowledge about the prepositions that introduce other syntactic functions, is considered an acceptable guaranty to decide on the clause argument status of prepositional phrases.

On the other hand, prepositional phrases might also function as **CIRCUMSTANCES** at clause-level analysis. To distinguish these from **Prepositional MODIFIERS**, we only have the preposition that introduces the prepositional phrase: if it is preposition *de* —the affix *preptype* takes value **DE**—, the prepositional phrase is considered a **Prepositional MODIFIER**, if it is another preposition —the affix *preptype* takes a value in the domain of *no_DE*—, the prepositional phrase is considered a **CIRCUMSTANCE**. The rewrite rule of **Prepositional Phrase MODIFIER**, introduced in Section 3.2.3.4.1.2, remains thus as follows:

```
Prepositional Phrase MODIFIER ( preposition ) :
  { preposition :: DE },
  ( HEAD Prepositional Adverb Phrase ( premodifier, preposition,
    premodifier1, postmodifier, PLACE, avsubtype, CH_comp,
    degree );
    (...);
    Prepositional Clause ( preposition, clause_type, INFINITIVE,
    person ) ).
```

4.1.2 Ambiguity within Constituents

Within the boundaries of one constituent, many ambiguities can be solved by means of order alone. Some examples are the following:

4.1.2.1 *el que*

The rewrite rules of clause constituents determine the preferred analysis when a subordinated clause introduced by *el que* —either a nominalized relative clause or a *that*-clause preceded by determiner *el*— cannot be disambiguated on the basis of its internal structural analysis (*El que no hubiera contestado no era admitido*

por nadie). The rewrite rule of **SUBJECT** in Section 3.2.3.6.1 gives preference to the relative clause interpretation.

4.1.2.2 (Quantifier) + (*más/menos*) + Adjective + Noun Quantifier + *más/menos*

The order of alternatives in the rewrite rules of clause constituents determines also whether, placed before a noun preceded by an adjective, *poco*, *mucho*, *bastante*, *demasiado*, sometimes followed by *más*, should be interpreted as adverbs, determiners or sequences of determiner and adverb (*bastante más buen vino*). The same situation arises when there is only *más/menos* preceding an NP constituted by a noun modified by an adjective (*más buenos vinos*). The rewrite rule of **SUBJECT** in Section 3.2.3.6.1 gives preference to the first interpretation. That is, if possible, the ambiguity is solved in favour of adverbs.

Quantifiers and comparatives *más/menos* when they form one constituent pose a very similar problem. Apart from the fact that they can also constitute an AvP, when they constitute a PrP, we still have to decide whether the comparative is the pronoun nucleus and the quantifier is the determiner or vice versa (the fact that *muchas más* can be related with both NPs *muchas más botellas* and *muchas botellas más* logically has this consequence within our descriptive frame, see Section 3.1.3.2). The order of alternatives in the relevant rewrite rules gives preference to the first possibility.

However, what appears to be particularly remarkable with respect to the ambiguities described in this Section is the fact that none of them can be solved on the basis of grammatical or lexical information. Only in certain cases, they can be solved on the basis of textual information. For instance, if the NP underlying the PrP constituted by a quantifier and a comparative is made explicit in a different point of the text. Nevertheless, this may not be the case and, no doubt, this is not a question for a formal grammar as the one described here. These are, in fact, very good examples to point out that not all the distinctions drawn within more or less traditional descriptive frames of grammars are appreciated by native speakers and might be corroborated by evidence from the computer. In these cases indeed, the computer seems to be corroborating that, on the contrary, some theoretical distinctions are not distinguishable.

4.1.2.3 Internal Structure of Modifiers

When a modifier is added to a nucleus, if the modifier can itself contain the same type of modifiers —prepositional NPs that function as prepositional modifiers are the best example—, the hierarchy of the whole structure is difficult to determine. For instance, in *la pierna del paciente que habría habido que tratar*, the relative clause *que habría habido que tratar* may be a modifier either of *paciente* or of *pierna del paciente*. More frequently, of course, only one possibility is correct, but more information of a different —semantic— nature is needed for disambiguation (*la mesa de madera del niño, la carta de la hija del vecino, la pregunta de la prueba a la que pretendía responder*).

In the current stage of the grammar, the analysis preferred is that which connects all the modifiers in question with the main nucleus, in our example the analysis that connects *que habría habido que tratar* with *pierna de ese paciente*. This is ensured by a group of guards that organizes the values possibly taken by affix *postmodifier* within the rules that rewrite the relevant **HEAD** or **HIDDEN RECURSIVE** non-terminals. For instance, the rewrite rule of **HEAD Noun Phrase**, introduced in Section 3.2.3.2.4.1, remains as follows:

```
HEAD Noun Phrase ( ntype, premodifier, postmodifier, CH_comp, degree,
  definiteness, person, gender, number ):
  ( { postmodifier :: RELCLAUSE };
    { postmodifier :: PRPPHRASE };
    { postmodifier :: AJPHRASE };
    { postmodifier :: HOMO_APP };
    { postmodifier :: NOTHING } ),
  Noun Phrase ( ntype, premodifier, postmodifier, CH_comp, degree,
    definiteness, person, gender, number );
HIDDEN FINAL Noun Phrase ( ntype, premodifier, postmodifier,
  CH_comp, degree, definiteness, person, gender, number ).
```

Contrary to this, in the case of nominalizations the absence of modifiers is given priority in order to force the analysis of the possible modifiers within the nominalized linguistic unit, instead of associating them with the nominalization itself. That is, in *la de la casa que compramos*, the relative clause *que compramos* is considered a modifier of noun *casa* within the NP *la casa que compramos*, which is after nominalized as a prepositional NP. The analysis that considers the relative clause as a modifier of the nominalized prepositional NP *de la casa* is

instead found later and, consequently, it is given lower priority. For this reason, in the rewrite rules of non-terminal **Nominalization NOMINAL I/4,5**, found in Sections 3.2.3.5.2 and 3.2.3.5.7, the alternative that rewrites this non-terminal into non-terminal **HIDDEN FINAL Nominalized Phrase**, which appears only under the condition that no modifiers are associated with the nominalized unit, is placed before the alternative that rewrites it into non-terminal **Nominalized Phrase**, which appears only under the condition that some modifier is associated with the nominalized unit.

4.1.3 Lexical Ambiguity

This is considered the core of the problem of ambiguity, and, as a matter of fact, a great number of the so-called structural ambiguities can be ultimately related to lexical ambiguities. Moreover, it is indeed very possible that, if we would run the parser after a tagger, over previously disambiguated text, not only the problems described in Sections 4.1.1.1, 4.1.1.2, 4.1.2.1 and 4.1.2.2, but also other problems not accounted for here, would all disappear. However, given the current state of the art for taggers, most of these problems would require manual intervention to be solved with more guarantees than those offered by the solutions integrated in the grammar in its current stage. The best disambiguator, we have to be aware of it, is the grammar itself. The reasons thus to include taggers in order to disambiguate the input for parsers are to be found, on the one hand, on the still partial development of the formal grammars and lexicons, and, on the other, on the improvement of the efficiency of current parsers to ensure the production of more, and more accurate, analyses for more linguistic sequences.

We will not include examples here, because they are very well known, of ambiguous lexical items. We will only remark the fact that almost all the possible pairs of lexical categories can be found in some lexical item, as well as certain groupings of three or four categories. In spite of this and although the evidence that the efficiency of the AGFL parsers continues to be a problematic issue, resolution of lexical ambiguity is in our proposal left to the grammar. Apart from the items involved in Sections 4.1.1.1, 4.1.1.2, 4.1.2.1 and 4.1.2.2, there is only one exception to this decision: the identification of unmodified and undetermined nouns as direct constituents of clauses is penalized to force, for adverb-noun ambiguous

lexical items, their interpretation as adverbs and circumstances. This causes, for instance, that *ayer* in *Ayer vino* will not be interpreted as a noun and the **SUBJECT** of verb *vino*. The rewrite rule for **HEAD Noun Phrase** remains then as follows:

```

HEAD Noun Phrase ( ntype, premodifier, postmodifier, CH_comp, degree,
  definiteness, person, gender, number ):
  ( { postmodifier :: RELCLAUSE };
    { postmodifier :: PRPPHRASE };
    { postmodifier :: AJPHRASE };
    { postmodifier :: HOMO_APP };
    { postmodifier :: NOTHING } ),
  Noun Phrase ( ntype, premodifier, postmodifier, CH_comp, degree,
    definiteness, person, gender, number );
$PENALTY ( 2 ),
  HIDDEN FINAL Noun Phrase ( ntype, premodifier, postmodifier,
    CH_comp, degree, definiteness, person, gender, number ).

```

In line with this, we have found that, for certain groupings of lexical categories involving closed classes of words, if we take into account the domain of the application, a simple solution may be the removal of the more infrequent lexicon definitions related with open classes of words. For instance, out of certain domains, simple removal of the ambiguity generated by *de* interpreted either as a preposition or as the name of the letter *d* might be the best solution.

Lexical ambiguity of feature values, on the other hand, considering that synthetic forms are encoded only once in the lexicon and that the AGFL system can handle subsets of affix values, should not generate ambiguity in the sense of multiple parse trees for one and the same input sequence, although ambiguity of features, of course, is maintained and reflected by the analysis obtained. For instance, noun *atleta* is assigned values **MASC|FEM** for *gender*, but even if context cannot help to disambiguate this value —imagine the sequence *Bastantes atletas llegaron con retraso*—, only one analysis is produced. In this analysis, both values **MASC|FEM** are assigned to all *gender* affixes related by unification with noun *atleta*. The main point here is the adequate design of the lexicon, so as to prevent the multiple encoding of equal wordforms when they can take more than one value for one or more affixes.

Lexical ambiguity of lemmas, finally, always generates ambiguity. To avoid this ambiguity, all equal wordform definitions in the lexicon should be reduced to

one lexicon definition. For instance, masculine plural noun *botones*, which can be related to either lemma *botón*, “small objects sewn on to pieces of clothes”, or lemma *botones*, “boy carrying bags in the hotels”, should be encoded only once in the lexicon. Also, different lexicon definitions should be reduced to one lexicon definition if one of them may account for all the contextualizations of the other one. For instance, lexicon definition for noun *árabe*, “people from Arabia”, is assigned values **MASC|FEM** for *gender*, so that we can eliminate lexicon definitions of noun *árabe*, “language spoken in Arabia”, which should be assigned value **MASC** for *gender*, since all the possible contextualizations of the name of the language, from the syntactic and morphological point of view, can be accounted for by the name of the people.

From what we have observed in this Section, we can conclude that grammar and lexicon are intimately related to each other, and that some aspects of ambiguity might already be solved by means of lexicon encoding. From our point of view, lexicon and grammar must be simultaneously developed in order to account for the mutual dependencies between them. The grammar, for certain extensions, can only be developed on the basis of information encoded in the lexicon, but lexicon encoding must also be flexible enough to allow easy refinement of the information encoded whenever this is required by the grammar.

4.2 Phrase Analysis

Phrase analysis identifies sequences of phrases⁵. To do this we make use of both recursion and the invisible non-terminals **PHRASE** and **PHRASES**. The basic AGFL structure that accounts for the recursive identification of phrases is the following:

```
PHRASES :
  PHRASE;
  PHRASE,
  PHRASES.

PHRASE :
  HEAD Noun Phrase ( ntype, prmd_general, postmodifier,
    NIL|PARTICULAR|PARTICULAR+, degree, definiteness, person,
    gender, number );
```

⁵An AGFL example of phrase-level analysis is found in Appendix G.3.

```

Prepositional Noun Phrase ( preptype, ntype, prmd_general,
    postmodifier, CH_comp, degree, definiteness, person, gender,
    number );
(...)
Periphrastic Verb Phrase ( mvtype, voice, person, number,
    mood );
Verb Phrase ( mvtype, voice, person, number, mood );
(...)
Clause ( QUE, non_imperative );
Prepositional Clause ( preposition, QUE, non_imperative );
Clause ( SUBORDINATED|EL_SUBORDINATED, mood, MVperson );
Prepositional Clause ( preposition, SUBORDINATED|EL_SUBORDINATED,
    INFINITIVE, person );
Interrogative Clause ( MAIN|SUBORDINATED, question_type,
    non_imperative );
Prepositional Interrogative Clause ( preposition, SUBORDINATED,
    question_type, non_imperative );
Relative Clause ( CUANTO_MAS|S|DO|IO|PR|PC|AG|C, preptype,
    person, gender, number );
Prepositional Relative Clause ( preposition, argument, person,
    gender, number );
(...).

```

In the first one of these rules non-terminal **PHRASES** is recursively called upon in order to identify all the possible consecutive phrases, each of which analysed as a non-terminal **PHRASE**, found in the input sequence of words. In the second rule, non-terminal **PHRASE** rewrites in turn as all the possible types of phrases described in Chapters 2 and 3. With respect to this typology of phrases, we have to observe that first, for phrase analysis, verb configurations are no longer relevant beyond the boundaries of the VP itself, given the fact that clause-level syntactic functions —clauses— will not be arrived at in this type of analysis. Secondly, we have to point out that the constituents described in Sections 3.1.6 and 3.2.3.6 are also actual possibilities for rewriting non-terminal **PHRASE**. In this case, these constituents are reduced to their main marker, the conjunction or the relative or interrogative phrase that introduces them. This means that, for instance, the rewrite rule of *that*-clauses and prepositional *that*-clauses is reduced in the following way:

```

Clause ( QUE, non_imperative ):
    Conjunctive LINKER ( QUE ).

```

Which ultimately means that only the element that introduces the subordinated clause is identified in its adequate hierarchical position, while the rest of the

constituents are identified at the same level of the constituents of the main clause. Of course, the identification of VPs, even without taking into account verb configurations, could be used to determine —assuming certain risks— the hierarchical position of the phrases that constitute subordinated clauses. However, this solution has not been adopted here because of coordination, which is not yet formalized and should have to be considered in combination with VPs to ensure the correct identification of the boundaries of clauses and the hierarchy of the constituents found in these. In addition to this, the IR application that this development of the grammar was especially dedicated to, could be also adequately satisfied by means of this simpler approach, and so the necessary extension of the grammar to define boundaries of clauses on the basis of VPs was temporarily left out of the project. Apart from *that*-clauses, other subordinated clauses, those described in Sections 3.1.6 and 3.2.3.6, infinitive, interrogative and relative clauses, are reduced in the same way. Infinitive clauses, for instance, are reduced as follows⁶:

```

Clause ( SUBORDINATED, INFINITIVE, MVperson ):
  NON FINITE PREDICATE ( mood ).

NON FINITE PREDICATE ( mood ):
  Periphrastic Non Finite Verb Phrase ( mvtype, voice, mood );
  { voice :: ACT|PASS },
  Non Finite Verb Phrase ( mvtype, voice, mood );
  Periphrastic Non Finite Verb Phrase ( mvtype, voice, mood,
    case_one, person_one, gender_one, number_one );
  ( { voice :: ACT|PASS };
    { voice :: MIDD }, { person_one :: THIRD } ),
  Non Finite Verb Phrase ( mvtype, voice, mood, case_one,
    person_one, gender_one, number_one );
  Periphrastic Non Finite Verb Phrase ( mvtype, voice, mood,
    case_one, person_one, gender_one, number_one, case_two,
    person_two, gender_two, number_two );
  ( { voice :: ACT };
    { voice :: MIDD }, { case_two :: DAT } ),
  Non Finite Verb Phrase ( mvtype, voice, mood, case_one,
    person_one, gender_one, number_one, case_two, person_two,
    gender_two, number_two ).

```

The rewrite rules for interrogative clauses, on the other hand, are the following:

```

Interrogative Clause ( SI, Yes_No, non_imperative ):

```

⁶The need for the guards that restrict the values of affix non-terminals *voice*, *person* and *case* will be clarified below.

Conjunctive LINKER (SI).

Interrogative Clause (MAIN|SUBORDINATED, Wh, non_imperative):
Interrogative CONSTITUENT (Wh).

Interrogative CONSTITUENT (Wh):

Nominal Phrase (ntype, prmd_general, INTG, dtno_declarative, definiteness, person, gender, number);
Prepositional Nominal Phrase (preptype, ntype, prmd_general, INTG, dtno_declarative, definiteness, person, gender, number);
HEAD Pronoun Phrase (prmd_general, postmodifier, INTG, prno_declarative, definiteness, person, gender, number);
Prepositional Pronoun Phrase (preptype, prmd_general, postmodifier, INTG, prno_declarative, definiteness, person, gender, number);
Pronominal Phrase (prmd_general, dtgeneral, dtsubtype, INTG, prno_declarative, definiteness, person, gender, number);
Pronominal Phrase (prmd_general, INTG, dtno_declarative, prgeneral, prsubtype, definiteness, person, gender, number);
Prepositional Pronominal Phrase (preptype, prmd_general, dtgeneral, dtsubtype, INTG, prno_declarative, definiteness, person, gender, number);
Prepositional Pronominal Phrase (preptype, prmd_general, INTG, dtno_declarative, prgeneral, prsubtype, definiteness, person, gender, number);
HEAD Adverb Phrase (prmd_general, postmodifier, INTG, avno_declarative, CH_comp, degree);
HEAD Prepositional Adverb Phrase (prmd_general, preptype, prmd_general1, postmodifier, INTG, avno_declarative, CH_comp, degree).

Only the first interrogative constituent is identified as the interrogative clause. This constituent is given a generic name, since syntactic functions are not taken into consideration in phrase analysis. Obviously, finite and non-finite interrogative clauses cannot be distinguished on the basis of this type of analysis.

The rewrite rule of relative clauses, finally, is the following:

Relative Clause (RCfunction, preptype, person, gender, number):
Relative CONSTITUENT (RCfunction, preptype, person, gender, number).

The rewrite rule of non-terminal **Relative CONSTITUENT** results from the addition of all the alternatives found in the rules for **Relative CONSTITUENT** showed in Sections 3.2.3.4.1.3, 3.2.3.5.8 and 3.2.3.6.5.

4.2.1 Largest Phrases

Phrase analysis is not simply aimed at the segmentation of the input sequences into the phrases contained in them, more specifically what it is aimed at is the segmentation of the input sequences into the phrases that match the syntactic constituents of the clauses found in the sequence in question. This means that the items or embedded phrases that could be identified as modifiers embedded in other phrases should be analysed as part of them and not as phrases by themselves. Therefore, the identification of the largest phrase possible should be always given priority, which is something that can be done by means of the addition of penalties⁷ each time that non-terminal **PHRASES** is recursively called upon by the rule:

```
PHRASES :  
  PHRASE ;  
  $PENALTY ( 10 ) ,  
  PHRASE ,  
  PHRASES .
```

However, as was observed earlier in this Chapter, the combination of consecutive segments to form one constituent is not desirable when one of these segments is a proper name, since proper names are only very rarely modified or determined. In order to balance this merging tendency in the case of proper names, we add a higher number of penalties whenever a proper name is analysed as part of an NP. The rewrite rule of predicate **ntype NOMINAL** within phrase analysis remains then as follows:

```
ntype NOMINAL ( COMMON ) : .  
ntype NOMINAL ( proper ) : $PENALTY ( 15 ) .
```

4.2.2 Optimization

In this Section we primarily deal with the necessary adaptations that allow for the grammar of phrases described in Chapters 2 and 3 to be used in a concrete IR application whose main interest with respect to the parser is the production of one analysis for each input sequence. This obviously means that ambiguity has to be removed from the output of the parser, and, wherever possible, from the grammar itself, ensuring in this manner that the parser will find an analysis very fast,

⁷The motivation for adding ten penalties will be clarified below (see Section 4.2.2.2.1).

without attempting alternative rewrite possibilities created by ambiguities that, as a matter of fact, cannot be solved on the basis of phrase analysis. For this reason, the grammar needs to be optimized, and therefore, while certain ambiguities are simply removed from it, for other, certain analyses are given priority by means of penalties.

4.2.2.1 Removed Ambiguities

4.2.2.1.1 VPs Verb schemes and configurations are obviously necessary for clause analysis, while phrase analysis can do nothing with this kind of information. Therefore, the assignment of verb schemes to verb forms in the lexicon, as described in Section 2.2.2.6, is not carried out when generating lexicons to be used in phrase analysis. This means, on the one hand, that values for *vtype* and *voice* of verb forms do not take a concrete value within lexicons for phrase analysis and, on the other, that prepositional requirements of verbs are not accounted for —so affix *preptype* is never realized for verb forms of main verbs contained in these lexicons.

In addition to this, a great deal of the ambiguities generated by main verbs that theoretically have all the possibilities for verb scheme contextualization cannot be solved on the basis of phrase analysis. In order to avoid the generation of multiple analyses caused by these ambiguities, VP formalization has been adapted so as to give always one and the same analysis for each combination of verb forms and clitic pronouns. As a consequence of this, VPs are indeed the best example of removal of ambiguities, the following reductions have been formalized for them:

se dice, se dicen The ambiguities related to clitic pronoun *se* and values **MIDD**, **ACT_IMP** and **ACT_IMP+PASS** of affix *voice* are solved in favour of value **MIDD**. In phrase analysis, in fact, value **ACT_IMP** is possible only if periphrasis *hay que* is involved, while value **ACT_IMP+PASS** is not possible whatsoever.

lo The ambiguity generated by values **PRTVE** and **ACC** of affix *case* of clitic pronoun *lo* is solved in favour of value **ACC**.

sigá, sigamos, sigan These verb forms might produce for all verbs two different VPs with either value **IMPERATIVE** or value **SUBJUNCTIVE** for affix

mood (VPs in which these verb forms are combined with clitic pronouns are not ambiguous, *sígase, se siga*).

To avoid multiple analyses of this kind, in phrase analysis all the rules that simultaneously describe **IMPERATIVE** and **SUBJUNCTIVE** VPs must be restricted to value **SECOND** for affix *person* in **IMPERATIVE** mood⁸. For instance, the first alternative of the rule for active voice **Verb Phrase/5**, see Section 2.2.3.5.1, is turned into the following one:

```
Verb Phrase ( mvtype, act, person, number, mood ):
## (no) salgo, (no) salga, sal
## (no) llueve, (no) llueva
  { act :: ACT },
    ( { person :: SECOND }, { mood :: IMPERATIVE };
      { mood :: INDICATIVE|SUBJUNCTIVE } ),
      Verb ( mvtype, act, person, number, tense, mood ).
```

no sigas, no siga, no sigamos, no sigáis, no sigan These verb forms might produce for all verbs two different VPs with either value **IMPERATIVE** or value **SUBJUNCTIVE** for affix *mood*.

In phrase analysis, all alternatives that describe negative **IMPERATIVE** VPs are removed, these VPs will be always identified as having value **SUBJUNCTIVE** for affix *mood*. For instance, the second alternative of the rule for active voice **Verb Phrase/5**, see Section 2.2.3.5.1, is removed.

decirme, decirte, decirnos, decirnos Non-finite VPs with values **INFINITIVE** or **GERUND** for affix *mood*, if there is one clitic pronoun with values **FIRST** or **SECOND** for affix *person*, might have middle voice (*Repetirme a mí mismo constantemente que no importaba no me ayudaba nada*) or active voice (*Repitiéndome constantemente que aquello no importaba consiguió que acabara por crearlo*).

In phrase analysis, middle voice target VPs, that is, VPs immediately under the dominance of a non-terminal **PHRASE** node, are only possible if they

⁸Value **ACT** needs to be assigned for affix *voice* in order to avoid the generation of VPs with the affix non-terminal *act* as the output value for affix *voice*. We have to do so because verb forms, see above in this Section, are not assigned concrete values for affix *voice* in lexicons for phrase analysis. This also means, obviously, that *salga* and *llueva* have exactly the same description in phrase analysis.

have value **THIRD** for affix *person* of the clitic pronoun (*decirse* is not ambiguous). This explains the guards that set the values of affixes *voice* and *person_one* in the fourth alternative of the rewrite rule of non-terminal **NON FINITE PREDICATE**, see above in Section 4.2⁹.

decírmelo, decírtelo, decírnoslo, decíroslo, decírsele Non-finite verb forms with values **INFINITIVE** or **GERUND** for affix *mood*, if they are combined with two clitic pronouns, the second one having value **ACC** for case, might produce middle voice VPs (*Repetírmelo a mí mismo constantemente no me ayudaba nada*) or active voice VPs (*Repetírsele constantemente no hizo que ella llegara a entenderlo*).

In phrase analysis, target middle voice VPs are only produced if they have also value **DAT** for affix *case* of the functional clitic pronoun (*entregársele* is not ambiguous). This explains the guards that set the values of affixes *voice* and *case_two* in the sixth alternative of the rewrite rule of non-terminal **NON FINITE PREDICATE**, see above in Section 4.2¹⁰.

se lo dice Verb forms with value **THIRD** for affix *person*, if they are combined with two clitic pronouns, the first one with value **THIRD** for affix *person* and the second one with value **ACC** for affix *case*, might produce middle voice VPs (*Se lo repetía a sí mismo constantemente*) or active voice VPs (*Él se lo repetía constantemente, pero ella nunca le hacía caso*).

In phrase analysis, middle voice VPs are only produced if they have value **DAT** for affix *case* of the functional clitic pronoun (*se le entrega* is not ambiguous). For instance, the second alternative of the rule for middle voice **Verb Phrase/9**, see Section 2.2.3.5.2, is turned into the following one:

```
Verb Phrase ( one_clitic, voice, person, number, mood,
              case_one, person_one, gender_one, number_one ):
    (...);
## créetelo
{ voice :: MIDD },
```

⁹Middle voice periphrastic groups, integrated in **Periphrastic Non Finite Verb Phrase/7**, eliminate this ambiguity (*ponerme a hacerme* is no doubt middle voice).

¹⁰Middle voice periphrastic groups, integrated in **Periphrastic Non Finite Verb Phrase/11**, reduce this ambiguity to only value **THIRD** for affix *person* (*ponerme a hacerme* is no doubt middle voice, but *ponerse a hacersele* is still ambiguous).

```

( { person :: FIRST|SECOND };
  { person :: THIRD }, { case_one :: DAT } ),
Verb_clitized ( one_clitic, voice, person, number,
  simple, mood, ONE|TWO ),
Clitic Pronoun ( person, gender, number, case ),
Clitic Pronoun ( person_one, gender_one, number_one,
  case_one, ONE_OF_ONE ),
two persons values ( mood, person, person_one ),
verb one ClPr ( one_clitic, case_one, person_one );
(...).

```

me decía Verb forms with values **FIRST|THIRD** for *person*, if they are combined with a clitic pronoun with value **FIRST** for affix *person* and value **SING** for affix *number*, might produce middle voice VPs (*Me decía a mí mismo que aquello no podía durar*) or active voice VPs (*Él siempre me decía la verdad*).

In phrase analysis, middle voice VPs are only produced for verb forms that have either value **FIRST** or value **THIRD** for affix *person* (*hago/hace, hice/hizo, haré/hará, he/ha hecho, hube/hubo hecho, habré/habrá hecho*).

For instance, the first alternative of the rule for middle voice **Verb Phrase/5**, see Section 2.2.3.5.1, is turned into the following one:

```

Verb Phrase ( mvtype, voice, person, number, mood ):
## (no) me visto, (no) me vista
## (no) se trata, (no) se trate
  { voice :: MIDD },
  ( ( ( { person :: FIRST }, { number :: PLU };
    { person :: SECOND|THIRD } ),
    { tense :: IMPERFECT|CONDITIONAL|PLUPERFECT|
      PAST_CONDITIONAL };
    { tense :: PRESENT|PAST|FUTURE|PRESENT_PERFECT|
      PAST_PERFECT|FUTURE_PERFECT } ),
    { mood :: INDICATIVE };
    ( { person :: FIRST }, { number :: PLU };
      { person :: SECOND|THIRD } ),
      { mood :: SUBJUNCTIVE } ),
    Clitic Pronoun ( person, gender, number, case ),
    Verb ( mvtype, voice, person, number, tense, mood );
  (...).

```

espera These verb forms might produce for all verbs two different VPs with either value **IMPERATIVE** or value **INDICATIVE** for affix *mood*. The first

is associated with value **SECOND** for affix *person*, while the second is associated with value **THIRD**. This ambiguity cannot be removed because there are some verbs that have different verb forms for these inflectional possibilities (*sal/sale*)¹¹.

This ambiguity is associated with ambiguity of *voice* when a clitic pronoun with value **SECOND** for affix *person* is attached to some non-finite verb form introduced by a semiauxiliary that starts the periphrastic group in a periphrastic VP (*Sigue haciéndote el tonto y verás lo que te pasa, Sigue haciéndote la pascua cada vez que tiene una oportunidad*)¹².

4.2.2.1.2 *el que* The ambiguity generated by *that*-clauses preceded by determiner *el* and nominalized relative clauses can never be solved on the basis of phrase analysis, thus in the rewrite rule of non-terminal **PHRASE** one of the possible analyses for clauses introduced by *el que* (nominalized relative clause or *that*-clause) should be blocked in order to avoid multiple parsings. We decided to block the analysis of the structure as a *that*-clause preceded by determiner *el*.

4.2.2.1.3 (**Quantifier**) + (*más/menos*) + **Adjective** + **Noun**. **Quantifier** + *más/menos* Wherever the ambiguous items that in these sequences precede the combination **Adjective** + **Noun** can be interpreted as a sequence of adverbs, other analyses are blocked (*más buena persona, bastante buena persona, bastante más buena persona*). If this is not possible, the analysis identifies the ambiguous items as a combination of determiners (*muchas más buenas personas*). The analysis always blocked is that which identifies a combination of one determiner followed by an adverb. Groups of guards must be introduced in various rewrite rules of non-terminal **Nominal Phrase** to ensure that only the analyses preferred can be produced. For instance, the rewrite rule in Section 3.2.3.3.4.1 for **COMPTVE** determiners is turned into the following one¹³:

¹¹VPs in which these verb forms are combined with clitic pronouns are no longer ambiguous (*síguete, se sigue*).

¹²This ambiguity disappears with middle voice periphrastic groups.

¹³A more specified group of guards should be included to account also for the behaviour of the **EQUAL** comparative *tanto*. The fact that we wanted to make this reduction as simple as possible, set against the observation that frequency of these examples was very low, lead us to restrict the extension of the rule to what is necessary in order to eliminate the ambiguity in question only in the case of **NoEQUAL** comparatives *más/menos*.

```

Nominal Phrase ( ntype, NOTHING, COMPTVE, comp, definiteness,
                person, gender, number ):
{ premodifier :: NOTHING },
  DETERMINER ( COMPTVE, comp, PRE, definiteness, gender,
              number ),
  Nominal Phrase NOMINAL I ( ntype, premodifier, postmodifier,
                             NIL|PARTICULAR|PARTICULAR+, degree, person, gender,
                             number ),
  Top Determiner Level ( number, COMPTVE, comp, ONE ),
  ntype NOMINAL ( ntype ),
  [ Second Term MODIFIER ( second_term_type, NIL|PARTICULAR+,
                          comp, QUE|COMO );
    Second Term MODIFIER ( COMPTVE, PARTICULAR+, comp, DE,
                          person, gender, number ) ].

```

The ambiguity generated by pronoun-determiner sequences constituted by the combination of quantifiers and comparatives within PrPs (*muchas más*) is removed by eliminating the corresponding alternatives of the predicates that control the combination of pronouns and determiners within PrPs (see Section 3.2.3.3.5.1). We have eliminated the alternative that allows for the combination of **QUANT** pronouns with **COMPTVE** postdeterminers, so that the analysis of, for instance, sequence *muchas más* will identify a comparative pronoun *más* preceded by a quantifier determiner *muchas*.

4.2.2.1.4 Relative Clauses and Prepositional Relative Clauses Nominalized relative clauses introduced by prepositional PrPs (*A los que hayamos ofendido con nuestra actitud deberían decírnoslo cuanto antes*) and prepositional nominalizations of relative clauses (*A los que se hayan ofendido por nuestra actitud les pedimos humildemente perdón*) cannot be distinguished on the basis of phrase analysis. Moreover, we have a potentially threefold ambiguous analysis if we take into consideration that the preposition is frequently shared by both the nominalized relative clause introduced by a prepositional PrP and the prepositional nominalization of such relative clause (*A los que hayamos ofendido con nuestra actitud les pedimos humildemente perdón*).

In phrase analysis the third possibility is removed by eliminating the first alternative in the rewrite rule of **Prepositional Nominalization/7** (see Section 3.2.3.7.6). This rule thus remains as follows:

```

Prepositional Nominalization ( preposition, MDcategory, definiteness,

```

```

    person, gender, number ):
Prepositional LINKER ( preposition ),
    LINKED Nominalization ( MDcategory, preptype, definiteness,
    person, gender, number );
Prepositional LINKER ( preptype, dttype, dtsubtype, position,
    definiteness, gender, number ),
    Nominalization NOMINAL I ( MDcategory, NIL, person, gender,
    number ).

```

The first possibility, on the other hand, is blocked by forcing the affix *preptype* to take the value **NIL** wherever non-terminal **Nominalization/6** is called upon by rewrite rules different from that of non-terminal **LINKED Nominalization/6**—here we need to allow for the analysis of two consecutive prepositions, one of them required by the prepositional nominalization and the other by the prepositional PrP that introduces the nominalized relative clause (*A aquellas horas, no se sabía aún nada de a los que, supuestamente, habían dejado marchar en primer lugar*). The rewrite rule of non-terminal **PHRASE** into **Nominalization/6** remains then as follows:

```

PHRASE:
    (...);
    Nominalization ( MDcategory, NIL, definiteness, person, gender,
    number );
    (...).

```

As a result of this, all sequences of preposition followed by *el que, lo que*, etc., are analysed—in phrase analysis—as prepositional nominalizations of relative clauses introduced by non-prepositional PrPs.

A parallel situation—similarly solved—arises in the case of, on the one hand, argumental relative clauses introduced by prepositional PrPs and, on the other, prepositional relative clauses (*A quienes hayamos ofendido con nuestra actitud deberían decirnoslo cuanto antes, A quienes se hayan ofendido por nuestra actitud les pedimos humildemente perdón, A quienes hayamos ofendido con nuestra actitud les pedimos humildemente perdón, A aquellas horas, no se sabía aún nada de a quienes, supuestamente, habían dejado marchar en primer lugar*).

4.2.2.1.5 Noun/Pronoun Phrase person Predicates **Noun Phrase person** and **Pronoun Phrase person** have already been described in previous Sections 3.2.3.2.4.2 and 3.2.3.2.5.2. In clause analysis, they were necessary to ensure

that, whenever values **FIRST** or **SECOND** were not contextually required by other constituents of the clause, priority would be given to the analyses that assign value **THIRD** for affix *person* of NPs and PrPs. However, in phrase analysis no context can be checked, so values **FIRST** or **SECOND** for affix *person* can simply be blocked¹⁴.

4.2.2.2 Ambiguities Added Penalties

4.2.2.2.1 Internal Structure of Modifiers In order to ensure that the connection of modifiers with the first possible nucleus, blocking at the same time other possible analyses, will be given priority, penalties are used in a way so that rewriting a phrase category that includes a certain modifier into a phrase category that includes the next modifier found in the hierarchy established by affix *postmodifier* is less penalized than rewriting it into a phrase category that includes some other modifier found below in such hierarchy (see Section 4.1.2.3). For instance, the rewrite rule of non-terminal **Noun Phrase NUCLEUS** when it includes a relative clause modifier, see Section 3.2.3.2.4.3, is turned into the following one:

```
Noun Phrase NUCLEUS ( ntype, NOTHING, RELCLAUSE, CH_comp, degree,
    definiteness, person, gender, number ):
  HIDDEN RECURSIVE Noun Phrase ( ntype, NOTHING, RELCLAUSE,
    CH_comp, degree, definiteness, person, gender, number );
  $PENALTY,
  HIDDEN RECURSIVE Noun Phrase ( ntype, NOTHING, PRPPHRASE,
    CH_comp, degree, definiteness, person, gender, number );
  $PENALTY ( 2 ),
  HIDDEN RECURSIVE Noun Phrase ( ntype, NOTHING, ADJPHRASE,
    CH_comp, degree, definiteness, person, gender, number );
  $PENALTY ( 3 ),
  HIDDEN RECURSIVE Noun Phrase ( ntype, NOTHING, NOTHING,
    CH_comp, degree, definiteness, person, gender, number ).
```

On the contrary, rewriting a **Nominalization** that includes a certain modifier into a **Nominalization** that includes the next modifier found in the hierarchy established by affix *postmodifier* is more penalized than rewriting it into a **Nominalization** that includes some other modifier found below in such hierarchy. We ensure in this manner that modifiers will be preferably connected with the nucleus

¹⁴For relative pronouns *que*, *quien*, these values are maintained, so that sequences as *Nosotros que...* can be identified as one constituent.

of the nominalized unit and not with the nucleus of the nominalization (see also Section 4.1.2.3).

In order to ensure that modifiers will not be separated from their nucleus for the integration of subsequent constituents of the input sequence, the use of the strategy described above requires also that the number of penalties added each time that a new phrase is identified during the analysis needs to be higher than the highest number of penalties added to optimize the identification of the internal structure of modifiers. This is indeed the explanation for the addition of ten penalties each time that non-terminal **PHRASES** is recursively called upon, see above Section 4.2.1.

4.2.2.2.2 Plural Proper Names In Section 3.2.3.1, the rewrite rule of **Noun** showed that, in the case of proper names, value **SING** for *number* was preferred to value **PLU**. Within phrase analysis, the second possibility is added a penalization, to ensure, on the one hand, that both analyses will not be completely explored, and, on the other, that *las Ardenas*, for instance, can be analysed.

4.2.2.3 Lexical Ambiguity

For the purposes of certain applications, assuming that a number of analyses will certainly fail, the identification of different lexical categories might be added different numbers of penalties, in order to ensure that only one analysis will be produced for all sequences of words, including those that contain ambiguous words. The following ones can be accepted as general principles if these risks are considered acceptable:

- Closed classes of words must be less penalized than open classes of words.
- Pronouns should be more penalized than adverbs.

Apart from this, the domain of the application must be taken into account to determine the number of penalties added to different lexical categories. A previous study of the domain might be very useful to ensure that this —quite rough— strategy will yield the best possible results.

Chapter 5

Summary

In the present study, within a descriptive framework that considers clauses, phrases and words as main syntactic units, phrases have been formally described, together with other linguistic structures that might also have a role as clause-level syntactic functions (SUBJECT, DIRECT OBJECT...). To achieve this task we have used the AGFL formalism, which has proved to be a suitable formalism for the formal description of such linguistic units.

VPs have been described in first place. With respect to them, we accounted for all the linguistic phenomena that combine graphical words for the constitution of VPs. Although in our view these linguistic phenomena are taken to behave as morphological phenomena, whose process of analysis must be hidden in parse trees, they have to be very carefully described because they have an enormous influence on the VP features actually used to organize the elements of the clause, i.e. these phenomena are responsible for the contextualization of verb schemes to produce configurations. Verb schemes and configurations are the central issues of the VP description, both of them are considered to be the conjunction of the four verb features verb type, voice, impersonality and prepositional requirements, being voice and impersonality the only ones that undergo variations during the process of contextualization. Nine possible flows from verb schemes to configurations, i.e. nine different contextualization flows, have been identified, while three AGFL affixes have been used to encode the information about verb features, namely *vtype*, *voice* and *preptype*, where *voice* simultaneously accounts for both voice and impersonality. BDS has been used as the source of the verb schemes associated with

the verb forms contained in the lexicon. 171 AGFL syntax rules have been after necessary to account for the contextualization of these verb schemes to produce actual configurations.

NPs, PrPs, AvPs and AjPs have been described in second place. These are taken to be syntagms that contextualize lexical categories —noun, pronoun, adverb and adjective— that are expanded by means of general mechanisms such as subordination and coordination. For both these mechanisms, general AGFL frames have been designed which might account for their formalization, although only subordination has been specifically described in this work. Two main operations have been identified that expand lexical categories by means of subordination: addition of modifiers and addition of determiners. The application of these two operations to nouns, pronouns, adverbs and adjectives has been described and formalized by means of the further design, taking the AGFL General Frame for the Description of Subordination as a starting point, of the AGFL General Frame for Addition of Modifiers and the AGFL General Frame for Addition of Determiners, which proved both to be suitable and adequate rule-schemata for the description of both operations for all categories. This fact finally illustrates the general claim that subordination and addition of modifiers or determiners are general mechanisms or operations whose behaviour, except for the productivity of the operation when associated to a particular lexical or structural category, is similar for all of them.

Other syntactic structures, finally, that might also be found to occupy clause-level syntactic functions, nominalizations or subordinated clauses, have been also accounted for. 364 AGFL rules have been necessary to account for them and the contextualization of nouns, pronouns, adverbs and adjectives for the constitution of phrases.

From a different point of view, the process leading to formalization has been found to face problems of a different nature at different stages of development. Formalization, in a strict sense, primarily deals with the definition of the general frames that must be after used to formally describe more particular cases. The exact determination of these cases —those that will be actually covered by the formal grammar— is, however, considered to be a secondary or subsequent problem, related to decisions that depend on the use that is to be made of the formal grammar, which, for this reason, must be developed so as to allow for a maximum

discrimination of the features involved, as well as for as easy as possible an addition/removal of alternatives. In the broadest sense of the task, formalization has to deal also with the problem of ambiguity, since formalization is here not merely carried out with the scope of validating formalisms and linguistic models, but to be used in real applications. The way in which this fact has influenced our formalization has been described in previous Chapter: order of alternatives and penalties, as well as careful control of the lexicon, have been identified as useful—but limited—resources to deal with ambiguity in two different contexts: a full-sentence grammar and a phrase-level one.

Future developments of this work are obvious. Our immediate concern will be the integration of argumental behaviour of nouns and adjectives and the extensive formalization of coordination. This should be sufficient to achieve the exhaustive segmentation of clauses into their direct constituents. These in turn might be the basis for the identification of syntactic functions within clauses, as well as for the testing of verb lexicons with respect to verb schemes. We plan also to enrich and refine the information about verb schemes extracted from BDS. Information about semantic and syntactic subcategorization might be integrated together with requirements as to what syntactic functions can be combined with what verbs, in order to reduce ambiguity for the correct identification of such syntactic functions. Integration of the phrase grammar with a clause-level module is already available, but further extensions of the grammar so as to include aspects of a semantic and pragmatic nature, as well as text structure and cohesion, are yet within our purposes for the future. For the time being we simply expect, just as this work was aimed at, to have demonstrated that the model of description proposed, as well as the AGFL formalism, have proven to be together suitable tools in order to account for the formalization of phrase-level analysis in Spanish.

Appendix A

Non-passive Semiauxiliaries

<i>acabar</i>	active	personal	gerund
<i>andar</i>	active	personal	gerund
<i>comenzar</i>	active	personal	gerund
<i>continuar</i>	active	personal	gerund
<i>empezar</i>	active	personal	gerund
<i>estar</i>	active	personal	gerund
<i>ir</i>	active	personal	gerund
<i>llevar</i>	active	personal	gerund
<i>proseguir</i>	active	personal	gerund
<i>quedar</i>	active	personal	gerund
<i>salir</i>	active	personal	gerund
<i>seguir</i>	active	personal	gerund
<i>terminar</i>	active	personal	gerund
<i>venir</i>	active	personal	gerund
<i>acabar de</i>	active	personal	infinitive
<i>acabar por</i>	active	personal	infinitive
<i>acostumbrar a</i>	active	personal	infinitive
<i>alcanzar a</i>	active	personal	infinitive
<i>cesar de</i>	active	personal	infinitive
<i>comenzar a</i>	active	personal	infinitive
<i>comenzar por</i>	active	personal	infinitive
<i>concluir por</i>	active	personal	infinitive
<i>dar en</i>	active	personal	infinitive
<i>deber</i>	active	personal	infinitive
<i>deber de</i>	active	personal	infinitive
<i>dejar de</i>	active	personal	infinitive
<i>echarse a</i>	middle	personal	infinitive
<i>echar a</i>	active	personal	infinitive

<i>echar de</i>	active	personal	infinitive
<i>empezar a</i>	active	personal	infinitive
<i>empezar por</i>	active	personal	infinitive
<i>estar a punto de en un tris de a pique de a un pelo de</i>	active	personal	infinitive
<i>estar por</i>	active	personal	infinitive
<i>haber de</i>	active	personal	infinitive
<i>haber que</i>	active	impersonal	infinitive
<i>ir a</i>	active	personal	infinitive
<i>llegar a</i>	active	personal	infinitive
<i>osar</i>	active	personal	infinitive
<i>parar de</i>	active	personal	infinitive
<i>pasar a</i>	active	personal	infinitive
<i>poder</i>	active	personal	infinitive
<i>ponerse a</i>	active	personal	infinitive
<i>romper a</i>	active	personal	infinitive
<i>soler</i>	active	personal	infinitive
<i>tener que</i>	active	personal	infinitive
<i>terminar de</i>	active	personal	infinitive
<i>terminar por</i>	active	personal	infinitive
<i>tornar a</i>	active	personal	infinitive
<i>venir a</i>	active	personal	infinitive
<i>volver a</i>	active	personal	infinitive

Table A.1: Non-passive semiauxiliaries.

Appendix B

One Clause Example from BDS

BDS was described in Section 2.2.2.6, we indicated there that it contains the analyses of approximately 150.000 clauses found in the ARTHUS corpus of Spanish. For each of these analyses 61 different fields of information are recorded. In the following screen we show the record containing the analysis of the example found in line 25 of page 42 of *El Sur*¹:

ABANDONAR [SUR: 42, 25]		Act.		
Car. grales.		Sujeto		C. dir.
Tipo	Claus. con que	Carácter	1	1
Función	Térm. de fprep.	Tipo		
Polaridad	Afirmativa	Clítico1		11
Modalidad	Declarativa	Clítico2		
Perífrasis	poder + Inf.	Marca		
Forma vbl.	llegaras	Unidad	22	
Form. dom.		Animac.	11	11
Persona	2 sing.	Determ.	1	
Pdo. comp.		Número	1	
Argumentos	2	Refnt.		
Orden	SV	Observ.		

42,22 tus cartas, tu proposición de volver con ella,
42,23 abandonándonos a nosotras. >Me equivoco? En mis ca-vilaciones
42,24 de niña sobre lo que yo consideraba tu secreto
42,25 nunca apareció la posibilidad de que tú pudieras aban-donarme.
42,26 Yo sabía tan poco de ti... Mi mirada era tan

¹A. García Morales. *El Sur (seguido de Bene)*. Anagrama, Barcelona, 1985.

42,27 corta.

42,28 Decidi visitar a aquella mujer. Ahora sabía que vivía

In the top of the screen, we can see the location of the example, together with what we consider the main data for the identification of the configuration, i.e. voice, in this case, active. In the bottom of the screen, the example is showed with various lines of context. In the central part of the screen, finally, the first two columns on the left indicate that the example in question, *de que tú pudieras abandonarme*, has the following characteristics:

- i) It is a *that*-clause.
- ii) It functions as a prepositional modifier.
- iii) It has affirmative polarity.
- iv) It has declarative mood.
- v) It includes periphrasis *poder + infinitive*.
- vi) The verb is in imperfect tense and subjunctive mood.
- vii) Tense and mood of the main clause are not relevant.
- viii) The verb is in second person singular.
- ix) The main verb is not a multiword one.
- x) It includes 2 syntactic arguments.
- xi) The arguments are found in the order S(SUBJECT)V(ERB).

The type of clause, *that*-clause, is encoded in BDS as a numerical key, 21. Other possibilities for this field of information are the following:

Independent clause	11
Coordinated clause	12
<i>That</i> -clause	21
Infinitive clause	22
Other direct object clauses	23
Relative clause	31
Nominalization of relative clause	32
Adverbial relative clause	4
Gerund clause	5

Participle clause	6
Clause in a sentence	7
Non-direct-object clause introduced by certain conjunctions	81
Indicative or subjunctive clause introduced by conjunctive phrases	82
Infinitive clause introduced by conjunctive phrases	83
Other types of clauses	9

As can be observed, the hierarchical organization of numerical keys is linguistically motivated, so that it is possible to search for various levels of linguistically relevant information very easily.

The three last columns on the right part of the middle of the screen particularly describe the two syntactic arguments identified in this example, a SUBJECT and a DIRECT OBJECT². The description of these syntactic arguments, indicated by means of numerical keys in the screen above, is verbalized in the following one:

ABANDONAR [SUR: 42, 25]	Act.
Sujeto	C. dir.
Carácter	Explícito
Tipo	
Clítico1	me
Clítico2	
Marca	
Unidad	Pron. 2 pers.
Animac.	Animado discont. Animado discont.
Determ.	Definido
Número	Singular
Refnt.	
Observ.	

42,22 tus cartas, tu proposición de volver con ella,
 42,23 abandonándonos a nosotras. >Me equivoco? En mis ca-vilaciones
 42,24 de niña sobre lo que yo consideraba tu secreto
 42,25 nunca apareció la posibilidad de que tú pudieras aban-donarme.
 42,26 Yo sabía tan poco de ti... Mi mirada era tan
 42,27 corta.
 42,28 Decidí visitar a aquella mujer. Ahora sabía que vivía

²An INDIRECT OBJECT, one or two PREPOSITIONAL COMPLEMENTS, an AGENT and a PREDICATIVE COMPLEMENT might also be specified here if they are actually found in the example in question.

The SUBJECT is described here as:

- i) An explicit argument.
- ii) Type of the argument, clitic pronoun, secondary clitic pronoun and preposition are not relevant fields for the SUBJECT (these are fields that concern PREPOSITIONAL COMPLEMENTS and DIRECT or INDIRECT OBJECTS).
- iii) It is a personal pronoun in second person.
- iv) The referent is animate and countable.
- v) The pronoun is a definite one.
- vi) It is in singular.
- vii) The *Referent* field is not relevant for the SUBJECT (this is relevant for PREDICATIVE COMPLEMENT only).
- viii) No additional specifications have been recorded.

The DIRECT OBJECT is in turn described as:

- i) Character and type of the argument are not relevant fields for DIRECT OBJECT (these concern SUBJECT and PREPOSITIONAL COMPLEMENT respectively).
- ii) It is represented by a clitic pronoun *me*.
- iii) Secondary clitic pronoun is not relevant for DIRECT OBJECT (this is relevant for INDIRECT OBJECT only).
- iv) Preposition and category of the syntactic unit are not encoded for direct object clitic pronouns.
- v) The referent is animate and countable.
- vi) Determination and number are not encoded for direct object clitic pronouns.
- vii) As before, the *Referent* field is not relevant for the DIRECT OBJECT (this is relevant for PREDICATIVE COMPLEMENT only).
- viii) No additional specifications have been recorded.

Appendix C

Verb Scheme Cluster of *abandonar*

Table C.1 shows the first and more important part of the process of extraction of *verb scheme clusters* from BDS, see Section 2.2.2.6, for verb lemma *abandonar*. Location of the 197 examples of verb *abandonar* found in the ARTHUS¹ corpus

¹Source ARTHUS texts for the examples are the following:

1VOZ *La Voz de Galicia*, 10/30/91.

2VOZ *La Voz de Galicia*, 11/22/91.

3VOZ *La Voz de Galicia*, 11/23/91.

BAIRES A.M. Barrenechea, editor. *El habla culta de la ciudad de Buenos Aires. Materiales para su estudio*, (vol. 2). Inst. de Filología y Literatura Hispánicas “Dr. Amado Alonso”, 1987.

CAIMAN A. Buero Vallejo. *Caimán*. Espasa-Calpe, Madrid, 1981.

CARTA A. Colinas. *Larga carta a Francesca*. Seix Barral, Barcelona, 1986.

CINTA M.M. Reina. *La cinta dorada*. Antonio Machado, Madrid, 1989.

COARTADA F. Fernán Gómez. *La coartada*. Antonio Machado, Madrid, 1987.

CRONICA G. García Márquez. *Crónica de una muerte anunciada*. Mondadori, Madrid, 1987.

DIEGO E. Poniatowska. *Querido Diego, te abraza Quiela y otros cuentos*. Alianza / Era, Madrid, 1987.

GLENDIA J. Cortázar. *Queremos tanto a Glenda*. Alfaguara, 3rd edition, Madrid, 1981.

JOVENES J.R. Aldecoa. *Porque éramos jóvenes*. Seix Barral, Barcelona, 1986.

HISTORIAS A. Bioy Casares. *Historias desaforadas*. Alianza, Madrid, 1986.

HOMBRE J. Salom. *Un hombre en la puerta*. Preyson, Madrid, 1984.

LABERINTO E. Mendoza. *El laberinto de las aceitunas*. Seix Barral, Barcelona, 1982.

LING M. Bunge. *Lingüística y filosofía*. Ariel, Barcelona, 1983.

MADRID M. Esgueva and M. Cantarero, editors. *El habla de la ciudad de Madrid. Materiales para su estudio*. CSIC (Miguel de Cervantes), Madrid, 1981.

MIRADA J.M. Guelbenzu. *La mirada*. Alianza, Madrid, 1987.

PAISAJES J. Goytisolo. *Paisajes después de la batalla*. Montesinos, Barcelona, 1982.

RATON R. Sánchez Ferlosio. *La homilía del ratón*. El País, Madrid, 1986.

is showed in the first two columns, while configurations documented by the BDS analyses of these examples are showed in the next eight columns², the scheme documented by the example, obtained after the application of the reductions described in Section 2.2.2.6, is recorded in the last column. So, for example, the verb form *abandonar* in SUR/42,25, see Appendix B, is considered an instance of scheme *SDO, ACT*, according to rules 1 and 3 of those described in Section 2.2.2.6.

LOCATION		BDS RELEVANT FIELDS								SCHEME
TEXT	PAGE_LINE	6	9	16	21	28	35	37	55	
2VOZ	49,1,6,11	1		2						S, ACT
1VOZ	48,1,1,22	1		1						S, ACT
SUR	27,8	1		2	1					SDO, ACT
SUR	42,23	1		2	1					SDO, ACT
SUR	42,25	1	27	1	1					SDO, ACT
LABERINTO	9,19	1		1	1					SDO, ACT
LABERINTO	57,22	1		2	1					SDO, ACT
LABERINTO	57,16	1		1	1					SDO, ACT
LABERINTO	45,32	1	27	2	1					SDO, ACT
JOVENES	4,15	1		1	1					SDO, ACT
JOVENES	46,28	1		2	1					SDO, ACT
SONRISA	3,11	1		1	1					SDO, ACT

SEVILLA M.A. de Pineda, editor. *Sociolingüística andaluza 2. Material de encuestas para el estudio del habla urbana culta de Sevilla*. Servicio de Publicaciones, Universidad de Sevilla, 1983.

SONRISA J.L. Sampedro. *La sonrisa etrusca*. Alfaguara, Madrid, 1985.

TERNURA I. Martínez de Pisón. *La ternura del dragón*. Anagrama, 3rd edition, Barcelona, 1988.

TIEMPO O. Paz. *Tiempo nublado*. Seix Barral, Barcelona, 1983.

USOS C. Martín Gaite. *Usos amorosos de la postguerra española*. Anagrama, 8th edition, Barcelona, 1988.

ZORRA F. Nieva. *Te quiero, zorra*. Antonio Machado, Madrid, 1988.

²The keys for the examples, already introduced in Section 2.2.2.6, are the following:

Field 6 Voice: 1, active, 2, middle, 3, passive.

Field 9 Periphrasis: 1, *acabar* + gerund, 4, *ir* + gerund, 10, *acabar de* + infinitive, 14, *comenzar a* + infinitive, 15, *deber* + infinitive, 23, *ir a* + infinitive, 24, *llegar a* + infinitive, 27, *poder* + infinitive, 30, *tener que* + infinitive, 33, *volver a* + infinitive.

Field 16 Character of the SUBJECT: 1, implicit, 2, explicit, 30, generic subject of infinitive or gerund, 34, *uno* (impersonal meaning).

Fields 21, 28, 35, 55 Character of the DIRECT OBJECT, INDIRECT OBJECT, PREPOSITIONAL COMPLEMENT, PREDICATIVE COMPLEMENT: 1, there is one.

Field 37 Preposition introducing a prepositional complement.

SONRISA	3,19	1		1	1					SDO, ACT
SONRISA	5,25	1		2	1					SDO, ACT
SONRISA	3,18	1		2	1					SDO, ACT
SONRISA	3,22	1		2	1					SDO, ACT
PAISAJES	40,3	1		2	1					SDO, ACT
PAISAJES	79,9	1		2	1					SDO, ACT
GLENDIA	40,25	1		1	1					SDO, ACT
MIRADA	9,30	1		2	1					SDO, ACT
CAIMAN	58,17	1		2	1					SDO, ACT
CARTA	2,9	1		1	1					SDO, ACT
CARTA	70,23	1		2	1					SDO, ACT
HISTORIAS	29,29	1		1	1					SDO, ACT
HISTORIAS	29,31	1		2	1					SDO, ACT
HISTORIAS	9,24	1		1	1					SDO, ACT
HISTORIAS	57,7	1		2	1					SDO, ACT
TIEMPO	95,2	1		2	1					SDO, ACT
CINTA	40,27	1		2	1					SDO, ACT
COARTADA	0,9	1	27	2	1					SDO, ACT
COARTADA	9,18	1		2	1					SDO, ACT
COARTADA	6,17	1		2	1					SDO, ACT
2VOZ	37,3,2,3	1		1	1					SDO, ACT
3VOZ	18,6,0,1	1		2	1					SDO, ACT
SUR	75,23	1		2	1					SDO, ACT
SUR	89,3	1		2	1					SDO, ACT
SUR	89,31	1		2	1					SDO, ACT
SUR	5,9	1		2	1					SDO, ACT
LABERINTO	55,10	1		1	1					SDO, ACT
LABERINTO	35,6	1		1	1					SDO, ACT
LABERINTO	69,22	1		2	1					SDO, ACT
LABERINTO	37,28	1		2	1					SDO, ACT
LABERINTO	37,18	1		2	1					SDO, ACT
LABERINTO	44,1	1		2	1					SDO, ACT
LABERINTO	56,31	1		2	1					SDO, ACT
LABERINTO	47,2	1		2	1					SDO, ACT
LABERINTO	62,8	1		2	1					SDO, ACT
LABERINTO	57,13	1	10	2	1					SDO, ACT
LABERINTO	55,2	1	30	2	1					SDO, ACT
LABERINTO	90,29	1	10	2	1					SDO, ACT
LABERINTO	49,30	1	30	2	1					SDO, ACT
CRONICA	02,16	1		1	1					SDO, ACT
CRONICA	14,7	1		2	1					SDO, ACT
CRONICA	3,10	1		2	1					SDO, ACT
CRONICA	5,12	1		2	1					SDO, ACT
CRONICA	66,1	1		2	1					SDO, ACT

DIEGO	64,5	1		1	1					SDO, ACT
DIEGO	3,4	1		1	1					SDO, ACT
DIEGO	67,11	1		2	1					SDO, ACT
DIEGO	1,7	1		2	1					SDO, ACT
DIEGO	8,25	1		2	1					SDO, ACT
DIEGO	0,14	1		2	1					SDO, ACT
JOVENES	54,39	1		1	1					SDO, ACT
JOVENES	1,30	1		1	1					SDO, ACT
JOVENES	44,30	1		1	1					SDO, ACT
JOVENES	65,35	1		1	1					SDO, ACT
JOVENES	76,10	1		2	1					SDO, ACT
JOVENES	42,1	1		2	1					SDO, ACT
JOVENES	21,29	1		2	1					SDO, ACT
JOVENES	49,34	1		2	1					SDO, ACT
JOVENES	49,4	1		2	1					SDO, ACT
JOVENES	50,23	1		1	1					SDO, ACT
SONRISA	99,3	1		2	1					SDO, ACT
SONRISA	5,22	1		2	1					SDO, ACT
SONRISA	95,10	1		2	1					SDO, ACT
TERNURA	75,28	1		1	1					SDO, ACT
TERNURA	7,33	1	15	2	1					SDO, ACT
TERNURA	4,24	1		2	1					SDO, ACT
PAISAJES	6,9	1		2	1					SDO, ACT
PAISAJES	88,14	1		1	1					SDO, ACT
PAISAJES	59,25	1		1	1					SDO, ACT
PAISAJES	17,3	1		2	1					SDO, ACT
PAISAJES	46,30	1		2	1					SDO, ACT
PAISAJES	86,5	1		2	1					SDO, ACT
PAISAJES	20,17	1		2	1					SDO, ACT
PAISAJES	35,9	1		2	1					SDO, ACT
PAISAJES	10,11	1	27	2	1					SDO, ACT
PAISAJES	44,22	1		2	1					SDO, ACT
PAISAJES	61,30	1		2	1					SDO, ACT
PAISAJES	73,5	1		2	1					SDO, ACT
GLENDA	59,26	1	24	1	1					SDO, ACT
GLENDA	33,18	1	33	2	1					SDO, ACT
GLENDA	50,20	1		2	1					SDO, ACT
MIRADA	4,2	1		2	1					SDO, ACT
MIRADA	0,13	1		2	1					SDO, ACT
MIRADA	4,15	1		2	1					SDO, ACT
MIRADA	88,13	1		2	1					SDO, ACT
MIRADA	26,21	1		2	1					SDO, ACT
MIRADA	90,2	1		2	1					SDO, ACT
MIRADA	45,24	1		2	1					SDO, ACT

1VOZ	42,1,7,11	1		2	1				SDO, ACT
3VOZ	58,2,3,27	1		1	1				SDO, ACT
3VOZ	38,2,1,2	1		1	1				SDO, ACT
3VOZ	60,1,0,2	1		2	1				SDO, ACT
3VOZ	18,4,1,17	1		2	1				SDO, ACT
MADRID	32,36	1		34	1				SDO, ACT
MADRID	33,1	1		1	1				SDO, ACT
MADRID	32,6	1		2	1				SDO, ACT
MADRID	9,28	1		2	1				SDO, ACT
SEVILLA	42,22	1		30	1				SDO, ACT
BAIRES	57,26	1		2	1				SDO, ACT
BAIRES	57,28	1		2	1				SDO, ACT
SUR	6,18	1		1	1				SDO, ACT
LABERINTO	8,10	1		1	1				SDO, ACT
SONRISA	1,29	1		1	1				SDO, ACT
GLENDIA	77,3	1		1	1				SDO, ACT
GLENDIA	63,12	1		1	1				SDO, ACT
GLENDIA	25,4	1		1	1				SDO, ACT
GLENDIA	40,25	1		2	1				SDO, ACT
MIRADA	98,27	1		2	1				SDO, ACT
CAIMAN	7,29	1		1	1				SDO, ACT
USOS	8,13	1	15	1	1				SDO, ACT
OCHENTA	64,29	1		1	1				SDO, ACT
JOVENES	73,26	1		1	1				SDO, ACT
LING	97,33	1	15	1	1				SDO, ACT
TIEMPO	6,18	1		2	1				SDO, ACT
TIEMPO	56,7	1		1	1				SDO, ACT
RATON	9,23	1	4	1	1				SDO, ACT
1VOZ	40,1,1,46	1		2	1				SDO, ACT
SUR	82,31	1		2	1	1	a		SDOPC, ACT, A
PAISAJES	163,5	1		1	1	1	a		SDOPC, ACT, A
3VOZ	18,6,1,3	1		1	1	1	en		SDOPC, ACT, EN
JOVENES	156,5	1		2	1	1	en		SDOPC, ACT, EN
SONRISA	249,10	1		2	1	1	en		SDOPC, ACT, EN
SEVILLA	227,24	1	30	2	1	1	en		SDOPC, ACT, EN
CAIMAN	43,2	1		2	1	1	sobre		SDOPC, ACT, SOBRE
GLENDIA	99,14	2		1					S, MIDD
SEVILLA	93,17	2		1					S, MIDD
3VOZ	67,1,3,41	2		1					S, MIDD
RATON	45,15	2		2		1	a		SPC, MIDD, A
SUR	39,15	2		1		1	a		SPC, MIDD, A
PAISAJES	92,22	2		2		1	a		SPC, MIDD, A
MIRADA	71,26	2		2		1	a		SPC, MIDD, A

HISTORIAS	65,5	2		2			1	a		SPC, MIDD, A
HISTORIAS	93,30	2		2			1	a		SPC, MIDD, A
ZORRA	14,7	2		2			1	a		SPC, MIDD, A
RATON	103,32	2		30			1	a		SPC, MIDD, A
LING	112,18	3		1						SDO, ACT
1VOZ	26,1,2,8	3		2						SDO, ACT
3VOZ	41,3,2,11	3		1						SDO, ACT
MADRID	137,4	3		2						SDO, ACT
3VOZ	35,5,1,4	3		2			1	a		SDOPC, ACT, A
2VOZ	28,4,1,8	3		2			1	en		SDOPC, ACT, EN

Table C.1: Schemes documented in BDS for verb lemma *abandonar*

The process of merging schemes is represented in Table C.2: the schemes documented—with various frequencies—in BDS for verb lemma *abandonar* are recorded in the first column, while in the second one the first reduction carried out by merging prepositions is showed. The third column finally records the second reduction performed on schemes by merging verb type. The third possible reduction, which should merge voice values, cannot be illustrated with verb *abandonar*, since, after the second reduction, there are not schemes left of this verb that are only different with respect to voice values.

Schemes of <i>abandonar</i>	Merging	
	of prepositions	of verb type
S, ACT	S, ACT	S SDO, ACT
SDO, ACT	SDO, ACT	
SDOPC, ACT, A	SDOPC, ACT, A EN SOBRE	SDOPC, ACT, A EN SOBRE
SDOPC, ACT, EN		
SDOPC, ACT, SOBRE		
S, MIDD	S, MIDD	S, MIDD
SPC, MIDD, A	SPC, MIDD, A	SPC, MIDD, A

Table C.2: *Verb scheme clusters* for verb lemma *abandonar*

The set of merged schemes recorded in the last column of the table above for each verb lemma constitutes the set of *verb scheme clusters* of this verb lemma. In the following stage, these *verb scheme clusters* are combined with each verb form of the *verb form cluster*, and, in the case of verb lemma *abandonar*, for *first*

person, singular, present, indicative, we finally obtain the following AGFL lexicon definitions:

```
VerbSt ( S|SDO, ACT, FIRST, SING, PRESENT, INDICATIVE ):
```

```
  "abandono".
```

```
VerbSt ( SDOPC, ACT, A|EN|SOBRE, FIRST, SING, PRESENT,  
        INDICATIVE ):
```

```
  "abandono".
```

```
VerbSt ( S, MIDD, FIRST, SING, PRESENT, INDICATIVE ):
```

```
  "abandono".
```

```
VerbSt ( SPC, MIDD, A, FIRST, SING, PRESENT, INDICATIVE ):
```

```
  "abandono".
```

Appendix D

Pronouns¹

PronounSt (INTG, QUE, DEFINITE, NEUT, SING): "qué". # "qué"
PronounSt (INTG, QUIEN, DEFINITE, MASC|FEM, SING): "quién".
"quién"
PronounSt (INTG, CUAL, DEFINITE, MASC|FEM, SING): "cuál". # "cuál"
PronounSt (INTG, CUANTO, DEFINITE, MASC|NEUT, SING): "cuánto".
"cuánto"
PronounSt (REL, QUE, DEFINITE, MASC|FEM|NEUT, SING|PLU): "que".
"que"
PronounSt (REL, QUIEN, DEFINITE, MASC|FEM, SING): "quien". # "quien"
PronounSt (REL, CUAL, DEFINITE, MASC|FEM|NEUT, SING): "cual".
"cual"
PronounSt (REL, CUANTO, DEFINITE, MASC|NEUT, SING): "cuanto".
"cuanto"
PronounSt (COMPTVE, EQUAL, UNDEFINITE, MASC|NEUT, SING): "tanto".
"tanto"
PronounSt (COMPTVE, NoEQUAL, UNDEFINITE, MASC|FEM|NEUT, SING|PLU):
"más". # "más"
PronounSt (COMPTVE, NoEQUAL, UNDEFINITE, MASC|FEM|NEUT, SING|PLU):
"menos". # "menos"
PronounSt (PERS, NIL, DEFINITE, FIRST, MASC|FEM, SING, NOM): "yo".
"yo"
Prepositional PronounSt (CON, PERS, NIL, DEFINITE, FIRST, MASC|FEM,
SING, PREP): "conmigo". # "conmigo"
PronounSt (PERS, NIL, DEFINITE, FIRST, MASC|FEM, SING, PREP): "mí".
"mí"
PronounSt (PERS, NIL, DEFINITE, FIRST, MASC, PLU, NOM|PREP):
"nosotros". # "nosotros"
PronounSt (PERS, NIL, DEFINITE, SECOND, MASC|FEM, SING, NOM): "tú".
"tú"

¹To reduce the size of this Appendix, forms inflected for *gender* **MASC** —and/or **NEUT**— and *number* **SING** have been selected as representatives for the whole inflectional paradigm of pronouns inflected for *gender* and *number*. In addition to this, for *prtypes* **CARD**, **ORD** and **PART**, only one lemma is showed from each set of lemmas with similar behaviour.

PrepositionalPronounSt (CON, PERS, NIL, DEFINITE, SECOND, MASC|FEM, SING, PREP): "contigo". #"contigo"
 PronounSt (PERS, NIL, DEFINITE, SECOND, MASC|FEM, SING, PREP): "ti". #"ti"
 PronounSt (PERS, NIL, DEFINITE, SECOND, MASC, PLU, NOM|PREP): "vosotros". #"vosotros"
 PronounSt (PERS, NIL, DEFINITE, THIRD, MASC, SING, NOM|PREP): "él". #"él"
 PronounSt (PERS, NIL, DEFINITE, THIRD, MASC|FEM|NEUT, SING|PLU, PREP): "sí". #"sí"
 PrepositionalPronounSt (CON, PERS, NIL, DEFINITE, THIRD, MASC|FEM|NEUT, SING|PLU, PREP): "consigo". #"consigo"
 PronounSt (PERS, NIL, DEFINITE, THIRD, MASC|FEM, SING, NOM|PREP): "usted". #"usted"
 PronounSt (GLOBAL, NIL, DEFINITE, MASC|NEUT, SING): "todo". #"todo"
 PronounSt (DEM, ESTE, DEFINITE, MASC, SING): "éste". #"éste"
 PronounSt (DEM, ESTE, DEFINITE, MASC, SING): "este". #"este"
 PronounSt (DEM, ESTE, DEFINITE, MASC, SING): "ése". #"ése"
 PronounSt (DEM, ESTE, DEFINITE, MASC, SING): "ese". #"ese"
 PronounSt (DEM, ESTE, DEFINITE, MASC, SING): "aquél". #"aquél"
 PronounSt (DEM, ESTE, DEFINITE, MASC, SING): "aquel". #"aquel"
 PronounSt (POSS, NIL, DEFINITE, FIRST|FOURTH, MASC|NEUT, SING): "nuestro". #"nuestro"
 PronounSt (POSS, NIL, DEFINITE, SECOND|FIFTH, MASC|NEUT, SING): "vuestro". #"vuestro"
 PronounSt (POSS, NIL, DEFINITE, FIRST, MASC|NEUT, SING): "mío". #"mío"
 PronounSt (POSS, NIL, DEFINITE, SECOND, MASC|NEUT, SING): "tuyo". #"tuyo"
 PronounSt (POSS, NIL, DEFINITE, THIRD|SIXTH, MASC|NEUT, SING): "suyo". #"suyo"
 PronounSt (QUANT, AMBOS, UNDEFINITE, MASC, PLU): "ambos". #"ambos"
 PronounSt (QUANT, VARIOS, UNDEFINITE, MASC, PLU): "varios". #"varios"
 PronounSt (QUANT, POCO, UNDEFINITE, MASC|NEUT, SING): "poco". #"poco"
 PronounSt (QUANT, MUCHO, UNDEFINITE, MASC|NEUT, SING): "mucho". #"mucho"
 PronounSt (QUANT, MUCHO, UNDEFINITE, MASC|FEM|NEUT, SING): "bastante". #"bastante"
 PronounSt (QUANT, DEMASIADO, UNDEFINITE, MASC|NEUT, SING): "demasiado". #"demasiado"
 PronounSt (DIS, NIL, UNDEFINITE, MASC, SING): "cada cual". #"cada cual"
 PronounSt (INDEF, UNO, UNDEFINITE, MASC, PLU): "unos". #"unos"
 PronounSt (INDEF|CARD, UNO, UNDEFINITE, MASC|NEUT, SING): "uno". #"uno"
 PronounSt (INDEF, QUIENQUIERA, UNDEFINITE, MASC, SING): "quienquiera". #"quienquiera"
 PronounSt (INDEF, OTRO, UNDEFINITE, MASC|NEUT, SING): "otro".

"otro"
 PronounSt (INDEF, ALGUNO, UNDEFINITE, MASC, SING): "alguno".
 # "alguno"
 PronounSt (INDEF, ALGUNO, UNDEFINITE, MASC, PLU): "unos cuantos".
 # "unos cuantos"
 PronounSt (INDEF, ALGUNO, UNDEFINITE, MASC, SING): "ninguno".
 # "ninguno"
 PronounSt (INDEF, ALGUIEN, UNDEFINITE, MASC, SING): "alguien".
 # "alguien"
 PronounSt (INDEF, ALGUIEN, UNDEFINITE, MASC, SING): "nadie".
 # "nadie"
 PronounSt (INDEF, ALGUIEN, UNDEFINITE, NEUT, SING): "algo". # "algo"
 PronounSt (INDEF, ALGUIEN, UNDEFINITE, NEUT, SING): "nada". # "nada"
 PronounSt (INDEF, CUALQUIERA, UNDEFINITE, MASC, SING):
 "cualquiera". # "cualquiera"
 PronounSt (INDEF, DEMAS, UNDEFINITE, MASC|FEM|NEUT, SING|PLU):
 "demás". # "demás"
 PronounSt (REFL, NIL, UNDEFINITE, MASC|NEUT, SING): "mismo".
 # "mismo"
 PronounSt (CARD, NoUNIT, DEFINITE, MASC|FEM, PLU): "cero". # "cero"
 PronounSt (ORD, CUARTO, UNDEFINITE, MASC|NEUT, SING): "primero".
 # "primero"
 PronounSt (ORD|PART, CUARTO, UNDEFINITE, MASC|NEUT, SING):
 "cuarto". # "cuarto"
 PronounSt (PART, MEDIO, UNDEFINITE, FEM, SING): "media". # "medio"
 PronounSt (PART, MEDIO, UNDEFINITE, MASC, SING): "medio". # "medio"
 PronounSt (PART, CUARTO, UNDEFINITE, MASC, SING): "tercio".
 # "tercio"

Appendix E

Adverbs

AdvSt (REL, DONDE, NIL, NIL, +STRESS): "donde". # "donde"
AdvSt (PLACE, DONDEQUIERA, NIL, NIL, +STRESS): "dondequiera".
"dondequiera"
AdvSt (INTG|EXCL, DONDE, NIL, NIL, +STRESS): "dónde". # "dónde"
Prepositional AdvSt (A, PLACE, DONDEQUIERA, NIL, NIL, +STRESS):
"adondequiera". # "adondequiera"
Prepositional AdvSt (A, REL, DONDE, NIL, NIL, +STRESS): "adonde".
"adonde"
Prepositional AdvSt (A, INTG|EXCL, DONDE, NIL, NIL, +STRESS):
"adónde". # "adónde"
AdvSt (MANNER, COMOQUIERA, NIL, NIL, +STRESS): "comoquiera".
"comoquiera"
AdvSt (REL, COMO, NIL, NIL, +STRESS): "como". # "como"
AdvSt (INTG|EXCL, COMO, NIL, NIL, +STRESS): "cómo". # "cómo"
AdvSt (REL, QUE, NIL, NIL, +STRESS): "que". # "que"
AdvSt (REL|TIME, MIENTRAS, NIL, NIL, +STRESS): "mientras".
"mientras"
AdvSt (REL, CUANDO, NIL, NIL, +STRESS): "cuando". # "cuando"
AdvSt (INTG|EXCL, CUANDO, NIL, NIL, +STRESS): "cuándo". # "cuándo"
AdvSt (REL, CUANTO, NIL, NIL, +STRESS): "cuanto". # "cuanto"
AdvSt (INTG|EXCL, CUANTO, NIL, NIL, +STRESS): "cuánto". # "cuánto"
AdvSt (EXCL, NIL, NIL, NIL, -STRESS): "qué". # "qué"
AdvSt (PLACE, AQUI, NIL, NIL, +STRESS): "ahí". # "ahí"
AdvSt (PLACE, AQUI, NIL, NIL, +STRESS): "allí". # "allí"
AdvSt (PLACE, AQUI, NIL, NIL, +STRESS): "aquí". # "aquí"
AdvSt (PLACE, ACA, NIL, POS, +STRESS): "allá". # "allá"
AdvSt (PLACE, ACA, NIL, POS, +STRESS): "acá". # "acá"
AdvSt (PLACE, DETRAS, DE, POS, +STRESS): "cerca". # "cerca"
AdvSt (PLACE, DETRAS, DE, POS, +STRESS): "lejos". # "lejos"
AdvSt (PLACE, DETRAS, DE, NIL, +STRESS): "alrededor". # "alrededor"
AdvSt (PLACE, DETRAS, DE, NIL, +STRESS): "debajo". # "debajo"
AdvSt (PLACE, DETRAS, DE, NIL, +STRESS): "delante". # "delante"
AdvSt (PLACE, DETRAS, DE, NIL, +STRESS): "dentro". # "dentro"

AdvSt (PLACE, DETRAS, DE, NIL, +STRESS): "detrás". #"detrás"
AdvSt (PLACE, DETRAS, DE, NIL, +STRESS): "encima". #"encima"
AdvSt (PLACE, DETRAS, DE, NIL, +STRESS): "enfrente". #"enfrente"
AdvSt (PLACE, DETRAS, DE, NIL, +STRESS): "fuera". #"fuera"
AdvSt (PLACE, ATRAS, NIL, POS, +STRESS): "afuera". #"afuera"
AdvSt (PLACE, ATRAS, NIL, POS, +STRESS): "adentro". #"adentro"
AdvSt (PLACE, ATRAS, NIL, POS, +STRESS): "adelante". #"adelante"
AdvSt (PLACE, ATRAS, NIL, POS, +STRESS): "abajo". #"abajo"
AdvSt (PLACE, ATRAS, NIL, POS, +STRESS): "atrás". #"atrás"
AdvSt (PLACE, ATRAS, NIL, POS, +STRESS): "arriba". #"arriba"
AdvSt (TIME, AHORA, NIL, NIL, +STRESS): "ahora". #"ahora"
AdvSt (TIME, SIEMPRE, NIL, NIL, +STRESS): "siempre". #"siempre"
AdvSt (TIME, NIL, NIL, NoEQUAL, +STRESS): "después". #"después"
AdvSt (TIME, NIL, NIL, NoEQUAL, +STRESS): "antes". #"antes"
AdvSt (TIME, NIL, NIL, POS, +STRESS): "pronto". #"pronto"
AdvSt (TIME, NIL, NIL, POS, +STRESS): "temprano". #"temprano"
AdvSt (TIME, NIL, NIL, POS, +STRESS): "tarde". #"tarde"
AdvSt (TIME, NIL, NIL, NIL, +STRESS): "apenas". #"apenas"
AdvSt (TIME, NIL, NIL, NIL, +STRESS): "todavía". #"todavía"
AdvSt (TIME, NIL, NIL, NIL, +STRESS): "ya". #"ya"
AdvSt (TIME, NIL, NIL, NIL, +STRESS): "aún". #"aún"
AdvSt (TIME, NIL, NIL, NIL, +STRESS): "primero". #"primero"
AdvSt (TIME, NIL, NIL, NIL, +STRESS): "mañana". #"mañana"
AdvSt (TIME, NIL, NIL, NIL, +STRESS): "hoy". #"hoy"
AdvSt (TIME, NIL, NIL, NIL, +STRESS): "entretanto". #"entretanto"
AdvSt (TIME, NIL, NIL, NIL, +STRESS): "enseguida". #"enseguida"
AdvSt (TIME, NIL, NIL, NIL, +STRESS): "ayer". #"ayer"
AdvSt (TIME, NIL, NIL, NIL, +STRESS): "anteanoche". #"anteanoche"
AdvSt (TIME, NIL, NIL, NIL, +STRESS): "anteayer". #"anteayer"
AdvSt (TIME, NIL, NIL, NIL, +STRESS): "anoche". #"anoche"
AdvSt (TIME, NIL, NIL, NIL, +STRESS): "luego". #"luego"
AdvSt (TIME, NIL, NIL, NIL, +STRESS): "entonces". #"entonces"
AdvSt (TIME, JAMAS, NIL, NIL, +STRESS): "jamás". #"jamás"
AdvSt (TIME, NUNCA, NIL, NIL, +STRESS): "nunca". #"nunca"
AdvSt (MANNER, ASI, NIL, NIL, +STRESS): "así". #"así"
AdvSt (MANNER, NIL, NIL, POS, +STRESS): "bien". #"bien"
AdvSt (MANNER, NIL, NIL, POS, +STRESS): "mal". #"mal"
AdvSt (MANNER, NIL, NIL, POS, +STRESS): "derecho". #"derecho"
AdvSt (MANNER, NIL, NIL, POS, +STRESS): "deprisa". #"deprisa"
AdvSt (MANNER, NIL, NIL, POS, +STRESS): "aprisa". #"aprisa"
AdvSt (MANNER, NIL, NIL, POS, +STRESS): "rápido". #"rápido"
AdvSt (MANNER, NIL, NIL, POS, +STRESS): "despacio". #"despacio"
AdvSt (MANNER, NIL, NIL, POS, +STRESS): "aposta". #"aposta"
AdvSt (MANNER, NIL, NIL, POS, +STRESS): "adrede". #"adrede"
AdvSt (MANNER, NIL, NIL, NIL, +STRESS): "infraganti". #"infraganti"
AdvSt (MANNER, NIL, NIL, NIL, +STRESS): "exprofeso". #"exprofeso"
AdvSt (MANNER, NIL, NIL, NoEQUAL, +STRESS): "mejor". #"mejor"
AdvSt (MANNER, NIL, NIL, NoEQUAL, +STRESS): "peor". #"peor"
AdvSt (COMPTVE, NoEQUAL, NIL, NoEQUAL, AVposition): "más". #"más"
AdvSt (COMPTVE, NoEQUAL, NIL, NoEQUAL, AVposition): "menos".

"menos"
 AdvSt (COMPTVE, EQUAL, NIL, EQUAL, -STRESS): "tan". # "tan"
 AdvSt (COMPTVE, EQUAL, NIL, EQUAL, +STRESS): "tanto". # "tanto"
 AdvSt (QUANT, POCO, NIL, POS+, AVposition): "un poco". # "un poco"
 AdvSt (QUANT, BASTANTE, NIL, POS++, AVposition): "bastante".
 # "bastante"
 AdvSt (QUANT, DEMASIADO, NIL, POS+++, AVposition): "demasiado".
 # "demasiado"
 AdvSt (QUANT, HARTO, NIL, SUP, -STRESS): "harto". # "harto"
 AdvSt (QUANT, ALGO, NIL, POS+, -STRESS): "algo". # "algo"
 AdvSt (QUANT, MUCHO, NIL, SUP, -STRESS): "muy". # "muy"
 AdvSt (QUANT, CASI, NIL, NIL, +STRESS): "casi". # "casi"
 AdvSt (QUANT, NIL, NIL, NIL, +STRESS): "sobremanera".
 # "sobremanera"
 AdvSt (QUANT, MUCHO, NIL, SUP, +STRESS): "mucho". # "mucho"
 AdvSt (QUANT, NADA, NIL, ZERO, AVposition): "nada". # "nada"
 AdvSt (QUANT, POCO, NIL, POS-, AVposition): "poco". # "poco"
 AdvSt (POL, NIL, NIL, NIL, +STRESS): "sólo". # "sólo"
 AdvSt (POL, NIL, NIL, NIL, +STRESS): "justo". # "justo"
 AdvSt (POL, NIL, NIL, NIL, +STRESS): "sí". # "sí"
 AdvSt (POL, NIL, NIL, NIL, +STRESS): "quizá". # "quizá"
 AdvSt (POL, NIL, NIL, NIL, +STRESS): "quizás". # "quizás"
 AdvSt (POL, NIL, NIL, NIL, +STRESS): "acaso". # "acaso"
 AdvSt (POL, NIL, NIL, NIL, +STRESS): "siquiera". # "siquiera"
 AdvSt (POL, NIL, NIL, NIL, +STRESS): "además". # "además"
 AdvSt (POL, NIL, NIL, NIL, +STRESS): "incluso". # "incluso"
 AdvSt (POL, NIL, NIL, NIL, +STRESS): "inclusive". # "inclusive"
 AdvSt (POL, NIL, NIL, NIL, +STRESS): "asimismo". # "asimismo"
 AdvSt (POL, NIL, NIL, NIL, +STRESS): "también". # "también"
 AdvSt (POL, NIL, NIL, NIL, +STRESS): "tampoco". # "tampoco"

Appendix F

Determiners

Prepositional DeterminerSt (A, ART, EL, PRE, DEFINITE, MASC, SING):
"al". # "al"
Prepositional DeterminerSt (DE, ART, EL, PRE, DEFINITE, MASC,
SING): "del". # "del"
DeterminerSt (INTG, QUE, PRE, DEFINITE, MASC|FEM|NEUT, SING|PLU):
"qué". # "qué"
DeterminerSt (INTG, CUANTO, PRE, DEFINITE, MASC|NEUT, SING):
"cuánto". # "cuánto"
DeterminerSt (REL, CUYO, PRE, DEFINITE, MASC|NEUT, SING): "cuyo".
"cuyo"
DeterminerSt (REL, CUAL, PRE, DEFINITE, MASC|FEM|NEUT, SING):
"cual". # "cual"
DeterminerSt (REL, CUANTO, PRE, DEFINITE, MASC|NEUT, SING):
"cuanto". # "cuanto"
DeterminerSt (COMPTVE, EQUAL, PRE, UNDEFINITE, MASC|NEUT, SING):
"tanto". # "tanto"
DeterminerSt (COMPTVE, NoEQUAL, position, UNDEFINITE, MASC|FEM|NEUT,
SING|PLU): "más". # "más"
DeterminerSt (COMPTVE, NoEQUAL, position, UNDEFINITE, MASC|FEM|NEUT,
SING|PLU): "menos". # "menos"
DeterminerSt (ART, EL, PRE, DEFINITE, MASC, SING): "el". # "el"
DeterminerSt (ART|CARD, UNO, PRE, UNDEFINITE, MASC|NEUT, SING):
"un". # "un"
DeterminerSt (ART|CARD, UNO, POST, UNDEFINITE, MASC|NEUT, SING):
"uno". # "un"
DeterminerSt (ART, UNO, PRE, UNDEFINITE, MASC, PLU): "unos". # "un"
DeterminerSt (POSS, NIL, PRE, DEFINITE, FIRST, MASC|FEM|NEUT,
SING): "mi". # "mi"
DeterminerSt (POSS, NIL, PRE, DEFINITE, SECOND, MASC|FEM|NEUT,
SING): "tu". # "tu"
DeterminerSt (POSS, NIL, PRE, DEFINITE, THIRD|SIXTH, MASC|FEM|NEUT,

SING): "su". #"su"
 DeterminerSt (POSS, NIL, position, DEFINITE, FOURTH, MASC|NEUT,
 SING): "nuestro". #"nuestro"
 DeterminerSt (POSS, NIL, position, DEFINITE, FIFTH, MASC|NEUT,
 SING): "vuestro". #"vuestro"
 DeterminerSt (POSS, NIL, POST, DEFINITE, FIRST, MASC|NEUT, SING):
 "mío". #"mío"
 DeterminerSt (POSS, NIL, POST, DEFINITE, SECOND, MASC|NEUT, SING):
 "tuyo". #"tuyo"
 DeterminerSt (POSS, NIL, POST, DEFINITE, THIRD|SIXTH, MASC|NEUT,
 SING): "suyo". #"suyo"
 DeterminerSt (REFL, NIL, position, UNDEFINITE, MASC|NEUT, SING):
 "mismo". #"mismo"
 DeterminerSt (ORD, CUARTO, PRE, UNDEFINITE, MASC|NEUT, SING):
 "primer". #"primer"
 DeterminerSt (ORD, CUARTO, POST, UNDEFINITE, MASC|NEUT, SING):
 "primero". #"primer"
 DeterminerSt (ORD, CUARTO, position, UNDEFINITE, MASC|NEUT, SING):
 "segundo". #"segundo"
 DeterminerSt (ORD|PART, CUARTO, position, UNDEFINITE, FEM, SING):
 "tercera". #"tercer|tercera"
 DeterminerSt (GLOBAL, NIL, PRE, DEFINITE, MASC|NEUT, SING): "todo".
 #"todo"
 DeterminerSt (DEM, TAL, PRE, DEFINITE, MASC|FEM, SING): "tal".
 #"tal"
 DeterminerSt (DEM, ESTE, position, DEFINITE, MASC, SING): "este".
 #"este"
 DeterminerSt (DEM, ESTE, position, DEFINITE, MASC, SING): "ese".
 #"ese"
 DeterminerSt (DEM, ESTE, position, DEFINITE, MASC, SING): "aquel".
 #"aquel"
 DeterminerSt (QUANT, SENDOS, PRE, UNDEFINITE, MASC, PLU): "sendos".
 #"sendos"
 DeterminerSt (QUANT, AMBOS, PRE, UNDEFINITE, MASC, PLU): "ambos".
 #"ambos"
 DeterminerSt (QUANT, VARIOS, PRE, UNDEFINITE, MASC, PLU): "varios".
 #"varios"
 DeterminerSt (QUANT, POCO, PRE, UNDEFINITE, MASC|NEUT, SING):
 "poco". #"poco"
 DeterminerSt (QUANT, MUCHO, PRE, UNDEFINITE, MASC|NEUT, SING):
 "mucho". #"mucho"
 DeterminerSt (QUANT, MUCHO, PRE, UNDEFINITE, MASC|FEM|NEUT, SING):
 "bastante". #"bastante"
 DeterminerSt (QUANT, DEMASIADO, PRE, UNDEFINITE, MASC|NEUT, SING):
 "demasiado". #"demasiado"
 DeterminerSt (INDEF, CIERTO, PRE, UNDEFINITE, MASC|NEUT, SING):
 "cierto". #"cierto"
 DeterminerSt (INDEF, OTRO, PRE, UNDEFINITE, MASC|NEUT, SING):
 "otro". #"otro"
 DeterminerSt (INDEF, ALGUNO, PRE, UNDEFINITE, MASC, PLU):

"unos cuantos". # "unos cuantos"
 DeterminerSt (INDEF, ALGUNO, PRE, UNDEFINITE, MASC|NEUT, SING):
 "algún". # "algún"
 DeterminerSt (INDEF, ALGUNO, POST, UNDEFINITE, MASC|NEUT, SING):
 "alguno". # "algún"
 DeterminerSt (INDEF, ALGUNO, PRE, UNDEFINITE, MASC|NEUT, SING):
 "ningún". # "ningún"
 DeterminerSt (INDEF, ALGUNO, POST, UNDEFINITE, MASC|NEUT, SING):
 "ninguno". # "ningún"
 DeterminerSt (INDEF, CUALQUIERA, PRE, UNDEFINITE, MASC|FEM|NEUT,
 SING): "cualquier". # "cualquier"
 DeterminerSt (INDEF, CUALQUIERA, POST, UNDEFINITE, MASC|FEM|NEUT,
 SING): "cualquiera". # "cualquier"
 DeterminerSt (INDEF, DEMAS, PRE, UNDEFINITE, MASC|FEM, PLU):
 "demás". # "demás"
 DeterminerSt (DIS, NIL, PRE, UNDEFINITE, MASC|FEM|NEUT, SING|PLU):
 "cada". # "cada"
 DeterminerSt (CARD, NoUNIT, position, UNDEFINITE, MASC|FEM, PLU):
 "cero". # "cero"
 DeterminerSt (CARD, NoUNIT+UNO, PRE, UNDEFINITE, MASC, PLU):
 "veintiún". # "veintiún"
 DeterminerSt (CARD, NoUNIT+UNO, POST, UNDEFINITE, MASC, PLU):
 "veintiuno". # "veintiún"
 DeterminerSt (PART, MEDIO, PRE, UNDEFINITE, MASC|NEUT, SING):
 "medio". # "medio"

Appendix G

Examples of Analysis

G.1 Examples from Section 3.1

The following AGFL parse trees correspond to the example sequences introduced in Section 3.1. These parse trees have been obtained using the parser generated with the grammar for phrase-level analysis described in Chapter 4. Therefore, certain examples, those which contain subordinated clauses, are simplified in that the constituents of such subordinated clauses, which would be placed by the analysis at the same level of phrases that have the main clause as their mother node, have been removed, leaving subordinated clauses reduced to their syntactic marker, a conjunction or a relative or interrogative phrase (see Section G.3). The same goes for exclamative phrases. As these are considered to be one of the possible markers of exclamative mood, they receive an analysis that includes a mother node labelled as an exclamative main clause.

For these and the rest of the examples in this appendix, the parser has been run with the option `-P 1`, i.e. it has been forced to return only one analysis. As can be observed, parsing times are quite heterogeneous. The 1.7.52 (BETA) version of the AGFL system used to develop the grammar described in this work and to generate the parser run over these examples cannot, in fact, be considered suitable for tests about parsing times. Our experience, at the moment, is that very little changes in the grammar might sometimes have unexpected consequences in parsing times. But, as already indicated in other parts of this work, AGFL is continuously evolving and being improved, and later versions of the system have

been developed taking this into account. Discontinuous lines connecting syntactic functions at the same level in the parse trees are not part of the AGFL output, but they have been introduced here to facilitate the reading of the analyses.

una mucho más vertiginosa subida de los precios que en Brasil\$EOS\$

parsing 1 (12.943 s.)

penalty: 10

UTTERANCE

Nominal Phrase (COMMON, NOTHING, ART|CARD, UNO, UNDEFINITE, THIRD, FEM, SING)

DETERMINER (ART|CARD, UNO, PRE, UNDEFINITE, FEM, SING)

| Determiner (ART|CARD, UNO, PRE, UNDEFINITE, FEM, SING)

| "una"

Nominal Phrase NOMINAL I (COMMON, QUANT, PRPPHRASE, NIL, SUP, THIRD, FEM, SING)

Noun Phrase (COMMON, QUANT, PRPPHRASE, NIL, SUP, UNDEFINITE, THIRD, FEM, SING)

Comparison Quantifier MODIFIER (QUANT, MUCHO, NIL, SUP)

| Adverb (QUANT, MUCHO, NIL, SUP, +STRESS)

| "mucho"

Noun Phrase NUCLEUS (COMMON, QUANT, PRPPHRASE, PARTICULAR+, NoEQUAL, UNDEFINITE, THIRD, FEM, SING)

Noun Phrase (COMMON, DEGREE, PRPPHRASE, PARTICULAR+, NoEQUAL, UNDEFINITE, THIRD, FEM, SING)

Degree MODIFIER (COMPTVE, NoEQUAL, PARTICULAR+, NoEQUAL)

| Adverb (COMPTVE, NoEQUAL, NIL, NoEQUAL, -STRESS)

| "más"

Noun Phrase NUCLEUS (COMMON, DEGREE, PRPPHRASE, NIL, POS, UNDEFINITE, THIRD, FEM, SING)

| Noun Phrase (COMMON, ADJECTIVE, PRPPHRASE, NIL, POS, UNDEFINITE, THIRD, FEM, SING)

| Adjective MODIFIER (QUAL, NIL, POS, FEM, SING)

| Adjective (QUAL, PRE, POS, FEM, SING)

| "vertiginosa"

Noun Phrase NUCLEUS (COMMON, ADJECTIVE, PRPPHRASE, NIL, NIL, UNDEFINITE, THIRD, FEM, SING)

Noun Phrase (COMMON, NOTHING, PRPPHRASE, NIL, NIL, UNDEFINITE, THIRD, FEM, SING)

Noun Phrase NUCLEUS (COMMON, NOTHING, PRPPHRASE, NIL, NIL, UNDEFINITE, THIRD, FEM, SING)

| Noun (COMMON, FEM, SING)

| "subida"

Prepositional Phrase MODIFIER (DE)

Prepositional Nominal Phrase (DE, COMMON, NOTHING, ART, EL, DEFINITE, THIRD, MASC, PLU)

Prepositional LINKER (DE)

```

|         | Preposition(DE)
|         | "de"
| LINKED Nominal Phrase(COMMON,NOTHING,ART,
| EL,DEFINITE,THIRD,MASC,PLU)
|         Nominal Phrase(COMMON,NOTHING,ART,EL,
|         DEFINITE,THIRD,MASC,PLU)
|         DETERMINER(ART,EL,PRE,DEFINITE,MASC,PLU)
|         | Determiner(ART,EL,PRE,DEFINITE,MASC,
|         | PLU)
|         | "los"
|         Nominal Phrase NOMINAL I(COMMON,NOTHING,
|         NOTHING,NIL,NIL,THIRD,MASC,PLU)
|         Noun(COMMON,MASC,PLU)
|         "precios"
Second Term MODIFIER(COMPTVE,PARTICULAR+,NoEQUAL,QUE)
Comparison Second Term(COMPTVE,PARTICULAR+,NoEQUAL,
QUE)
Comparison LINKER(PARTICULAR+,NoEQUAL,QUE)
| Conjunction(QUE)
| ConjunctionSt(QUE)
| "que"
LINKED Second Term(QUE)
Second Term(QUE)
Prepositional Noun Phrase(EN,Alphabetic,
NOTHING,NOTHING,NIL,NIL,DEFINITE,THIRD,
MASC|FEM,SING)
Prepositional LINKER(EN)
| Preposition(EN)
| "en"
LINKED Noun Phrase(Alphabetic,NOTHING,
NOTHING,NIL,NIL,DEFINITE,THIRD,MASC|FEM,SING)
Noun(Alphabetic,MASC|FEM,SING)
"Brasil"

```

printing time was: (0.003 s.)

all: 1 (prescan: 0.147 s., parsing: 64.615 s., printing: 0.004 s.)

penalty: 10

una manifestación evidente de esquizofrenia a la que\$EOS\$

parsing 1 (2.273 s.)

penalty: 14

UTTERANCE

```

Nominal Phrase(COMMON,NOTHING,ART|CARD,UNO,UNDEFINITE,THIRD,FEM,
SING)
DETERMINER(ART|CARD,UNO,PRE,UNDEFINITE,FEM,SING)

```

```

| Determiner(ART|CARD,UNO,PRE,UNDEFINITE,FEM,SING)
|  "una"
Nominal Phrase NOMINAL I(COMMON,NOTHING,RELCLAUSE,NIL,POS,THIRD,
FEM,SING)
  Noun Phrase(COMMON,NOTHING,RELCLAUSE,NIL,POS,UNDEFINITE,THIRD,
FEM,SING)
    Noun Phrase NUCLEUS(COMMON,NOTHING,RELCLAUSE,NIL,POS,
|UNDEFINITE,THIRD,FEM,SING)
      | Noun Phrase(COMMON,NOTHING,PRPPHRASE,NIL,POS,UNDEFINITE,
|  THIRD,FEM,SING)
        | Noun Phrase NUCLEUS(COMMON,NOTHING,PRPPHRASE,NIL,POS,
|  |UNDEFINITE,THIRD,FEM,SING)
          | | Noun Phrase(COMMON,NOTHING,AJPHRASE,NIL,POS,
|  | |UNDEFINITE,THIRD,FEM,SING)
            | | Noun Phrase NUCLEUS(COMMON,NOTHING,AJPHRASE,NIL,NIL,
|  | | |UNDEFINITE,THIRD,FEM,SING)
              | | | Noun(COMMON,FEM,SING)
                | | |  "manifestación"
                  | | | Adjective Phrase MODIFIER(QUAL,NIL,POS,FEM,SING)
                    | | | Adjective(QUAL,POST,POS,FEM,SING)
                      | | |  "evidente"
                        | | Prepositional Phrase MODIFIER(DE)
                          | | | Prepositional Noun Phrase(DE,COMMON,NOTHING,NOTHING,
|  | | |NIL,NIL,UNDEFINITE,THIRD,FEM,SING)
                            | | | Prepositional LINKER(DE)
                              | | | | Preposition(DE)
                                | | | |  "de"
                                  | | | LINKED Noun Phrase(COMMON,NOTHING,NOTHING,NIL,NIL,
|  | | |UNDEFINITE,THIRD,FEM,SING)
                                    | | | Noun(COMMON,FEM,SING)
                                      | | |  "esquizofrenia"
                                        | Relative Clause MODIFIER(AGR_MOD,THIRD,FEM,SING)
                                          | Relative Clause(AGR_MOD,A,THIRD,FEM,SING)
                                            | Relative CONSTITUENT(AGR_MOD,A,THIRD,FEM,SING)
                                              | Prepositional Pronominal Phrase(A,NOTHING,ART,EL,REL,
|  | | |QUE,DEFINITE,THIRD,FEM,SING)
                                                | Prepositional LINKER(A)
                                                  | | Preposition(A)
                                                    | |  "a"
                                                      | LINKED Pronominal Phrase(NOTHING,ART,EL,REL,QUE,
|  | | |DEFINITE,THIRD,FEM,SING)
                                                        | Pronominal Phrase(NOTHING,ART,EL,REL,QUE,DEFINITE,
|  | | |THIRD,FEM,SING)
                                                          | DETERMINER(ART,EL,PRE,DEFINITE,FEM,SING)
                                                            | | Determiner(ART,EL,PRE,DEFINITE,FEM,SING)
                                                              | |  "la"
                                                                | Pronominal Phrase NOMINAL I(REL,QUE,NOTHING,
|  | | |NOTHING,DEFINITE,THIRD,FEM,SING)
                                                                  | Pronoun(REL,QUE,DEFINITE,FEM,SING)
                                                                    |  "que"

```

printing time was: (0.002 s.)

all: 1 (prescan: 0.067 s., parsing: 6.157 s., printing: 0.003 s.)

penalty: 14

muy pocos días de expansión\$EOS\$

parsing 1 (0.162 s.)

penalty: 10

UTTERANCE

Nominal Phrase(COMMON,DEGREE,QUANT,POCO,UNDEFINITE,THIRD,MASC,PLU)
Degree MODIFIER(QUANT,MUCHO,NIL,SUP)
| Adverb(QUANT,MUCHO,NIL,SUP,-STRESS)
| "muy"
Nominal Phrase NUCLEUS I(QUANT,POCO,UNDEFINITE,COMMON,NOTHING,
PRPPHRASE,NIL,NIL,THIRD,MASC,PLU)
Nominal Phrase I(QUANT,POCO,UNDEFINITE,COMMON,NOTHING,
PRPPHRASE,NIL,NIL,THIRD,MASC,PLU)
DETERMINER(QUANT,POCO,PRE,UNDEFINITE,MASC,PLU)
| Determiner(QUANT,POCO,PRE,UNDEFINITE,MASC,PLU)
| "pocos"
Nominal Phrase NOMINAL I(COMMON,NOTHING,PRPPHRASE,NIL,NIL,
THIRD,MASC,PLU)
Noun Phrase(COMMON,NOTHING,PRPPHRASE,NIL,NIL,UNDEFINITE,
THIRD,MASC,PLU)
Noun Phrase NUCLEUS(COMMON,NOTHING,PRPPHRASE,NIL,NIL,
|UNDEFINITE,THIRD,MASC,PLU)
| Noun(COMMON,MASC,PLU)
| "días"
Prepositional Phrase MODIFIER(DE)
Prepositional Noun Phrase(DE,COMMON,NOTHING,NOTHING,
NIL,NIL,UNDEFINITE,THIRD,FEM,SING)
Prepositional LINKER(DE)
| Preposition(DE)
| "de"
LINKED Noun Phrase(COMMON,NOTHING,NOTHING,NIL,NIL,
UNDEFINITE,THIRD,FEM,SING)
Noun(COMMON,FEM,SING)
"expansión"

printing time was: (0.001 s.)

all: 1 (prescan: 0.035 s., parsing: 0.225 s., printing: 0.002 s.)

penalty: 10

muy pocos que\$EOS\$

parsing 1 (0.054 s.)

penalty: 9

UTTERANCE

Pronoun Phrase(DEGREE,RELCLAUSE,QUANT,POCO,UNDEFINITE,THIRD,
MASC,PLU)

Degree MODIFIER(QUANT,MUCHO,NIL,SUP)

| Adverb(QUANT,MUCHO,NIL,SUP,-STRESS)

| "muy"

Pronoun Phrase NUCLEUS(DEGREE,RELCLAUSE,QUANT,POCO,UNDEFINITE,
THIRD,MASC,PLU)

Pronoun Phrase(NOTHING,RELCLAUSE,QUANT,POCO,UNDEFINITE,THIRD,
MASC,PLU)

Pronoun Phrase NUCLEUS(NOTHING,RELCLAUSE,QUANT,POCO,

|UNDEFINITE,THIRD,MASC,PLU)

| Pronoun(QUANT,POCO,UNDEFINITE,MASC,PLU)

| "pocos"

Relative Clause MODIFIER(AGR_MOD,THIRD,MASC,PLU)

Relative Clause(AGR_MOD,NIL,THIRD,MASC,PLU)

Relative CONSTITUENT(AGR_MOD,NIL,THIRD,MASC,PLU)

Pronoun(REL,QUE,DEFINITE,MASC,PLU)

"que"

printing time was: (0.001 s.)

all: 1 (prescan: 0.014 s., parsing: 0.060 s., printing: 0.002 s.)

penalty: 9

ahora que\$EOS\$

parsing 1 (0.012 s.)

penalty: 5

UTTERANCE

Adverb Phrase(NOTHING,RELCLAUSE,TIME,AHORA,NIL,NIL)

Adverb Phrase NUCLEUS(NOTHING,RELCLAUSE,TIME,AHORA,NIL,NIL,NIL)

| Adverb(TIME,AHORA,NIL,NIL,+STRESS)

| "ahora"

Relative Clause MODIFIER(AdvP)

Relative Clause(AdvP,NIL,NIL,NIL,NIL)

Relative CONSTITUENT(AdvP,NIL,NIL,NIL,NIL)

Adverb(REL,QUE,NIL,NIL,+STRESS)

"que"

printing time was: (0.001 s.)

all: 1 (prescan: 0.008 s., parsing: 0.026 s., printing: 0.001 s.)

penalty: 5

tanto más cerca de la pantalla que nosotros como ellos\$EOS\$

parsing 1 (0.171 s.)

penalty: 5

UTTERANCE

```
Adverb Phrase (QUANT, PRPPHRASE, PLACE, DETRAS, NIL, EQUAL)
  Comparison Quantifier MODIFIER (COMPTVE, EQUAL, NIL, EQUAL)
  | Adverb (COMPTVE, EQUAL, NIL, EQUAL, +STRESS)
  | "tanto"
Adverb Phrase NUCLEUS (QUANT, PRPPHRASE, PLACE, DETRAS, DE,
| PARTICULAR+, NoEQUAL)
| Adverb Phrase (DEGREE, PRPPHRASE, PLACE, DETRAS, DE, PARTICULAR+,
| NoEQUAL)
| Degree MODIFIER (COMPTVE, NoEQUAL, PARTICULAR+, NoEQUAL)
| | Adverb (COMPTVE, NoEQUAL, NIL, NoEQUAL, -STRESS)
| | "más"
| Adverb Phrase NUCLEUS (DEGREE, PRPPHRASE, PLACE, DETRAS, DE, NIL,
| | POS)
| | Adverb Phrase (NOTHING, PRPPHRASE, PLACE, DETRAS, DE, NIL, POS)
| | | Adverb Phrase NUCLEUS (NOTHING, PRPPHRASE, PLACE, DETRAS,
| | | | DE, NIL, POS)
| | | | Adverb (PLACE, DETRAS, DE, POS, +STRESS)
| | | | "cerca"
| | Prepositional Phrase MODIFIER (DE)
| | | Prepositional Nominal Phrase (DE, COMMON, NOTHING, ART,
| | | | EL, DEFINITE, THIRD, FEM, SING)
| | | | Prepositional LINKER (DE)
| | | | | Preposition (DE)
| | | | | "de"
| | | LINKED Nominal Phrase (COMMON, NOTHING, ART, EL,
| | | | DEFINITE, THIRD, FEM, SING)
| | | | Nominal Phrase (COMMON, NOTHING, ART, EL, DEFINITE,
| | | | | THIRD, FEM, SING)
| | | | | DETERMINER (ART, EL, PRE, DEFINITE, FEM, SING)
| | | | | | Determiner (ART, EL, PRE, DEFINITE, FEM, SING)
| | | | | | "la"
| | | | Nominal Phrase NOMINAL I (COMMON, NOTHING, NOTHING,
| | | | | NIL, NIL, THIRD, FEM, SING)
| | | | | Noun (COMMON, FEM, SING)
| | | | | "pantalla"
| Second Term MODIFIER (COMPTVE, PARTICULAR+, NoEQUAL, QUE)
| Comparison Second Term (COMPTVE, PARTICULAR+, NoEQUAL, QUE)
| Comparison LINKER (PARTICULAR+, NoEQUAL, QUE)
```

```

|      | Conjunction(QUE)
|      |   ConjunctionSt(QUE)
|      |     "que"
|      LINKED Second Term(QUE)
|      Second Term(QUE)
|      Pronoun(PERS,NIL,DEFINITE,FIRST,MASC,PLU,NOM|PREP)
|      "nosotros"
Second Term MODIFIER(COMPTVE,NIL,EQUAL,COMO)
Comparison Second Term(COMPTVE,NIL,EQUAL,COMO)
Comparison LINKER(NIL,EQUAL,COMO)
| Conjunction(COMO)
|   ConjunctionSt(COMO)
|     "como"
LINKED Second Term(COMO)
Second Term(COMO)
Pronoun(PERS,NIL,DEFINITE,THIRD,MASC,PLU,NOM|PREP)
"ellos"

```

printing time was: (0.002 s.)

all: 1 (prescan: 0.098 s., parsing: 13.119 s., printing: 0.003 s.)

penalty: 5

qué poco delicado\$EOS\$

parsing 1 (0.001 s.)

penalty: 3

UTTERANCE

```

Exclamative Clause(MAIN,non_imperative)
Exclamative CONSTITUENT
Adjective Phrase(EXCL,QUAL,NIL,NIL,NEUT|MASC,SING)
Degree MODIFIER(EXCL,NIL,NIL,NIL)
| Adverb(EXCL,NIL,NIL,NIL,-STRESS)
|   "qué"
Adjective Phrase NUCLEUS(EXCL,NOTHING,QUAL,NIL,POS-,
NEUT|MASC,SING)
Adjective Phrase(DEGREE,NOTHING,QUAL,NIL,POS-,NEUT|MASC,
SING)
Degree MODIFIER(QUANT,POCO,NIL,POS-)
| Adverb(QUANT,POCO,NIL,POS-,-STRESS)
|   "poco"
Adjective Phrase NUCLEUS(DEGREE,NOTHING,QUAL,NIL,POS,
NEUT|MASC,SING)
Adjective(QUAL,POST,POS,NEUT|MASC,SING)
"delicado"

```

printing time was: (0.001 s.)

all: 1 (prescan: 0.013 s., parsing: 0.121 s., printing: 0.003 s.)

penalty: 3

todos los demás participantes\$EOS\$

parsing 1 (0.113 s.)

penalty: 4

UTTERANCE

Nominal Phrase (COMMON, NOTHING, GLOBAL, NIL, DEFINITE, THIRD, MASC, PLU)

DETERMINER (GLOBAL, NIL, PRE, DEFINITE, MASC, PLU)

| Determiner (GLOBAL, NIL, PRE, DEFINITE, MASC, PLU)

| "todos"

Nominal Phrase NOMINAL III (ART, EL, COMMON, NOTHING, NOTHING, NIL, NIL, THIRD, MASC, PLU)

Nominal Phrase II (ART, EL, COMMON, NOTHING, NOTHING, NIL, NIL, THIRD, MASC, PLU)

DETERMINER (ART, EL, PRE, DEFINITE, MASC, PLU)

| Determiner (ART, EL, PRE, DEFINITE, MASC, PLU)

| "los"

Nominal Phrase NOMINAL II (INDEF, DEMAS, COMMON, NOTHING, NOTHING, NIL, NIL, UNDEFINITE, THIRD, MASC, PLU)

Nominal Phrase I (INDEF, DEMAS, UNDEFINITE, COMMON, NOTHING, NOTHING, NIL, NIL, THIRD, MASC, PLU)

DETERMINER (INDEF, DEMAS, PRE, UNDEFINITE, MASC, PLU)

| Determiner (INDEF, DEMAS, PRE, UNDEFINITE, MASC, PLU)

| "demás"

Nominal Phrase NOMINAL I (COMMON, NOTHING, NOTHING, NIL, NIL, THIRD, MASC, PLU)

Noun (COMMON, MASC, PLU)

"participantes"

printing time was: (0.001 s.)

all: 1 (prescan: 0.026 s., parsing: 0.132 s., printing: 0.001 s.)

penalty: 4

mucha más importancia de la que\$EOS\$

parsing 1 (0.005 s.)

penalty: 7

UTTERANCE

Nominal Phrase (COMMON, NOTHING, QUANT, MUCHO, UNDEFINITE, THIRD, FEM, SING)

```

DETERMINER(QUANT,MUCHO,PRE,UNDEFINITE,FEM,SING)
| Determiner(QUANT,MUCHO,PRE,UNDEFINITE,FEM,SING)
|  "mucha"
Nominal Phrase NOMINAL II(COMPTVE,NoEQUAL,COMMON,NOTHING,
NOTHING,PARTICULAR+,NIL,UNDEFINITE,THIRD,FEM,SING)
  Nominal Phrase I(COMPTVE,NoEQUAL,UNDEFINITE,COMMON,NOTHING,
  NOTHING,PARTICULAR+,NIL,THIRD,FEM,SING)
    DETERMINER(COMPTVE,NoEQUAL,PRE,UNDEFINITE,FEM,SING)
      | Determiner(COMPTVE,NoEQUAL,PRE,UNDEFINITE,FEM,SING)
      |  "más"
    Nominal Phrase NOMINAL I(COMMON,NOTHING,NOTHING,NIL,NIL,
    |THIRD,FEM,SING)
      | Noun(COMMON,FEM,SING)
      |  "importancia"
    Second Term MODIFIER(COMPTVE,PARTICULAR+,NoEQUAL,DE,THIRD,
    FEM,SING)
      Prepositional Nominalization(DE,RELCLAUSE,DEFINITE,THIRD,
      FEM,SING)
        Prepositional LINKER(DE)
          | Preposition(DE)
          |  "de"
        LINKED Nominalization(RELCLAUSE,NIL,DEFINITE,THIRD,
        FEM,SING)
          Nominalization(RELCLAUSE,NIL,DEFINITE,THIRD,FEM,SING)
            DETERMINER(ART,EL,PRE,DEFINITE,FEM,SING)
              | Determiner(ART,EL,PRE,DEFINITE,FEM,SING)
              |  "la"
            Nominalization NOMINAL I(RELCLAUSE,NIL,THIRD,FEM,
            SING)
              Relative Clause(NOMINAL,NIL,THIRD,FEM,SING)
                Relative CONSTITUENT(NOMINAL,NIL,THIRD,FEM,
                SING)
                  Pronoun(REL,QUE,DEFINITE,FEM,SING)
                    "que"

```

printing time was: (0.001 s.)

all: 1 (prescan: 0.030 s., parsing: 0.557 s., printing: 0.003 s.)

penalty: 7

los otros cuatro\$EOS\$

parsing 1 (0.047 s.)

penalty: 3

UTTERANCE

```

Pronominal Phrase(NOTHING,ART,EL,CARD,NoUNIT,DEFINITE,THIRD,
MASC,PLU)

```

```

DETERMINER(ART,EL,PRE,DEFINITE,MASC,PLU)
| Determiner(ART,EL,PRE,DEFINITE,MASC,PLU)
| "los"
Pronominal Phrase NOMINAL II(INDEF,OTRO,THIRD,MASC,PLU,CARD,
NoUNIT,NOTHING,NOTHING)
Pronominal Phrase I(INDEF,OTRO,THIRD,MASC,PLU,CARD,NoUNIT,
NOTHING,NOTHING)
DETERMINER(INDEF,OTRO,PRE,UNDEFINITE,MASC,PLU)
| Determiner(INDEF,OTRO,PRE,UNDEFINITE,MASC,PLU)
| "otros"
Pronominal Phrase NOMINAL I(CARD,NoUNIT,NOTHING,NOTHING,
DEFINITE,THIRD,MASC,PLU)
Pronoun(CARD,NoUNIT,DEFINITE,MASC,PLU)
"cuatro"

```

printing time was: (0.001 s.)

all: 1 (prescan: 0.018 s., parsing: 0.061 s., printing: 0.001 s.)

penalty: 3

G.2 Full Sentence Analysis

The following AGFL parse trees correspond to the example sequences in Section 4.1.1.2 and 4.1.1.3.

```
dijo que cogiera eso otro.$EOS$
```

```
-----
parsing 1 (0.005 s.)
```

penalty: 9

UTTERANCE

```

Declarative(INDICATIVE)
  Clause(MAIN,INDICATIVE)
    PREDICATE(SDO,ACT,THIRD,SING,INDICATIVE)
    | Verb Phrase(SDO,ACT,THIRD,SING,INDICATIVE)
    |   Verb(SDO,ACT,THIRD,SING,PAST,INDICATIVE)
    |   "dijo"
    DIRECT OBJECT(const_category,THIRD,NEUT,SING,THIRD)
    Clause(QUE,SUBJUNCTIVE)
      Conjunctive LINKER(QUE)
      | Conjunction(QUE)
      |   ConjunctionSt(QUE)
      |   "que"
    LINKED Clause(SUBORDINATED,SUBJUNCTIVE)
      Clause(SUBORDINATED,SUBJUNCTIVE)
        PREDICATE(SDO,ACT,FIRST|THIRD,SING,SUBJUNCTIVE)

```

```

| Verb Phrase(SDO,ACT,FIRST|THIRD,SING,
| SUBJUNCTIVE)
|   Verb(SDO,ACT,FIRST|THIRD,SING,IMPERFECT,
|     SUBJUNCTIVE)
|     "cogiera"
DIRECT OBJECT(const_category,THIRD,NEUT,SING,
FIRST|THIRD)
  Pronominal Phrase(NOTHING,DEM,ESTE,INDEF,OTRO,
DEFINITE,THIRD,NEUT,SING)
    DETERMINER(DEM,ESTE,PRE,DEFINITE,NEUT,SING)
      | Determiner(DEM,ESTE,PRE,DEFINITE,NEUT,SING)
      |   "eso"
    Pronominal Phrase NOMINAL I(INDEF,OTRO,NOTHING,
NOTHING,UNDEFINITE,THIRD,NEUT,SING)
      Pronoun(INDEF,OTRO,UNDEFINITE,NEUT,SING)
        "otro"

```

","

printing time was: (0.001 s.)

Break: maximum number of parses (1) reached
all: 1 (prescan: 0.039 s., parsing: 0.867 s., printing: 0.005 s.)

penalty: 9

dijo que les había hecho algunas preguntas a las que no
habían podido responder.\$EOS\$

parsing 1 (1.839 s.)

penalty: 13

```

UTTERANCE
  Declarative(INDICATIVE)
    Clause(MAIN,INDICATIVE)
      PREDICATE(SDO,ACT,THIRD,SING,INDICATIVE)
        | Verb Phrase(SDO,ACT,THIRD,SING,INDICATIVE)
        |   Verb(SDO,ACT,THIRD,SING,PAST,INDICATIVE)
        |     "dijo"
      DIRECT OBJECT(const_category,THIRD,NEUT,SING,THIRD)
        Clause(QUE,INDICATIVE)
          Conjunctive LINKER(QUE)
            | Conjunction(QUE)
            |   ConjunctionSt(QUE)
            |     "que"
          LINKED Clause(SUBORDINATED,INDICATIVE)
            Clause(SUBORDINATED,INDICATIVE)
              PREDICATE(SDOIO,ACT,FIRST|THIRD,SING,INDICATIVE,DAT,

```

```

|THIRD,MASC|FEM,PLU)
| Verb Phrase(SDOI0,ACT,FIRST|THIRD,SING,
| INDICATIVE,DAT,THIRD,MASC|FEM,PLU)
| Clitic Pronoun(THIRD,MASC|FEM,PLU,DAT,ONE_OF_ONE)
| "les"
| Verb(SDOI0,ACT,FIRST|THIRD,SING,PLUPERFECT,
| INDICATIVE)
| Verb(AUX,ACT,FIRST|THIRD,SING,IMPERFECT,
| INDICATIVE)
| "había"
| Participle Verb(SDOI0,ACT,PARTICIPLE,MASC,SING)
| "hecho"
DIRECT OBJECT(const_category,THIRD,FEM,PLU,
FIRST|THIRD)
Nominal Phrase(COMMON,NOTHING,INDEF,ALGUNO,
UNDEFINITE,THIRD,FEM,PLU)
DETERMINER(INDEF,ALGUNO,PRE,UNDEFINITE,FEM,PLU)
| Determiner(INDEF,ALGUNO,PRE,UNDEFINITE,FEM,PLU)
| "algunas"
Nominal Phrase NOMINAL I(COMMON,NOTHING,
RELCLAUSE,NIL,NIL,THIRD,FEM,PLU)
Noun Phrase(COMMON,NOTHING,RELCLAUSE,NIL,NIL,
UNDEFINITE,THIRD,FEM,PLU)
Noun Phrase NUCLEUS(COMMON,NOTHING,RELCLAUSE,
|NIL,NIL,UNDEFINITE,THIRD,FEM,PLU)
| Noun(COMMON,FEM,PLU)
| "preguntas"
Relative Clause MODIFIER(AGR_MOD,THIRD,FEM,
PLU)
Relative Clause(AGR_MOD,A,THIRD,FEM,PLU)
Relative PREPOSITIONAL COMPLEMENT(
|const_category,AGR_MOD,A,THIRD,FEM,PLU)
| Prepositional Pronominal Phrase(A,
| NOTHING,ART,EL,REL,QUE,DEFINITE,THIRD,
| FEM,PLU)
| Prepositional LINKER(A)
| | Preposition(A)
| | "a"
| LINKED Pronominal Phrase(NOTHING,ART,
| EL,REL,QUE,DEFINITE,THIRD,FEM,PLU)
| Pronominal Phrase(NOTHING,ART,EL,
| REL,QUE,DEFINITE,THIRD,FEM,PLU)
| DETERMINER(ART,EL,PRE,DEFINITE,
| FEM,PLU)
| | Determiner(ART,EL,PRE,DEFINITE,
| FEM,PLU)
| | "las"
| Pronominal Phrase NOMINAL I(REL,
| QUE,NOTHING,NOTHING,DEFINITE,
| THIRD,FEM,PLU)

```

```

|           Pronoun(REL,QUE,DEFINITE,FEM,
|           PLU)
|           "que"
PREDICATE(SPC,ACT,A,THIRD,PLU,INDICATIVE)
Periphrastic Verb Phrase(SPC,ACT,A,
THIRD,PLU,INDICATIVE)
Periphrasis(INFINITIVE,ACT,THIRD,PLU,
INDICATIVE)
"no"
Verb(INFINITIVE,ACT,THIRD,PLU,
PLUPERFECT,INDICATIVE)
Verb(AUX,ACT,THIRD,PLU,IMPERFECT,
INDICATIVE)
"habían"
Participle Verb(INFINITIVE,ACT,
PARTICIPLE,MASC,SING)
"podido"
Non Finite Verb Phrase(SPC,ACT,A,
INFINITIVE)
Verb(SPC,ACT,A,PRESENT,INFINITIVE)
"responder"

```

","

printing time was: (0.003 s.)

Break: maximum number of parses (1) reached
all: 1 (prescan: 0.184 s., parsing: 78.878 s., printing: 0.010 s.)

penalty: 13

G.3 Phrase Analysis

The following example is extracted from the test documents provided by the company *Informática El Corte Inglés*, the spanish user involved in the DoRo project. The text analysed is the following one:

Cualquiera de las partes podrá resolver el mismo, comunicando por escrito a la otra parte, su intención en tal sentido con una antelación mínima de tres meses. En el caso de que la resolución del contrato se efectuase antes de la finalización del período contratado y a instancias de I.E.C.I.S.A, esta sociedad vendrá obligada a devolver a GRUPO ASNKF EQUIFAX el importe proporcional al período liberado. En el caso de que la cancelación del contrato fuese a instancias de GRUPO

ASNEF EQUIFAX, I.E.C.I.S.A no vendrá obligado a devolver cantidad alguna. Cualquier cambio en las condiciones de este contrato exige la firma de uno nuevo que anule al vigente.

This is a brief paragraph of a document, but it is very representative of the domain. The type of analysis obtained can be considered simple, but it is suitable for a Document Routing application as that involved in the DoRo project. Two main points have to be highlighted in the analyses, these also indicate future further extensions of the grammar: the coordination *antes de la finalización del período contratado y a instancias de I.E.C.I.S.A*, and the adjective modifier in *proporcional al período liberado*, which are not covered. Apart from these, the structural ambiguity found in *la firma de uno nuevo que*, where the relative clause may be a modifier of either *firma de uno nuevo* or *uno nuevo* is also interesting. According to the preferences established in Section 4.2.2.2.1, the first possibility is showed in the analysis.

cualquiera de las partes podrá resolver el mismo, comunicando por escrito a la otra parte, su intención en tal sentido con una antelación mínima de tres meses.\$EOS\$

parsing 1 (4.640 s.)

penalty: 149

UTTERANCE

Pronoun Phrase(NOTHING,PRPPHRASE,INDEF,CUALQUIERA,UNDEFINITE,THIRD,MASC,SING)

Pronoun Phrase NUCLEUS(NOTHING,PRPPHRASE,INDEF,CUALQUIERA,UNDEFINITE,THIRD,MASC,SING)

| Pronoun(INDEF,CUALQUIERA,UNDEFINITE,MASC,SING)

| "cualquiera"

Prepositional Phrase MODIFIER(DE)

Prepositional Nominal Phrase(DE,COMMON,NOTHING,ART,EL,DEFINITE,THIRD,FEM,PLU)

Prepositional LINKER(DE)

| Preposition(DE)

| "de"

LINKED Nominal Phrase(COMMON,NOTHING,ART,EL,DEFINITE,THIRD,FEM,PLU)

Nominal Phrase(COMMON,NOTHING,ART,EL,DEFINITE,THIRD,FEM,PLU)

DETERMINER(ART,EL,PRE,DEFINITE,FEM,PLU)

```

    | Determiner(ART,EL,PRE,DEFINITE,FEM,PLU)
    | "las"
Nominal Phrase NOMINAL I(COMMON,NOTHING,NOTHING,NIL,NIL,
THIRD,FEM,PLU)
    Noun(COMMON,FEM,PLU)
    "partes"
Periphrastic Verb Phrase(mvtype,ACT,THIRD,SING,INDICATIVE)
Periphrasis(INFINITIVE,ACT,THIRD,SING,INDICATIVE)
    Verb(INFINITIVE,ACT,THIRD,SING,FUTURE,INDICATIVE)
    "podrá"
Non Finite Verb Phrase(mvtype,ACT,INFINITIVE)
    Verb(mvtype,ACT,PRESENT,INFINITIVE)
    "resolver"
Pronominal Phrase(NOTHING,ART,EL,REFL,NIL,DEFINITE,THIRD,MASC,
SING)
    DETERMINER(ART,EL,PRE,DEFINITE,MASC,SING)
    | Determiner(ART,EL,PRE,DEFINITE,MASC,SING)
    | "el"
Pronominal Phrase NOMINAL I(REFL,NIL,NOTHING,NOTHING,UNDEFINITE,
THIRD,MASC,SING)
    Pronoun(REFL,NIL,UNDEFINITE,MASC,SING)
    "mismo"
Punctuation Mark
    ","
Clause(SUBORDINATED,INFINITIVE,PERSON)
    NON FINITE PREDICATE(GERUND)
    Non Finite Verb Phrase(mvtype,ACT,GERUND)
    Verb(mvtype,ACT,PRESENT,GERUND)
    "comunicando"
Adverb(MANNER,NIL,NIL,NIL,+STRESS)
    "por escrito"
Prepositional Nominal Phrase(A,COMMON,NOTHING,ART,EL,DEFINITE,i
THIRD,FEM,SING)
    Prepositional LINKER(A)
    | Preposition(A)
    | "a"
LINKED Nominal Phrase(COMMON,NOTHING,ART,EL,DEFINITE,THIRD,FEM,
SING)
    Nominal Phrase(COMMON,NOTHING,ART,EL,DEFINITE,THIRD,FEM,SING)
    DETERMINER(ART,EL,PRE,DEFINITE,FEM,SING)
    | Determiner(ART,EL,PRE,DEFINITE,FEM,SING)
    | "la"
Nominal Phrase NOMINAL II(INDEF,OTRO,COMMON,NOTHING,NOTHING,
NIL,NIL,UNDEFINITE,THIRD,FEM,SING)
    Nominal Phrase I(INDEF,OTRO,UNDEFINITE,COMMON,NOTHING,
NOTHING,NIL,NIL,THIRD,FEM,SING)
    DETERMINER(INDEF,OTRO,PRE,UNDEFINITE,FEM,SING)
    | Determiner(INDEF,OTRO,PRE,UNDEFINITE,FEM,SING)
    | "otra"
Nominal Phrase NOMINAL I(COMMON,NOTHING,NOTHING,NIL,

```

NIL,THIRD,FEM,SING)
 Noun(COMMON,FEM,SING)
 "parte"
 Punctuation Mark
 ", "
 Nominal Phrase(COMMON,NOTHING,POSS,NIL,DEFINITE,THIRD,FEM,SING)
 DETERMINER(POSS,NIL,PRE,DEFINITE,FEM,SING)
 | Determiner(POSS,NIL,PRE,DEFINITE,SIXTH|THIRD,FEM,SING)
 | "su"
 Nominal Phrase NOMINAL I(COMMON,NOTHING,NOTHING,NIL,NIL,THIRD,
 FEM,SING)
 Noun(COMMON,FEM,SING)
 "intención"
 Prepositional Nominal Phrase(EN,COMMON,NOTHING,DEM,TAL,DEFINITE,
 THIRD,MASC,SING)
 Prepositional LINKER(EN)
 | Preposition(EN)
 | "en"
 LINKED Nominal Phrase(COMMON,NOTHING,DEM,TAL,DEFINITE,THIRD,
 MASC,SING)
 Nominal Phrase(COMMON,NOTHING,DEM,TAL,DEFINITE,THIRD,MASC,
 SING)
 DETERMINER(DEM,TAL,PRE,DEFINITE,MASC,SING)
 | Determiner(DEM,TAL,PRE,DEFINITE,MASC,SING)
 | "tal"
 Nominal Phrase NOMINAL I(COMMON,NOTHING,NOTHING,NIL,NIL,
 THIRD,MASC,SING)
 Noun(COMMON,MASC,SING)
 "sentido"
 Prepositional Nominal Phrase(CON,COMMON,NOTHING,ART|CARD,UNO,
 UNDEFINITE,THIRD,FEM,SING)
 Prepositional LINKER(CON)
 | Preposition(CON)
 | "con"
 LINKED Nominal Phrase(COMMON,NOTHING,ART|CARD,UNO,UNDEFINITE,
 THIRD,FEM,SING)
 Nominal Phrase(COMMON,NOTHING,ART|CARD,UNO,UNDEFINITE,THIRD,
 FEM,SING)
 DETERMINER(ART|CARD,UNO,PRE,UNDEFINITE,FEM,SING)
 | Determiner(ART|CARD,UNO,PRE,UNDEFINITE,FEM,SING)
 | "una"
 Nominal Phrase NOMINAL I(COMMON,NOTHING,PRPPHASE,NIL,POS,
 THIRD,FEM,SING)
 Noun Phrase(COMMON,NOTHING,PRPPHASE,NIL,POS,UNDEFINITE,
 THIRD,FEM,SING)
 Noun Phrase NUCLEUS(COMMON,NOTHING,PRPPHASE,NIL,POS,
 |UNDEFINITE,THIRD,FEM,SING)
 | Noun Phrase(COMMON,NOTHING,AJPHRASE,NIL,POS,
 | UNDEFINITE,THIRD,FEM,SING)
 | Noun Phrase NUCLEUS(COMMON,NOTHING,AJPHRASE,NIL,

```

| |NIL,UNDEFINITE,THIRD,FEM,SING)
| | Noun(COMMON,FEM,SING)
| | "antelación"
| Adjective Phrase MODIFIER(QUAL,NIL,POS,FEM,SING)
| Adjective(QUAL,POST,POS,FEM,SING)
| "minima"
Prepositional Phrase MODIFIER(DE)
  Prepositional Nominal Phrase(DE,COMMON,NOTHING,CARD,
    NoUNIT,UNDEFINITE,THIRD,MASC,PLU)
    Prepositional LINKER(DE)
      | Preposition(DE)
      | "de"
LINKED Nominal Phrase(COMMON,NOTHING,CARD,NoUNIT,
  UNDEFINITE,THIRD,MASC,PLU)
    Nominal Phrase(COMMON,NOTHING,CARD,NoUNIT,
      UNDEFINITE,THIRD,MASC,PLU)
        DETERMINER(CARD,NoUNIT,PRE,UNDEFINITE,MASC,PLU)
          | Determiner(CARD,NoUNIT,PRE,UNDEFINITE,MASC,
            | PLU)
            | "tres"
          Nominal Phrase NOMINAL I(COMMON,NOTHING,NOTHING,
            NIL,NIL,THIRD,MASC,PLU)
            Noun(COMMON,MASC,PLU)
              "meses"

```

Punctuation Mark
 "."

printing time was: (0.002 s.)

all: 1 (prescan: 0.076 s., parsing: 4.647 s., printing: 0.003 s.)

penalty: 149

en el caso de que la resolución del contrato se efectuase antes
 de la finalización del período contratado y a instancias de
 I.E.C.I.S.A, esta sociedad vendrá obligada a devolver a GRUPO
 ASNKF EQUIFAX el importe proporcional al período liberado.\$EOS\$

 parsing 1 (4.108 s.)

penalty: 193

UTTERANCE

```

  Prepositional Nominal Phrase(EN,COMMON,NOTHING,ART,EL,DEFINITE,
    THIRD,MASC,SING)
    Prepositional LINKER(EN)
      | Preposition(EN)

```

```

| "en"
LINKED Nominal Phrase(COMMON,NOTHING,ART,EL,DEFINITE,THIRD,
MASC,SING)
  Nominal Phrase(COMMON,NOTHING,ART,EL,DEFINITE,THIRD,MASC,SING)
    DETERMINER(ART,EL,PRE,DEFINITE,MASC,SING)
      | Determiner(ART,EL,PRE,DEFINITE,MASC,SING)
      | "el"
  Nominal Phrase NOMINAL I(COMMON,NOTHING,PRPPHRASE,NIL,NIL,
THIRD,MASC,SING)
    Noun Phrase(COMMON,NOTHING,PRPPHRASE,NIL,NIL,UNDEFINITE,
THIRD,MASC,SING)
      Noun Phrase NUCLEUS(COMMON,NOTHING,PRPPHRASE,NIL,NIL,
|UNDEFINITE,THIRD,MASC,SING)
        | Noun(COMMON,MASC,SING)
        | "caso"
      Prepositional Phrase MODIFIER(DE)
        Prepositional Clause(DE,QUE,non_imperative)
          Prepositional LINKER(DE)
            | Preposition(DE)
            | "de"
          LINKED Clause(QUE,non_imperative)
            Clause(QUE,non_imperative)
              Conjunctive LINKER(QUE)
                Conjunction(QUE)
                  ConjunctionSt(QUE)
                    "que"
  Nominal Phrase(COMMON,NOTHING,ART,EL,DEFINITE,THIRD,FEM,SING)
    DETERMINER(ART,EL,PRE,DEFINITE,FEM,SING)
      | Determiner(ART,EL,PRE,DEFINITE,FEM,SING)
      | "la"
  Nominal Phrase NOMINAL I(COMMON,NOTHING,PRPPHRASE,NIL,NIL,
THIRD,FEM,SING)
    Noun Phrase(COMMON,NOTHING,PRPPHRASE,NIL,NIL,UNDEFINITE,
THIRD,FEM,SING)
      Noun Phrase NUCLEUS(COMMON,NOTHING,PRPPHRASE,NIL,NIL,
|UNDEFINITE,THIRD,FEM,SING)
        | Noun(COMMON,FEM,SING)
        | "resolución"
      Prepositional Phrase MODIFIER(DE)
        Prepositional Nominal Phrase(DE,COMMON,NOTHING,ART,EL,
DEFINITE,THIRD,MASC,SING)
          Prepositional LINKER(DE,ART,EL,PRE,DEFINITE,MASC,SING)
            | Prepositional Determiner(DE,ART,EL,PRE,DEFINITE,MASC,
| SING)
            | "del"
          Nominal Phrase NOMINAL I(COMMON,NOTHING,NOTHING,NIL,
NIL,THIRD,MASC,SING)
            Noun(COMMON,MASC,SING)
              "contrato"
  Verb Phrase(mvtype,MIDD,THIRD,SING,SUBJUNCTIVE)

```

Clitic Pronoun (THIRD, NEUT | MASC | FEM, SING, VERB)
 "se"
 Verb (mvtype, MIDD, THIRD, SING, IMPERFECT, SUBJUNCTIVE)
 "efectuase"
 Adverb Phrase (NOTHING, NOTHING, TIME, NIL, PARTICULAR+, NoEQUAL)
 Adverb Phrase NUCLEUS (NOTHING, NOTHING, TIME, NIL, NIL, PARTICULAR+,
 | NoEQUAL)
 | Adverb (TIME, NIL, NIL, NoEQUAL, +STRESS)
 | "antes"
 Second Term MODIFIER (COMPTVE, PARTICULAR+, NoEQUAL, DE)
 Comparison Second Term (COMPTVE, PARTICULAR+, NoEQUAL, DE)
 LINKED Second Term (DE)
 Second Term (DE)
 Prepositional Nominal Phrase (DE, COMMON, NOTHING, ART, EL,
 DEFINITE, THIRD, FEM, SING)
 Prepositional LINKER (DE)
 | Preposition (DE)
 | "de"
 LINKED Nominal Phrase (COMMON, NOTHING, ART, EL, DEFINITE,
 THIRD, FEM, SING)
 Nominal Phrase (COMMON, NOTHING, ART, EL, DEFINITE, THIRD,
 FEM, SING)
 DETERMINER (ART, EL, PRE, DEFINITE, FEM, SING)
 | Determiner (ART, EL, PRE, DEFINITE, FEM, SING)
 | "la"
 Nominal Phrase NOMINAL I (COMMON, NOTHING, PRPPHRASE,
 NIL, NIL, THIRD, FEM, SING)
 Noun Phrase (COMMON, NOTHING, PRPPHRASE, NIL, NIL,
 UNDEFINITE, THIRD, FEM, SING)
 Noun Phrase NUCLEUS (COMMON, NOTHING, PRPPHRASE,
 | NIL, NIL, UNDEFINITE, THIRD, FEM, SING)
 | Noun (COMMON, FEM, SING)
 | "finalización"
 Prepositional Phrase MODIFIER (DE)
 Prepositional Nominal Phrase (DE, COMMON,
 NOTHING, ART, EL, DEFINITE, THIRD, MASC, SING)
 Prepositional LINKER (DE, ART, EL, PRE,
 | DEFINITE, MASC, SING)
 | Prepositional Determiner (DE, ART, EL, PRE,
 | DEFINITE, MASC, SING)
 | "del"
 Nominal Phrase NOMINAL I (COMMON, NOTHING,
 AJPHRASE, NIL, POS, THIRD, MASC, SING)
 Noun Phrase (COMMON, NOTHING, AJPHRASE, NIL,
 POS, UNDEFINITE, THIRD, MASC, SING)
 Noun Phrase NUCLEUS (COMMON, NOTHING,
 | AJPHRASE, NIL, NIL, UNDEFINITE, THIRD,
 | MASC, SING)
 | Noun (COMMON, MASC, SING)
 | "período"

```

                                Adjective Phrase MODIFIER(PPL, NIL,
                                POS, MASC, SING)
                                Participle Verb(mvtype, ACT|MIDD,
                                PARTICIPLE, MASC, SING)
                                "contratado"

Conjunction(NIL)
  ConjunctionSt(NIL)
  "y"

Prepositional Noun Phrase(OTHER, Alphabetic, NOTHING, NOTHING, NIL,
NIL, DEFINITE, THIRD, MASC|FEM, SING)
  Prepositional LINKER(OTHER)
  | Preposition(OTHER)
  | "a instancias de"
LINKED Noun Phrase(Alphabetic, NOTHING, NOTHING, NIL, NIL, DEFINITE,
THIRD, MASC|FEM, SING)
  Noun(Alphabetic, MASC|FEM, SING)
  "I"
  "."
  "E"
  "."
  "C"
  "."
  "I"
  "."
  "S"
  "."
  "A"

Punctuation Mark
", "

Nominal Phrase(COMMON, NOTHING, DEM, ESTE, DEFINITE, THIRD, FEM, SING)
  DETERMINER(DEM, ESTE, PRE, DEFINITE, FEM, SING)
  | Determiner(DEM, ESTE, PRE, DEFINITE, FEM, SING)
  | "esta"
  Nominal Phrase NOMINAL I(COMMON, NOTHING, NOTHING, NIL, NIL, THIRD,
  FEM, SING)
  Noun(COMMON, FEM, SING)
  "sociedad"

Verb Phrase(mvtype, ACT, THIRD, SING, INDICATIVE)
  Verb(mvtype, ACT, THIRD, SING, FUTURE, INDICATIVE)
  "vendrá"

Participle Clause(FEM, SING)
  PARTICIPLE PREDICATE(FEM, SING)
  Participle Verb Phrase(mvtype, ACT|MIDD, PARTICIPLE, FEM, SING)
  Participle Verb(mvtype, ACT|MIDD, PARTICIPLE, FEM, SING)
  "obligada"

Prepositional Clause(A, SUBORDINATED, INFINITIVE, person_two)
  Prepositional LINKER(A)
  | Preposition(A)
  | "a"
  LINKED Clause(SUBORDINATED, INFINITIVE, person)

```

```

Clause(SUBORDINATED,INFINITIVE,person)
  NON FINITE PREDICATE(INFINITIVE)
    Non Finite Verb Phrase(mvtype,ACT,INFINITIVE)
      Verb(mvtype,ACT,PRESENT,INFINITIVE)
        "devolver"
Prepositional Noun Phrase(A,Alphabetic,NOTHING,NOTHING,NIL,NIL,
DEFINITE,THIRD,MASC|FEM,SING)
  Prepositional LINKER(A)
    | Preposition(A)
    | "a"
LINKED Noun Phrase(Alphabetic,NOTHING,NOTHING,NIL,NIL,DEFINITE,
THIRD,MASC|FEM,SING)
  Noun(Alphabetic,MASC|FEM,SING)
    "GRUPO"
    "ASNKF"
    "EQUIFAX"
Nominal Phrase(COMMON,NOTHING,ART,EL,DEFINITE,THIRD,MASC,SING)
  DETERMINER(ART,EL,PRE,DEFINITE,MASC,SING)
    | Determiner(ART,EL,PRE,DEFINITE,MASC,SING)
    | "el"
Nominal Phrase NOMINAL I(COMMON,NOTHING,AJPHRASE,NIL,POS,THIRD,
MASC,SING)
  Noun Phrase(COMMON,NOTHING,AJPHRASE,NIL,POS,UNDEFINITE,THIRD,
MASC,SING)
    Noun Phrase NUCLEUS(COMMON,NOTHING,AJPHRASE,NIL,NIL,
|UNDEFINITE,THIRD,MASC,SING)
    | Noun(COMMON,MASC,SING)
    | "importe"
    Adjective Phrase MODIFIER(QUAL,NIL,POS,MASC,SING)
    Adjective(QUAL,POST,POS,MASC,SING)
    "proporcional"
Prepositional Nominal Phrase(A,COMMON,NOTHING,ART,EL,DEFINITE,
THIRD,MASC,SING)
  Prepositional LINKER(A,ART,EL,PRE,DEFINITE,MASC,SING)
    | Prepositional Determiner(A,ART,EL,PRE,DEFINITE,MASC,SING)
    | "al"
Nominal Phrase NOMINAL I(COMMON,NOTHING,AJPHRASE,NIL,POS,THIRD,
MASC,SING)
  Noun Phrase(COMMON,NOTHING,AJPHRASE,NIL,POS,UNDEFINITE,THIRD,
MASC,SING)
    Noun Phrase NUCLEUS(COMMON,NOTHING,AJPHRASE,NIL,NIL,
|UNDEFINITE,THIRD,MASC,SING)
    | Noun(COMMON,MASC,SING)
    | "período"
    Adjective Phrase MODIFIER(PPL,NIL,POS,MASC,SING)
    Participle Verb(mvtype,ACT|MIDD,PARTICIPLE,MASC,SING)
    "liberado"
Punctuation Mark
  "."

```

printing time was: (0.002 s.)

all: 1 (prescan: 0.088 s., parsing: 4.155 s., printing: 0.003 s.)

penalty: 193

en el caso de que la cancelación del contrato fuese a instancias de GRUPO ASNEF EQUIFAX, I.E.C.I.S.A no vendrá obligado a devolver cantidad alguna.\$EOS\$

parsing 1 (4.110 s.)

penalty: 138

UTTERANCE

Prepositional Nominal Phrase(EN,COMMON,NOTHING,ART,EL,DEFINITE,THIRD,MASC,SING)
 Prepositional LINKER(EN)
 | Preposition(EN)
 | "en"
LINKED Nominal Phrase(COMMON,NOTHING,ART,EL,DEFINITE,THIRD,MASC,SING)
 Nominal Phrase(COMMON,NOTHING,ART,EL,DEFINITE,THIRD,MASC,SING)
 DETERMINER(ART,EL,PRE,DEFINITE,MASC,SING)
 | Determiner(ART,EL,PRE,DEFINITE,MASC,SING)
 | "el"
 Nominal Phrase NOMINAL I(COMMON,NOTHING,PRPPHRASE,NIL,NIL,THIRD,MASC,SING)
 Noun Phrase(COMMON,NOTHING,PRPPHRASE,NIL,NIL,UNDEFINITE,THIRD,MASC,SING)
 Noun Phrase NUCLEUS(COMMON,NOTHING,PRPPHRASE,NIL,NIL,UNDEFINITE,THIRD,MASC,SING)
 | Noun(COMMON,MASC,SING)
 | "caso"
 Prepositional Phrase MODIFIER(DE)
 Prepositional Clause(DE,QUE,non_imperative)
 Prepositional LINKER(DE)
 | Preposition(DE)
 | "de"
 LINKED Clause(QUE,non_imperative)
 Clause(QUE,non_imperative)
 Conjunctive LINKER(QUE)
 Conjunction(QUE)
 ConjunctionSt(QUE)
 "que"
 Nominal Phrase(COMMON,NOTHING,ART,EL,DEFINITE,THIRD,FEM,SING)
 DETERMINER(ART,EL,PRE,DEFINITE,FEM,SING)
 | Determiner(ART,EL,PRE,DEFINITE,FEM,SING)

```

| "la"
Nominal Phrase NOMINAL I (COMMON, NOTHING, PRPPHRASE, NIL, NIL, THIRD,
FEM, SING)
  Noun Phrase (COMMON, NOTHING, PRPPHRASE, NIL, NIL, UNDEFINITE, THIRD,
  FEM, SING)
    Noun Phrase NUCLEUS (COMMON, NOTHING, PRPPHRASE, NIL, NIL,
    | UNDEFINITE, THIRD, FEM, SING)
      | Noun (COMMON, FEM, SING)
        | "cancelación"
  Prepositional Phrase MODIFIER (DE)
    Prepositional Nominal Phrase (DE, COMMON, NOTHING, ART, EL,
    DEFINITE, THIRD, MASC, SING)
      Prepositional LINKER (DE, ART, EL, PRE, DEFINITE, MASC, SING)
        | Prepositional Determiner (DE, ART, EL, PRE, DEFINITE, MASC,
        | SING)
          | "del"
      Nominal Phrase NOMINAL I (COMMON, NOTHING, NOTHING, NIL, NIL,
      THIRD, MASC, SING)
        Noun (COMMON, MASC, SING)
          "contrato"
Verb Phrase (mvtype, ACT, FIRST | THIRD, SING, SUBJUNCTIVE)
  Verb (mvtype, ACT, FIRST | THIRD, SING, IMPERFECT, SUBJUNCTIVE)
  "fuese"
Prepositional Noun Phrase (OTHER, Alphabetic, NOTHING, NOTHING, NIL,
NIL, DEFINITE, THIRD, MASC | FEM, SING)
  Prepositional LINKER (OTHER)
  | Preposition (OTHER)
  | "a instancias de"
LINKED Noun Phrase (Alphabetic, NOTHING, NOTHING, NIL, NIL, DEFINITE,
THIRD, MASC | FEM, SING)
  Noun (Alphabetic, MASC | FEM, SING)
  "GRUPO"
  "ASNEF"
  "EQUIFAX"
Punctuation Mark
", "
Noun (Alphabetic, MASC | FEM, SING)
"I"
"."
"E"
"."
"C"
"."
"I"
"."
"S"
"."
"A"
Negative Particle
"no"

```

Verb Phrase(mvtype,ACT,THIRD,SING,INDICATIVE)
 Verb(mvtype,ACT,THIRD,SING,FUTURE,INDICATIVE)
 "vendrá"
 Participle Clause(NEUT|MASC,SING)
 PARTICIPLE PREDICATE(NEUT|MASC,SING)
 Participle Verb Phrase(mvtype,ACT|MIDD,PARTICIPLE,NEUT|MASC,
 SING)
 Participle Verb(mvtype,ACT|MIDD,PARTICIPLE,NEUT|MASC,SING)
 "obligado"
 Prepositional Clause(A,SUBORDINATED,INFINITIVE,person_two)
 Prepositional LINKER(A)
 | Preposition(A)
 | "a"
 LINKED Clause(SUBORDINATED,INFINITIVE,person)
 Clause(SUBORDINATED,INFINITIVE,person)
 NON FINITE PREDICATE(INFINITIVE)
 Non Finite Verb Phrase(mvtype,ACT,INFINITIVE)
 Verb(mvtype,ACT,PRESENT,INFINITIVE)
 "devolver"
 Nominal Phrase(COMMON,NOTHING,INDEF,ALGUNO,UNDEFINITE,THIRD,FEM,
 SING)
 Nominal Phrase NOMINAL I(COMMON,NOTHING,NOTHING,NIL,NIL,THIRD,
 |FEM,SING)
 | Noun(COMMON,FEM,SING)
 | "cantidad"
 DETERMINER(INDEF,ALGUNO,POST,UNDEFINITE,FEM,SING)
 Determiner(INDEF,ALGUNO,POST,UNDEFINITE,FEM,SING)
 "alguna"
 Punctuation Mark
 "."

printing time was: (0.002 s.)

all: 1 (prescan: 0.099 s., parsing: 4.157 s., printing: 0.003 s.)

penalty: 138

cualquier cambio en las condiciones de este contrato exige la firma
 de uno nuevo que anule al vigente.\$EOS\$

 parsing 1 (2.535 s.)

penalty: 98

UTTERANCE

Nominal Phrase(COMMON,NOTHING,INDEF,CUALQUIERA,UNDEFINITE,THIRD,
 MASC,SING)
 DETERMINER(INDEF,CUALQUIERA,PRE,UNDEFINITE,MASC,SING)
 | Determiner(INDEF,CUALQUIERA,PRE,UNDEFINITE,MASC,SING)

```

| "cualquier"
Nominal Phrase NOMINAL I (COMMON, NOTHING, NOTHING, NIL, NIL, THIRD,
MASC, SING)
  Noun (COMMON, MASC, SING)
    "cambio"
Prepositional Nominal Phrase (EN, COMMON, NOTHING, ART, EL, DEFINITE,
THIRD, FEM, PLU)
  Prepositional LINKER (EN)
  | Preposition (EN)
  | "en"
LINKED Nominal Phrase (COMMON, NOTHING, ART, EL, DEFINITE, THIRD, FEM,
PLU)
  Nominal Phrase (COMMON, NOTHING, ART, EL, DEFINITE, THIRD, FEM, PLU)
  DETERMINER (ART, EL, PRE, DEFINITE, FEM, PLU)
  | Determiner (ART, EL, PRE, DEFINITE, FEM, PLU)
  | "las"
  Nominal Phrase NOMINAL I (COMMON, NOTHING, PRPPHRASE, NIL, NIL,
THIRD, FEM, PLU)
  Noun Phrase (COMMON, NOTHING, PRPPHRASE, NIL, NIL, UNDEFINITE,
THIRD, FEM, PLU)
  Noun Phrase NUCLEUS (COMMON, NOTHING, PRPPHRASE, NIL, NIL,
| UNDEFINITE, THIRD, FEM, PLU)
  | Noun (COMMON, FEM, PLU)
  | "condiciones"
  Prepositional Phrase MODIFIER (DE)
  Prepositional Nominal Phrase (DE, COMMON, NOTHING, DEM,
ESTE, DEFINITE, THIRD, MASC, SING)
  Prepositional LINKER (DE)
  | Preposition (DE)
  | "de"
  LINKED Nominal Phrase (COMMON, NOTHING, DEM, ESTE,
DEFINITE, THIRD, MASC, SING)
  Nominal Phrase (COMMON, NOTHING, DEM, ESTE, DEFINITE,
THIRD, MASC, SING)
  DETERMINER (DEM, ESTE, PRE, DEFINITE, MASC, SING)
  | Determiner (DEM, ESTE, PRE, DEFINITE, MASC, SING)
  | "este"
  Nominal Phrase NOMINAL I (COMMON, NOTHING, NOTHING,
NIL, NIL, THIRD, MASC, SING)
  Noun (COMMON, MASC, SING)
    "contrato"
Verb Phrase (mvtype, ACT, THIRD, SING, INDICATIVE)
  Verb (mvtype, ACT, THIRD, SING, PRESENT, INDICATIVE)
    "exige"
Nominal Phrase (COMMON, NOTHING, ART, EL, DEFINITE, THIRD, FEM, SING)
  DETERMINER (ART, EL, PRE, DEFINITE, FEM, SING)
  | Determiner (ART, EL, PRE, DEFINITE, FEM, SING)
  | "la"
  Nominal Phrase NOMINAL I (COMMON, NOTHING, RELCLAUSE, NIL, NIL, THIRD,
FEM, SING)

```

```

Noun Phrase(COMMON,NOTHING,RELCLAUSE,NIL,NIL,UNDEFINITE,THIRD,
FEM,SING)
  Noun Phrase NUCLEUS(COMMON,NOTHING,RELCLAUSE,NIL,NIL,
|UNDEFINITE,THIRD,FEM,SING)
  | Noun Phrase(COMMON,NOTHING,PRPPHRASE,NIL,NIL,UNDEFINITE,
  | THIRD,FEM,SING)
  |   Noun Phrase NUCLEUS(COMMON,NOTHING,PRPPHRASE,NIL,NIL,
  |   |UNDEFINITE,THIRD,FEM,SING)
  |   | Noun(COMMON,FEM,SING)
  |   |   "firma"
  |   Prepositional Phrase MODIFIER(DE)
  |     Prepositional Pronoun Phrase(DE,NOTHING,AJPHRASE,
  |     INDEF|CARD,UNO,UNDEFINITE,THIRD,MASC,SING)
  |     Prepositional LINKER(DE)
  |     | Preposition(DE)
  |     |   "de"
  |     LINKED Pronoun Phrase(NOTHING,AJPHRASE,INDEF|CARD,
  |     UNO,UNDEFINITE,THIRD,MASC,SING)
  |     Pronoun Phrase(NOTHING,AJPHRASE,INDEF|CARD,UNO,
  |     UNDEFINITE,THIRD,MASC,SING)
  |     Pronoun Phrase NUCLEUS(NOTHING,AJPHRASE,
  |     |INDEF|CARD,UNO,UNDEFINITE,THIRD,MASC,SING)
  |     | Pronoun(INDEF|CARD,UNO,UNDEFINITE,MASC,SING)
  |     |   "uno"
  |     Adjective Phrase MODIFIER(QUAL,NIL,POS,MASC,
  |     SING)
  |     Adjective(QUAL,POST,POS,MASC,SING)
  |     "nuevo"
  Relative Clause MODIFIER(AGR_MOD,THIRD,FEM,SING)
  Relative Clause(AGR_MOD,NIL,THIRD,FEM,SING)
  Relative CONSTITUENT(AGR_MOD,NIL,THIRD,FEM,SING)
  Pronoun(REL,QUE,DEFINITE,FEM,SING)
  "que"
Verb Phrase(mvtype,ACT,FIRST|THIRD,SING,SUBJUNCTIVE)
Verb(mvtype,ACT,FIRST|THIRD,SING,PRESENT,SUBJUNCTIVE)
"anule"
Prepositional Nominalization(A,AJPHRASE,DEFINITE,THIRD,MASC,SING)
Prepositional LINKER(A,ART,EL,PRE,DEFINITE,MASC,SING)
| Prepositional Determiner(A,ART,EL,PRE,DEFINITE,MASC,SING)
|   "al"
Nominalization NOMINAL I(AJPHRASE,THIRD,MASC,SING)
Adjective(QUAL,POST,POS,MASC,SING)
"vigente"
Punctuation Mark
","

```

printing time was: (0.002 s.)

all: 1 (prescan: 0.068 s., parsing: 2.539 s., printing: 0.002 s.)

penalty: 98

Bibliography

- [1] E. Alarcos. *Estudios de gramática funcional del español*. Gredos, Madrid, 3rd edition, 1980.
- [2] E. Alarcos. *Gramática de la lengua española*. Real Academia Española, Colección Nebrija y Bello. Espasa-Calpe, Madrid, 1994.
- [3] A. Alcalá, C. Rodríguez, L. de Sopena, and C. Villar. Elaboración de una codificación sintáctica en ordenador de los verbos castellanos. In C. Martín Vide, editor, *Actas del V Congreso de Lenguajes Naturales y Lenguajes Formales, Vilafranca del Penedès, 18-22 de septiembre, 1989*, pages 283–296, University of Barcelona, 1989.
- [4] J. Alcina Franch and J.M. Blecua. *Gramática española*. Ariel, Barcelona, 1975.
- [5] A. Alonso and P. Henríquez Ureña. *Gramática castellana*. Losada, Buenos Aires, 26th (1st Course)-24th (2nd Course) edition, 1971.
- [6] C. Álvarez, P. Alvariño, A. Gil, T. Romero, M.P. Santalla, and S. Sotelo. AVALON, una gramática formal basada en corpus. *Revista de la Sociedad Española para el Procesamiento del Lenguaje Natural*, 23:132–139, 1998.
- [7] A.T. Arampatzisk, T. Tsoris, and C.H.A. Koster. IRENA: Information retrieval engine based on natural language analysis. In *Proceedings of RIAO'97 Computer-Assisted Information Searching on Internet*, pages 159–175, McGill University, Montreal, 1997.

- [8] T. Badia. *Aspectes del sintagma nominal en català des de la perspectiva de la traducció automàtica*. Publicacions de l'Abadia de Montserrat, Barcelona, 1994.
- [9] A. Bello. *Gramática de la lengua castellana destinada al uso de los americanos*. Arco/Libros, Madrid, 1988. Edition prepared by R. Trujillo.
- [10] I. Bosque and V. Demonte, editors. *Gramática descriptiva de la lengua española*. Real Academia Española, Colección Nebrija y Bello. Espasa-Calpe, Madrid, 1999.
- [11] R. Casajuana and C. Rodríguez. Clasificación de los verbos castellanos para un diccionario en ordenador. In C. Martín Vide, editor, *Actas del I Congreso de Lenguajes Naturales y Lenguajes Formales, Barcelona, 14-18 de octubre, 1985*, pages 221–244, University of Barcelona, 1985.
- [12] P.-A. Coppen. The use of AGFL in sequential modular NLP systems. In C.H.A. Koster and E. Oltmans, editors, *Proceedings of the first AGFL Workshop, october, 1995. Technical Report CSI-R9604*, pages 143–151, Catholic University of Nijmegen, 1996.
- [13] C. Dekkers, C.H.A. Koster, M.-J. Nederhof, and A. van Zwol. The Grammar WorkBench: A first step towards lingware engineering. In W.G. ter Stal, A. Nijholt, and H.J. op den Akker, editors, *Linguistic Engineering: Tools and Products, Proceedings of the second Twente Workshop on Language Technology. Memoranda Informatica 92-29*, pages 103–115, Twente, 1992.
- [14] C. Dekkers, C.H.A. Koster, M.-J. Nederhof, and A. van Zwol. *Manual for the Grammar WorkBench Version 1.5*. Computing Science Institute, Department of Informatics, Faculty of Mathematics and Informatics, Catholic University of Nijmegen, 1992.
- [15] C. Derksen. *Manual for the GEN Parser Generator Version 1.5*. Computing Science Institute, Department of Informatics, Faculty of Mathematics and Informatics, Catholic University of Nijmegen, 1995.
- [16] C. Derksen and J. Potjer. *Manual for the AGFL System Version 1.8*. Computing Science Institute, Department of Informatics, Faculty of Mathematics

and Informatics, Catholic University of Nijmegen, 1999. Preliminary version.
URL://www.cs.kun.nl/agfl.

- [17] E. Ditters. An extended affix grammar for the noun phrase in modern standard Arabic. In J. Aarts and W. Meijs, editors, *Corpus Linguistics II. New Studies in the Analysis and Exploitation of Computer Corpora*, pages 47–77. Rodopi, Amsterdam, 1986.
- [18] E. Ditters. Affix grammars and the description of Arabic syntax. In C.H.A. Koster and E. Oltmans, editors, *Proceedings of the first AGFL Workshop, october, 1995. Technical Report CSI-R.9604*, pages 75–91, Catholic University of Nijmegen, 1996.
- [19] M.V. Escandell. *Los complementos del nombre*. Arco/Libros, Madrid, 2nd edition, 1997.
- [20] E. Farkas, C.H.A. Koster, P. Köves, and M. Naszódi. Towards an affix grammar for the Hungarian language. In A. György, editor, *Proceedings CIS'91 Conference on Intelligent Systems*, pages 433–446, John von Neumann Society for Computing Sciences, Budapest, 1991.
- [21] S. Fernández Ramírez. *Gramática española. 4. El verbo y la oración*. Arco/Libros, Madrid, 1986. Edition prepared by I. Bosque.
- [22] S. Fernández Ramírez. *Gramática española. 3.1. El nombre*. Arco/Libros, Madrid, 1987. Edition prepared by J. Polo.
- [23] S. Fernández Ramírez. *Gramática española. 3.2. El pronombre*. Arco/Libros, Madrid, 1987. Edition prepared by J. Polo.
- [24] J.M. García-Miguel. La voz media en español: Las construcciones pronominales con verbos transitivos. *Verba*, 12:307–410, 1985.
- [25] J.M. García-Miguel. Corpus de textos analizados sintácticamente. In J. Gómez, editor, *Aplicaciones lingüísticas de la informática*, pages 19–33. Tórculo, Santiago de Compostela, 1994.

- [26] J.M. García-Miguel. *Las relaciones gramaticales entre predicado y participantes*. Lalia Series Maior, 2. University of Santiago de Compostela, Santiago de Compostela, 1995.
- [27] J.M. García-Miguel. *Transitividad y complementación preposicional en español*. Annex of *Verba*, 40. University of Santiago de Compostela, Santiago de Compostela, 1995.
- [28] A. García Velasco. Análisis computacional de los sintagmas de un texto y de su relación con el verbo de la oración. Aplicaciones al procesamiento del lenguaje natural. In C. Martín Vide, editor, *Actas del VIII Congreso de Lenguajes Naturales y Lenguajes Formales, Gerona, 21-25 de septiembre, 1992*, pages 335–342, University of Barcelona, 1992.
- [29] S. Gili Gaya. *Curso Superior de Sintaxis Española*. Biblograf, Barcelona, 9th edition, 1970.
- [30] J. Hallebeek. Aspectos de una gramática formal de la frase verbal en español. In C. Martín Vide, editor, *Actas del III Congreso de Lenguajes Naturales y Lenguajes Formales, Sitges, 28 de septiembre a 2 de octubre, 1987*, pages 545–558, University of Barcelona, 1988.
- [31] J. Hallebeek. Una gramática de afijo extendida para el español. *Boletín de la SEPLN*, 7:147–156, 1989.
- [32] J. Hallebeek. Nominalización funcional y análisis sintáctico automatizado. In C. Martín Vide, editor, *Actas del IV Congreso de Lenguajes Naturales y Lenguajes Formales, Tarragona, 17-21 de septiembre, 1990*, pages 621–640, University of Barcelona, 1990.
- [33] J. Hallebeek. *A Formal Approach to Spanish Syntax*. Rodopi, Amsterdam-Atlanta, 1992.
- [34] J. Hallebeek. Un nuevo tipo de *affixgrammar* para el español. In C. Martín Vide, editor, *Actas del X Congreso de Lenguajes Naturales y Lenguajes Formales, Sevilla, 26-30 de octubre, 1994*, pages 419–426, University of Barcelona, 1994.

- [35] H. van Halteren and N. Oostdijk. Towards a syntactic database: The TOSCA analysis system. In J. Aarts, P. de Haan, and N. Oostdijk, editors, *English Language Corpora: Design, Analysis and Exploitation. Papers from the thirteenth International Conference on English Language Research on Computerized Corpora*, pages 145–162. Rodopi, Nijmegen, 1992.
- [36] C. Hernández Alonso. *Sintaxis española*. Published by the author, Valladolid, 3rd edition, 1970.
- [37] T. Jiménez Juliá. La caracterización de las unidades sintácticas en español. In *Actas del III Congreso de la Asociación Asiática de Hispanistas, Universidad de Seisen, Tokio, 8-10 de enero, 1993*, pages 146–156, Tokio, 1993.
- [38] T. Jiménez Juliá. Clases de construcciones, tipos de unidad y oraciones. In F. Hernández Paricio, editor, *Perspectivas sobre la oración*, pages 145–183. Universidad de Zaragoza, Zaragoza, 1994.
- [39] T. Jiménez Juliá. Endocentrismo y exocentrismo en un marco funcional. In M. Iglesias Bango, editor, *Gramma-temas 2*, Colección Contextos, pages 297–331. Centro de Estudios metodológicos y disciplinares. Universidad de León, León, 1997.
- [40] H. Kontouli and A. Mountzia. Morphological and syntactical description of the Greek language with the AGFL system. Master’s thesis, University of Patras, 1992.
- [41] C.H.A. Koster. Affix grammars. In J.E.L. Peck, editor, *Proceedings of the IFIP Working Conference on ALGOL69 Implementation*, pages 95–109. North Holland Publishing Company, Amsterdam, 1971.
- [42] C.H.A. Koster. Affix grammars for natural languages. In H. Albas and B. Melichar, editors, *Attribute Grammars, Applications and Systems, International Summer School SAGA, Prague, June, 1991, Lecture Notes in Computer Science 545*, pages 469–484. Springer-Verlag, Heidelberg, 1991.
- [43] C.H.A. Koster. Developments in AGFL. In C.H.A. Koster and E. Oltmans, editors, *Proceedings of the first AGFL Workshop, october, 1995. Technical Report CSI-R9604*, pages 45–57, Catholic University of Nijmegen, 1996.

- [44] C.H.A. Koster. The family of affix grammars. In C.H.A. Koster and E. Oltmans, editors, *Proceedings of the first AGFL Workshop, october, 1995. Technical Report CSI-R9604*, pages 5–15, Catholic University of Nijmegen, 1996.
- [45] C.H.A. Koster and C.P.A. Tiberius. Affix grammars for full-text information retrieval. In A. Ralli et al., editors, *Working Papers in Natural Language Processing*, pages 139–154. Ekdotis Diavlos, Athens, 1997.
- [46] C.H.A. Koster and R. Willems. Towards an affix grammar for Turkish. In M. Baray and B. Ozguc, editors, *Computer and Information Science VI*, pages 1067–1076. Elsevier, Nijmegen, 1991.
- [47] R. Lenz. *La oración y sus partes*. Centro de Estudios Históricos, Madrid, 3rd edition, 1935.
- [48] F. Marcos Marín. *Curso de gramática española*. Cincel, Madrid, 1980.
- [49] L.G.L.T. Meertens and C.H.A. Koster. An affix grammar for a part of the English language. In *Euratom Colloquium, University of Amsterdam, 1962*, 1962. Unpublished, reprints may be obtained from the authors.
- [50] M.-J. Nederhof and C.H.A. Koster. A customized grammar workbench. In J. Aarts, P. de Haan, and N. Oostdijk, editors, *English Language Corpora: Design, Analysis and Exploitation. Papers from the thirteenth International Conference on English Language Research on Computerized Corpora*, pages 163–179. Rodopi, Nijmegen, 1992.
- [51] E. Oltmans. AMAZON in AGFL. Een contextvrije herschrijfgrammatica voor de structurele component van het AMAZON/CASUS-systeem, beschreven in het AGFL-formalisme. Master's thesis, Catholic University of Nijmegen, 1994.
- [52] E. Oltmans. Developing a robust grammar in AGFL based on linguistic principles. In C.H.A. Koster and E. Oltmans, editors, *Proceedings of the first AGFL Workshop, october, 1995. Technical Report CSI-R9604*, pages 59–74, Catholic University of Nijmegen, 1996.

- [53] N. Oostdijk. Extended affix grammar and English syntax: A formal description of the English noun phrase. Master's thesis, Catholic University of Nijmegen, 1984.
- [54] N. Oostdijk. An extended affix grammar for the English noun phrase. In J. Aarts and W. Meijs, editors, *Corpus Linguistics. Recent Developments in the Use of Computer Corpora in English Language Research*, pages 95–122. Rodopi, Amsterdam, 1984.
- [55] N. Oostdijk. Coordination and gapping in corpus analysis. In J. Aarts and W. Meijs, editors, *Corpus Linguistics II. New Studies in the Analysis and Exploitation of Computer Corpora*, pages 177–202. Rodopi, Amsterdam, 1986.
- [56] N. Oostdijk. *Corpus Linguistics and the Automatic Analysis of English*. Rodopi, Amsterdam-Atlanta, 1991.
- [57] J. Pena. La voz en español. Intento de caracterización. *Verba*, 9:215–252, 1982.
- [58] Real Academia Española. *Gramática de la lengua española*. Espasa-Calpe, Madrid, 1931.
- [59] Real Academia Española. *Esbozo de una nueva gramática de la lengua española*. Espasa-Calpe, Madrid, 1973.
- [60] J. Roca Pons. *Introducción a la gramática*. Teide, Barcelona, 1970.
- [61] C. Rodríguez Margo, L. de Sopena, and C. Villar. Caracterización sintáctica de los verbos castellanos para una gramática en ordenador. *Boletín de la SEPLN*, 7:99–104, 1989.
- [62] G. Rojo. Sobre la coordinación de adjetivos en la frase nominal y cuestiones conexas. *Verba*, 2:193–224, 1975.
- [63] G. Rojo. *Cláusulas y oraciones*. University of Santiago de Compostela, Santiago de Compostela, 1978.
- [64] G. Rojo. *Aspectos básicos de sintaxis funcional*. Ágora, Málaga, 1983.

- [65] G. Rojo. El futuro *diccionario de construcciones verbales del español actual*. In C. Martín Vide, editor, *Actas del VIII Congreso de Lenguajes Naturales y Lenguajes Formales, Gerona, 21-25 de septiembre, 1992*, pages 41–50, University of Barcelona, 1992.
- [66] G. Rojo. La base de datos sintácticos del español actual. *Español Actual*, 59:15–20, 1993. The volume was published in 1995.
- [67] G. Rojo. Problemas lingüísticos e informáticos en los diccionarios de construcción y régimen. In *Actas del Congreso de la Lengua Española, Sevilla, 7-10 de mayo, 1992*, pages 307–315, Instituto Cervantes, Madrid, 1994.
- [68] G. Rojo, M.C. Álvarez, P. Alvariño, A. Gil, M.P. Santalla, and S. Sotelo. An architecture for document routing in spanish: Two language components, pre-processor and parser. In M. Gavrilidou, G. Carayannis, S. Markantonatou, S. Piperidis, and G. Stainhaouer, editors, *Proceedings of the Second International Conference on Language Resources and Evaluation*, volume II, pages 675–682. National Technical University of Athens Press, Athens, 2000.
- [69] G. Rojo and T. Jiménez Juliá. *Fundamentos del análisis sintáctico funcional*. Lalia, 2, Serie Lingüística. University of Santiago de Compostela, Santiago de Compostela, 1989.
- [70] M.P. Santalla. An AGFL formal grammar for the Spanish noun phrase. URL://www.cs.kun.nl/agfl, 1994.
- [71] M.P. Santalla and J. Cloeren. A practical example for the use of AGFL in the automatic syntactic analysis of Spanish. In C.H.A. Koster and E. Oltmans, editors, *Proceedings of the first AGFL Workshop, october, 1995. Technical Report CSI-R9604*, pages 123–141, Catholic University of Nijmegen, 1996.
- [72] R. Seco. *Manual de gramática española*. Aguilar, Madrid, 11th edition, 1988.
- [73] M. Seutter. Informal introduction to the Extended Affix Grammar formalism and its compiler. Technical Report 93-19, Computing Science Institute, Department of Informatics, Faculty of Mathematics and Informatics, Catholic University of Nijmegen, 1993.

- [74] G. Tabuteau. Integration of AGFL in an information retrieval system. In C.H.A. Koster and E. Oltmans, editors, *Proceedings of the first AGFL Workshop, october, 1995. Technical Report CSI-R9604*, pages 143–151, Catholic University of Nijmegen, 1996.
- [75] C.P.A. Tiberius. An annotated AGFL grammar for English. Unpublished, reprints may be obtained from the author, 1997.
- [76] T.G. Triantopoulou. A description of the Modern Greek noun phrase using the “Affix Grammar over a Finite Lattice” formalism. *Literary and Linguistic Computing*, 12(2):119–133, 1997.

