UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE FILOSOFIA E CIÊNCIAS HUMANAS

MARÍA INÉS CORBALÁN

FROM GENERATIVE LINGUISTICS TO CATEGORIAL GRAMMARS: OVERT SUBJECTS IN CONTROL INFINITIVES

DA LINGUÍSTICA GERATIVA À GRAMÁTICA CATEGORIAL: SUJEITOS LEXICAIS EM INFINITIVOS CONTROLADOS

CAMPINAS
2018

MARÍA INÉS CORBALÁN

FROM GENERATIVE LINGUISTICS TO CATEGORIAL GRAMMARS: OVERT SUBJECTS IN CONTROL INFINITIVES

DA LINGUÍSTICA GERATIVA À GRAMÁTICA CATEGORIAL: SUJEITOS LEXICAIS EM INFINITIVOS CONTROLADOS

Thesis presented to the Institute of Philosophy and Human Sciences of the University of Campinas in partial fulfilment of the requirements for the degree of Doctor in Philosophy.

Tese apresentada ao Instituto de Filosofia e Ciências Humanas da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Doutora em Filosofia.

Supervisor: MARCELO ESTEBAN CONIGLIO
Co-supervisor: SONIA MARIA LAZZARINI CYRINO

ESTE EXEMPLAR CORRESPONDE À VERSÃO FINAL DA TESE DEFENDIDA PELA ALUNA MARÍA INÉS CORBALÁN, E ORIENTADA PELO PROFESSOR MARCELO ESTEBAN CONIGLIO.

CAMPINAS
2018

Agência(s) de fomento e nº(s) de processo(s): FAPESP, 2013/08115-1; FAPESP, 2015/09699-2; CAPES
ORCID: https://orcid.org/0000-0001-8036-715X

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Filosofia e Ciências Humanas
Cecília Maria Jorge Nicolau - CRB 8/3387

C81f

Corbalán, María Inés, 1978-
From generative linguistics to categorial grammars : overt subjects in control infinitives / María Inés Corbalán. – Campinas, SP : [s.n.], 2018.

Orientador: Marcelo Esteban Coniglio.
Coorientador: Sonia Maria Lazzarini Cyrino.
Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Filosofia e Ciências Humanas.

1. Controle (Linguística). 2. Línguas românicas - Pronome. 3. Anáfora (Linguística). 4. Gramática categorial. 5. Gramática gerativa. I. Coniglio, Marcelo Esteban, 1963-. II. Cyrino, Sonia Maria Lazzarini, 1957-. III. Universidade Estadual de Campinas. Instituto de Filosofia e Ciências Humanas. IV. Título.

Informações para Biblioteca Digital

Título em outro idioma: Da linguística gerativa à gramática categorial : sujeitos lexicais em infinitivos controlados
Palavras-chave em inglês:
Control (Linguistics)
Romance languages - Pronoun
Anaphora (Linguistics)
Categorial grammar
Generative grammar
Área de concentração: Filosofia
Titulação: Doutora em Filosofia
Banca examinadora:
Marcelo Esteban Coniglio [Orientador]
Marco Antonio Caron Ruffino
Marcelo Barra Ferreira
Luiz Arthur Pagani
Marcelo Finger
Data de defesa: 27-02-2018
Programa de Pós-Graduação: Filosofia

UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE FILOSOFIA E CIÊNCIAS HUMANAS

A Comissão Julgadora dos trabalhos de Defesa de Tese de Doutorado, composta pelos Professores Doutores a seguir descritos, em sessão pública realizada em 27 de fevereiro de 2018, considerou a candidata María Inés Corbalán aprovada.

Prof. Dr. Marcelo Esteban Coniglio

Prof. Dr. Marco Antonio Caron Ruffino

Prof. Dr. Marcelo Barra Ferreira

Prof. Dr. Luiz Arthur Pagani

Prof. Dr. Marcelo Finger

A Ata de Defesa, assinada pelos membros da Comissão Examinadora, consta no processo de vida acadêmica da aluna.

a ti, gt

Agradecimientos

Agradezco a la *Fundação de Amparo à Pesquisa do Estado de São Paulo* (FAPESP) la concesión de dos becas de estudios, con números de proceso 2013/08115-1 y 2015/09699-2. Gracias a la segunda beca pude realizar una estancia de un año en la *Universitat Politècnica de Catalunya* (UPC) bajo la orientación de Glyn Morrill. Glyn ha sido un excelente orientador, y me ha apoyado tanto académica como personalmente. Desde nuestro primer encuentro en Barcelona, inclusive antes del comienzo de mi beca, él ha estado dispuesto a trabajar y a discutir los problemas de mi investigación. Y más importante aún, él me ha acompañado emocionalmente durante toda la estancia en Barcelona. Su personalidad y su compromiso ante el trabajo ciertamente constituyen un modelo para mí. Agradezco también a Oriol Valentín, colaborador y amigo. A Oriol le agradezco las discusiones de trabajo en la UPC, y las horas de charlas sobre todo y nada en nuestros viajes en metro. Agradezco también a Sonia M. L. Cyrino, mi co-orientadora del doctorado. Sonia ha acompañado mis años de estudio y formación en el área de sintaxis generativista. Sus comentarios pertinentes a partir de su lectura atenta del proyecto, sus sugerencias bibliográficas, sus clases, su apoyo ante mi primer congreso importante de lingüística me permitieron avanzar con mayor seguridad en el estudio de la sintaxis. Es a ella a quien le debo mi creciente interés en esta área.

A Marcelo E. Coniglio, orientador del doctorado, le agradezco que me haya aceptado como su orientanda. No habría tenido yo la oportudidad de pasar estos cinco años estudiando temas de mi interés si no hubiera sido por él.

Muchas otras personas han sido importantes durante estos años de doctorado. Agradezco las horas de charla y bebida a mis amigxs de Campinas: Edgar, Leandro, Tami, Lau, Emiliano, May, Oscar & flia. A Maria, *in memoriam*. A Fede le agradezco su continua amistad durante más de 25 años. A Mónica, por haberme regalado mi primer libro de gramática categorial, donde toda la idea del proyecto inició, casi 10 años antes. A mis amigxs de las tierras catalanas también agradezco. A Antonio y a Chris, por nuestras charlas de balcón en casa, por nuestra excelente convivencia. A Regi, por nuestras conversaciones hasta la madrugrada. A Serena, por las risas. A María, por haberme abierto su casa; la tranquilidad de Aspa propició el comienzo de escritura de la tesis. A Mercè y a Maite, por mi recuerdo de los Pirineos.

A mi familia le agradezco su constante apoyo a la distancia y por recibir siempre mis avances con orgullo.

Giulia, a ti te agradezco por haber estado conmigo durante todo este año de escritura de la tesis, y por — a pesar de ello — continuar.

Abstract

The present thesis lies at the interface of logic and linguistics; its object of study are control sentences with overt pronouns in Romance languages (European and Brazilian Portuguese, Italian and Spanish). This is a topic that has received considerably more attention on the part of linguists, especially in recent years, than from logicians. Perhaps for this reason, much remains to be understood about these linguistic structures and their underlying logical properties. This thesis seeks to fill the lacunas in the literature — or at least take steps in this direction — by way of addressing a number of issues that have so far been under-explored. To this end we put forward two key questions, one linguistic and the other logical. These are, respectively: What is the syntactic status of the surface pronoun? And: What are the available mechanisms to reuse semantic resources in a contraction-free logical grammar? Accordingly, the thesis is divided into two parts: generative linguistics and categorial grammar. Part I starts by reviewing the recent discussion within the generative literature on infinitive clauses with overt subjects, paying detailed attention to the main accounts in the field. Part II does the same on the logical grammar front, addressing in particular the issues of control and of anaphoric pronouns. Ultimately, the leading accounts from both camps will be found wanting. The closing chapter of each of Part I and Part II will thus put forward alternative candidates, that we contend are more successful than their predecessors. More specifically, in Part I we offer a linguistic account along the lines of Landau's T/Agr theory of control. In Part II we present two alternative categorial accounts: one based on Combinatory Categorial Grammar, the other on Type-Logical Grammar. Each of these accounts offers an improved, more fine-grained perspective on control infinitives featuring overt pronominal subjects. Finally, we include an Appendix in which our type-logical proposal is implemented in a categorial parser/theorem-prover.

Keywords: Control (Linguistics); Romance Languages - Pronoun; Anaphora (Linguistics); Categorial grammar; Generative grammar

Resumo

A presente tese situa-se na interface da lógica e da linguística; o seu objeto de estudo são os pronomes lexicais em sentenças de controle em três línguas Românicas: Português, Italiano e Espanhol. Esse assunto tem recebido mais atenção na linguística gerativa, especialmente nos anos recentes, do que na gramática de cunho lógico. Talvez como consequência disso, há ainda muito a ser entendido sobre essas estruturas linguísticas e as suas propriedades lógicas. Essa tese tenta preencher as lacunas na literatura — ou, pelo menos, avançar nessa direção — colocando questões que não foram suficientemente exploradas até agora. Para tal efeito avançamos duas perguntas-chaves, uma linguística e a outra lógica. Elas são, respectivamente: Qual é o estatuto sintático dos pronomes lexicais em estruturas de controle? E: Quais são os mecanismos disponíveis, em uma gramática lógica livre de contração, para se reusar recursos semânticos? A tese divide-se, consequentemente, em duas partes: linguística gerativa e gramática categorial. Na Parte I revisamos algumas das principais teorias de controle gerativistas e a recente discussão acerca das cláusulas infinitivas com sujeito lexical. Na Parte II revisamos a literatura categorial, atendendo principalmente às propostas acerca das estruturas de controle e dos pronomes anafóricos. Em última instância, mostraremos que as propostas linguísticas e lógicas prévias precisam ser modificadas para se explicar o fenômeno linguístico em questão. Com efeito, nos capítulos finais de cada uma das partes avançamos propostas alternativas que, a nosso ver, resultam mais adequadas que as suas rivais. Mais específicamente, na Parte I avançamos uma proposta linguística na linha do cálculo de controle T/Agr de Landau. Na Parte II apresentamos duas propostas categoriais, uma na linha do cálculo categorial combinatório e a outra, na gramática lógica de tipos. Finalmente mostramos a implementação da última proposta em um analisador sintático e de demonstração categorial (*categorial parser/theorem-prover*).

Palavras-chaves: Controle (Linguística); Línguas Românicas - Pronome; Anáfora (Linguística); Gramática categorial; Gramática gerativa

List of Figures

Figure 5.1 – Cancellation schema . . . 87
Figure 5.2 – Rules of the **AB** system . . . 88
Figure 5.3 – Structural rules for the Classical **LK** sequent logic . . . 89
Figure 5.4 – Associativity left rules . . . 90
Figure 5.5 – Binary trees and Associativity . . . 91
Figure 5.6 – Rules of the (associative) **L** system . . . 95
Figure 5.7 – Labelled sequent calculus for **L** . . . 103
Figure 5.8 – Labelled derivation for *John gives Mary the prize* in **L** . . . 104
Figure 5.9 – Labelled Contraction, Weakening and Expansion rules . . . 105
Figure 5.10 – Expansion and Weakening, Empty Abstraction and Combinator **K** . . . 105
Figure 5.11 – Contraction, Multiple Binding and Combinator **W** . . . 105
Figure 5.12 – Derivation for object relativization in CCG . . . 108
Figure 5.13 – Rules of Combinatory Categorial Grammars . . . 109
Figure 6.1 – Derivation for: *John tries to leave* in **LJ** . . . 113
Figure 6.2 – Derivation for: *John persuades Bill to leave* in **LJ** . . . 114
Figure 6.3 – Equi-verb rule . . . 115
Figure 6.4 – Labelled derivation of Equi in **LJ** . . . 115
Figure 6.5 – Derivation for: *John wants to leave* in $\mathbf{S}_W$ . . . 115
Figure 6.6 – Derivation for: *John persuades Bill to leave* in $\mathbf{S}_W$ . . . 117
Figure 6.7 – Analysis trees for: *John persuades Bill to leave* in Bach's proposal . . . 122
Figure 7.1 – Derivation for: *John admires himself* in Szabolcsi's proposal . . . 132
Figure 7.2 – Derivation for: *John*$_1$ *presents Bill himself*$_1$ in Szabolcsi's proposal . . . 133
Figure 7.3 – Derivation for: **John presents himself*$_1$ *Bill*$_1$ in Szabolcsi's proposal . . . 133
Figure 7.4 – Derivation for non-peripheral subject-oriented reflexives in Morrill's proposal . . . 134
Figure 7.5 – (Forward, Backward and Generalized) G Rules . . . 136
Figure 7.6 – (Forward, Backward and Mixed) Z Rules . . . 137
Figure 7.7 – Derivation for: *He lost* in $\mathbf{J}_{GZ}$ . . . 137
Figure 7.8 – Derivation for: *Mary admires him* in $\mathbf{J}_{GZ}$ . . . 137
Figure 7.9 – Derivation for: *John*$_1$ *said he*$_2$ *lost* in $\mathbf{J}_{GZ}$ . . . 137
Figure 7.10 – Derivation for: *She admires him* in $\mathbf{J}_{GZ}$ . . . 138
Figure 7.11 – Derivation for: *John*$_1$ *said he*$_1$ *lost* in $\mathbf{J}_{GZ}$ . . . 139
Figure 7.12 – Derivation for: *John*$_1$ *admires his*$_1$ *mother* in $\mathbf{J}_{GZ}$ . . . 139
Figure 7.13 – Logical rules of **LA** . . . 140
Figure 7.14 – Structural rules of **LA** . . . 141
Figure 7.15 – Derivation for: *John*$_1$ *admires himself*$_1$ in **LA** . . . 142

Figure 7.16–Derivation for: *John*$_1$ *says he*$_1$ *runs* in **LA** 142
Figure 7.17–Derivation for: *John*$_1$ *says he*$_2$ *runs* in **LA** 142
Figure 7.18–Derivation for: *John*$_1$ *washes his*$_2$ *car* in **LA** 143
Figure 7.19–Derivation for: *He*$_1$ *drinks and he*$_1$ *smokes* in **LA** 144
Figure 7.20–Derivation for: **John*$_1$ *says Peter*$_2$ *admires himself*$_1$ in **LA** 144
Figure 7.21–Derivation for: **The man*$_1$ *says Peter*$_2$ *admires him*$_2$ in **LA** 145
Figure 7.22–Logical rules of **LLC** 148
Figure 7.23–Derivation of $|L$, where $Y = A$, in **LA** and in **LLC** 149
Figure 7.24–Splitting $|R$, where $n = 1$ 150
Figure 7.25–Derivations of $|R$, when $n = 1$ in **LA** and in **LJ** 150
Figure 7.26–Derivation of $|R$ in **LJ** 150
Figure 7.27–Derivation for: *He*$_1$ *drinks and he*$_1$ *smokes* in **LLC** 152
Figure 7.28–Derivation for: *He*$_1$ *drinks and she*$_2$ *smokes* in **LLC** 152
Figure 7.29–Derivation for: *He*$_1$ *smokes and John*$_2$ *hates him*$_1$ in **LLC** 152
Figure 7.30–Derivation for: **He*$_1$ *hates him*$_1$ in **LLC** 152
Figure 7.31–Illicit Derivation for $\mathbf{Z} >$ and $\mathbf{Z} \lessgtr$ in **LLC** 153
Figure 8.1 – Derivation for: *Julia quería telefonear* in $\mathrm{S}_{\mathbf{W}}$ 158
Figure 8.2 – Derivation for: *Julia quería telefonear ella* in $\mathbf{J}_{GZ}$ 159
Figure 8.3 – Schematic derivation for: *Julia prometió a Marta encargarse ella del asunto* in $\mathbf{J}_{GZ}$ 160
Figure 8.4 – Schematic Derivation for: **O menino quer Maria namorar ele* in $\mathbf{J}_{GZ}$. 162
Figure 8.5 – (Backward and Generalized) G rules of $\mathbf{J}_{G^2Z}$ 163
Figure 8.6 – Free subject and object pronouns in $\mathbf{J}_{G^2Z}$ 164
Figure 8.7 – (Forward and Backward) G rules of $\mathbf{J}_{G^4Z}$ 165
Figure 8.8 – Derivation for: *Julia prometió a Marta encargarse ella del asunto* en $\mathbf{J}_{G^4Z}$ 166
Figure 8.9 – Derivation for: *Gianni ha deciso di intervenire lui* in $\mathbf{J}_{G^4Z}$ 166
Figure 8.10–Derivation for: *A polícia forçou os manifestantes a eles saírem* in $\mathbf{J}_{G^4Z}$ 166
Figure 8.11–(Alternative) Left rule of **LLC** 169
Figure 8.12–Simplified left rule of **LLC** 169
Figure 8.13–Left rule for bound pronouns of $\mathbf{LLC}^2$ 170
Figure 8.14–Left rule for (subject-oriented) reflexives of $\mathbf{LLC}^2$ 170
Figure 8.15–(Schematic) Derivation for: *The father of John*$_1$ *admires him*$_1$ in $\mathbf{LLC}^2$ 170
Figure 8.16–(Schematic) Derivation for: *[The father of John]*$_1$ *admires himself*$_1$ in $\mathbf{LLC}^2$ 170
Figure 8.17–(Illicit) Derivation for: *He*$_1$ *admires himself*$_1$*/*(him*$_1$*)* in $\mathbf{LLC}^2$ 171
Figure 8.18–Rules for the semantically inactive disjunction type-constructor $\sqcup$ . . . 172
Figure 8.19–Derivation for: *Os meninos querem eles embrulhar(em) os presentes* in $\mathbf{LLC}^2$ 173
Figure 8.20–Left rules for long-distance $\upharpoonright$ and discontinuous $\downarrow$ type-constructors . . 174

Figure 8.21–Derivations for the VSO and SVO word-order in $\mathbf{LLC}^2$ 174
Figure 8.22–Derivation for the VOS word-order in $\mathbf{LLC}^2$ 175
Figure 8.23–Rules for bracket modalities $\langle\ \rangle$ and $\lceil\ \rceil$ 176
Figure 8.24–Illicit derivations for the VOS word-order in $\mathbf{LLC}^{\langle 2\rangle}$ 176
Figure 8.25–Derivation for an infinitive nominal in $\mathbf{LLC}^{\langle 2\rangle}$ 177
Figure 8.26–Illicit derivation for pronouns in embedded infinitive domains in $\mathbf{LLC}^{\langle 2\rangle}$ 178

List of Tables

Table 2.1 – Possible combinations for [T, Agr] features on the C head 42
Table 5.1 – Combinators and λ-terms 101
Table 5.2 – Combinators, λ-terms, (logical and structural) rules and **LJ**-sequents . . 106

Contents

1 Introduction **15**

1.1 A Semantic-Syntactic Mismatch in Control Clauses 15

1.2 A New Challenge: Control Infinitives with Overt Pronouns 18

1.3 Outline and Proposal 20

I Generative Linguistics **23**

2 The Syntax of Control from a Generative Perspective **24**

2.1 Introduction 24

2.2 Selected Theories of Control: Deletion, Binding, Movement and Agreement 26

2.2.1 Equi-NP Deletion 26

2.2.2 Control in Government and Binding Theory 27

2.2.3 PRO in the Minimalist Program 31

2.2.4 Control as Movement 34

2.2.5 Landau's T/Agr Calculus 36

2.3 Summing Up 47

3 Overt Subjects in Infinitive Clauses from Romance Languages **48**

3.1 Inflected and Personal Infinitives 48

3.2 Null Subjects in (Portuguese) Inflected Infinitives 50

3.2.1 Case and Reference for *pro* and PRO 50

3.2.2 Null and Overt Subjects in Complement Clauses 54

3.3 Overt Subjects in (Spanish and Italian) Personal Infinitives 57

4 Proposal: Overt Infinitive Subjects in Generative Linguistics **70**

4.1 F(ree)-Infinitives 72

4.1.1 Inflected F-Infinitives 76

4.2 C(ontrolled)-Infinitives 80

4.3 Conclusions 84

II Categorial Grammars **86**

5 Lambek Calculus **87**

5.1 Introduction: Logical Grammar 87

5.2 Lambek Calculus 93

5.2.1 Syntax 93

5.2.2 Semantics 98

5.3 (Typed) λ-Calculus 99

5.4 Labelled Lambek Calculus 102

5.5 Combinatory and Type-Logical Grammars 105
6 Control in Categorial Grammar 111
6.1 Introduction 111
6.2 Control and the Combinator W 112
6.3 Control Verbs and Wrap 118
6.4 Summing Up 124
7 Anaphoric Pronouns in Categorial Grammar 126
7.1 Introduction 126
7.2 Routes to Treating Anaphoric Pronouns 130
7.2.1 Multiple-Binding in the Lexicon 130
7.2.2 A Bridge between Combinatory and Type-Logical Approaches . . . 135
7.3 On Contraction: Duplication in the Syntax 139
7.3.1 **LA** Calculus: Multi-structured Sequents and Modal Contraction . . 139
7.3.2 **LLC** Calculus: Encoding Limited Contraction 146
7.4 Summing Up 153
8 Proposal: Overt Anaphoric Control in Categorial Grammars 154
8.1 Overt Controlled Subjects 154
8.1.1 Anaphoric Pronouns or Anaphoric Verbs 157
8.2 A Combinatory Proposal 158
8.2.1 Preliminary Proposal: $\mathbf{J}_{\mathrm{GZ}}$ 158
8.2.2 Modifiying $\mathbf{J}_{\mathrm{GZ}}$: $\mathbf{J}_{\mathrm{G^2Z}}$ and $\mathbf{J}_{\mathrm{G^4Z}}$ 162
8.2.2.1 Final Remarks 167
8.3 A Type-Logical Proposal 168
8.3.1 Modifying **LLC**: **LLC**^{2} 168
8.3.2 Anaphoric Control in $\mathbf{LLC^2}$ 171
9 Conclusions 180

Bibliography 183
Appendix A. Overt Pronouns in Control Infinitives 201
Appendix B. Proofs in CatLog2 204

1 Introduction

1.1 A Semantic-Syntactic Mismatch in Control Clauses

Control sentences, such as (1) and (2) below, represent a challenge for most theories of grammar. This is because, as we shall see, they present an apparent mismatch between surface syntax (and phonology) and semantics.

(1) John promised Peter [to make him a tuna sandwich].

(2) Someone forced Mary [to do it].

A control sentence is a complex sentence containing two clauses: the matrix clause and the embedded clause. The matrix clause contains a finite control verb, such as *promise*, *force*, *try*, *persuade*, *order*, *condescend*, *want*, *propose*, *suggest*, *advise*. The embedded clause is selected by the matrix verb and contains a verb in infinitive form.[1] The infinitive complement clause is said to be *controlled* by the matrix clause because the semantic, phonologically null, subject of the former is understood as being identical with the nominal — subject or object — argument of the latter. Sentences of the first kind are called subject control clauses; those of the second, object control clauses.

Sentence (1) above is a textbook example of the first kind. We have a finite matrix clause (*John promised Peter*) and an embedded infinitive clause (*to make him a tuna sandwich*); the latter is selected by the matrix verb *promised*. Most importantly, the syntactic subject of the matrix clause — *John* — also appears to be the semantic subject of the infinitive, *despite its absence at the surface level*; i.e., despite being phonologically null. In both subject and object control clauses, the matrix argument that is used as antecedent for (or to ascribe reference to) the subject of the infinitive clause, is called the *controller*. In the examples above, this role is fulfilled by *John* and *Mary*, respectively. The covert (or unexpressed) subject of the infinitive complement clause is called the *controllee* (or: controlled). Thus, the controllee is understood as being coreferential with the controller.

Since the semantic value of the embedded subject depends on the reference of the matrix nominal, the controller has a double semantic function: it is interpreted as an argument of the matrix predicate and also as the external argument of the infinitive embedded verb phrase.[2] Put differently: since the semantic embedded subject does not have a surface representation, a single surface syntactic constituent is interpreted in two

[1] We signal the infinitive clause by placing it inside square brackets.

[2] The syntactic subject of an (active) clause is the external argument of the verb phrase; the syntactic object is selected by (or is a complement of) the verb and is an internal argument of the same.

different argument positions: namely, the controller and controllee of the control sentence. Hence, a mismatch between the semantic and surface syntactic levels is generated.

Indeed, there is no lexical item in the embedded subject position which could provide the semantic value of the subject of the embedded predicate. Assuming that the linguistic surface material constitutes the entirety of our resources, the mismatch displayed by a control sentence can be seen either as a deficit of syntactic resources or, conversely, as a surplus of semantic resources.

At least two options are available in order to dissolve this mismatch between syntax and semantics, and balance deficit against surplus. A theory of grammar could avoid this surplus of semantic resources by claiming that only one semantic resource is consumed in a control sentence, as opposed to two. The upshot: no embedded semantic subject resource is consumed by the embedded verb phrase. There are then two ways to further develop this first approach: either hold that the embedded verb phrase does not denote a propositional function (i.e. a function from entities to propositions); or, hold that the embedded predicate does denote a function, but is also semantically unsaturated. Importantly, both routes are compatible with the so-called Property Theory, according to which the denotation of a controlled infinitive clause is a property. For, it is widely assumed that a property is a semantic correlate of a predicate — and therefore, an unsaturated structure: that is, a (propositional) function from the set of entities into the set of truth-values (but see [Chi85] for properties as individuals). Regardless of which concept of property is adopted, both options agree on one key fact: no (subject) semantic resource is consumed by the embedded controlled clause. Crucially, the control relationship may still be guaranteed under the Property Theory, by means of a semantic rule or meaning postulates. The semantic-syntactic mismatch may thus be dissolved.

Conversely, a theory of grammar could avoid the syntactic deficit by assuming there is not just one syntactic resource, but two: a surface resource and a non-surface resource. This second approach to resolving the mismatch assumes that the semantic reading is — completely or partially — obtained from a level of syntactic representation other than the surface level. This approach is therefore available only to those grammars that admit different levels of syntactic representation. Such grammars can postulate an embedded subject, coreferential with the controller matrix nominal, in some non-phonological level of syntactic representation; i.e. a phonologically null subject. The semantic resource supplied by this non-surface syntactic subject could then be consumed by the embedded predicate. Thus, a grammar of this kind could endorse the so-called Propositional Theory, according to which the denotation of the control infinitive clause is a proposition. Once again, the semantic-syntactic mismatch is dissolved.

Having sketched the general ideas behind the two approaches in question, let us now add some flesh to the bone, and introduce the two frameworks that are the subject of this thesis: Categorial Grammars and Generative Grammars.

One of the Categorial Grammars we will be working with is the so-called Type-Logical grammar. This is a formal Categorial Grammar based on the Lambek **L** Calculus [Lam58]. **L** is a substructural logic, that is, a logic that rejects some of the structural rules of the Classic Sequent Calculus **LK**: Weakening (and so also Expansion), Contraction, and Permutation (or: Exchange). The sequent calculus for **L** has no structural rules; **L** only contains the Identity axiom and a restricted version of the Intuitionistic Cut rule.[3] Since they lack the Weakening and Contraction rules, Categorial Grammars are resource-conscious grammars; the linguistic analyses they provide are resource-conscious proofs. In a nutshell, in a resource-conscious proof all formulae occurring in the proof have to be used at least once and cannot be reused. Moreover, a Categorial Grammar is a monostratal framework. Grammars of this kind implicitly or explicitly adopt what is called the *hypothesis of direct surface compositionality*: the compositional semantics directly assigns each (and only) surface syntactic expressions a model-theoretic interpretation (cf. [Jac96a]; [BJ07]). Categorial Grammars only consider the resources that occur on the surface level of representation as the input of the compositional semantics. Consequently, resorting to a non-surface level to avoid the apparent deficit/surplus of resources in control clauses is not available for grammars of this kind. Therefore from a categorial point of view, since embedded infinitive clauses in control structures do not contain a syntactic subject, they do not contain a semantic subject either. Instead, infinitive clauses in control structures are syntactically and semantically subjectless clauses; they denote a property, not a proposition (cf. [Dow85]).

In contrast, Generative Grammars assume that the semantic reading of a clause is obtained not only or directly from the surface level of syntactic representation. In fact, Generative Grammars generally admit more than one level of syntactic representation: deep structure or logical form on the one hand, and surface structure on the other, for example. Thus, grammars of this kind can avoid the syntactic-semantic mismatch displayed by control structures via the second route we sketched earlier: that is, by taking some form of a non-phonological nominal constituent as the syntactic subject of the infinitive embedded clause. In some generative theories it is assumed that a deleted copy of the overt matrix — subject or object — controller occupies the embedded subject position in some non-phonological level of syntactic representation (cf. for example [Ros65]; [Hor99]); others assume that it is the (phonologically) null subject PRO which occupies the infinitive subject position (cf., for example, [Cho81], [Cho93]; [Mar01]). Whichever the non-surface nominal category chosen, these proposals assume that the infinitive controlled clause does have a syntactic subject. Assuming that the syntactic subject supplies a denotation saturating the embedded predicate, most generative theories contend that the embedded

[3] See Figure 5.3 on page 89 for the structural rules of **LK** and Figure 5.6 on page 95 for the rules of the **L** system.

clause in control structures denotes a proposition (but see [Lan15]).

In sum, both Categorial and Generative Grammars can circumvent the semantic-syntactic mismatch apparently displayed by prototypical control structures. We now turn to the non-prototypical instances of these structures, and thereby introduce the central topic of this thesis.

1.2 A New Challenge: Control Infinitives with Overt Pronouns

Whether we adopt the property approach as in a Categorial Grammar, or the propositional approach as in most Generative Grammars, the semantic-syntactic mismatch threatens to reappear in a special kind of control sentences of certain *pro*-drop languages.[4] The occurrence of overt, semantically controlled, focused pronouns within the (inflected or uninflected) infinitive clause in languages such as European and Brazilian Portuguese (EP/BP), Spanish (ES) and Italian (IT) causes once again both kinds of grammars to face a mismatch. Examples of this phenomenon are given in the sentences in (3)–(7) below. For future reference, note that sentences of this kind will be our main topic of linguistic and logical research (for more examples, see Appendix 9).[5]

(3) Eu exigi aos alunos [**ELES** fazer**em** um trabalho]. (EP)
I forced to-the students [they.NOM do.INFL.3PL a work]
'I forced the students to do their homework.' [Zwa90]

(4) O João decidiu [resolver **ELE** o problema]. (EP)
the John decided solve.INF.(3SG) the problem
'John decided to solve the problem himself.' [Bar10]

(5) Pedro quer [**ELE** chegar cedo]. (BP)
Peter wants [he.NOM arrive.INF early]
'Peter wants to arrive early.' [GM15]

(6) Juan prometió a su profesor [hacer **EL** los deberes]. (ES)
John promised to his teacher [do.INF he.NOM the homework]
'John promised his teacher to do the homework by himself.' [Her11]

(7) Gianni ha deciso [di intervenire **LUI**]. (IT)
John has decided [COMP intervene.INF he.NOM]
'John has decided to intervene himself.' [Car99]

As we will soon see, sentences like the above call for new responses to old questions. For, in contrast with typical control sentences, these examples contain an overt (or: lexical)

[4]In very simple terms, a *pro*-drop language is a language that allows for referential pronouns to be omitted (or be phonologically null). See Chapter 3 for more details about *pro*-drop languages.

[5]The infinitive subjects and the inflection on the infinitive verb are marked in bold. Focused pronouns are marked in uppercase as standard. In what follows, however, we will drop the use of the uppercase for readability.

pronoun in the pre- or post-verbal position within the complement clause. The pronoun in the infinitive clause is semantically controlled by a nominal expression in the matrix clause. In other terms, this pronoun necessarily acquires its reference through a relation with a — subject or object — matrix argument. Whichever the mechanism (variable binding, coreference or meaning postulates, for example) by which the overt pronoun obtains its semantic value, will depend on the theoretical framework adopted; in particular, it will depend on which theoretical assumptions about the pronoun and the infinitive clause are in place. In any case, it is clear that the reference of the matrix controller is in some way (re)used in order for the pronoun to acquire its semantic value.

This kind of non-prototypical control sentence has become quite relevant in the recent generative literature because it challenges several of its basic assumptions, and because it presents a conflict with specific theories on control. Indeed, non-prototypical control sentences are an ongoing topic of research, as they raise a number of worries. For instance: why do *pro*-drop languages admit an overt pronoun within an infinitive clause, while non-*pro*-drop languages do not? If PRO were assumed as the syntactic subject of the infinitive clauses, what would be the syntactic role of the overt pronoun in these infinitive controlled clauses? Is the pronoun the genuine syntactic subject of the infinitive clause? If the infinitive subject were a deleted copy of the matrix nominal controller in typical control sentences, why does the non-deleted copy have a pronominal form in these non-prototypical control structures? In other terms, what would be the mechanism through which a nominal expression in a non-phonological level of syntactic representation acquires a pronominal form in the surface level? Could there be expressions other than pronominals in these same syntactic contexts?

The linguistic data presented above — which has not received categorial treatment so far — also poses a number of semantic and technical worries for Categorial Grammars. Namely, is the infinitive phrase containing a pronoun, a sentence? If so, does the sentence necessarily denote a proposition or could it denote a property, like other infinitives? In other terms, does the pronoun saturate the external argument position of the infinitive verb phrase? Answers to these questions depend, in turn, on other semantic issues: What is the denotation or the semantic value of a pronoun in general, and within this kind of control sentences in particular? Are they instances of anaphoric pronouns, when occurring in control sentences? Does an anaphoric pronoun denote an individual? These semantic questions in turn raise some old technical, logical questions: How can we deal with anaphoric expressions from a categorial perspective? How can we address the problem of reusing a semantic resource in a Categorial Grammar? Certainly, the linguistic phenomenon of anaphora presents a serious challenge for this kind of grammar, in that the antecedent resource is multiplied in the semantics but not also in the syntax. This duplication, which corresponds to the structural rule of Contraction, turns out to be particularly troubling for a resource-conscious grammar such as a Categorial Grammar.

In sum: a Generative Grammar can overcome the semantic-syntactic mismatch displayed by typical control sentences by assuming that there is a syntactic subject in some non-phonological level of representation. When faced with non-prototypical control sentences, however, it must first explain: What is the syntactic status of the surface pronoun? This is the **key linguistic question** that will concern us in this thesis. The more straightforward answer — that these overt pronouns are the genuine subjects of the infinitive clause — goes against traditional generative assumptions. Similarly, though a Categorial Grammar can avoid the surplus-deficit of resources displayed by typical control sentences assuming the Property Theory, it is confronted with an old challenge in a new context: reusing semantic resources in control sentences with overt anaphoric pronouns. This in turn motivates the **key logical question** we shall subsequently be addressing: What are the available mechanisms for reusing semantic resources in a contraction-free logical grammar?

1.3 Outline and Proposal

This doctoral thesis addresses both the linguistic and the logical challenges posed by the non-prototypical control sentences that are the topic of our research: that is, **sentences with an overt anaphoric pronoun in the** (apparent) **subject position of the infinitive complement clause**. The thesis will be divided into two parts, addressing each of the above key questions in turn.

In Part I of the thesis — **Generative Linguistics** — we critically examine the basic assumptions of Generative Grammar with respect to the phenomenon of control and the occurrence of overt (referentially free or controlled) subjects in infinitive complements. After reviewing and dismantling the theoretical assumptions that cause problems for control sentences with overt pronouns in such a framework, we argue for our own **linguistic thesis**. In a nutshell, we claim that the more straightforward answer to the key linguistic question is in fact both admissible and plausible: **overt anaphoric pronouns in control sentences from Romance languages are the real syntactic-semantic subject arguments of the infinitive clause**.

Part I in turn is divided into three chapters. In Chapter 2 — The Syntax of Control from a Generative Perspective — we review different generative theories about control structures and about the phonologically null controlled subject: Equi NP-Deletion [Ros65], Government and Binding [Cho81], [Cho93], the Minimalist Program [Cho95], the Movement Theory of Control as Movement [Hor99] and the T/Agr Calculus [Lan00], [Lan04], [Lan06]. We highlight how — by putting aside certain generative assumptions — these last two theories on control clear the way for admitting overt forms of infinitive subjects in a generative framework. The T/Agr calculus will deserve special attention, as we shall later adopt some of its assumptions in our own linguistic account.

Chapter 3 — Overt Subjects in Infinitive Clauses from Romance Languages — is guided by the question of which kind of null pronoun — PRO or *pro* — is allowed in the subject position of inflected infinitives from Portuguese. We also review the extant discussion about whether a null pronoun can alternate with an overt (or lexical) Nominal/Determiner Phrase in the subject position of an infinitive clause from Romance languages. To this end, we first characterize the null subject in finite contexts — *pro* — in Portuguese and in other *pro*-drop Romance languages. Subsequently, we review a few proposals about control and overt subjects in non-finite clauses from Portuguese, Spanish and Italian.

In Chapter 4 — Proposal: Overt Subjects from a Generative Perspective — we present our own linguistic thesis about the status of overt pronouns within infinitive complements of control structures and, in general, about the occurrence of overt subjects in infinitive complements of Portuguese, Italian and Spanish. We show how our proposal can meet the challenges faced by other accounts.

In Part II — **Categorial Grammars** — we adopt a logical perspective; we address the problem of reusing semantic resources in a Categorial Grammar. This problem is separately posed by control verbs and anaphoric pronouns, and jointly posed by anaphoric pronouns in control structures. In response, we present a somewhat different logical account. In brief, the idea is to use some of the categorial tools provided by Combinatory and Type-Logical approaches to anaphoric pronouns and control verbs or, in more general terms, some of the logical tools proposed in the categorial framework to trigger duplication of semantic resources.

Hence, firstly, in Chapter 5 — Lambek Calculus — we review the Lambek **L** system as well as its predecessor, the **AB** system [Ajd35]; [BH53]. In so doing, we focus on the complete lack of structural rules in **L**; specifically, we emphasize the consequences of the rejection of the Contraction rule in this system. We also present some basic concepts of the Typed λ-calculus, which will be necessary to understand the labelled Lambek system. In sum, this chapter sets things up for the work to be carried out in the subsequent chapters.

Chapter 6 — Control in Categorial Grammar — reviews some categorial proposals on control verbs. We focus on two concerns. Firstly, we look at the problem of ensuring the control relationship in logical grammars suffering from the lack of Contraction. We present two combinatory categorial proposals due to Steedman: a syntactic approach [Ste88] and a lexical approach [Ste94], [SB11]. Secondly, we concentrate on the issue of explaining the different behaviour of three-place control verbs. We review Bach's [Bac79] distinction between subject and object control verbs, and his treatment of the latter in terms of a discontinuous prosodic operation.

Chapter 7 — Anaphora in Categorial Grammar — turns its attention to some of the categorial systems developed to account for the linguistic phenomenon of anaphoric pronouns. We critically review some of these systems in order to determine which, if any, can be used to solve the problem posed by overt anaphoric pronouns in control clauses. In general terms, two kinds or lines of response have been developed to solve the problem of reusing resources in a Categorial perspective. We will see that some proposals treat multiple-binding in the lexicon; certain lexical items are then identified as responsible for resource multiplication. Others use a syntactic approach: they extend the **AB** or **L** system in order to allow a syntactic-semantic operation to do the work of reusing lexical material. In light of the above, we examine two families of accounts for treating anaphoric pronouns in a Categorial framework: a lexical [Sza89], [Sza92] and a syntactic approach [Jac99] in Combinatory Categorial Grammar; and a lexical [MV10a] and also a syntactic approach [Jäg05] in Type-Logical Grammar. Two Type-Logical systems in particular will deserve our attention: the Lambek system with Anaphora **LA** [Jäg97], [Jäg98] and the Lambek system with Limited Contraction **LLC** [Jäg05]. Both systems extend **L** with restricted forms of Contraction, and so they appear to be good candidates to adequately handle the phenomenon of reusing lexical resources in control sentences.

In Chapter 8 — Proposal: Overt Anaphoric Control in Categorial Grammars — we show how we can deal with overt controlled pronouns in a categorial framework. After summarizing the most important findings of our linguistic research, we offer two accounts: one Combinatory and the other Type-Logical. The first will be based on Steedman and Jacobson's approaches; the second, on Jaeger's system **LLC**. In a nutshell, our two logical proposals will be guided by the following linguistic considerations: i) overt nominals in the subject position of the infinitive clause are free pronouns; ii) the mechanism for duplicating the value of the controller will be attached to the control verb, rather than to the pronoun itself; iii) control is a semantic relationship that can be triggered by either the lexical or the derivational semantics of the control verb; iv) embedded clauses in control structures denote a property — an unsaturated function — in spite of the overt occurrence of a pronoun. In sum, this chapter shows how our two accounts, produced by way of answering the key logical question, might address the challenge posed by the new semantic-syntactic mismatch.

Finally, Chapter 9 reviews and sums up the discussion from both Part I and II, and draws our conclusions from the same.

Closing the thesis are two Appendices. Appendix 9 displays several instances of the kind of constructions that are the topic of our research. Appendix 9 displays the implementation of our logical analysis provided in CatLog2, which is a Prolog parser/theorem-prover for a (type) logical (categorial) grammar. The CatLog program series is currently being developed by G. Morrill; the web interface for CatLog2 — available in `www.cs.upc.edu/~droman/index.php` — was developed by D. Roman (cf. [Mor12], [Mor17b]).

Part I

Generative Linguistics

2 The Syntax of Control from a Generative Perspective

2.1 Introduction

Obligatory control, according to the Generative Grammar orthodoxy, is a semantic phenomenon occurring in specific syntactic configurations: namely, where an infinitive clause is selected as the complement of the main verb of a complex clause. Verbs such as *permit*, *want*, *expect*, *persuade*, *condescend*, *try*, *propose*, *promise* are prototypical verbs that select an infinitive clause; they are prototypical instances of the so-called control verbs. Obligatory (or exhaustive) control structures involve a single overt nominal (Determiner Phrase – DP) that appears to have two semantic roles: one as the subject or the object argument of the main verb, and the other as the subject argument of the embedded infinitive clause. In control structures a single overt DP is interpreted in two semantic positions, even while it occupies only one surface syntactic position — in the matrix clause. The following sentences (8) and (9) exhibit the phenomenon of subject and object obligatory control, respectively:

(8) The doctor condescended [to examine John].

(9) Barnett persuaded the doctor [to examine Tilman].

The sentence in (8) is an example of subject control in that the matrix subject DP — *the doctor* — plays the role of the controller; it is also interpreted as the semantic subject of the infinitive embedded clause, that is, as the controllee, due to the meaning of the main verb *condescend*. Sentence (9) is an example of object control; it is the object matrix DP *the doctor* that works as the controller, and thus gives the (phonologically null) infinitive subject its reference.

Generative Grammars are multi-stratal grammars, as they assume more than one level of linguistic representation. Hence, generative theories can postulate that although the controllee lacks a representation at the syntactic surface level, and thus is phonologically null, there is at least some non-phonological level where it has a syntactic form. Thus, the null semantic subject of the infinitive verb also has a syntactic (non-surface) representation. Several generative theories have been proposed in order to explain the control relation. Among such theories the key distinguishing factors are, on the one hand, the theoretical category to which the controlled is assigned; on the other, which syntactic mechanism licenses the control relation. A further difference between control theories may result from the very definition of obligatory control (cf. [Wur02]).

Types of control

There is a broad consensus among scholars that subject and object control are subclasses of the subtype of obligatory exhaustive control. That which is called obligatory exhaustive control is a type of semantic or lexical control in which the semantic content of the null infinitive subject coincides with the denotation of a single DP in a particular syntactic position in the matrix sentence. The null infinitive subject obtains its total semantic value from this matrix DP.[6] Obligatory control is induced by predicates exhibiting semantic or lexical control: intentional, desired, directive and manipulative predicates. There are a number of verbs that trigger exhaustive subject control: for instance, *attempt*, *condescend*, *desire*, *hope*, *intend*, *promise*, *refuse*, *try*; others give rise to exhaustive object control, such as *accuse*, *allow*, *cause*, *coax*, *convince*, *encourage*, *forbid*, *force*, *let*, *order*, *permit*, *persuade*, *prevent*, *request*, *tell*, *urge* (cf. [DD08]).

Declarative, epistemic and factive verbs can also select an infinitive complement and control the null subject. However, in these cases the control relationship does not arise from the semantics of the main verb, but results from the semantic content and syntactic structure of the complement clause. In general, verbs that give rise to *structural* control display non-exhaustive control (cf. [Sti07]). There is no wide agreement about the type of control displayed by the subtypes of structural control, such as split, variable or shift, partial and implicit control. Some scholars regard them as subclasses of obligatory control (cf. [Lan00] and subsequent work; [Sti07]); others, as subclasses of non-obligatory control (cf. [Wil80]; [Wur02]).[7] In split control both of the matrix arguments jointly control the infinite null subject, which is semantically plural (see example in (11) below).[8] In variable or shift control, the controller can be either one or the other of the matrix arguments (see (12)). In partial control, the controller is semantically plural; the controller includes both the reference of a matrix argument and a contextually determined entity. In partial control structures the controller may be implicit, in that one of the controllers is not syntactically expressed in the matrix clause. Partial control depends on a collective predicate in the embedded controlled clause (see the example in (13)).

(10) Mary$_1$ thought that John$_2$ didn't know where [— $_{1+2}$ to go together]. [Lan00]

(11) Chris$_1$ proposed to his officemate$_2$ [— $_{1+2}$ to introduce themselves$_{1+2}$ to the new boss].

(12) Tommy$_1$ begged his mother$_2$ [— $_{1/2}$ to stand next to her$_2$/him$_1$]. [Lan13]

(13) The chair$_1$ preferred [— $_{1+n}$ to gather at 6].

[6] In *pro*-drop languages, such as Spanish, Italian, European Portuguese and Catalan among others, the subject controller can be null.

[7] Non-obligatory control includes, as other subclasses, long distance control and arbitrary control.

[8] Since the syntactic category assigned to the null subject depends on the control theory being adopted, in these examples we use a line to indicate the subject position and its semantic value.

Summing up, with these other types of control the matrix controller for the null subject does not need to be unique, but at least one matrix controller DP must be syntactically expressed. By contrast, in the case of exhaustive control the controller must be unique and syntactically explicit.

In the remainder of this chapter we briefly describe some of the more representative generative theories on exhaustive control.

2.2 Selected Theories of Control: Deletion, Binding, Movement and Agreement

2.2.1 Equi-NP Deletion

One of the first generative control theories was presented by Rosenbaum [Ros65] in terms of Chomsky's Standard Theory [Cho65]. In the Standard Theory, the surface syntactic structure of a sentence is obtained by applying transformational rules — deleting, insertion of non-lexical items, for example — to the deep structure obtained, in turn, by inserting the lexical items within the constituent structure. In particular, to derive a control structure Rosenbaum uses a special rule called Equivalent Noun Phrase Deletion (Equi-NP Deletion). This rule erases the lower copy of the duplicated nominal phrase (NP), which functions as the syntactic subject in the embedded clause. As an example, observe the schematic derivation of the subject control sentence in (14):

(14) The doctor condescended [to examine John].
a. the doctor condescended [$_S$ [$_{NP}$ **the doctor**] examine John].
b. the doctor condescended [$_S$ for [$_{NP}$ **the doctor**] to examine John].
c. the doctor condescended [$_S$ for to examine John].
d. the doctor condescended [$_S$ to examine John].

The first step — in (14a) — represents the deep structure of (the syntactic surface representation of) the sentence in (14); in (14b) the complementizers *for* and *to* are inserted; (14c) is obtained by applying the Equi-NP deletion rule, which deletes the second occurrence of the NP subject, that is, the duplicated NP that works as the subject of the embedded infinitive complement; finally, (14d) is obtained from (14c) by deleting the complementizer *for*. The derivation of an object control sentence is analogous, the difference being that, in object control, it is the duplicated NP object which is erased by the Equi-NP rule from the embedded subject position.[9]

[9] In passing, we observe that the Equi-NP deletion transformation is an instance of the Contraction rule applied to non-adjacent NPs (or: long-distance Contraction). We have noted that Categorial Grammar does not admit the structural rule of Contraction. Thus, the derivation of control structures proposed by Rosenbaum cannot be simulated in this substructural grammar (see Chapter 6 for a formal proof of control sentences in a calculus that admits Permutation and Contraction).

The semantic representation of a sentence in the Standard Theory is obtained from the deep structure, and thus is prior to the syntactic transformations by which the surface structure is obtained. So, although there is a mismatch between the semantic and the surface syntactic levels of representation in control sentences, there is no such mismatch between the semantic representation and the deep structure. Both semantic and deep structures contain a subject in the controlled clause. Since the semantic representation of a control sentence takes into account the occurrence of the infinitive subject, and since this subject is a denotative nominal phrase, the infinitive controlled clause is a sentence and denotes a proposition. From this theoretical perspective, the control relation is a coreference relation between the NP duplicated subject and the NP matrix controller.

In closing, note that following the formulation of the Extended Standard Theory [Cho73], subsequent accounts abandoned the idea that the subject embedded position of a controlled clause is filled by an NP semantically and syntactically equivalent to a — subject or object — matrix NP; although we will see one exception to this in Section 2.2.4.[10]

2.2.2 Control in Government and Binding Theory

According to both the Government and Binding Theory (GB) [Cho81], [Cho86] and the Minimalist Program (MP) [Cho95], the empty (or phonologically null) category PRO is the subject of the infinitive clause; in particular, therefore, PRO is the subject of the controlled infinitive clause.[11] The rationale for this assumption is given by the so-called Thematic Theory (Θ-Theory), as we will see in a moment. According to the Θ-Theory, all and only the thematic positions — which are semantically determined and specified by corresponding lexical entries — have to be filled in the deep structure by the lexical items from the lexicon (or vocabulary).[12] To see this consider, for example, the sentence in (15) below:

(15) Peter persuaded his friends [to buy a beer].

[10]The analysis of constituents such as *this car* or *my best friend*, etc. as a nominal category headed by a noun, i.e. an NP, was similarly set aside. Instead, phrases of this kind came to be analyzed as determiner phrases (DPs) — the determiner being the head of the phrase.

[11]In all but Exceptional Case Marked (ECM) constructions, in which the non-finite complement clause has a non-null subject, or in raising constructions. In ECM structures, the overt subject of the embedded clause is not nominative-marked, as is usual in other constructions, but accusative-marked. In raising constructions, it is assumed that the null subject is a deleted copy (or a trace) of the matrix subject (See Section 2.2.3 for a more detailed exposition of these constructions.) For the case of ECM constructions consider, for example, the sentence in i.:

i. I believe [**him** to be honest].

[12]The GB framework assumes four levels of representation: deep structure (DS), surface structure (SS), phonological form (PF) and logical form (LF). The architecture of the grammar, and thus, the relations between these levels can be schematized as a T-shaped graph: a derivation begins at the DS, the SS is obtained from the DS, the PF of a sentence is read from its SS, and the PF is separated from the LF, which represents the meaning of the sentence.

The lexical entry for the verb *persuade* specifies that it takes two internal arguments — a DP and a propositional complement — and an external argument — a DP subject. Besides the quantity and syntactic category of the arguments, the lexical entry for the verb states their semantic or thematic role: the target or goal of the action — *his friends* — , the theme, and the agent of the action — *Peter*.[13] Analogously, the lexical entry for the verb *buy* specifies that it takes a single internal argument — the DP *a beer* — with the semantic role of theme; the verb phrase (VP) headed by the verb *buy* assigns the role of agent to its external argument.[14] Now recall that by the Θ-Theory all the thematic positions, even that of the subject of the infinitive embedded clause, have to be filled at the deep structure level by items from the lexicon. This means that thematic positions cannot be saturated by movement of DPs previously inserted at the deep structure level. The movement of constituents only starts when all the thematic positions have been syntactically saturated.[15] Thus, in order to satisfy the Θ-Theory the infinitive clause requires a syntactic subject at this non-surface syntactic level.

As we also noted earlier, the pronominal null category PRO is assumed as the syntactic subject of the infinitive clauses. Despite lacking a semantic value, moreover — since it is a variable — PRO receives an interpretation by way of being co-indexed with a nominal constituent. As a result, PRO is also the semantic subject of the infinitive clause.[16] As an illustration, consider the sentence in (16) below, where PRO is co-indexed with the matrix object — *his friends*.[17] Since the null subject PRO is co-indexed with the object, the nominal phrase *his friends* appears to take two semantic roles: theme (persuaded) and agent (buyer). And yet, although the denotation of the DP controller is

[13]Strictly speaking, the thematic role for the external argument is not assigned by the verb, but by the verbal phrase formed by the verb and its internal arguments.

[14]VP is the maximal projection of the verb V, which is its head. In more general terms, according to the X-bar theory, for a functional category X, XP is the maximal projection of the head X that includes the specifier position [Spec-XP] on the left, and its complements on the right. Roughly schematized: $[_{XP}$ Specifier position $[_{X'}$ head position complement position]]. The X-bar schema are commonly drawn in terms of syntax trees, since these more clearly express the hierarchical relationships between the nodes of the tree such as *dominance*, *sisterhood*. The latter in turn are used to formally define theoretical concepts such as *c-command* and *m-command*, among many others. The ordering of the constituents not only expresses a linear order but also a hierarchical structure.

[15]Although every thematic position is an argument position (A-position), the inverse is not true: the specifier position of the inflectional head, that is [Spec-IP], is an A-position, but not a thematic one. Since [KS88], [KS91] it is standard to assume the so-called *VP-internal subject hypothesis*: the subject originates in the [Spec-VP] position and (usually) then moves to [Spec-IP].

[16]From the GB perspective, variables are syntactic objects. Categorial grammars, on the contrary, assume a variable-free semantics; they do not admit variables as a component of the grammar [Jac99]. We shall return to the role of variables in these frameworks in Section 7.1.

[17]At the syntactic level, the choice of the controller for PRO is made in terms of the Minimal Distance Principle (MDP). This principle expects PRO will be controlled by the c-commanding closest antecedent [Ros65], [Ros70].
c-command:$_{def}$ a node α c-commands a node β if and only if: i) α and β are different nodes; ii) α does not dominate β, iii) β does not dominate α, and iv) the first (i.e. lowest) branching node that dominates α also dominates β. Also the Equi-NP rule satisfies the PMD. The control verb *promise* seems to challenge this principle (cf. [Lar91]).

in fact (re)used to give a semantic value to the syntactic subject PRO, it occupies not two different syntactic positions, but one.

(16) Peter persuaded his friends$_1$ [PRO$_1$ to buy a beer].

We thus see that according to the GB theory, control is a semantic dependence relation resulting from a syntactic — binding — relation between a matrix DP and the phonologically null subject PRO.

Recall now that the empty category PRO is taken to be the subject not only of the controlled clause, but of the infinitive clause in general. Since infinitive clauses can occur in contexts other than controlled complements, PRO can also receive a non-controlled, free reading, as exemplified in (17–19) below.

(17) John wondered how [PRO to win the game].

(18) John thinks that [PRO to eat a bagel] would be fun.

(19) [PRO to vote for oneself] would be a mistake. [Hor03]

Much like in control sentences, in (17) too the infinitive clause is selected by the matrix verb. By contrast, in this context PRO can receive an arbitrary reading. In (18) PRO occurs within an infinitive clause that is the subject of the embedded sentence selected by the propositional verb *thinks*; thus, PRO's interpretation is neither lexically nor syntactically determined. In (19) the interpretation of PRO is arbitrary due to the fact that the infinitive clause is the subject of the sentence.

In sum, in some infinitive contexts PRO receives a specific interpretation that is coreferential with some matrix nominal phrase; there, PRO has to be bound. In others, the interpretation of PRO is not determinate; there, PRO is syntactically free. In these cases, PRO is said to be obligatory controlled or non-(obligatory) controlled, respectively [Wil80]. Because of its varied distribution, PRO is formally characterized by the [+anaphoric, +pronominal] features in GB.[18] This formal characterization in terms of semantic features, together with some well-protected theoretical hypotheses of the GB theory, speak to the most important properties of PRO: non-case-markedness, nullness and complementary distribution with overt DPs.

Now due to its mutually exclusive features, it seems that PRO has to simultaneously satisfy both Principles A and B of Binding Theory, which establishes the distribution

[18]Binding Theory classifies the nominals according to two binary features: [±anaphoric] and [±pronominal]. Thus, a nominal can carry one of the four possible combinations: [+anaphoric, −pronominal], [−anaphoric, −pronominal], [−anaphoric, +pronominal] or [+anaphoric, +pronominal]. A nominal that carries [+anaphoric, −pronominal] features is an anaphora (reflexive or reciprocal), such as *himself* and *each other*; an anaphora satisfies Principle A of Binding Theory. A nominal that carries [−anaphoric, +pronominal] features is a personal pronoun, such as *she* or *him*; personal pronouns respect Principle B of this theory. A nominal with [−anaphoric, −pronominal] features is an R-(eferential) expression, and satisfies Principle C: an R-expression cannot have an antecedent that c-commands it.

of nominals. Principle A stipulates that an anaphora must be bound in its governing category (roughly, it must have a c-commanding local antecedent).[19] Principle B stipulates that a pronoun must be free (i.e. not bound) within its governing category.[20] In order to bypass the contradiction between the two conditions — to be simultaneously bound and free — it is stipulated that PRO lacks a governor and therefore also a governing category.[21]

In GB theory, nominative, accusative, dative, and oblique cases for the DPs are assigned under a government relation by their governing heads at the surface syntactic level. Thus, for example, an object complement is governed by the verb, which is the head of the VP and assigns to this complement the accusative case; the complement of a preposition receives its oblique case from its prepositional head. Both the DPs in complement positions and the DP in a specifier position can be case-marked by their respective heads. The subject of a clause, which occupies the [Spec-IP] at the surface level, receives the nominative case from the finite Inflectional head.[22]

In light of the above and given that PRO, by assumption, lacks a governor, it can only occur in a non-case position. Further, assuming the infinitive Inflectional head is not a governor, the distribution of PRO is determined by and limited to this non-finite context: PRO can only occupy the specifier position of the infinitive head. And, since the Case Filter of the Case Theory requires for all overt DPs in an argument position to be

[19] For an accurate definition of *governing category* see [Hae94].

[20] A pronoun can have *a non-local* (or non c-commanding) *antecedent*.

[21] The assumption that PRO lacks a governor was — surprisingly — called PRO *Theorem*. It is even more striking that scholars like Martin [Mar01] endorse the *logic of the deduction* of the PRO Theorem: "the logic of the deduction is straightforward: if PRO is an anaphora and a pronoun, it must simultaneously be bound and free in its governing category. The only way for PRO to meet these contradictory requirements is to not have a governing category. An auxiliary assumption ensures that the only way to not have a governing category is to be ungoverned." As is clear — and contrary to Martin's claim — there is no way to avoid the contradiction resulting from the features attached to PRO and Principles A and B, if we assume that having a governing category is a *necessary condition* for an expression to be an anaphora or to be a pronoun. The best we can do, *after deriving the contradiction*, is to try to overcome it by using the logical Principle of Explosion (PE). Nevertheless, if we recognize the contradiction, then by PE, we can prove not only the statement that PRO is ungoverned, but we can derive any other statement as well. Obviously, this is an undesirable result for any theory. To derive the so-called PRO Theorem without using the Principle of Explosion we have to assume, not only that the governing categories for an anaphora and a pronoun are similar, but also that having a governing category is not a necessary but a *sufficient* condition. In other terms, we have to assume the following premises: i) if α is an anaphora, and if α has a governing category γ, then α has to be bound in γ; ii) if β is a pronoun, and if β has a governing category γ, then β has to be free in γ. But from these premises we can derive not only that a nominal bearing both [+anaphoric] and [+pronominal] features could lack a governing category, but also that a nominal bearing either [+anaphoric] or [+pronominal] features does not require it. In other terms, if having a governing category is not a necessary but a sufficient condition, then the theory at least in principle allows for [+anaphoric] or [+pronominal] nominals to not have a governing category. Obviously, this thesis would be extremely problematic for the generative theory, and consequently it cannot be maintained. As a result, the most "straightforward" way to derive the PRO Theorem is by forcing the contradiction to explode.

[22] Government is a structural relationship between a governor and a governee, defined in terms of m-command. For a precise definition of *government*, *governor* and *intervening node* see [Riz90]; [Hae94]:Ch.2. **m-command**$_{def}$: a node α m-commands a node β if and only if: i) α and β are different nodes; ii) α does not dominate β nor β dominates α, iii) the maximal projection of α does dominate β.

case-marked, PRO cannot be overt; it must be phonologically null. Conversely, an overt DP cannot occupy the [Spec-IP] position of an infinitive clause. Thus, the Case Filter and the concept of government explain the absence of a phonological form for PRO, as well as the complementary distribution between overt DPs and PRO. In sum, the properties of PRO — nullness, caselessness, and complementary distribution with respect to overt DPs — , are indirect consequences of its opposing features: PRO is *both* an anaphora *and* a pronoun.[23]

2.2.3 PRO in the Minimalist Program

While the concept of government was given up in the Minimalist Program (MP), in this most current version of the generative framework the empty category PRO continues to be the subject of infinitive controlled clauses. Replacing the notion of government as an *explanans* of the structural case-assignment in GB is a checking relation between a specifier and its head in MP. Despite these conceptual and technical changes, the distribution of PRO still continues to be determined by the Case Theory. In contrast with GB, in MP the non-finite T(ense) head really can check the case of the subject in [Spec-TP];[24] although the case checking for this T head is not the nominative one as in finite clauses, but a special so-called null Case. Thus, in MP the occurrence of PRO is limited to the subject position of the non-finite T(ense) head (i.e. [Spec-TP]). Assuming that only PRO can check the null Case in [Spec-TP], where T is non-finite, the complementary distribution between PRO and lexical DPs is explained once again as a consequence of the Case Theory. Thus, in both GB and MP, PRO does not occur in the same contexts as overt DPs.

The minimalist take on the distribution of PRO was later improved by Martin [Mar01]. As he observes, not every non-finite T allows the occurrence of PRO in its specifier position. Besides control verbs, other types of verbs select an infinitive complement: for instance, raising and Exceptional Case Marked (ECM) verbs.[25]

[23] Ever since its formulation, the PRO Theorem has been the target of various objections. Some challenge the consequences of the theorem; others, its premises. Hornstein [Hor99], for example, notes that PRO is ambiguous: in some contexts PRO is anaphoric, and so bears [+anaphoric, −pronominal] features; in others, it is pronominal, and so bears [−anaphoric, +pronominal] features. Thus, Hornstein rejects the claim that PRO simultaneously carries both [+anaphoric, +pronominal] features. Similarly, Landau [Lan00] distinguishes between obligatory controlled PRO and non-obligatory controlled PRO.

[24] Since Pollock [Pol89], the IP projection has been split into several layers: AgrP, TP (and NegP). AgrP was dispensed with in the minimalist framework. The basic minimalist sentence structure is shown in i.:

i. $[_{CP}\ [_{TP}\ [_{(NegP)}\ [_{VP}\ ]]]]$

[25] ECM constructions are analyzed as 'raising-to-object' constructions by Martin [Mar01].

Other non-finite complements: ECM, raising

In order to distinguish between control complements and raising (to-subject and to-object) complements, and so to explain why PRO cannot check the null Case in raising complements, Martin pays attention to the temporal properties of the infinitives selected by these classes of verbs. Following Stowell [Sto82], he notes that the event time of non-finite control complements is, in some sense, either future or unrealized with respect to the event of the matrix clause, while the tense of non-finite raising complements is identical or simultaneous with the tense of the matrix clause (cf. [Mar01]). The following English sentences are instances of control sentences (see (20–21) below), ECM constructions (see (22–23)) and raising (to-subject) (see (24)):

(20) Sara convinced Bill [to go to the party].

(21) Bob wants [to buy a new camera].

(22) Zagallo believed [Ronaldo to be the best].

(23) The doctor showed Bill [to be sick].

(24) The defendant seemed [to the DA to be guilty]. [Mar01]

Although raising verbs, like control verbs, select a non-finite clause as their complement, these two classes of verb are different with respect to the semantic (or thematic) selection: raising verbs do not impose a semantic restriction on their syntactic subject. Because of this, in English raising verbs (e.g.: *seems*), but not control verbs (e.g.: *hopes*), admit the expletive *it* (or *there*) in their subject position:

(25) a. John seems [to leave early].
b. It seems that John leaves early.

(26) a. John hopes [to leave early].
b. *It hopes that John leaves early.

Other differences between control and raising constructions are consequences of the fact that control verbs, but not raising verbs, select a semantic external argument. In GB, the embedded subject position is occupied by a trace t co-indexed with the subject matrix DP; this trace marks the argument base position the DP occupies in the deep structure, that is, the embedded thematic position that the movement — raising — of the subject DP arises from. Traces are co-indexed variables, and thus they are syntactic objects in the GB framework.[26] Despite the fact that traces are not admitted in the more current version of the Generative Grammar MP, another non-phonological category is assumed in raising structures: namely, a deleted copy of the raised DP. Whether we have a trace or a

[26] In standard generative semantics, variables are assumed to denote an individual under an assignment function. Hence, the denotation of a phrase containing a trace (or a variable in general) will be relative to an assignment function as well (cf. [HK98]).

deleted copy of the DP controller as syntactic entities, the syntactic category occupying the subject position within the infinitive complement of raising verbs is not PRO (see (27a-b) below):

(27) John seems [to leave early].
a. John$_1$ seems [t_1 to leave early].
b. John seems [~~John~~ to leave early].

To formalize the temporal difference between infinitive complements of control verbs and those of raising verbs, in Martin's proposal the former are assumed to carry a [+tense] feature and the latter a [−tense] feature. Assuming that only non-finite [+tense] T heads can check the null Case, Martin explains why PRO occupies the specifier position only in the non-finite T head of control sentences. By contrast, non-finite T of ECM and raising infinitives have no case at all and, consequently, the main finite T head has to check the case of their embedded subjects.

As an aside, note that a similar situation occurs in other languages besides English. In Portuguese, for instance, ECM verbs (see (28–29)) and raising verbs (examples (30–31)) can also select a non-finite complement:

(28) Vi [**os soldados** cair].
the soldiers fall.INF
'I saw the soldiers fall.' [Sit02]

(29) O juiz mandou os Sem-terra [sair(**em**) da fazenda].
the judge ordered the Landless go-out.(INFL.3PL) of-the farm
'The judge ordered the Landless to leave the farm.' [DS09]

(30) João parece [cantar].
John seems sing.INF
'John seems to sing.'

(31) A bondade costuma [ser escassa em tempos de crise].
the kindness tend be.INF limited in times of crisis
'Kindness tends to be scarce in times of crisis.' [MSL13]:153

In the above, we see that ECM verbs select a single internal argument: a clause whose Θ-role is an event. Though the DP *os soldados* 'the soldiers' in (28) above checks the accusative Case from the matrix verb *ver*, *the soldiers* is not the object of the perception verb, but the syntactic subject (external argument) of the infinitive embedded verb phrase. This fact can be witnessed in those cases where this DP appears in a clitic form (see (32) below).[27] In ECM structures, the syntactic subject of the non-finite complement clause is not the null subject PRO, but an overt accusative-case marked DP.

(32) Os vi [cair].
CL.ACC saw fall.INF
'I saw them fall.'

2.2.4 Control as Movement

The Movement Theory of Control (MTC) proposed by Hornstein [Hor99] is based on the minimalist copy theory of movement. Roughly, in MP *movement* is a mechanism of copy, internal merge (of copies), chain formation and copy deletion (i.e. chain reduction). In somewhat more concrete terms, control structures are obtained by copying a DP and raising it to the controlled landing position for the controller. Thus, the matrix controller DP is an overt raised copy of the deleted *in situ* controlled DP.

From a logical point of view, it is important to remark that movement in MP is not a result of a genuine modification of the order of syntactic constituents; movement is not related to the structural rule of Permutation, but an apparent phenomenon produced by the Expansion rule along with (a long-distance form of) Contraction.[28]

The derivation of a subject control sentence can be schematized as in (33a-e) below:

(33) John hopes [to win].
a. [$_{vP}$ John win]
b. [$_{TP}$ John to [$_{vP}$ John win]]
c. [$_{vP}$ John hopes [$_{TP}$ John to [$_{vP}$ John win]]]
d. [$_{TP}$ John [$_{vP}$ John hopes [$_{TP}$ John to [$_{vP}$ John win]]]]
e. [$_{TP}$ John [$_{vP}$ ~~John~~ hopes [$_{TP}$ ~~John~~ to [$_{vP}$ ~~John~~ win]]]]

In the first step of the derivation the DP *John* is inserted in the [Spec-vP] position, where it checks the Θ-role of the functional v head and gets the Θ-role of agent. Subsequently this DP is copied and the copy is merged in the [Spec-TP] position, where it checks the D-feature (or EPP-feature) of the non-finite T head. However this DP cannot check the nominative case feature due to the properties of the T head.[29] After inserting the matrix v head, the DP *John* is copied once again and this copy is merged in [Spec-vP], where the

[27] Causative and perception verbs also admit a mono-clausal structure in European Portuguese. In mono-clausal structures the perceptive and causative verbs select an internal argument (or direct object) (cf. [Sci04].)

[28] See Fig. 5.3 in Chapter 5.

[29] In highly simplified terms, the Extended Projection Principle (EPP) states that every [Spec-TP] position has to be filled regardless of the argument structure of the predicate [Cho82]. EPP has received much attention in the literature and it is debated whether such a condition can be derived from the formal properties of the T head, combined with the more basic theoretical assumptions of the Minimalist Program (cf. [Las01]; [CL93]; [Oku94]; [Boš02]).

DP gets a second Θ-role from the *v* head *hope*.[30] When this copy is merged in [Spec-TP], it checks the nominative case with the finite T head. Finally, all the lower copies of the sequence of copies (or chain) are deleted.

Accordingly, the controller and the subject controllee in control sentences are not two coreferential DPs, but a single DP occupying two different Θ-positions. This syntactic approach to control has an important semantic consequence. The type of predication relation established between two coreferential DPs is different from predication on copies of a single DP. In the former case, an entity — denoted by coreferential expressions — satisfies two independent properties; in the latter, an entity — denoted by the same DP — satisfies a reflexive property. In fact, in Hornstein's theory, a sentence containing a reflexive pronoun is obtained by the same mechanism of copying and raising whereby a control sentence is derived. The sentences in (34) and (35) are both assumed to be derived in the same way, although the first has a non-surface object and the second has a reflexive object pronoun (cf. [Hor99]):

(34) John shaved.

(35) John shaved himself.

Both sentences are derived by raising the DP *John* from the object to the subject position. Hence, this DP is assigned two Θ-roles: agent and theme. In both sentences, the lower copy is deleted; however in the derivation for the second sentence a reflexive pronoun is inserted. According to Hornstein, the reflexive pronoun is not a lexical item in the numeration, but the phonological realization of the grammatical operation of copying a DP (cf. [Hor06]; [Mar11]).

Hornstein's theory paves the way for admitting non-finite subjects other than null (or covert) ones, since it recognizes the overt realization of copies in sentences with reflexive pronouns. In fact, the movement theory of control has been used to explain several non-prototypical forms of control, like forward, backward and copy control. In this last form of control, both the controller and the controlled are overt. The phenomenon of copy control is attested in Hungarian and Bangla [Bis14], Tongan, San Lucas Zapotec [PP06], old English [McF08], Dravidian languages as Assamese, Tamil and Malabar [MS11]; [SM10]. In Zapotec, for example, the controlled subject can be overtly realized not only by a pronoun, but also by a referential expression having a reflexive meaning (cf. [Lee03] and see example in (36) below).[31]

(36)	r-cáá a'z	Gye'eilhlly	g-auh	(**Gye'eilhlly**)	bxaady.
	want.HAB	Gye'eilhlly	eat.IRR	(Gye'eilhlly)	grasshopper

[30] We remind the reader that the Θ-Theory prevents a single DP from being assigned two θ-roles. Hence, the movement theory of control departs from the mainstream generative framework.

[31] Abbreviations used in this example: HAB (habitual mood), IRR (*irrealis* mood).

'Gye'eilhlly wants to eat grasshopper.' [PP06]:183

We maintain that copy control is also attested in Romance languages such as Portuguese, Italian and Spanish. We also claim that the so-called emphatic pronouns in infinitive sentences from these languages (cf. [Riz82]; [PG87]) could be explained as instances of copy control sentences (cf. [GM15]). Even so, for reasons that will soon become clear our preferred analysis will ultimately align with a different proposal, due to Landau [Lan00].[32]

2.2.5 Landau's T/Agr Calculus

Recall that according to the Government and Binding theory (GB) and the Minimalist Program (MP), the occurrence of PRO is limited to non-finite defective domains, as the distribution of PRO is determined by the defectiveness of the operation of assigning or checking (nominative) Case. And recall that although in the Movement Theory of Control (MTC) [Hor99] PRO is not admitted as the subject of control structures, the phenomenon of control is similarly explained as a consequence of some form of defectiveness related to the case properties of the non-finite embedded domain. Like GB and MP, Landau's account (2000 and subsequent work) admits the theoretical category of PRO as the controllee subject. By contrast, the distribution of PRO is now severed from the case properties of the non-finite contexts in his theory of control.

Landau compiles data from Icelandic, Russian, German, Hungarian and Korean in order to give support to the controversial claim that PRO can be marked with a case other than the null case: nominative, accusative, genitive, dative, and so on (cf. for example [Sig08]). This entails in other words that in infinitive clauses, PRO can occur in a position other than [Spec-IP]. (Note that Landau returns to the pre-Pollock notation: thus I° once again denotes the inflectional head, whereas T and Agr denote the tense and agreement features on C° and I°, respectively.) Now since PRO is assumed to be case-marked, it is natural to wonder whether it could occur in the prototypical nominative position — i.e., [Spec-IP] — in finite clauses.[33] Put differently, the question arises of whether finite contexts might feature control.

Based on the study of subjunctive clauses from Balkan languages such as Romanian, Bulgarian, Albanian, Macedonian and Modern Greek, Landau assumes that controlled PRO is licensed in defective domains. Importantly, however, in his proposal this

[32] Although the syntactic operations of copying and deleting could be implemented in an extension of the substructural Lambek Calculus by adding some restricted form of the expansion and contraction rules, the operation whereby a syntactic category (the copy of a DP) is transformed or replaced by another (a reflexive pronoun or a personal pronoun) does not seem to be easy to perform in a monostratal grammar.

[33] Traditionally, in Generative Grammar, [Spec-TP] is restricted to nominal phrases, to overt (referentially free) pronouns, to overt expletives, such as *it* and *there* in English, or to the null pronoun *pro* of the *pro*-drop languages.

defectiveness is not a consequence of some structural defectiveness of the controlled clause. In Landau's T/Agr Calculus controlled clauses are assumed to be complementizer phrases (CPs); as such, they are structurally complete.[34] In addition, the defectiveness is not related to a morphological defectiveness of the controlled verb. Since in some languages the I(nflectional) head of the controlled clauses presents a complete set of ϕ-features (or inflectional features), the controlled clause can be morphologically complete (cf. [Lan00] and subsequent work). In Landau's theory, defectiveness is related to a kind of weakness of the [T] and/or [Agr] features on the C and I heads of the embedded CP. In a nutshell, defectiveness, and so control, may come about in one of two ways. That is, control may either result from the [T] and/or [Agr] features taking a *negative value* on either head; or it may result from the *absence* of either feature on C° or I° in the complement clause.

The [T] feature

In Landau's T/Agr calculus, [T] is not a morphological but a semantic (tense) feature. For this reason the [T] feature is not directly related to the inflectional morphology. In the case of a complement clause, its semantic tense is determined by its relation with the tense of the main clause. An embedded tense can be interpreted with respect either to the speech (or utterance) time or, alternatively, to the time of the main clause situation, which in turn is evaluated considering the utterance's point of view. In this sense, an embedded tense may or may not be independent of the matrix tense. For a complement clause to be interpreted regardless of the utterance time of the main clause, it must itself be anchored to the speech time.

According to Landau the semantic tense of a complement clause can be identical (or anaphoric) with, or dependent or independent of, the tense of the matrix clause. The tense a specific complement clause bears is determined by testing the temporal relationship between the matrix and its embedded complement clause. Specifically, whether the embedded clause carries its own tense operator can be tested by the possibility of generating a mismatch in tense between the matrix and embedded clauses by using deictic adverbials, i.e. adverbials that refer to the speech time.

In particular, an infinitive clause is not anchored to a speech time (i.e. it lacks deictic anchorage), and so its tense cannot be directly evaluated with respect to the same. Infinitive clauses cannot bear an independent tense; they can only bear an anaphoric or a dependent tense. The anaphoric (also called empty) tense is identical to that of the matrix clause; a complement clause carrying anaphoric tense lacks a tense operator, it cannot circumscribe an own temporal domain, and thus cannot host a deictic adverbial. The anaphoric tense amounts to the presence of an overlap between the matrix situation

[34] The functional projections are linearly embedded ordered: $[_{CP}\ [_{IP}\ [_{vP}\ [_{VP}]]]]$, where *CP* stands for Complementizer Phrase, *IP* stands for Inflectional Phrase, *vP*, for light verb Phrase and *VP*, for verb Phrase.

(event or state) and the embedded situation. The following infinitive complements are examples of such tenseless complements:

(37) Yesterday, John managed [to solve the problem (*tomorrow)].

(38) Juan quiso [ir al cine (*mañana)].
John wanted [go.INF to-the cinema tomorrow]
'John wanted to go to the cinema (*tomorrow).'

(39) Io volevo [andare a vedere la partita (*domani)].
I wanted [go.INF to watch.INF the game tomorrow]
'I wanted to go watch the game (*tomorrow).'

But it is not only infinitive clauses, which are morphologically defective as they lack inflectional features, that can bear an anaphoric tense. Some kinds of subjunctive complements from Balkan languages also lack deictic anchorage; they cannot license an own independent temporal domain, and thus exhibit an anaphoric tense dependence between the subjunctive and the main clauses.[35] Despite displaying a finite morphology, subjunctive complements are in some sense tenseless in the same way that infinitive complements are.

(40) Am reuşit să plec (*mîine).
Aux.1SG managed [SBJ leave.1SG tomorrow]
'I managed to leave (*tomorrow).' [Alb04]: 63

By contrast, a clause bearing a dependent tense can locate an event or state in a temporal domain different from the temporal domain of the main clause. However, the situation described in a dependent-tense clause is still evaluated considering the utterance time introduced by the main clause. In other words, the event time of a dependent-tense clause is indirectly related to the time of utterance through the matrix event clause. Since a clause bearing dependent tense can define its own temporal domain, it can host a deictic adverbial associated with the time expressed by the event described in it. As the sentences in (41) and (42) below show, some infinitive complements do allow this kind of temporal mismatch with the matrix clause.

(41) Yesterday, John wanted [to solve the problem tomorrow]. [Lan00]

(42) Io ho permesso a loro [di mangiare al ristorante domani].
I have permitted to them [COMP eat.INF to-the restaurant tomorrow]
'I allowed them to eat at the restaurant tomorrow.'

[35] Landau calls these complements C(ontrolled) subjunctives. Alboiu [Alb04] does not consider this kind of Romanian subjunctives a CP, but a TP, because they do not admit the complementizer *ca*.

(43) Ricordo [di aver messo il portafoglio nel cassetto].
Remember.1SG [COMP have.INF put.PP the wallet in-the drawer]
'I remember that I put the wallet into the drawer.' [Bia07]

(44) Dichiaro [di aver ricevuto dal signor Pallino la somma di euro 300].
declare.1SG [COMP have.INF receive.PP de-of mister Pallino the sum of euro 300]
'I declare that I have received euros 300 from Mister Pallino.'

The sentences in (42–44) are examples of compound infinitives from Italian. Compound infinitives formed by the auxiliary verb *haber/avere/ter* 'have' plus the participle in Spanish, Italian and Portuguese express a temporal priority relation between the event they describe and the situation described by the matrix verb. In these infinitive clauses the auxiliary verb induces a shift to the past from the matrix tense; compound infinitives in these languages describe an event which is temporally located before the matrix event. Thus, the tense domain of these embedded infinitive clauses is different from, but dependent on, the matrix tense domain.

Finally, in a complement clause with independent tense an utterance time different from the main one is available to evaluate the embedded situation. The speech time of an independent-tense clause is not determined by the matrix speech time. Hence, both matrix tense and embedded tense are considered from their own utterance's point of view. Thus, it is clear that embedded clauses bearing an independent tense can host their own deictic adverbial. Typically, this semantic tense is exhibited in indicative complement clauses:

(45) Yesterday he said he wants me back today.

The type of semantic tense—anaphoric, dependent or independent—an embedded clause bears depends first and foremost on the semantic properties of the matrix verb that select it; but it also varies according to the specific structure of the clause. In general, propositional, factive and desiderative verbs select an embedded clause that admits a tense mismatch. Among these, some verbs introduce a temporal orientation of the embedded event; nevertheless the only admissible tense in these cases is a dependent tense. By contrast, implicative, aspectual and modal verbs select an embedded clause with anaphoric tense [Lan13].[36]

In Landau's calculus the semantic tense of a clause is formalized by the [T] feature, which ranges over two values: positive or negative. The value for the [T] feature on C° and I° is fixed by the rule 1 below (cf. [Lan04]:870):[37]

[36] In Landau's [Lan15] terminology these kinds of verbs trigger the so-called *obligatory predicative control*.

[37] For the sake of precision, we use the notation [$\emptyset_x$], where x stands for T or for Agr, when a head lacks [T] or [Agr] features, respectively.

1. a. Anaphoric (or "empty") tense ⇒ [−T] on I°/C°;
   b. Dependent tense ⇒ [+T] on I°/C°;
   c. Independent tense ⇒ [+T] on I°, [$\emptyset_T$] on C°. [T on C°/I° rule]

Note that since the [T] feature is semantic and not morphological, it can be negatively specified ([−T]) in finite clauses and positively specified ([+T]) in (un)inflected infinitive clauses. As the above rule shows, the [T] feature is only specified on the C head when the embedded tense is not independent of the matrix tense, that is, when the tense of the complement clause is selected or restricted by the matrix predicate. Now recall that the state or event of anaphoric and dependent-tense clauses has to be evaluated with respect to the speech time of the main clause, since it is not anchored to a time of utterance. This seems to be mirrored in rule 1 above by the presence of the [±T] feature on C°. Conversely, recall that in complement clauses with an independent tense the embedded event is not evaluated with respect to the matrix speech time. Again, this seems to be formalized in Landau's rule by the lack of a [T] feature on C°. Further, the fact that both dependent and independent-tense clauses carry a [+T] feature on I° seems to indicate that these clauses can define their own temporal domain. In contrast anaphoric clauses, which are not anchored to the speech time and also do not define an own temporal domain, carry a [−T] feature on both functional heads.

Finally note that from a theoretical point of view, the [T] feature is uninterpretable on the C head; according to minimalist assumptions, it must therefore be checked and eliminated. Specifically, the uninterpretable [T] feature on C° is checked off by an Agree relation — involving feature matching, checking and deletion [Cho00] — established with the interpretable [T] feature on I°.[38]

The [Agr] feature

The second formal component of Landau's calculus of control is the [Agr] feature. Like [T], [Agr] is carried by the functional I and C heads; unlike [T], it is uninterpretable on both. As we will see, an I head carrying a [±Agr] feature enters into an Agree relationship

[38] Agree is a syntactic relationship between a probe and a goal. Lexical and functional items are inserted in the derivation carrying a number of features; some of them interpretable, others uninterpretable (*u*F). The *u*F are properties that play no role in the semantic interpretation (in more technical terms, they are illegible to the semantic component). The ϕ-features on the T head and the *wh*-feature on C are examples of *u*F; the ϕ-features on a DP are an example of interpretable features. The uninterpretable features must be checked and also eliminated. Whereas all features must be checked, only uninterpretable features must be eliminated. To do so the *u*F seeks a goal. Thus, the *u*F acts as a probe of the Agree relation. This relation is, in a certain sense, local; the Agree relation is established within a very restricted syntactic domain: phase. It is widely assumed that a simple sentence (in the active voice of a transitive or inergative verb) presents two phases: *v*P, the weak phase, and CP, the strong phase. In sum, Agree is a syntactic relationship established between (at least) one uninterpretable active feature F on a functional head (C, T, *v*), which works as a probe of the corresponding feature located in an item or lexical phrase that operates as a goal in the restricted c-command domain of the goal.

with a nominal phrase. On the other hand, a [+Agr] feature on C° is checked off by an Agree relation with a corresponding [+Agr] feature on I° (ditto for [–Agr] features).

To determine the specific value of [Agr] on these embedded heads, Landau proposes two independent rules. The [Agr] feature on I° is morphologically determined; its value expresses the presence or absence of ϕ-features, as per the rule in 2 below:[39]

2. a. overt agreement ⇒ [+Agr];
   b. abstract agreement ⇒ [−Agr];
   c. non agreement ⇒ [$\emptyset_{Agr}$]. [Agr on I° rule]

The value of [Agr] on C° is stipulated by the rule in 3 below, which establishes a dependency relationship between the [Agr] and [T] features on the C head:

3. a. [+Agr] ⇒ [+T];
   b. otherwise ⇒ [$\emptyset_{Agr}$].[40] [Agr on C° rule]

We should note, before anything else, that both the above rule and Landau's discussion of the same seem to leave room for somewhat different interpretations of its content. Here we present what appears to us to be the most plausible of these, that best coheres with the wider framework of his account.

With this in mind, it is fairly clear that the first clause of the rule in 3 establishes a dependence of the [+T] feature on the [+Agr] feature. To illustrate this dependence Landau [Lan04]:840 appeals to a metaphor: "[+Agr] is parasitic on [+T], [although] only in the sense that the latter is a necessary condition for the former."[41] As we understand it, the metaphor implies that the presence of [+Agr]—i.e., the host—shows up the presence of [+T]—i.e., the parasite—on C°.

The second clause of the rule seems to limit the scope of the first by selectively blocking its logical consequences. Taken on its own, clause 3a would classically entail that if [T] takes a non-positive value, either [−Agr] or [$\emptyset_{Agr}$] would follow. Instead, the second clause says that the only permissible consequence of [T] having a non-positive value on C° is the null value [$\emptyset_{Agr}$]. In other terms, where the [T] feature has a value other than a positive one, C° lacks the [Agr] feature. Thus, while the first clause has as its (sufficient) condition an [Agr] specification, the second clause breaks the expected parallelism and puts a [T] specification as its condition.[42]

[39] Where [−Agr] is [Agr] with no morphological realization (cf. [Lan04]: 839).

[40] $\emptyset$ in Landau's notation.

[41] Landau returns to this metaphor in subsequent works. For instance, he writes: "Recall that [+T] on Comp may "parasitically" introduce [+Agr]" [Lan06]. And "PRO-control is always available; C-control is only available when C contains a ϕ-set (normally parasitic on a [Tense] feature)" [Lan08].

[42] At first sight, the dependence of [+T] on [+Agr] seems incompatible with the rule in 1, whereby [T] features on I° and C° are determined. It seems somewhat odd that [+T] on C° should depend on

It is important to note that the combinations [+T, −Agr] and [+T, $\emptyset_{Agr}$] on the C head satisfy the first clause of the rule, as they vacuously satisfy the antecedent and are not blocked by the second clause.[43] Table 2.1 displays the five combinations of [T, Agr] features on the C head obtained from the rule in 3.

	C°		
	[T]	**[Agr]**	**Examples**
1.	[$\emptyset_T$]	[$\emptyset_{Agr}$]	indicative clauses
2.	[+T]	[$\emptyset_{Agr}$]	factive complements (European Portuguese), partial control
3.	[−T]	[$\emptyset_{Agr}$]	C-subjunctives (Balkan languages), exhaustive control
4.	[+T]	[+Agr]	F-subjunctives (Balkan languages), subjunctive (Hebrew)
5.	[+T]	[−Agr]	

Table 2.1 – Possible combinations for [T, Agr] features on the C head

The first four combinations have been used by Landau to give an account of controlled and non-controlled complement clauses in different languages.[44] By contrast, the fifth combination ([+T], [−Agr]) has so far remained unexplored in the literature. With this thesis we propose to fill precisely this lacuna. Thus, in Chapter 4 we show how the [+T, −Agr] combination not only finds witnesses in Romance languages, but is also key to giving an account of infinitive non-controlled clauses selected by propositional verbs, in Spanish and Italian.

The [R] feature

The third component of Landau's framework has to do with the distribution of embedded subjects, and as such is the nub of his theory of control. Landau groups nominal phrases according to their referential properties, as is standard; the latter are codified by [±R]. On his account, overt or null nominal phrases bearing ϕ-features carry a [+R] feature; anaphoric DPs, which lack any inherent specification for ϕ-features and require an antecedent for identification, bear a [−R] feature.[45] Finally, the [R] feature on nominal phrases is interpretable. The assignment rule for [±R] is schematized in 4:

the presence of a complementizer or some other [+Agr] constituent on C°, as [+T] on C° expresses the fact that the tense of the embedded clause is dependent on the main tense. Nevertheless, if we assume [+Agr] on C° as a consequence of the presence of an (inflected) auxiliary verb on C°, this dependency relationship can be expected.

[43] Reed [Ree12]:131 uses another (although incorrect) interpretation of the rule in 3. According to her, if C is [+T], then C is either [+Agr] or it is unspecified for [Agr]; if C is [−T] or unspecified for [T], then it is unspecified for [Agr]. Although Landau's notation [$\emptyset$], which is used in the consequent of the second clause of this rule, is ambiguous between [$\emptyset_{Agr}$] and [$\emptyset_T$] — and as such could give rise to misunderstanding — Reed surprisingly changes the first clause of the rule, which is clearly and unambiguously formalized by Landau [Lan04]:804. The feature combination [+T, −Agr] turns out to be problematic under Reed's interpretation, since it is blocked by the first clause, while it is a logical consequence of the second.

[44] It seems that Reed's interpretation of the [Agr] rule is inferred from the feature combinations that Landau has actually used, since it covers all and only these four combinations of [T, Agr] features on the C head.

[45] Landau assumes that *pro*, unlike PRO, bears ϕ-features that need to be identified.

4. a. lexical DP, *pro* $\Rightarrow$ [+R];
   b. PRO $\Rightarrow$ [−R]. [±R rule]

Recall that on Landau's account the distribution of embedded subjects is not determined by their case properties; his theory does however agree with GB and MP in accepting the complementary distribution between lexical DPs and *pro* on the one hand, and PRO on the other. This is because, as we can see, the rule above ensures that [+R] and [−R] nominal classes are mutually exclusive. In addition, Landau's theory preserves the nullness of the controllee subject. In fact, although the [R] feature formalizes a semantic or referential distinction, and not a phonological one, Landau does not assume an overt form of [−R] nominal: null PRO is the only nominal carrying a [−R] feature (cf. [Lan04]:841).[46] Indeed, as he claims in later work ([Lan06]:162) no overt anaphoric DP — not even a SE-anaphor — can occur in the same syntactic contexts in which the null PRO occurs.[47] In fact, Landau holds that PRO is a radically impoverished null SE-anaphor; though unlike a SE-anaphor, PRO is unspecified for any and all ϕ-features: number, gender as well as person. The slots for each unspecified ϕ-feature of PRO are valued under agreement with the functional head that licenses the controller.

Besides nominal phrases, the C and I heads can also host an [R] feature. In these cases, whether [R] takes a positive or a negative value depends on the values of the [T, Agr] features, as per the rule in 5:

5. For X $^{\circ}_{[\alpha T, \beta Agr]} \in \{I^{\circ}, C^{\circ}, \ldots\}$:
   a. [+R]/X$^{\circ}_{[.]}$, if $\alpha = \beta = +$;
   b. [−R]/elsewhere. [±R assignment rule]

The first clause stipulates that a [+R] feature is borne by functional heads carrying positive values for both [T] and [Agr] features. By the second clause, when at least one of [T] or [Agr] has a negative value, the corresponding head bears a negative [R] feature as well. In addition, lack of [T] and/or [Agr] features on the C or I head renders the

[46] Landau departs from Reinhart & Reuland [RR93], who assign the [−R] feature to two types of anaphors — SELF and SE-anaphors — and the [+R] feature to R-expressions and pronouns.

[47] Reinhart & Reuland [RR93] make a syntactic distinction between two types of reflexive in Dutch (German): a weak reflexive *zich* (*sich*) and a strong reflexive *ziszelf* (*sich selbst*), called a SE-anaphor and a SELF-anaphor respectively. SELF-anaphors but not SE-anaphors are specified for ϕ-features for person, number and gender in Dutch. Although SE-anaphors may preserve person features in many languages, they always lack number and gender features. According to Reinhart & Reuland, a SELF-anaphor is an operator that can reflexivize a transitive predicate, while a SE-anaphor is not an operator. SE-anaphors in Dutch can only occur in necessarily or inherently reflexive predicates. Pragmatically, *zich* can never be phonetically focused, while *zichzelf* is used in contrastive situations. We remind the reader that in GB theory reflexives are constrained by Principle A of Binding Theory: a reflexive must be bound in its local domain. SELF-anaphors, but not SE-anaphors, respect this principle.

R-assignment rule inapplicable, i.e., no [R] value is assigned on the functional head in which one (or both) of these features is absent.[48]

When carried by either functional head, the [R] feature is uninterpretable; it must therefore be checked off and deleted. To check off the uninterpretable [R] feature on C°, an Agree relation between C° and I° is established. In fact, this Agree relation is responsible for the distribution of controlled [−R] and non-controlled [+R] embedded subjects.

Distribution of PRO

Broadly speaking, there are two ways to check off the uninterpretable [R] feature on the I head. One possibility is for it to be checked off through an Agree relation with the interpretable [±R] feature on a nominal embedded phrase carrying the same — positive or negative — value. This derivational path is exhibited by the so-called C(ontrolled) and F(ree)-subjunctives in Balkan languages:

6. C-subjunctives: [–R] subject

$$[_{CP}\ C^{\circ}_{[-T]}\quad [_{IP}\ PRO_{[-R]}\quad [_{I'}\ I^{\circ}_{[-T,+Agr,-R]}\quad [_{VP}\ t_{PRO[-R]}\quad]]]]]$$

Agree [+Agr,–R]

Agree [-T]

7. F-subjunctives: [+R] subject

$$[_{CP}\ C^{\circ}_{[+T,+Agr,+R]}\quad [_{IP}\ [_{I'}\ I^{\circ}_{[+T,+Agr,+R]}\quad [_{VP}\ DP/pro_{[+R]}\quad]]]]]$$

Agree [+T,+Agr,+R] *Agree* [+Agr,+R]

In the first case — C-subjunctive — the controlled [−R] subject PRO is licensed through the Agree relation with the I head bearing [−T, +Agr] features, and consequently a [−R] feature. In the second case — F-subjunctive — a referentially free [+R] subject is licensed through the Agree relation with the I head bearing [+T, +Agr] features, and consequently a [+R] feature. The sentences in (46–47) instantiate the derivational paths in 6 and 7, respectively.

(46) I Maria$_1$ prospathise PRO$_1$ na divasi.
the Mary tried PRO SBJ read.3.SG
'Mary tried to read.' [Ter97] (*apud* [Lan04]:826)

[48] For the sake of precision, the general condition of the rule should also explicitly state that $\alpha, \beta \in \{+, -\}$. It is not clear to us what other possible functional heads other than C and I could be included in the set $\{I^{\circ}, C^{\circ}, \ldots\}$.

(47) Ion$_1$ vrea ca Dan/*pro*$_2$ să rezolve problema.
Ion wants that Dan/*pro* SBJ solve the-problem
'Ion wants Dan/him to solve the problem.' [Far85] (*apud* [Lan04]:828)

The other option is for [R] to be checked off by entering in an Agree relation with C° through a "conspiracy Agree relation" (cf.[Lan04]:844). The idea here seems to be the following: first, C° enters the Agree relation with I°, to check off the uninterpretable [T] feature on C°; this is standard, and a given. But now in addition to this, I° also enters the Agree relation with C°: as a result, the [R] features on I° and C° effectively check each other off, leaving only [+Agr] to be dealt with. The "conspiracy" mechanism is illustrated below:

8. F-subjunctives: [−R] subject

Agree [+R]

$[_{CP}\ C^{\circ}_{[+T,+Agr,+R]} \quad [_{IP}\ PRO_{[-R]} \quad [_{I'}\ I^{\circ}_{[+T,+Agr,+R]} \quad [_{VP}\ t_{PRO[-R]} \quad]]]]]$

Agree [+Agr]

Agree [+T,+Agr,+R]

As we can see, here the uninterpretable [+T, +Agr, +R] features on C° and the uninterpretable [+R] feature on I° are mutually checked by the Agree relation. Further, the uninterpretable [+Agr] feature on I° is checked through an Agree relation with a [−R] subject PRO. This second option is available for F(ree)-subjunctives in Balkan languages, which license not only a referentially free [+R] subject, but also a controlled [−R] subject despite the [+R] feature on the I head. In the particular case in which the embedded subject of F-subjunctives is phonologically null, this subject could be controlled [+R] PRO or non-controlled [–R] *pro*, as exemplified in (48) from Romanian:

(48) Ion$_1$ vrea ca Dan/*pro*$_2$/PRO$_1$ să rezolve problema.
Ion wants that Dan/*pro*/PRO SBJ solve problem
a. 'Ion wants to solve the problem.'
b. 'Ion wants Dan/him to solve the problem.' [Far85] (*apud* [Lan04]:828)

We now have a clearer picture of the syntactical aspects of Landau's account. We conclude this section with a brief overview of the semantics of PRO.

Interpretation of PRO

As noted earlier, on Landau's account PRO lacks both ϕ-feature and case features. Due to this fact, Landau [Lan15] recognizes that PRO may also be seen to fall within the class of so-called *minimal pronouns*: that is, those pronouns that are born without (a complete set of) features (Kratzer [Kra09]). For an example of a minimal pronoun, consider the two readings for the second occurrence of *you* in the following sentence:

(49) Only you eat what you cook. [Kra09]:188
a. Nobody else eats what you cook.
b. Nobody else eats what s/he cooks.

On the second interpretation, the ϕ-features for the second occurrence of the personal pronoun *you* are not interpreted. This occurrence is interpreted as having the ϕ-features of the third person singular pronoun *he* or *she*, as it is bound by the quantifier *nobody*.

Definition 2.1 (Minimal Pronoun)**.** X is a minimal pronoun iff X = [D, $u\phi$], where [$u\phi$] stands for unvalued ϕ-features [Kra09].

In general, minimal pronouns acquire their missing features, and hence their morphophonological shape, via local agreement relations established in the syntax. In the case of PRO, an Agree relation is established with a matrix DP; by this means, the latter contributes its ϕ-features to the former; and the control relation between PRO and matrix DP is formed.[49]

Even though the ϕ-features of PRO become valued, PRO is not a variable, and consequently does not denote an individual under any assignment function. PRO has no semantic value of its own, yet its presence affects the semantic value of the constituent in which it occurs because it functions as a λ-operator upon movement [Lan15]:28. In effect, when PRO is raised, it leaves a copy behind; the deleted copy of PRO serves as the variable being abstracted from.[50] Raising PRO is then an instance of a type-shift rule;

[49] In more precise terms, the relationship between the matrix DP controller and controllee PRO is not direct, but mediated by a functional head F° located within the matrix clause. In subject control constructions, T° is the functional head that links controller and PRO; in object control structures, the functional head is (light verb) v°. Hence, the matrix head F — T° or v° — enters into two agreement relations. First, it establishes an Agree relation with the DP — subject or object — controller. Through this first Agree, the T/v head gets the ϕ-features from this matrix DP. When the probe of the Agree operation is T°, the uninterpretable unvalued case feature of the goal DP is in turn also valuated under this Agree relation, as T° bears a complete set of ϕ-features (i.e., T° is ϕ-complete). Second, having valued its uninterpretable ϕ-features, the matrix head F enters into another Agree relation with the null subject PRO. Under this second Agree relationship, interpretable ϕ-features (including case) of PRO are valued. Control is a product of a multiple-Agree relation or, in other terms, a mediated Agree relation between a (controller) matrix DP and the (controllee) subject PRO.

There is another type of control, C-control, which gives rise to partial control. In this type of control, the matrix head F enters into an Agree relation with the embedded C head which in turn enters into an Agree relation with the null, controllee subject PRO.

[50] The embedded clausal projection of control structures comes to denote a λ-predicate because of the movement of PRO to the edge of this projection: [Spec-FinP]. In order to force raising of PRO from [Spec-TP] to [Spec-FinP] it is assumed that the Fin head of obligatory control complements bears a [uD] feature that acts as a probe. Since PRO is the closest category bearing an interpretable [D] feature, the

on applying it, the semantic type of the embedded clause that contains PRO, changes. As a result, the embedded phrase containing the moved PRO does not denote an open proposition $VP'\ x$ (i.e. a proposition containing the unbound variable PRO), but rather the function $\lambda x.(VP'\ x)$ from the set D_e of entities to the set D_t of truth values. Hence, although Landau takes PRO to be the syntactic subject of controlled clauses, PRO does not saturate the subject position of the predicate. It follows that the embedded controlled clause does not denote a proposition, but a property of individuals.

2.3 Summing Up

In this chapter we have seen that a great many generative accounts present an almost-united front on the matter of whether an infinitive clause can admit a lexical subject. The majority consensus is motivated by widely accepted assumptions — Case Theory, θ-Theory, Binding Theory — and may be schematically captured by two key claims: that controlled subjects are obligatorily null, and that PRO, and only PRO, is this null subject. Yet, as we have also seen, not all theories of control subscribe to these assumptions. In departing from the norm, these outliers have paved the way for recognizing overt forms of case-marked controllee subjects. We examined two such cases. We saw that by abandoning the θ-Theory, the Movement Theory of Control postulates a nominal category other than the null PRO in the controllee subject position. And we saw that by rejecting the Null-Case Theory, the T/Agr Calculus of Control can postulate a nominative-marked subject in the controllee position of a complement clause.

Further, the concept of minimal pronoun used to categorize PRO opens the door to postulating an *overt* nominal bearing [−R] features, alongside the covert [−R] PRO. To see why this is important, recall that the concept of minimal pronoun is a semantic one, and that there is wide consensus in the literature on the existence of overt minimal pronouns. Together, these observations lend at least provisional support to the hypothesis that there is an overt form of controlled subject after all.

[*u*D] feature on Fin° attracts PRO to its specifier position (cf. [Lan15]:28). By moving PRO from [Spec-TP] to [Spec-FinP], it creates an operator-variable configuration that is interpreted as a property; PRO in [Spec-FinP] works as the λ-operator and the trace — or deleted copy — of PRO located on [Spec-TP] serves as the variable being abstracted over.

3 Overt Subjects in Infinitive Clauses from Romance Languages

3.1 Inflected and Personal Infinitives

Towards the end of Chapter 2 we saw that in principle at least, the idea that infinitive subjects are necessarily null enjoys less support than might be thought. The present chapter seeks to explore this possibility by gathering linguistic evidence to the effect that Romance languages — at least in some syntactic contexts — admit lexical subjects in infinitive clauses. We will be focusing in particular on the so-called *pro*-drop languages: those languages in which the subject position can be filled with the null pronominal category *pro*.[51] *Pro*-drop languages — such as Portuguese, Italian, Catalan and Spanish — therefore admit two null pronoun categories: *pro* and PRO. It is widely accepted that these pronominal categories are in complementary distribution, that is, they do not occupy the same syntactic subject position. The complementary distribution between *pro* and PRO is based on their semantic features and, as a consequence, on their case properties. As briefly mentioned in Section 2.2.2, according to Binding Theory *pro* is characterized by [+pronominal, −anaphoric] features; consequently, its distribution respects Principle B of that theory. By contrast, PRO is characterized by the — jointly problematic — [+anaphoric, +pronominal] features; therefore, its distribution should simultaneously respect Principles A and B of Binding Theory. In order to avoid the incompatible requirements on binding conditions posed by these principles, it is assumed that PRO cannot be (nominative) case-marked. As a result PRO is banned from the [Spec-TP] position of finite clauses; this position is filled instead by *pro*.

A separate argument to this effect has been made in connection with the so-called null subject parameter, which establishes formal licensing and identification conditions for *pro*. According to Rizzi [Riz82], these are respectively that *pro* is governed by an inflectional head, and that it is identified through the ϕ-features on the inflectional head. As a result, it has traditionally been assumed that the nominative case tends to correlate with finiteness or agreement, and indeed the nominative case is assigned on [Spec-TP] by a finite T head. Hence, it is assumed that *pro*, rather than PRO, is the null pronoun that is found in finite environments, where an overt nominative subject can also occur.

Surprisingly, however, Romance languages show instances of overt nominative sub-

[51] Also referred to as Null subject languages. This nomenclature does not commit itself to admitting the theoretical category of *pro*. See [HNS09]; [Hol10] for details on the distinction between consistent or canonical *pro*-drop, partial *pro*-drop and radical *pro*-drop languages. From this perspective, BP is a partial *pro*-drop language whereas EP is a consistent *pro*-drop language.

jects in several infinitive contexts (cf. [Pou95]; [Men00]; [Sit02]; [Sci04]; [Sch07]; [Sza09]; [Paz13], among many others). Some of these languages present an inflected form of infinitive, that is, an infinitive that carries morphemes for number and personal agreement; these are: Portuguese, Galician, Old Neapolitan, Old Leonese, Mirandese and Sardinian.[52] The remaining languages — Italian, Spanish, Romanian, Sicilian, French, Occitan and Catalan — also allow an overt nominative subject, despite lacking agreement morphemes on the infinitive form; they feature the so-called personal infinitives.[53] Both groups of languages admit infinitive clauses with an overt nominative subject in several syntactic positions, such as adjunct, subject, verb complement, attributive and appositive clauses. The following sentences are instances of an infinitive clause with an overt subject appearing respectively in the subject position (50–52), as an adjunct clause (53–55), and as a verb complement (56–58), from Spanish, Italian and Portuguese:[54]

(50) [Ir **yo** mañana a la facultad] va a ser imposible.
go.INF I.NOM tomorrow to the school go.3SG to be.INF impossible
'Going to school tomorrow will be impossible.' [Dem77]:185

(51) El non è vergogna [a essere **l'uomo** abbattuto].
it not is shame to be.INF the man knocked-down
'It is not a shame for a man to be knocked down.' [Men00]:19

(52) [**Os meninos** sair**em** à noite] preocupa suas maẽs.
the boys go-out.INFL.3PL at night worries their mothers
'The boys going out at night worry their mothers.'

(53) María salió de casa sin [**yo** verla].
Mary left the house without I.NOM see.INF-her
'Mary left the house without me seeing her.' [Pou95]:14

(54) Prima di [morire **papà**], mamma era felice.
Before of die.INF dad, mum was happy
'Before dad died, mum was happy.' [Sit02]:72

(55) Eu comprei esse livro pr[**os meninos** ler**em**].
I bought this book to-the boys read.INFL.3PL
'I bought this book for the boys to read.'

(56) señala, [habérselo referido **el reverendo maestre fray Martin de Córdoba**].
note.3SG have.INF.CL relate.PP the reverend master fray Martin de Córdoba

[52] In Sardinian the inflected infinitive is introduced by prepositions *a* 'to' or *de* 'of'. Complements introduced by complementizers *ki* 'that' or *ca* 'that' are considered, by some scholars, as (imperfect) subjunctive forms (cf. [Sci04]).

[53] Some scholars do not distinguish between these two groups and jointly call them *Infinitive with Subject languages* (cf. [Pou95]; [Van13]).

[54] We remind the reader that we use bold font to mark the (supposed) infinitive subjects and the inflection on the infinitive verb.

'He noted that the Reverend Master Fray Martin de Córdoba had related it.' [PyF53a]

(57) Maria mi ha chiesto [di parlare **io** con Gianni].
maria CL has asked of talk.INF I.NOM with Gianni
'Maria asked if I would talk to Gianni.' [Bel05]:21

(58) O governo admite [**eles** vender**em** os imóveis à pessoa].
The government admits they.NOM sell.INFL.3PL the properties to-the person
'The government admits that they sold the properties to the person.' [Gor00]:98

Although personal infinitives from Romance languages lack agreement morphemes for number and person, they can license a nominative-marked subject, as evidenced by the case-marked pronouns within the infinitive clauses in (50), (53), (57) and (58) above. Consequently, personal and inflected infinitives call into question the assumed correspondence between finiteness, inflection and mechanisms for assigning nominative case.

Furthermore, the occurrence of overt nominative subjects within infinitive clauses indirectly poses a challenge for explaining the distribution of PRO based on its — distinguished — case properties; consequently, it also challenges the complementary distribution between the two available types of null pronouns in a *pro*-drop language: PRO and *pro*. This last topic has been extensively studied for languages with inflected infinitives, as we shall see in the next section.[55]

3.2 Null Subjects in (Portuguese) Inflected Infinitives

3.2.1 Case and Reference for *pro* and PRO

In *pro*-drop languages like Portuguese the question of which kind of null pronoun can occupy the [Spec-TP] position of infinitives carrying morphological agreement features — non-finite non-nominative-marked PRO or finite nominative-marked *pro* — has received wide attention.

As seen in the previous section, an inflected infinitive can host an overt nominative subject. In fact, the licensing of an overt (nominative) subject is closely related to the presence of inflectional morphology: if the infinitive does not bear inflectional morphology, the overt subject is banned.[56] As the following (un)grammatical sentences (59–61) show,

[55]As we explained in the previous chapter, working with data from different languages Landau rejects this widely assumed relation.

[56]There is no complete agreement among scholars on whether the presence of overt inflectional morphology is necessary in ECM constructions like those obtained from perceptive and causative verbs (cf. [Mod07] and [MSL13] for two opposed approaches):

i. Eu vi [as folhas caír(**em**)].
I saw the leaves fall.(INFL.3PL)

the ϕ-features on the infinitive T head seem to be a necessary condition for licensing overt nominative subjects in Portuguese:

(59) O João lamenta [**eles** ter*(**em**) gastado esse dinheiro para nada].
the John regrets they.NOM have.INFL(3PL) spent that money for nothing
'John regrets their having spent that money for nothing.' [RV80]:76

(60) Eu entrei em casa sem [**os meninos** ver*(**em**)].
I went-into the house without the children see.(INFL.3PL)
'I went into the house without the children seeing.' [Rap87]:87

(61) O júri anunciou [não preencher*(**em**) **três candidatos** as condições].
the jury announced not meet.(3.PL) three candidates the requirements
'The Jury announced that three candidates did not meet the requirements.' [**?**]

For this reason, and in particular if one takes case properties to be responsible for differentiating *pro* and PRO, it is assumed that the null subject that alternates with lexical DPs in an inflected infinitive clause is the finite nominative-marked *pro*.

A Note about *pro* in Brazilian Portuguese

There is in fact an exception here: for, although there is wide agreement that *pro* is the null subject of inflected infinitives in European Portuguese, there is no complete consensus among scholars for the case of Brazilian Portuguese. This is due to the fact that referentially-free *pro* is no longer licensed in BP.

Indeed, it has been noted that the distribution and referential characteristics of the null finite subject *pro* in European Portuguese (EP) are different from those of *pro* in Brazilian Portuguese (BP). EP, as well as Italian and Spanish, is a consistent or canonical *pro*-drop language, while BP is a partial *pro*-drop language; in BP *pro* is licensed in more restricted finite contexts compared to EP. It is thought that this is because BP is undergoing a change in the null subject parameter. This ongoing process of change is manifested in the impoverishment of ϕ-features — particularly, the loss for the [person] feature on T — of the finite morphology, the loss of the post-verbal position for the subject, and the propensity towards filling the subject position both in matrix and embedded finite clauses (cf. [**?**]).[57] The following sentences illustrate this last point:

'I saw the leaves fall.'

ii. O juiz mandou [os sem-terra saír(**em**) da fazenda].
The judge ordered the Landless leave.(INFL.3PL) to-the farm
'The judge ordered the Landless to leave the farm.' [San09]:26

[57]BP also shows a degradation of morphology in the case of inflected infinitives.

(62) A Maria disse que comi carne ontem. (?/*BP - ✓EP)
the Mary said that ate.1SG meat yesterday
'Mary said that I ate meat yesterday.'

(63) A Maria disse que eu comi carne ontem. (BP - EP)
the Mary said that I ate meat yesterday
'Mary said that I ate meat yesterday.'

The loss of the null subject *pro* in the subject position of the matrix clause seems to be a consequence of a more general condition on licensing: whenever it is referential, i.e. not arbitrary, the null subject requires a grammatical antecedent; without such an antecedent, the null subject with a referential interpretation is no longer licensed in BP, and *pro* receives an arbitrary interpretation. Hence, while in EP the null matrix subject in the sentence in (64) below is interpreted as a third-person singular pronoun, in BP it is interpreted arbitrarily:

(64) Diz que Dilma vai mudar de estratégia.
say.3SG that Dilma is-going to-change of strategy
a. 'S/He says that Dilma is going to change strategy.' (EP)
b. 'It is said that Dilma is going to change strategy.' (BP)

In order for the subject to obtain the third person singular reading in BP, it has to be overtly realized, as exemplified in (65). In order to obtain the arbitrary reading, a consistent *pro*-drop language like EP uses a clitic, as shown in (66):

(65) Ele diz que Dilma vai mudar de estratégia. (BP)
he says that Dilma is-going to-change of strategy
'He says that Dilma is going to change strategy.'

(66) Diz-se que Dilma vai mudar de estratégia. (EP)
say.3SG-CL that Dilma is-going to-change the strategy
'It is said that Dilma is going to change strategy.'

The null subject pronoun with a non arbitrary interpretation can still be licensed in embedded finite sentences in BP; this notwithstanding, in the Brazilian variation the interpretation of *pro* is not as free as in canonical or consistent *pro*-drop languages. In BP the finite null subject *pro* in embedded sentences can be interpreted either as coreferential with the matrix subject (anaphoric type control) or as coreferential with a matrix topic (variable type control), as exemplified in (67–68) respectively (cf. [FS93]:77):[58]

(67) O José disse que comeu carne ontem.
the Joseph said that ate.3SG meat yesterday

[58]In consistent *pro*-drop languages *pro* picks up its semantic value from a (possibly null) matrix topic: aboutness-shift topic (cf. [Fra07]).

'Joseph$_1$ says that he$_1$ ate meat yesterday.'

(68) A Maria, o José disse que comeu carne ontem.
the Mary, the Joseph said that ate.3SG meat yesterday
'Joseph said about Mary$_1$ that she$_1$ ate meat yesterday.'

Since the interpretation of *pro* in complement finite clauses from BP is dependent on some local, overt DP antecedent, some scholars claim that embedded *pro* is not free in this language; *pro*, as well as PRO, is controlled.

Assuming that (the reference of) *pro* is controlled — by the matrix subject or by a topic — in finite embedded clauses in BP, the road is clear for — controlled — *pro* to be admitted as the subject within inflected infinitive complement clauses, as well (cf. [Nun08]). As the sentence in (69) shows, the overt ϕ-features are not sufficient for licensing the referentially-free null subject *pro*. The null subject of the inflected infinitive clause — be it *pro* or PRO — needs an antecedent despite the overt ϕ-features on the verb. Consequently, if *pro* is the null subject of an inflected infinitive from BP, it has to be assumed that *pro* can be controlled, just like PRO.

(69) *Eu entrei em casa sem [***pro*/PRO** ver**em**].
I went into house without see.INFL.3PL
'I went into the house without anyone seeing.'

(70) Eu$_1$ entrei em casa sem [***pro*$_1$/PRO$_1$** ver].
I went into house without see.INF
'I went into the house without me seeing it.'

In sum, while *pro* is widely accepted as the null subject within inflected infinitive clauses from EP, in the case of BP this assumption depends on having previously accepted *pro* as the null subject within finite clauses. Those who do not recognize the category of controlled *pro* as the null subject within finite embedded clauses in BP, also do not admit *pro* — but rather, PRO — as the subject of inflected (as well as uninflected) infinitive complement clauses (cf. [Mod11]). Since it is widely acknowledged that an overt nominative subject cannot be licensed within uninflected infinitives in both European and Brazilian Portuguese, the question of which kind of null subject occurs within uninflected infinitive clauses also receives an answer by common consensus: namely, PRO.

The assumption of PRO as the subject of uninflected clauses is in itself unproblematic; but, combined with the claim that *pro* is the null subject of inflected infinitives, it gives rise to a somewhat odd picture. In contexts of coreference between the null subject and a matrix nominal phrase, the inflection on the infinitive verb can (but need not) be deleted, i.e. it is optional. In these cases it appears we have to admit the possibility of both PRO and *pro* occurring. Compare the sentences in (71a-b).

(71) a. Também Pedro Machado e Carlos Encarnação$_1$ entendem que o uso
Also Pedro Machado and Carlos Encarnação understand that the use
da base aérea de Monte Real$_2$ por empresas de *low cost*$_3$ é uma boa
of-the base air of Monte Real by companies of low cost is a good
perspectiva, por ***pro***$_3$ estar**em** em franca expansão.
prospect because be.INFL.3PL in clear expansion

b. Também Pedro Machado e Carlos Encarnação$_1$ entendem que o uso
Also Pedro Machado and Carlos Encarnação understand that the use
da base aérea de Monte Real$_2$ por empresas de *low cost*$_3$ é uma boa
of-the base air of Monte Real by companies of low cost is a good
perspectiva, por **PRO**$_3$ estar em franca expansão.
prospect because be.INF in clear expansion
'Also Pedro Machado and Carlos Encarnação understand that the use of the air base of Monte Real by low cost companies is a good prospect because of the marked companies' expansion.' [Van13]:118,120

The alternation between *pro* and PRO that results from the optionality of the inflection is more striking in cases where the coreference between the null subject and a matrix nominal is obligatory, like in some infinitive complement clauses.

3.2.2 Null and Overt Subjects in Complement Clauses

The presence of overt ϕ-features on the infinitive in Portuguese seems to be related to two factors: on the one hand the presence of an overt subject within the infinitive clause, on the other the non-coreference between the (possibly null) infinitive subject and the main finite subject. Consequently, the ϕ-features on the infinitive can be deleted when the infinitive clause does not contain an overt subject or when the (null) infinitive subject is coreferential with the subject of the main clause. Hence, the presence of overt morphology on the infinitive verb is related to the possibility that the clause hosts its own (null or overt) referential subject, which in turn depends on the syntactic independence of the infinitive clause with respect to the finite main clause. Therefore, in the particular case of an infinitive clause that works as a verb complement, i.e., an infinitive complement, the occurrence of the inflection basically depends on the class of verb that selects it.

Infinitive complements selected by propositional verbs, such as *pensar* 'think', *afirmar* 'assert', and factive verbs, such as *lamentar* 'regret', can host a (null or overt) non-coreferential subject, as a result of which the infinitive verb presents inflectional features (see examples in (72–79) below).[59]

[59] In EP, but not in BP, only the so-called Aux-to-Comp word-order is admitted in infinitive compound

(72) A Maria pensa [**os meninos** ter**em** corrido].
the Maria thinks the boys have.INFL.3PL run.PP
'Maria thinks that the boys have run.' [Luz93]:22

(73) A maioria afirma [**eu** ter sido uma mãe na vida escolar].
the most affirms I have.INFL.1SG be.PP a mom in-the life scholar
'Most of them affirm that I was a mom during their student life.' [LM14]:118

(74) Eu afirmo [ter**em** **os deputados** trabalhado pouco].
I affirm have.INFL.3PL the representatives work.PP not-much
'I affirm that the representatives have not worked much.' [Rap87]:87

(75) O cientista acredita [ter**em** descoberto a cura do câncer].
the scientist believes have.INFL.3PL discover.PP the cure of-the cancer
'The scientist believes that they have discovered the cure to cancer.' [Mod10a]

(76) Os cientistas acreditam [ter(**em**) descoberto a cura do câncer].
the scientists believe have.INF.(3PL) discover.PP the cure of-the cancer
'The scientists$_1$ believe that they$_1$ have discovered the cure to cancer.' [Mod10a]

(77) O João lamenta [**eles** ter**em** gastado esse dinheiro para nada].
the John regrets they.NOM have.INFL.3PL spend.PP that money for nothing
'John regrets that they have spent that money for nothing.' [RV80]:76

(78) Os pais da Maria admitiram [morar**em** **eles** nos EUA].
the parents of-the Mary admitted live.INFL.3PL they in-the EUA
'Mary's parents admitted that they live in the EUA.' [Rab04]:67

(79) Os pais lamentaram [ir**mos** ao cinema amanhã].
the parents regretted go.INFL.1PL to-the cinema tomorrow
'The parents regretted the fact that we will go to the cinema tomorrow.' [GSD14]:169

Since the subject of an object control complement is coreferential with the matrix object, the inflection on the infinitive verb is optional when the embedded subject is

complements of propositional verbs. In EP the subject in compound complements of factive verbs can surface in a position other than the post-auxiliary. In more general terms, overt subjects in (simple or compound) infinitive propositional complements must appear in a post-verbal position, while in factive complements the subject can surface in a pre-verbal position. Modesto [Mod11] does not admit overt subjects in inflected infinitives selected by propositional verbs in BP.

i. *O júri declarou [**dois candidatos** ir**em** à final].
The jury declared two candidates go.INFL.3PL to-the final
'The jury declared that two candidates will go to the final.' [Rap87]

ii. O júri lamentou [**poucos candidatos** ir**em** à final].
The jury regretted few candidates go.INFL.3PL to-the final
'The jury regretted the fact that few candidates (would) go to the final.' [GSD14]:165

phonologically null, as exemplified in (80) below. When the controlled subject is overt, as in the examples in (81–82), the inflection is also overt.

(80) Eu obriguei os meninos a [ler(**em**) esse livro].
I forced the boys to read.(INFL.3PL) this book
'I forced the boys to read this book.' [Rab04]:56

(81) Eu exigi aos alunos [**eles** fazer**em** um trabalho].
I forced to-the students they do.INFL.3PL a work
'I forced the student to do some homework.' [Zwa90]:186

(82) Acusa os colegas de [**eles** ser**em** corruptos].
accuse.3SG the colleagues of they be.INFL.3PL corrupt
'She accuses her colleagues of being corrupt.' [Luf10]:34

On the other hand, complement clauses selected by subject control verbs host a subject that is necessarily coreferential with the matrix subject. The sentences in (83–84) below show that neither an overt nor a null independent subject can be licensed in these complements despite the presence of the ϕ-features on the infinitive verb.

(83) *Os pais quiseram [ir**em** **os meninos** ao cinema].
the parents wanted go.INFL.3PL the children to-the cinema
'The parents wanted for the children to go to the cinema.' [GSD14]:164

(84) Os ministros quiseram [suspender(***mos**) a lei].
the ministers wanted suspend.(INFL.1PL) the law
'The ministers wanted for us to suspend the law.'

Since the inflection is deleted in case of coreference between the embedded and the matrix subject, infinitive complements of subject control verbs are supposed not to be inflected.[60] That would also be the case for complements selected by interrogative, implicative, modal and aspectual verbs.

(85) Os pais quiseram [ir(***em**) ao cinema].
the parents wanted go.INFL.3PL to-the cinema
'The parents wanted to go the cinema.' [GSD14]:164

(86) Os meninos tentaram [abrir(***em**) o cofre].
the children tried open.INFL.3PL the safe
'The children tried to open the safe.' [Mod11]

[60] Some scholars admit a partial control reading in complements of subject control verbs. The infinitive inflection is overt when this non-exhaustive reading for the embedded subject is licensed:

i. A Dani$_1$ não quis [**PRO**$_{1+}$ dormir**em** um em cima do outro]. (BP)
the Dani not wanted Pro sleep.INFL.3PL one on top of-the other
'Dani did not want (for them) to sleep one on top of each other.' [Mod10b]:86

Consequently, due to the lack of inflection an overt subject could not be licensed in these kinds of infinitive complements. Nevertheless, as the following sentences show, the inflection is not always deleted in cases of coreference, and controlled (pre- or post-verbal) pronominal subjects can surface in (un)inflected infinitive complements from European and Brazilian Portuguese:

(87) Eles tentaram não [ir**em** todos ao mesmo tempo].
they tried not go.INFL.3PL all.PL.MASC at-the same time
'They tried not to all go at once.' [Mil02]:85 (*apud* [Mod11])

(88) Roberto, eu tentei [**eu** enviar meu convite a você]. (BP)
Robert, I tried I send.INFL.(1SG) my invitation to you
'Robert, I tried myself to send you an invitation.' [Cyr11]:12

(89) Os pais quiseram [ir **eles** de comboio]. (EP)
the parents wanted go.INF they by train
'The parents wanted to go by train.'

(90) Os meninos querem [**eles** embrulhar**em** os presentes]. (BP)
the children want they wrap.INFL.3PL the gifts
'The boys want to wrap the gifts themselves.' [Neg87] (*apud* [Mod10b]:87)

(91) O João decidiu [resolver **ele** o problema]. (EP)
the John decided solve.INF.(3SG) the problem
'John decided to solve the problem himself.' [Bar10]

(92) Não quero [**eu** também ser falso moralista]. (BP)
Not want.1SG I also be.INF false moralist
'I don't want it to be the case that I too am a false moralist.' [Sza09]:35

3.3 Overt Subjects in (Spanish and Italian) Personal Infinitives

The assumption that there is no overt (coreferential) subject within infinitive complement clauses is also endorsed by a number of scholars working on Personal Infinitives from Romance Languages like Spanish and Italian (cf. [PG87]; [Lag87]; [Tor98]; [Sit02]; [Ort02]). Different theories have been proposed in order to explain the overt occurrence of free subjects in infinitive clauses in adjunct position, or in subject clauses [PG87]. Some of these theories are also able to cover cases of overt free subjects in infinitives in the complement position [Riz82]; [Men00]. Nevertheless, they cannot be extended to explain the presence of both controlled pronouns and free overt subjects within infinitive complements from Spanish and Italian in a uniform way (but see [Her15]). That is, those proposals explaining overt free subjects cannot be consistently extended to explain coreferential overt pronouns, *et vice versa*.

Piera [PG87] analyzes overt pronouns in infinitive clauses in subject and verb complement positions from Spanish, like those exemplified in (93–94) respectively:

(93) [Telefonear **tú** primero] sería un error.
phone.INF you first would-be a mistake
'If you phone first, it would be a mistake.' [PG87]

(94) Julia quería [telefonear **ella**].
Julia wanted phone.INF she
'Julia wanted to phone herself.' [PG87]

According to Piera, the overt pronoun within an infinitive clause is not the genuine — semantic and syntactic — subject of the clause, but rather a duplicate of the actual subject PRO. His thesis is based on two assumptions widely admitted in the GB era: i) infinitive clauses from Spanish present the same word-order [NP VP] as finite clauses; ii) the infinitive verb cannot assign the nominative case to the subject; consequently, PRO is invariably the subject of infinitive clauses. According to these assumptions, the structure of the control sentence in (95) above would not be (95a), but (95b):

(95) Julia prometió a Marta [encargarse ella del asunto].
Julia promised to Martha deal.INF-CL she of-the matter
a. [Julia$_1$ prometió a Marta [encargarse **ella**$_1$ del asunto].
b. [Julia$_1$ prometió a Marta [**PRO**$_1$ encargarse ella$_1$ del asunto].

'Julia promised Martha that she would deal with the matter' [PG87]:161

In support of the thesis that the overt post-verbal pronoun within the infinitive clause is not the genuine subject of the complement clause, Piera presents the finite sentence in (96) below, in which the post-verbal pronoun duplicates the referential subject:

(96) Julia telefoneó ella.
Julia phoned she
'Julia phoned herself.' [PG87]:160

Assuming that the referential pre-verbal expression is the semantic and syntactic subject of the finite clause, and since the verb lacks a complement, Piera argues that the overt nominative-marked pronoun *ella* 'she' in (96) above works as a duplicate of the pre-verbal subject *Julia*.[61] Following Ronat [Ron79] and Rizzi [Riz82], Piera regards the pronouns duplicating the — overt or null — syntactic-semantic subject of a — finite or infinitive — clause as emphatic pronouns. These are pronouns having an antecedent in the same clause and standing in a non-argument position (or A'-position) within the clause; they are *adjunct* pronouns. In the particular case of an infinitive clause, PRO operates as the antecedent of the emphatic adjunct pronoun; these pronouns co-occur with PRO, which is their actual antecedent and the genuine subject of the infinitive.

[61] But see [Car99] for a different view.

The two theoretical assumptions underwriting Piera's proposal were revisited in subsequent stages of Generative Grammar. The first assumption concerns the position of the subject with respect to that of the verb: according to Piera, the canonical word-order for a Spanish — finite and infinite — clause is SV (subject + verb). In light of the data from different stages of Spanish this assumption has since been amended. For instance, Mensching [Men00]:25 shows that the prosodic order subject + infinitive is found in infinitive clauses from Old Spanish (see examples in (97–98)) despite the fact that the reverse word-order already prevailed during this stage, and that its frequency is even higher in Classical Spanish.

(97) los quales creerian [**yo** no haber leido las reglas]
the which believe I.NOM not have.INF read.PP the rules
'who would believe that I had not read the rules.' [Men00]:26

(98) Todo por culpa de cosas que pasaron antes de [**yo** nacer].
All by fault of things that happened before of I.NOM be-born.INF
'All because of things that happened before I was born.' [Men00]:152

Starting from 18th century Spanish pre-verbal subjects are almost banned from infinitive environments, while the post-verbal position becomes canonical:

(99) este testigo no afirma [haber **yo** dicho que se podían traer
this witness no assert [have.INF I say.PP that CL could bring
interpretaciones nuevas].
interpretations new]
'This witness did not assert that I said it was possible to bring new interpretations.' [Nav47]:335

In Modern Spanish nominative subjects within infinitive clauses almost always stand in a post-verbal position; the pre-verbal position is only admitted in some very particular adjunct infinitive clauses (cf. [Sit02]; [Sch07]; [VD13]). Such restrictions may be explained by the fact that the licensing of a pre-verbal subject within an adverbial infinitive clause depends on several conditions. These are: the position of the adverbial clause with respect to the main clause, the preposition that introduces the infinitive clause, and the type of nominal phrase (pronominal or referential) that works as the infinitive subject (see the contrast between (101b) and (102) below).

(100) a. *[**Tú** telefonear primero] sería un error.
You.NOM phone.INF first would-be a mistake

b. [Telefonear **tú** primero] sería un error.
phone.INF you.NOM first would-be a mistake

'If you phone first, it would be a mistake.'

(101) a. Al [abrir **yo** los ojos], mi hermano ya estaba cerca.
to-the open.INF I.NOM the eyes my brother already was nearby

b. *Al [**yo** abrir los ojos], mi hermano ya estaba cerca.
to-the I.NOM open.INF the eyes my brother already was nearby
' When I opened my eyes, my brother was already nearby.

(102) Sin [**tú** saberlo], Julián te observaba y te estudiaba.
without you.NOM know.INF-CL Julian CL observed and CL analyze
'Without you knowing, Julian was observing and analyzing you.'

Piera's analysis, however, focuses on overt pronouns within infinitives functioning as the subject or as the complement of the clause; and in these cases, the post-verbal position of the so-called emphatic pronouns is the canonical one. Contra Piera, that is, the typical word-order here is VS.

The second assumption of Piera's proposal, also popular in the GB era, is that infinitive verbs cannot assign the nominative case to their subjects. Nevertheless, scholars working on personal infinitives from Romance languages show that overt subjects in infinitive clauses are nominative-marked. In particular, where the personal infinitive is introduced by a preposition, as in adverbial clauses, the preposition does not license the case of the subject. In effect, as the contrast between (103a) and (103b) below evidences, the overt pronoun does not surface in the oblique but rather in the nominative case.

(103) a. Maria salió de la sala sin [**yo** verla].
Maria goes of the room without I.NOM see.INF-her.CL

b. *Maria salió de la sala sin [mi verla].
Maria goes of the room without I.OBL see.INF-her.CL
'Maria went out of the room without me seeing her.'

If one adopts Piera's thesis, it is always possible to argue that overt nominative anaphoric pronouns in infinitive clauses from Spanish, and also from Italian, are duplicates of PRO. But we can now see that the claim that PRO is invariably the subject of the infinitive clause is itself a consequence of the two generative assumptions that were recently revisited. In addition, the occurrence of a referential expression in the subject position of an infinitive clause — exemplified in (104) below — cannot be explained by Piera's thesis; thus an alternative explanation is called for.

(104) En diversas letras apostólicas se declara [haber sido **Manfredo**
In various letters apostolic CL declare.3SG have.INF be.PP Manfredo
bastardo].
bastard
'In various letters from the apostles it is declared that Manfredo was a bastard.'

We now turn to a different proposal, due to Mensching. In order to demarcate his topic of research, Mensching [Men00]:14 notes that a *prototypical* subject presents the following properties: i) pragmatically, it works as the topic of the clause; ii) semantically, it saturates a semantic θ-role, which in turn is paradigmatically the agent role; iii) morphologically, it is case-marked; indeed it is paradigmatically nominative-case marked; and iv) syntactically, it agrees in person and number with the predicate. Given that the overt pronouns that appear in infinitive complements from Modern Spanish display all of these prototypical characteristics, the question arises of why we should refuse them the status of genuine subject.

Although Mensching adopts from Piera the label *emphatic pronouns* to refer to the overt anaphoric pronouns that occur within infinitive complements from Spanish and Italian, he does not deny them the status of genuine subjects. According to Mensching [Men00]:62 emphatic pronouns are not doubles of PRO, but rather they overtly realize PRO. Thus, like PRO, emphatic pronouns stand in an A-position: [Spec-VP]. Since Mensching adheres to the GB characterization of PRO in terms of [+anaphoric, +pronominal] features, emphatic pronouns are not case-marked.[62] His proposal differentiates overt anaphoric pronouns, which co-occur with PRO, from overt free subjects, which are assumed to occur in the same configuration as *pro*, and consequently are case-marked.[63] Hence, in particular, emphatic pronouns are distinguished from the free nominative-marked pronouns that are licensed in infinitive complements selected by propositional verbs. Thus, whereas the instance of the pronoun *lui* 'he' is emphatic in (105), the one in (106) is not:

[62]From this perspective, it is surprising that they look like nominative-marked pronouns. According to Piera, emphatic pronouns surface in the nominative case as a consequence of being coindexed with a c-command subject.

[63]Nevertheless, following Piera, he also regards as emphatic certain pronouns co-occurring with a null free pronoun and a referential expression in infinitive clauses in subject position, as exemplified in (i–ii) below:

i. [Andarci (**noi**)] sarebbe un errore.
go.INF (we) would-be a mistake
'If we go, it would be a mistake.'

ii. [Telefonear **tú/Julia** primero] sería un error.
phone.INF you/Julia first would-be a mistake
'If you/Julia phone/s first, it would be a mistake.'

These infinitive contexts, where a null subject co-occurs with an (overt or covert) free subject, would justify the [+pronominal] feature on the emphatic pronoun.

(105) Gianni$_1$ mi ha promesso [di farlo **lui**$_1$].
John CL has promised COMP do.INF he
'John$_1$ has promised me that he$_1$ will do it.' [Riz82]

(106) Credevo [essere **lui** arrivato].
thought.1SG be.INF he arrive.PP
'I thought he had arrived.'

In Italian, infinitive complements with a free subject pronoun present the so-called Aux-to-Comp structure. In a nutshell, an Aux-to-Comp structure is a compound-verb clause in which it is assumed that the auxiliary verb (*essere* 'be' or *avere* 'have') in infinitive form is raised to the Comp(lement) position. Structures of this kind are selected by propositional verbs; the assumed movement for the auxiliary verb is triggered by a [+Agr] feature on the C head of this particular type of propositional complement (cf. [Riz82]; [Men00]:144). It is from the Comp position that the auxiliary infinitive verb can assign the nominative case to its post-auxiliary subject (in [Spec-AgrP]).[64] Thus, the [+Agr] feature on C° not only triggers the Aux-to-Comp movement, but also functions as the nominative-case assigner. Besides free overt (i.e. non-local bound) pronouns, the infinitive Aux-to-Comp complement of propositional verbs actually hosts both referential expressions and null free pronouns, as exemplified in the following sentences (107)–(110):[65]

(107) <Giovanni Dondi dall'Orologio>$_1$ (...) L'Abate de Sade$_2$ aggiunge, che il
<Giovanni Dondi dall'Orologio> (...) the-Abbot of Sade adds that the
Petrarca$_3$ afferma [aver **lui**$_1$ avuto il cognome dall'Orologio
Petrarca asserts [have.INF he.NOM receive.PP the surname of-the-Orologio
dal libro che scritto aveva]
of-the book that written had]

[64] Following Pollock [Pol89], Mensching adopts the hypothesis of split-Infl, whereby the node Infl is divided into AgrP and TP projections.

[65] In Italian, like in Spanish, a referential expression within an infinitive complement of a propositional verb can surface in a position other than the post-auxiliary, as exemplified in (i–iii):

i. quello che negli Atti si riferisce [avere detto **l'Apostolo** agli Ateniensi:]
that which in-the Acts CL report have.INF said.PP the.Apostle to-the Athenians
'that which in the Acts it is reported that the Apostle said to the Athenians:' [Il Mediatore 1862]

ii. Orarono, [essere venuti **gli uomini Lombardi** a congratularsi].
prayed.3PL be.INF come.PP.3PL the men Lombardi to congratulate
'They prayed for the Lombards to come to congratulate themselves.' [Bot34]:162

iii. señala, [habérselo referido **el reverendo maestre fray Martin de Córdoba**].
note.3SG have.INF-CL report.PP the reverend master fray Martin de Córdoba
'He noted that the Reverend Master Fray Martin de Córdoba has reported it.' [PyF53a]

This particular position for the referential subject could be explained by appealing to the hypothesis of the so-called heavy DPs. According to Paz [Paz13]:77 the heavier a subject is, the higher the probability that it occupies a post-participle position.

'About Giovanni Dondi, the Abbot of Sade adds, that Petrarca asserts that he had received the surname 'dall'Orologio' from the book that he had written.' [Tir07]:233

(108) il Panciroli, che afferma [aver **Alberto** composto il suo libro in Bologna]
the Panciroli that assert have.INF Alberto compose.PP the his book in Bologna
'Panciroli, who asserted that Alberto composed his book in Bologna.' [Tir83]:227

(109) il quale afferma [aver **Niccolò** dipinto (...) nella cappella].
the who assert have.INF Niccolò paint.PP in-the chapel
'Who asserted that Niccolò has painted in the chapel.' [Fio35]

(110) Quell'altro Bernardino$_1$ che stampava nel 1565 in Vercelli e che *pro*$_2$ suppongo [avere ***pro***$_1$ stampato nel 1577 in Torino]
that-other Bernardino, who printed in-the 1565 in Vercelli and that suppose.1sg have.INF print.PP in-the 1577 in Turin
'That other Bernardino who printed in 1565 in Vercelli and who, I suppose, printed in 1577 in Turin.' [VdF59]:270

Like in Italian, compound infinitive complements selected by propositional verbs from Spanish also allow an overt or a null free pronoun in the post-auxiliary position. As an example, consider the following sentences in (111)–(113):

(111) Digo que este testigo no afirma [haber **yo** dicho que se podían traer interpretaciones nuevas].
say.1SG that this witness not assert have.INF I.NOM say.PP that CL can bring interpretations new
'This witness does not assert that I said that it was possible to bring new interpretations.' [Nav47]:335

(112) <Tancredo>$_1$ Y no falta autor antiguo$_2$ que afirma [haber **él**$_1$ puesto en un monasterio a Constanza].
<Tancredo> and not lack author old that asserts have.INF he.NOM put.PP in a monastery to Constanza
'About Tancredo, there are some old authors that assert that he made Constanza enter a monastery.' [PyF53b]:175

(113) Uno de los cuales se llamó Luso$_1$, y una de las mujeres Lissa$_2$, que dice el mismo Marco Varrón$_3$ [haber ***pro***$_1$ dado el nombre á la parte de Portugal, que antiguamente llamaban Lusitania].
one of the which CL called Luso and one of the women Lissa that says the same Marco Varrón have.INF give.PP the name to the part of Portugal that formerly called Lusitania

> 'One of whom was called Luso, and one of the women, Lissa; and the same Marco Varrón says that Luso gave the name to the part of Portugal once known as Lusitania.' [dM64]:123

Despite the fact that the infinitive complements of propositional verbs from Spanish and Italian exhibit similar word-order, Mensching does not think that the Aux-to-Comp structure licenses (covert or overt) free subjects from the former.[66] On his view, the post-verbal subject in a compound infinitive complement of a propositional verb from Spanish remains in [Spec-VP].[67] This specific proposal for free subjects from Spanish looks a bit odd when we consider his uniform proposal for the case of emphatic pronouns from Spanish and Italian. As we have seen, one of the basic differences between emphatic and non-emphatic subjects in infinitival complements lies in their case properties, which in turn are based on the different [Spec-XP] position they occupy in the structure. In Italian it is claimed, on the one hand, that [Spec-AgrP] is the landing site for non-emphatic (nominal or pronominal) subjects; in this position, they are nominative-case marked by the auxiliary verb on the C head.[68] On the other hand, emphatic pronominal subjects remain in their base position [Spec-VP], and consequently they are not case-marked.[69] Differently from Italian, both emphatic and non-emphatic pronouns in infinitive clauses from Spanish are assumed to stand in the same base position [Spec-VP]. Nevertheless, while non-emphatic subjects are nominative-marked, emphatic pronouns are assumed to be non-case-marked.[70]

Hence, it seems we have two options. We could either assume that infinitive complements with overt free subjects from Spanish instantiate the Aux-to-Comp structure, as in Italian. Or, we could assume that emphatic pronouns, like non-emphatic pronouns, are nominative-marked in Spanish. Whichever we choose, Mensching's proposal cannot explain the occurrence of both free (or non-emphatic) and anaphoric (or emphatic) subjects

[66] According to this author, the question of whether Old and Classical Spanish had infinitival Aux-to-Comp constructions is rather difficult to answer; it is not clear if non-finite Agr° could assign the nominative case under government to its post-auxiliary subject. In Modern Spanish the word-order exhibited by Aux-to-Comp structures is banned (cf. [Men00]:116,147).

[67] It is claimed that the movement of T° (+V°) to Agr° is obligatory (cf. [Men00]:146).

[68] [Spec-AgrP] in (non-Old) Italian is assumed to be an A-position. Thus, movement to this position is an A-movement triggered by the case filter. In Old and Classical Spanish, but not in Modern Spanish, [Spec-AgrP] is an A'-position, and hence, it could function as a landing site only for DPs that are already case-marked (cf. [Men00]:150).

[69] Propositional verbs select either infinitive constructions with an overt complementizer or infinitive constructions with an empty complementizer. When the infinitive complement is headed by a preposition (*di* or *a*), only an emphatic subject, but not a free pronoun, can be licensed. It is assumed that the preposition prevents the auxiliary from being raised to the Comp position. Thus, only the infinitive construction with an empty complementizer allows Aux-to-Comp movement and licenses overt free subjects.

[70] Even the more updated proposal in terms of null case looks enigmatic, since it is assumed that the null case is assigned to the null subject PRO and to the emphatic pronouns by the (governing) T head of infinitives from both Italian and Spanish complements (cf. [Men00]:16). As Mensching claims, whereas nominative assignment to the subject in [Spec-VP] is impossible in non-finite clauses from (non-Old) Italian, the nominative case is assigned by the (governing) T head in Spanish.

in Italian and Spanish in a uniform way. On the one hand, if we admit that the Aux-to-Comp structure licenses free subjects in Spanish, we would still maintain the distinction between emphatic and non-emphatic pronouns. On the other, if we admit that emphatic pronouns can be nominative-marked in Spanish, we must also admit that emphatic pronouns in this language are different from their counterpart in Italian. It is clear that both options bear consequences for the contrast between the two types of null pronouns — PRO and *pro* — and, consequently, for their assumed complementary distribution in Spanish. For, if we follow the first option, we could still maintain the complementary distribution between *pro* and PRO; whereas in the latter case, the complementary distribution becomes blurred.

In fact, not only does Mensching's proposal seem internally problematic; it is also strongly dependent on certain mechanisms and concepts, adopted to explain case marking, which have been revised in more current stages of the generative framework.

Working within the Minimalist Program, Sitaridou [Sit02]:182 explicitly sets the conditions for an infinitive complement clause to license an overt subject. These are: A) having an own tense and B_1) displaying Agr features or B_2) a Comp position filled with overt material. The European Portuguese sentences in (114) and (115) satisfy A and B_1: the tense of the complement clause is different from the matrix tense, and the infinitive verb presents an inflected morphology.

(114) Eu lamento [**os deputados** perder**em** los documentos].
I regret the deputies lost.INFL.3PL the documents
'I regret the deputies' having lost the documents.'

(115) Penso [ter**em** **as crianças** mentido].
think.3SG have.INFL.3PL the child lie.PP
'I think the children have lied.'

The Sardinian sentence in (116) below satisfies the A-condition as well as both of the alternative B-conditions. Indeed, the complement clause selected by the epistemic verb *crédere* 'believe' presents an own tense (condition A), is headed by an overt complementizer *de* (condition B_2), and the auxiliary infinitive verb *essere* 'be' displays morphology for number and person agreement (condition B_1).

(116) Non credo [de ésserе**t** ghiratu **Juanne**].
not believe.1SG COMP be.INFL.3SG return.PP John
'I don't believe that John has returned.'

In contrast with Sardinian, Spanish verb complements are not headed by an overt complementizer and the infinitive verbs do not display agreement morphology, as the next two examples show. Assuming that at least one among the conditions B_1–B_2 is *necessary* for a language to license a personal infinitive in complement clauses, Sitaridou's proposal could then explain why infinitive complements from Spanish do not surface with an overt subject (cf.[LR95]; [Tor98]).

(117) *Odia [jugar **Pablo** a las cartas].
Hate.3SG play.INF Paul to the cards
'He hates for Paul to play cards.' [Tor98]: 210

(118) *Lamenta [perder **ellos** los documentos].
regret.3SG lost.INF they.NOM the documents
'She regrets their having lost the documents.' [Sit02]:189

It is less clear on the other hand why the Italian sentences in (119)–(121) deserve almost no attention in Sitaridou's account as instances of personal infinitives. This appears odd seeing as they are headed by the complementizer *di* — as in Sardinian — and as such would satisfy condition B_2.[71]

(119) Gianni credeva [di essere arrivato **lui**].
John believed COMP be.INF arrive.PP he
'John$_1$ believed that he$_1$ had arrived.'

(120) Maria mi ha chiesto [di parlare **io** con Gianni].
Mary CL has ask COMP talk.INF with John
'Mary has asked me to talk with John.'

(121) Gianni mi ha promesso [di farlo **lui**].
John CL has promise COMP do.INF-CL he
'John$_1$ has promised me that he$_1$ will do it.

Although the complements in Sardinian and Italian display a similar structure, we should note an important difference between them with respect to the semantic value of their overt subjects: namely, while in the former the reference of the overt subject is free, in the latter it is controlled by a matrix nominal. Thus, since Sitaridou does not consider Italian controlled pronouns in great detail, it seems she is mostly interested in overt non-emphatic subjects. In other words, it seems that in talking of *personal infinitives* Sitaridou effectively refers to infinitives with overt nominative *referentially free* subjects. If, as we are suggesting, Sitaridou's monograph is mainly concerned with free (i.e., non-controlled) subjects, this would explain why she does not consider the sentence in (122) as an instance of a personal infinitive from Spanish; nor even mentions the cases of emphatic

[71] There is only a single, brief reference to controlled pronouns from Italian complements: "a personal infinitive in complement position can only be found in languages that — among other things — have a fairly pronounced C-system, where *de/di* introduces a number of control infinitives (e.g. Italian dialects)" (cf. [Sit02]:194).

(or anaphoric) pronouns from Spanish — as exemplified in (123) — despite the similarities to the previous (121) from Italian.[72]

(122) No sabemos [si firmar **nosotros** la carta].
does-not know.3PL if sign.INF we the letter
'We do not know whether to sign the letter.' [Tor96]

(123) Juan me ha prometido [hacerlo **él**].
John CL has promise do.INF-CL he
'John$_1$ has promised me that he$_1$ will do it.'

As we have seen, the instances of overt subjects in Aux-to-Comp complements from Italian, contrary to the overt pronouns that occur in *di*-complement clauses, present a disjunctive reference with respect to the main subject. Although Sitaridou makes reference to Aux-to-Comp complements, she notes that it is questionable whether their subjects are nominative or accusative-marked, as this kind of propositional complement is very similar to the *accusativus cum infinitivo* construction from Latin. Thus, although Aux-to-Comp propositional complements from literary Italian are headed by an overt complementizer — and as such could satisfy conditions A and B_2 — they do not satisfy the very definition of *personal infinitives*: infinitive constructions that involve an overt *nominative* subject but in which no agreement shows up (cf. [Sit02]:18).

Although the overt subjects in propositional complements from Spanish are nominative-marked, and thus should be considered as instances of personal infinitives, they cannot be covered by Sitaridou's general proposal due to the lack of an overt complementizer. In order to explain complements of this kind, which represent a challenge for her generalization, Sitaridou [Sit02]:243 adheres to the thesis that they are learned constructions, and thus are governed by stylistic criteria [Pou95]:

> Given that Old Castilian personal infinitives as complements are not introduced by complementizers (as indeed is the case in Sardinian), it becomes obvious, if not disconcerting, that the above generalization cannot accommodate this type of personal infinitive since it essentially rules them out. At this stage there are two possible options: (a) either the generalization is wrong; (b) the construction is a learnèd construction.

In sum, although Sitaridou presents detailed and seemingly comprehensive conditions for a complement clause to license instances of personal infinitives, her proposal does not give a uniform account of the occurrences of overt free and anaphoric subjects in verb complement clauses from Italian and Spanish. More specifically, on the one hand Aux-to-Comp structures from stylistic Italian cannot be paired with learned personal infinitives

[72] In contrast with Italian, the cases of Spanish pronouns studied by Piera [PG87] occur in a complement clause with an empty head. Note that these differences in clause structure did not prevent Mensching [Men00] from giving a parallel treatment in terms of doubles of PRO.

from non-Modern Spanish because of their different source and, consequently, the different case of their overt subjects. In fact, according to Sitaridou Aux-to-Comp structures are not even instances of personal infinitives. On the other hand, even though Sitaridou recognizes that controlled pronouns from Italian complements are instances of personal infinitives, this claim is not also extended to the case of Spanish overt anaphoric pronouns. Hence, anaphoric pronouns from Modern Italian and Spanish, which were taken to comprise an homogeneous class — of emphatic pronouns — by Piera [PG87] and Mensching [Men00], do not run in parallel in Sitaridou's account.

Anaphoric, controlled or the so-called emphatic pronouns from Italian, Brazilian Portuguese and Spanish (among other *pro*-drop languages) are the topic of Szabolcsi's [Sza09] account, which gives a uniform treatment of the same. Barbosa [Bar09] extends Szabolcsi's proposal to the case of European Portuguese. Their accounts deal only with instances of overt pronouns in complement clauses selected by control (and raising) verbs, which are necessarily coreferential with the main (overt or null) subject. This phenomenon is exemplified by the Italian and European Portuguese sentences in (124)–(126) below:

(124) Gianni$_1$ odierebbe [andare solo **lui**$_1$ a Milano].
Gianni would-hate.3SG go.INF only he to Milan
'Gianni$_1$ would hate it if only he$_1$ went to Milan.' [Sza09]

(125) Anche io odierei [andare solo **io** a Milano].
also I would-hate.1sg go.INF only I to Milan
'I too would hate it if only I went to Milan.'

(126) Detesta [ir só **ele** ao mercado].
hate.3SG go.INF only he to-the market
'He hates it if only he goes to the market.' [Bar09]:99

The first goal of Szabolcsi and Barbosa's proposals is to show that overt pronouns are the very same semantic-syntactic subjects of the infinitive embedded clause and not, as previously argued by Piera and Rizzi, adjuncts to or doubles of the null subject PRO. Their second goal is to explain how the overt pronominal subjects can be licensed in uninflected infinitive controlled clauses (in some *pro*-drop languages, but not in others). Thus, a mechanism for long distance agreement and a multi-agreement parameter are called upon to deal with overt controlled pronouns (cf. [Sza09]:6):

Long-distance agreement A sufficient condition for nominative subjects in infinitival complements to be overt is if the relevant features of a superordinate finite inflection are transmitted to the former.

Multi-agreement parameter Languages vary as to whether a single finite inflection may share features with more than one nominative DP.

Barbosa focuses on the correlation, already noted by Szabolcsi, between the *pro*-drop and the multi-agreement parameters: consistent *pro*-drop languages admit an overt pronominal subject within uninflected infinitive control complements.[73]

These proposals present evidence for the claim that an overt anaphoric pronoun can be the real subject argument of a controlled uninflected infinitive clause. Thus, they depart from the traditional generative assumption that only PRO could be the infinitive subject. Since these scholars adopt a generative framework, they are themselves concerned with the theoretical devices that can explain the licensing of nominative-marked subjects in a clause whose T head does not present overt agreement morphology. In a nutshell, overt embedded nominative pronouns are licensed through an Agree relation with the matrix inflection. Given the mechanism assumed in their explanations, Szabolcsi and Barbosa can only account for the occurrences of overt pronouns that are coreferential with the matrix subject. Moreover, their accounts fail to encompass either cases of object controlled clauses or cases of free subjects, like those licensed in uninflected infinitive clauses selected by propositional verbs from stylistic Italian and Spanish.[74] Hence, in these as in the other generative proposals reviewed in this chapter, it seems that overt nominative anaphoric pronouns in controlled complements and free nominative subjects in propositional complements are not licensed through a uniform mechanism.

[73]The key features of her proposal are that, first, the finite and infinitival agreement affixes work as a pronominal clitic; secondly, that the agreement affix in the matrix clause establishes a long distance Agree relation with the *in situ* subject of the embedded clause. A finite clitic checks the EPP in the matrix clause and an infinitival clitic checks the EPP in the embedded clause. Since the EPP feature is checked by the affix, it is assumed that a pre-verbal subject in the main clause occupies, not [Spec-TP], but a position in the left periphery of the matrix clause. Thus, the embedded subject does not need to raise to [Spec-TP] and may stay in its base (post-verbal) position [Spec-VP]. The set of ϕ-features on finite T is valued and interpretable; the set of ϕ-features on infinite T is also interpretable but is unvalued. Thus, whereas ϕ-features on finite T work as a probe, and they look for an active goal — the *in situ* subject — to enter into an Agree relation, those of the infinite T work as an active goal. These active ϕ-features are valued through an Agree relation with the ϕ-features on matrix T (or v), and thus they are co-indexed with those of the antecedent. The Agree relation between (finite or infinite) T and its argumental (null or overt) subject on [Spec-VP] is mediated by the agreement affix.

[74]In fact, Szabolcsi [Sza09]:18 argues that the fact that object control verbs do not license overt nominative subjects in Hungarian gives support to the long distance agreement (with the matrix finite verb) mechanism.

4 Proposal: Overt Infinitive Subjects in Generative Linguistics[75]

As we observed in the previous chapter, some generative proposals ascribe the status of subject to the overt anaphoric pronouns occurring in the post-verbal position in uninflected infinitive clauses from Contemporary Italian and/or Spanish ([Men00]; [Sit02]; [Sza09]). We also saw that uninflected infinitive clauses selected by propositional verbs in 18-19th century Spanish and in literary Italian can host an overt non-anaphoric (post-verbal) subject (cf. [Riz82]; [Men00]). Therefore, the overt pronouns in the following complement clauses would all be instances of infinitive subjects, despite the fact that some of them display a kind of controlled reference (sentences in (127–131)) and others are certainly referentially free (examples in (132–133)).

(127) Saez se lamenta [haber sido **él** la nueva víctima de este acto vandálico].
Saez CL regrets have.INF be.PP he the new victim of this act hooliganism
'Saez regrets having been the new victim of this act of hooliganism.' [wwwa]

(128) Non vuole [andare solo **lui** a Milano].
not want.3SG go.INF only he to Milan
'He does not want to be the only one going to Milan.' [Sza09]

(129) Gianni ha deciso [di consegnare **lui** il pacco a Maria.]
John has decided COMP deliver.INF he the package to Mary
'John decided to deliver the package to Mary himself.' [Car99]

(130) No niego [haber **yo** ido con intención de hacérsele como él merece].
not deny.1SG have.INF I go.PP with intention of have-CL as he deserve
'I do not deny that I went with the intention of giving him what he deserves.' [DO38]:532

(131) E Sant'Agostino (...) afferma, [aver **lui** veduti molti Monasteri in Roma].
and Saint'Augustine (...) asserts have.INF he see.PP many monasteries in Rome
'And Saint Augustine asserts that he saw many monasteries in Rome.' [Bia46]:336

(132) Este documento prueba [haber **tú** nacido en 1938].
this document evidence have.INF you be-born.PP in 1938
'This document evidences that you were born in 1938.' [Van13]

[75] Previous versions of this chapter were presented at *XXIX Congreso Internacional de la Asociación de Jóvenes Lingüistas*, Murcia-Spain; *Going Romance 2014*, Lisbon-Portugal and *I Colóquio de Semântica Referencial*, São Carlos-Brazil.

(133)

il	quale	afferma	[aver	**noi**	per	guida	S. Agostino].
the	which	affirms	have.INF	we	for	guide	S. Augustine

'who affirms that we have Saint Augustine as our guide.' [Puj80]:11

These sentences from Italian and Spanish thus behave much like sentences (134–136) below from Portuguese. The main difference between them is that the infinitive verb appearing in the propositional complements from the latter may present inflected morphology (see examples (134) and (136)).

(134)

A	polícia	forçou	os	manifestantes	a	[(todos)	**eles**	saír(**em**)].
the	police	forced	the	protesters	to	(all)	they	leave.INFL.3PL

'The police forced all of the protesters to leave.' [Rab04]:49,57

(135)

O	João	decidiu	[resolver	**ele**	o	problema].
the	John	decided	solve.INF	he	the	problem

'John decided to solve the problem himself.' [Bar10]

(136)

Penso	[ter**em**	**as crianças**	mentido].
think.1SG	have.INFL.3PL	the children	lie.PP

'I think that the children have lied.' [Sit02]:185

Despite these similarities, the literature to date offers separate, language-specific accounts of these linguistic structures. We maintain that it is not merely desirable but in fact possible to give a uniform and unified account of infinitive subjects from these languages, whether referentially free or anaphoric. In this chapter we adopt Landau's theory of control [Lan00], [Lan04] to achieve precisely this goal. We start in Section 4.1 by showing that Landau's analysis can be adapted to account for infinitives with referentially free subjects in literary Italian and 19th century Spanish; we then show how this proposal can be extended to the case of inflected infinitives from Portuguese, in Section 4.1.1. Finally, we show that infinitives with controlled pronominal subjects may similarly be treated within the same framework.

In two footnotes, Landau observes that infinitive complements from literary Italian, which present the Aux-to-Comp structure, represent a challenge for his theory. He suggests analyzing them by ascribing a [+Agr] feature on the I head, although he does not then develop this idea any further (cf. [Lan04]:868, [Lan06]:167). However, as Landau himself observes, this suggestion is not consistent with his overall proposal. Effectively, the mapping between [Agr] features and morphological spell-out would be lost if uninflected infinitive verbs were to induce the presence of a [+Agr] feature on I°.

Here we propose and develop an alternative solution route. We will adopt Landau's analysis for F-subjunctives in Balkan languages, and show it can be extended to the case of infinitive complements from 19th century Spanish and literary Italian. We call these F(ree)-infinitives; note that F-infinitives include the so-called Aux-to-Comp complements from

Italian. We then show that this kind of infinitive complement can be fully subsumed under the general framework of Landau's theory; this will allow us to overcome his challenge in a way that is consistent with all the basic assumptions of his T/Agr calculus. As the first and main step towards this goal, in the next section we shift the focus to the features of C°, and argue for our first thesis: that *F-infinitives bear [+T, –Agr] features on C°*.

4.1 F(ree)-Infinitives

The infinitive clauses we call F-infinitives are selected by propositional — declarative or epistemic — verbs only, such as *decir/dire* 'say', *señalar* 'mark', *afirmar/affermare* 'affirm', *pensar/ritenere* 'think', *probar* 'demonstrate', *credere* 'believe', *recordar/ricordare* 'remember', *declarar* 'declare', among others. Propositional verbs select an embedded clause that allows a tense mismatch with deictic (as well as non-deictic) adverbials. The infinitive embedded clause can be simple, or it can be compounded by an auxiliary (infinitive) verb. Simple infinitives selected by propositional verbs describe an event which overlaps with the event or state described by the matrix clause, as exemplified in (137) below.

(137) Dice/Ricorda di [lavorare a un progetto segreto].
say/recall.3SG COMP work.INF in a project secret
'He says/recalls that he works on a secret project.'

Infinitives selected by propositional verbs in 19th century Spanish and literary Italian are compound infinitives. As we saw in Section 2.2.5, an infinitive compound clause describes an event which is temporally located before the event of the matrix clause, as the infinitive auxiliary verb induces a shift to the past from the matrix tense. Infinitive complements of propositional verbs in these languages carry a tense that is dependent on, but not identical to, the matrix tense. Hence any temporal adverb appearing in the infinitive clause should respect this dependence:

(138) Ayer dijo [haber ido al médico el mes anterior/*hoy].
yesterday say.3SG have.INF go.PP to-the doctor the month previous/*today
'Yesterday he said he went to the doctor the previous month/*today.'

Recall that according to the rule for the [T] feature (Sec. 2.2.5, rule 1), an embedded clause carrying a dependent tense bears a [+T] feature on its C and I heads. In light of the above, we see that this condition is also met by F-infinitives; since we are working within the T/Agr framework ourselves, we may therefore conclude that both heads of an F-infinitive bear a [+T] feature, taking us one step closer to our desired conclusion.

A second theoretical reason supporting our thesis that F-infinitives carry [+T, –Agr] features comes from Landau's own analysis of F-subjunctives in Balkan languages.

As the reader will recall, the latter exemplify the parasite-host relationship established between [+T] and [+Agr], according to which [+T] on Comp may "parasitically" introduce [+Agr] (cf. [Lan06]:163). In Landau's analysis, F- and C-subjunctives bear mismatching [T] values as a result of the landing position for the embedded (finite) verb. Further, in F- but not C-subjunctives, the embedded verb overtly raises to the C head (cf. [Var93]). By raising the inflected subjunctive verb to the Comp position, a dependent tense is licensed in F-subjunctives; thus, the C head of these complements bears a [+T] feature. Now since a movement is not optional, but triggered by an uninterpretable feature [Cho95]; and since, according to Landau, [+T] is parasitic on the uninterpretable [+Agr] feature on C°, it seems we would have to assume that what actually forces the verb movement up to the Comp position in F-subjunctives is a [+Agr] feature on C°. Put more concisely, it would appear that the dependent tense exhibited by F-subjunctives is an indirect consequence of the verb movement to the C position, which in turn is a direct consequence of the [+Agr] feature on this head. We now test this hypothesis for the case of F-infinitives.

In the so-called Aux-to-Comp complements from stylistic Italian, it is the uninflected auxiliary verb that moves up to the Comp position. It is widely accepted that Aux-to-Comp structures depend on the properties of C° (cf. [Riz82]; [RR89]; [Rap89]; [Men00]). As we have just noted, these Italian complements express a temporal dependency with respect to the main clause, much like F-subjunctives from Balkan languages. Thus a seemingly plausible hypothesis is that Aux-to-Comp complements also bear a [+Agr] feature on C° (cf. [Riz82]; [Men00]). In contrast to F-subjunctives, however, the verb that raises to the Comp position in Aux-to-Comp structures is uninflected. For this reason, and since the criterion for selecting the [Agr] feature on the I head is a morphological one, I° comes to bear a [−Agr] feature. Following this line of reasoning, an Aux-to-Comp structure would thus be characterized by the feature set [+T, +Agr] — and therefore [+R] — on C°, and [+T, −Agr] — and therefore [−R] — on I°. However the derivation for such a structure, carrying conflicting [R]-values, could converge only if a null subject PRO were to check off the [−R] feature on I°, and the matrix F head were to check off the [+R] feature on C° (cf. [Lan04]:840). In other terms, the hypothesis that Aux-to-Comp complements bear [+Agr] on C° would entail that a DP/*pro* — [+R] — subject should be barred from these same complements, since it could not check the [−R] feature on the I head. Given that the chief distinguishing characteristic of Aux-to-Comp complements is to allow for referentially free subjects, we are forced to reject that hypothesis after all. We are thus led to conclude that positing an [Agr] feature on C° does indeed explain the verb raising phenomenon; in which case it follows that the value for this feature in Aux-to-Comp structures is negative, contrary to expectations.

To recap: by adapting the analysis given for F-subjunctives to the case of F-infinitives, we argued that it is the [−Agr] feature on C° that forces the movement of the infinitive auxiliary verb up to the Comp position in these Italian complements. By raising the verb to the Comp position, the dependent tense — [+T] — is licensed despite the negative value for the [Agr] feature on C head of the F-infinitives. In sum, we claim that what triggers the verb movement is the uninterpretable [Agr] feature on the C head, independently of its specific value: [+Agr] in F-subjunctives, [−Agr] in F-infinitives.[76]

It is important to note that the feature combination [+T, −Agr] on C° is not only consistent with the Agr on C° rule in Sec. 2.2.5, which establishes a dependency relationship between the [Agr] and [T] features on C°; it is also a logical consequence of its second clause. For ease of reference, we reproduce this rule in 9 below:

9. a. [+Agr] ⇒ [+T];
   b. otherwise ⇒ [$\emptyset_{Agr}$]. [Agr on C° rule]

In effect, according to the second clause — or at least the most plausible interpretation of the same — both [$\emptyset_T$] and [−T] imply [$\emptyset_{Agr}$] on C°. By contraposition it should then follow that [+Agr] and [−Agr] both imply [+T] on C°. In other words, [+T] on C° is not merely parasitic on [+Agr], but on both [±Agr] on C°. Once again, contrary to what seems to be the received view, we come to the conclusion that the C head of an infinitive complement clause selected by a propositional verb can bear the [+T, −Agr] feature combination.

Having shown that our thesis enjoys the required support in the case of stylistic Italian, we now show that it can be easily extended to the case of 19th century Spanish; the case of Portuguese is deferred to the next subsection. As we saw in Section 3.3, propositional verbs from 19th century Spanish can also select a compound infinitive complement with an overt or null free nominative subject. These infinitive complements exhibit the auxiliary-subject-participle word-order typical of Aux-to-Comp structures.[77]

(139) le declaró [haberte **Dios** permitido aquella enfermedad por ciertas faltas].
CL.DAT declared.3SG have.INF-CL God permit.PP that disease because certain faults

[76] An anonymous review has claimed that we should present an empirical and independent reason for assuming this negative value for the [Agr] feature on C°. First, we have to recognize that unfortunately we have no such reason to offer. Secondly, we want to clarify that our proposal only intends to be a uniform, theoretical model coherent with all of Landau's assumptions. The review also questions why we do not assume [+T] as the feature that triggers the verb movement. We answer that in Landau's proposal [+T] is a necessary condition, not a sufficient one, for the [+Agr] feature. If [+T] were sufficient for the verb moving to C°, then the feature combination [+T, $\emptyset_{Agr}$], which Landau uses to explain partial control, would not be admitted. In addition, we observe that our proposal for Aux-to-Comp structures is in accordance with other minimalist proposals. In effect, in order to adapt the GB analysis of Aux-to-Comp structures in minimalist terms, Mensching [Men00]:187 reinterprets [+Agr] features as strong [ϕV] features on C°. When a complementizer with strong [ϕV] features is selected, the verb raises overtly to C°. Note that Sitaridou [Sit02]:182ss also assumes that what triggers the obligatory T-to-C movement in propositional complements is not an uninterpretable [T] feature, but an uninterpretable F feature similar to a *v* feature that triggers Aux-to-C movement in questions in English.

[77] Nevertheless, in that section we noted that an R-expression usually occupies a post-participle position in an infinitive complement clause.

'He declared to him that God permits you that disease because of certain faults.' [Paz13]:176

(140) este testigo no afirma [haber **yo** dicho que se podian traer
this witness not affirms have.INF I say.PP that CL could bring
interpretaciones nuevas].
interpretations new
'This witness does not affirm I said that new interpretations could be brought.' [Nav47]:335

(141) la conversación en que asegura [haber **yo** dicho que él habia
the conversation in which assure.3SG have.INF I say.PP that he had
usurpado...]
usurped...
'The conversation in which he assures I said that he had usurped...' [Oli20]:2

It is important to note that this word-order—auxiliary verb+subject+participle—is found in Classical and 19th century Spanish clauses other than infinitive complements, such as finite interrogative clauses (142–143), gerundial clauses (144) and in infinitive clauses in adjunct position (145):

(142) ¿No han vuestras mercedes leido?
not have.3PL your grace read.PP
'Has Your Grace not read?' [dCS71] (*apud* [Men00]:147)

(143) ¿Has tú creído en mí deshonra?
Have.1SG you believe.PP in my dishonor?
'Did you believe in my dishonor?' [CM53]

(144) habiendo él poco antes con un azote echado los mercaderes del
Have.GRD he shortly before with a scourge chase-out.PP the merchant of-the
templo.
temple
'having chased the merchant out of the temple with a scourge shortly before.' [Zár48]:606

(145) él solo ha sufrido la pena por nosotros sin [haber **él** cometido
he only has suffered the pain for us without have.INF he.NOM commit.PP
demérito].
demerit
'he only suffered the pain for us without having committed any demerit.' [Ref58]:449

This evidence would suggest that compound infinitive complements with overt subjects in earlier stages of Spanish are similar to stylistically marked complements from Italian. As a result, it would be consistent to assume that the T-to-C movement is admitted in Spanish as it was in Italian (cf. [Paz13]:75,178).[78] According to the foregoing, we conclude that — finite or infinitive — verb movement to the C head is triggered by an uninterpretable [Agr] feature on C°. Taking into account our previous reasoning for the case of Italian F-infinitives, we are led to the conclusion that the C head of F-infinitives from 19th century Spanish also bears an [−Agr] feature.

We now show that this proposal can be consistently extended to deal with inflected F-infinitives from Portuguese. Along the way, we will also discuss the issue of how a [+R] subject is licensed in personal F-infinitives from all three languages.

4.1.1 Inflected F-Infinitives

As we noted in Section 3.2.2, Portuguese infinitive propositional complements can also host an overt or null non-anaphoric subject. In cases where the free subject is overt, the infinitive verb exhibits overt inflectional morphology. Although factive complements from Portuguese can also host an overt non-anaphoric subject, in propositional infinitives the overt subject must occupy the post-verbal position. As a consequence, propositional compound infinitives from (European) Portuguese, but not factive infinitives, necessarily display the Aux-to-Comp word-order [Rap87].[79] Like the auxiliary verbs *haber/avere* in Spanish and Italian respectively, the auxiliary verb *ter* 'have' used in compound infinitive complements from Portuguese induces a shift to the past from the matrix tense; these (inflected) complements thus describe an event which is temporally located before the matrix event. That (un)inflected infinitive complements of propositional or factive verbs from Portuguese bear a dependent tense, is a particular case of a more general fact. Indeed, based on data from different languages, Landau [Lan00]:6, [Lan04]:835s and [Lan06]163 draws a very general conclusion: the C head of propositional, factive, interrogative and desiderative (simple or compound) infinitive complements bears a [+T] feature because of the presence of a tense operator.

Surprisingly — despite explicitly acknowledging the universality of this thesis — Landau does not assume it for his own analysis of propositional infinitives in (European) Portuguese. Instead, here he posits a distinction between the tense of factive and propositional complements: factive predicates would select complements with dependent tense, while propositional complements would bear an independent tense [Lan04]:850s. Thus, given the T on C°/I° rule, and given this alleged disparity among tenses, he is forced to assume an (optional) [+T] feature on the C head of factive infinitives, and [$\emptyset_T$] for propositional infinitives.

[78] Following McFadden [McF12], Paz argues that verb movement in infinitive complements from Spanish is triggered by a prosodic restriction: the left edge of an intonation phrase cannot be empty.

[79] We have already observed there is no consensus on licensing overt subjects in inflected propositional infinitives from Brazilian Portuguese.

Landau's reasons for assuming a divergent analysis for the specific case of Portuguese inflected infinitives remain unclear to us. First of all, it is not clear why propositional and factive complements from Portuguese would be an exception to his own generalization: both propositional and factive infinitive complements bear a dependent tense. Secondly, it is also not clear why the [+T] feature on C head of factive complements would be optional. Does an optional [+T] feature mean that factive verbs can select a clause bearing either an independent or a dependent tense, and consequently a clause with either a [+T] or a [$\emptyset_T$] feature on its C head? Or does it mean that the adverbial test is adequate to distinguish between an anaphoric and a non-anaphoric tense, but not enough to decide between independent and dependent tenses? Third, it is unclear how the possibly different [T] feature on the C head of factive and propositional infinitives can be used to explain the different word-order in these Portuguese compound complements. Since it is claimed that the [+T] feature is optional in factive complements, it seems we have to assume that the [$\emptyset_T$] feature on the C head, and thus [$\emptyset_{Agr}$], trigger the typical Aux-to-Comp word-order which is optional in factive, but obligatory in propositional complements from Portuguese. Indeed, it makes sense to think that if the verb movement is optional, so is the uninterpretable feature that induces it. But once again, this conclusion does not fit well into the general T/Agr theory of control. Since the verb movement is triggered by an uninterpretable feature on the C head, it is not clear why the auxiliary inflected verb in factive and propositional complements would raise to a C position bearing no [T, Agr] features.[80]

As we saw in the course of reviewing the T/Agr calculus, F(ree)-subjunctives in Balkan languages also bear a dependent tense. According to Landau, the dependent tense exhibited by F-subjunctives is a consequence of the verb movement to the [+T, +Agr] Comp position. Thus, it seems that at least one of [+T], [+Agr] must be displayed in C° in order not only to explain the semantic dependent tense of the clause, but also to induce the verb raising to the Comp position. For this reason, in order to explain the Aux-to-Comp propositional inflected F(ree)-infinitive from Portuguese it seems we have to assume some of these uninterpretable features on its C°. Hence, departing from Landau, we assume that the uninterpretable feature which is optional in factive infinitives but necessary in propositional complements is not [+T], but [+Agr]. That is, we assume that propositional verbs select a complement with a [+Agr] feature on C°, while factive verbs select a complement with either the [+Agr] or the [$\emptyset_{Agr}$] feature. Assuming [+Agr] on C° in propositional complements and (possibly) [$\emptyset_{Agr}$] on C° in factive complements we can explain why the inflected auxiliary verb is necessarily moved up to C° only in propositional complements. In our proposal the [+T] feature is assumed on the C head of both propositional and factive infinitive complements. Thus, as per Landau's general

[80]Although Landau [Lan04]:851 notes the possibly different word-order in propositional and factive infinitive complements, he does not devote any further attention to this fact.

proposal, both propositional and factive complements bear dependent tense. Hence, our proposal for inflected infinitives from Portuguese turns out to be consistent with all of Landau's previous analyses of non-controlled clauses. Furthermore, our proposal for Portuguese is also consistent with our own account for propositional infinitives from Italian and Spanish: the different value of the [Agr] feature on the C head for Portuguese propositional *inflected* F-infinitives on the one hand, and for Italian and Spanish *uninflected* F-infinitives on the other, simply depicts a difference in the inflectional morphology of the auxiliary verb.

Now, the difference in the [Agr] feature on the I and C heads of these infinitive complements results in a difference in the uninterpretable [R] feature. In effect, according to the rule in 5 in Sec. 2.2.5, only the [+T,+Agr] combination implies a [+R] feature on either C° or I°. We summarize our proposal in more schematic terms:

10. Inflected propositional F-infinitives (EP):
    $[_{CP}$ propositional-V $[_{CP}$ $C^{\circ}_{[+T,+Agr,+R]}$ $[_{IP}$ $I^{\circ}_{[+T,+Agr,+R]}$ $[_{VP}$ V° $\ldots]]]]$
11. Uninflected propositional F-infinitives (IT/ES):
    $[_{CP}$ propositional-V $[_{CP}$ $C^{\circ}_{[+T,-Agr,-R]}$ $[_{IP}$ $I^{\circ}_{[+T,-Agr,-R]}$ $[_{VP}$ V° $\ldots]]]]$
12. Inflected factive F-infinitives (EP):
    a. $[_{CP}$ factive-V $[_{CP}$ $C^{\circ}_{[+T,+Agr,+R]}$ $[_{IP}$ $I^{\circ}_{[+T,+Agr,+R]}$ $[_{VP}$ V° $\ldots]]]]$
    b. $[_{CP}$ factive-V $[_{CP}$ $C^{\circ}_{[+T,\emptyset_{Agr}]}$ $[_{IP}$ $I^{\circ}_{[+T,+Agr,+R]}$ $[_{VP}$ V° $\ldots]]]]$

Given the presence of [+T,+Agr] on the C and I head of inflected infinitives from Portuguese, it is clear that a [+R] subject can be licensed. Nevertheless, a [+R] subject can also be licensed in uninflected infinitives. In effect, given that C° and I° in uninflected propositional F-infinitives have the same set of features, namely [+T,−Agr,−R], the two derivational paths used to check off the uninterpretable [R] feature in F-subjunctives can be adapted to the case of these F-infinitives. Thus, there are two possible Agree relations to check off the uninterpretable [−R] on I°: it can be checked off by the [−R] feature on the embedded subject, or it can be checked off by the [−R] feature on C°. If I° enters in an Agree relation with the [−R] feature on PRO, an obligatory control structure is obtained, as delineated in (13) below.[81] If I° enters in an Agree relation with the [−R] feature on C°, a no control structure is derived, as delineated in (14). Effectively, because the [−R] uninterpretable features on both of the functional heads are mutually checked off, a referentially free [+R] subject may be licensed.

13. Uninflected F-infinitives: [–R] subjects

[81]PRO can be a goal for the Agree operation because of its anaphoric [−R] feature (cf. [Lan04]:843).

$$[_{CP}\ C^{\circ}_{[+T,-Agr,-R]} \quad [_{IP}\ PRO_{[-R]} \quad [_{I'}\ I^{\circ}_{[+T,-Agr,-R]} \quad [_{VP}\ t_{PRO[-R]} \quad]]]]]$$

Agree [–Agr,–R]

Agree [+T,–Agr,–R]

14. Uninflected F-infinitives: [+R] subjects

Agree [-R]

$$[_{CP}\ C^{\circ}_{[+T,-Agr,-R]} \quad [_{IP} \quad [_{I'}\ I^{\circ}_{[+T,-Agr,-R]} \quad [_{VP}\ DP/pro_{[+R]} \quad]]]]]$$

Agree [+T,–Agr,–R] *Agree* [–Agr]

These two derivational paths predict an alternation between referentially free and controlled subjects for F-infinitives from Classical Spanish and stylistic Italian. Thus, in particular, when the subject is phonologically null it can be *either PRO or pro*:

(146) <Tancredo> … <autores sicilianos> … y hay alguno es de opinión
<Tancredo> … <writer sicilian> … and there-is some is of opinion
[haber ***pro*** sido nieto del rey Guillelmo].
have.INF *pro* be.PP grandson of-the king Guillelmo
'Some Sicilian writer considers that Tancredo was the grandson of the King Guillelmo.' [Cuesta 1853]

(147) L'amico credo [essere ***pro*** arrivato puntualmente].
the-friend believe.1SG be.INF *pro* arrive.PP on-time
'About the friend, I think that he has arrived on time.' [Men00]:192

(148) Credo [essere **PRO** arrivato puntualmente].
believe.1SG be.INF PRO arrive.PP on-time
'I believe that I have arrived on time.' [Men00]:192

(149) Ítem declaró [haber **PRO** oído á algunos estudiantes, que no se acuerda quiénes son, que …].
Item declared have.INF PRO hear.PP from some students, that not CL remember.3SG who are, that …
'Item$_1$ declared that he$_1$ heard from some students, which he$_1$ does not remember who they are, that …' [LMyS99]:xviii

We have thus achieved the first of our goals: we gave an analysis for infinitive complements of propositional verbs from 19$^{\text{th}}$ century Spanish and literary Italian that does not assume the [+Agr] feature on the I head. Therefore, our analysis is fully consistent with the basic assumptions of Landau's calculus. To complete our proposal, we now turn to the case of Controlled infinitives.

4.2 C(ontrolled)-Infinitives

The data presented in Ch. 3 shows that infinitive complements from Contemporary Spanish and non-literary Italian cannot host an (overt nor null) referentially free subject; they can only host an (overt or null) controlled, and thus pronominal, subject. In addition, the contrast between propositional infinitive complements from Italian (cf. (150–151) below) suggests that the kind of subject that can be licensed depends on the material that fills the Comp position: a complementizer or a verb. Here, the overt complementizer *di* exclusively appears with (overt or null) anaphoric pronouns, while the auxiliary infinitive verb appears with a referential subject.

(150) Credevo [di aver vinto **io**].
believe.1SG COMP have.INF win.PP I
'I believed that it was me who had won.'

(151) Credevo [essere **lui** arrivato ieri].
believe.1SG be.INF he arrive.PP yesterday
'I believe that he had arrived yesterday.'

In order to maintain the parallelism between the subjunctive complements from Balkan languages and these infinitive complements from Romance languages, we baptize the complements licensing only an anaphoric subject by the name of C(ontrolled)-infinitives.

In our analysis of F-infinitives we argued that a propositional verb selects a CP with a [−Agr] on its head; due to this fact, the infinitive verb is thus located in the Comp position. We will now argue that infinitives headed by an overt complementizer *di* bear [$\emptyset_{Agr}$] on C°.

To begin with, observe that in some varieties of Contemporary Spanish the Comp position of an infinitive compound clause selected by a propositional or a factive verb can be filled by an overt particle *de*, similar to the *di* complementizer from Italian (cf. [DT11]:185). Note that following [Luj80], we assume that *de*, as it is used in these sentences, is not a real preposition but rather a functional element.

(152) Juan lamenta [de no haber ido al cine más a menudo].
John regrets COMP not have.INF go.PP to-the cinema more to often
'John regrets not going more often to the cinema.' [Ber13]:22

(153) Mi sono lamentato [di partire domani].
CL be.1SG complain.PP COMP leave.INF tomorrow

'I complained about leaving tomorrow.'

(154) Agustí afirma [de estar sorprendido].
August asserts COMP be.INF surprise.PP
'August asserts that he is surprised.' [@pl13]

(155) La polizia afferma [di aver ritrovato una mia impronta digitale].
the police claim COMP have.INF find.PP one my print finger
'The police claim to have found one of my fingerprints.'

Indeed, the same holds for simple infinitive complements selected by control verbs, in this variety of Spanish as in Italian:

(156) José trató [de leerlo].
Joe tried to read.CL
'Joe tried COMP read it.' [Nis13]

(157) No permito a mis hijos [de llegar tarde].
not permit.1SG to my children COMP arrive.INF late
'I do not allow my children to be late.' [Ber13]:17

(158) Permettetemi [di presentarvi il mio caro amico].
allow-me COMP introduce.INF-CL the my dear friend
'Allow me to introduce my dear friend.'

(159) Televisa intenta [de salir de su propia caldera del diablo].
Televisa tries COMP escape.INF of its own cauldron of-the evil
'Televisa tries to escape its own cauldron of evil.' [www99]

(160) Cercava [di persuadermi a partire].
try.3SG COMP persuade.INF-CL to leave
'He was trying to persuade me to leave.'

(161) <Bonifacio III> Prohibió [de ocuparse de la elección del nuevo Papa].
<Bonifacio III> prohibited COMP take-on.INF of the election of-the new Pope
'Bonifacio III prohibited taking on the election of the new Pope.' [wwwb]

(162) Ti proibisco [di uscire].
CL forbid.1SG COMP go-out.INF
'I forbid you from going out.'

On the basis of this data, and since the Comp position of embedded infinitive clauses selected by propositional, factive and control verbs from non-literary Italian and Contemporary Spanish can be filled with an overt *di* and a — not necessarily — null *de* complementizer, we claim that these complements carry $[\emptyset_{Agr}]$ on C°, despite bearing a dependent tense.[82] In more schematic terms, we propose the following feature combination

[82]Observe that this proposal of a C head without an [Agr] feature is similar to our claim for factive complements from Portuguese.

for complements of this kind:

15. C-infinitives (non-literary IT/Contemporary ES):
$[_{CP}$ V $[_{CP}$ $C^{\circ}_{[\pm T, \emptyset_{Agr}]}$ $[_{IP}$ $I^{\circ}_{[\pm T, -Agr, -R]}$ $[_{VP}$ V° ...]]]]

Crucially, this now allows us to explain the properties of referential subjects within C-infinitives, *regardless* of whether [T] takes a positive or a negative value. Indeed, if these complements carry $[\emptyset_{Agr}]$ on C°, and given the R-assignment rule in 5, there can be no [R] feature on C°; only I° would bear a [−R] feature. As a consequence, the only possible path to check off the uninterpretable [−R] feature on the I head of C-infinitives would be by means of an Agree relation with the [−R] feature on the nominal phrase in the embedded subject position. We can see this in (16), below:

16. C-infinitives: [–R] subject

$$[_{CP}\ \underbrace{C^{\circ}_{[\pm T]}\quad [_{IP}\ PRO_{[-R]}}_{Agree\ [\pm T]}\quad [_{I'}\ \underbrace{I^{\circ}_{[\pm T, -Agr, -R]}\quad [_{VP}\ t_{PRO[-R]}}_{Agree\ [-Agr, -R]}\ ]]]]]$$

Since we have claimed that the [±Agr] features trigger the verb movement, our proposal for C-infinitives predicts that auxiliary uninflected verbs do not raise to the Comp embedded position not only in non-literary Italian, but also in Contemporary Spanish. In fact, there is no data for anaphoric pronouns in post-auxiliary position, nor many instances of (propositional and factive) compound infinitives with anaphoric pronouns. For instance, the controlled pronoun occurring in the following compound factive infinitive from Spanish stands in the post-participle position:[83]

(163) Saez se lamenta [haber sido **él** la nueva víctima de este acto vandálico].
Saez CL regrets have.INF be.PP he the new victim of this act hooliganism

[83] In addition, observe that the auxiliary-subject-verb word order became ungrammatical in finite sentences in Italian and also in Contemporary Spanish (cf. [Men00]; [Bel05]; [Paz13]). Thus, a nominal or a pronominal subject can surface in pre-auxiliary position or in the post-participle position, but cannot stand between the finite auxiliary and the participle verbs.

i. (Mario) Ha (*Mario) accettato (*Mario) di aiutarci.
(Mario) has (*Mario) accepted (*Mario) COMP help.INF-CL

‘Mario has accepted to help us.’

ii. ¿(María) Ha (*María) visto (María) la película?
(Mary) has (*Mary) seen (Mary) the movie

‘Has Mary seen the movie?’

iii. Lo que (él) ha (*él) hecho (él) es increíble.
that which (he) has (*he) done (he) is incredible

‘What he has done is incredible.’

'Saez$_1$ regrets that he$_1$ has been the new victim of this act of hooliganism.'

As is widely known, consistent *pro*-drop languages admit a post-verbal subject. In particular, overt anaphoric pronouns are restricted to the post-verbal position in consistent *pro*-drop languages.[84] Post-verbal subjects are typically (part of) the focus of the (finite or infinitive) sentence, and are assumed to be located in a low focus position, between *v*P and TP (cf. [Car99]; [Bel05]). In particular, the post-verbal position in infinitive clauses is associated with contrastive focus. In fact, as Piera observes, the meaning of the sentence in (164), containing the so-called emphatic pronoun, is similar to that of the pseudo-cleft sentence in (165).

(164) Julia quería [telefonear **ella**].
Julia wanted phone.INF she.NOM
'Julia wanted to phone herself.'

(165) Julia quería [ser **ella** la que telefoneara].
Julia wanted be.INF she.NOM the that phone
'Julia wanted to be the one who phoned.' [PG87]:160

A pseudo-cleft sentence expresses emphasis or contrastive focus on the constituent that follows the verb *be*. A widespread notion of focus-marking specifies that when a constituent is focus-marked in a sentence α, it triggers the presupposition that the context contains a set of propositions minimally differing from the one expressed by α in just the value of the focus-marked constituent. Thus, in (164), the contrastive focus on the infinitive controlled subject excludes the possibility that x — in the embedded predicate $\lambda x.(phone'\ x)$ — takes any value other than *Julia*. Thus, when this predicate combines with the matrix function $\lambda A.\lambda x.((wanted'\ (A))\ x)$ we do not obtain the function $\lambda x.((wanted'\ (phone'\ x))\ j)$, but rather the sentence $((wanted'\ (phone'\ j))\ j)$. Nevertheless, as claimed by Landau for the null subject PRO, we assume that the overt pronoun *ella* 'she' does not saturate the embedded controlled predicate in which it occurs, as it does not denote an entity. Thus, both *telefonear* 'to phone' and *telefonear ella* 'to phone herself' denote the function $\lambda x.(phone'\ x)$.

Recall from Section 2.3 that the distribution of PRO is severed from Case Theory in the T/Agr Calculus; and that PRO nonetheless retains a special status in this calculus. Despite this latter fact, the adoption of the concept of minimal pronoun to categorize PRO in Landau's more recent proposals clears the way for admitting an overt nominal bearing

[84] Brazilian Portuguese is a partial *pro*-drop language (cf. [HNS09]). In this language the anaphoric pronoun can be located in a pre-verbal position:

i. Os meninos querem [**eles** embrulhar(**em**) os presentes].
the boys want they wrap.INFL.(3PL) the gifts
'The boys want to wrap the gifts themselves.' [Neg87] (*apud* [Mod10b])

a [−R] feature alongside the null [−R] PRO. After all, the [−R] feature expresses the fact that the nominal bearing it is one incapable of independent reference (cf. [Lan04]:841). In light of this, we conclude that despite the different morpho-phonological properties and landing position, *anaphoric pronouns, such as PRO, are the thematic-syntactic subjects of the infinitive clause* (cf. [AOD00]; [Men00]; [Liv11]; [Her15]). And, when overt, the [−R] anaphoric pronoun stands in a post-verbal low focus position and carries a contrastive focus (cf. [Her15]; [Bar16]).

4.3 Conclusions

In the previous chapters we reviewed some of the generative assumptions that prevented an overt pronoun from being considered as the actual semantic-syntactic subject of the infinitive clause. Following the analysis of subjunctive complements in Balkan languages due to Landau, we adopted his T/Agr calculus to distinguish between two kinds of infinitive complements in Romance languages: Free infinitives and Controlled infinitives. We then collected data from literary Italian and Spanish to argue for this distinction. In fact, we showed that in some stages of Spanish and Italian, different kinds of referentially free subjects were licensed in infinitive complements despite the lack of inflection on the infinitive verb. In cases where the free subject is a pronoun, the latter exhibits the same shape as pronouns occurring in controlled clauses; in particular, they are nominative-marked.

Since there is no morphological distinction between pronouns in infinitive propositional complements and those within infinitive controlled clauses, there seems to be no reason to consider them as different types of pronouns (i.e. [–R] vs [+R]). The fact that in some infinitive contexts a pronoun must be controlled, while in others it can be free, is better explained as a consequence of the syntactic and/or semantic properties of the verb that selects its (infinitive) complement clause, rather than as an inherent characteristic of the pronoun itself. More specifically, we claimed that the crucial feature that distinguishes F- from C-infinitive complements is the [Agr] feature on the C embedded head. On the one hand, propositional verbs from literary Spanish and Italian can select an infinitive complement with a [−Agr] feature — an F-infinitive — or with no [Agr] feature on the C head — a C-infinitive. On the other hand, some verbs — the so-called control verbs — can *only* select an infinitive complement with no [Agr] feature on the C head. When this same infinitive is uninflected, the I head will bear a [−Agr] feature, and *because of this* only a subject carrying a [−R] feature can be licensed.

In addition, we noted that our proposal can be consistently extended to analyse inflected infinitive complements from Portuguese. We observed that our proposal, contrary to Landau's, fully respects all of the assumptions of the T/Agr calculus. Finally, we noted that the feature combination we used to give an analysis of (uninflected) F-infinitives instantiates the last possible combination of T/Agr features, obtained from the Agr on C° rule in 3 and displayed in Table 2.1 on page 42.

Part II

Categorial Grammars

5 Lambek Calculus

5.1 Introduction: Logical Grammar

In ‘The Mathematics of Sentence Structure’, Lambek [Lam58] proposes an algorithm to distinguish, among the expressions of natural and formal languages, those that are sentences from those that are not. The syntactic calculus **L** he devised for this purpose is based on the previous logical work of Ajdukiewicz [Ajd35] and Bar-Hillel [BH53]. Both Ajdukiewicz and Bar-Hillel assign syntactic types or categories to linguistic expressions, and they formalize the syntactic relationships among these expressions in terms of functional application between syntactic types. The basic intuition underlying the structure of (both formal and natural) languages in Ajdukiewicz’s proposal is that linguistic signs may be complete or incomplete, combined with the idea that grammatical composition is the process of completing incomplete linguistic signs. In Ajdukiewicz’s formal proposal, where A and B are (meta)syntactic categories, we have that $B|A$ is a functional category that, given an argument of category A, returns an expression of category B. The linguistic expressions are recognized as syntactically well-formed in his grammar if they are connected in terms of the functional application (or cancellation) schema in Fig. 5.1. Such a syntactic connectivity relationship is intended to be a necessary condition for well-formedness, but not a sufficient one, because it does not take word-order into account.

$$\frac{B}{A}\ A \Rightarrow B$$

Figure 5.1 – Cancellation schema

Thus, for example, we may consider a nominal expression like *John* to be complete, since its linguistic function does not depend on another linguistic expression. Similarly, a sentence is a complete sign insofar as it denotes a proposition carrying a truth-value that does not depend on the linguistic context. By contrast, an intransitive verb is incomplete in this sense. For, from a semantic perspective, an intransitive verb is a function that turns an individual into a proposition; and from a syntactic perspective, it needs a subject. Thus a verb like *run*, being an incomplete sign that turns a nominal expression of syntactic type n into a complete sentence of syntactic type s, is assigned the functional type $s|n$ in Ajdukiewicz’s formalism. As a result, the string *john runs* (and also *runs john*) satisfies the syntactic connection criterion because at least some ordering of the types assigned to the words can be reduced to the (distinguished) type s by means of the cancellation schema.

Bar-Hillel is the first logician to propose an application of Ajdukiewicz's categorial formalism for natural language. Motivated by a particular interest in natural language, Bar-Hillel modifies the previous formalism by splitting the functional connective | into two. In its place, he proposes two functional types or connectives \ and /, called *under* and *over* (respectively).[85] The resulting system, known as **AB** (for Ajdukiewicz-Bar-Hillel), features two functional application (or: directional *modus ponens*) rules for grammatical composition:

$$\frac{B/A \qquad A}{B}\ /\mathrm{E} \qquad\qquad \frac{A \qquad A\backslash B}{B}\ \backslash\mathrm{E}$$

Figure 5.2 – Rules of the **AB** system

In words, a syntactic type B/A ($A\backslash B$) takes an input argument A to the right (left), to yield an output value B. This directionality reflects the directionality in the formation of constituents in natural languages: typically, these languages impose constraints on word-order, such that some words have to be concatenated with others exclusively to the right or exclusively to the left.

Since **AB** may define directional types, an intransitive verb like *run* would be assigned the type $n\backslash s$, and a transitive verb like *eat*, the type $(n\backslash s)/n$. In addition, **AB** can accurately recognize the correct order for connecting a noun with an adjective in English and in Portuguese, say. Thus, if we assume the basic syntactic type *cn* to a common noun like *house* 'casa', we may attribute the category cn/cn to an English adjective such as *new*, and $cn\backslash cn$ to the Portuguese 'nova'.[86] By using the rules in Fig. 5.2 we may then prove that both sequences of syntactic types assigned to the strings *new house* and *casa nova* are of type *cn*. This last example also evidences that while the assignment of syntactic types to linguistic expressions may be language-dependent, the syntactic component of the grammar (i.e., the rules) is assumed to remain invariant across languages.

The system just described was later adapted by Lambek to produce the so-called **L** calculus, which is a conservative extension of **AB**. Like its predecessor, **L** was devised for the purpose of analyzing natural languages, and as such was developed with an eye to the empirical adequacy of the logical grammar therein contained.

The **L** system is a substructural logic, just like Intuitionistic, Relevant, Linear and Łukasiewicz multivalued logics, among others. As is well known, these logics are called substructural because, when formulated in a Gentzen sequent format, they lack certain structural rules enjoy by the Classical sequent logic **LK**: Permutation (P), Contraction (C) and Weakening (W) (See figure 5.3) (cf. [Ono03]; [GJKO07]).[87]

[85]These connectives are also called forward and backward looking slashes, respectively.
[86]Unlike English, Portuguese also allows an adjective to stand in a pre-nominal position.
[87]Permutation is also called Exchange, and Weakening is also called Monotonicity. Expansion (E) is a

$$\frac{X, A, B, Y \Rightarrow Z}{X, B, A, Y \Rightarrow Z}\ \mathrm{P}\Rightarrow \qquad\qquad \frac{X \Rightarrow Z, C, D, W}{X \Rightarrow Z, D, C, W} \Rightarrow \mathrm{P}$$

$$\frac{X, A, A, Y \Rightarrow Z}{X, A, Y \Rightarrow Z}\ \mathrm{C}\Rightarrow \qquad\qquad \frac{X \Rightarrow Z, D, D, W}{X \Rightarrow Z, D, W} \Rightarrow \mathrm{C}$$

$$\frac{X, Y \Rightarrow Z}{X, A, Y \Rightarrow Z}\ \mathrm{W}\Rightarrow \qquad\qquad \frac{X \Rightarrow Z, Y}{X \Rightarrow Z, C, Y} \Rightarrow \mathrm{W}$$

Figure 5.3 – Structural rules for the Classical **LK** sequent logic

A sequent of **LK** consists of a sequence of formulae $A_1, \ldots, A_n, n \geq 0$, (of the **LK** language) — the *antecedent* — and a sequence of formulae $C_1, \ldots, C_m, m \geq 0$ — the *consequent*. Antecedent and consequent are connected by the deduction symbol $\Rightarrow$. An **LK**-sequent $A_1, \ldots, A_n \Rightarrow C_1, \ldots, C_m$ expresses that some of the formulae C_i can be derived from the antecedent formulae $A_1, \ldots, A_n$. An inference rule asserts that if the premise sequent is derivable, then the conclusion sequent is derivable. Given that **LK** admits Permutation and Contraction, the ordering of formulae and the multiplicity of occurrences of the same formula into (the right or the left part of) a sequent both come to be superfluous. Hence, these two structural rules may be built into the very notion of sequent of **LK** by defining the latter in terms of sets of formulae.

Among the substructural systems mentioned earlier, one that deserves special mention here is the so-called **LJ** calculus for intuitionistic logic. Sequents of the **LJ** sequent calculus are expressions of the form $A_1, \ldots, A_n \Rightarrow C$, where $n \geq 0$ and the consequent may be empty. The rules of **LJ** are obtained from those of **LK**, first by deleting both $\Rightarrow\ C$ and $\Rightarrow\ P$ and then by changing the other rules so as to meet the no-more-than-one-formula condition in the consequent of the sequent.

The **L** sequent calculus does not contain either right nor left structural rules. In addition, like in **LJ**, the consequent of a sequent of **L** does not contain more than a single formula; moreover the antecedent sequence cannot be empty in **L**. Thus, an **L**-sequent has the form $A_1, \ldots, A_n \Rightarrow C$, where its antecedent $A_1, \ldots, A_n$ is a finite, non-empty list of

restricted version of Weakening, in that the conclusion is augmented by adding a copy of a formula that is already available from the premise:

$$\frac{X, A, Y \Rightarrow C}{X, A, A, Y \Rightarrow C}\ \mathrm{E}\Rightarrow \qquad\qquad \frac{X \Rightarrow Z, C, Y}{X \Rightarrow Z, C, C, Y} \Rightarrow \mathrm{E}$$

syntactic types and its consequent C is a single syntactic type.[88] By reading $A_1, \ldots, A_n$, where $n > 0$, as a sequence or list (of syntactic types), **L** implicitly encodes Associativity left rules (see figure 5.4).[89]

$$\frac{X, (A, B), D, Y \Rightarrow C}{X, A, (B, D), Y \Rightarrow C} A_1 \Rightarrow \qquad\qquad \frac{X, A, (B, D), Y \Rightarrow C}{X, (A, B), D, Y \Rightarrow C} A_2 \Rightarrow$$

Figure 5.4 – Associativity left rules

The rejection of all structural rules in **L** comes from the use of this calculus as a logical grammar. Structural rules affect grouping (Associativity), ordering (Permutation), multiplicity (Weakening/Expansion) and occurrence (Contraction) of linguistic resources. As a result, they cannot be accepted as rules for the management of linguistic material in a global fashion ([KO03]:Ch.9). We now take a closer look at these in turn.

The left Permutation rule ($P \Rightarrow$) licenses changes in the order of the antecedent formulae. In linguistic terms, it expresses the fact that there are no order constraints on concatenating expressions of a language. If this rule were admitted in the logical grammar for English, it would imply that, for example, (166a) and (166b) should both be judged as grammatical. Nevertheless, inferences in a grammar should be sensitive to the linear order of the linguistic material.

(166) a. John runs.
b. *Runs John.

In a system in which the Associative rules ($A_{1/2} \Rightarrow$) are (implicitly or explicitly) admitted, there is no place for (hierarchical) grouping of formulae (cf.[Moo11]:113). By assuming the Associativity rules, the linguistic notion of *constituency* is relaxed, insofar as chains of words can be freely rebracketed; consequently, non-standard constituent analysis can be carried out. Grammar asymmetrical relations like c-command and dominance, however, cannot be defined in an associative logical grammar. To put it in more visual terms, note that the first two binary trees in Fig. 5.5 collapse into the third ternary, flat tree in an associative system. In such a system there would be no way to establish a hierarchical relation between, for example, two prepositional verb complements (see sentences in (167)).

(167) a. John talked to Mary about herself.
b. *John talked about Mary to herself.

[88]In Section 5.2.1 we shall present formal definitions for *language of **L**, sequent, inference rule, proof*, etc.

[89]The non-associativity Lambek calculus **NL** is presented in [Lam61].

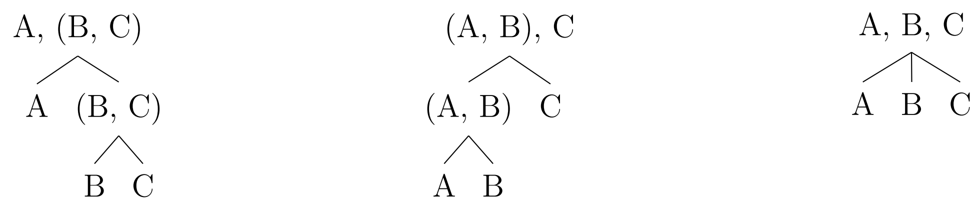


Figure 5.5 – Binary trees and Associativity

The left Contraction rule ($C \Rightarrow$) deletes an occurrence of a duplicated formula in the antecedent sequence. By admitting this rule in a grammar, it would be possible to erase repeated occurrences of lexical material.[90] Consequently, each of the sentences displayed below would — incorrectly — be typed as s:

(168) a. Anybody who laughed laughed about John.
b. *Anybody who laughed about John.

(169) a. Not everything which is is necessary.
b. *Not everything which is necessary.

(170) a. What John is is proud of himself.
b. *What John is proud of himself.

Finally, the left Weakening rule ($W \Rightarrow$) warrants that the conclusion of a sequent is preserved even when a formula is inserted in the sequence of — antecedent — formulae. In linguistic terms, this rule implies that the syntactic type a string of words is assigned to is preserved, regardless of any insertion of linguistic material. However, the addition of lexical items is not free in a language and can affect the syntactic type, and even the grammaticality, of a linguistic string of words:

(171) a. John runs.
b. *John runs Mary.

The rejection of Weakening in a grammar warrants that all of the lexical material that occurs in a sequent is used at least once; lexical material is not superfluous or irrelevant. The rejection of Contraction, in turn, avoids reusing linguistic material; each instance of any lexical item may be used only once. Because of the rejection of both Weakening and Contraction, the syntactic types in an **L**-sequent must all be used in a proof, after which they are no longer available to be reused. In this sense the Lambek Calculus, like Linear logic, is a resource-conscious (or a resource-sensitive) logic.

[90] The Contraction rule only allows deleting of *contiguous*, identical formulae; thought, by using Permutation, the adjacency condition can be straightforwardly satisfied.

Despite the fact that natural languages are certainly not completely free-word-order, and that multiplicity and structuring of linguistic resources would affect grammatical well-formedness in some cases, the following sentences suggest that some controlled form of these structural operations has to be available for a logical grammar to manage the linguistic material.

(172) a. John talks to Mary about Bill.
b. John talks about Bill to Mary. [Permutation]

(173) a. Students are required to speak the Chinese language only in the school.
b. Students are required to speak only the Chinese language in the school. [Permutation]

(174) a. Mary spoke with some boy, but I don't know with who.
b. Mary spoke with some boy, but I don't know who. [Contraction]

(175) a. John washed his car, and Bill washed his car.
b. John washed his car, and Bill did, too. [Contraction]

(176) the paper that John filed without reading. [Contraction]

(177) a. John saw Mary and Bill heard Mary.
b. John saw and Bill heard Mary. [(Non-)Associativity]

(178) a. I did it.
b. I did it myself. [Weakening]

(179) a. John believes to be unproductive.
b. John believes himself to be unproductive. [Contraction or Weakening]

(180) Pablo le dio un libro a María.
Paul CL gave a book to Mary
'Paul gave a book to Mary.' [Weakening]

As we shall show in the following chapters, different strategies have been proposed in the categorial framework to deal with linguistic phenomena that seem to require using these structural rules. Several strategies have been pursued to obtain controlled structural extensions of the **AB** and **L** systems, and thus to adequately increase the empirical scope of Categorial Grammars. In particular, we shall concentrate on two ways to extend these systems in order to deal with the problem of reusing resources (and multiple-binding) generated by the lack of the Contraction rule.

5.2 Lambek Calculus

5.2.1 Syntax

In Categorial Grammars, a grammar is presented as a logical system for reasoning about linguistic resources; roughly speaking, Categorial Grammars reduce grammar to logic. In these grammars in general, and in the Lambek Calculus in particular, lexical units of a natural language are assigned syntactic types inductively built out of basic types plus a small number of type-forming connectives (or type-constructors). Determining whether a linguistic expression is well-formed amounts to presenting a derivation (i.e. a proof) of an associated logical statement in the logical system for these connectives. In this and the following subsection we present formal definitions of key syntactic and semantic notions of the Lambek calculus.

Definition 5.1 (Syntactic types of **L**)**.** Let a finite set P of basic, atomic or primitive types be given. The *set* $F_{\mathbf{L}}$ *of syntactic types (or: type formulae) of* $\boldsymbol{L}$ is the smallest set such that:

1. $\mathrm{P} \subseteq \mathrm{F}_{\mathbf{L}}$;
2. $if A, B \in \mathrm{F}_{\mathbf{L}},\ then\ (B/A) \in \mathrm{F}_{\mathbf{L}}$;
3. $if A, B \in \mathrm{F}_{\mathbf{L}},\ then\ (A\backslash B) \in \mathrm{F}_{\mathbf{L}}$;
4. $if A, B \in \mathrm{F}_{\mathbf{L}},\ then\ (A \bullet B) \in \mathrm{F}_{\mathbf{L}}$.

Informally, basic types categorize expressions one can think of as 'complete'. Most linguistic applications assume the following basic types: s for sentence, n for proper (or: referential) name, *cn* for common noun, and *pp* for prepositional phrase. However, this is not essential and other basic types can be assumed. Compound types are freely generated from basic types by the binary type-constructors \ (under), / (over) and • (product).[91]

Example 5.1.

n/cn

$(n\backslash s)/n$

$(n\backslash s)/(n \bullet pp)$

$n\backslash(n/n)$

The interpretation of syntactic types of **L** is in terms of semigroups.[92]

[91] As usual, we shall delete the most external parentheses hereinafter.
[92] For an interpretation of **L**-types in terms of frame semantics see, for example, [Moo11]:101.

Definition 5.2 (Basic prosodic algebra)**.** A *basic prosodic algebra* is an algebra $(L, +)$ of arity (2) which is a free semigroup [Lam88], i.e. L is a set and $+$ is a binary operation on L such that for all $s_1, s_2, s_3 \in L$:

$$s_1 + (s_2 + s_3) = (s_1 + s_2) + s_3 \quad \text{associativity}$$

The associative binary operation + is called concatenation. In applying the basic prosodic algebra to a natural language, the set L is assumed as the set Σ^+ of (non-empty) strings over a vocabulary (or: alphabet) Σ; L is $\Sigma^* - \{t\}$, where t is the empty string. In other words, the prosodic algebra is a semigroup freely generated by the words of the language under concatenation. Each formula A of **L** is interpreted as a subset of L, i.e. as a subset of non-empty strings. Given such a mapping for atomic types, it is extended to the compound types.

Definition 5.3 (Prosodic intrepretation of **L**)**.** A *prosodic interpretation of* ***L*** is a function $[\![\cdot]\!]$ mapping each type $A \in F_{\mathbf{L}}$ into a subset of L such that:

$$\begin{aligned}
[\![A \backslash B]\!] &= \{s_2 | \forall s_1 \in [\![A]\!], s_1 + s_2 \in [\![B]\!]\} \\
[\![C/B]\!] &= \{s_1 | \forall s_2 \in [\![B]\!], s_1 + s_2 \in [\![C]\!]\} \\
[\![A \bullet B]\!] &= \{s_1 + s_2 | s_1 \in [\![A]\!],\ s_2 \in [\![B]\!]\}
\end{aligned}$$

Note that the product $\bullet$ is an associativity type-constructor, as it inherits associativity from the basic prosodic algebra. The product $\bullet$ is, however, an ordered-sensitive conjunction as the binary operation + of concatenation is not commutative.

Definition 5.4 (Sequent of **L**)**.** A *sequent of* ***L*** is an expression of the form $A_1, \ldots, A_n \Rightarrow A$, where $n > 0$ and $A_1, \ldots, A_n$ is a (finite) sequence of syntactic types of **L** — the antecedent, and A is a single type of **L** — the consequent.

According to the previous definition, an **L**-sequent $A_1, \ldots, A_n \Rightarrow A$ can be read as asserting that for any elements $s_1, \ldots, s_n$ in $A_1, \ldots, A_n$ respectively, $s_1 + \ldots + s_n$ is in A. The relevant prosodic operations are thus encoded by the linear ordering of the antecedents in the **L**-sequent.

Definition 5.5 (Validity of Sequents in **L**)**.** An ***L****-sequent* $A_1, \ldots, A_n \Rightarrow A$ is *valid* if and only if $[\![A_1]\!], \ldots, [\![A_n]\!] \subseteq [\![A]\!]$ in every (prosodic) interpretation.

A Gentzen sequent calculus comprises a right and a left rule for each connective; they are called logical rules. A left rule expresses the sufficient conditions for a connective to be used; thus it is also called *rule of use.* A right rule for a connective expresses the necessary conditions for its proof, and is also called *rule of proof.* Thus, the sequent calculus for **L** contains a rule of use and a rule of proof for the type-constructors $/, \backslash$ and

$\bullet$. Each logical rule introduces a single connective from premises to conclusion. A left rule introduces a (formula built out of a) type-constructor in the antecedent sequence of the conclusion of the rule. A right rule in turn introduces a connective in the consequent of the conclusion of the rule. Besides logical rules, the sequent calculus for **L** comprises the Identity and Cut rules, which establish the reflexivity and transitivity of the derivability relation, respectively.[93] Associative left rules are implicitly admitted in **L**. Each rule of inference in the **L** sequent calculus has the form $\frac{\Pi_1 \cdots \Pi_n}{\Pi_0}, n \geq 0$, where each of Π_i is a sequent schemata. Each rule asserts that if the (premises) $\Pi_1, \ldots, \Pi_n$ are derivable, then the conclusion Π_0 is derivable. The rules for the **L** system are given in Fig. 5.6.

$$\frac{}{A \Rightarrow A}\,\mathrm{Id} \qquad\qquad \frac{X \Rightarrow A \qquad Y, A, W \Rightarrow C}{Y, X, W \Rightarrow C}\,\mathrm{Cut}$$

$$\frac{X \Rightarrow A \qquad Y, B, Z \Rightarrow C}{Y, B/A, X, Z \Rightarrow C}\,/\mathrm{L} \qquad\qquad \frac{X, A \Rightarrow C}{X \Rightarrow C/A}\,/\mathrm{R}$$

$$\frac{X \Rightarrow A \qquad Y, B, Z \Rightarrow C}{Y, X, A\backslash B, Z \Rightarrow C}\,\backslash\mathrm{L} \qquad\qquad \frac{A, X \Rightarrow C}{X \Rightarrow A\backslash C}\,\backslash\mathrm{R}$$

$$\frac{X, A, B, Y \Rightarrow C}{X, A \bullet B, Y \Rightarrow C}\,\bullet\mathrm{L} \qquad\qquad \frac{X \Rightarrow A \qquad Y \Rightarrow B}{X, Y \Rightarrow A \bullet B}\,\bullet\mathrm{R}$$

Figure 5.6 – Rules of the (associative) **L** system

Definition 5.6 (Proof of a sequent in **L**)**.** A *proof* $\mathcal{P}$ *of a sequent* $\Gamma \Rightarrow A$ in **L** — an **L**-proof — is a finite tree of sequents satisfying:

1. All sequents in a leaf of $\mathcal{P}$ are instances of the conclusion of the Identity rule;
2. Every sequent in a node of $\mathcal{P}$ is a conclusion of an instance of a rule of **L** obtained from upper sequents in $\mathcal{P}$;
3. $\Gamma \Rightarrow A$ is the lowest sequent (or: endsequent) in $\mathcal{P}$.

Definition 5.7 (Sequent derivable in **L**)**.** A *sequent* $\Gamma \Rightarrow A$ is *derivable* in **L** if and only if $\Gamma \Rightarrow A$ is the endsequent of some proof in **L**. (In this case, we write $\vdash_{\mathbf{L}} \Gamma \Rightarrow A$; otherwise we write $\nvdash_{\mathbf{L}} \Gamma \Rightarrow A$.)

[93] The Cut rule is admissible in **L**, i.e. adding the Cut rule does not give rise to any new theorems. **L** satisfies decidability and the subformula property.

Definition 5.8 (Theorem of **L**)**.** A *theorem of* $\boldsymbol{L}$ is a sequent which is derivable in the **L** calculus.

Proposition 5.1 (Soundness of **L**)**.** *In* $\boldsymbol{L}$*, every theorem is valid.*

Proof. Straightforward induction on the length of sequent proofs. □

Proposition 5.2. *The following sequents are theorems of* $\boldsymbol{L}$ *(cf. [Lam88]:303; [Moo11]:111):*

$(B/A) \bullet A \Rightarrow B$	$[ForwardApplication]$
$A \bullet (A \backslash B) \Rightarrow B$	$[BackwardApplication]$
$A \bullet (B \bullet C) \Rightarrow (A \bullet B) \bullet C$	$[Associativity]$
$A/(B \bullet C) \Rightarrow (A/C)/B$	$[Currying]$
$(A \backslash B)/C \Rightarrow A \backslash (B/C)$	$[Restructuring]$
$A/B, B/C \Rightarrow A/C$	$[Composition]$
$A/B \Rightarrow (A/C)/(B/C)$	$[Division\ or\ Geach\text{-}rule]$
$A \Rightarrow B/(A \backslash B)$	$[Type\text{-}Raising\ or\ Lifting]$
$A \Rightarrow (B/A) \backslash B$	$[Type\text{-}Raising\ or\ Lifting]$

Proposition 5.3. *The following laws of residuation are derivable in* $\boldsymbol{L}$*. Thus,* $(\bullet, /)$ *and* $(\bullet, \backslash)$ *each form residuated pairs in* $\boldsymbol{L}$*.*

$$A \bullet B \Rightarrow C \textit{ iff } A \Rightarrow C/B$$
$$A \bullet B \Rightarrow C \textit{ iff } B \Rightarrow A \backslash C.$$

The inference rules given in Fig. 5.6 on the previous page above provide the core syntax of a Categorial Grammar based on **L**; they are assumed to be the same for all such grammars. Categorial Grammars may however differ with respect to the type-constructors they use; several conservative extensions of **L** have been proposed in the literature (cf., for example, [MVpt]). In applying Categorial Grammars to specific languages, further assumptions will be required to account for the basic types assumed in the grammar. Thus, the parametric or non-fixed part of a Categorial Grammar is witnessed by the lexicon.

A (categorial) lexicon assigns one or more types to the atomic elements of a language. In more formal terms, the lexicon of a Categorial Grammar is a mapping that assigns a finite set of syntactic types to some finite set of strings built out of an alphabet. In yet more formal terms:

Definition 5.9 (Lexicon)**.** Let an alphabet Σ and a finite set P of basic types be given. A lexicon LEX is a finite relation between the set Σ^+ (of non-empty strings over Σ) and the set $F_{\mathbf{L}}$.

Thus, a lexicon LEX is a (finite) set of ordered pairs $\langle s, A\rangle$, where $s \in \Sigma^+$ and $A \in F_{\mathbf{L}}$. For the sake of readability, however, we display the lexicon as a list of items of the form $\mathbf{s} : A$, instead of writing $\langle s, A\rangle \in LEX$. Moreover, we write $\mathbf{s}_1/\ldots/\mathbf{s}_n : A$ when each string $\mathbf{s}_i$ is of the same type A. Consider the following sample lexicon for English:

Example 5.2.

a/the : n/cn

give : $(n\backslash s)/(n \bullet pp)$

give : $(n\backslash s)/(n \bullet n)$

John/Mary : n

mother : cn

of : $(n\backslash n)/n$

prize : cn

say : $(n\backslash s)/s$

walk / lose / win : $n\backslash s$

walk / lose / win : $(n\backslash s)/n$

Definition 5.10 (Lambek Grammar)**.** Let an alphabet Σ be given. A *Lambek grammar* G is a triple $\langle P, LEX, D\rangle$, where P is a finite set (the basic types), LEX is a finite sub-relation of $\Sigma^+ \times F_{\mathbf{L}}$, and D is a finite subset of $F_{\mathbf{L}}$ (that is, the designated syntactic types).

Definition 5.11 (Language of **L**)**.** Let $G = \langle P, LEX, D\rangle$ be a Lambek grammar over the alphabet Σ. Then $\alpha \in L(G)$ iff there are $s_1, \ldots, s_n \in \Sigma^+, A_1, \ldots, A_n \in F_{\mathbf{L}}$, and $S \in D$ such that:

1. $\alpha = s_1 + \ldots + s_n$,

2. For all i such that $1 \leq i \leq n : \langle s_i, A_i\rangle \in LEX$, and

3. $\vdash_{\mathbf{L}} A_1, \ldots, A_n \Rightarrow S$.

The language of **L** consists of all the concatenations of strings $s_1 + \ldots + s_n \in \Sigma^+$ such that for some sequence of types $A_1, \ldots, A_n$ assigned to the strings $s_1, \ldots, s_n$ in the lexicon LEX, a distinguished type $S \in F_{\mathbf{L}}$ is assigned by the grammar G. And the grammar assigns a distinguished type S, if the sequent $A_1, \ldots, A_n \Rightarrow S$ is derivable in the **L** system. Usually, the set D of designated categories is tacitly assumed to be the singleton $\{s\}$.

Example 5.3. *Consider $\alpha = s_1 + \ldots + s_5$, where $s_1 =$ **John**, $s_2 =$ **gives**, $s_3 =$ **Mary**, $s_4 =$ **the** and $s_5 =$ **prize**. Also consider the lexicon given in the example 5.2 above. The **L**-proof given below shows that **John+gives+Mary+the+prize** is recognized by the Lambek grammar.*

$$\frac{\dfrac{n \Rightarrow n \qquad \dfrac{cn \Rightarrow cn \quad n \Rightarrow n}{n/cn, cn \Rightarrow n}\,/\ \mathrm{L}}{n, n/cn, cn \Rightarrow n \bullet n}\,\bullet\ \mathrm{R} \qquad \dfrac{n \Rightarrow n \quad s \Rightarrow s}{n, n\backslash s \Rightarrow s}\,\backslash\ \mathrm{L}}{n, (n\backslash s)/(n \bullet n), n, n/cn, cn \Rightarrow s}\,/\ \mathrm{L}$$

5.2.2 Semantics

Definition 5.12 (Semantic types of **L**)**.** Let a finite set δ of basic (or: primitive) semantic types be given. The set γ of *semantic types* is the smallest set such that:

1. $\delta \subseteq \gamma$; [basic type]
2. if $\alpha, \beta \in \gamma$, then $(\alpha \& \beta) \in \gamma$; [product type]
3. if $\alpha, \beta \in \gamma$, then $(\alpha \rightarrow \beta) \in \gamma$. [functional type]

In linguistic applications, the semantic types e and t are usually assumed as (the) basic semantic types. The semantic type of a sign indicates what kind of object the sign denotes. Semantic categories correspond to ontological domains. Each semantic type α is associated with a semantic domain $Dom(\alpha)$.

Definition 5.13 (Type Domains)**.** The *type domain* $Dom(\alpha)$ of each semantic type $\alpha \in \gamma$ is a function defined on the basis of an assignment d from non-empty sets (i.e. the basic type domains) to basic semantic categories δ as follows:

1. $Dom(\alpha) = d(\alpha)$ for $\alpha \in \delta$
2. $Dom(\alpha \rightarrow \beta) = Dom(\beta)^{Dom(\alpha)}$ [functional formation]
3. $Dom(\alpha \& \beta) = Dom(\alpha) \times Dom(\beta)$ [Cartesian product]

The previous definition does not restrict the assignment of domains to basic types beyond the non-emptiness condition. Nevertheless, it is usually assumed that the set E of entities and the set $\{0, 1\}$ of truth-values are the basic semantic domains for the basic (semantic) types e and t, respectively. In other terms, $d(e) = E$ and $d(t) = \{0, 1\}$. Given some basic domains, the domains for semantic types are generated by function formation $\pi_1^{\pi_2}$ (the set of all functions from π_2 into π_1) and pair formation $\pi_1 \times \pi_2$ (the set of all ordered pairs such that π_2 follows π_1).

Definition 5.14 (Semantic type map)**.** The *semantic type map* for **L** is a mapping τ from syntactic types of $F_{\mathbf{L}}$ to semantic types of γ. Given such a type map τ for basic syntactic types P, it is extended to compound syntactic types by:

1. $\tau(A \bullet B) = \tau(A) \,\&\, \tau(B)$ [product type map]

2. $\tau(A\backslash B) = \tau(B/A) = \tau(A) \to \tau(B)$ [functional type map]

The previous map establishes a correspondence between semantic categories and syntactic types of **L**.[94] Given that a linguistic term is assigned a syntactic type in the lexicon, each lexical entry provides syntactic and semantic information: it encodes the syntactic combinatory possibilities of the term and also specifies which kind of denotation the linguistic sign has. Thus, for example, the denotation of a type $n\backslash s$, generally assigned to intransitive verbs, will be a function of semantic type $(e \to t)$ (also written: $< e, t >$), i.e. a function from the set of entities to the set of truth-values. And a particular intransitive verb will then denote a function $\phi \in \{0, 1\}^E$.

An interpreted lexicon LEX will then be a (finite) set of triples of the form $\langle s, A, \phi\rangle$, where s and A are as before and ϕ is a term of semantic type $\tau(A)$ (i.e. $\phi \in Dom(\tau(A))$). As before, we shall display the lexicon as a list of lexical entries each of the form $\mathbf{s} : A : \phi$, rather than $\langle s, A, \phi\rangle \in LEX$. Following the tradition in Categorial Grammars, we shall use (closed terms of) the language of the (typed) λ-calculus for lexical semantics.

In Sec. 5.4 we will be presenting the syntactic Lambek calculus complete with a derivational semantics. The resulting system is known as labelled Lambek calculus: briefly, in this system syntactic types are labelled with terms of the so-called λ-calculus, and proofs are labelled with operations on these terms. By way of background, the next section gives an overview of the basic λ-calculus.

5.3 (Typed) λ-Calculus

The (typed) λ-calculus is a formal language for functions; its expressions are called (typed) λ-terms. Giving a detailed presentation of the entire theory would take us far beyond the

[94]Note that, while cn is usually assumed as a basic syntactic type, it is also assumed that it denotes a function of semantic type $(e \to t)$ (cf. [Mor92]:10; [Jac03]:86). Thus, while $cn \in P, \tau(cn) \notin \delta$; $Dom(\tau(cn)) = \{0, 1\}^E$.

scope of this thesis; the interested reader is referred to [HS86]; [Bar84]; [Fer09]; [BDS13]. Here we merely introduce those definitions and rules that are necessary in order to present the labelled Lambek Calculus and, in particular, to exhibit the semantic consequence of the lack of Contraction (and also Weakening) in **L**.

Definition 5.15 (λ-terms)**.** Let V be an infinite, countable set of variables $v_0, v_1, \ldots$ and let C be an infinite, countable set of constants $c_0, c_1, \ldots$. The set Λ of *λ-terms* is defined inductively as follows ([Bar84]:22; [Fer09]:34):

1. if $c \in C$, $c \in \Lambda$; [constants]
2. if $v \in V$, $v \in \Lambda$; [variables]
3. if $v \in V$ and $M \in \Lambda$, then $(\lambda v.M) \in \Lambda$; [functional abstraction]
4. if $M, N \in \Lambda$, then $(MN) \in \Lambda$; [(functional) application]
5. if $M, N \in \Lambda$, then $\langle M, N\rangle \in \Lambda$; [pairing]
6. if $M, N \in \Lambda$, then $\pi_1(\langle M, N\rangle) \in \Lambda$; [first-projection]
7. if $M, N \in \Lambda$, then $\pi_2(\langle M, N\rangle) \in \Lambda$. [second-projection]

Definition 5.16 (Set of free variables in a term)**.** $FV(M)$ is the *set of free variables in* M and can be defined inductively as follows:

1. $FV(c) = \emptyset$;
2. $FV(x) = \{x\}$;
3. $FV(\lambda x.M) = FV(M) - \{x\}$;
4. $FV((MN)) = FV(M) \cup FV(N)$;
5. $FV(\langle MN\rangle) = FV(M) \cup FV(N)$.

If $FV(M) = \emptyset$, then M is a *closed term* (or: *combinator*).

Definition 5.17 (Free occurrence of a variable)**.** A *variable x occurs free* in a λ-term M if $x \in FV(M)$; otherwise, x occurs bound.

Table 5.1 exhibits a few examples of closed λ-terms that are frequently used in (Combinatory) Categorial Grammars ([KO03]:244; [Gre15]:286).[95]

[95] λ-terms are defined modulo α-equivalence; so, for example, $\lambda x.x$ and $\lambda y.y$ are the same λ-term. For a precise definition of α-equivalence, see [Fer09]:38.

In the typed λ-calculus, λ-terms are assigned types, which are syntactic objects. If M is such a term and a type τ is assigned to M, then we say that M has type τ. By assigning types to the λ-terms, the domain of a function is specified. Thus, in a typed λ-calculus the application of a function to an argument is not allowed unless the type of the argument is the same as the domain of the function.

Definition 5.18 (Set $\mathbb{T}$ of types)**.** Let a finite set δ of basic (or: primitive) semantic types be given. The *set* $\mathbb{T}$ *of types* (of $\lambda^{\to,\times}$) is inductively defined as follows ([Bar93]:34; [Meu14]:370):

1. if $\alpha \in \delta$, then $\alpha \in \mathbb{T}$; [basic types]
2. if $\sigma, \tau \in \mathbb{T}$, then $(\sigma \to \tau) \in \mathbb{T}$; [function space types]
3. if $\sigma, \tau \in \mathbb{T}$, then $(\sigma \times \tau) \in \mathbb{T}$. [Cartesian product types]

Definition 5.19 (Typed λ-terms)**.** The sets Λ_τ of typed λ-terms for any type $\tau \in \mathbb{T}$ are defined on the basis of a set C_τ of constants of type τ and an enumerable infinite set V_τ of variables of type τ for each type τ as follows:

1. if $c \in C_\tau$, then $c \in \Lambda_\tau$;
2. if $v \in V_\tau$, then $v \in \Lambda_\tau$;
3. if $x \in \Lambda_\tau$ and if $M \in \Lambda_\sigma$, then $\lambda x.M \in \Lambda_{\tau\to\sigma}$;
4. if $M \in \Lambda_{\tau\to\sigma}$ and $N \in \Lambda_\tau$, then $(MN) \in \Lambda_\sigma$;
5. if $M \in \Lambda_\tau$ and $N \in \Lambda_\sigma$, then $\langle M, N\rangle \in \Lambda_{\tau\times\sigma}$.
6. if $M \in \Lambda_{\sigma\times\tau}$, then $\pi_1(M) \in \Lambda_\sigma$.
7. if $M \in \Lambda_{\sigma\times\tau}$, then $\pi_2(M) \in \Lambda_\tau$.

We write $M[N/T]$ to denote the result of substituting N for T in M (also written: $M\{T \mapsto N\}$ in [Car97]; [Fer09] and $M[T := N]$ in [Bar93]).

I(dentity)	$\lambda x.x$	$\mathbf{I}x = x$
A(pply)	$\lambda x.\lambda y.(x\ y)$	$\mathbf{A}xy = xy$
Type-Raising	$\lambda x.\lambda y.(y\ x)$	$\mathbf{CI}xy = yx$
C(ommutor)	$\lambda x.\lambda y.\lambda z.((x\ z)\ y)$	$\mathbf{C}xyz = xzy$
K (constant)	$\lambda x.\lambda y.x$	$\mathbf{K}x = x$
W (duplicator)	$\lambda x.\lambda y.((x\ y)\ y)$	$\mathbf{W}xy = xyy$
B (change structure or composition)	$\lambda x.\lambda y.\lambda z.(x\ (y\ z))$	$\mathbf{B}xyz = x(yz)$
S(trong composition)	$\lambda x.\lambda y.\lambda z.((x\ z)\ (y\ z))$	$\mathbf{S}xyz = xz(yz)$

Table 5.1 – Combinators and λ-terms

Proposition 5.4. *The typed λ-calculus contains — among others — the following reduction rules on typed λ-terms:*[96]

1. $\lambda x.M \rightsquigarrow_\alpha \lambda y.M[y/x]$ *for* $y \notin FV(M)$ *and* y *is free for* x *in* M *[α-reduction]*
2. $((\lambda x.M)N) \rightsquigarrow_\beta M[N/x]$ *provided* N *is free for* x *in* M *[β-reduction]*
3. $\lambda x.(Mx) \rightsquigarrow_\eta M$ *provided* $x \notin FV(M)$ *[η-reduction]*
4. $\langle \pi_1(M), \pi_2(M)\rangle \rightsquigarrow_\eta M$ *[pairing]*
5. $\pi_1(\langle M, N\rangle) \rightsquigarrow_\beta M$ *[left- (or: first-) projection]*
6. $\pi_2(\langle M, N\rangle) \rightsquigarrow_\beta N$ *[right- (or: second-) projection]*

Here, we are assuming that the λ-terms are appropriately typed. So for example, in α-reduction, x and y have to be variables of the same type; for β-reduction, we implicitly require x and N to be of the same type, and so on.

The typed λ-calculus provides a language to represent meanings of (standard and non-standard) constituents of natural language. Thus, for example, a nominal like *John* may be represented by an individual constant of type e; thus it may be interpreted as the individual $j' \in Dom(e)$. An intransitive verb like *run* may be represented as a functional type $\lambda x.(run'\ x)$ of type $e \to t$. Nevertheless, since $\lambda x.(run'\ x) \rightsquigarrow_\eta run'$, *run* is a constant of type $e \to t$. Thus, $\lambda x.((run'\ x)\ j')$ and $run'\ j'$ are terms of type t.

5.4 Labelled Lambek Calculus

As is well known, the Curry-Howard correspondence maps Intuitionistic logic into a (particular version of) the simply-typed lambda calculus. A little more specifically: formulae of the Intuitionistic logic correspond to types of the simply-typed λ-calculus; (natural deduction) proofs map onto terms of the λ-calculus; and Cut elimination corresponds to λ-reduction ([BDS13]; [Jun04]; [HS86]; [Roj15]). It follows from the Curry-Howard correspondence that logical proofs acquire a semantic reading, i.e. a derivational semantics. This correspondence, also known as the "proofs-as-terms" or the "formulae-as-types" correspondence, was proved for other substructural logics. In particular, the semantic interpretation of **L**-proofs along the lines of the Curry-Howard correspondence was proved by van Benthem [vB88]. From this perspective, a type A can be viewed as a formula and a term M of type A, as a proof of this formula.[97]

[96]For the interpretation of (typed) λ-terms $M \in \Lambda_\tau$ by means of the interpretation function $[\![.]\!]$, which validates the previous conversion-rules of the λ-calculus see, for example, [Car97]:Ch.2; [Jäg05]:32s; [Mor11]:131.

[97]For a friendly, step-by-step proof of the Curry-Howard correspondence for **LJ**, we refer the reader to [SU06]:Ch.4 and [Alb10]:Ch.11.

Lambek originally presented his type logic as a calculus of syntactic types. Syntactic types, however, correspond to semantic types via the semantic type map $\tau : F_{\mathbf{L}} \rightarrow \gamma$. Thus, each syntactic type of **L** may be labelled by a λ-term. Operations on syntactic types correspond then to operations on terms of the typed λ-calculus.

As is usual in Categorial Grammar, we write $M : A$ when M is a typed λ-term and A is a syntactic type (of **L**). However we omit the types on λ-terms for the sake of readability.

A semantically labelled **L**-sequent $x_1 : A_1, \ldots, x_n : A_n \Rightarrow M : C$, where $n > 0$, is a sequent in which the antecedent type occurrences $A_1, \ldots, A_n$ are labelled by *distinct* variables $x_1, \ldots, x_n$ of types $\tau(A_1), \ldots, \tau(A_n)$ respectively, and the consequent type C is labelled by a term of type $\tau(A)$ with free variables drawn from $x_1, \ldots, x_n$ (cf. [Mor17a]:126). Figure 5.7 displays the semantically labelled **L** sequent calculus.[98]

$$\frac{}{x : A \Rightarrow x : A}\ \mathrm{Id} \qquad \frac{X \Rightarrow M : A \qquad Y, x : A, W \Rightarrow N : C}{Y, X, W \Rightarrow N[M/x] : C}\ \mathrm{Cut}$$

$$\frac{X \Rightarrow M : A \qquad Y, x : B, Z \Rightarrow N : C}{Y, y : B/A, X, Z \Rightarrow N[(yM)/x] : C}\ /\mathrm{L} \qquad \frac{X, x : A \Rightarrow N : C}{X \Rightarrow \lambda x.N : C/A}\ /\mathrm{R}$$

$$\frac{X \Rightarrow M : A \qquad Y, x : B, Z \Rightarrow N : C}{Y, X, y : A\backslash B, Z \Rightarrow N[(yM)/x] : C}\ \backslash\mathrm{L} \qquad \frac{x : A, X \Rightarrow N : C}{X \Rightarrow \lambda x.N : A\backslash C}\ \backslash\mathrm{R}$$

$$\frac{X, x : A, y : B, Y \Rightarrow N : C}{X, z : A \bullet B, Y \Rightarrow N[\pi_1(z)/x][\pi_2(z)/y] : C}\ \bullet\mathrm{L} \qquad \frac{X \Rightarrow M : A \qquad Y \Rightarrow N : B}{X, Y \Rightarrow \langle M, N\rangle : A \bullet B}\ \bullet\mathrm{R}$$

Figure 5.7 – Labelled sequent calculus for **L**

On a top-bottom reading, a labelled **L**-proof starts with instances of the identity rule labelled with the same variable in the antecedent and consequent. Furthermore, each rule of inference corresponds to an operation over their (typed) λ-terms. Each step in the proof is recorded as an operation in the semantics of the consequent of the sequent. More specifically, the rules of use for type-constructors $\backslash$ and $/$ correspond to functional

[98] In a semantically labelled sequent, the prosodic operation + of string concatenation is left implicit in the ordering of antecedent syntactic types. The implicit coding of this prosodic operation is, however, not sufficiently expressive to account for some linguistic discontinuous phenomena, the analysis of which require some form of Permutation. For the **L** sequent calculus with semantic and prosodic labels see, for example, [MG92]; [Jäg05].

application, while their rules of proof correspond to functional abstraction (or: λ-binding). The rule of proof for $\bullet$ corresponds to pairing; its rule of use, to (first and second) projection. The Cut rule, in turn, corresponds to substitution. Thus, the consequent of the endsequent of an **L**-proof records the semantic operations corresponding with the syntactic operations on types; the λ-term that labels the endsequent is a recipe for semantic composition. Thus, composition of linguistic form and meaning composition (derivational semantics) become aspects of one and the same process of grammatical inference.

As a first linguistic illustration of the semantically labelled **L** calculus, consider once again example 5.3 on page 98 and the lexicon below (recall that types $A/(B \bullet C)$ and $(A/C)/B$ are interderivables in **L**). We shall give more examples of semantically labelled proofs in Chapters 6 and 7. Following [Car97], we assign the description operator ι to the definite article; the former returns the unique individual denoted by its common noun argument.

the : $n/cn : \iota$

give : $((n\backslash s)/n)/n : \lambda x.\lambda y.\lambda z.(((give'\ (x))\ y)\ z)$

John : $n : j'$

Mary : $n : m'$

prize : $cn : prize'$

$$\dfrac{z : n \Rightarrow z : n \qquad \dfrac{\dfrac{y : cn \Rightarrow y : cn \qquad z : n \Rightarrow z : n}{v : n/cn, y : cn \Rightarrow z[(vy)/z] : n}\,/\,\mathrm{L} \qquad \dfrac{w : n \Rightarrow w : n \qquad u : s \Rightarrow u : s}{w : n, z : n\backslash s \Rightarrow u[(zw)/u] : s}\,\backslash\,\mathrm{L}}{w : n, x : (n\backslash s)/n, v : n/cn, y : cn \Rightarrow (zw)[(x(vy))/z] : s}\,/\,\mathrm{L}}{w : n, u : ((n\backslash s)/n)/n, z : n, v : n/cn, y : cn \Rightarrow ((x(vy))w)[(uz)/x] : s}\,/\,\mathrm{L}$$

Figure 5.8 – Labelled derivation for *John gives Mary the prize* in **L**

After substituting, the **L**-proof displayed in Fig. 5.8 delivers $(((uz)(vy))w)$ as the semantics. By replacing the variable u by the lexical meaning of *give*, the variable z by the lexical meaning of *Mary* — m' — and so on, we then obtain the functional term $((((((\lambda x.\lambda y.\lambda z.(((give'\ x)y)z))))\ m')(\iota\ prize))\ j')$. By successive β-conversions we are led to the semantic reading $(((give'\ m')\ \iota\ prize')\ j')$.

Unlike **LJ**, **L** does not contain any structural rules. Because of this, **L**-proofs corresponds to λ-terms without multiple-binding and without empty abstraction. Indeed, multiple-bind λ-terms correspond to the structural rule of Contraction. Moreover, λ-terms with empty (or: vacuous) abstraction correspond to Weakening and Expansion (see

Fig. 5.9).[99] In addition, the antecedent sequent of **L** cannot be empty. Thus, the only free variables contained in the term that labels the consequent of a sequent in an **L**-proof are those that also occur in the antecedent sequent. Figures 5.10 and 5.11 display two labelled **LJ**-proofs that exhibit the relation between these structural rules and (closed) λ-terms (or: combinators).[100] Table 5.2 on the following page displays a few **LJ**-sequents and their corresponding λ-terms and combinators.

$$\frac{X, x : A, y : A, Y \Rightarrow M : B}{X, y : A, Y \Rightarrow M[y/x] : B}\,\mathrm{C} \Rightarrow \qquad \frac{X, x : A, Y \Rightarrow M : C}{X, x : A, y : A, Y \Rightarrow M : C}\,\mathrm{E} \Rightarrow \qquad \frac{X, Y \Rightarrow M : C}{X, x : A, Y \Rightarrow M : C}\,\mathrm{W} \Rightarrow$$

Figure 5.9 – Labelled Contraction, Weakening and Expansion rules

$$\dfrac{\dfrac{x : A \Rightarrow x : A}{y : A, x : A \Rightarrow x : A}\,\mathrm{E} \Rightarrow}{x : A \Rightarrow \lambda y.x : A \to A}\,\to\mathrm{R} \qquad \dfrac{\dfrac{x : A \Rightarrow x : A}{y : B, x : A \Rightarrow x : A}\,\mathrm{W} \Rightarrow}{x : A \Rightarrow \lambda y.x : B \to A}\,\to\mathrm{R}$$

Figure 5.10 – Expansion and Weakening, Empty Abstraction and Combinator **K**

$$\dfrac{\dfrac{\dfrac{\dfrac{v : A \Rightarrow v : A \qquad \dfrac{x : A \Rightarrow x : A \qquad y : B \Rightarrow y : B}{z : A \to B, x : A \Rightarrow y[(zx)/y] : B}\,\to\mathrm{L}}{u : A \to (A \to B), v : A, x : A \Rightarrow (zx)[(uv)/z] : B}\,\to\mathrm{L}}{u : A \to (A \to B), v : A \Rightarrow ((uv)x)[v/x] : B}\,\mathrm{C} \Rightarrow}{v : A, u : A \to (A \to B) \Rightarrow ((uv)v) : B}\,\mathrm{P} \Rightarrow}{u : A \to (A \to B) \Rightarrow \lambda v.((uv)v) : A \to B}\,\to\mathrm{R}$$

Figure 5.11 – Contraction, Multiple Binding and Combinator **W**

5.5 Combinatory and Type-Logical Grammars

As we saw in the previous sections, **L** is a conservative extension of the **AB** categorial calculus. **L** extends **AB** in two ways: firstly, **L** increases the *set of rules* for the **AB** connectives. That is, along with the rules of use (or: elimination rules) for the directional slashes

[99] Of course, in a system that admits Expansion and/or Weakening, they can be obtained by assuming a form of the identity axiom with (possibly) non-empty antecedent contexts:

$$\frac{}{X, A, Y \Rightarrow A}\,\mathrm{Id}$$

[100] Note that the proof of the last sequent does not necessarily require Contraction; it may be obtained, alternatively, by using the following *dependent-context* left rule due to Ketonen [Ket44] (cf. [Pao02]:23).

$$\frac{X \Rightarrow A \qquad B, X \Rightarrow C}{A \to B, X \Rightarrow C}\to \mathrm{L'}$$

In Sec. 7.3.2 we will see that the left rule for the anaphoric type-constructor proposed by Jaeger [Jäg05] encodes (long-distance) Contraction in a similar way.

I(dentity)	$\lambda x.x$	Id	$A \Rightarrow A$
A(pply)	$\lambda x.\lambda y.(xy)$	$\to$L	$A, A \to B \Rightarrow B$
CI	$\lambda x.\lambda y.(yx)$	$\to$L+$\to$R	$A \Rightarrow (B \to A) \to B$
B (composition)	$\lambda x.\lambda y.\lambda z.x(yz)$	A$\Rightarrow$	$A \to B, B \to C \Rightarrow A \to C$
C(ommutor)	$\lambda x.\lambda y.\lambda z.(xz)y$	P$\Rightarrow$	$A \to (B \to C) \Rightarrow B \to (A \to C)$
K (constant)	$\lambda x.\lambda y.x$	E$\Rightarrow$	$A \Rightarrow (A \to A)$
K (constant)	$\lambda x.\lambda y.x$	W$\Rightarrow$	$A \Rightarrow (B \to A)$
W (duplicator)	$\lambda x.\lambda y.((xy)y)$	C$\Rightarrow$	$A \Rightarrow (A \to (A \to B)) \to B$
W (duplicator)	$\lambda x.\lambda y.((xy)y)$	C$\Rightarrow$+P$\Rightarrow$	$A \to (A \to B) \Rightarrow A \to B$
S(trong composition)	$\lambda x.\lambda y.\lambda z.(xz)(yz)$	P$\Rightarrow$+C$\Rightarrow$+A$\Rightarrow$	$A \to (B \to C) \Rightarrow (A \to B) \to (A \to C)$

Table 5.2 – Combinators, λ-terms, (logical and structural) rules and **LJ**-sequents

$\backslash$ and $/$, **L** also contains rules of proof (or: right rules) for the same. But **L** also extends the *language* of **AB**; in addition to $\backslash$ and $/$, **L** contains a product type-constructor $\bullet$. This two-fold extension of the Basic Categorial Grammar gives rise to two categorial approaches: a *rule-based* approach and a *type-logical* (or: deductive) approach [RTO88]; [SA96]; [Pag06]; [Moo11].

In a rule-based categorial approach, the **AB** system is extended by adding several new rules for its two connectives. On such an approach, the functional application ($A >$, $A <$) basis is augmented with a finite set of rules that are postulated as primitives; for example, Type-lifting, Functional Composition, Substitution. What makes a grammar of this kind a *Combinatory* Categorial Grammar (CCG) is that the semantic operation of each syntactic rule is equivalent to a combinator: **CI** for Type-lifting, **B** for Functional Composition, **S** for Substitution, and so on. Figure 5.13 on page 109 displays the principal rules that have been proposed in the CCG literature (cf. [Jac90]; [Sza92]:249; [Ste93], [Ste94], [Ste96]; [Jac96b] among many others).

A note on combinatory and type-logical notation: Henceforth, we use λ-terms instead of combinators to express the operations performed in the semantics. Thus, for example, we write $\lambda x.\lambda y.\lambda z.(x(yz))$ rather than $\mathbf{B}xyz$. Then, as above, $N : A$ stands for a λ-term N of syntactic category A.[101] In addition, for the sake of uniformity, we shall use the type-logical notation rather than the directional combinatory notation, even when we introduce a combinatory theory. For instance, we write $B/(A\backslash B)$ instead of $B/(B\backslash A)$ (as in Steedman's combinatory notation) in the rule $T >$ below and, analogously, we write $(B/A)\backslash B$ instead of $B\backslash(B/A)$ in $T <$. Indeed, note that a type $(X\backslash Y)$ in the combinatory notation stands for a functional type $(Y\backslash X)$ in the type-logical notation, that is, a type that takes an argument type Y on its left and returns a value of type X. In both notations a type (X/Y) stands for a functional type that takes an argument Y on its right and returns a value X. In more general terms, in the combinatory directional notation the leftmost type always stands for the value of the functional type and the slashes indicate the (right or left) position of its arguments.

[101] In the following chapters, we will simply use $>$ and $<$ instead of $A >$ and $A <$, respectively. Furthermore, we shall drop $>$ and $<$ in the other combinatory rules.

The combinatory rules employed by CCGs make it possible to significantly extend the linguistic scope of the basic **AB** grammar.[102] The type-lifting rule turns arguments into functions over functions over such arguments. Then, for example, a nominal may lift from the type n to a functional type $s/(n\backslash s)$ by means of T. In addition, this rule has also been used to treat quantifier scope ambiguities.

Thus, with the help of T, a lifted subject can compose (via the B rule) with a transitive verb to form a non-standard constituent of type s/n. The manoeuvre jointly performed by T and B then licenses a combinatory analysis of the phenomenon of extraction from the object position (see example in (181) and Fig. 5.12 below; cf. [Bal02]:28). In conjunction with T, B has been used in connection with coordination of (contiguous but) non-standard constituents such as *right node raising* constructions (see example in (182)). Similarly in the case of mixed $B \gtrless$ for extraction of non-peripheral arguments, as in (188), and for the so-called *heavy NP-shift* (see sentence in (187)). This is because directional mixed versions of **B** are non-order-preserving: each of them encodes a form of Permutation. With $B \gtrless$ (and also T) available in the system, constructions containing discontinuities — such as those triggered by separable phrasal verbs — can be handled (see sentences in (184)). In fact, by lifting the object type from n to $((n\backslash s)/n)/(n\backslash s)$ and by assigning the type $((n\backslash s)/n)/pr$ to the phrasal verb, the type $(n\backslash s)/pr$ may be obtained by $B \gtrless$. A generalized version of B, in turn, licenses constructions known as *left node raising*, as exemplified in (185) and (186) below.

Rule S, in conjunction with B, has been proposed to treat constructions with multiple (or *parasitic*) *gaps* [Sza89]; that is, constructions with two nominal positions bound by the same nominal. By means of the combinator **S**, the relative clause in (189) may receive the double-bind semantics $\lambda x.((filed'\ x)(without\ reading'\ x))\ j'$. In fact, like **W**, **S** acts as a duplicator: both encode a form of Contraction. For this reason, as we will soon see, combinators like these recur throughout the combinatory literature on anaphoric expressions.

(181) the team$_1$ that Brazil defeated $-_1$ [Bal02]:20

(182) Keat steals, and Chapman eats, apples. [Ste96]:33

(183) Keat cooks, and might eat, some apples.

(184) a. Marcos picked up the ball.
b. Marcos picked the ball up. [Bal02]:35

(185) I gave a teacher an apple and a policeman a flower. [Ste96]:45

[102] They also increase the context-free power of **AB** grammars. The class of languages generated by classical Categorial Grammars consists of the context-free languages not containing the empty string (cf. [Car05]:23). Analogously, the language generated by a finite lexicon under the **L**-rules is context-free. For the issue of generative capacity and also for the computational complexity of **AB** and their (CCG and TLG) extensions, see [HGS60]; [BMVB88]:Ch.2; [Car97]:151; [Bus03]:326; [Car05].

(186) John sent a letter to Mary and a book to Sue. [Lar88]:345

(187) Max sent to me the longest letter anyone had ever seen. [Lar88]:347

(188) a cake which I will buy on Saturday and eat on Sunday. [Ste92]:19

(189) ...the book$_1$ that John filed $-_1$ without reading $-_1$.

In the rule-based system CCG, parsing involves a process of derivation. Each leaf of a CCG-derivation tree is labelled with the syntactic and semantic information provided in the lexicon. Leaves are labelled with lexical semantics or, alternatively, with variables that are later replaced with lexical semantics. The root of the tree expresses the syntactic and semantic information derived by applying, in a top-bottom direction, the combinatory rules. Thus, grammatical reasoning leads from lexical categorizations to a syntactic type, which is labelled with a recipe for semantic composition; the latter is obtained from applying the combinators corresponding to the rules. In Fig. 5.12 below we give an example of a CCG-derivation for the relative clause in (181) above.

$$\dfrac{\dfrac{that}{\lambda x.x : (n\backslash n)/(s/n)}\,\text{lex} \qquad \dfrac{\dfrac{\dfrac{Brazil}{b' : n}\,\text{lex}}{\lambda u.(u\ b') : s/(n\backslash s)}\,\text{T} \quad \dfrac{defeated}{\lambda y.\lambda z.((defeat'\ y)\ z) : (n\backslash s)/n}\,\text{lex}}{\lambda v.\lambda u.(u\ b')(\lambda y.\lambda z.((defeat'\ y)\ z)\ v) : s/n}\,\text{B}}{\lambda v.((defeat'\ v)\ b') : n\backslash n}\,>$$

Figure 5.12 – Derivation for object relativization in CCG

While CCGs provide adequate derivations for several linguistic constructions that cannot be handled in **AB** grammars, the former can suffer the opposite problem: overgeneration. Since in CCG, like in **AB**, there are only two type constructors, it is not possible to control the application of these rules via lexical specification (but see [Bal02]). And, certainly, an unrestricted application of these rules can overgenerate (cf. [Dow88]). As Steedman [Ste96]:43s, [Ste92]:20 himself recognizes, it is "a key assumption in CCG that languages are free to restrict these combinatory rules to certain categories". Thus, for example, the type A in both T rules — in a grammar for English — is restricted to syntactic categories that are allowed as arguments of English verbs (cf. [Ste96]:36). And, since extraction mainly uses the B rule, island constraint on extraction is ensured by imposing a condition on the bridge-type which the rule applies to (specifically, $B \neq n$). Even more drastically, languages may exclude a given rule altogether. In the English grammar, for example, mixed $B \lessgtr$ is generally assumed to be inactive, as it can induce an ungrammatical word-order [Bal02]:31.

Whereas order-preserving versions of T, B, and D rules can all be derived in **L**, those that encode some structural rule other than associativity cannot. Thus, the combinatory analysis that uses only the former rules can also be given in **L**.

$$\frac{M : C/A \qquad N : A}{(MN) : C}\ A > \qquad\qquad \frac{N : A \qquad M : A\backslash C}{(MN) : C}\ A <$$

(Forward and Backward) Functional Application

$$\frac{M : A}{\lambda x.(xM) : B/(A\backslash B)}\ T > \qquad\qquad \frac{M : A}{\lambda x.(xM) : (B/A)\backslash B}\ T <$$

(Forward and Backward) Type-Raising/Lifting

$$\frac{M : C/B \qquad N : B/A}{\lambda x.(M(Nx)) : C/A}\ B > \qquad\qquad \frac{N : A\backslash B \qquad M : B\backslash C}{\lambda x.(M(Nx)) : A\backslash C}\ B <$$

$$\frac{M : C/B \qquad N : A\backslash B}{\lambda x.(M(Nx)) : A\backslash C}\ B \lesseqgtr \qquad\qquad \frac{M : B/A \qquad N : B\backslash C}{\lambda x.(M(Nx)) : C/A}\ B \gtreqless$$

(Forward, Backward and Mixed) Function Composition

$$\frac{M : B/A}{\lambda x.\lambda y.(M(xy)) : (B/C)/(A/C)}\ D > \qquad \frac{M : A\backslash B}{\lambda x.\lambda y.(M(xy)) : (A\backslash C)\backslash(B\backslash C)}\ D <$$

(Forward and Backward) Division (or Geach) Rules

$$\frac{M : (A/B)/C \qquad N : B/C}{\lambda x.((Mx)(Nx)) : A/C}\ S > \qquad\qquad \frac{N : C\backslash B \qquad M : C\backslash(B\backslash A)}{\lambda x.((Mx)(Nx)) : C\backslash A}\ S <$$

$$\frac{N : A/B \qquad M : (A\backslash C)/B}{\lambda x.((Mx)(Nx)) : C/B}\ S \gtreqless \qquad\qquad \frac{M : B\backslash(C/A) \qquad N : B\backslash A}{\lambda x.((Mx)(Nx)) : B\backslash C}\ S \lesseqgtr$$

(Forward and Mixed) Substitution (or Connection) Rules

Figure 5.13 – Rules of Combinatory Categorial Grammars

In order to increase the linguistic scope of **L**, the type-logical approach augments the set of **L**-type constructors. As in **L**, a rule of use and a rule of proof are given for each of the new connectives. These new connectives can be defined so as to encode some of the structural rules that are not available in the weaker system. Thus, for example, while $/$ and $\backslash$ are the order-sensitive slashes of **L**, a non-directional implication $\div$ can be defined in a conservative extension. Similarly for a commutative product $\circ$. Alternatively, in some Type-Logical Grammars structural rules are make available for a new unary connective. Thus, for example, Contraction, Permutation and/or Weakening may be assumed for a unary connective (! in [Gir87] and $\Box$ in [Mor90]; [Hep92], for example). For such modalized types only, all or some of the structural rules are allowed. Naturally, the availability of Permutation in a system is useful to give an analysis of non-peripheral extraction, for example. By means of (at least a restricted form of) Contraction, Type-Logical Grammars can handle parasitic gaps. Weakening, in turn, can be used to deal with cases of clitic duplication (in Spanish, for instance; see sentence in (180) above).

Once restricted versions of the structural rules are at hand, a type-logical grammar highly increases the empirical scope of the calculus. Due to the extension of the **L**-language, then, control of the structural rules can be made via lexical specification. This strategy depends on identifying the lexical items responsible for the more flexible behaviour. Thus, ditransitive verbs in English, say, can be typed by using a type-constructor that encodes Permutation; in Spanish, they may require a type built out of a connective encoding Weakening (along with Permutation).

In the following two chapters we turn to these two categorial approaches for extending the basic **AB** and **L** grammars. In particular, we focus on the different strategies for obtaining controlled forms of Contraction so as to deal with two kinds of anaphoric expressions: control verbs and pronouns.

6 Control in Categorial Grammar

6.1 Introduction

As we have extensively explained, control sentences seem to display a semantic-syntactic mismatch. In control sentences, the embedded subject position seems to receive a semantic value even though there is no lexical material at the surface level to contribute this value or, put differently, to fill the subject gap. In spite of this, the semantic subject of a controlled complement clause is understood as identical with one of the matrix arguments. As a result, in control structures a single lexical item of type NP appears to offer up its semantic value to fill two different argument slots: the controller and the subject controllee. More specifically, subject control verbs — e.g. *try*, *promise*, *want*, etc. — select a clause in which the semantic value of the embedded controllee subject coincides with the semantic value of the matrix subject. In object control structures, by contrast, it is the object complement which controls, and gives its semantic value to, the embedded subject. Hence, the adequate semantic representations of subject and object control sentences like (190) and (191) seem to be (190a) and (191a), respectively, despite the lack of an NP in the surface subject position of the infinitive *to*-clause.

(190) John tries [to leave].
a. $((try'\ (leave'\ j'))\ j')$

(191) John persuades Bill [to leave].
a. $(((persuade'\ (leave'\ b'))\ b')\ j')$

As we saw in Chapter 1, there are at least two ways to avoid generating such a mismatch between semantic and syntactic resources in subject and object control sentences: by changing the syntactic representation of the complement *to*-clause or, alternatively, by changing its semantic representation. With regard to the first option, a non-monostratal grammar could in fact countenance a non-surface level of syntactic representation in which the *to*-clause does have a syntactic, semantically controlled, subject. At this level, a *to*-infinitive clause would have the structure [NP VP], just like a finite clause. Such a grammar could then interpret the embedded infinitive clause in (190) as denoting the open proposition $(leave'\ x)$. And, by adopting some mechanism for binding the free variable x, the controlled clause in (190) would then denote the sentence $(leave'\ j')$.[103] The

[103]Nevertheless, the claim about a non-surface syntactic subject does not force this propositional representation. Remember that according to Landau, although PRO is the infinitive subject, the controlled *to*-clause denotes an unsaturated function: $\lambda x.(VP'\ x)$. In the generative tradition — but not in the Minimalist Program — variables are syntactic objects whose reference depend on a variable assignment

mismatch is then dismantled: the semantic level and the non-surface level of syntactic representation would thus converge once again.

However, this first option is not available to Categorial Grammars. For, insofar as they are monostratal, it appears that semantic representation is obtained from the meaning provided by surface lexical items only. This carries a number of consequences for a Categorial Grammar. First, it would immediately follow that there is no embedded semantic subject to be consumed by the *to*-infinitive clause.[104] This seems to further imply that the infinitive does not denote a proposition, but a property: $\lambda x.(VP'\ x)$. In syntactic terms, this in turn seems to imply that the complement of a control verb is not a sentence, but a *to*-VP. But if a control verb takes a VP denoting a property as its complement, the surface resources displayed in control structures are enough for semantic composition after all; the mismatch would also be dissolved from a categorial perspective.

But there is a further option available to Categorial Grammars, in addition to the one just described: to accept the semantic-syntactic mismatch and then compensate the deficit of semantic resources by adopting an appropriate mechanism for syntactic composition. More specifically, the idea is that by making available a syntactic mechanism — other than functional application — for duplicating lexical resources, a Categorial Grammar could then allow control verbs to take sentential arguments. With such a mechanism at hand it would then be possible to reuse the semantic value of the surface controller in order to assign it to the controllee "missed" subject.

6.2 Control and the Combinator W

In Section 5.4 we saw that the availability of a mechanism for duplicating the meaning of a lexical item in a logical grammar depends on the availability of the Contraction rule. This rule offers a procedure for deleting identical (contiguous) syntactic resources and, consequently, for multiple-binding at the semantic level. For this reason, the mismatch displayed by control sentences can be easily solved in a grammar that admits Contraction as a syntactic operation. To see this a little more clearly, let us examine the following

function. Variable binding is an operation that removes the dependency on assignment functions (see Section 7.1).

[104]The rejection of the first option is not just based on the architecture of these grammars. Adopting a semantic perspective, Chierchia [Chi84a], [Chi85] also rejects that an infinitive VP can combine with an NP subject. According to him, the sequence [NP *to*-VP] cannot be a syntactic constituent (in any point of the derivation) because it cannot — under any circumstances — be a semantic constituent. The meaning of a *to*-VP cannot combine with the meaning of an NP to form a proposition. This is because, according to Chierchia, infinitives do not denote a propositional function; in fact, they do not denote a function at all but a special sort of entities. Infinitive VPs are nominalized propositional functions and, as such, cannot take an NP as their argument. Although infinitives are syntactically VPs, they are semantically associated with a special kind of individuals: properties that have lost their unsaturatedness: nominalized properties.

intuitionistic (**LJ**-)labelled sequent proofs for the control sentences in (190–191) above.[105] For the purpose of this example, let the lexicon be as given below, where s_{to} stands for the type of a *to*-infinitive clause. In addition, note that by the η-reduction rule in Proposition 5.4, we have that $\lambda x.(leave'\ x) \rightsquigarrow_\eta leave'$, and so on.[106]

Bill : $n : b'$

John : $n : j'$

persuade : $((n\backslash s)/n)/s_{to} : persuade'$

to-leave : $n\backslash s_{to} : leave'$

try : $(n\backslash s)/s_{to} : try'$

$$
\dfrac{\dfrac{\dfrac{\dfrac{x : n \Rightarrow x : n \qquad y : s_{to} \Rightarrow y : s_{to}}{x : n, z : n\backslash s_{to} \Rightarrow y[zx/y] : s_{to}}\,\backslash\mathrm{L} \qquad \dfrac{\vdots}{v : n, u : n\backslash s \Rightarrow uv : s}}{v : n, w : (n\backslash s)/s_{to}, x : n, z : n\backslash s_{to} \Rightarrow (uv)[(w(zx))/u] : s}\,/\mathrm{L}}{v : n, x : n, w : (n\backslash s)/s_{to}, z : n\backslash s_{to} \Rightarrow ((w(zx))v) : s}\,\mathrm{P}\Rightarrow}{x : n, w : (n\backslash s)/s_{to}, z : n\backslash s_{to} \Rightarrow ((w(zx))v)[x/v] : s}\,\mathrm{C}\Rightarrow
$$

Figure 6.1 – Derivation for: *John tries to leave* in **LJ**

From the displayed **LJ**-proof we obtain the functional application-term $((w(zx))x)$; and, once lexical meanings are inserted, we obtain the semantic reading $((try'\ (leave'\ j'))\ j')$. It is important to note that it is only due to Contraction that the infinitive *to*-clause gets: i) a propositional reading and ii) an infinitive subject identical with the matrix one.[107]

Similarly, after inserting lexical meanings into the term $(((v(tx))x)u)$ corresponding to the **LJ**-proof in 6.2, we obtain $(((persuade'\ (leave'\ b'))\ b')\ j)$, where the infinitive *to*-clause also receives a controlled subject.

However, neither Permutation nor Contraction are fully available in a grammar based on **AB** or **L**.[108] Still, Categorial Grammars can strengthen their deductive capacity

[105] Though **LJ**, unlike **L**, does not have two slash connectives but only one (because of Permutation), we will use \ and / instead of $\rightarrow$ in our examples in order to highlight the comparison between **LJ** and Categorial proofs.

[106] As we indicated before, there is no agreement among (categorial) scholars on the type-assignment for (complement of) control and infinitive verbs. As Dowty [Dow85] notes, a (generative) tradition that goes back to Rosenbaum [Ros65] assumes that control verbs select a proposition. He calls this tradition the R-analysis of infinitive complements. Instead, the (categorial) analysis that starts from Montague's works, the M-analysis, assumes that subjectless complements are VPs denoting a property. From this perspective, a control verb would denote a relation between individuals and properties.

[107] In fact, this is the very same semantic mechanism as that behind the long-distance contraction rule used by Rosenbaum [Ros70]: his Equi-NP Deletion rule (see Section 2.2.1.)

[108] Before continuing, notice that if we instead assumed the type $((n\backslash s)/s_{to})/n$ for *persuade*, we would not use Permutation in the **LJ**-derivation of the object control sentence in (191). In this case the

$$
\dfrac{\dfrac{\dfrac{x : n \Rightarrow x : n \qquad \dfrac{y : s_{to} \Rightarrow y : s_{to} \qquad \begin{array}{c}\vdots\\ u : n, w : ((n\backslash s)/n), z : n \Rightarrow ((wz)u) : s\end{array}}{u : n, v : ((n\backslash s)/n)/s_{to}, y : s_{to}, z : n \Rightarrow ((wz)u)[(vy)/w] : s}\,/\mathrm{L}}{u : n, v : ((n\backslash s)/n)/s_{to}, x : n, t : n\backslash s_{to}, z : n \Rightarrow (((vy)z)u)[(tx)/y] : s}\,\backslash\mathrm{L}}{u : n, v : ((n\backslash s)/n)/s_{to}, x : n, z : n, t : n\backslash s_{to} \Rightarrow (((v(tx))z)u) : s}\,\mathrm{P}\Rightarrow}{u : n, v : ((n\backslash s)/n)/s_{to}, x : n, t : n\backslash s_{to} \Rightarrow (((v(tx))z)u)[x/z] : s}\,\mathrm{C}\Rightarrow
$$

Figure 6.2 – Derivation for: *John persuades Bill to leave* in **LJ**

by extending the basic systems **AB** or **L** with some restricted mechanism for duplicating semantic resources in the grammar. In Combinatory Categorial Grammar, this is the strategy pursued by Steedman [Ste88], who extends **AB** with (a version of) the combinator **W** to give a syntactic treatment of control structures: specifically, those generated from the verb *want*. In what follows we present Steedman's proposal in more detail. Recall, meanwhile, that the combinator **W** is a function which duplicates its second argument, and corresponds to the closed lambda-term $\lambda x.\lambda y.((xy)y)$, where the λ-operator binds two occurrences of the same variable (see Sec. 5.3).

The verb *want* gives rise to a subject control structure, as in (192), where the semantic subject of the infinitive verb does not have an overt realization. It also generates an ECM structure as in (193), where an overt nominal occupies the subject position of the infinitive verb. Despite this double syntactic structure generated by *want*, Steedman [Ste88] assigns it a single syntactic category, which takes an infinitive clause as its complement: $(n\backslash s)/s_{to}$. Semantically, *want* then denotes a relation between individual and propositions. [109] Given that the syntactic type of the complement of *want* — s_{to} — does not coincide with the type of the infinitive complement — $n\backslash s_{to}$ — the rule of functional application cannot be applied; another rule is called for to combine these types. Steedman uses the Equi rule in 6.3 specifically to deal with control structures (cf. also [SA96]). Note that, like the combinator **W**, the Equi-verb rule generates multiple-binding.[110] We label Steedman's proposal $\mathbf{S}_W$ for future reference.

(192) John wants [to leave].

(193) John wants [(for) Bill to leave].

to-leave : $n\backslash s_{to}$: $leave'$

nominal argument of the matrix verb (i.e. the controller) and the nominal argument of the embedded verb (i.e. the controllee) would already have been in the adequate adjacent position to apply Contraction. Nevertheless, under this type-assignment, the sentence would have obtained the semantic reading $(((persuade'\ b')(leave'\ b'))\ j')$.

[109] In this sense, Steedman departs from the categorial tradition, which generally assumes a VP argument (denoting a property): the M-analysis.

[110] Strictly speaking, the Equi-verb rule is a non-directional rule. Steedman's proposal also includes many other non-directional combinatory rules (see Section 5.5).

want : $(n\backslash s)/s_{to}$: $want'$

$$\frac{(A\backslash B)/C \qquad A\backslash C}{A\backslash B}\,\text{Equi}$$

Figure 6.3 – Equi-verb rule

$$\cfrac{\cfrac{\cfrac{\cfrac{\begin{array}{c}\vdots\\ x:A, y:A\backslash C \Rightarrow yx:C\end{array} \qquad \begin{array}{c}\vdots\\ z:A, w:A\backslash B \Rightarrow wz:B\end{array}}{z:A, u:(A\backslash B)/C, x:A, y:A\backslash C \Rightarrow (wz)[(u(yx))/w]:B}\,/\mathrm{L}}{x:A, z:A, u:(A\backslash B)/C, y:A\backslash C \Rightarrow ((u(yx))z):B}\,\mathrm{P}\Rightarrow}{z:A, u:(A\backslash B)/C, y:A\backslash C \Rightarrow ((u(yx))z)[z/x]:B}\,\mathrm{C}\Rightarrow}{u:(A\backslash B)/C, y:A\backslash C \Rightarrow \lambda z.((u(yz))z):A\backslash B}\,\backslash\mathrm{R}$$

Figure 6.4 – Labelled derivation of Equi in **LJ**

$$\cfrac{\cfrac{\cfrac{john}{n}}{s/(n\backslash s)}\,{>}\mathrm{T} \qquad \cfrac{\cfrac{wants}{(n\backslash s)/s_{to}} \quad \cfrac{to\text{-}leave}{n\backslash s_{to}}}{n\backslash s}\,\text{Equi}}{s}\,{>}$$

Figure 6.5 – Derivation for: *John wants to leave* in $\mathbf{S}_W$

The Equi rule combines the finite two-place control verb *want* and its *to*-uninflected verb complement, and returns the category of an intransitive verb $n\backslash s$; the latter then combines with the lifted matrix subject by forward application. Thus, the Equi rule allows us to combine the finite control verb and the infinitive complement despite the absence of an overt embedded subject. Put another way, this rule simulates the effect of duplicating the subject matrix argument. In fact, the labelled **LJ**-proof of Equi in Fig. 6.4 makes it explicit that the Equi rule corresponds to the semantic operation of multiple-binding. Thus, the semantic operation behind the syntactic operation performed by Equi supplies the infinitive clause with the controlled semantic subject.[111] The meaning of the intransitive complex verb phrase *wants to leave* obtained by the Equi rule will then be $\lambda z.((want' \ (leave' \ z)) \ z)$. In contrast, in the case of the ECM structure, the infinitive complement first combines with its own semantic subject by functional application before combining with the matrix finite verb. By using two different rules — Equi and $<$ — but a single lexical entry for *want*, Steedman's proposal may adequately recognize the double structure supported by this particular control verb. Nevertheless, as Steedman himself accepts, the Equi rule overgenerates, as it can be applied to any categories A, B, C without

[111] Because the proof of the Equi rule, like the proof of combinator **W**, uses $P \Rightarrow$ and $C \Rightarrow$, it cannot be derived in the **L** calculus.

restriction.[112] To see this even more clearly, consider the verb *believe* and the examples below:

(194) John believes [Bill to be sick].

(195) John believes (that) Bill leaves.

(196) *John believes leaves.

Just like *want*, this verb produces an ECM structure when it selects an infinitive complement, as in (194). In addition, *believe* may also take a finite complement clause—see (195). For this reason, in this case it seems that a single lexical entry is not adequate to recognize the possible combinations of *believe* and its different types of complements; we need more than one type-assignment. Nevertheless, if we assume the standard lexical entry indicated below alongside Steedman's, we may derive the grammatical examples by using functional application, and also the ungrammatical one in (196) as a consequence of the Equi rule.[113] In effect, the type $n\backslash s$ can be obtained by Equi from the types $(n\backslash s)/s$ and $n\backslash s$.

believe : $(n\backslash s)/s_{to}$

believe : $(n\backslash s)/s$

leave : $n\backslash s$

Steedman [Ste88] does not extend this syntactic, combinatory proposal for other control verbs, such as *try*, *persuade* and *promise*. Unlike *want*, these control verbs do not alternate with ECM constructions; consequently, they do not admit the possible occurrence of an overt nominal in the subject position of the infinitive complement. For this reason, in [Ste94] and [SB11] control verbs are assigned a category that no longer selects an infinitive clause, of type s_{to}, but rather an uninflected verb phrase $n\backslash s_{to}$.[114] As a consequence of this lexical assignment, the combination of a control verb and its infinitive complement may be obtained by functional application only (see Fig. 6.6).[115] In order to

[112] In fact, given that the slashes are non-directional, the Equi rule proposed by Steedman is even more powerful than that in Fig. 6.3.

[113] In order to derive (i) below by using Equi, it seems we need to assign the category $n\backslash s_{to}$ to the (small) clause *to-be sick*.

i. John believes to be sick.

[114] In Steedman's directional notation: **john** : $s/(s\backslash n)$, **persuade** : $((s\backslash n)/(s_{to}\backslash n))/n$, **to-leave** : $s_{to}\backslash n$, **want** : $(s\backslash n)/(s_{to}\backslash n) : \lambda x.\lambda y.(want'\ (xana'\ y))\ y)$, and so on.

[115] Actually, Steedman uses other combinatory operations besides functional application. He assumes, for example, combinator B, and uses it to combine the lexically lifted subject type and the matrix verb, giving then a incremental analysis of the sentence. This notwithstanding, our somewhat simplified presentation is sufficient to convey the main gist of his analysis of control.

$$\cfrac{\cfrac{\mathit{john}}{n} \qquad \cfrac{\cfrac{\cfrac{\mathit{persuade}}{((n\backslash s)/(n\backslash s_{to}))/n} \quad \cfrac{\mathit{bill}}{n}}{(n\backslash s)/(n\backslash s_{to})}< \qquad \cfrac{\mathit{to\text{-}leave}}{n\backslash s_{to}}}{n\backslash s}<}{s}>$$

Figure 6.6 – Derivation for: *John persuades Bill to leave* in $\mathbf{S}_W$

ensure the semantic control relation, Steedman proposes a lexical approach: he encodes into the lexical meaning of these verbs the duplicating operation previously performed by the Equi rule at the syntactic level.[116]

persuade : $((n\backslash s)/(n\backslash s_{to}))/n : \lambda x.\lambda y.\lambda z.(((persuade'\ (y\ (x)))\ x)\ z)$

promise : $((s\backslash n)/(n\backslash s_{to}))/\diamond n : \lambda x.\lambda y.\lambda z.(((promise'\ (y\ (z)))\ x)\ z)$

to-leave : $n\backslash s_{to} : \lambda x.(leave'\ x)$

try : $(n\backslash s)/(n\backslash s_{to}) : \lambda x.\lambda y.((try'\ (x\ (y)))\ y)$

want : $(n\backslash s)/(n\backslash s_{to}) : \lambda x.\lambda y.((want'\ (x\ (y)))\ y)$

Given that the syntactic rules applied in the previous derivation correspond to the semantic operation of functional application, and given that multiple-binding is encoded into the lexical meaning of the control verb, the sentence gets the expected controlled reading: $(((persuade'\ (leave'\ b'))\ b')\ j')$.

Unlike his syntactic combinatory account, Steedman's lexical account might be accepted in other non-Combinatory Categorial Grammars. At this point, it is important to recall that the multiple-binding restriction on λ-terms does not apply in the lexicon, but at the syntactic level of representation, since it is a consequence of the lack of Contraction (or any other rule for duplication).[117] Thus, at a first glance, the latter of Steedman's accounts seems only to require rules that are already theorems of the **L** calculus. Nevertheless, a more careful examination of the derivation given above reveals that functional application rules can only be used under a specific type-assignment to the object (and

[116]Steedman [SB11] defines *c-command* and Principles of Binding Theory at the level of logical form. $ana'x$ and $pro'x$ are proforms in which the argument x is identical to some other node in the logical form. In particular, the subject of the infinitive verb at the level of logical form in (object) subject control structures is the proterm ana' bound to the matrix (object) subject. Thus, for example: **persuade** : $((s\backslash n)/(s_{to}\backslash n))/n : \lambda x.\lambda y.\lambda z.(persuade'\ (y\ (ana'\ x))\ xz)$ and **want** : $(s\backslash n)/(s_{to}\backslash n) : \lambda x.\lambda y.(want'\ (x\ (ana'\ y))\ y)$ (cf. also [Ste94]). The diamond operator $\diamond$ in the type of *promise* intends to block the passive construction. Also Chierchia [Chi88] builds Control Agreement Principle into the lexical entries of the control verbs: **force** : $((n\backslash s)/n_i)/(n_i\backslash s)$ and **promise** : $((n_i\backslash s)/(n_i\backslash s))/n$, where the index i ranges over the set of number, gender and person features.

[117]In other terms, the λ-terms associate with **L**-proofs are without multiple-binding, but not necessarily so the λ-terms that represent the meaning of the lexical items.

subject transitive) control verbs: $((n\backslash s)/(n\backslash s_{to}))/n$. Indeed, if we instead assign the type $((n\backslash s)/n)/(n\backslash s_{to})$ to *persuade* (or to *promise*), the derivation will no longer be possible by using only the **L** rules for / and \.

The first syntactic type does however have an advantage over the second: it correctly mirrors the linear order of the arguments of the control verb. Under the former type-assignment and by applying only the rules of **L**, a Categorial Grammar can derive the adequate prosodic chain: **persuade+Bill+to-leave**. In contrast, under the second type-assignment, another logical rule for syntactic (and prosodic) composition will be required in order to combine the control verb with its arguments and — simultaneously — to obtain the correct linear word-order, as opposed to **persuade+to-leave+Bill**.

Despite their different argument ordering, both types suffer the same disadvantage: they cannot display a hierarchical relation between the arguments themselves, nor therefore with respect to the functor. Both types — $((n\backslash s)/(n\backslash s_{to}))/n$ and $((n\backslash s)/n)/(n\backslash s_{to})$ — express the fact that the nominal (internal) argument — of type n — and the infinitive complement — of type $n\backslash s_{to}$ — maintain the same structural relation with respect to the control verb. This is because a type $A/(B \bullet C)$ is equivalent in **L** to the type $(A/C)/B$. Both functional types — simultaneously or consecutively — take arguments B and C *in this linear order* to return a type A; the type constructors $\bullet$ and / of **L** only express the ordering of concatenating lexical items.

6.3 Control Verbs and Wrap

Unlike Steedman, Bach [Bac79] assigns different types to the subject control verb *promise* and to the object control verb *persuade*, despite the fact that both subcategorize an NP and an infinitive *to*-VP (197a–197b) (see also [Chi88]). As we will be seeing shortly, one of the main reasons for this is that the operations whereby subject and object control verbs combine with their nominal complement (if there is one) are essentially different. Furthermore, the separate type-assignment for *promise* and *persuade* is also motivated by well-known linguistic data. In more specific terms, their behaviour is at odds with several linguistic phenomena: passives (198a–198b), extraction of the *to*-infinitive (199a–199b) and pseudoclefts (200a–200b), binding (201a–201b), heavy NP shift [Ros67] (202a–202c), extraction of the NP-object (203a–203b), among others (cf. [Lar91]).[118]

[118] In fact, by noting the different behaviour of *promise* on the one hand, and *persuade/force* on the other with respect to double-object constructions (ia-b) and oblique dative complements (iia-b), Larson regards *promise* as a dative verb.

i. a. *John persuades Bill a car.
   b. John promises Bill a car.

ii. a. *John persuades a car to Bill.
   b. John promises a car to Bill.

(197) a. John persuades Bill to buy a car.

b. John promises Bill to buy a car.

(198) a. Bill was persuaded to buy a car (by John).

b. *Bill was promised to buy a car (by John).

(199) a. *What did John persuade Bill?

b. What did John promise Bill?

(200) a. *What John persuaded Bill was to buy a car.

b. What John promised Bill was to buy a car.

(201) a. John persuades the men [to like themselves/each other/*himself].

b. John promises the men [to like himself/*themselves/*each other].

(202) a. John persuades to do the homework the three kids that were hanging around the house.

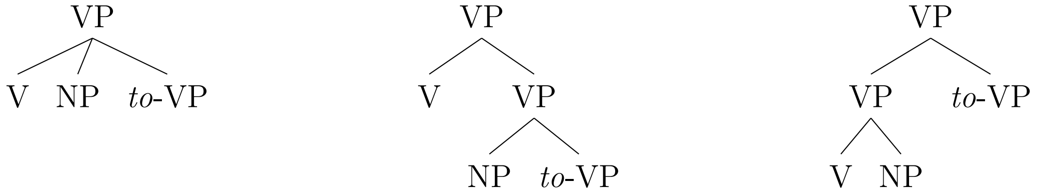
b. *John promises to do the homework the three kids that were hanging around the house.

c. ?John promises to do the homework to the three kids that were hanging around the house.

(203) a. Who do you think John persuades to leave?

b. *Who do you think John promises to leave?

The above contrasts suggest that *promise* on the one hand, and *persuade* on the other, maintain different relationships with their respective complements. While *promise* seems to have a closer connection with its nominal complement than with its infinitive complement, the opposite appears to be true of *persuade*. This in turn implies that the NP complement on the one hand, and the infinitive *to*-VP complement on the other, stand in different structural relationships with each of these control verbs. The appearance of a uniform, flat and common structure [V – NP – *to*-VP] then disappears. In light of this, only two more possible structures are left: namely, [V [NP – *to*-VP]] or [[V – NP] *to*-VP]. The three alternatives are represented in tree format below:



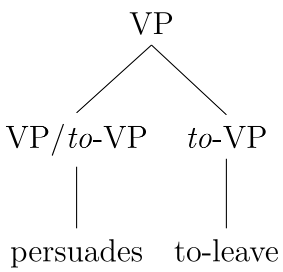
The second diagram illustrates how the infinitive *to*-VP first combines with its nominal (external) complement NP, and then this complex constituent combines with the matrix V.[119] The third expresses the fact that first the main verb combines with its nominal (internal) argument, and then this syntactic constituent combines with the infinitive *to*-VP. It is important to note that, despite the differences between these three structures, they are all (binary or ternary) continuous chains; the syntactic operation whereby a functional type takes its arguments mirrors the prosodic operation of concatenation of strings.

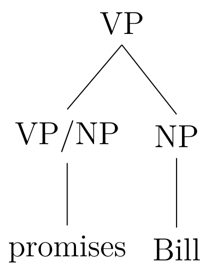
Now note that in control structures, the NP is the (internal) argument of the control verb, not of the infinitive; as a result, the second diagram is not adequate to also categorize control verbs. Bach notes that while the third structure correctly represents the relationship established between the subject control verb *promise* and its two complements, an alternative structure is needed to categorize *persuade*. In order to display the syntactic relationship between the latter and its complements, Bach proposes a discontinuous structure: namely, [[V *to*-VP] NP]. This conveys the fact that the object control verb combines with the *to*-VP before combining with the object NP argument. In sum, although both verbs take an NP and a *to*-VP arguments, they do so one in the opposite order to the other: while the first argument of *promise* is the nominal, that of *persuade* is the infinitive verb. This notwithstanding, the rule underpinning these combinations is the same, namely *RCon* [Bac88].[120] We present this and its complement rule *LCon* below.

Let $a \in X : \mathbf{s}$ stay for a linguistic expression a of syntactic category X and prosodic form $\mathbf{s}$, then:

If $a \in X/Y : \mathbf{s}_1$ and $b \in Y : \mathbf{s}_2$, then $RCon(a, b) \in X : \mathbf{s}_1{+}\mathbf{s}_2$ [RCon Rule]

If $a \in Y \backslash X : \mathbf{s}_1$ and $b \in Y : \mathbf{s}_2$, then $LCon(a, b) \in X : \mathbf{s}_2{+}\mathbf{s}_1$ [LCon Rule]





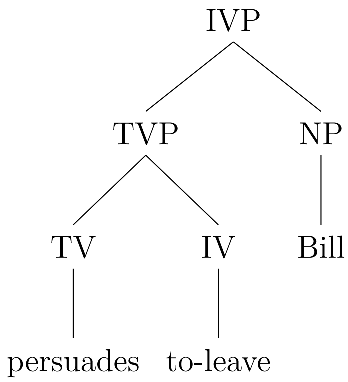
Although there is only one rule here, two kinds of semantic predicate and syntactic phrase emerge as a result. Indeed, a string of words like *promise Bill* is a predicative

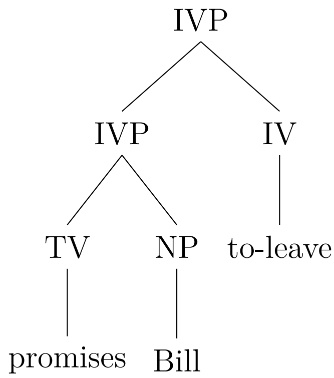
[119]We remind the reader that according to Chierchia [Chi85], the sequence NP *to*-VP does not even form a semantic unit, since the latter does not denote a function, but a (special sort of) individual (an action or a state), and semantic composition requires that a constituent be obtained by functional application. As a consequence, such a sequence does not also form a syntactic unit. Nevertheless, the sequence NP *to*-VP could form a semantic-syntactic constituent if there were some operator, like *for*, that denominalizes the infinitive VP. Thus, for example, *for Bill to leave* is a constituent.

[120]This rule is also used in the analysis of control structures generated by *try*.

intransitive verb phrase (IVP); whereas the constituent *persuade to leave* is a predicative transitive verb phrase (TVP).[121] The latter is the phrasal counterpart of a transitive verb (TV); the former, of an intransitive verb (IV) (cf. also [Dow82]).[122] Consequently, the complex *persuades to leave* is a semantic predicative unit and a syntactic constituent, even though it does not directly correspond to any phrase at the surface level (i.e. the prosodic level): it is a discontinuous constituent.

We mentioned earlier on that functional application rules—and therefore $RCon$—are not enough for the discontinuous transitive verb phrase *persuade to leave* to be combined with its NP object *Bill* and, simultaneously, to reach the correct representation at the prosodic level: **persuade+Bill+to-leave**. By using $RCon$ only, the best we can do is obtain **persuade+to-leave+Bill** and, of course, **promise+Bill+to-leave**.





In order to combine a transitive predicative verb phrase with its nominal object, Bach then assumes a new, discontinuous, wrap operation ($RWrap$) along with the aforementioned concatenative, continuous operation $RCon$:

17. a. If $a \in X/Z : \mathbf{s}_1+\mathbf{s}_2$ and $b \in Z : \mathbf{s}_3$, then $RWrap(a, b) \in X : \mathbf{s}_1+\mathbf{s}_3+\mathbf{s}_2$.
    b. If $a \in X/Z : \mathbf{s}_1$ and $b \in Z : \mathbf{s}_2$, then $RWrap(a, b) = RCon(a, b) \in X : \mathbf{s}_1+\mathbf{s}_2$.

[RWrap rule]

Observe the following analysis trees (see [Dow82]) and type assignment. We use Lambek's notation for the latter, corresponding to Bach's categories for NP, Vt/ViP and Vi.[123]

[121] According to Larson [Lar91], the constituent *persuade to leave* is a small predicate and the nominal object *Bill* is its subject. Thus, he assigns the non-surface structure [$_{VP}$Bill [$_{V'}$persuade to leave]] to the VP.

[122] According to Bach, TVP, unlike IVP, can undergo a passive transformation. Passivization can apply to transitive VPs, not only to transitive verbs. From this it follows that *persuades to leave* can passivize, but *promise Mary* cannot. Bach's proposals can then account for *Visser's generalization*: subject control verbs do not passivize.

[123] In Bach's proposal, each syntactic category is represented as a pair formed by: i) the lexical category of the element (N for noun, V for verb, Adj for adjective, and so on); ii) the categorial index which

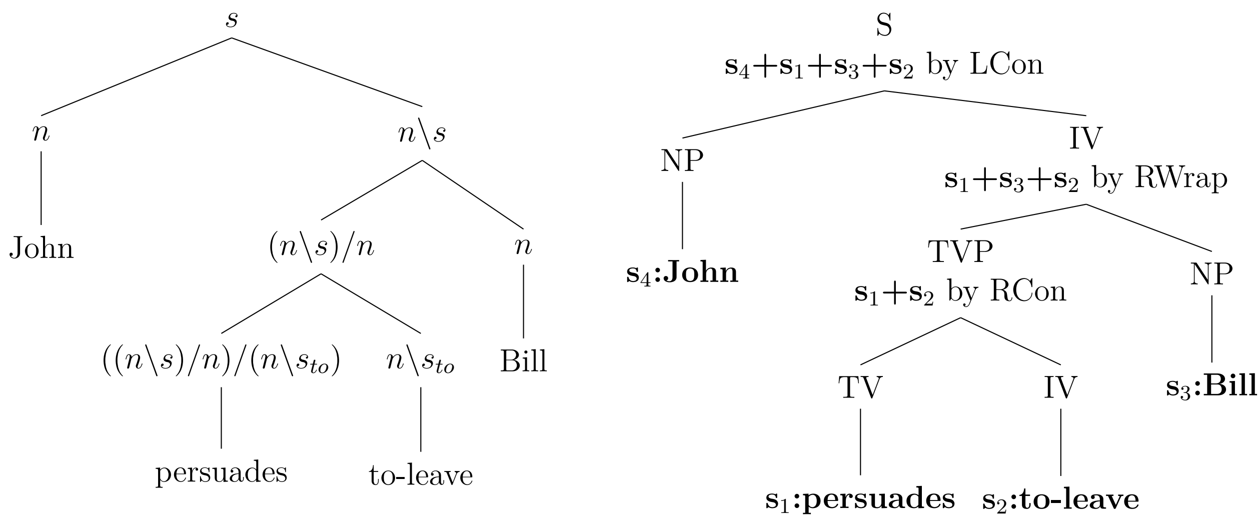


Figure 6.7 – Analysis trees for: *John persuades Bill to leave* in Bach's proposal

Bill : n

John : n

persuade : $((n\backslash s)/n)/(n\backslash s_{to})$

to-leave : $n\backslash s_{to}$

The crucial step in this derivation is that in which the functional transitive verb-type $(n\backslash s)/n$ combines with the argument type n. The syntactic (and semantic) operation behind this step is the standard functional application rule incorporated into the RWrap rule. The result of this is the intransitive verb-type $n\backslash s$ with the semantic $((persuade'\ (leave'))\ b')$. Importantly, at the prosodic level $RWrap$ performs a non-standard, non-concatenative operation: it intercalates the string $\mathbf{s}_3$, corresponding to the nominal argument of the function, between the two constituents of the prosodic complex chain $\mathbf{s}_1+\mathbf{s}_2$ corresponding to the functor. To "wrap" the functor expression around the object means that the string $\mathbf{s}_3$ is placed to the right of the first constituent of the complex string $\mathbf{s}_1+\mathbf{s}_2$ to form the string $\mathbf{s}_1+\mathbf{s}_3+\mathbf{s}_2$. Hence, $RWrap$ performs a discontinuous

determines the semantic type, where the functional category t/e corresponds to the intensional expression $<< s, e >, t >$, where $< s, e >$ is a function from possible words to the set of entities e and t denotes the set of truth values [Mon74]. With these lexical and semantic basic categories at hand, Bach may distinguish among different verbal categories and semantic predicates. Consider, as an example, the following:

Vi (intransitive verb) (ex. *run*): $[V, (t/e)]$

Vt (transitive verb) (ex. *kiss*): $[V, (t/e)/[N, t/[V, (t/e)]]]$

Vto (verb with a to-infinitive complement) (ex. *try*): $[V, (t/e)/[V, (t/e)]]$

Then, Bach's Vi category corresponds to $n\backslash s$ in Lambek's notation, Vt corresponds to $(n\backslash s)/n$, and Vto, to $(n\backslash s)/(n\backslash s_{to})$. Moreover *promise* is assigned the category Vto/NP, which corresponds to $((n\backslash s)/(n\backslash s_{to}))/n$, just like in Steedman's proposal.

prosodic operation. Whether $RWrap$ is seen as a rule for inserting a string into a discontinuous or separable prosodic chain or, alternatively, as a rule for enclosing the former by the discontinuous chain, $Wrap$ allows for the order of the arguments to be different from the surface order of the constituents.[124] Consequently, $RWrap$ allows us to obtain the correct surface ordering **persuade+Bill+to-leave** despite following the reverse order for combining functional and argument types.[125]

If the infinitive and nominal arguments are combined in any other order, the dissimilar control properties exhibited by the three-place verbs *persuade* and *promise* can be explained from a single assumption: the NP that controls the "missed" infinitive subject is the nominal argument of the verb phrase that is taken immediately after the infinitive predicative phrase. The different (subject or object) control patterns displayed by *promise* and *persuade* then follow from the different syntactic structure of these verbs. In the case of *persuade*, the object NP is added immediately after the infinitive, and is therefore the controller. In the case of *promise*, the object NP is taken before the infinitive; here, it is the NP subject that combines immediately after the infinitive, and thus achieves control. Therefore in both cases the controller of the "missed" subject is the nearest nominal argument to the verb phrase that contains the subjectless infinitive.

[124] In fact, the (right) wrap operation goes hand at hand with its reverse, (left) infix operation: $RWrap(a, b) = LInfix(b, a)$. In more strict terms, since the complex string $\mathbf{s}_1+\mathbf{s}_2$ that undergoes the wrap operation is assumed as a discontinuous chain, it should be written as $\mathbf{s}_1 \ldots \mathbf{s}_2 = (\mathbf{s}_1, \mathbf{s}_2)$ (cf. [Hep94]; [MM96]:135; [SA96]:180). This discontinuous chain, of course, cannot be obtained by the functional application rule. In more recent and contemporary type-logical grammars, the wrap operation is defined in terms of the discontinuous type-constructor $\uparrow$.

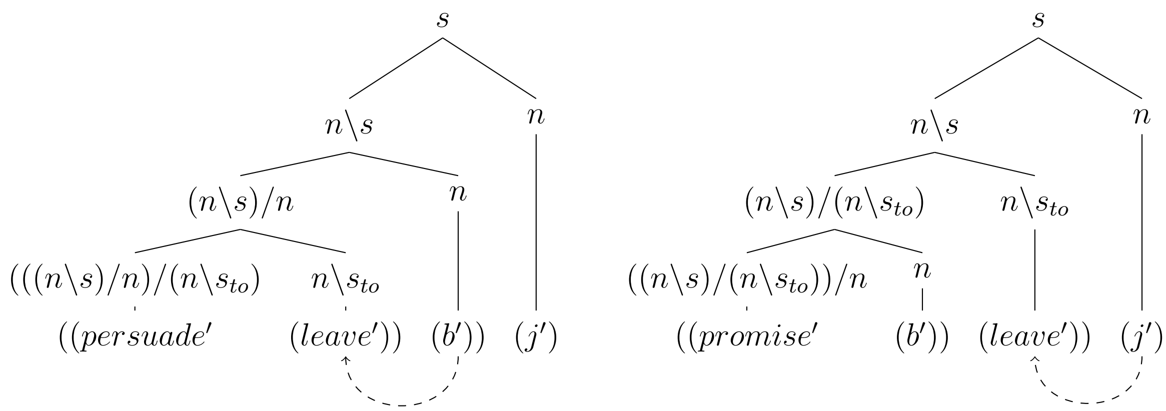
[125] Moortgat [Moo88] defines two discontinuous type-constructors: extraction $\uparrow$ and infixation $\downarrow$. A type $B \uparrow A$ features a discontinuous functional type that wraps around its argument A to form a continuous type B, and a type $B \downarrow A$ features a functor than infixes itself in its A argument to form a type B. These type-constructors have a non-concatenative interpretation in the string algebra. For the infixation type-constructor $\downarrow$ and for the extraction type-constructor $\uparrow$ a rule of use ($L \downarrow$) and a rule of proof ($R \uparrow$), respectively, are given. These rules constitute the partial logic of discontinuity [Moo88]. In order to formulate the remaining two rules (i.e. $L \uparrow$ and $R \downarrow$) the structure of the sequent has to be enriched with prosodic strings (cf. [SA92]; [Moo96a]; [MSA93]).

$$\frac{X, Z \Rightarrow A \qquad Y, B, W \Rightarrow C}{Y, X, B \downarrow A, Z, W \Rightarrow C} \downarrow \mathrm{L} \qquad\qquad \frac{X, A, Y \Rightarrow B}{X, Y \Rightarrow B \uparrow A} \uparrow \mathrm{R}$$

These type-constructors have been successfully applied in type-logical grammar to deal with numerous linguistic phenomena: quantifier scope, reflexives, parasitic gaps, non-peripheral extraction, non-constituent coordination, among many others [Dow97]. The categorial literature about discontinuity is vast and ever-growing (cf., for example, [Hep94]; [Mor94b], [Mor95], [Mor95], [MV10b], [Mor11]; [MVF11]; [MV12]; [Val13]; [MV14], [MV15]; [Mor17a]). While much interesting work is being done in this area, devoting any more attention to it would take us too far afield. We will briefly come back to this topic in Section 7.2.1, where we will examine a lexical proposal for reflexives based on discontinuous type-constructors.

Another approach to word-order variation is given by a long-distance form of functional type $A \nearrow B$ (or: $A//B$ in Bach's [Bac84]:273 notation). A long-distance type $A \nearrow B$ features a functor that looks anywhere to the right; it can then be applied to a non-adjacent or distant argument B to result in A (cf. also [SA96]:102). This type-constructor so encodes Permutation.

$$\frac{X \Rightarrow M : A \qquad Y, x : B, Z \Rightarrow N : C}{Y, y : B \nearrow A, Z, X \Rightarrow N[(yM)/x] : C} \nearrow \mathrm{L} \qquad\qquad \frac{X, x : A, Y \Rightarrow M : B}{X, Y \Rightarrow \lambda x.M : B \nearrow A} \nearrow \mathrm{R}$$



6.4 Summing Up

In this chapter we examined three categorial proposals on the linguistic phenomenon of control. In particular, we focused on two concerns. First, we looked at the problem of ensuring the control relationship in Contraction-free logical grammars. In light of Steedman's proposals we showed that there are (at least) two ways to compensate for the lack of Contraction and still account for control. The first is to adopt a combinatory rule — the Equi-rule — that allow us to duplicate resources at the syntactic level. The second is to ensure the identity of controller and controllee in the lexical entry of the control verb. We then turned our attention to three-place subject and object control verbs, such as *promise* and *persuade/force*. We saw that these verbs exhibit rather different patterns of behaviour, and that for this reason Bach assigns them different categories, thus breaking the initial appearance of syntactic uniformity. All three proposals are able to account for the phenomenon of control; needless to say, each comes with its own set of background assumptions and commitments. In the remainder of the section we sketch the ramifications of each route, and simultaneously draw a few important lessons from the foregoing discussion.

1. If (the infinitive argument of) the control verb is assigned a type that is different from that of the infinitive verb, functional application cannot be used to combine them. In this case, another logical rule must be used. In particular, a syntactic rule that simulates type-duplication, such as the combinator **W**, is very useful for obtaining a propositional, multiple-bounded, controlled reading.

2. Thus, conversely, if functional application is the only available syntactic operation to combine the control verb and its (infinitive) subjectless complement, the complement of the control verb and the infinitive itself have to be assigned the same functional type — $n\backslash s_{to}$.

3. In this case however the control relation cannot be the result of a syntactic operation,

and must instead be encoded into the lexical semantics of the control verb (or through meaning postulates).

4. The control relation can be encoded into the lexical meaning of the control verb only if infinitive VPs are assigned a functional type $n\backslash s_{to}$, denoting a function from the set of individuals to the set of the denotations of s_{to}.

5. Even though the same functional syntactic type $n\backslash s_{to}$ is assigned to both the infinitive argument position of the control verb and the infinitive itself, the availability of the concatenation rule of functional application also depends on the order of the complements in the object control type. Specifically, if the object control verb is assigned the functional type $((n\backslash s)/n)/(n\backslash s_{to})$, it cannot directly take its infinitive complement as its argument and, simultaneously, result in the correct word-order.

6. Therefore, if this same functional type is assigned to object control verbs, as opposed to the type $((n\backslash s)/(n\backslash s_{to}))/n$, a rule for prosodic composition other than concatenation has to be assumed.

7. While this is the route chosen by Bach, it is not available from a type-logical perspective. Instead, the language of **L** can be enriched with a discontinuous functional type-constructor and a discontinuous prosodic operation as its interpretation.

7 Anaphoric Pronouns in Categorial Grammar

7.1 Introduction

In Generative semantics pronouns (and traces) are considered variables. Variables are syntactic objects (at some non-surface level of syntactic representation) whose semantic value may vary across assignment functions. A (variable) assignment function a is a partial function from the set $\mathbb{N}$ of natural numbers into the set D_e of entities. Thus, such a function assigns potentially different individuals to different numerical indices. Furthermore, the interpretation function $[\![.]\!]$ assigns appropriate denotations to phrase structure trees. By means of these functions, pronouns may receive a value by the following (Trace and) Pronouns Rule [HK98]:[126]

8. If α is a pronoun (or a trace), a is a variable assignment, and $i \in dom(a)$, then $[\![\alpha_i]\!]^a = a(i)$. [Pronouns Rule]

According to this rule, pronouns (and traces) do denote an entity but only under an assignment function a.

Definition 7.1 (Variables)**.** A terminal symbol α is a *variable* iff there are variable assignment functions a and a' such that $[\![\alpha]\!]^a \neq [\![\alpha]\!]^{a'}$.

Definition 7.2 (Constant)**.** A terminal symbol α is a *constant* iff for any two variable assignment functions a and a', $[\![\alpha]\!]^a = [\![\alpha]\!]^{a'}$.

An occurrence of a variable may be free (in a tree) or it may be bound. Variable binding is a semantic mechanism whereby an expression binds an occurrence of a free variable.[127] As such, variable binding reduces the dependency of variables on assignment functions. Thus, the semantic value of a bound variable, as opposed to a free variable, is invariant across assignment functions. Some linguistic expressions are responsible for binding (an occurrence of) a variable, thereby for reducing the assignment dependency of the latter. Insofar as they are variables, pronouns too may be free or bound. Typically,

[126] A phrase structure tree consists of a finite set of (labelled) linearly ordered nodes. As an example of interpretation function, consider: $[\![NP]\!] \in D_e$ and $[\![IV]\!] \in D_{<e,t>}$, and so on.

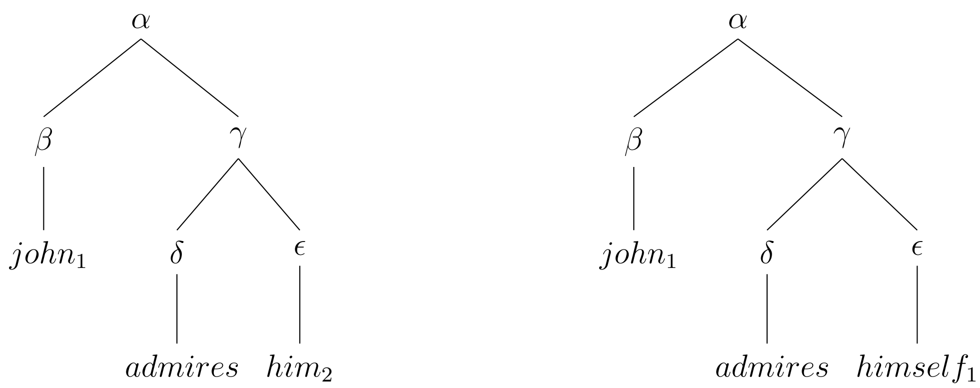
[127] This notwithstanding, the inputs to semantic interpretation are phrase structure trees provided by the *syntax*. Because of this, the formal definition of *variable binder* in this framework implies some configurational relationships such as, for example, *immediate constituents* on (*binary*) *trees*. As a consequence of the definition, a hierarchical relation is established between the binder and the bindee; in particular, the former must *c-command* the latter.

quantifier phrases bind pronouns (and relative pronouns bind traces); though pronouns may also be bound by nominal phrases.[128] Nevertheless, according to Principle A of GB Theory, anaphors must be bound. Therefore, in particular, reflexive pronouns are variables that are necessarily bound. By contrast, personal pronouns must be free (in some syntactic domains). Thus while the semantic value of the sentence (204) does not depend on an assignment function, that of (205) does.

(204) John_1 admires himself_1.

(205) John_1 admires him_2.

If α is a terminal node of a tree that is occupied by a lexical item, the interpretation of α under an assignment function a — $[\![\alpha]\!]^a$ — is specified in the lexicon. Note that lexical items are in the domain of the function $[\![.]\!]^a$ for every assignment a. As a particular consequence, lexical items are in the domain of $[\![.]\!]^{\emptyset} = [\![.]\!]$. Assume then that the meaning of the terminal node δ, occupied by the transitive verb *admires*, is the function $\lambda x.\lambda y.((admire'\ x)\ y)$; and that the meaning of β, occupied by *John*, is the individual j'. Finally, let the semantic value of the terminal node ϵ be given by the Pronouns Rule above: $[\![him_2]\!]^a = a(2)$ and $[\![himself_1]\!]^a = a(1) = j'$. Then, the interpretation (for the α node) of (204) and (205) can be obtained by functional application rules, as shown in the third tree diagram below:



[128] Strictly speaking, pronouns are assumed to be bound by the trace left by the moved (raised) quantifier antecedent. The crucial difference between variables in general, and traces in particular, is that the latter are assumed to be formed via (raising) movement. The movement (of a relative pronoun, for example) creates a λ-operator/trace configuration where the former binds the trace. As a consequence, there is no trace in the lower projection independently of the higher operator, i.e. its binder; traces are variables necessarily bound.

$$\llbracket \alpha \rrbracket^a = \llbracket \gamma \rrbracket^a \ (\llbracket \beta \rrbracket^a)$$

$$\llbracket \beta \rrbracket^a = \llbracket \beta \rrbracket \qquad \llbracket \gamma \rrbracket^a = \llbracket \delta \rrbracket \ (\llbracket \epsilon \rrbracket^a)$$

$$\llbracket \delta \rrbracket^a = \llbracket \delta \rrbracket \quad \llbracket \epsilon \rrbracket^a = \llbracket \epsilon \rrbracket^a$$

Application of the same diagram to each of the other two trees delivers the following, respectively:
$\llbracket John_1 \ admires \ him_2 \rrbracket^a = \llbracket \lambda x.\lambda y.((admire' \ x) \ y) \rrbracket \ (\llbracket him_2 \rrbracket^a) = \llbracket \lambda y.((admire' \ (a(2))) \ y) \rrbracket = ((admire' \ (a(2))) \ (a(1))) = ((admire' \ (a(2))) \ j')$.

$\llbracket John_1 \ admires \ himself_1 \rrbracket^a = \llbracket \lambda x.\lambda y.((admire' \ x) \ y) \rrbracket \ (\llbracket himself_1 \rrbracket^a) = \llbracket \lambda y.((admire' \ (a(1))) \ y) \rrbracket = ((admire' \ (a(1))) \ j') = ((admire' \ j') \ j')$.

Given that the argument $a(2)$ of the function $\lambda x.\lambda y.((admire' \ x) \ y)$ is, in turn, a function, the sentence in (205) receives a truth-value only if the domain of the assignment function a includes the index 2. The meaning of this sentence is then not a truth value, but a function: specifically, a function from the variable assignment function a to the set of truth-values. In other terms, a sentence containing a free occurrence of a pronoun may denote a truth-value only with respect to an assignment function.

In contrast, Categorial Grammars adopt a variable-free semantics (or: no-variable semantics). Variables are not considered syntactic objects in a semantics of this kind. Consequently, Categorial Grammars also do away with mechanisms for variable binding and variable assignment functions (cf. [Sza89]; [Jac99]). These grammars implicitly or explicitly adopt what is called the *hypothesis of direct surface compositionality*: the compositional semantics directly assigns each (and only) syntactic expression a model-theoretic interpretation [Jac96b]. The only level of syntactic representation that contributes to the compositional semantics is the surface structure. This, in turn, implies the *hypothesis of local interpretation*: each surface syntactic expression has a meaning [Jac96a].

In light of these semantic principles and the assumed semantic-syntactic correspondence, any categorial approach to pronouns must address the following questions:

- If variables are not syntactic objects, which syntactic category do pronouns belong to?
- Given the semantic-syntactic correspondence, what is the denotation of a pronoun?
- What is the meaning of an expression containing a free pronoun?
- Given that an anaphoric pronoun gets its semantic value from its antecedent, what is the mechanism by which a multiple-binding representation is obtained?

- What is the difference, if any, between anaphoric and bound pronouns?
- Is it possible to encode the Principles of Binding Theory in a Categorial Grammar?

In order to see more clearly how Categorial Grammars handle anaphoric pronouns, it is important to keep a few important points in mind. The first is that variables, and terms of the λ-calculus more generally, are commonly used for semantic representation in Categorial Grammars. Variables are only a part of the language used to represent meanings; thus, it would be possible to dispense with variables by using combinators. Secondly, expressions of the λ-calculus play a role at two levels of representation in these logical grammars: they are used for derivational representation (semantic-syntactic level) and also for lexical representation (lexicon). In addition, while the λ-terms associated with proofs of Categorial Grammars do not admit multiple-binding (due to the lack of Contraction), there is no such restriction on the semantic representation given in the lexicon. In fact, we saw in Chapter 6 that there are two ways to obtain a multiple-binding representation for control sentences, one lexical and one syntactic. These two routes are also available when dealing with anaphoric pronouns in variable-free Categorial Grammars. One option is to assign a multiple-binding λ-term to the meaning of the pronoun in the lexicon. The other is to yield a multiple-binding λ-term as the derivational representation of the clause in which the anaphoric pronoun occurs. This latter option implies, of course, that the logical calculus has to be extended in order to encode some form of Contraction. As we saw in Section 5.5, this extension may go by way of adding new rules to the **AB** system or, alternatively, by extending the vocabulary (and so also the set of rules) of **L**. By combining these alternatives, we obtain four routes to treating anaphoric pronouns:

Lexical approach in Combinatory Categorial Grammar: assigns a multiple-binding representation, in the lexicon, to the syntactic pronominal type built out of the $/$ and $\backslash$ connectives.

Lexical approach in Type-Logical Grammar: assigns a multiple-binding representation, in the lexicon, to the syntactic pronominal type built out of some new type-constructors.

Syntactic approach in Combinatory Categorial Grammar: obtains multiple-binding by means of the addition of new (semantic and syntactic) operations on the $/$ and $\backslash$ connectives.

Syntactic approach in Type-Logical Grammar: obtains multiple-binding as a result of (logical or structural) operations on some new type-constructors.

In the following sections we sketch the general ideas behind each of these categorial routes. We will then focus on two Type-Logical systems developed by Jaeger: **LA** [Jäg97], [Jäg98] and **LLC** [Jäg05].[129]

7.2 Routes to Treating Anaphoric Pronouns

7.2.1 Multiple-Binding in the Lexicon

Szabolcsi's [Sza89], [Sza92] Combinatory account and Morrill's [Mor00], [Mor03], [MV10a], [MV14] Type-Logical lexical account to reflexives both go back to Montague's [Mon74] earlier proposal of treating all noun phrases as generalized quantifiers. A similar idea can be traced back to Lambek's [Lam58], where functional — lifted or raised — types are assigned to, specifically, anaphoric nominals.

Recall that **L**, unlike **AB**, additionally contains right rules for \ and /. Therefore, any expression of category A must also be in categories $B/(A\backslash B)$ and $(B/A)\backslash B$, for any type B (see Proposition 5.2). Consequently, when A is a nominal basic type n and B is a type s, the former can also receive a derived lifted category $s/(n\backslash s)$ (as well as $(s/n)\backslash s$). Given that Combinatory Grammars, which are based on **AB**, generally admit rules for type-raising — >T and <T — , these higher syntactic types can also be derived in such systems.[130] Therefore, a lifted type may either be assigned to a nominal in the lexicon or be derived from the basic type n by using the lifting-type rules of the logical system.

The functional type $s/(n\backslash s)$ may then be used to categorize nominative (or preverbal) pronouns, and subjects in general; it expresses the fact that a nominal may combine with an intransitive verb-type $n\backslash s$ on its right, to form a sentential type s. By the syntactic-semantic correspondence, moreover, a lifted syntactic type corresponds to a functional semantic category. Specifically, $s/(n\backslash s)$ denotes a function of type $< t, < e, t >>$. Thus for example, in addition to the entity j', the nominal *John* also denotes the function $\lambda x.(x\ j')$, where the variable x is of type $< e, t >$.[131] In a similar way, the lifted type $((n\backslash s)/n)\backslash(n\backslash s)$ appears to be useful to categorize a nominal object, since it can combine with — or saturate the object argument position of — a transitive verb-type $(n\backslash s)/n$ on its left to return an intransitive verb-type $n\backslash s$.

Syntactically, pronouns behave like nominal phrases, as they may occupy nominal positions. In this sense, pronouns may also be assigned lifted syntactic types. Their se-

[129] For a good critical review of each approach, except for Jaeger's, see [Jäg05] (cf. also [KO03]: Ch. 10).

[130] Recall that type-raising operations correspond to the combinator **CI** of Combinatory Logic: $\mathbf{CI}xy = yx$ (cf. [Gre15]:284, and see Table 5.1).

[131] A sentence like *John runs* receives the single representation $(run'\ j')$ despite the fact that the latter results, in **L**, from two different λ-terms: $\lambda x.(run'\ x)(j')$, which corresponds to the final sequent $n, n\backslash s \Rightarrow s$, and $\lambda x.(x\ j')(\lambda y.(run'\ y))$, which corresponds to $s/(n\backslash s), n\backslash s \Rightarrow s$. The problem of assigning the same semantic reading to two different derivations is known as *spurious ambiguity* (cf. [Car97]; [Mor00]).

mantic behaviour, however, differs from nominals inasmuch as they lack an independent semantic value. In order to refer, pronouns have to pick up their denotation from another linguistic denotative expression: the binder. Thus, when it is bound, a (personal, possessive, reflexive) pronoun gets its denotation by reusing the semantic value of its binder. In this sense a bound pronoun acts like a duplicator of the binder's semantics. Accordingly, and given that reflexive pronouns are necessarily bound, Szabolcsi and Morrill respectively assign the combinator $\mathbf{W}$ or, equivalently, the function $\lambda x.\lambda y.((xy)y)$ to the *lexical meaning* of these pronouns. The combinator $\mathbf{W}$—and therefore the reflexive—returns the function $\mathbf{W}f$ from a (non-necessary) two-place function f by identifying its arguments, i.e: $\mathbf{W}fx = ((fx)x)$.

To this extent, then, the two accounts are indistinguishable from a semantic point of view. The main difference between these lexical combinatory and type-logical approaches rests on the syntactic type assigned to anaphoric pronouns. While Szabolcsi assigns them a category built out of **AB** type-constructors, Morrill opts for categories built out of type-constructors other than Lambek's: namely, a discontinuous type $(A \uparrow B) \downarrow A$. In general terms, a sign a has category $C \downarrow A$ iff wrapping a discontinuous constituent of category C around a returns a continuous constituent of type A. A sign a has category $A \uparrow B$ iff it is a discontinuous constituent which results a continuous constituent of type A if it is wrapped around a sign of type B. By way of briefly illustrating their proposals, consider the following combinatory and type-logical assignments, respectively, for subject-oriented reflexives:

himself : $((n\backslash s)/n)\backslash(n\backslash s) : \mathbf{W}$

himself : $((n\backslash s) \uparrow n) \downarrow (n\backslash s) : \lambda x.\lambda y.((x\ y)\ y)$

Accordingly, the reflexive pronoun *denotes* the function $\lambda x.\lambda y.((x\ y)\ y)$, where x is a functional type (for example: $< e, t >$) and y is of type e. In more linguistic terms, this means that the subject-oriented *himself* is itself a function that identifies both arguments of the transitive verb it applies to. As a linguistic illustration of Szabolcsi's assignment, consider the following simple derivation:[132]

This derivation shows, in the first step, that when the reflexive pronoun syntactically combines with the transitive verb *admires*, which denotes the two-place function $f : \lambda x.\lambda y.((admire'\ x)\ y)$, the semantic operation $\mathbf{W}f$ performed by the reflexive on the transitive verb equals the multiple-binding function $\lambda z.((admire'\ z)\ z)$. Finally, when this function applies to the subject-denotation j', it returns the reading $((admire'\ j')\ j')$. It is

[132] We remind the reader that $\lambda x.T(T') \rightsquigarrow_\beta T(T'/x)$, if every occurrence of the term T' is free in $T(T'/x)$, where $T(T'/x)$ stands for the substitution of variables x for T' in T. Then, we may spell out the semantic operation in the following β-reduction steps: $\lambda z.\lambda w.((z\ w)\ w)(\lambda x.\lambda y.((admire'\ x)\ y) \rightsquigarrow_\beta \lambda w.((\lambda x.\lambda y.((admire'\ x)\ y)\ w)\ w) \rightsquigarrow_\beta \lambda w.(\lambda y.((admire'\ w)\ y)\ w) \rightsquigarrow_\beta \lambda w.((admire'\ w)\ w)$.

$$\dfrac{\dfrac{John}{j' : n} \qquad \dfrac{\dfrac{admires}{\lambda x.\lambda y.((admire'\ x)\ y) : (n\backslash s)/n} \qquad \dfrac{himself}{\lambda z.\lambda w.((z\ w)\ w) : ((n\backslash s)/n)\backslash(n\backslash s)}}{\lambda w.((admire'\ w)\ w) : n\backslash s}>}{((admire'\ j')\ j') : s}>$$

Figure 7.1 – Derivation for: *John admires himself* in Szabolcsi's proposal

important to note that the reflexive does not really duplicate the denotation of its binder, but rather identifies both — subject and object — *argument slots* of the two-place verb.

In light of the above, Szabolcsi's assignment faces a first obstacle when the reflexive combines with (i.e. applies to) a three-place verb, such as *present, send, give, show.*[133] Verbs of this kind may select two nominal phrases as complements, and thus give rise to double-object constructions (see sentence in (206)). These structures tolerate another pattern of reflexivization: reflexives bound by a nominal in an object position. Thus, double-object constructions allow not just for subject-oriented but also for object-oriented reflexives (see sentences in (207–208)).

(206) Mary showed/gave/sent John a gift.

(207) John_1 sent/gave himself_1 a gift.

(208) John_1 showed/presented Bill_2 $\text{himself}_{1/2}$.

In order to clearly present the problem posed by three-place verbs, we only briefly review the key aspects of Szabolcsi's lexical combinatory proposal for (subject-oriented) reflexives: 1) Syntactically, a reflexive is assigned a *lifted-type.* 2) Semantically, a reflexive *denotes* **W**. 3) As a consequence, a reflexive is not an argument of a transitive verb, but rather a *a function* that takes the latter as its argument. 4) The reflexive is not directly *bound by* the binder, but by *the nominal* (subject) *slot.* 5) As a result, the *bound reading* of the (subject-oriented) reflexive *is necessarily obtained* when the multiple-binding function $((\mathbf{W}f)x)$ or $\lambda x.((f'\ x)\ x)$ finally applies to the subject-denotation. 6) Thus, the *binder* will be the *second nominal argument* of the two-place verb which the reflexive combines with. 7) This guarantees a specific case of the *prominence condition*: the binder cannot be situated below the bindee in the hierarchy.

Although Szabolcsi's lexical treatment appears to accurately deal with subject-oriented reflexives in transitive verbs, it does not work as well for these same reflexives in double-object constructions, nor for object-oriented reflexives. To see why, let us put the penultimate point in more general terms. In order for a reflexive to be bound, it must combine with the function that contains the binder gap *before* the function combines with the actual binder. Since three-place verbs select two internal (or object) arguments, a

[133]For *promise* see footnote [118].

reflexive has to combine with the verb before the multiple-binding function containing the verb immediately combines with the subject or object binder; this is in order for the former to become subject- or object-oriented, respectively. For subject-oriented reflexives, this only works if the reflexive stands in a peripheral position, and so if the non-reflexive object first combines with the double-object verb (see Fig. 7.2).

$$\dfrac{\dfrac{\mathit{John}}{j' : n} \qquad \dfrac{\dfrac{\dfrac{\dfrac{\mathit{presents}}{\lambda x.\lambda y.\lambda z.(((\mathit{present}'\ x)\ y)\ z) : ((n\backslash s)/n)/n} \quad \dfrac{\mathit{Bill}}{b' : n}}{\lambda y.\lambda z.(((\mathit{present}'\ b')\ y)\ z) : (n\backslash s)/n}\,< \qquad \dfrac{\mathit{himself}}{\lambda u.\lambda w.((u\ w)\ w) : ((n\backslash s)/n)\backslash(n\backslash s)}}{\lambda w.(((\mathit{present}'\ b')\ w)\ w) : n\backslash s}\,>}{}}{(((\mathit{present}'\ b')\ j')\ j') : s}\,>$$

Figure 7.2 – Derivation for: *John*$_1$ *presents Bill himself*$_1$ in Szabolcsi's proposal

While non-peripheral subject-oriented reflexives, as well as object-oriented reflexives, can syntactically combine with the three-place verb by changing the lifted-type of the reflexive, other problems arise. In particular, object-oriented reflexives face the familiar word-order problem reappears (see Fig. 7.3).

himself : $(((n\backslash s)/n)/n)\backslash((n\backslash s)/n) : \lambda x.\lambda y.((x\ y)\ y)$

present : $(((n\backslash s)/n)/n : \lambda x.\lambda y.\lambda z.(((\mathit{present}'\ x)\ y)\ z)$

$$\dfrac{\dfrac{\mathit{John}}{j' : n} \qquad \dfrac{\dfrac{\dfrac{\mathit{presents}}{\lambda x.\lambda y.\lambda z.(((\mathit{present}'\ x)\ y)\ z) : ((n\backslash s)/n)/n} \quad \dfrac{\mathit{himself}}{\lambda u.\lambda w.((u\ w)\ w) : (((n\backslash s)/n)/n)\backslash((n\backslash s)/n)}}{\lambda w.\lambda z.(((\mathit{present}'\ w)\ w)\ z) : (n\backslash s)/n}\,> \qquad \dfrac{\mathit{Bill}}{b' : n}}{\lambda z.(((\mathit{present}'\ b')\ b')\ z) : n\backslash s}\,<}{(((\mathit{present}'\ b')\ b')\ j') : s}\,>$$

Figure 7.3 – Derivation for: **John presents himself*$_1$ *Bill*$_1$ in Szabolcsi's proposal

To sum up: in Szabolcsi's account, semantic composition triggered by **W** requires the functional application ordering $(\lambda x.\lambda y.((x\ y)\ y)(\lambda z.\lambda w.\lambda u.(((\mathit{verb}'\ z)\ w)\ u)))(\mathit{binder}')$, which corresponds to the syntactic combination ordering ((Verb, reflexive), binder). However, this conflicts with the fact that the prosodic chain must exhibit the reverse complement's ordering [[Verb+binder]+reflexive].

The reader may recall that we encounter a similar situation in Section 6.3, while discussing the case of the three-place control verb *persuade*. The strategy we adopted then seems apt to the present case: namely, employing a wrap-like operation so as to obtain the correct word-order, and simultaneously satisfy the prominence condition. This is where the discontinuous type-constructor comes in. Under the assignment-type $((n\backslash s)\uparrow n)\downarrow(n\backslash s)$, non-peripheral subject-oriented reflexives work in double-object constructions. Here we

$$\dfrac{\dfrac{\dfrac{\dfrac{\vdots}{n, (n\backslash s)/(n \bullet n), n, n \Rightarrow s}}{(n\backslash s)/(n \bullet n), n, n \Rightarrow n\backslash s}\,\backslash\mathrm{R}}{(n\backslash s)/(n \bullet n), [\,], n \Rightarrow (n\backslash s) \uparrow n}\,\uparrow\mathrm{R} \qquad \dfrac{\vdots}{n, n\backslash s \Rightarrow s}}{n, (n\backslash s)/(n \bullet n), ((n\backslash s) \uparrow n) \downarrow (n\backslash s), n \Rightarrow s}\,\downarrow\mathrm{L}$$

Figure 7.4 – Derivation for non-peripheral subject-oriented reflexives in Morrill's proposal

do not present the rules for the discontinuous type-constructors $\uparrow$ and $\downarrow$; in their stead, please consider the following schematic syntactic derivation in Fig. 7.4. In this derivation, note that while $\uparrow$R creates a nominal gap (marked by introducing the separator [ ]) in the antecedent sequent, $\downarrow$L fills this gap. Given the multiple-binding lexical assignment for the reflexive, such a proof could deliver the adequate semantic reading $(((send'\ j')\ \iota\ gift')\ j')$ for the sentence in (207) above (cf. [MV10a]).

Since double-object constructions alternate with oblique dative structures, the same conflict resurfaces when the object-oriented reflexive occupies the prepositional argument. As a result, the lexical combinatory and type-logical treatment of reflexives calls for several syntactic assignments depending on their orientation and their syntactic position. Though the lexical multiple-binding proposal may be adapted to deal with anaphoric pronouns, this option needs to augment the inventory of syntactical types — and so, lexical entries — further still. In fact, in their earlier proposals Szabolcsi and Morrill extended their lexical account to also deal with anaphoric and free pronouns. Thus, on such accounts free and bound pronouns receive different semantic treatments, despite being morphologically identical.

Here we shall not examine these authors' approaches any further, however, for two reasons. The first is that Szabolcsi and Morrill's accounts of anaphoric pronouns fall out of — and are therefore conceptually dependent on — their accounts of reflexives. Thus a full and satisfactory evaluation of the former necessarily presupposes an in-depth analysis of the latter. But this, in turn, would lead us much too far off from the topic of this thesis: overt anaphoric pronouns.[134] The second reason is that most recently, these scholars have opted for a syntactic treatment of pronouns; they respectively adopt a Combinatory and Type-Logical syntactic approaches that follow those developed by Jacobson [Jac96a], [Jac96b], [Jac99], [Jac03] and Jaeger [Jäg05], respectively (cf. [Sza03]; [MV10a], [MV14]).

Thus, in the following section we devote ourselves to reviewing Jacobson's Combinatory approach, which is in turn the starting point of Jaeger's Type-Logical approach.

[134] Although we shall not spend any more time on the topic of reflexives here, this should not be taken to imply that we regard it as devoid of interest. Indeed, and on the contrary, we are already planning a post-doctoral research proposal precisely on this topic (cf. also [Cor17a]).

7.2.2 A Bridge between Combinatory and Type-Logical Approaches

One of the most important aspects of Jacobson's proposal, from the perspective of our research agenda, is that its focus is on pronouns rather than on reflexives. Moreover, Jacobson does not distinguish either semantically or syntactically between free pronouns on the one hand, and bound pronouns on the other; on the contrary, she presents a uniform account of the two. Free and bound pronouns are then assigned the same meaning and the same syntactic category in the lexicon. In particular, personal pronouns denote the identity function $\lambda x.x$ over individuals and carry the functional type n^n.[135] The latter is an instance of a new syntactic type X^Y, which results from extending the set of **AB** types. In general, if X and Y are types, then X^Y is also a type in Jacobson's system.

Jacobson assumes that the difference between free and bound pronouns only rests on the fact that the former never chances upon a binder, so to speak. On the other hand when a pronoun denoting $\lambda x.x$ happens to be bound, it denotes an entity: specifically, whichever entity is denoted by its binder. On Jacobson's account, pronouns are bound as a result not of a (multiple-binding) lexical assignment, but of the application of a logical rule. In other terms, the duplication mechanism is not ascribed to the meaning of the bound pronoun itself, but rather to the syntax. Among other thing, this tells us that Jacobson adopts a syntactic approach to binding.

In fact, her account is not merely syntactic but also *combinatory*: indeed, she augments the inferential power of **AB** grammars with two new combinatory rules, **G** and **Z** (cf. [Jac03]:61).[136] Let us then name Jacobson's proposal $\mathbf{J}_{GZ}$ for short. Given an expression δ, let $[\delta]$ (roughly) indicate the phonological form of δ, and let δ' indicate its meaning.

9. Let α be an expression of the form $< [\alpha]; A/B; \alpha' >$. Then there is an expression β of the form $< [\alpha]; A^C/B^C; \lambda x.\lambda y.(\alpha'(x(y))) >$, where x is a variable of type $< C', B' >$ and y is of type C'. [**G** rule]

10. Let α be an expression of the form $< [\alpha]; (A/n)/B; \alpha' >$. Then there is an expression β of the form $< [\alpha]; (A/n)/B^n; \lambda x.\lambda y.(\alpha'(x(y))(y)) >$. [**Z** rule]

The **G** rule owes its name to the Geach rules; in effect, the semantic operations performed by **G** and by the Geach rules are identical (see Fig. 5.13). Syntactically, **G** is also a division rule but, unlike the Geach rules, the former divides both the result and the argument types by using the new type A^C. **G** is a bridge-like combinator: it allows a standard **AB** syntactic type to combine with the new pro-form type A^C. Assuming that

[135] In her semantic proposal, Jacobson follows Hepple [Hep90], [Hep92]; nevertheless, Hepple assigns different modal syntactic types to reflexives and pronouns: $n/\ominus n$ and $\Box n/\ominus n$, respectively.

[136] Jacobson also admits type-raising and function composition rules.

any expression A that contains (or is itself) a pro-form C is of type A^C, **G** allows any functional expression A/B to take the pro-form B^C as its argument — for instance, via the functional application rule. In this case, the resulting type — A^C — will display the fact that it contains a pro-form. Thus, the resulting type inherits the following functional syntactic and semantic information: the functional pro-form A^C contains a gap of syntactic type C with unsaturated semantics $\lambda y.(\alpha' \ (\beta' \ (y)))$, where y is of semantic type C'. Thus, in particular, any expression A that contains a pronominal gap will be of type A^n with semantics $\lambda y.(\alpha' \ (\beta' \ (y)))$, where y is of type e. Yet more specifically, a sentence containing a (free) pronoun will be of type s^n denoting the function $\lambda y.(\varsigma' \ y)$, since β' is simply the identity function over individuals. In general, a sentence s containing i free pronouns will be of type s^{n_i}. Thus, the (generalized version of the) **G** rule syntactically and semantically displays a record of how many free pronominal gaps a sentence contains. Sentences with pronominal gaps do not denote a proposition, but rather a propositional function, that is, a function from the set of entities to the set of truth values.[137]

Before we examine the $\mathbf{J}_{GZ}$ system at work, a notational remark is in order: the single slash / used in our previous formulation of the **G** and **Z** rules actually stands for \ and / in **L** (or: $/_L$ and $/_R$ in Jacobson's notation). Consequently, those rules abbreviate the directional versions, which we display in a tree-format for the sake of readability and uniformity (see Figs 7.5 and 7.6 on the following page). In addition, hereinafter, we shall use Jaeger's notation $A|C$ instead of Jacobson's A^C.[138]

$$\frac{M : A/B}{g_0(M) = \lambda x.\lambda y.(M(x(y))) : (A|C)/(B|C)}\ G_0 > \qquad \frac{M : B\backslash A}{g_0(M) = \lambda x.\lambda y.(M(x(y))) : (B|C)\backslash(A|C)}\ G_0 <$$

$$\frac{M : A|C}{g_n(M) = \lambda z.(g_{n-1}(Mz)) : G_{n-1} > (A)|C}\ G_n > \qquad \frac{M : A|C}{g_n(M) = \lambda z.(g_{n-1}(Mz)) : G_{n-1} < (A)|C}\ G_n <$$

Figure 7.5 – (Forward, Backward and Generalized) G Rules

In order to illustrate the role of **G**, as well as its interaction with other combinators, we display a few simple derivations (see Figs. 7.7–7.10; and cf. [Jac96a]:117, [Jac99]). In particular, the derivation in Fig 7.10 on page 138 shows the use of the generalized form of **G**, which is needed in order to derive sentences containing more than one free pronoun.

(s)he/her/him : $n|n : \lambda x.x$

[137] We remind the reader that in generative semantics, a sentence containing a free pronoun also denotes a function, not a proposition. Nevertheless, in this framework free pronouns, and variables in general, are interpreted via variable assignment functions (see Sec. 7.1). The variable-free semantics adopted by Jacobson can strip the assignment function away.

[138] This notation should not be confused with Jacobson's $A|C$, which stands for a type A with a C-type extraction gap [Jac03]:80.

$$\dfrac{M : (A/n)/B}{\lambda x.\lambda y.(M(x(y))(y)) : (A/n)/(B|n)}\; Z > \qquad \dfrac{M : B\backslash(n\backslash A)}{\lambda x.\lambda y.(M(x(y))(y)) : (B|n)\backslash(n\backslash A)}\; Z <$$

$$\dfrac{M : (n\backslash A)/B}{\lambda x.\lambda y.(M(x(y))(y)) : (n\backslash A)/(B|n)}\; Z \gtrless \qquad \dfrac{M : B\backslash(A/n)}{\lambda x.\lambda y.(M(x(y))(y)) : (B|n)\backslash(A/n)}\; Z \lessgtr$$

Figure 7.6 – (Forward, Backward and Mixed) Z Rules

$$\dfrac{\dfrac{He}{\lambda x.x : n|n} \qquad \dfrac{\dfrac{lost}{lose' : n\backslash s}}{\lambda y.\lambda z.(lose'\ (y\ (z))) : (n|n)\backslash(s|n)}\; \mathrm{G}_0 <}{\lambda z.(lose'\ z) : s|n}\; <$$

Figure 7.7 – Derivation for: *He lost* in $\mathbf{J}_{GZ}$

$$\dfrac{\dfrac{\dfrac{\dfrac{Mary}{m' : n}}{\lambda y.(y\ m') : s/(n\backslash s)}\; \mathrm{T} >}{\lambda v.\lambda u.((\lambda y.(y\ m'))\ (v\ (u))) : (s|n)/((n\backslash s)|n)}\; \mathrm{G}_0 > \qquad \dfrac{\dfrac{\dfrac{admires}{admire' : (n\backslash s)/n}}{\lambda z.\lambda w.(admire'\ (z\ (w))) : ((n\backslash s)|n)/(n|n)}\; \mathrm{G}_0 > \qquad \dfrac{him}{\lambda x.x : n|n}}{\lambda w.(admire'\ w) : ((n\backslash s)|n)}\; >}{\lambda w.((admire'\ w)\ m') : s|n}\; >$$

Figure 7.8 – Derivation for: *Mary admires him* in $\mathbf{J}_{GZ}$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{John}{j' : n}}{\lambda x.(x\ j') : s/(n\backslash s)}\; \mathrm{T} > \qquad \dfrac{said}{say' : (n\backslash s)/s}}{\lambda z.((say'\ z)\ j') : s/s}\; \mathrm{B} >}{\lambda x.\lambda y.((say'\ (x\ (y)))\ j') : (s|n)/(s|n)}\; \mathrm{G}_0 > \qquad \dfrac{\dfrac{he}{\lambda y.y : n|n} \qquad \dfrac{\dfrac{lost}{lose' : n\backslash s}}{\lambda z.\lambda w.(lose'\ (z\ (w))) : (n|n)\backslash(s|n)}\; \mathrm{G}_0 <}{\lambda w.(lose'\ w) : s|n}\; <}{\lambda y.((say'\ (lose'\ y))\ j') : s|n}\; >$$

Figure 7.9 – Derivation for: *John*$_1$ *said he*$_2$ *lost* in $\mathbf{J}_{GZ}$

It is in order to obtain a bound pronominal reading that the **Z** rule enters this combinatory scene. In a nutshell, this is because the semantic operation triggered by the type-shift **Z** rule yields a multiple-binding λ-term. Let us break this down into two steps. First, observe that **Z** performs a twofold shift: one on the syntactic type of the two-place function it applies to, and one on its semantics. Thus, while the type $(n\backslash A)/B$ denotes the function M with — at least — two arguments, $(n\backslash A)/(B|n)$ denotes the function $\mathbf{Z}(M)$, whose first argument is itself a function. In addition, note that besides lifting the syntactic-semantic type of the first argument, **Z** also identifies two argument positions: namely, the second argument position of the function that **Z** applies to, and the new argument position. Thus, by applying **Z** to the function $\lambda x.\lambda y.(M(x(y))$, we obtain the λ-term $\lambda x.\lambda y.(M(x(y))(y))$, in which the variable y occurs twice. Since y corresponds to the

$$\cfrac{\cfrac{\cfrac{\cfrac{\mathit{She}}{n|n}}{(s|n)/((n|n)\backslash(s|n))}\,\mathrm{T}>}{(s|n)|n/((n|n)\backslash(s|n))|n}\,\mathrm{G}_0> \qquad \cfrac{\cfrac{\cfrac{\cfrac{\mathit{admires}}{(n\backslash s)/n}}{((n\backslash s)|n)/(n|n)}\,\mathrm{G}_0> \quad \cfrac{\mathit{him}}{n|n}}{(n\backslash s)|n}\,>}{((n|n)\backslash(s|n))|n}\,\mathrm{G}_1<}{(s|n)|n}\,>$$

Figure 7.10 – Derivation for: *She admires him* in $\mathbf{J}_{GZ}$

syntactic type n, y is a variable of semantic type e. Put differently, in $\lambda x.\lambda y.(M(x(y))(y))$ not just one but two nominal argument positions happen to be bound.

To cast this explanation in more linguistic terms, consider a two-place predicate, for instance a transitive or a propositional verb: say, $(n\backslash s)/n$ and $(n\backslash s)/s$, respectively. Then, **Z** simultaneously performs two semantic-syntactic operations: first, **Z** builds a pronominal type on its second—lower—nominal argument: $(n\backslash s)/(n|n)$ and $(n\backslash s)/(s|n)$, respectively. Second, **Z** identifies the pro-form and its nominal antecedent. As a consequence, this second — higher — nominal will be the one binding the pronominal expression in the lower position. Thus, **Z** ensures the prominence condition of the binder. Needless to say, **Z** cannot meet the anti-locality condition (or: Principle B of GB) for pronouns. As in the lexical approaches seen earlier on, **Z** does not duplicate the meaning of the actual binder, but rather its nominal gap. Since the second argument of the predicate which the **Z** rule applies to must be of type n, the rule ensures that the binder — whatever it happens to be — will be a nominal expression. This notwithstanding, Jacobson herself recognizes that other expressions, in addition to nominals, could also work just as well as binders.

Since our research is only concerned with nominal binders, the rules of $\mathbf{J}_{GZ}$ already seem to suffice for our purposes and, as a result, also to evaluate the scope of Jacobson's proposal in the specific case of overt pronouns in control sentences.

To wrap up our exposition, let us run through a couple of simple derivations (see Figs. 7.11 and 7.12 on the following page). The first of these illustrates the way in which the two new combinators work in tandem. Observe that while **G** is responsible for allowing combinations with the pronominal expression, the **Z** rule is responsible for the duplicating operation needed to finally obtain a bound reading for the pronoun. Consider the following lexical entry for the possessive pronoun (cf. [MV14]:402):

his : $(n|n)/cn : \lambda x.\lambda y.((of'\ y)\ x)$

In closing, it is important to draw attention to one of the major innovations behind Jacobson's proposal. For, while her system is combinatory in its design, it is in fact Type-Logical in its inception. Indeed, instead of augmenting only the set of combinatory rules

$$\dfrac{\dfrac{\textit{John}}{j' : n} \qquad \dfrac{\dfrac{\dfrac{\textit{said}}{say' : (n\backslash s)/s}}{\lambda x.\lambda u.(say'\ (x\ u)\ u) : (n\backslash s)/(s|n)}\,\mathrm{Z} \qquad \dfrac{\dfrac{\textit{he}}{\lambda y.y : n|n} \qquad \dfrac{\dfrac{\textit{lost}}{lose' : n\backslash s}}{\lambda z.\lambda w.(lose'\ (z\ (w))) : (n|n)\backslash(s|n)}\,\mathrm{G}_0 <}{\lambda w.(lose'\ w) : s|n}\,<}{\lambda u.((say'\ (lose'\ u))\ u) : n\backslash s}\,>}{((say'\ (lose'\ j'))\ j') : s}\,<$$

Figure 7.11 – Derivation for: *John*$_1$ *said he*$_1$ *lost* in $\mathbf{J}_{GZ}$

$$\dfrac{\dfrac{\textit{John}}{j' : n} \qquad \dfrac{\dfrac{\dfrac{\textit{admires}}{admire' : (n\backslash s)/n}}{\lambda u.\lambda z.(admire'\ (u\ (z))(z)) : (n\backslash s)/(n|n)}\,\mathrm{Z} \qquad \dfrac{\dfrac{\textit{his}}{\lambda x.\lambda y.((of'\ y)\ x) : (n|n)/cn} \qquad \dfrac{\textit{mother}}{mother' : cn}}{\lambda y.((of'\ y)\ mother') : n|n}\,<}{\lambda z.((admire'\ ((of'\ z)\ mother'))\ z) : n\backslash s}\,<}{((admire'\ ((of'\ j')\ mother'))\ j') : s}\,<$$

Figure 7.12 – Derivation for: *John*$_1$ *admires his*$_1$ *mother* in $\mathbf{J}_{GZ}$

of **AB** — as is typically done in a combinatory grammar — $\mathbf{J}_{GZ}$ also extends the set of syntactic categories, just like a Type-Logical approach. Nevertheless, whereas these new rules are specifically devised for manipulating the new syntactic type, they also feature slashes other than $|$. In sum, Jacobson effectively builds a bridge between the Combinatory and the Type-Logical approaches. Indeed, as we shall see in Section 7.3.2, at the heart of the Type-Logical system **LLC** [Jäg05] are none other than the combinators **G** and **Z**.

7.3 On Contraction: Duplication in the Syntax

7.3.1 LA Calculus: Multi-structured Sequents and Modal Contraction

Jaeger [Jäg97, Jäg98] formulates the Lambek system with Anaphora **LA** to give a uniform treatment of two linguistic phenomena which require reusing semantic resources: ellipsis of VP and anaphoric pronouns (see examples (209) and (210), respectively).

(209) John washed his car, and Bill did, too.

(210) John washed his car, and Bill waxed it.

The **LA** system is a conservative extension of **L** obtained by adding three new type-constructors: $\sim, \hookrightarrow$ and $\hookleftarrow$ to the language for the latter. Formally, in Backus-Naur form:

Definition 7.3 (syntactic types of **LA**). Where P is a set of basic types, *the set F of types of **LA*** is defined as follows:

$$F \ ::= \ P \mid F\backslash F \mid F/F \mid F \bullet F \mid F \hookrightarrow F \mid F \hookleftarrow F \mid F \sim F$$

Like the sequent calculus **L**, in **LA** each of the anaphoric connectives $\hookleftarrow, \hookrightarrow$, and $\sim$ are characterized by a right and a left rule. These rules are given in figure 7.13.[139] Despite the syntactic and semantic similarities between the Lambek connectives $\backslash, /, \bullet$ and Jaeger's $\hookrightarrow, \hookleftarrow, \sim$, the left rules for the **LA** slashes create a new structure $\{\ \}$ on the left-hand side of an **LA**-sequent.

$$\frac{X \Rightarrow M : A \qquad Y, x : B, Z \Rightarrow N : C}{Y, \{X, y : A \hookrightarrow B\}, Z \Rightarrow N[yM/x] : C} \hookrightarrow \mathrm{L} \qquad\qquad \frac{\{x : A, X\} \Rightarrow M : B}{X \Rightarrow \lambda x.M : A \hookrightarrow B} \hookrightarrow \mathrm{R}$$

$$\frac{X \Rightarrow M : A \qquad Y, x : B, Z \Rightarrow N : C}{Y, \{y : B \hookleftarrow A, X\}, Z \Rightarrow N[yM/x] : C} \hookleftarrow \mathrm{L} \qquad\qquad \frac{\{X, x : A\} \Rightarrow M : B}{X \Rightarrow \lambda x.M : B \hookleftarrow A} \hookleftarrow \mathrm{R}$$

$$\frac{X, \{x : A, y : B\}, Y \Rightarrow N : C}{X, z : A \sim B, Y \Rightarrow N[(z)_0/x][(z)_1/y] : C} \sim \mathrm{L} \qquad\qquad \frac{X \Rightarrow M : A \qquad Y \Rightarrow N : B}{\{X, Y\} \Rightarrow \langle M, N\rangle : A \sim B} \sim \mathrm{R}$$

Figure 7.13 – Logical rules of **LA**

Thus, unlike **L**, **LA** is a multi-modal system: it encodes two modes for combining (or concatenating) syntactic types. Indeed, since **LA** is a conservative extension of **L**, it already contains the associative form of combination (,), which corresponds to the (associative) product connective $\bullet$. In addition, **LA** contains a new form of combination: $\{\ ,\ \}$, corresponding to the other product connective $\sim$. For this new structure $\{\ ,\ \}$ only, the structural rules of Contraction and Permutation are explicitly admitted (see Fig. 7.14).[140]

As a result, the left-hand side of a sequent $X \Rightarrow C$ of **LA** may contain both an (associative) linear structure (,) of syntactic types as well as a non-linear (and associative) structure $\{\ ,\ \}$. In other words, the structure (A, B) stands for a sequence of types A, B, whereas the structure $\{A, B\}$ stands for a set of these types.

[139] As the reader can check, these connectives, like those of **L**, form a residuated triple. Indeed, the following residuation relation holds in **LA**:

$$A \sim B \Rightarrow C \text{ is derivable iff } B \Rightarrow A \hookrightarrow C \text{ is derivable iff } A \Rightarrow C \hookleftarrow B \text{ is derivable.}$$

[140] Despite the presence of Contraction, the **LA** system enjoys Cut Elimination and is decidable (cf. [Jäg97]). The **LA** system exposed in [Jäg97] contains, in addition to those, the following structural rule, called Index Percolation (IP):

$$\frac{X_1, (\{Y, Z\}, W), X_2 \Rightarrow B}{X_1, \{Y, (Z, W)\}, X_2 \Rightarrow B} \text{IP} \Rightarrow$$

$$\frac{X, \{x : A, y : A\}, Y \Rightarrow M : B}{X, y : A, Y \Rightarrow M[y/x] : B}\ \mathrm{C} \Rightarrow \qquad \frac{X_1, \{Y, \{Z, W\}\}, X_2 \Rightarrow B}{X_1, \{Z, \{Y, W\}\}, X_2 \Rightarrow B}\ \mathrm{P} \Rightarrow$$

$$\frac{X_1, (Y, \{Z, W\}), X_2 \Rightarrow B}{X_1, (\{Z, Y\}, W), X_2 \Rightarrow B}\ \mathrm{IM} \Rightarrow$$

Figure 7.14 – Structural rules of **LA**

In the system for anaphora **LA**, the syntactic type $n \hookrightarrow n$ is assigned to personal pronouns (pronominals and reflexives), with the identity function $\lambda x.x$ as their semantic value. This syntactic and semantic assignation rests on the intuition that a personal pronoun behaves as a proper noun (of type n), in the presence of an antecedent of type n. In turn, the type $n \hookrightarrow (n/cn)$ is assigned to a possessive pronoun like *his*, with the function $\lambda x \lambda y.((of'\ y)\ x)$ as its semantic value. This syntactic type expresses the fact that a possessive works as a definite determiner in the presence of a nominal antecedent.

To illustrate how the **LA** system works, consider the following sentences, lexicon and proofs:

(211) John admires himself.

(212) John_1 says $\text{he}_{1/2}$ runs.

(213) John_1 washes $\text{his}_{1/2}$ car.

admires : $(n\backslash s)/n : admires'$

he : $n \hookrightarrow n : \lambda y.y$

him : $n \hookrightarrow n : \lambda y.y$

himself : $n \hookrightarrow n : \lambda y.y$

his : $n \hookrightarrow (n/cn) : \lambda x.\lambda y.((of'\ y)\ x)$

John : $n : j'$

man : $cn : man'$

Peter : $n : p'$

runs : $n\backslash s : run'$

says : $(n\backslash s)/s : say'$

the : $n/cn : \iota$

washes : $(n\backslash s)/n : wash'$

As a result of an **LA**-proof (Fig. 7.15 on the next page) for the sentence in (211), we obtain the λ-term $((v(zu))x)[x/u]$, which is equivalent to $((v(zx))x)$ once we effect variable substitution. After inserting the lexical meanings in $((v(zx))x)$, where the variable x corresponds to the lexical entry for **John**, **admires** replaces v, and **himself** replaces z, we then perform β-reduction (i.e., apply the relevant functions to their arguments). Thus, we obtain the reading $((admire'\ j')\ j')$. In symbols:

$$((v(zx))x) = ((admire' \ ((\lambda y.y) \ j')) \ j') \rightsquigarrow_\beta ((admire' \ j') \ j')$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{u : n \Rightarrow u : n \qquad \dfrac{w : n \Rightarrow w : n \qquad \dfrac{x : n \Rightarrow x : n \qquad z : s \Rightarrow z : s}{x : n, y : n\backslash s \Rightarrow z[(yx)/z] : s}\,\backslash L}{x : n, v : (n\backslash s)/n, w : n \Rightarrow (yx)[(vw)/y] : s}\,/L}{x : n, v : (n\backslash s)/n, \{u : n, z : n \hookrightarrow n\} \Rightarrow ((vw)x)[(zu)/w] : s}\,\hookrightarrow L}{x : n, \{u : n, v : (n\backslash s)/n\}, z : n \hookrightarrow n \Rightarrow ((v(zu))x) : s}\,\mathrm{IM} \Rightarrow}{\{u : n, x : n\}, v : (n\backslash s)/n, z : n \hookrightarrow n \Rightarrow ((v(zu))x) : s}\,\mathrm{IM} \Rightarrow}{x : n, v : (n\backslash s)/n, z : n \hookrightarrow n \Rightarrow ((v(zu))x)[x/u] : s}\,\mathrm{C} \Rightarrow$$

Figure 7.15 – Derivation for: *John*$_1$ *admires himself*$_1$ in **LA**

$$\dfrac{\dfrac{\begin{array}{c}\dfrac{\dfrac{z : n \Rightarrow z : n \qquad \dfrac{\begin{array}{c}\vdots\\ x : n, w : n\backslash s \Rightarrow wx : s\end{array} \qquad \begin{array}{c}\vdots\\ y : n, u : n\backslash s \Rightarrow uy : s\end{array}}{y : n, v : (n\backslash s)/s, x : n, w : n\backslash s \Rightarrow (uy)[(v(wx))/u] : s}\,/L}{y : n, v : (n\backslash s)/s, \{z : n, t : n \hookrightarrow n\}, w : n\backslash s \Rightarrow ((v(wx))y)[(tz)/x] : s}\,\hookrightarrow L}{\vdots}\,\mathrm{IM} \Rightarrow\end{array}}{\{z : n, y : n\}, v : (n\backslash s)/s, t : n \hookrightarrow n, w : n\backslash s \Rightarrow ((v(w(tz)))y) : s}}{y : n, v : (n\backslash s)/s, t : n \hookrightarrow n, w : n\backslash s \Rightarrow ((v(w(tz)))y)[y/z] : s}\,\mathrm{C} \Rightarrow$$

Figure 7.16 – Derivation for: *John*$_1$ *says he*$_1$ *runs* in **LA**

$$\dfrac{\begin{array}{c}\vdots\\ y : n, v : (n\backslash s)/s, \{z : n, t : n \hookrightarrow n\}, w : n\backslash s \Rightarrow ((v(wx))y)[(tz)/x] : s\end{array}}{y : n, v : (n\backslash s)/s, t : n \hookrightarrow n, w : n\backslash s \Rightarrow \lambda z.((v(w(tz)))y) : n \hookrightarrow s}\,\hookrightarrow R$$

Figure 7.17 – Derivation for: *John*$_1$ *says he*$_2$ *runs* in **LA**

For the bound reading of the sentence in (212), we insert the lexical meanings in the λ-term $((v(w(ty)))y)$, which substitutes for $((v(w(tz)))y)[y/z]$ (see Fig. 7.16). After applying β-reduction, we obtain the reading $((said' \ (run' \ j')) \ j')$. For the free reading—$\lambda.z((said' \ (run' \ z)) \ j')$—see Fig. 7.17.

Finally, we insert the lexical meanings in the term $\lambda y.((v((wy)u))x)$ corresponding to an **LA**-proof (Fig. 7.18 on the following page) for the sentence in (213).

$$\lambda y.((wash' \ ((\lambda w.((of' \ x) \ w) \ y) \ car')) \ j') \rightsquigarrow_\beta \lambda y.((wash' \ ((of' \ car') \ y)) \ j')$$

Observe, first, that the derivation of a sentence containing pronominal expressions uses the left rule for the anaphoric type-constructor, $\hookrightarrow L$. It is then this left rule which creates the syntactic domain $\{\ ,\ \}$, where types are allowed to be permuted and also contracted. Thus, to derive a reflexive or the bound-reading of a personal pronoun, the

$$\dfrac{\dfrac{y : n \Rightarrow y : n \qquad \begin{array}{c}\vdots\\ x : n, v : (n\backslash s)/n, z : n/cn, u : cn \Rightarrow ((v(zu))x) : s\end{array}}{x : n, v : (n\backslash s)/n, \{y : n, w : n \hookrightarrow (n/cn)\}, u : cn \Rightarrow ((v(zu))x)[(wy)/z] : s}\hookrightarrow L}{x : n, v : (n\backslash s)/n, w : n \hookrightarrow (n/cn), u : cn \Rightarrow \lambda y.((v((wy)u))x) : n \hookrightarrow s}\hookrightarrow R$$

Figure 7.18 – Derivation for: *John*$_1$ *washes his*$_2$ *car* in **LA**

idea is to make a copy of the nominal antecedent; this copy is then moved (i.e. permuted) in order to be finally deleted (i.e. contracted). Also note that the Contraction rule used to derive anaphoric expressions secures a condition on the order of precedence for the antecedent of the anaphoric pronoun: the antecedent precedes the anaphora. This is because $IM \Rightarrow$ moves the non-linear domain, where Contraction can be applied, only to the left. Therefore, $IM \Rightarrow$ ensures that, when Contraction is applied, the antecedent of the anaphoric pronoun stands to the left of the latter.

In contrast, the derivation of a sentence containing (a single) free pronoun does not use any form of Permutation (i.e. $IM \Rightarrow$ or $P \Rightarrow$) nor Contraction. In addition, note that the application of Contraction does not only erase an instance of the repeated syntactic type n, but also removes the non-linear structure $\{\ ,\ \}$. Alternatively, this structure may be removed by applying $\hookrightarrow R$. In this case, the free reading of the pronoun is obtained; the functional type $n \hookrightarrow s$ corresponds to a sentence with a pronominal gap. Given that Contraction and also $\hookrightarrow R$ delete the non-linear structure which is required for both rules to be applied, they appear to be mutually exclusive. There are, however, some contexts where these rules can both be used. As an example, consider the **LA**-proof (Fig. 7.19 on the next page) for the sentence in (214) below, with the free reading $\lambda x.((drink'\ x) \wedge (smoke'\ x))$.[141]

(214) He$_1$ drinks and he$_1$ smokes.

As we saw in Section 5.3, as a consequence of applying Contraction, the λ-term corresponding to the proof contains more than one occurrence of the same variable. By contrast, in the λ-term of a proof in which Contraction has not been applied, each variable occurs at least once. This is because, without Contraction, each (occurrence of each) assumption is used at most once in deriving a conclusion.

While **LA** enjoys some desired theoretical properties and may offer a unified account for different linguistic phenomena which require reusing semantic resources, it runs into the problem of overgeneration. This becomes evident when we pay closer attention to the treatment of personal pronouns. First, note that Principles A and B of Binding Theory (see Section 2.1; a brief recap is also given below) are not encoded into the Type-Logical

[141]For the definition of conjunction as the generalized meet operation, see [Jäg05]; [Jac96b].

$$
\dfrac{\begin{matrix}\dfrac{\dfrac{n \Rightarrow n \quad \dfrac{n \Rightarrow n \quad \dfrac{\overset{\vdots}{n, n\backslash s \Rightarrow s} \quad \dfrac{\overset{\vdots}{n, n\backslash s \Rightarrow s} \quad s \Rightarrow s}{n, n\backslash s, s\backslash s \Rightarrow s}\,\backslash\mathrm{L}}{n, n\backslash s, (s\backslash s)/s, n, n\backslash s \Rightarrow s}\,/\mathrm{L}}{n, n\backslash s, (s\backslash s)/s, \{n, n \hookrightarrow n\}, n\backslash s \Rightarrow s}\,{\hookrightarrow}\mathrm{L}}{\{n, n \hookrightarrow n\}, n\backslash s, (s\backslash s)/s, \{n, n \hookrightarrow n\}, n\backslash s \Rightarrow s}\,{\hookrightarrow}\mathrm{L}}{\vdots}\,\mathrm{IM} \Rightarrow \\ \dfrac{\dfrac{\{n, \{n, n \hookrightarrow n\}\}, n\backslash s, (s\backslash s)/s, n \hookrightarrow n, n\backslash s \Rightarrow s}{\{\{n, n\}, n \hookrightarrow n\}, n\backslash s, (s\backslash s)/s, n \hookrightarrow n, n\backslash s \Rightarrow s}\,\mathrm{IM} \Rightarrow}{\{n, n \hookrightarrow n\}, n\backslash s, (s\backslash s)/s, n \hookrightarrow n, n\backslash s \Rightarrow s}\,\mathrm{C} \Rightarrow \end{matrix}}{n \hookrightarrow n, n\backslash s, (s\backslash s)/s, n \hookrightarrow n, n\backslash s \Rightarrow n \hookrightarrow s}\,{\hookrightarrow}\mathrm{R}
$$

Figure 7.19 – Derivation for: *He*$_1$ *drinks and he*$_1$ *smokes* in **LA**

$$
\begin{matrix}\dfrac{\dfrac{\dfrac{u : n \Rightarrow u : n \quad \overset{\vdots}{x : n, y : (n\backslash s)/s, z : n, w : (n\backslash s)/n, v : n \Rightarrow ((y((wv)z))x) : s}}{x : n, y : (n\backslash s)/s, z : n, w : (n\backslash s)/n, \{u : n, t : n \hookrightarrow n\} \Rightarrow ((y((wv)z))x)[(tu)/v] : s}\,{\hookrightarrow}\mathrm{L}}{x : n, y : (n\backslash s)/s, z : n, \{u : n, w : (n\backslash s)/n\}, t : n \hookrightarrow n \Rightarrow ((y((w(tu))z))x) : s}\,\mathrm{IM} \Rightarrow}{\vdots}\,\mathrm{IM} \Rightarrow \\ \dfrac{\{u : n, x : n\}, y : (n\backslash s)/s, z : n, w : (n\backslash s)/n, t : n \hookrightarrow n \Rightarrow ((y((w(tu))z))x) : s}{x : n, y : (n\backslash s)/s, z : n, w : (n\backslash s)/n, t : n \hookrightarrow n \Rightarrow ((y((w(tu))z))x)[x/u] : s}\,\mathrm{C} \Rightarrow \end{matrix}
$$

Figure 7.20 – Derivation for: **John*$_1$ *says Peter*$_2$ *admires himself*$_1$ in **LA**

Grammar **LA**. Consequently, **LA** recognizes not only those sentences which satisfy Principles A and B, but also several others that do not obey them. As an example, consider the **LA**-proofs for the first two ungrammatical sentences below:

(215) John$_1$ says Peter$_2$ admires himself$_{*1/2}$.

(216) The man$_1$ says Peter$_2$ admires him$_{1/*2}$.

(217) John$_1$ expects *he/him$_2$/himself$_1$ to be admired.

In fact, after inserting the lexical meanings in the term $((y((w(tx))z))x)$ corresponding to a proof for the sentence in (215), we obtain the reading $((\mathit{say}'\ ((\mathit{admire}'\ j')\ p'))\ j')$, which does not meet Principle A for reflexives (Fig. 7.20).

Once again, after inserting the lexical meanings in $((y((w(rz))z))(xt))$, which corresponds to a proof (Fig. 7.21 on the following page) for the sentence in (216), we obtain the reading $((\mathit{say}'\ ((\mathit{admire}'\ p')\ p'))(\iota\ \mathit{man}'))$. In this case, the reading does not meet Principle B for pronouns.

$$
\begin{array}{cl}
 & \vdots \\
u : n \Rightarrow u : n \qquad x : n/cn, t : cn, y : (n\backslash s)/s, z : n, w : (n\backslash s)/n, v : n \Rightarrow ((y((wv)z))(xt)) : s & \\
\hline
x : n/cn, t : cn, y : (n\backslash s)/s, z : n, w : (n\backslash s)/n, \{u : n, r : n \hookrightarrow n\} \Rightarrow ((y((wv)z))(xt))[(ru)/v] : s & \hookrightarrow\mathrm{L} \\
\hline
x : n/cn, t : cn, y : (n\backslash s)/s, z : n, \{u : n, w : (n\backslash s)/n\}, r : n \hookrightarrow n \Rightarrow ((y((w(ru))z))(xt)) : s & \mathrm{IM} \Rightarrow \\
\hline
x : n/cn, t : cn, y : (n\backslash s)/s, \{u : n, z : n\}, w : (n\backslash s)/n, r : n \hookrightarrow n \Rightarrow ((y((w(ru))z))(xt)) : s & \mathrm{IM} \Rightarrow \\
\hline
x : n/cn, t : cn, y : (n\backslash s)/s, z : n, w : (n\backslash s)/n, r : n \hookrightarrow n \Rightarrow ((y((w(ru))z))(xt))[z/u] : s & \mathrm{C} \Rightarrow
\end{array}
$$

Figure 7.21 – Derivation for: **The man$_1$ says Peter$_2$ admires him$_2$* in **LA**

As we know, from a generative perspective the licensing of pronominal expressions such as *he*, *him* and *himself* is determined by Principles A and B of the Binding Theory [Cho81]. Principle A stipulates that an anaphor (reflexives and reciprocals) must be bound in its governing category (roughly, it must have a c-commanding local antecedent). Principle B stipulates that a pronoun must be free (i.e. not bound) within its governing category; nonetheless, a pronoun can be bound from outside this syntactic domain. Thus a pronoun, unlike an anaphor, also admits a free reading. Principles A and B jointly imply a strict complementary distribution between pronouns and reflexives in some syntactic domains: in certain domains an anaphora can occur, but a bound pronoun cannot (and vice-versa).

In Jaeger's proposal, however, *he, him* and *himself* are treated as if they were identical. In **LA** personal pronouns and reflexives are uniformly assigned the syntactic type $n \hookrightarrow n$ with the semantic $\lambda x.x$. Consequently, the rules for the anaphoric type-constructor $\hookrightarrow$ can be applied without discrimination to pronouns as well as reflexives. But this turns out to be problematic. Consider, first, the rule of proof $\hookrightarrow R$. It creates a constituent with a pronominal gap, which means that the constituent has an open slot for a pronoun. In semantic terms, such a constituent denotes an unsaturated function (from the set of entities to another set). In **LA**, the rule of use—$\hookrightarrow L$—for a pronoun goes hand in hand with its rule of proof. In effect, since $\hookrightarrow L$ creates the required $\{\ ,\ \}$ domain for $\hookrightarrow R$ to be applied, the dependence between the two rules seems adequate: the presence of a pronoun $n \hookrightarrow n$ in a constituent X is reflected in the consequent functional type $n \hookrightarrow X$ assigned to the latter. In more linguistic terms, and given the uniform approach of **LA**, this however means that *both* pronouns and reflexives may receive a free reading in the Type-Logical Grammar **LA**. The problem is that a reflexive, unlike a pronoun, must necessarily be bound. This inadequacy may be easily overcome by splitting the anaphoric connective $\hookrightarrow$ into two anaphoric type-constructors: a reflexive one, which would not be defined by the right rule, and a pronominal one, defined by the left and also by the right rule of **LA**.

Consider now the $C \Rightarrow$ rule. Once again, this rule (as well as $IM \Rightarrow$) may be applied indiscriminately in the logical analysis of sentences containing a reflexive or a pronoun. Note, however, that once $C \Rightarrow$ has been applied, the non-linear domain is deleted. As a consequence, $\hookrightarrow R$ cannot be applied afterwards. Thus, as a result of $C \Rightarrow$, a constituent is left without a pronominal gap, even though the constituent might contain a pronominal expression. In semantic terms, as we have said before, a proof in which $C \Rightarrow$ has been applied corresponds to a non-linear λ-term, that is: a term containing repeated occurrences of the same variable. Hence, these syntactic and semantic operations seem correct for reflexives and bound pronouns, but certainly not for free pronouns.

Although at first sight it might seem a good idea to restrict the application of $C \Rightarrow$ to reflexives and bound pronouns, here the situation seems more complicated than in the case of $\hookrightarrow R$. First, note that free and bound pronouns on the one hand, unlike reflexives and pronouns on the other, are morphologically identical. Second, note that according to Principles A and B, pronouns and reflexives must satisfy some locality (and also hierarchical) conditions in order to be bound.

Finally, consider the interaction between the structural rules $IM \Rightarrow$ and $C \Rightarrow$. The latter requires a certain locality condition to be applied: in order for two occurrences of the same syntactic type to be contracted, they have to be adjacent. However, this locality condition may be easily satisfied. Indeed, $IM \Rightarrow$, which moves the copy of the antecedent of the reflexive or of the pronoun leftwards in the structure, can do so an unlimited number of times. In other terms, $IM \Rightarrow$ may be applied without any restriction within the non-linear structure $\{,\}$. Consequently, neither rule really imposes any kind of locality condition or hierarchical ordering on the antecedent for the bound expression. Thus, there is no room to encode a hierarchical ordering (such as c-command) nor to stipulate syntactic (local) domains for binding in a grammar based on the **LA** calculus.

7.3.2 **LLC** Calculus: Encoding Limited Contraction

The Lambek calculus with Limited Contraction (**LLC**) proposed by Jaeger [Jäg05] is designed to treat different linguistic phenomena related to anaphora. Specifically, the **LLC** calculus intends to cover cases of personal and possessive pronouns bound by *wh*-operators and quantifiers, and cases of bound pronouns in ellipsis of VP, as exemplified below:

(218) a. Every Englishman$_1$ loves his$_1$ mother.

b. *His$_1$ mother loves every Englishman$_1$.

(219) The man who asked for John$_1$ met him$_1$.

(220) John revised his paper, and Bill did (too).
= John$_1$ revised his$_1$ paper, and Bill$_2$ revised his$_1$ paper. (strict reading)
= John$_1$ revised his$_1$ paper, and Bill$_2$ revised his$_2$ paper. (sloppy reading)

= John_1 revised his_2 paper, and Bill_3 revised his_2 paper. (free+sloppy reading)

The phenomenon of anaphora manifests in contexts where a lexical item takes its reference from another linguistic expression. Therefore, in order to handle anaphors from a syntactic type-logical perspective, **L** must be expanded with a mechanism for multiplying semantic resources. In **LLC** this task is performed by the logical rules of the new type-constructor $|$, which is added to the vocabulary of **L**.

Definition 7.4 (Syntactic types of **LLC**)**.** Where P is a set of basic types, the *set F of syntactic types of **LLC*** is defined as follows:

$$F ::= \; P \mid F\backslash F \mid F/F \mid F \bullet F \mid F|F$$

Definition 7.5 (Semantic types)**.** The *set γ of semantic types* is defined on the basis of a set δ of primitive semantic types by:

$$\gamma ::= \; \delta \mid \gamma \,\&\, \gamma \mid \gamma \rightarrow \gamma$$

As in **L**, the product type-constructor is semantically interpreted as Cartesian product, and implication, as function space formation. So, the category-to-type correspondence for **LLC** is given as follows:

Definition 7.6 (Semantic type map for **LLC**)**.** The *semantic type map for **LLC*** is a mapping τ from syntactic types F to semantic types γ such that:

1. $\tau\,(A \bullet B) = \tau\,(A) \;\&\, \tau\,(B)$
2. $\tau\,(A\backslash B) = \tau\,(B/A) = \tau\,(B|A) = \tau\,(A) \rightarrow \tau\,(B)$

Given that **LLC** is a conservative extension of the Lambek **L** Calculus, the sequent rules for the product $\bullet$ and the slash connectives $\backslash$ and $/$ are as in **L**. The left and right rules for the anaphoric slash $|$ are as in Fig. 7.22 below.

In **LLC** anaphoric expressions are assigned a type $B|A$: it works as a type B in the presence of an antecedent of type A. Like in **LA**, in **LLC** the mechanism for reusing resources operates in the syntax. However, unlike **LA** and like **L**, **LLC** is free of structural rules. Thus, in **LLC** multiple binding is a semantic consequence of logical as opposed to structural rules. Specifically, it is the left and right rules for the anaphoric connective $|$ which encode a non-local version of Contraction. We shall motivate this claim in two steps, focusing first on $|L$ and then on $|R$.

The $|L$ rule expresses the fact that for an anaphoric expression $B|A$ to be bound it needs an *antecedent* A in the same premise. Thus, this rule imposes a precedent-linear

$$\frac{Y \Rightarrow M : A \qquad X, x : A, Z, y : B, W \Rightarrow N : C}{X, Y, Z, z : B|A, W \Rightarrow N[M/x][(zM)/y] : C}\ |\mathrm{L}$$

$$\frac{X, x_1 : B_1, Y_1, \dots, x_n : B_n, Y_n \Rightarrow N : C}{X, y_1 : B_1|A, Y_1, \dots, y_n : B_n|A, Y_n \Rightarrow \lambda z.N[(y_1 z)/x_1] \dots [(y_n z)/x_n] : C|A}\ |\mathrm{R}$$

where $n \geq 1$

Figure 7.22 – Logical rules of **LLC**

condition on the antecedent for the anaphor. In addition, this logical rule subsumes two structural rules: Contraction and also Permutation.[142] Indeed, the antecedent A for the anaphoric type $B|A$ occurs *twice*, both times in *some aleatory position* in the two premises of $|L$. To see this even more clearly, compare the following sequents. As we shall see in the next paragraph, underlying the first five sequents is the Contraction rule, whereas the last three employ some form of Permutation (as well as Contraction):

Proposition 7.1.

$\vdash_{\boldsymbol{LLC}} A, B|A \Rightarrow A \bullet B$

$\vdash_{\boldsymbol{LA}} A, A \hookrightarrow B \Rightarrow A \bullet B$

$\nvdash_{\boldsymbol{L}} A, A\backslash B \Rightarrow A \bullet B$

$\vdash_{\boldsymbol{L}} A, A, A\backslash B \Rightarrow A \bullet B$

$\vdash_{\boldsymbol{LA}} \{A, A\}, A \hookrightarrow B \Rightarrow A \bullet B$

$\vdash_{\boldsymbol{LLC}} A, C, B|A \Rightarrow C \bullet B \bullet A$

$\nvdash_{\boldsymbol{L}} A, C, A\backslash B \Rightarrow C \bullet B$

$\vdash_{\boldsymbol{LA}} \{A, A\}, C, A \hookrightarrow B \Rightarrow C \bullet B \bullet A$

The first two sequents exhibit the fact that the functional anaphoric types $B|A$ and $A \hookrightarrow B$ (of **LLC** and **LA** respectively) use, but do not consume, their antecedent A. In **L**, however, this is not the case; this is witnessed by the third sequent. In order to obtain the product type $A \bullet B$ from the functional type $A\backslash B$ in **L**, repeated occurrences of its antecedent A are needed. Both in **LLC** and in **LA**, unlike in **L**, these duplicated

[142]Despite incorporating this structural rule, **LLC**, as well as Lambek system, enjoys Cut elimination, decidability and the subformula property. For the latter, note that all the formulas that occur in the premises of the two new rules for the anaphoric type-constructor are subformulas of the formulas that occur in their conclusion.

$$\dfrac{\dfrac{\dfrac{A \Rightarrow A \qquad X, A, Z, B, W \Rightarrow C}{X, A, Z, \{A, A \hookrightarrow B\}, W \Rightarrow C} \hookrightarrow\text{L}}{X, \{A, A\}, Z, A \hookrightarrow B, W \Rightarrow C} \text{IM} \Rightarrow}{X, A, Z, A \hookrightarrow B, W \Rightarrow C} \text{C} \Rightarrow \qquad \dfrac{A \Rightarrow A \qquad X, A, Z, B, W \Rightarrow C}{X, A, Z, B|A, W \Rightarrow C} |\text{L}$$

Figure 7.23 – Derivation of $|L$, where $Y = A$, in **LA** and in **LLC**

occurrences can be deleted. While in **LA** this task is accomplished by applying the Contraction rule into the non-linear domain, in **LLC** all that is required is the logical rule $|L$. The last three sequents show that anaphoric types of **LLC** and **LA**, unlike $A\backslash B$ in **L**, can take a long-distance antecedent or, in other terms, that there can be lexical material between the antecedent and the functional anaphoric type, which consumes the former. There is however a clear difference between Jaeger's two systems. Once again, while in **LA** the structural rule $IM \Rightarrow$ comes into play before $C \Rightarrow$, no structural rule is needed in **LLC** (see Fig. 7.23). In **LLC**, the left rule for the anaphoric type-constructor | implicitly performs both tasks at once: it exchanges the order of the antecedent type and deletes one of its occurrences. In other words, $|L$ deletes non-adjacent occurrences of a syntactic type.

In sum, the implicit presence of long-distance Contraction in $|L$ allows us on the one hand to reuse a lexical item, and on the other to reuse it in spite of the presence of lexical material between the antecedent for the anaphor and the functional anaphoric type.

Analogously, $|R$ also encodes these structural rules: Contraction and Permutation. This rule is non-standard, for three reasons. First, it simultaneously constructs an anaphoric type on both sides of the sequent. In addition, it allows us to construct several anaphoric types $B_n|A$, where $n > 0$, with the same antecedent A. Finally, the antecedent for the anaphoric types does not necessarily occur in the premise of the rule. Because of these special features, the implicit presence of these structural rules is as obvious here. Still, the presence of Permutation can easily be confirmed by examining the case of $n = 1$. Note, first, that in this particular case $|R$ can be split into a left and a right rule ($|L_1$ and $|R_1$ in Fig. 7.24), in a similar way to Jaeger's system **LA**; hence, $|R$ is to some extent comparable to \R. However, unlike the latter $|R$ does not require the type in the premise to stand in the left-periphery of the antecedent sequent, since it does not require the antecedent of either anaphoric type to occur in the premise at all.[143] Additionally, where $n > 1$ the rule reveals an underlying form of Contraction along with Permutation. Indeed, as the reader can see below in Fig. 7.26, the intuitionistic proof of $|R$ requires changing

[143]Also note that, while antecedent-type A for the functional type $A\backslash B$ has to stand in the left periphery of the antecedent sequent of \R, the antecedent for the anaphoric type $A \hookrightarrow B$ of **LA** has to be left-peripheral into the non-linear domain (see Fig. 7.13). This fact shows, then, that $|R$ encode the Permutation rule (compare derivations in Fig. 7.25).

$$\dfrac{\dfrac{A \Rightarrow A \qquad X, B, Y \Rightarrow C}{A, X, B|A, Y \Rightarrow C}\,|\mathrm{L}_1}{X, B|A, Y \Rightarrow C|A}\,|\mathrm{R}_1$$

Figure 7.24 – Splitting $|R$, where $n = 1$

$$\dfrac{\dfrac{A \Rightarrow A \qquad X, B, Y \Rightarrow C}{X, \{A, A \hookrightarrow B\}, Y \Rightarrow C}\,\hookrightarrow\mathrm{L}}{X, A \hookrightarrow B, Y \Rightarrow A \hookrightarrow C}\,\hookrightarrow\mathrm{R} \qquad\qquad \dfrac{\dfrac{\dfrac{A \Rightarrow A \qquad X, B, Y \Rightarrow C}{X, A, A\backslash B, Y \Rightarrow C}\,\backslash\mathrm{L}}{A, X, A\backslash B, Y \Rightarrow C}\,\mathrm{P} \Rightarrow}{X, A\backslash B, Y \Rightarrow A\backslash C}\,\backslash\mathrm{R}$$

Figure 7.25 – Derivations of $|R$, when $n = 1$ in **LA** and in **LJ**

$$\dfrac{\dfrac{\dfrac{\dfrac{A \Rightarrow A \qquad \dfrac{A \Rightarrow A \qquad X, B_1, Y_1, \ldots, B_n, Y_n, Z \Rightarrow C}{X, A, A\backslash B_1, Y_1, \ldots, B_n, Y_n, Z \Rightarrow}\,\backslash\mathrm{L}}{\vdots}\,\backslash\mathrm{L}}{\dfrac{X, A, A\backslash B_1, Y_1, \ldots, A, A\backslash B_n, Y_n, Z \Rightarrow C}{\vdots}\,\mathrm{P} \Rightarrow}}{\dfrac{A, \ldots, A, X, A\backslash B_1, Y_1, \ldots, A\backslash B_n, Y_n, Z \Rightarrow C}{\vdots}\,\mathrm{C} \Rightarrow}}{\dfrac{A, X, A\backslash B_1, Y_1, \ldots, A\backslash B_n, Y_n, Z \Rightarrow C}{X, A\backslash B_1, Y_1, \ldots, A\backslash B_n, Y_n, Z \Rightarrow A\backslash C}\,\backslash\mathrm{R}}$$

Figure 7.26 – Derivation of $|R$ in **LJ**

the order as well as deleting repeated occurrences of the antecedent formula.

In **LLC** personal pronouns are assigned the syntactic anaphoric type $n|n$, given that they work as a nominal in the presence of a nominal antecedent. In semantic terms, a pronoun denotes the identity function $\lambda x.x$ over individuals; the reference of a pronoun is identical with the reference of its nominal antecedent.[144] Hence, the logical analysis of sentences containing pronouns makes use of $|R$ and $|L$. By applying them, we alternatively get free and bound readings for pronominal expressions. Even though personal, possessive and also reflexive pronouns are similar in that they all take their reference from a nominal antecedent, Jaeger explicitly states that his account is not intended to encompass reflexive pronouns (nor reciprocals). The reason for this restriction on the scope of his logical grammar is twofold. On the one hand, the structural relationship of c-command, which seems to be a necessary condition to bind reflexives, does not also seem necessary for personal and possessive anaphoric pronouns. Although in many prototypical cases of binding, the binder does in fact c-command (and also precedes) the pronoun, the following

[144]Possessive pronouns, however, are not assigned a category of their own. Instead, Jaeger treats a complex expression like *his X*, where X is of type *cn*, as a single lexical unit of category $n|n$, which denotes the (Skolem) function X' mapping individuals to their Xs.

sentences show that personal and possessive pronouns, but not reflexives, can be bound despite not being c-commanded by their binder.

(221) Everybody$_1$'s mother loves him$_1$.

(222) The policemen turned a citizen of each state$_1$ over to its$_1$ governor.

(223) *John$_1$'s father admires himself$_1$.

Jaeger claims then that bound pronouns cannot be restricted to configurations where they are c-commanded by their binder. According to him, an adequate logic for pronominal binding need only impose one constraint: the antecedent has to precede the pronoun. This of course immediately tells us that his proposal is not intended to meet the anti-locality condition (or: Principle B) for pronouns.

The second reason is that the c-command condition is not encoded into **LLC**, due to Jaeger's own interest in giving a unified proposal for pronominal anaphora and ellipsis. In effect, elliptical constructions do not also satisfy the c-command restriction on the position of the ellipsis antecedent. It follows that a uniform treatment for pronouns and ellipsis must not be based on the c-command relationship between the anaphoric item and its binder. Because of this, the only structural constraint on binding formalized by the rule $|L$ of **LLC** is the requirement that the binder precede the anaphoric element.[145]

It should hopefully be clear at this point that the key difference between Jaeger's systems lies in the fact that the analysis of bound pronouns requires structural rules in **LA**, whereas these rules may be left implicit in **LLC**. As a result, those sentences which can be proved in **LA** by means only of the logical rules for anaphor (i.e. $\hookrightarrow L$ and $\hookrightarrow R$), can also be proved in **LLC** without using structural rules; in this case a single logical rule is sufficient, namely $|R$. By contrast, sentences containing bound pronouns, which are proved in **LA** by using both $IM \Rightarrow$ and $C \Rightarrow$, can be obtained in **LLC** by using only $|L$. In addition, since the $|R$ rule of **LLC** is more general than its **LA**-counterpart, it can be used to analyze sentences which contain co-anaphoric pronouns (i.e. pronouns depending on the same antecedent). By contrast, in **LA** this requires the structural rule $C \Rightarrow$ (compare the **LA**- and **LLC**-proofs given in Figs. 7.19 on page 144 and 7.27, respectively). The following **LLC**-proofs, corresponding to the sentences in (224–227), help illustrate the role of $|R$ in the analysis of free pronouns.

(224) He$_1$ drinks and he$_1$ smokes.

[145] As we have said, c-command seems a necessary condition for anaphors (reflexive and reciprocal pronouns), Thus, this pure linear order condition on the binder for anaphoric pronouns is not sufficient for reflexives, as the following well-known examples show:

i. a. John talked to Mary$_1$ about herself$_1$.
b. *John talked about Mary$_1$ to herself$_1$.

$$\dfrac{\dfrac{\begin{array}{c}\vdots\\ n, n\backslash s \Rightarrow s\end{array} \quad \dfrac{\begin{array}{c}\vdots\\ n, n\backslash s \Rightarrow s\end{array} \quad s \Rightarrow s}{n, n\backslash s, s\backslash s \Rightarrow s}\backslash L}{n, n\backslash s, (s\backslash s)/s, n, n\backslash s \Rightarrow s}/L}{n|n, n\backslash s, (s\backslash s)/s, n|n, n\backslash s \Rightarrow s|n}|R$$

Figure 7.27 – Derivation for: *He*$_1$ *drinks and he*$_1$ *smokes* in **LLC**

$$\dfrac{\dfrac{\begin{array}{c}\vdots\\ n, n\backslash s, (s\backslash s)/s, n, n\backslash s \Rightarrow s\end{array}}{n|n, n\backslash s, (s\backslash s)/s, n, n\backslash s \Rightarrow s|n}|R}{n|n, n\backslash s, (s\backslash s)/s, n|n, n\backslash s \Rightarrow (s|n)|n}|R$$

Figure 7.28 – Derivation for: *He*$_1$ *drinks and she*$_2$ *smokes* in **LLC**

$$\dfrac{\begin{array}{c}\vdots\\ n, n\backslash s, (s\backslash s)/s, n, (n\backslash s)/n, n \Rightarrow s\end{array}}{n|n, n\backslash s, (s\backslash s)/s, n, (n\backslash s)/n, n|n \Rightarrow s|n}|R$$

Figure 7.29 – Derivation for: *He*$_1$ *smokes and John*$_2$ *hates him*$_1$ in **LLC**

$$\dfrac{\begin{array}{c}\vdots\\ n, (n\backslash s)/n, n \Rightarrow s\end{array}}{n|n, (n\backslash s)/n, n|n \Rightarrow s|n}|R$$

Figure 7.30 – Derivation for: **He*$_1$ *hates him*$_1$ in **LLC**

(225) He$_1$ drinks and she$_2$ smokes.

(226) He$_1$ smokes and John$_2$ hates him$_1$.

(227) *He$_1$ hates him$_1$.

All of these derivations exhibit the fact that a constituent becomes an unsaturated syntactic category $X|n$ as a result of the occurrence of (at least) a free pronoun within the former. The first and third derivations, in addition, exhibit the fact that pronouns may be co-anaphoric even though no antecedent is present in the clause. The second derivation, in turn, returns a syntactic type that records how many free pronouns the clause contains. The syntactic type obtained as a result of the $|R$ rule cannot however display the position (or the syntactic function) in which the free pronoun occurs, given that subject and object pronouns are all assigned the type $n|n$. As a result, the type $s|n$ can be ambiguously assigned to a sentence containing a free subject (224) or a free object pronoun (226).[146] This notwithstanding, the syntactic function of a pronoun will be reflected in the λ-term corresponding to the proof. Finally, as the proof for (227) shows, the type $s|n$ can even be assigned to a sentence containing a pronoun that does not satisfy Principle B.

7.4 Summing Up

Both of Jaeger's systems **LA** and **LLC**, in contrast with **L**, include some restricted form of Permutation and Contraction. In **LA**, these are explicitly encoded by the rules $IM \Rightarrow$ and $C \Rightarrow$, respectively. In **LLC**, a long-distance form of Contraction is implicitly encoded by the rules for the type-constructor $|$. As a result, Jaeger's systems are both better able to account for linguistic phenomena that require reusing semantic resources, such as bound and co-anaphoric pronouns. However, neither **LA** nor **LLC** impose any kind of restriction on the distance between the pronoun and its antecedent, nor a hierarchical ordering on the antecedent of the anaphoric expression. Put differently, neither system encodes structural relationships nor stipulates syntactic domains for binding.

In this sense, **LLC** and Jacobson's combinatory system $\mathbf{J}_{GZ}$ are very much alike. Note that both directional versions of Jacobson's **G**, as well as their generalized versions, can be derived in **LLC** by using $|$R. Furthermore, two of the four directional versions of **Z** can be obtained by using $|$L. While those versions in which the second — nominal — argument stands to the left of the premise function — $B\backslash(n\backslash A)$ and $(n\backslash A)/B$ — can be derived in **LLC**, the remaining two require Permutation (or, alternatively, a cataphoric rule instead an anaphoric one; see Fig. 7.31 (and cf. [Jäg99], [Jäg01]).

$$\dfrac{\dfrac{\dfrac{\dfrac{B \Rightarrow B \qquad \begin{matrix}\vdots\\ A/n, n \Rightarrow A\end{matrix}}{(A/n)/B, B, n \Rightarrow A}\,/\mathrm{L}}{(A/n)/B, B|n, n \Rightarrow A}\,{}^{*}|\mathrm{L}}{(A/n)/B, B|n \Rightarrow A/n}\,/\mathrm{R}}{(A/n)/B \Rightarrow (A/n)/(B|n)}\,/\mathrm{R} \qquad\qquad \dfrac{\dfrac{\dfrac{\dfrac{B \Rightarrow B \qquad \begin{matrix}\vdots\\ A/n, n \Rightarrow A\end{matrix}}{B, B\backslash(A/n), n \Rightarrow A}\,/\mathrm{L}}{B|n, B\backslash(A/n), n \Rightarrow A}\,{}^{*}|\mathrm{L}}{B|n, B\backslash(A/n) \Rightarrow A/n}\,/\mathrm{R}}{B\backslash(A/n) \Rightarrow (B|n)\backslash(A/n)}\,\backslash\mathrm{R}$$

Figure 7.31 – Illicit Derivation for $\mathbf{Z}>$ and $\mathbf{Z}\lessgtr$ in **LLC**

LLC and $\mathbf{J}_{GZ}$ are alike also insofar as they suffer certain empirical inadequacies, due to the fact that they both disregard Principle B of GB in the logic for pronouns. Still, overall both systems more than compensate for their flaws; so much so, in fact, that they will be the starting points for our own combinatory and type-logical analysis of pronouns.

[146]Note that, as a matter of fact, the sequence Y_i in $|R$ can contain a type B.

8 Proposal: Overt Anaphoric Control in Categorial Grammars[147]

8.1 Overt Controlled Subjects

Over the course of this chapter we will bring together the key results from Part I and Part II; these will be our stepping stones towards a new logical account of overt controlled subjects.

Accordingly, we begin by reviewing the key themes and results of the discussion up to this point. To this end we first briefly return to the linguistic strand of our inquiry, so as to remind the reader of the main assumptions that were put forward and defended in Chs. 2–4. In particular, we remind the reader that in Section 4.3 we argued that *overt anaphoric pronouns occurring in control sentences from Romance languages are the real syntactic-semantic subject of the infinitive clause*. We saw that this claim appears to be plausible in its own right; once we give up certain generative assumptions, moreover, any remaining counterarguments appear to lose much of their force.

Subsequently we return to the two strategies for addressing the problem of reusing resources in a logical grammar that were encountered in Chs. 6 and 7. In Sections 8.2 and 8.3, we evaluate the previously discussed Combinatory and Type-Logical syntactic proposals *vis à vis* the phenomenon of controlled overt pronouns in Romance languages.

We now review the following topics from Part I:

I. Propositional vs. control verbs

II. Infinitive subject types

III. Propositions vs. properties

IV. Infinitive subject position

I. Propositional vs. control verbs

In Chapter 3 we saw that propositional verbs from literary 19$^{\text{th}}$ century Spanish and Italian can select an — uninflected — infinitive complement as well as, of course, a finite

[147] Previous versions of this chapter were presented at *Manchester Forum in Linguistics*, Manchester-England; *I Colóquio de Semântica Referencial*, São Carlos-Brazil [Cor15]; *Conference on Formal Grammar 2016*, Bozen/Bolzano-Italy [CM16]; *Formal Grammar 2017*, Toulouse-France [Cor17a] and *14th International Conference on Distributed Computing and Artificial Intelligence*, Porto-Portugal [Cor17b].

sentence. This is also the case for European and Brazilian Contemporary Portuguese, a language that contains an inflected form of infinitive along with the standard uninflected form. In this respect these four languages behave differently from others, like English, which only admit finite sentences as complements of propositional verbs. Further and even more importantly, in Romance languages propositional infinitive complements can host an overt subject; not so in English.

We also saw that the so-called control verbs select infinitive complements, though not finite ones.[148] Moreover in these Romance languages — again, unlike English — infinitive complements of control verbs may host an overt subject as well as a covert one. Yet, this subject must observe strict semantic and therefore morphological restrictions.

II. Infinitive subject types

In Ch. 4, we adopted Landau's account of control for Romance languages. In the course of that discussion, we identified a class of propositional infinitive complements whose subject can be referentially free or coreferential with another expression; we called these F-infinitives. Following Landau's terminology, we saw that F-infinitives can host either [+R] or [−R] infinitive subjects: namely, R-expressions, personal pronouns, *pro* and PRO. F-infinitives thus tolerate different syntactic-semantic types of subjects and behave much like finite complements.

In contrast, infinitives selected by control verbs — C-infinitives — may only host a [−R] subject. As we have seen, Landau [Lan04], [Lan06] does not assume any overt form of a [−R] nominal. In his theory the null PRO is the only nominal carrying a [−R] feature, despite the fact that [R] mainly formalizes a semantic or referential distinction, and not a phonological one. Of course, this particular thesis of Landau's conflicts with the linguistic phenomenon we are interested in: pronouns that behave like obligatory controlled PRO, despite being overt. As the reader will recall, we addressed this problem by arguing that bearing a [−R] feature is a necessary condition for PRO, but not also a sufficient one. And we argued that, as a result, pronominals other than the phonologically null PRO can bear the [−R] feature.

In the next subsection, we will be sketching two ways for dealing with overt [−R] pronouns in control clauses from a categorial perspective.

III. Propositions vs. properties

Infinitives selected by propositional verbs behave much like finite complements in these Romance languages. Their subject position may certainly be occupied by an overt, ref-

[148] Most of these verbs may also select a subjunctive clause and give rise to an obviation (or: anti-control) effect [HSM01]: the embedded — overt or covert — subject displays a disjunctive reference with respect to the matrix subject.

erentially independent nominal expression. Thus, the only apparent reason to not admit that the former may denote a proposition seems to be based on the semantic assumption that infinitives do not have a propositional nature. Our data, however, directly undermines this semantic claim.

The situation is somewhat less clear in the case of infinitives selected by control verbs, which may only contain an — overt or covert — pronominal subject. In order for an infinitive to become saturated and denote an (open or closed) proposition, we would have to admit: i) that the infinitive itself denotes a propositional function (but see [Chi84a]); and ii) that the pronominal subject denotes an individual of some kind. Thus, the propositional semantics for controlled complements does not sit well within the categorial semantic framework. As we have seen, in Categorial Grammars nominal expressions and pronominal expressions belong to different syntactic and semantic categories. Furthermore, in a free-variable semantics, free and bound pronominal expressions are assumed to denote not individuals, but rather some kind of function. As a consequence, overt pronouns in control clauses may not denote an entity inside the complement clause, even though they necessarily end up being bound by a nominal argument in the matrix (i.e. higher) clause. In other words, even when the semantic representation of the entire clause displays two occurrences of the same nominal, this fact does not force us to assume that the infinitive clause itself denotes a saturated function.

Thus, we follow a well-established categorial tradition in assuming that infinitive complements selected by control verbs denote a property (of some kind; see [Chi84b]) in spite of the presence of a bound surface pronoun within the infinitive complement. The specific property that the infinitive clause denotes will depend on the specific categorial treatment of bound pronouns and control verbs we adopt. More specifically, but still approximately: $\lambda x.(VP'\ x)$ or a multiple-binding function $\lambda x.((VP'\ x)\ x)$.

IV. Infinitive subject position

Finally, we remind the reader of a relevant prosodic fact concerning infinitive clauses. Although infinitive clauses from Romance languages may take their own overt subject, much like finite clauses, infinitive and finite clauses generally exhibit a reciprocally inverse word-order with respect to their subject. Albeit with some exceptions, a pattern in particular has been observed in this connection.

In Spanish and Italian, the subjects of (adverbial or subject) infinitive constructions necessarily occupy the post-verbal position (cf. Section 3.3). In European Portuguese subjects within infinitive clauses of different kinds normally occur in the post-verbal position, but the pre-verbal position can also be admitted (in, for example, infinitives selected by factive verbs). In contrast, in Brazilian Portuguese such subjects normally occupy the pre-verbal position (cf. Section 3.2.2). In fact, this word-order is the same as the one we observe in cases of infinitives selected by control verbs. Hence, the (pre- or post-) nominal argument position that we take to syntactically categorize the infinitive verb in control structures receives support from other infinitive contexts.

8.1.1 Anaphoric Pronouns or Anaphoric Verbs

Having argued that overt pronouns occurring in controlled clauses are instances of [−R] pronominals, we now sketch some possible lexical and syntactic categorial treatments for the same.

A first option for a lexical treatment would be to adapt Szabolcsi's and Morrill's proposals on bound pronouns. This would entail assigning **a multiple-binding meaning to the overt obligatory anaphoric pronoun** that occurs in controlled contexts. Nevertheless, given that these anaphoric pronouns are morphologically indistinguishable from free — nominative — personal pronouns, we should concede that for each [+R] personal pronoun, there is a morphologically identical counterpart that bears a [−R] feature. But then we would be assigning different lexical meanings to two morphologically identical linguistic expressions, which seems a bit excessive. For this reason, we shall not be pursuing this lexical categorial route any further.

Fortunately, there are other routes to take if we bear in mind that, after all, a nominal bearing the [−R] feature is one incapable of independent reference (cf. [Lan04]:841). Thus, as a second option, we could assume that [−R] is not an inherent — lexical — feature of the pronominal expression, but rather that it is triggered by the syntactic-semantic context in which the latter occurs. In other words, [±R] would be a semantic feature that the pronoun carries *because of the verb* that selects the clause in which it occurs. But then the inability for an overt pronoun in control clauses to have an independent reference — despite being specified for all of its ϕ-features — would not be a defect of the pronoun itself, but rather an effect of the control verb.

In a Categorial Grammar we could then encode the constraints on the kind of subject that the embedded infinitive clause can host in **the syntactic-semantic type of the main verb**, rather than in that of the pronominal. In other words, instead of focusing on the anaphoric pronominal subject, we could treat the control verb as the *de facto* anaphoric expression. After all, the control relationship is none other than the one that duplicates the controllee.

Just as there are two main categorial routes to deal with anaphoric pronouns, so there are two ways to manage anaphoric verbs: namely, as lexical and derivational meanings. First, the duplication of the matrix — subject or object — nominal resource could be encoded in **the lexical meaning of the control verb**. Alternatively, the duplication effect could be attached to **the semantic — functional — operation triggered by**

the syntax of the control verb. In fact, we have seen several syntactic operations that yield a multiple-binding semantics: for instance, the **W** and **Z** operations, and the Equi and |L rules.

In the following, we apply some of the categorial tools that were previously reviewed, to our linguistic object. And we sketch two proposals for treating anaphoric verbs: a Combinatory account and a Type-Logical account.

8.2 A Combinatory Proposal

8.2.1 Preliminary Proposal: $\mathbf{J}_{\mathsf{GZ}}$

As we saw in Chapter 6, alongside his lexical account Steedman outlines a preliminary syntactic combinatory approach to (English) control verbs. He proposes the combinatory Equi rule to specifically account for the control relation triggered by *want* (see Fig. 6.3). Semantically, this rule corresponds to the multiple-binding **W** combinator. Syntactically, the Equi rule allows the finite control verb *want* to combine with the *to*-infinitive complement despite the lack of an overt embedded subject.

This rule could also be used for Romance forms of *want* 'querer'/'volere', which may select a subjectless infinitive clause. Let us then sketch a preliminary combinatory proposal by combining Steedman's verbal assignment with Jacobson's pronominal type. Note, in passing, the rightward direction of the infinitive-type in the Spanish lexicon. The derivation for the Spanish sentence in (228) is given in Figure 8.1. There, note that Equi delivers the multiple-bound term $\lambda z.((want'\ (phone'\ z))\ z)$.

(228) Julia quería [telefonear].
Julia wanted phone.INF
'Julia wanted to phone.'

ella : $n|n : \lambda x.x$

querer : $(n\backslash s)/s_{to} : \lambda x.\lambda y.(want'\ (x\ y))$

telefonear : $s_{to}/n : \lambda x.(phone'\ x)$

$$
\dfrac{\dfrac{Julia}{j' : n} \qquad \dfrac{\dfrac{quer\acute{\imath}a}{want' : (n\backslash s)/s_{to}} \quad \dfrac{telefonear}{phone' : s_{to}/n}}{\lambda z.((want'\ (phone'\ z))\ z) : n\backslash s}\,\text{Equi}}{((want'\ (phone'\ j'))\ j') : s}\,>
$$

Figure 8.1 – Derivation for: *Julia quería telefonear* in $\mathrm{S}_{\mathbf{W}}$

In order to derive the sentence in (229), where the infinitive clause contains a pronominal subject, we use the rules of $\mathbf{J}_{ZG}$ instead. In this case the expected multiple-binding semantics is obtained by using **Z**, which ensures that the variables corresponding to the nominal matrix subject and to the infinitive pronominal subject are both bound by the same λ-operator.

(229) Julia quería [telefonear **ella**].
Julia wanted phone.INF she.NOM
'Julia wanted to be the one who phoned.' [PG87]:160

$$
\dfrac{\dfrac{Julia}{j' : n} \quad \dfrac{\dfrac{\dfrac{quer\acute{\imath}a}{want' : (n\backslash s)/s_{to}}}{\lambda u.\lambda w.(want'\ (u\ w))\ w) : (n\backslash s)/(s_{to}|n)}\ \mathrm{Z} \gtrless \quad \dfrac{\dfrac{\dfrac{\dfrac{telefonear}{phone' : s_{to}/n}}{\lambda y.\lambda z.(phone'\ (y\ z)) : (s_{to}|n)/(n|n)}\ \mathrm{G} < \quad \dfrac{ella}{\lambda x.x : n|n}}{\lambda z.(phone'\ z) : s_{to}|n}\ >}{}}{\lambda w.((want'\ (phone'\ w))\ w) : n\backslash s}\ >}{((want'\ (phone'\ j'))\ j') : s}\ >
$$

Figure 8.2 – Derivation for: *Julia quería telefonear ella* in $\mathbf{J}_{GZ}$

In fact, in the previous derivation **G** is applied to the type s_{to}/n to subsequently allow the infinitive (intransitive) verb to combine with its own overt pronominal subject. Functional application then yields the functional type $s_{to}|n$. When **Z** is applied to the control verb-type, we obtain $(n\backslash s)/(s_{to}|n)$; thus the latter can combine with the infinitive clause-type containing the pronominal subject by functional application, delivering the multiple-bound term $\lambda w.((want'\ (phone'\ w))\ w)$.

Analogously, we can ensure the multiple-binding semantics, and so the control relation, by applying **Z** to the three-place verb *promise* 'prometer'/'promettere (di)'. In this case, the control verb combines with its nominal complement before **Z** is applied. The resulting lifted category can then take the infinitive complement containing the pronominal gap as its argument. The relevant steps of the proof for (230) are given in Figure 8.3.[149]

(230) Julia (le) prometió a Marta [encargarse **ella** del asunto].
Julia CL promised to Marta deal.INF-CL she.NOM of-the matter
'Julia promised Marta that she would deal with the matter.' [PG87]:161

prometer : $((n\backslash s)/s_{to})/n : \lambda x.\lambda y.\lambda z.(((((promise'\ y))\ x))\ z)$

As the two last derivations show, in order for a subject control relation to be triggered by the intransitive verb *querer* and the transitive verb *prometer*, the directional

[149]For a categorial treatment of (Catalan) clitics and clitic doubling, see [MG92]; [SA96]; [Val13]. For a treatment of movement of clitic pronouns in Italian in a calculus of pregroups see [CL01], [CKM14]. Later on, we shall present a more detailed lexicon in which we spell out the syntax and semantics of prepositions and of transitive infinitive verbs.

$$\dfrac{\dfrac{\dfrac{\dfrac{\textit{prometió (a)}}{\lambda x.\lambda y.\lambda z.((((promise'\ y)\ x))\ z) : ((n\backslash s)/s_{to})/n} \quad \dfrac{Marta}{m' : n}}{\lambda y.\lambda z.((((promise'\ y)\ m'))\ z) : (n\backslash s)/s_{to}}>}{\lambda y.\lambda z.((((promise'\ (y\ (z)))\ m'))\ z) : (n\backslash s)/s_{to}|n}\mathrm{Z} \gtreqless \qquad \dfrac{\vdots}{\lambda w.((\textit{deal-with}'\ a')\ w) : s_{to}|n}}{\lambda z.((((promise'\ ((\textit{deal-with}'\ a')\ z))\ m'))\ z) : n\backslash s}>$$

Figure 8.3 – Schematic derivation for: *Julia prometió a Marta encargarse ella del asunto* in $\mathbf{J}_{GZ}$

mixed $\mathbf{Z} \lesseqgtr$ rule is applied to the very same type: $(n\backslash s)/s_{to}$. In contrast to *promise*, most three-place verbs induce an object control relation. A brief inspection of **Z** (see Fig. 7.6) reveals that none of its directional versions can be applied to a verb V of type $((n\backslash s)/s_{to})/n$ to yield $((n\backslash s)/(s_{to}|n))/n$ with the suitable multiple-binding semantics $\lambda x.\lambda y.\lambda z.(((V'\ ((y\ (x)))\ x))\ z)$. This is because **Z** guarantees a prominence-like condition on the binder: the binder cannot be an argument lower than the bindee in the argument structure. Thus, an object nominal argument in the lower (or: rightmost) position cannot be the binder of a higher pronominal gap. So, if an object control verb were assigned the same type as *promise*, the control relation could not be ensured (by **Z**). The object control relation and so the **Z** rule require another type-assignment for verbs of this kind. If, following Jacobson [Jac03]:61, we assume a wrap operation $/_W$, we could then assign $((n\backslash s)/_W n)/s_{to}$ to these verbs. Under this argument-inverse assignment, a three-place verb can now be syntactically and semantically lifted by $\mathbf{Z} >$. The wrap operation in turn ensures the correct word-order of the nominal and infinitive complements.

Despite being preliminary, this proposal based on $\mathbf{J}_{GZ}$ already allow us to formalize certain expected properties at the syntactic-semantic level. Notably, **G** allows an infinitive verb to host an overt subject pronoun; moreover, **Z** ensures the control of the embedded pronominal gap. The control relation can be formalized as a semantic operation performed by the main verb, since it undergoes the **Z** operation. In addition, with the help of the new type-constructor $|$, it is possible to ensure that the pronoun remains free within the infinitive clause, and consequently that the infinitive clause receives a functional syntactic-semantic type. The pronoun, however, gets the expected bound reading. In sum, under the previous verbal-type assignments based on Steedman's proposal, the system $\mathbf{J}_{GZ}$ appears at first glance to be suitable to account for overt pronouns in control clauses.

However, these lexical assignments run into a familiar problem, afflicting all of the above-mentioned verbs: overgeneration. Here we detail the nature of the problem with reference only to *want*, since the other verbs (*promise*, *persuade*) follow the same pattern. In particular, the type $(n\backslash s)/s_{to}$, which Steedman uses to account for the double — ECM and control — structure triggered by *want*, is inadequate for the Romance counterparts of the latter.[150] This is because these verbs, unlike *want*, do not admit a nominal infinitive

[150]By the same token, this type assignment is inadequate for the ECM *want*, once the **G** and in

subject. At first sight it might appear straightforward to change the type of the infinitive verb so it can no longer take a nominal subject. But recall that infinitives from Romance languages take nominal arguments in complement clauses different from those selected by control verbs. Thus, the verb that is responsible for restricting the kind of subject that the infinitive clause may host is not the infinitive, but the main — finite — verb. So instead of changing the type of the infinitive verb, we should actually change the sentential argument of the control verb. In order to avoid the problem of overgeneration, we might opt for lifting the syntactic and semantic type of the control verb. By adopting a lexical perspective we then assign the **Z** type and the **Z** meaning to the control verb. For example:

querer : $(n\backslash s)/(s_{to}|n)$: $\lambda x.\lambda y.((want'\ (x\ y))\ y)$

Of course, once we do this the opposite problem arises: namely, undergeneration. In fact, under this lexical assignment, the type $(n\backslash s)/(s_{to}|n)$ cannot take the type s_{to}/n as its argument. As a result the control verb cannot combine with the infinitive clause unless the infinitive firstly combines with its own pronominal subject. Consequently, prototypical control sentences can no longer be derived.

To overcome this problem we may opt to duplicate the lexical entry for the control verb. In one of these entries, the control verb will be assigned the type $(n\backslash s)/(s_{to}|n)$; in the other, the type $(n\backslash s)/(s_{to}/n)$ (or $(n\backslash s)/(n\backslash s_{to})$, depending on the particular Romance language). Both types limit their infinitive complement by taking only a functional type. In spite of their differences, both syntactic types may be assigned the same double-binding semantics. Alternatively, these two possible unsaturated arguments of the control verb — $s_{to}|n$ and s_{to}/n — may be encoded into a single lexical entry. In a Type-Logical Grammar this can be done by extending the vocabulary of the system so as to add a disjunctive connective as a type-constructor (see [Mor94b]165; [Dow85]). We shall return to this approach in Section 8.3.

For now, let us turn our attention to other worries, which arise as the structure of the control clause becomes more complex and more nominal gaps occur within it. In fact, we have just seen that three-place (object) control verbs call for a wrap operation (or a wrap type-constructor). Obviously, infinitive verbs may also be such three-place functions. As an example, consider the Brazilian Portuguese sentence in (231) below, whose infinitive clause displays the SVO word-order:

(231) Os meninos querem [**eles** embrulhar(**em**) os presentes].
the boys want they wrap.INFL.(3PL) the gifts
'The boys want to wrap the gifts themselves.' [Negrão 1986]

particular the **Z** rules are available.

Since the infinitive clause contains a free overt pronoun, the type $s_{to}|n$ can be derived by using **G** as previously shown. The main difference with the previous derivation in Fig. 8.2 is that here the two-place infinitive verb-type first combines with its nominal object before applying **G** to the former. The type $s_{to}|n$ will then combine by functional application with the (lexically) lifted **Z**-type $(n\backslash s)/(s_{to}|n)$ assigned to the control verb *querer*.

Once the infinitive verb can take two arguments, the infinitive clause in the sentence (232) below can be recognized as being of type $s_{to}|n$. Since the latter can subsequently combine with the type of the main verb, this sentence can be recognized by the $\mathbf{J}_{GZ}$ system despite being ungrammatical. Like the previous grammatical sentences, here the infinitive clause contains a free pronoun; nevertheless, unlike the grammatical examples above, here the overt infinitive subject is not that free pronoun, but a referential expression. The relevant steps of the proof for (232) are given in Fig 8.4 (see also Fig. 7.8).

(232) *O menino quer [**Maria** namorar ele].
the boy wants Mary date.INFL.(3SG) him
'The boy wants for Mary to date him.'

$$\dfrac{\dfrac{\dfrac{\dfrac{Maria}{m' : n}}{\lambda y.(y\ m') : s_{to}/(n\backslash s_{to})}\,\mathrm{T}{>}}{\lambda v.\lambda u.((\lambda y.(y\ m'))\ (v\ (u))) : (s_{to}|n)/((n\backslash s_{to})|n)}\,\mathrm{G}{>} \qquad \dfrac{\dfrac{\dfrac{namorar}{date' : (n\backslash s_{to})/n}}{\lambda z.\lambda w.(date'\ (z\ (w))) : ((n\backslash s_{to})|n)/(n|n)}\,\mathrm{G}{>} \qquad \dfrac{ele}{\lambda x.x : n|n}}{\lambda w.(date'\ w) : (n\backslash s_{to})|n}\,{>}}{((\lambda w.(date'\ w))\ m') : s_{to}|n}\,{>}$$

Figure 8.4 – Schematic Derivation for: **O menino quer Maria namorar ele* in $\mathbf{J}_{GZ}$

In brief, the problem arises because the type $s_{to}|n$ does not distinguish between a subject and an object pronominal gap; it only expresses that the infinitive sentence contains *a* pronominal gap. Because of this, an infinitive clause containing a free pronoun in an embedded position within the nominal subject can also be assigned the type $s_{to}|n$. Certainly, the constituent *o amigo d'ele* will be of type $n|n$. As a consequence, the sentence in (233) can also be recognized by the $\mathbf{J}_{GZ}$ system despite being ungrammatical.

(233) *João quer [**o amigo d'ele** ganhar].
John wants the friend of-him win.INFL.(3SG)
'John wants his friend to win.'

8.2.2 Modifiying $\mathbf{J}_{GZ}$: $\mathbf{J}_{G^2Z}$ and $\mathbf{J}_{G^4Z}$

We might hope to overcome this syntactic *double-face* by distinguishing between two anaphoric type-constructors: $\|$ and $|$. With two type-constructors at hand, a subject pronoun could then be assigned the type $n\|n$ while object pronouns continue to be assigned $n|n$. Subject pronouns, like subjects in general, usually stand to the left of the functor;

objects normally stand to the right of the latter (in nominative-accusative languages). Thus, a first step to amend the $\mathbf{J}_{GZ}$ system is to assign only one of the directional versions of **G** to each type: **G** $<$ for $\|$ and **G** $>$ for $|$ (analogously for the generalized versions $\mathbf{G}_n <$ and $\mathbf{G}_n >$). Let us name the modified system $\mathbf{J}_{G^2Z}$. While $G >$ and $G_n >$ are as in $\mathbf{J}_{GZ}$, the remaining two **G** rules will introduce the new anaphoric, pre-functional type $\|$. For simplicity, here we display a particular instance of G_n.

$$\frac{M : A\backslash B}{\lambda x.\lambda y.(M(x(y))) : (A\|C)\backslash(B\|C)}\, G < \qquad \frac{M : (A\backslash B)|C}{\lambda w.\lambda x.\lambda y.(M(x(y))w) : ((A\|D)\backslash(B\|D))|C}\, G_1 <$$

Figure 8.5 – (Backward and Generalized) G rules of $\mathbf{J}_{G^2Z}$

Since in $\mathbf{J}_{G^2Z}$ the type $s_{to}\|n$ corresponds to an infinitive sentence containing a free subject pronoun, we can now give a more accurate lexical syntactic assignment for control verbs. For example:

quer : $(n\backslash s)/(s_{to}\|n) : \lambda x.\lambda y.((want'\ (x\ y))\ y)$

Of course, the proposed system is not without its problems; this will become evident when we apply it to linguistic structures that do not exhibit the SVO word-order. But first let us emphasize its positive traits. To begin with, $\mathbf{J}_{G^2Z}$ does not recognize the ungrammatical sentence in (232) above. This is because though the system will still assign the type $s_{to}|n$ to strings like *Maria namorar ele*, the functional application rule will be blocked. Indeed, $(n\backslash s)/(s_{to}\|n)$ cannot combine with $s_{to}|n$. Second, $\mathbf{J}_{G^2Z}$ also does not recognize the ungrammatical sentence in (233) above. In effect, assuming that *ele* within the constituent *o amigo d'ele* is not of type $n\|n$, the system can no longer assign the type $s_{to}\|n$ to the infinitive clause in which this constituent occurs. Third, $\mathbf{J}_{G^2Z}$ can still derive the grammatical sentence in (231), since it will assign the type $s_{to}\|n$ to the string *eles embrulhar(em) os presentes*. Certainly, in this case, functional application is not ruled out. Fourth, a control clause containing only an overt free object pronoun can also be derived in $\mathbf{J}_{G^2Z}$. Fifth, $\mathbf{J}_{G^2Z}$ records not just the number of free pronominal occurrences, but also the specific positions these pronouns occupy. Thus, for example, the type $(s\|n)|n$ corresponds to a sentence with two pronominal gaps, one in the object and the other in the subject position. Like in $\mathbf{J}_{GZ}$, the $\mathbf{J}_{G^2Z}$-proof of a sentence with more than one pronominal gap requires using the generalized versions of **G**. To see this, compare the syntactic $\mathbf{J}_{G^2Z}$-proofs given in Fig. 8.6, where V_1 stands for an intransitive control verb and V_2 stands for a transitive infinitive verb. With regard to V_1, recall that we have assumed two entries for each control verb, in order to cover both complement possibilities.

$$\dfrac{n \qquad \dfrac{\dfrac{V_1}{(n\backslash s)/(s_{to}\|n)} \qquad \dfrac{n\|n \qquad \dfrac{\dfrac{\dfrac{V_2}{(n\backslash s_{to})/n} \qquad n}{n\backslash s_{to}}\,{>}}{n\|n\backslash(s_{to}\|n)}\,\mathrm{G}{<}}{s_{to}\|n}\,{>}}{n\backslash s}\,{>}}{s}\,{<}$$

$$\dfrac{\dfrac{\dfrac{n}{s_{to}/(n\backslash s_{to})}\,\mathrm{T}{>}}{(s_{to}|n)/((n\backslash s_{to})|n)}\,\mathrm{G}{>} \qquad \dfrac{\dfrac{\dfrac{V_1}{(n\backslash s)/(n\backslash s_{to})}}{((n\backslash s)|n)/((n\backslash s_{to})|n)}\,\mathrm{G}{>} \qquad \dfrac{\dfrac{\dfrac{V_2}{(n\backslash s_{to})/n}}{((n\backslash s_{to})|n)/n|n}\,\mathrm{G}{>} \qquad n|n}{(n\backslash s_{to})|n}\,{>}}{(n\backslash s)|n}\,{>}}{s|n}\,{<}$$

$$\dfrac{\dfrac{\dfrac{n}{\vdots}}{(s|n)/((n\backslash s)|n)} \qquad \dfrac{\dfrac{\dfrac{V_1}{(n\backslash s)/(s_{to}\|n)}}{((n\backslash s)|n)/((s_{to}\|n)|n)}\,\mathrm{G}{>} \qquad \dfrac{\dfrac{\dfrac{n\|n}{(s_{to}\|n)/((n\|n)\backslash(s_{to}\|n))}\,\mathrm{T}{>}}{((s_{to}\|n)|n)/(((n\|n)\backslash(s_{to}\|n))|n)}\,\mathrm{G}{>} \qquad \dfrac{\dfrac{\dfrac{\dfrac{V_2}{(n\backslash s_{to})/n}}{((n\backslash s_{to})|n)/n|n}\,\mathrm{G}{>} \qquad n|n}{(n\backslash s_{to})|n}\,{>}}{((n\|n)\backslash(s_{to}\|n))|n}\,\mathrm{G}_g{<}}{(s_{to}\|n)|n}\,{>}}{(n\backslash s)|n}\,{>}}{s|n}\,{<}$$

Figure 8.6 – Free subject and object pronouns in $\mathbf{J}_{G^2Z}$

As we have mentioned, with the exception of Brazilian Portuguese (BP) the Romance languages we are treating do not exhibit the SVO word-order in infinitive clauses, but rather the VSO order. From a categorial perspective, this means that transitive infinitive verbs in these languages should be assigned the type $(s_{to}/n)/n$, instead of $(n\backslash s_{to})/n$. The former type, of course, can no longer undergo $\mathbf{G} <$ nor $\mathbf{G}_g <$ of $\mathbf{J}_{G^2Z}$; so the type $((s_{to}\|n)/(n\|n))/n$ cannot be derived in this system. In more linguistic terms, this implies that infinitive complements containing a pronominal subject in a post-verbal position are not recognized by our modified system $\mathbf{J}_{G^2Z}$. Needless to say, this spells trouble for us. It appears that we have to revert to the initial $\mathbf{J}_{GZ}$ system or, at least, restore the $\mathbf{G} >$ rule for the new pre-functional anaphoric type-constructor $\|$. We now explore this second option.

Earlier we distinguished two pronominal types: $n\|n$ for subject pronouns and $n|n$ for object pronouns. In these Romance languages, as well as in English, pronouns are morphologically case-marked. We might then think of $n\|n$ as being not merely a subject pronoun, but rather a nominative subject pronoun; similarly for the accusative object pronoun $n|n$. From this perspective, $G <$ and $G >$ of $\mathbf{J}_{G^2Z}$ allow a functional type — a verb, for example — to take an accusative pronoun on its right and a nominative pronoun on its left. They do not allow, however, the former to take a nominative pronoun on its right nor an accusative pronoun on its left. Although these rules are enough to analyse several linguistic structures, certain marked (or non-standard) constructions show that this is not always the case. Consider once again ECM structures, for example (in English and other languages; see 2.2.3). In ECM constructions a (free) *accusative* pronoun may be the subject of an infinitive clause and so it may occupy the *pre-verbal position*. Thus, in order to analyse infinitive clauses from Romance languages as well as other marked structures we need to relax our previous restrictions so as to restore the directional versions of **G** for the anaphoric type-constructors, as in $\mathbf{J}_{GZ}$. In addition to $G >$ and $G <$ of $\mathbf{J}_{GZ}$, the resulting new system $\mathbf{J}_{G^4Z}$ will contain similar G rules for two new directional slash connectives $/\!\!/$ and $\backslash\!\!\backslash$:

$$\frac{M : A /\!\!/ B}{\lambda x.\lambda y.(M'(x(y))) : (A\|C) /\!\!/ (B\|C)}\, G \gg \qquad \frac{M : B \backslash\!\!\backslash A}{\lambda x.\lambda y.(M'(x(y))) : (B\|C) \backslash\!\!\backslash (A\|C)}\, G \ll$$

Figure 8.7 – (Forward and Backward) G rules of $\mathbf{J}_{G^4Z}$

Of course, as for the / and \ type-constructors of a combinatory system based on **AB**, we have to assume other combinatory rules for the slash type-constructors $/\!\!/$ and $\backslash\!\!\backslash$ in $\mathbf{J}_{G^4Z}$: for example, Functional Application and Type-Lifting. Furthermore, a Wrap operation will also be needed to deal with object control verbs, as in Jacobson's proposal.

With these new **G** rules in place we are now ready to handle free subject pronouns, regardless of their pre- or post-verbal position. Figures. 8.8–8.10 display three $\mathbf{J}_{G^4Z}$-proofs for Romance control sentences containing either a pre-verbal or a post-verbal subject pronoun. As an illustration, consider also the simplified preliminary lexicon given below. And observe, in particular, that the syntactic-semantic **Z**-type assigned to the control verbs in the lexicon will ensure that subject pronouns will be free within the infinitive clause but semantically bound by a matrix nominal.[151]

(234) Gianni ha deciso [di intervenire **lui**]. (IT)
John has decided [COMP intervene.INF he.NOM]
'John has decided to intervene himself.' [Car99]

(235) A polícia forçou os manifestantes a [**eles** saírem]. (BP)
the police forced the protesters to they leave.INFL.3PL
'The police forced the protesters to leave.' [Rab04]

decide (di) : $(n \backslash\!\!\backslash s)/(s_{to}\|n) : \lambda x.\lambda y.((decide'\ (x\ y))\ y)$

[151]Alternatively, if the language contains $\nearrow$ as a long-distance functional type-constructor, an object control verb can be assigned the types $((n\backslash s)/n) \nearrow (s_{to}\|n)$ and $((n\backslash s)/n) \nearrow (n\backslash s_{to})$. Similarly, if the language contains the wrap type constructor $\uparrow$, the types $((n\backslash s) \uparrow n)/(s_{to}\|n)$ and $((n\backslash s) \uparrow n)/(n\backslash s_{to})$ can be assigned to this kind of verb.

ella : $n\|n : \lambda x.x$

eles : $n\|n : \lambda x.x$

encargarse (de) : $(s_{to} \mathbin{/\!\!/_W} n)/n : \lambda x.\lambda y.((deal\ with'\ y)\ x)$

forçou (a) : $((n \mathbin{\backslash\!\!\backslash} s)/_W n)/(s_{to}\|n) : \lambda x.\lambda y.\lambda z.(((force'\ (x\ y))\ y)\ z)$

intervenire : $s_{to} \mathbin{/\!\!/} n : \lambda x.(intervene'\ x)$

lui : $n\|n : \lambda x.x$

prometió (a) : $((n \mathbin{\backslash\!\!\backslash} s)/(s_{to}\|n))/n : \lambda x.\lambda y.\lambda z.(((promise'\ (x\ z))\ y)\ z)$

sair(em) : $n \mathbin{\backslash\!\!\backslash} s_{to} : \lambda x.(leave'\ x)$

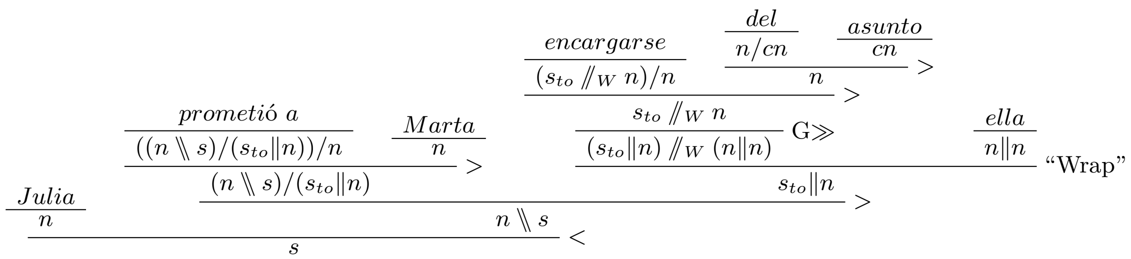


Figure 8.8 – Derivation for: *Julia prometió a Marta encargarse ella del asunto* en $\mathbf{J}_{G^4Z}$

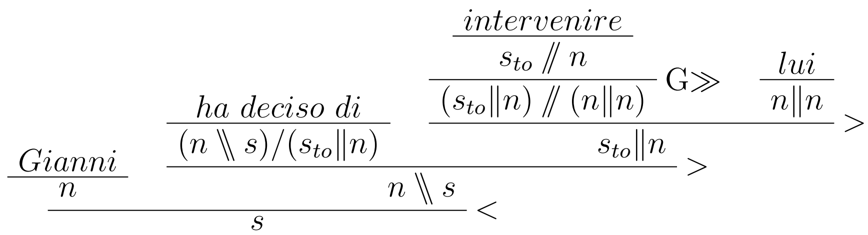


Figure 8.9 – Derivation for: *Gianni ha deciso di intervenire lui* in $\mathbf{J}_{G^4Z}$

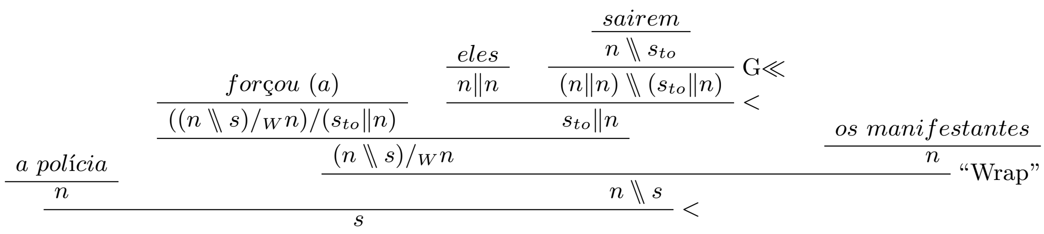


Figure 8.10 – Derivation for: *A polícia forçou os manifestantes a eles saírem* in $\mathbf{J}_{G^4Z}$

As the reader should have observed our previous lexicon is too simplistic. Firstly, it should be noted that in spoken BP, unlike in European Portuguese, the accusative 3SG/PL pronoun does not commonly take a clitic form *(l)o/(l)a*; instead the nominative form *ele(s)/a(s)* is typically used for the former.[152] Consequently, nominative and accusative 3SG/PL pronouns of BP cannot be distinguished in the lexicon. We should therefore have two different entries for a pronoun of this kind. Secondly, in our previous lexicon complementizers and prepositions were not assigned a lexical entry of their own; rather they were attached to the lexical entries of the verbs that select them. We have assumed, for example, that *forçar a* is a complex verbal constituent that selects the infinitive clause, instead of assuming that it is the preposition *a* which performs this selection. Nevertheless, we may split that lexical entry: we can assign the identity function — over functions — as the meaning of the preposition *a* (i.e. $\lambda x.x$, where x is of type $< e, t >$). Then, though the preposition selects the infinitive clause, the control relation is still ensured by the semantics of the control verb. In a similar fashion, we can separate the Italian verb *decide* from its complementizer *di*. Indeed, it is this complementizer which selects the infinitive complement. By assuming the basic categories *pp* (for prepositional phrases) and *cp* (for complementizers), we can now give a more fine-grained lexicon:

a : $pp/(s_{to} \| n) : \lambda x.x$

a : $pp/n : \lambda x.x$

decide : $(n \setminus\!\!\setminus s)/cp : \lambda x.\lambda y.((decide'\ (x\ y))\ y)$

di : $cp/(s_{to} \| n) : \lambda x.x$

ele : $n \| n : \lambda x.x$

ele : $n | n : \lambda x.x$

forçou : $((n \setminus\!\!\setminus s)/_W n)/pp : \lambda x.\lambda y.\lambda z.(((force'\ (x\ y))\ y)\ z)$

prometió : $((n \setminus\!\!\setminus s)/(s_{to} \| n))/pp : \lambda x.\lambda y.\lambda z.(((promise'\ (x\ z))\ y)\ z)$

As the reader can check, the previous derivations may still be obtained in $\mathbf{J}_{G^4Z}$ even when the lexicon is refined in this way.

8.2.2.1 Final Remarks

In this section we presented a combinatory proposal to treat overt anaphoric pronouns following Jacobson's and Steedman's combinatory approaches to pronoun and control

[152] The following Portuguese sentences display this contrast:

i. Visitei-o ontem. (EP)

ii. Visitei ele ontem. (BP)

'[I] visited him yesterday.'

verbs, respectively. We settled on the final version of our account once we made sure that it preserved the following two key features of our linguistic phenomenon. First, a subject pronoun within the infinitive clause, if there is one, has to be free inside the latter. Second, the semantic subject of the infinitive clause, whether or not it has a syntactic surface representation, is semantically controlled by one of the nominal arguments of the main verb. In order to ensure the semantic relation of control, we encoded the **Z** operation in the lexical semantics of control verbs rather than in their derivational semantics. This is entirely analogous to Steedman's second proposal. In order to ensure that the controlled infinitive subject may only surface in a pronominal form, we split up Jacobson's type-constructor as well as the corresponding **G** rules. Thus, the new **G** rules of $\mathbf{J}_{G^4Z}$, like those of $\mathbf{J}_{GZ}$, echo the Geach rules even for the extended logical language. In this respect, our proposal closely follows Jacobson's "bridge" approach.

Once we extend the language of the basic logical system **AB**, we may opt to cross our metaphorical bridge and take a Type-Logical route. As we prepare to venture across, however, the reader might be wondering what happened to the **Z** rules of $\mathbf{J}_{GZ}$.

In fact, once we encoded the multiple-binding operation performed by **Z** in the lexical entries for control verbs, no further mention was made of this rule. Nevertheless, it might continue to be operative in our modified system $\mathbf{J}_{G^4Z}$. And, given that subject and object pronouns may be bound, we certainly need a **Z**-like rule for them. However if **Z** were allowed to apply to a type $(n \setminus\!\!\setminus s)/\ n$ in $\mathbf{J}_{G^4Z}$, the anaphoric object type $n|n$ could end up having a reflexive meaning, as it did in $\mathbf{J}_{GZ}$. Certainly, **Z** does not satisfy the anti-locality condition for pronouns. Unfortunately, we do not have a combinatory proposal to offer that is capable of encoding syntactic domains for binding. We continue our analysis by examining this same problem from the other side of the categorial route.

8.3 A Type-Logical Proposal

8.3.1 Modifying **LLC**: **LLC**^{2}

As we saw in Sec. 7.3.2, the rule $|L$ of **LLC** yields a multiple-binding reading for pronouns. This rule, like **Z** of $\mathbf{J}_{GZ}$, does not meet the anti-locality condition (or: Principle B) for pronouns, and so it overgenerates. This is because the account is intended to cover cases of anaphoric expressions beyond merely pronouns. In order to overcome this problem, we present a first provisional modified version of $|L$, which we label $|L_1$ (see Fig. 8.11 and compare it with Fig. 7.22). As the reader can check, the main difference between $|L$ and $|L_1$ is that we require the occurrence of a type A in the left premise of the latter. The two rules are nevertheless still very much alike. Indeed, given that neither Weakening nor even Expansion are admitted in **LLC**, the sequence Y in the left premise of $|L$ certainly contains a type A in *some previous step* of the derivation (i.e. $Y[x : A]$). Thus, in $|L_1$ we replace Y by the sequence $Y_1, x : A, Y_2$, where both Y_1 and Y_2 may be the empty sequence ϵ. In this latter case, the left premise of $|L_1$ is an instance of the Identity rule, so the rule can be simplified as per Fig. 8.12.

$$\frac{Y_1, x : A, Y_2 \Rightarrow M : A \qquad X, y : A, Z, w : B, W \Rightarrow N : C}{X, Y_1, x : A, Y_2, Z, z : B|A, W \Rightarrow N[M/y](zM)/w] : C}\ |\mathrm{L}_1$$

Figure 8.11 – (Alternative) Left rule of **LLC**

$$\frac{X, y : A, Z, w : B, W \Rightarrow N : C}{X, y : A, Z, z : B|A, W \Rightarrow N[(zy)/w] : C}\ |\mathrm{L}$$

Figure 8.12 – Simplified left rule of **LLC**

It is known that pronouns in an argument position cannot be bound by a co-argument binder. Thus, for example, when a pronoun stands in the object argument position it cannot be bound by the subject of the clause; an object pronoun can, however, be bound by a nominal inside the subject clause. The reverse is true of reflexive pronouns.

(236) [John$_1$'s father]$_2$ admires him$_{1/*2}$/himself$_{2/*1}$.

(237) [The father of John$_1$]$_2$ admires him$_{1/*2}$/himself$_{2/*1}$.

The lack of Weakening and Expansion rules in **LLC** does not only guarantee that the antecedent sequence Y of the left premise of |L contains a type A in some step of the derivation, but also that the latter has to be an argument of some functional type in the sequence Y, if Y is non-empty. From this observation, we introduce three modifications in the rule $|L_1$ in order to formulate a more accurate rule for bound pronouns (see Fig. 8.13). Though we are not presently working with reflexive pronouns, we also suggest a preliminary rule for (subject-oriented) ones (see Fig. 8.14). As this modified system contains two rules — and so two connectives, $|_p$ and $|_r$ — each encoding a form of Contraction and of Cut, we name the system **LLC**^{2}.

As a first modification, we impose a side condition on the sequences Y_1 and Y_2 in the rule for bound pronouns: $Y_1 \neq \epsilon$ or $Y_2 \neq \epsilon$ in $|_pL$. As a result, $|_pL$ cannot be applied, but $|_rL$ can be, if the left premise of this rule is an instance of the Identity rule.[153] This condition captures the fact that the antecedent of the bound pronoun, but not that of the reflexive, stands in an embedded position. Second, we drop the sequence X in the right premise of $|L$; thus the type A in the right premise of both $|_pL$ and $|_rL$ must be

[153]Of course, these restrictions are too strong; they block the bound reading for a pronoun in embedded sentences as well as in prepositional complements. They also block this reading in cases where both $Y_1 \neq \epsilon, Y_2 \neq \epsilon$, and $Y_1, Y_2 \Rightarrow A$ but A only occurs as a subtype in Y_1, Y_2. By using structural modalities, in [Cor17a] we show how the empty restriction on Y_1, Y_2 could be relaxed.

$$\frac{Y_1, x : A, Y_2 \Rightarrow M : A \qquad y : A, Z, w : B, W \Rightarrow N : C}{Y_1, x : A, Y_2, Z, z : B|A, W \Rightarrow N[M/y][(zx)/w] : C} |_p\text{L}$$

where $Y_1 \neq \epsilon$ or $Y_2 \neq \epsilon$

Figure 8.13 – Left rule for bound pronouns of **LLC**^{2}

$$\frac{Y_1, x : A, Y_2 \Rightarrow M : A \qquad y : A, Z, w : B, W \Rightarrow N : C}{Y_1, x : A, Y_2, Z, z : B|A, W \Rightarrow N[M/y](zM)/w] : C} |_r\text{L}$$

Figure 8.14 – Left rule for (subject-oriented) reflexives of **LLC**^{2}

peripheral. So, if Y_1 and Y_2 are both empty, the antecedent A of the reflexive can only be an argument in the sequence Z. Third, since A must be an argument of some functional type in Y_1 or in Y_2 we take the meaning of *this* A as the antecedent of the bound pronoun. For reflexives, this will also be the case when Y_1 and Y_2 are both empty. Thus, what distinguishes $|_r L$ and $|_p L$ of **LLC**^{2} is the meaning of the type A to which the function $z : B|A$ applies: respectively, (zM) in $|_r L$ and (zx) in $|_p L$.

As an example of $|_r L$ and $|_p L$, compare the following schematic derivations (Figs 8.15–8.16) for the sentences in (236–237) above. For simplicity, consider the following approximated lexical entry for the preposition:

of : $(n\backslash n)/n : \lambda x.\lambda y.((of'\ y)\ x)$

$$\frac{\begin{array}{c}\vdots\\ v : n/cn, u : cn, w : (n\backslash n)/n, t : n \Rightarrow M : n\end{array} \qquad \begin{array}{c}\vdots\\ x : n, y : (n\backslash s)/n, z : n \Rightarrow N : s\end{array}}{v : n/cn, u : cn, w : (n\backslash n)/n, t : n, y : (n\backslash s)/n, r : n|_p n \Rightarrow N[M/x][(rt)/z] : s} |_p\text{L}$$

Figure 8.15 – (Schematic) Derivation for: *The father of John*$_1$ *admires him*$_1$ in **LLC**^{2}

$$\frac{\begin{array}{c}\vdots\\ v : n/cn, u : cn, w : (n\backslash n)/n, t : n \Rightarrow M : n\end{array} \qquad \begin{array}{c}\vdots\\ x : n, y : (n\backslash s)/n, z : n \Rightarrow N : s\end{array}}{v : n/cn, u : cn, w : (n\backslash n)/n, t : n, y : (n\backslash s)/n, r : n|_r n \Rightarrow N[M/x][(rM)/z] : s} |_r\text{L}$$

Figure 8.16 – (Schematic) Derivation for: *[The father of John]*$_1$ *admires himself*$_1$ in **LLC**^{2}

The first proof delivers the semantics $((admire'\ j')\ ((of'\iota\ father')\ j'))$. This is because the meaning of the pronoun (i.e. $\lambda x.x$) takes the meaning of an embedded nominal (i.e. j') as it argument. The second proof, in which the meaning of the reflexive (i.e. $\lambda x.x$) applies to the meaning of the entire nominal phrase (i.e. $((of'\iota\ father')\ john')$, delivers the semantics $((admire'\ ((of'\iota\ father')\ j'))\ ((of'\iota\ father')'\ j'))$.

$$\dfrac{\dfrac{\begin{array}{c}\vdots\\ x : n, y : (n\backslash s)/n, z : n \Rightarrow y(z(x)) : s\end{array}}{x : n, y : (n\backslash s)/n, w : n|n \Rightarrow y(z(x))[(wx)/z] : s}\,|_r\text{L}\ {}^*|_p\text{L}}{u : n|n, y : (n\backslash s)/n, w : n|n \Rightarrow \lambda v.(y((wx)(x)))[(uv)/x] : s|n}\,|\text{R}$$

Figure 8.17 – (Illicit) Derivation for: *He*$_1$ *admires himself*$_1$*/*(him*$_1$*)* in $\mathbf{LLC}^2$

Finally, note that from our $|_pL$ and $|_rL$ rules, and with the help of $|R$ from **LLC**, we can block the proof for **He*$_1$ *admires him*$_1$ but, at the same time, admit the proof for *He admires himself*, with type $s|n$ and semantics $\lambda v.((admire'\ v)\ v)$. Hereinafter, we include $|R$ as a rule for free pronouns in $\mathbf{LLC}^2$. Once again, however, we will be duplicating this rule.

Jaeger's **LLC** system, like Jacobson's $\mathbf{J}_{GZ}$, contains a single anaphoric type-constructor: $|$. Thus, as in our combinatory proposal based on $\mathbf{J}_{GZ}$, here we propose to extend **LLC** by adding the new type-constructor $\|_p$ for proforms. The language of $\mathbf{LLC}^2$ thus contains three pronominal type-constructors: $|_r$, $|_p$ and $\|_p$. The right and left rules for $\|_p$ are the same as those for $|_p$. With this new syntactic type $B\|_pA$ at hand, here too we can differentiate between an expression containing an object pronoun $B|_pn$ and an expression containing a subject pronoun $B\|_pn$. As we discussed in Sec. 7.3.2, the **G** rules of $\mathbf{J}_{GZ}$ are all derivable in **LLC**. Analogously, all of the **G** rules of $\mathbf{J}_{G^4Z}$ are derivable in $\mathbf{LLC}^2$. Consequently, the previous combinatory derivations that use different versions of **G** (and other combinators derivable in **L**) can also be obtained in $\mathbf{LLC}^2$. At the same time, there is an important difference between our combinatory and type-logical treatments due to the inherent features of a type-logical system. In such systems, each logical rule only contains a single type-constructor. In particular, given that $|_pR$ and $\|_pR$ do not include connectives other than $|_p$ and $\|_p$, we can do away with the subject and object slash type-constructors. Since we shall not be working with reflexive pronouns $n|_rn$, we drop the subscript on both $|_p$ and $\|_p$ type-constructors.

8.3.2 Anaphoric Control in $\mathbf{LLC}^2$

In our previous combinatory proposal, we assigned two entries to each control verb (or to the particles selected by them), in order to cover both of their possible complements in Romance languages: namely, $n\backslash s_{to}$ and $s_{to}\|n$. In a Type-Logical Grammar, these two entries can be unified into one with the help of the semantically inactive disjunction type-constructor $\sqcup$ [Mor94b].

Definition 8.1 (Syntactic types of $\mathbf{LLC}^2$)**.** Where P is a set of basic types, the *set F of syntactic types of* $\boldsymbol{LLC}^2$ is defined as follows:

$$F ::= \ P \mid F\backslash F \mid F/F \mid F \bullet F \mid F|_r F \mid F|_p F \mid F\|_p F \mid F \sqcup F$$

Definition 8.2 (Semantic types of $\mathbf{LLC}^2$)**.** Where δ is a set of basic semantic types, the *set γ of semantic types of $\boldsymbol{LLC}^2$* is defined as follows:

$$\gamma ::= \delta \mid \gamma \rightarrow \gamma \mid \gamma\&\gamma \mid \gamma + \gamma$$

Definition 8.3 (Type domains)**.** The *type domain $Dom(\gamma)$* for each semantic type γ is defined on the basis of an assignment d from non-empty sets (i.e. the basic type domains) to basic semantic categories δ as follows:

1. $Dom(\gamma) = d(\gamma)$, for $\gamma \in \delta$;
2. $Dom(\alpha \rightarrow \beta) = Dom(\beta)^{Dom(\alpha)}$;
3. $Dom(\alpha\&\beta) = Dom(\alpha) \times Dom(\beta)$;

Definition 8.4 (Semantic type map)**.** The *semantic type map for $\boldsymbol{LLC}^2$* is a mapping τ from syntactic types F to semantic types δ such that:

1. $\tau(A \bullet B) = \tau(A)\ \&\ \tau(B)$
2. $\tau(A\backslash B) = \tau(B/A) = \tau(B|_{p/r}A) = \tau(B\|_p A) = \tau(A) \rightarrow \tau(B)$
3. $\tau(A \sqcup B) = \tau(A) = \tau(B)$

$$\frac{\Gamma_1, x : A, \Gamma_2 \Rightarrow M : C \qquad \Gamma_1, x : B, \Gamma_2 \Rightarrow M : C}{\Gamma_1, x : A \sqcup B, \Gamma_2 \Rightarrow M : C} \sqcup L$$

$$\frac{\Gamma \Rightarrow M : A}{\Gamma \Rightarrow M : A \sqcup B} \sqcup R_1 \qquad\qquad \frac{\Gamma \Rightarrow N : B}{\Gamma \Rightarrow N : A \sqcup B} \sqcup R_2$$

Figure 8.18 – Rules for the semantically inactive disjunction type-constructor $\sqcup$

The optionality of the overt controlled pronoun within an infinitive clause can now be captured using such a disjunction type-constructor. We can account for prototypical and non-prototypical control sentences simply by assigning a semantically inactive disjunction type to the complement argument of the control verb or to the prepositions and complementizers selected by them. As an example, consider the following lexical entries and the following simple derivation in Fig. 8.19:

quer/queria : $(n\backslash s)/((n\backslash s_{to}) \sqcup (s_{to}\|n))$

di : $cp/((n\backslash s_{to}) \sqcup (s_{to}\|n))$

a : $pp/((n\backslash s_{to}) \sqcup (s_{to}\|n))$

prometió : $((n\backslash s)/((n\backslash s_{to}) \sqcup (s_{to}\|n)))/pp$

$$
\dfrac{\dfrac{\dfrac{\begin{matrix}\vdots\\ n, (n\backslash s_{to})/n, n/cn, cn \Rightarrow s_{to}\end{matrix}}{n\|n, (n\backslash s_{to})/n, n/cn, cn \Rightarrow s_{to}\|n}\,\|\mathrm{R}}{n\|n, (n\backslash s_{to})/n, n/cn, cn \Rightarrow (n\backslash s_{to}) \sqcup (s_{to}\|n)}\,\sqcup\mathrm{R}_2 \qquad \begin{matrix}\vdots\\ n/cn, cn, n\backslash s \Rightarrow s\end{matrix}}{n/cn, cn, (n\backslash s)/((n\backslash s_{to}) \sqcup (s_{to}\|n)), n\|n, (n\backslash s_{to})/n, n/cn, cn \Rightarrow s}\,/\mathrm{L}
$$

Figure 8.19 – Derivation for: *Os meninos querem eles embrulhar(em) os presentes* in **LLC**^{2}

When working with control structures from a combinatory perspective, we saw that connectives other than those for subject and object pronouns are needed to obtain the correct word-order in infinitive clauses from Romance languages. Recall also that while Brazilian Portuguese displays the SVO word-order, Italian and Spanish display the VSO order in infinitive contexts; whereas European Portuguese allows both. Thus, also from a type-logical perspective we need either a long-distance or a discontinuous type-constructor for the lexical assignments to the infinitive verbs (as well as for object-control verbs, for independent reasons). This implies extending the vocabulary of **LLC**^{2}.[154] To see this, consider the following alternative entries for a transitive infinitive verb, where $\downarrow$ is the infix operator of Moortgat [Moo88], $\upharpoonright$ is the (right) long-distance (or: far-right) operator of Bach [Bac84]:270–273, and the nondeterministic division type $C \div A =_{def} (A\backslash C) \bullet (C/A)$. In Fig. 8.20 below we recall the corresponding left rules for the former two connectives.[155]

fazer(em) : $(s_{to}/n) \downarrow n : fazer'$

resolver : $(s_{to}/n) \upharpoonright n : resolver'$

resolver : $(s_{to} \div n) \upharpoonright n : resolver'$

[154] Alternatively, we could assign a continuous or a discontinuous lifted type to pronouns. Thus, for example, we could have assigned the discontinuous type $(s_{to}\|n) \downarrow (s_{to} \uparrow n)$ to cover both the pre- or the post-verbal position for a subject pronoun within an infinitive clause. Nevertheless, the available positions for a subject in each of these Romance languages depend on the fact that they are selected by a verb in an infinitive form, rather than on the subject itself. In fact, in finite clauses a subject *normally* occupies the pre-verbal position in all of these languages. Then, if we were to opt for assigning some lifted type to a subject pronoun, we would duplicate the lexical entries for the latter.

[155] In Bach's [Bac84]:273 notation: $< a//b, \varphi, b > \Rightarrow < a, \varphi >$, for $a \neq b$ and where φ contains no instance of b. To encode this last condition in our formulation of the rule, we should preclude the sequence W from containing any instance of the type A.

The long-distance type $(s_{to}/n) \uparrow n$ is intended to express the fact that the nominal object argument may stand in a non-adjacent position with respect to the verb. The infix type $(s_{to}/n) \downarrow n$ in turn expresses the fact that the nominal subject may stand in a pre- or in a post-verbal position. The long-distance, nondeterministic type $(s_{to} \div n) \uparrow n$ covers both possibilities. The last two types license the SVO and the VSO word-order; they appear to be adequate to derive infinitive sentences containing a pre- or a post-verbal subject. Consider, for example, the EP infinitive phrases *eles fazerem um trabalho*, which exhibits the SVO order, and *resolver ele o problema*, which displays the VSO word-order. In figure 8.21 we show that both strings may be of type $s_{to}\|n$, which is an admissible argument of a control verb according to our lexicon.

$$\frac{X \Rightarrow M : A \qquad Y, x : B, W \Rightarrow N : C}{Y, y : B \uparrow A, W, X \Rightarrow N[(yM)/x] : C} \uparrow \mathrm{L} \qquad \frac{X, Z \Rightarrow M : A \qquad Y, x : B, W \Rightarrow N : C}{Y, X, y : B \downarrow A, Z, W \Rightarrow N[(yM)/x] : C} \downarrow \mathrm{L}$$

Figure 8.20 – Left rules for long-distance $\uparrow$ and discontinuous $\downarrow$ type-constructors

$$\frac{\dfrac{\vdots}{n/cn, cn \Rightarrow n} \qquad \dfrac{\dfrac{\vdots}{s_{to}/n, n \Rightarrow s_{to}}}{s_{to}/n, n\|n \Rightarrow s_{to}\|n} \|\mathrm{R}}{(s_{to}/n) \uparrow n, n\|n, n/cn, cn \Rightarrow s_{to}\|n} \uparrow\mathrm{L} \qquad \frac{\dfrac{n \Rightarrow n \qquad \dfrac{\vdots}{s_{to}/n, n/cn, cn \Rightarrow s_{to}}}{n, (s_{to}/n) \downarrow n, n/cn, cn \Rightarrow s_{to}} \downarrow\mathrm{L}}{n\|n, (s_{to}/n) \downarrow n, n/cn, cn \Rightarrow s_{to}\|n} \|\mathrm{R}$$

Figure 8.21 – Derivations for the VSO and SVO word-order in **LLC**^{2}

The types assigned to transitive infinitive verbs also allow us to derive the VOS word-order (see Fig. 8.22). Briefly, this is because none of them expresses the syntactic function of their two nominal arguments. The VOS order, however, is ungrammatical in controlled infinitive clauses in (some) Romance languages (cf. [Car99]; [Her12]).[156] In these infinitive contexts, the controlled subject must be adjacent to the verb despite the fact that it can stand either on the right (in IT, EP and ES) or on the left (in BP and EP) of the infinitive verb.

(238) Il Rettore ha deciso [di aprire (**lui**) il convegno (***lui**)].
the Dean has decided COMP open he.NOM the conference he.NOM
'The Dean has decided to open the conference.' [Car99]:79

(239) Le prometí a Juan [escribir **yo** un libro (?/***yo**)]
CL promised.1SG to John write.INF I.NOM a book (I.NOM)
'I promised John to write a book.' [Her12]

Thus, we need a mechanism to block a post-object infinitive pronominal subject, but still grant the post-verbal position for the pronoun in EP, IT and ES.

[156] We disagree with Herbeck about his grammaticality judgements of the word-order in Spanish.

$$\dfrac{\dfrac{n \Rightarrow n \qquad \overset{\vdots}{s_{to}/n, n/cn, cn \Rightarrow s_{to}}}{(s_{to}/n) \uparrow n, n/cn, cn, n \Rightarrow s_{to}} \uparrow L}{(s_{to}/n) \uparrow n, n/cn, cn, n\|n \Rightarrow s_{to}\|n} \|R \qquad \dfrac{\overset{\vdots}{n/cn, cn \Rightarrow n} \qquad \dfrac{\overset{\vdots}{s_{to}/n, n \Rightarrow s_{to}}}{s_{to}/n, n\|n \Rightarrow s_{to}\|n} \|R}{(s_{to}/n) \downarrow n, n/cn, cn, n\|n \Rightarrow s_{to}\|n} \downarrow L$$

Figure 8.22 – Derivation for the VOS word-order in $\mathbf{LLC}^2$

To this end, we extend the $\mathbf{LLC}^2$ system once again. Following [Mor92] and [Moo96b], we add the bracket (or existential) $\langle\ \rangle$ (though we change Morrill's prefixed notation) and the (anti)bracket (or universal) $\lceil\ \rceil$ modalities ($[\ ]^{-1}$ in Morrill's [Mor92] notation and $\Box\downarrow$ in Moortgat's [Moo96b]). Figure 8.23 displays the rules for these modalities.

Definition 8.5 (Syntactic types of $\mathbf{LLC}^{\langle 2\rangle}$)**.** Where P is a set of basic types, the *set F of types of $\boldsymbol{LLC}^{\langle 2\rangle}$* is defined as follows:[157]

$$F ::= \ P \mid F\backslash F \mid F/F \mid F \bullet F \mid F|_r F \mid F|_p F \mid F\|_p F \mid F \sqcup F \mid \langle F\rangle \mid \lceil F\rceil$$

A partially ordered bracket semigroup (p.o. b-semigroup) is a structure $(L, +, b; \leq)$ of arity $(2, 1; 2)$ such that $(L, +)$ is a semigroup (i.e. the operation $+$ is associative in L); $(L; \leq)$ is a partial order and $\leq$ is compatible with $+$ and b, i.e.:

$s_1 + s_3 \leq s_2 + s_4$ if $s_1 \leq s_2$ and $s_3 \leq s_4$

$b(s_1) \leq b(s_2)$ if $s_1 \leq s_2$

The syntactic (or: prosodic) interpretation $[\![.]\!]$ of bracket types $\lceil A\rceil$ and $\langle A\rangle$ is given in a p.o. b-semigroup as follows (cf. [Mor11]:85):[158]

$[\![\langle A\rangle]\!] = \{s_1 | \exists s_2 \in [\![A]\!], s_1 \leq b(s_2)\}$

$[\![\lceil A\rceil]\!] = \{s_1 | \forall s_2 \leq b(s_1), s_2 \in [\![A]\!]\}$

For the semantic interpretation of the new types, note that $\tau(\lceil A\rceil) = \tau(\langle A\rangle) = \tau(A)$.

These bracket modalities have been applied in Type-Logical Grammar to delimit syntactic domains and, in particular, to create islandhoods. They have been used, for example, to inhibit extraction from island domains (cf. [Hep90]; [Mor92], [Mor94a], [Mor94b]; [MVF11]). Like [Mor94a], we use $\langle\ \rangle$ to mark the subject argument position of a verb in

[157] We also assume a long-distance $\uparrow$ or a discontinuous type-constructor $\uparrow$. But this choice is orthogonal to our proposal. Moreover, a sequent of $\mathbf{LLC}^{\langle 2\rangle}$ may be of form $[A_1, \ldots, A_n] \Rightarrow C$, where $n \geq 1$. Accordingly, also the definition of *validity of sequents* has to be extended to include antecedent configurations of form $[\Gamma]$.

[158] For an interpretation in terms of frame-based models see [Moo02], [Moo11]:134.

$$\frac{X, [A], Y \Rightarrow C}{X, \langle A \rangle, Y \Rightarrow C}\ \langle\ \rangle\mathrm{L} \qquad\qquad \frac{X \Rightarrow A}{[X] \Rightarrow \langle A \rangle}\ \langle\ \rangle\mathrm{R}$$

$$\frac{X, A, Y \Rightarrow C}{X, [\lceil A \rceil], Y \Rightarrow C}\ \lceil\ \rceil\mathrm{L} \qquad\qquad \frac{[X] \Rightarrow A}{X \Rightarrow \lceil A \rceil}\ \lceil\ \rceil\mathrm{R}$$

Figure 8.23 – Rules for bracket modalities $\langle\ \rangle$ and $\lceil\ \rceil$

its lexical entry. Nevertheless, unlike [Mor94a], we only mark the infinitive subject position. With these new type-constructors we may first of all block the ungrammatical VOS word-order. Secondly, we can ensure that the controlled pronoun in control clauses does not occur in an embedded context.

Word-Order

Consider the following sample of bracket lexicon. The argument bracketing $\langle\ \rangle$ in the verb-type will be responsible for blocking the VOS word-order (see Fig. 8.24). In fact, by marking the subject argument it will be possible to distinguish internal and external verb arguments despite the fact that they both stand in a post-verbal position. Thus, although the associative system cannot express hierarchical relationships, we can partially recover non-associativity by means of structured antecedent sequents.

ele/lui : $n\|n : \lambda x.x$

fazer(em) : $(s_{to}/n) \downarrow \langle n \rangle : fazer'$

aprire/escribir : $(s_{to}/\langle n \rangle) \nearrow n : aprire'$

resolver : $(s_{to}/n) \div \langle n \rangle : resolver'$

$$\dfrac{\dfrac{\dfrac{n \Rightarrow n}{[n] \Rightarrow n}\,{*\langle\ \rangle\mathrm{R}} \quad \dfrac{\dfrac{n \Rightarrow n}{n \Rightarrow \langle n \rangle}\,{*\langle\ \rangle\mathrm{R}} \quad s_{to} \Rightarrow s_{to}}{s_{to}/\langle n \rangle, n \Rightarrow s_{to}}\,/\mathrm{L}}{(s_{to}/\langle n \rangle) \nearrow n, n, [n] \Rightarrow s_{to}}\,{\nearrow}\mathrm{L}}{(s_{to}/\langle n \rangle) \nearrow n, n, [n\|n] \Rightarrow s_{to}\|n}\,\|\mathrm{R} \qquad \dfrac{\dfrac{\dfrac{n \Rightarrow n}{n \Rightarrow \langle n \rangle}\,{*\langle\ \rangle\mathrm{R}} \quad \dfrac{\dfrac{n \Rightarrow n}{[n] \Rightarrow n}\,{*\langle\ \rangle\mathrm{R}} \quad s_{to} \Rightarrow s_{to}}{s_{to}/n, [n] \Rightarrow s_{to}}\,/\mathrm{L}}{(s_{to}/n) \downarrow \langle n \rangle, n, [n] \Rightarrow s_{to}}\,{\downarrow}\mathrm{L}}{(s_{to}/n) \downarrow \langle n \rangle, n, [n\|n] \Rightarrow s_{to}\|n}\,\|\mathrm{R}$$

Figure 8.24 – Illicit derivations for the VOS word-order in $\mathbf{LLC}^{\langle 2 \rangle}$

Whereas the first and third verb-types each allows the VSO as well as the SVO word-order, the second verb-type only allows the VSO word-order. However, all three block the VOS word-order.[159] Thus the only relevant adequacy criterion for choosing one

[159] Thus, to recover both VSO and VOS word-orders, we may assign the type $s_{to}/(\langle n \rangle \circ n)$, where $\circ$ is the non-deterministic (or: commutative) semantically inactive product type-constructor, i.e. $(A \circ B) \Leftrightarrow (B \circ A)$ and $\tau(A \circ B) = \tau(A) = \tau(B)$.

over the other will be which (Romance) language happens to be in play. In Appendix 9 we shall use the third verb-type for EP.

Pronouns and embedded pronouns

In Jaeger's and in Jacobson's proposals a constituent containing a free pronoun is of the same type, and denotes the same function, as the pronoun itself. As Jacobson puts it [Jac03]:60, "any expression containing a pronoun which is unbound within that expression (*including a pronoun itself*) is actually a function from individuals to something else." Therefore, in each of Jaeger's, Jacobson's and our proposals, any nominal phrase containing a free pronoun will be a function from individuals to individuals — just like a pronoun. Thus, for example, the phrase *the fact/claim that he wins*, which contains a free pronoun, and the pronoun *he* itself, will both be (lexically or derivationally) assigned the same functional type $n|n$ (or: n^n in Jacobson's notation).

Now, at first sight it seems appropriate to assume that both expressions denote a function from individuals to individuals. It also seems reasonable to assume that several syntactic contexts in which a free pronoun occurs, can also host a nominal phrase containing a free pronoun. But this is not always so. In particular, this is not the case in control structures. Consider, for example, the following recalcitrant example (due to Glyn Morrill). Here, a subject pronoun certainly occurs in the subject position of an infinitive verb; however, the pronoun occurs *within* the infinitive nominal phrase that occupies the subject position. Thus, the infinitive subject of the controlled clause is not the pronoun, but the nominal containing the embedded pronoun.

(240) João quer [**o fato de** [**ele** **ganhar**] ser comemorado].
John wants the fact of he.NOM win.INF be.INF celebrated
'John wants the fact that he wins to be celebrated.'

In order to analyze this apparent counterexample to our formal proposal, first note that, unlike Jacobson's and Jaeger's systems, in $\mathbf{LLC}^{\langle 2\rangle}$ we have more syntactic (and so prosodic) resources to distinguish between a constituent containing a subject pronoun and the subject pronoun itself.

$$\dfrac{\dfrac{\begin{matrix}\vdots\\ [n], \langle n\rangle\backslash s_{to} \Rightarrow s_{to}\end{matrix} \qquad \begin{matrix}\vdots\\ n/cn, cn, n\backslash n \Rightarrow n\end{matrix}}{n/cn, cn, (n\backslash n)/s_{to}, [n], \langle n\rangle\backslash s_{to} \Rightarrow n}\,/L}{n/cn, cn, (n\backslash n)/s_{to}, [n\|n], \langle n\rangle\backslash s_{to} \Rightarrow n\|n}\,\|R$$

Figure 8.25 – Derivation for an infinitive nominal in $\mathbf{LLC}^{\langle 2\rangle}$

Despite the fact that a subject pronoun and a nominal containing a subject infinitive pronoun are both of type $n\|n$ (see Fig. 8.25), the latter cannot be taken as a subject of another infinitive verb. Indeed, compare (the last step of) the derivation in Fig. 8.25 with (the second step of) that given in Fig. 8.26.

$$\dfrac{n \Rightarrow n \qquad \dfrac{\dfrac{\dfrac{\vdots}{n/cn, cn, (n\backslash n)/s_{to}, [n], \langle n\rangle\backslash s_{to} \Rightarrow n}}{[n/cn, cn, (n\backslash n)/s_{to}, [n], \langle n\rangle\backslash s_{to}] \Rightarrow \langle n\rangle}\langle\,\rangle R \qquad s_{to} \Rightarrow s_{to}}{[n/cn, cn, (n\backslash n)/s_{to}, [n], \langle n\rangle\backslash s_{to}], \langle n\rangle\backslash s_{to} \Rightarrow s_{to}}\backslash L}{\dfrac{[n/cn, cn, (n\backslash n)/s_{to}, [n\|n], \langle n\rangle\backslash s_{to}], (\langle n\rangle\backslash s_{to})/n, n \Rightarrow s_{to}\|n}{[n/cn, cn, (n\backslash n)/s_{to}, [n\|n], \langle n\rangle\backslash s_{to}], (\langle n\rangle\backslash s_{to})/n, n \Rightarrow s_{to}\|n}\|^{*}R}/L$$

Figure 8.26 – Illicit derivation for pronouns in embedded infinitive domains in $\mathbf{LLC}^{\langle 2\rangle}$

Note that whereas the nominal phrase *o fato de ele ganhar* may receive the type $n\|n$, the constituent *o fato de ele ganhar ser comemorado* cannot be assigned the type $s_{to}\|n$.[160] This is because the rule $\|$R cannot be applied inside the bracketing antecedent structure triggered by the second infinitive verb-type. In order to apply this rule we should first delete the antecedent bracket configuration (by $\lceil\ \rceil$R). But this is not possible, as some types find themselves outside the scope of the structural operator [ ]. As a result, that linguistic constituent cannot be taken as a complement of the control verb *querer*. Therefore, the sentence in (240) above is not recognized by the $\mathbf{LLC}^{\langle 2\rangle}$ grammar, as desired. In addition, since we do not mark the subject position of a finite verb, the grammatical finite version of the previous sentence (i.e. *o fato de ele ganhar alegra Maria*) can still be recognized by our system.[161]

To finally close our formal proposal we give an overview of the principal characteristics of the lexicon we offer in Appendix 9. There our formal proposal is implemented in the parser/theorem-prover CatLog2; due to this fact lexical entries are labelled with inflectional and (possibly quantifier) personal features. In the lexicon, Sf and Si stand for finite and infinitive sentences, respectively; $Nt(s(m))$, for example, stands for a third-person singular male nominal, while the quantifier type $\exists aNa$ leaves these features unspecified. The rules for quantifiers are the standard ones. Our proofs cover cases of strings containing: subject and object control verbs, two- and three-place object control verbs, transitive and

[160] The reader may be wondering about the lexical assignment for the preposition *de* 'of'; this preposition is usually assumed as being of type $(n\backslash n)/n$. In fact, we have to agree with Chierchia [Chi84a] that there seems to be a nominalized-like meaning behind (at least some kind of) infinitive phrases, after all.

[161] If, following [Mor94a], we were also marking the subject position of finite verbs, we would need some mechanism for licensing structural relaxation. The structural rules above may spread (in a top-bottom direction) bracket domains over the antecedent sequent. Note that these rules resemble the Index Percolation rule for domains { } of **LA** given in [Jäg97]; see footnote [140]. In [Moo96b]:375, [Moo11]:137 similar structural rules, called weak distributivity (also: percolation) $K_{1/2}$, are given. K_1 is required to derive the modal sequent $\Diamond(A \bullet B) \Rightarrow \Diamond A \bullet B$ and K_2, to derive $\Diamond(A \bullet B) \Rightarrow A \bullet \Diamond B$.

$$\frac{X, [Y] \Rightarrow B}{[X, Y] \Rightarrow B}\,[D_1] \Rightarrow \qquad\qquad \frac{[Y], Z \Rightarrow B}{[Y, Z] \Rightarrow B}\,[D_2] \Rightarrow$$

intransitive infinitive verbs, and pre- and post-verbal embedded pronouns in EP, BP, IT and ES. The lexicon we offer brings together the results of our logical inquiry. To sum up, and by way of guiding the reader through Appendix 9:

- We encode the control — multiple-binding — relation in the lexical semantics.
- We use the semantically inactive disjunctive type-constructor to express the fact that the control verb is polymorphic between two infinitive, unsaturated complement-types.
- Subject and object pronouns are assigned the same syntactic types with the same identity function as lexical semantics.
- By assigning a non-directional (or: non-lifted) type for subject pronouns, we let them occur in a pre- or in a post-verbal position.
- However, we encode the prosodic VSO and/or SVO order in the infinitive verb-type.
- So, we give alternative entries for transitive infinitive verbs according to the language.
- Nevertheless, all of these types ensure the ((VO)S) semantic structure of the infinitive clause.
- We also mark the subject position of infinitive verbs with the bracket modality.

So as not to try the reader's patience and endurance any further, we move now to the final chapter of the thesis. There, we bring together the results of our two-part inquiry and draw some forward-looking conclusions.

9 Conclusions

The starting points of this thesis, recall, were the following questions:

Key linguistic question: What is the syntactic status of the surface pronoun in infinitive clauses selected by control verbs?

Key logical question: What are the available mechanisms for reusing semantic resources in a contraction-free logical grammar?

We now sum up our findings in connection with each of these in turn.

In order to answer the linguistic question, in Part I of the thesis we critically examined the theoretical generative assumptions according to which a nominative subject cannot appear in infinitive contexts. Eventually, these assumptions were found to be wanting. For, we showed that infinitive clauses from Romance languages do host different kinds of nominals in their subject position. Adapting Landau's proposal, we distinguished two classes of infinitive clauses: F-infinitives and C-infinitives. We argued that what distinguishes one class from the other are the features on the C and I functional heads. Specifically, we saw that some propositional verbs can select an infinitive complement with either a [–Agr] or a null [Agr] feature; whereas control verbs can *only* select an infinitive complement with no [Agr] feature on the C head. And, we saw that whether the [Agr] features on the C functional head take a null or a negative value is determined by the semantics of the matrix verb.

This in turn led us to acknowledge, *contra* Chierchia, that — insofar as they can have a subject — infinitives may in fact denote something other than individuals. In particular, infinitives selected by propositional verbs seem to denote a propositional function; while those selected by control verbs denote a property (of sorts) that never happens to be saturated.

Against this backdrop, Part II moved to address our key logical question. We now review the main landmarks of that discussion. By way of preamble note that the linguistic structure that is the object of our study — namely, control infinitives with overt nominative pronouns — is not found in the English language. By contrast, the vast majority of existing categorial accounts of control verbs and pronouns focus almost exclusively on English. As a result, a formal analysis of control clauses containing overt subjects is nowhere to be found in the literature. With this thesis we sought to fill this lacuna; we did so by adapting —

both selectively and creatively — existing accounts to the case of Romance languages.

Finally, it is helpful to bear in mind that the focal points of our analysis correspond to the three main components of control structures: pronominal subjects, infinitive complements and control clauses. We now very briefly review each of these in turn.

Our first step was to adopt the thesis, endorsed by Jacobson and Jaeger, according to which the meaning of a personal pronoun is given by the identity function. We thereby agreed with Jacobson that a constituent containing a pronoun that is not bound *within* that constituent denotes a function (from individuals to whatever kind of meaning the constituent would have, if it did not contain the pronoun). Crucially, however, we also departed from these accounts on the syntactic front. For, unlike Jacobson and Jaeger, we drew a distinction between subject and object pronouns, by assigning them two different types.

This strategy allowed us to perform a more fine-grained analysis of the linguistic structures in question. In particular, it allowed us to formally express the fact that a constituent can contain, not merely a free pronoun, but a free *subject* pronoun. Crucially for our purposes, this entails that subject pronouns within infinitive clauses remain syntactically free, as desired. Mirroring this is a semantic result: namely that infinitive clauses denote unsaturated properties, even when their subject position is occupied by an overt pronoun. We will come back to this shortly.

Our pronominal distinction allows us to give a formalization of infinitive clauses that is both syntactically and semantically adequate; as it stands, however, it falls short of guaranteeing word-order adequacy. For, unlike English, Romance languages feature semi-free word-orders. In particular, most of them exhibit the VS word-order in infinitive clauses. In the case of transitive infinitives, the subject pronoun may thus appear between the verb and the direct object, despite being the higher or second argument of the former.

This analysis is threatened, we found, once we assume Associativity in the logical system. For then the syntactic relation of constituency becomes merely apparent, and *first* and *second* are no more than labels expressing the linear order of syntactic composition of the verb and its nominal arguments. This is particularly problematic in cases where both nominal arguments of a transitive verb stand in a post-verbal position; without a hierarchical distinction, we are unable to derive the VSO word-order and also block the VOS word-order. By marking the subject position of the infinitive verb with a bracket modality we sought to recover precisely this distinction between nominal arguments. Thus on our account, the lexical meaning attached to the infinitive verb constrains the subject to occupying the lower semantic position — even when the latter precedes the object.

Now, in light of the above, we have that infinitive clauses with free pre- or post-verbal pronominal subjects denote unsaturated properties. These are of course precisely the complements of our non-prototypical control clauses; and so we arrive at the third and final stage of our analysis. Importantly, this semantic result concerning infinitive clauses remains entirely compatible with the hypothesis that control sentences denote saturated properties — that is, propositions. Steedman was able to guarantee this in the case of English by assigning a multiple-binding meaning to the lexical semantics of the control verb. In a number of Romance languages, however, there are certain control verbs that do not directly select the infinitive clause; rather they select a preposition which, in turn, selects the infinitive controlled clause. We were nonetheless able to adapt Steedman's account here. We did so by assigning the multiple-binding meaning to the main verb and the identity function to the preposition. This allowed us to separate the main verb from the selected preposition and still give a uniform treatment *à la* Steedman that would encompass such verbs. We thus arrived at our desired result: we were able to guarantee that while subject pronouns are free inside the infinitive clause, they are semantically bound by the outside controller.

The more general key linguistic and logical results of our investigation may be summed up as follows:

- Our linguistic proposal is consistently integrated into Landau's T/Agr calculus for control.
- Our linguistic proposal instantiates the only remaining — and until now unexplored — combination of T/Agr features, thus filling the conceptual gap left by Landau.
- Our Type-Logical proposal is a conservative extension of **L** for which Cut-elimination, decidability and other desired meta-theoretical properties have already been proved.
- Our Type-Logical proposal can be implemented in a parser/theorem-prover.

Bibliography

[Ajd35] Kazimierz Ajdukiewicz. Die syntaktische konnexitat. Polish Logic, 1935.

[Alb04] Gabriela Alboiu. Shared arguments in control. Toronto Working Papers in Linguistics, 22, 2004.

[Alb10] Hugo Cardoso Albuquerque. Lógica proposicional via Lindenbaum-Tarski e Curry-Howard, 2010.

[AOD00] Luis Alonso-Ovalle and Francesco D'Introno. Full and null pronouns in Spanish: the zero pronoun hypothesis. Hispanic Linguistics at the Turn of the Millenium. Somerville, MA.: Cascadilla Press, 2000.

[Bac79] Emmon Bach. Control in Montague grammar. Linguistic Inquiry, 10(4):515–531, 1979.

[Bac84] Emmon Bach. Some generalizations of Categorial Grammars. In The formal complexity of natural language, pages 251–279. Springer, 1984.

[Bac88] Emmon Bach. Categorial grammars as theories of language. In Categorial grammars and natural language structures, pages 17–34. Springer, 1988.

[Bal02] Jason Baldridge. Lexically specified derivational control in Combinatory Categorial Grammar. PhDthesis, College of Science and Engineering. School of Informatics. University of Edinburgh., 2002.

[Bar84] Hendrik Pieter Barendregt. The Lambda calculus, volume 2. North-Holland Amsterdam, 1984.

[Bar93] Hendrik Pieter Barendregt. Lambda calculi with types. Handbook of logic in computer science (vol. 2): background: computational structures, 1993.

[Bar09] Maria do Pilar Pereira Barbosa. Sujeitos explícitos em orações infinitivas de controlo e elevação. Actas do XXIV Encontro da Associação Portuguesa de Linguística: Textos seleccionados, 1:97–114, 2009.

[Bar10] Pilar Barbosa. Overt subjects in raising and control complements and the null subject parameter. Linguistics Society of America Annual Meeting Extended Abstracts, 2010.

[Bar16] Pilar Pereira Barbosa. Overt subjects of raising and control infinitives and the mode of composition of subjects in the consistent null subject languages. 2016.

[BDS13] Henk Barendregt, Wil Dekkers, and Richard Statman. Lambda calculus with types. Cambridge University Press, 2013.

[Bel05] Adriana Belletti. Extended doubling and the vp periphery. Probus, 17(1):1–35, 2005.

[Ber13] Bruno Camus Bergareche. On deísmo: Another case of variation in spanish complementatio. Catalan journal of linguistics, (12):13–39, 2013.

[BH53] Yehoshua Bar-Hillel. A quasi-arithmetical notation for syntactic description. Language, 29(1):47–58, 1953.

[Bia46] Giovanni A. Bianchi. Della potestà e della politia della chiesa: trattati due contro le nuove opinioni di Pietro Giannone, volume 4. Salamoni, 1746.

[Bia07] Valentina Bianchi. Wh-infinitives and the licensing of "anaphoric tense". Proceedings of the" XXXII Incontro di Grammatica Generativa", Firenze, 2-4 March 2006, 2007.

[Bis14] Priyanka Biswas. The role of tense and agreement in the licensing of subjects: Evidence from participial clauses in bangla. Natural Language & Linguistic Theory, 32(1):87–113, 2014.

[BJ07] Chris Barker and Pauline I. Jacobson. Direct compositionality. Number 14. Oxford University Press, 2007.

[BMVB88] Wojciech Buszkowski, Witold Marciszewski, and Johan Van Benthem. Categorial grammar, volume 25. John Benjamins Publishing, 1988.

[Boš02] Željko Bošković. A–movement and the EPP. Syntax, 5(3):167–218, 2002.

[Bot34] Carlo Botta. Storia d'Italia dal 1789 al 1814. 1834.

[Bus03] Wojciech Buszkowski. Type logics in grammar. In Trends in logic, pages 337–382. Springer, 2003.

[Car97] Bob Carpenter. Type-logical semantics. MIT press, 1997.

[Car99] Anna Cardinaletti. Italian emphatic pronouns are post-verbal subjects. Working Papers in Linguistics, 9(1-2), 1999.

[Car05] Gabriel D. Carroll. Formal properties of Categorial Grammars. BSthesis, Department of Mathematics and Department of Linguistics. Harvard University, 2005.

[Chi84a] Gennaro Chierchia. Anaphoric properties of infinitives and gerunds. In Proceedings of the Third West Coast Conference on Formal Linguistics, pages 28–39, 1984.

[Chi84b] Gennaro Chierchia. Topics in the Semantics of Infinitives and Gerunds. PhDthesis, University of Massachusetts, Amherst. Published in 1989 by Garland, New York, 1984. Published in 1989 by Garland, New York.

[Chi85] Gennaro Chierchia. Formal semantics and the grammar of predication. Linguistic Inquiry, 16(3):417–443, 1985.

[Chi88] Gennaro Chierchia. Aspects of a categorial theory of binding. In Oehrle Richard T., Bach Emmon, and Wheeler Deirdre, editors, Categorial Grammars and Natural Language Structures, volume 32, pages 125–151. Springer, 1988.

[Cho65] Noam Chomsky. Aspects of the theory of syntax. Cambridge, Massachusetts: MIT Press, 1965.

[Cho73] Noam Chomsky. Conditions on transformations. a festschrift for Morris Halle, 1973.

[Cho81] Noam Chomsky. Lectures on Government and Binding. Kluwer, Dordrecht, 1981.

[Cho82] Noam Chomsky. Some concepts and consequences of the theory of government and binding, volume 6. MIT press, 1982.

[Cho86] Noam Chomsky. Knowledge of language: Its nature, origin, and use. Greenwood Publishing Group, 1986.

[Cho93] Noam Chomsky. Lectures on government and binding: The Pisa lectures. Number 9. Walter de Gruyter, 1993.

[Cho95] Noam Chomsky. The minimalist program. Current Studies in Linguistics 28, 1995.

[Cho00] Noam Chomsky. Minimalist inquiries: the framework. step by step. essays on minimalist syntax in honor of Howard Lasnik, edited by roger martin, david michaels & juan uriagereka, 89-155, 2000.

[CKM14] Claudia Casadio and Aleksandra Kiślak-Malinowska. Italian clitic patterns in Pregroup Grammar: State of the art. In Categories and Types in Logic, Language, and Physics, pages 156–171. Springer, 2014.

[CL93] Noam Chomsky and Howard Lasnik. The theory of principles and parameters. Syntax: An international handbook of contemporary research, 1:506–569, 1993.

[CL01] Claudia Casadio and Joachim Lambek. An algebraic analysis of clitic pronouns in Italian. In International Conference on Logical Aspects of Computational Linguistics, pages 110–124. Springer, 2001.

[CM53] Francisco Camprodón and Nicolau Manent. Tres para una: zarzuela en tres actos y en verso arreglada a la escena española por D. Francisco Camprodón; música del maestro D. Nicolás Manent. Barcelona (calle de Pons, 4) Imprenta de Pons y Cᵻ, 1853.

[CM16] María Inés Corbalán and Glyn Morrill. Overtly anaphoric control in Type Logical Grammar. In 21st International Conference on Formal Grammar, pages 183–199. Springer, 2016.

[Cor15] María Inés Corbalán. Podem as clausulas com sujeito denotar propriedades? o problema dos sujeitos controlados na perspectiva da Gramatica Categorial. Revista da ABRALIN, 14(3), 2015.

[Cor17a] María Inés Corbalán. Binding domains: anaphoric and pronominal pronouns in Categorial Grammar. In 22nd International Conference on Formal Grammar, page (forthcoming). Springer, 2017.

[Cor17b] María Inés Corbalán. A syntactic treatment for free and bound pronouns in a Type-Logical Grammar. In International Symposium on Distributed Computing and Artificial Intelligence, pages 223–230. Springer, 2017.

[Cyr11] Sonia Cyrino. On complex predicates in Brazilian Portuguese. Iberia: An International Journal of Theoretical Linguistics, 2(2), 2011.

[dCS71] Miguel de Cervantes Saavedra. El ingenioso hidalgo Don Quijote de la Mancha. Sáenz de Jubera, hermanos, 1871.

[DD08] William D. Davies and Stanley Dubinsky. The grammar of raising and control: A course in syntactic argumentation. John Wiley & Sons, 2008.

[Dem77] Violeta Demonte. La subordinación sustantiva. Cátedra Madrid, 1977.

[dM64] Diego Hurtado de Mendoza. Obras, coleccionadas por Nicolas del Paso y Delgado, volume 1 of Biblioteca de escritores granadinos. 1. na, 1864.

[DO38] Eugenio De Ochoa. Tesoro del teatro español, desde su origen (año de 1356) hasta nuestros dias, volume 1. Baudry, 1838.

[Dow82] David R. Dowty. Grammatical relations and Montague grammar. In The nature of syntactic representation, pages 79–130. Springer, 1982.

[Dow85] David R. Dowty. On recent analyses of the semantics of control. Linguistics and Philosophy, 8(3):291–331, 1985.

[Dow88] David Dowty. Type raising, functional composition, and non-constituent conjunction. In Categorial grammars and natural language structures, pages 153–197. Springer, 1988.

[Dow97] David Dowty. Non-constituent coordination, wrapping, and multimodal categorial grammars. In Structures and norms in science, pages 347–368. Springer, 1997.

[DS09] Paola Dos Santos. Orações infinitivas. Da seleção ao controle. MAthesis, Universidade Federal do Rio Grande do Sul, 2009.

[DT11] Ángela Di Tullio. Infinitivos introducidos por de. Cuadernos de la ALFAL, 3:176–187, 2011.

[Far85] Donka Farkas. Intensional descriptions and the Romance subjunctive mood. Taylor & Francis, 1985.

[Fer09] Maribel Fernández. Models of Computation: An Introduction to Computability Theory. Springer Science & Business Media, 2009.

[Fio35] Giuseppe Fiocco. Pizzolo, Niccolò. In Giovanni Treccani, editor, Enciclopedia Italiana. http://www.treccani.it, 1935.

[Fra07] Mara Frascarelli. Subjects, topics and the interpretation of referential pro. Natural Language & Linguistic Theory, 25(4):691–734, 2007.

[FS93] Maria Cristina Figuereido Silva. Some notes on null subjects in the Brasilian Portuguese tensed sentences. Working Papers in Linguistics, 3(1):69–82, 1993.

[Gir87] Jean-Yves Girard. Linear logic. Theoretical computer science, 50(1):1–101, 1987.

[GJKO07] Nikolaos Galatos, Peter Jipsen, Tomasz Kowalski, and Hiroakira Ono. Residuated lattices: an algebraic glimpse at substructural logics, volume 151. Elsevier, 2007.

[GM15] Maximiliano Guimarães and Gesoel Mendes. Interação entre foco, morfologia e controle em PB: evidências adicionais para a Teoria de Controle por Movimento. Revista Linguística, 9(1), 2015.

[Gor00] Edair Gorski. Variação no uso do infinitivo pessoal. Organon, 14(28-29), 2000.

[Gre15] H. Gregory. Language and Logics. Edinburgh University Press, 2015.

[GSD14] Anabela Gonçalves, Ana Lúcia Santos, and Inês Duarte. (Pseudo-)Inflected infinitives and control as Agree, volume 6 of Romance Languages and Linguistic Theory, chapter 9, pages 161–180. John Benjamins Publishing Company, 2014.

[Hae94] Liliane Haegeman. Introduction to Government and Binding theory. Wiley-Blackwell, 1994.

[Hep90] Mark Hepple. The grammar and processing of order and dependency: A categorial approach. PhDthesis, University of Edinburgh, 1990.

[Hep92] Mark Hepple. Command and domain constraints in a categorial theory of binding. Proceedings of the Eight Amsterdam Colloquium, pages 253–270, 1992.

[Hep94] Mark Hepple. Discontinuity and the Lambek calculus. In Proceedings of the 15th conference on Computational linguistics-Volume 2, pages 1235–1239. Association for Computational Linguistics, 1994.

[Her11] Peter Herbeck. Overt subjects in Spanish control infinitives and the theory of empty categories. Generative Grammar in Geneva, 7:1–22, 2011.

[Her12] Peter Herbeck. How much pro-drop is control? evidence from Spanish, Catalan, and Italian. In The 22nd Colloquium on Generative Grammar, Facultat de Filosofia i Lletres, Universitat Autònoma de Barcelona, 2012.

[Her15] Peter Herbeck. Overt PRO in Romance. In Hispanic Linguistics at the Crossroads: Theoretical linguistics, language acquisition and language contact. Proceedings of the Hispanic Linguistics Symposium 2013, volume 4, page 25. John Benjamins Publishing Company, 2015.

[HGS60] Y. Bar Hillel, Chaim Gaifman, and Eli Shamir. On categorial and phrase structure grammars. Bulletin of the Research Council of Israel, 9, 1960.

[HK98] Irene Heim and Angelika Kratzer. Semantics in generative grammar, volume 13. Blackwell Oxford, 1998.

[HNS09] Anders Holmberg, Aarti Nayudu, and Michelle Sheehan. Three partial null-subject languages: a comparison of Brazilian Portuguese, Finnish and Marathi. Studia Linguistica, 63(1):59–97, 2009.

[Hol10] Anders Holmberg. Null subject parameters. Parametric variation: Null subjects in minimalist theory, pages 88–124, 2010.

[Hor99] Norbert Hornstein. Movement and control. Linguistic Inquiry, 30(1):69–96, 1999.

[Hor03] Norbert Hornstein. On control. Minimalist Syntax, pages 6–81, 2003.

[Hor06] Norbert Hornstein. Pronouns in a minimalist setting. University of Maryland Working Papers in Linguistics, (14):47–80, 2006.

[HS86] J. Roger Hindley and Jonathan P. Seldin. Introduction to Combinators and (lambda) Calculus, volume 1. CUP Archive, 1986.

[HSM01] Norbert Hornstein and Itziar San Martin. Obviation as anti-control. Anuario del Seminario de Filología Vasca" Julio de Urquijo", 35(1):367–384, 2001.

[Jac90] Pauline Jacobson. Raising as function composition. Linguistics and Philosophy, 13(4):423–475, 1990.

[Jac96a] Pauline Jacobson. The locality of interpretation: the case of binding and coordination. In Proceedings from Semantics and Linguistic Theory VI, CLC Publications, Ithaca, pages 111–135, 1996.

[Jac96b] Pauline Jacobson. The syntax/semantics interface in Categorial Grammar. In Shalom Lappin, editor, The Handbook of Contemporary Semantic Theory, chapter 4, pages 89–116. Blackwell Reference, 1996.

[Jac99] Pauline Jacobson. Towards a variable-free semantics. Linguistics and Philosophy, 22(2):117–185, 1999.

[Jac03] Pauline Jacobson. Binding without pronouns (and pronouns without binding). In Resource-Sensitivity, Binding and Anaphora, pages 57–96. Springer, 2003.

[Jäg97] Gerhard Jäger. Anaphora and ellipsis in type-logical grammar. In Proceedings of the Eleventh Amsterdam Colloquium, ILLC, University of Amsterdam, pages 175–180, 1997.

[Jäg98] Gerhard Jäger. A multi-modal analysis of anaphora and ellipsis. University of Pennsylvania Working Papers in Linguistics, 5(2):2, 1998.

[Jäg99] Gerhard Jäger. Deconstructing jacobson's z. Proceedings of the 12th Amsterdam Colloquium, 1999.

[Jäg01] Gerhard Jäger. Anaphora and quantification in categorial grammar. In Selected Papers from the Third International Conference, on Logical Aspects of Computational Linguistics, LACL '98, pages 70–89, London, UK, UK, 2001. Springer-Verlag.

[Jäg05] Gehard Jäger. Anaphora and Type Logical Grammar. Springer, 2005.

[Jun04] Achim Jung. A short introduction to the Lambda Calculus. School of Computer Science, The University of Birmingham, 2004.

[Ket44] Oiva Ketonen. Untersuchungen zum prädikatenkalkül. Annales Academiae Scientiarum Fennicae, Ser. A, I Mathematica-physica, (23):71pp, 1944.

[KO03] Geert-Jan M Kruijff and Richard T Oehrle. Resource-sensitivity, binding and anaphora, volume 80. Springer Science & Business Media, 2003.

[Kra09] Angelika Kratzer. Making a pronoun: Fake indexicals as windows into the properties of pronouns. Linguistic Inquiry, 40(2):187–237, 2009.

[KS88] Hilda Koopman and Dominique Sportiche. Subjects, ms. Los Angeles: UCLA, 1988.

[KS91] Hilda Koopman and Dominique Sportiche. The position of subjects. Lingua, 85(2-3):211–258, 1991.

[Lag87] Marina Fernández Lagunilla. Los infinitivos con sujetos léxicos en Español. In Sintaxis de las lenguas Románicas, pages 125–147. Ediciones El Arquero, 1987.

[Lam58] Joachim Lambek. The mathematics of sentence structure. The American Mathematical Monthly, 65(3):154–170, 1958.

[Lam61] Joachim Lambek. On the calculus of syntactic types. Structure of Language and its Mathematical Aspects, 166:C178, 1961.

[Lam88] Joachim Lambek. Categorial and categorical grammars. In Categorial grammars and natural language structures, pages 297–317. Springer, 1988.

[Lan00] Idan Landau. Elements of Control: Structure and Meaning in Infinitival Constructions. Studies in natural language and linguistic theory. Kluwer Academic Publishers, 2000.

[Lan04] Idan Landau. The scale of finiteness and the calculus of control. Natural Language & Linguistic Theory, 22(4):811–877, 2004.

[Lan06] Idan Landau. Severing the distribution of pro from case. Syntax, 9(2):153–170, 2006.

[Lan08] Idan Landau. Two routes of control: Evidence from case transmission in russian. Natural Language & Linguistic Theory, 26(4):877–924, 2008.

[Lan13] Idan Landau. Control in generative grammar: A research companion. Cambridge University Press, 2013.

[Lan15] Idan Landau. A two-tiered theory of control, volume 71. MIT Press, 2015.

[Lar88] Richard K. Larson. On the double object construction. Linguistic inquiry, 19(3):335–391, 1988.

[Lar91] Richard K. Larson. Promise and the theory of control. Linguistic Inquiry, 22(1):103–139, 1991.

[Las01] Howard Lasnik. A note on the EPP. Linguistic Inquiry, 32(2):356–362, 2001.

[Lee03] Felicia Lee. Anaphoric R–expressions as bound variables. Syntax, 6(1):84–114, 2003.

[Liv11] Inna Livitz. Incorporating PRO: A defective-goal analysis. NYU Working Papers in Linguistics (NYUWPL), 3:95–119, 2011.

[LM14] Maria Elena Pena Louredo Martins. Caiu do céu. Autor independente, 2014.

[LMyS99] Luis de León and Gregorio Mayans y Siscar. Escritores del siglo xvi, tomo segundo, obras del maestro fray luisde león precedelas su vida, escrita por d. gregorio mayans y siscar. 1899.

[LR95] Marina Fernández Lagunilla and Alberto Anula Rebollo. Observaciones sobre la flexión de los infinitivos. In De grammatica generativa, pages 185–198, 1995.

[Luf10] Celso Pedro Luft. Dicionário prático de regência verbal. Ática, 2010.

[Luj80] María Luján. Clitic promotion and mood in Spanish verbal complements. Linguistics, 18(5-6):381–484, 1980.

[Luz93] Geraldo Antônio da Luz. O Infinitivo Pessoal: Uma análise do Português contemporâneo. PhDthesis, Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, 1993.

[Mar01] Roger Martin. Null case and the distribution of pro. Linguistic Inquiry, 32(1):141–166, 2001.

[Mar11] Marijana Marelj. Bound-variable anaphora and left branch condition. Syntax, 14(3):205–229, 2011.

[McF08] Thomas McFadden. Overt subjects of infinitives and for-to in the history of english. In Talk presented at the 10th Diachronic Generative Syntax Conference DiGS X, Cornell, 2008.

[McF12] Thomas McFadden. For in English infinitives has nothing to do with case. In Linguists of tomorrow: Selected papers from the 1st Cyprus Postgraduate Student Conference in Theoretical and Applied Linguistics, pages 131–155, 2012.

[Men00] Guido Mensching. Infinitive constructions with specified subjects: A syntactic analysis of the romance languages, 2000.

[Meu14] Gerard Meurant. Formal models and semantics, volume 2. Elsevier, 2014.

[MG92] Glyn Morrill and Anna Gavarró. Catalan clitics. Word Order in Categorial Grammar, pages 211–232, 1992.

[Mil02] D. Gary Miller. Nonfinite structures in theory and change. Oxford University Press on Demand, 2002.

[MM96] Glyn Morrill and Josep-Maria Merenciano. Generalising discontinuity. TAL. Traitement automatique des langues, 37(2):119–143, 1996.

[Mod07] Marcello Modesto. Inflected infinitives in Brazilian Portuguese as an argument both contra and in favor of a movement analysis of control. Revista Letras, 72, 2007.

[Mod10a] Marcello Modesto. Inflected infinitives in bp and the structure of nonfinite complements. Paper presented at IV Romania Nova, Campos do Jordão, São Paulo, Brazil, 2010.

[Mod10b] Marcello Modesto. What Brazilian Portuguese says about control: remarks on boeckx & hornstein. Syntax, 13(1):78–96, 2010.

[Mod11] Marcello Modesto. Infinitivos flexionados em Português Brasileiro e sua relevância para a teoria do controle. Estudos da linguagem, casamento entre temas e perspectivas, pages 63–87, 2011.

[Mon74] Richard Montague. Formal Philosophy: Selected Papers of Richard Montague. Yale University Press, 1974.

[Moo88] Michael Moortgat. Categorical investigations. Logical and linguistic aspects of the Lambek calculus., volume 9 of Groningen-Amsterdam studies in semantics. Foris Publications, Dordrecht and Providence, 1988.

[Moo96a] Michael Moortgat. Generalized quantifiers and discontinuous type, volume 6, page 181. Walter de Gruyter, 1996.

[Moo96b] Michael Moortgat. Multimodal linguistic inference. Journal of Logic, Language and Information, 5(3-4):349–385, 1996.

[Moo02] Michael Moortgat. Encyclopedia of cognitive science, volume 1, chapter: Categorial grammar and formal semantics, 2002.

[Moo11] Michael Moortgat. Categorial Type Logics. Handbook of Logic and Language. Second Edition Elsevier, 2011.

[Mor90] Glyn Morrill. Intensionality and boundedness. Linguistics and Philosophy, 13(6):699–726, 1990.

[Mor92] Glyn Morrill. Categorial formalisation of relativisation: Pied piping, islands, and extraction sites. Report de Recerca LSI-92-23-R, Departament de Llenguatges i Sistemes Informàtics, Universitat Politècnica de Catalunya, 1992.

[Mor94a] Glyn Morrill. Structural facilitation and structural inhibition, 1994.

[Mor94b] Glyn Morrill. Type Logical Grammar. Categorial Logic of Signs. Springer Netherlands, 1994.

[Mor95] Glyn Morrill. Discontinuity in categorial grammar. Linguistics and Philosophy, 18(2):175–219, 1995.

[Mor00] Glyn Morrill. Type-logical anaphora. resreport, Research Report LSI-00-77-R, Departament de Llenguatges i Sistemes Informatics, Universitat de Politecnica de Catalunya, 2000.

[Mor03] Glyn Morrill. On bound anaphora in Type Logical Grammar. In Resource-sensitivity, binding and anaphora, pages 159–177. Springer, 2003.

[Mor11] Glyn Morrill. Categorial grammar: Logical syntax, semantics, and processing. Oxford University Press, 2011.

[Mor12] Glyn Morrill. Catlog: A categorial parser/theorem-prover. Logical Aspects of Computational Linguistics: System Demonstrations. Nantes, France: University of Nantes, pages 13–16, 2012.

[Mor17a] Glyn Morrill. Grammar logicised: relativisation. Linguistics and Philosophy, 40(2):119–163, 2017.

[Mor17b] Glyn Morrill. Parsing logical grammar: CatLog3. In Workshop on Logic and Algorithms in Computational Linguistics 2017 (LACompLing2017), page 107, 2017.

[MS11] Thomas McFadden and Sandhya Sundaresan. Nominative case is independent of finiteness and agreement. Ms., Universitetet i Tromsø and Universität Stuttgart, 2011.

[MSA93] Glyn Morrill and M. Teresa Solias Arís. Tuples, discontinuity, and gapping in categorial grammar. In Proceedings of the sixth conference on European chapter of the Association for Computational Linguistics, pages 287–296. Association for Computational Linguistics, 1993.

[MSL13] Carlos Mioto, Maria Cristina Figueiredo Silva, and Ruth Elisabeth Vasconcellos Lopes. Novo manual de sintaxe. Ed. Contexto, 2013.

[MV10a] Glyn Morrill and Oriol Valentín. On anaphora and the binding principles in categorial grammar. In International Workshop on Logic, Language, Information, and Computation, pages 176–190. Springer, 2010.

[MV10b] Glyn Morrill and Oriol Valentín. On calculus of displacement. In Proceedings of the 10th International Workshop on Tree Adjoining Grammars and Related Formalisms, pages 45–52, 2010.

[MV12] Glyn Morrill and Oriol Valentín. Generalized discontinuity. In Formal Grammar, pages 146–161. Springer, 2012.

[MV14] Glyn Morrill and Oriol Valentín. Displacement logic for anaphora. Journal of Computer and System Sciences, 80(2):390–409, 2014.

[MV15] Glyn Morrill and Oriol Valentín. Computational coverage of TLG: Displacement. Proceedings of Empirical Advances in Categorial Grammar, pages 132–161, 2015.

[MVF11] Glyn Morrill, Oriol Valentín, and Mario Fadda. The displacement calculus. Journal of Logic, Language and Information, 20(1):1–48, 2011.

[MVpt] Glyn Morrill and Oriol Valentín. Type Logical Syntax and Semantics. manuscript.

[Nav47] Colección de documentos inéditos para la historia de españa, 1847.

[Neg87] Esmeralda Vailati Negrão. Anaphora in Brazilian Portuguese complement structures. PhD dissertation, University of Wisconsin, Madison, 1987.

[Nis13] Chiyo Nishida. Restructuring and clitic climbing in Romance: A categorial grammar analysis. Local Modelling of Non-Local Dependencies in Syntax, 547:371–400, 2013.

[Nun08] Jairo Nunes. Inherent case as a licensing condition for a-movement: The case of hyper-raising constructions in Brazilian Portuguese. Journal of Portuguese Linguistics, 7(2), 2008.

[Oku94] Satoshi Oku. The extended projection principle effect in the minimalist. Science, 43(1):127–154, 1994.

[Oli20] Matías Oliden. El Camarista Oliden contesta por lo que a él toca al papel publicado por el General Pueyrredon. Imprenta de Phocion, 1820.

[Ono03] Hiroakira Ono. Substructural logics and residuated lattices—an introduction. Trends in Logic, 20:177–212, 2003.

[Ort02] Iván Ortega. Overt subjects of infinitives in Spanish: Abstract agreement and the EPP. In Selected proceedings of the 12h Annual Graduate and Professional Symposium on Hispanic and Luso-Brazilian Literature, Language and Culture at the University of Arizona, pages 21–28, 2002.

[Pag06] Luiz Arthur Pagani. Impressões lingüísticas sobre duas axiomatizações para a Gramática Categorial. Revista Letras, 70, 2006.

[Pao02] Francesco Paoli. Substructural Logics: A Primer, volume 13. Springer Science & Business Media, 2002.

[Paz13] Justin Paz. La distribución e interpretación de los sujetos léxicos de infinitivo en Español: un caso problemático para la Teoría del Caso. The University of Arizona, 2013.

[PG87] Carlos Piera Gil. Sobre la estructura de las cláusulas de infinitivo. In Sintaxis de las lenguas románicas, pages 148–166. Ediciones El Arquero, 1987.

[@pl13] @planetacb. Available at `https://twitter.com/planetacb/status/336924708085854209`, june 2013.

[Pol89] Jean-Yves Pollock. Verb movement, universal grammar, and the structure of ip. Linguistic Inquiry, 20(3):365–424, 1989.

[Pou95] Christopher J. Pountain. Infinitives with overt subjects: a pragmatic approach. Portuguese, Brazilian and African Studies: studies presented to Clive Willis on his retirement, pages 11–25, 1995.

[PP06] Maria Polinsky and Eric Potsdam. Expanding the scope of control and raising. Syntax, 9(2):171–192, 2006.

[Puj80] Giuseppe M. Pujati. Difficoltà proposte all'ex gesuita signor canonico Luigi Mozzi sopra le sue riflessioni critico-dogmatiche. Lettera seconda. 1780.

[PyF53a] Fernando Patxot y Ferrer. Las glorias nacionales: Grande historia universal de todos los reinos, provincias, islas, y colonias de la monarquía española, desde los tiempos primitivos hasta el año de 1852, volume 5. L. Tasso, 1853.

[PyF53b] Fernando Patxot y Ferrer. Las glorias nacionales: Grande historia universal de todos los reinos, provincias, islas, y colonias de la monarqúia española, desde los tiempos primitivos hasta el año de 1852, volume 4. L. Tasso, 1853.

[Rab04] Poliana Camargo Rabelo. Sobre a questão do controle com o infinitivo flexionado Português, 2004.

[Rap87] Eduardo Raposo. Case theory and infl-to-comp: The inflected infinitive in european portuguese. Linguistic Inquiry, 18(1):85–109, 1987.

[Rap89] Eduardo Raposo. Prepositional infinitival constructions in European Portuguese. In Osvaldo Jaeggli and Kenneth J. Safir, editors, The null subject parameter, pages 277–305. Springer, 1989.

[Ree12] Lisa Reed. A semantic constraint on tenseless clausal complementation. Studia Linguistica, 66(3):286–321, 2012.

[Ref58] Reformistas antiguos españoles. Number v. 14, parte 1 in Reformistas antiguos españoles. 1858.

[Riz82] Luigi Rizzi. Issues in Italian syntax, volume 11. Walter de Gruyter, 1982.

[Riz90] Luigi Rizzi. Relativized minimality. The MIT Press, 1990.

[Roj15] Raúl Rojas. A tutorial introduction to the lambda calculus. arXiv preprint arXiv:1503.09060, 2015.

[Ron79] Mitsou Ronat. Pronoms topiques et pronoms distinctifs. Langue française, (44):106–128, 1979.

[Ros65] Peter S. Rosenbaum. The grammar of English predicate complement constructions. PhDthesis, Massachusetts Institute of Technology, 1965.

[Ros67] John Robert Ross. Constraints on variables in syntax. PhDthesis, Massachusetts Institute of Technology, 1967.

[Ros70] Peter S. Rosenbaum. A principle governing deletion in English sentential complementation. Readings in English Transformational Grammar, pages 20–29, 1970.

[RR89] Luigi Rizzi and Ian Roberts. Complex inversion in French. Probus, 1(1):1–30, 1989.

[RR93] Tanya Reinhart and Eric Reuland. Reflexivity. Linguistic inquiry, 24(4):657–720, 1993.

[RTO88] Richard T. Oehrle Emmon Bach Deirdre Wheeler (eds.) Richard T. Oehrle, Emmon Bach (auth.). Categorial grammars and natural language structures. Studies in Linguistics and Philosophy 32. Springer Netherlands, 1 edition, 1988.

[RV80] Alain Rouveret and Jean-Roger Vergnaud. Specifying reference to the subject: French causatives and conditions on representations. Linguistic Inquiry, pages 97–202, 1980.

[SA92] M. Teresa Solias Arís. Gramáticas categoriales, coordinación generalizada y elisión. PhDthesis, Universidad Autnoma de Madrid. Madrid, 1992.

[SA96] M. Teresa Solias Arís. Gramática categorial: modelos y aplicaciones. Editorial Síntesis, 1996.

[San09] Paola Junqueira Pinto dos Santos. Orações infinitivas: da seleção ao controle, 2009.

[SB11] Mark Steedman and Jason Baldridge. Combinatory Categorial Grammar, pages 181–224. Wiley-Blackwell, 2011.

[Sch07] Kim Schulte. Prepositional infinitives in Romance: A usage-based approach to syntactic change, volume 3. Peter Lang, 2007.

[Sci04] Emily E. Scida. The inflected infinitive in Romance languages. Routledge, 2004.

[Sig08] Halldór Ármann Sigurðsson. The case of PRO. Natural Language & Linguistic Theory, 26(2):403–450, 2008.

[Sit02] Ioanna Sitaridou. The synchrony and diachrony of Romance infinitives with nominative subjects. PhDthesis, University of Manchester, 2002.

[SM10] Sandhya Sundaresan and Thomas McFadden. Subject distribution and finiteness in tamil and other languages: Selection vs. case. Journal of South Asian Linguistics, 2(1), 2010.

[Ste88] Mark Steedman. Combinators and grammars. In Richard. T. Oehrle, Emmon W. Bach, and Weeler Deirdre, editors, Categorial grammars and natural language structures, number 32 in Studies in Linguistics and Philosophy 32, pages 417–442. Reidel Publishing Company, 1 edition, 1988.

[Ste92] Mark Steedman. Categorial grammar. resreport, Department of Computer Information Science Technical Reports (CIS), MS-CIS-92-52 LINC LAB 230, University of Pennsylvania, 1992.

[Ste93] Mark Steedman. Categorial grammar. Lingua, 90(3):221 – 258, 1993.

[Ste94] Mark Steedman. Binding and control in ccg and its relatives. techreport MS-CIS-94-47, University of Pennsylvania Department of Computer and Information Science, 1994.

[Ste96] Mark Steedman. Surface Structure and Interpretation, volume 30 of Linguistic Inquiry Monographs. MIT Press, 1996.

[Sti07] Barbara Stiebels. Towards a typology of complement control. ZAS Papers in Linguistics, 47(8), 2007.

[Sto82] Tim Stowell. The tense of infinitives. Linguistic Inquiry, 13(3):561–570, 1982.

[SU06] Morten Heine Sørensen and Pawel Urzyczyn. Lectures on the Curry-Howard isomorphism, volume 149. Elsevier, 2006.

[Sza89] Anna Szabolcsi. Bound variables in syntax (are there any?). In Renate Bartsch, Johan van Benthem, and Peter van Emde Boas, editors, Semantics and Contextual Expressions, pages 295–318. Dordrecht: Foris, 1989.

[Sza92] Anna Szabolcsi. Combinatory grammar and projection from the lexicon, volume 24 of Lecture Notes, chapter 9, pages 241–269. CSLI Publications, 1992.

[Sza03] Anna Szabolcsi. Binding on the fly: Cross-sentential anaphora in variable-free semantics. In Resource-Sensitivity, Binding and Anaphora, pages 215–227. Springer, 2003.

[Sza09] Anna Szabolcsi. Overt nominative subjects in infinitival complements cross-linguistically: Data, diagnostics, and preliminary analyses. Papers in syntax, 2, 2009.

[Ter97] Arhonto Terzi. PRO and null case in finite clauses. Linguistic review, 14(4):335–360, 1997.

[Tir83] Girolamo Tiraboschi. Storia della letteratura italiana, volume 6. Salvioni, 1783.

[Tir07] Girolamo Tiraboschi. Storia Della Letteratura Italiana del cavaliere abate Girolamo Tiraboschi, volume 5.1, Parte prima. 1807.

[Tor96] Esther Torrego. On quantifier float in control clauses. Linguistic inquiry, pages 111–126, 1996.

[Tor98] Esther Torrego. Nominative subjects and pro-drop INFL. Syntax, 1(2):206–219, 1998.

[Val13] Oriol Valentín. Theory of Discontinuous Lambek Calculus. PhDthesis, Universitat Autònoma de Barcelona, 2013.

[Van13] Clara Vanderschueren. Infinitivo y sujeto en portugués y español: Un estudio empírico de los infinitivos adverbiales con sujeto explícito, volume 377. Walter de Gruyter, 2013.

[Var93] Spyridoula Varlokosta. Control in modern Greek. MAthesis, Dept. of Linguistics. University of Maryland at College Park, 1993.

[vB88] Johan van Benthem. The semantics of variety in categorial grammar. Categorial grammar, 25:37–55, 1988.

[VD13] Clara Vanderschueren and Kevin Diependaele. The portuguese inflected infinitive: an empirical approach. Corpus linguistics and linguistic theory, 9(1):161–186, 2013.

[VdF59] Giuseppe Vernazza di Freney. Dizionario dei tipografi e dei principali correttori ed intagliatori che operarono negli Stati Sardi di terraferma e piu specialmente in Piemonte sino all'anno 1821: Opera e estampa che rimasta imperfetta per la morte dell'autore Barone Vernazza di Freney viene in luce per cura d'una societa anonima. 1859.

[Wil80] Edwin Williams. Predication. Linguistic Inquiry, 11(1):203–238, 1980.

[Wur02] Susi Wurmbrand. Syntactic versus semantic control. In Studies in comparative Germanic syntax: Proceedings from the 15th workshop on comparative Germanic syntax, volume 53, page 93. John Benjamins Publishing, 2002.

[wwwa] El Austral. diario de la Araucanía. Previously available at `http://www.australtemuco.cl/prontus4noticias/site/artic/20090417/pags/20090417000305.html`.

[wwwb] Oremos juntos. Previously available at `http://www.oremosjuntos.com/Papa/BonifacioIII.html`.

[www99] El tiempo. Available at `http://www.eltiempo.com/archivo/documento/MAM-866757`, february 1999.

[Zár48] Hernando de Zárate. Discursos de la paciencia cristiana, volume 27. Rivadeneyra, 1948.

[Zwa90] Jan-Wouter Zwart. On case theory and control in Portuguese and Modern Greek, volume 1 of LCJL, pages 181–196. Leiden University Press, 1990. Series continued as ConSOLE Proceedings.

Appendix A. Overt Pronouns in Control Infinitives

European Portuguese

(241) Eu exigi aos alunos [**eles** fazer**em** um trabalho].
I forced to-the students they.NOM do.INFL.3PL a work
'I forced the student to do some homework.' [Zwa90]:186

(242) Ontem os pais obrigaram as crianças a [fazer **elas** a cama].
yesterday the parents forced the children to make they the bed
'Yesterday their parents forced the children to make their bed themselves.' [Bar16]

(243) Os pais quiseram [ir **eles** de comboio].
the parents wanted go.INF they by train
'The parents wanted to go by train.' [Bar09]

(244) O João decidiu [resolver **ele** o problema].
the John decided solve.INF.(3SG) the problem
'John decided to solve the problem himself.' [Bar10]

Brazilian Portuguese

(245) Pedro quer [**ele** chegar cedo].
Peter wants [he.NOM arrive.INF early]
'Peter wants to arrive early.' [GM15]

(246) Os meninos querem [falar **eles** com o diretor].
the kids want to talk they with the director
'The kids want that they themselves speak with the director.' [Bel05]:22

(247) Acusa os colegas de [**eles** ser**em** corruptos].
accuse.3SG the colleagues of they.NOM be.INFL.3PL corrupt
'She accuses her colleagues of being corrupt.' [Luf10]:34

(248) Os meninos querem [**eles** embrulhar(**em**) os presentes].
the boys want they.NOM wrap.INFL.(3PL) the gifts
'The boys want to wrap the gifts themselves.' [Neg87]

(249) A polícia forçou os manifestantes a [todos **eles** saírem].
the police forced the protesters to all they.NOM leave.INFL.3PL
'The police forced all of the protesters to leave.' [Rab04]:49,57

(250) Nós convencemos os estudantes a [todos **eles** irem à festa].
We convinced the students to all they.NOM go.INFL.3PL to-the party
'We convinced the students to go to the party.' [Rab04]:57

(251) Roberto, eu tentei [**eu** enviar meu convite a você].
Robert, I tried I send.INFL.(1SG) my invitation to you
'Robert, I tried to send you an invitation.' [Cyr11]:12

(252) Não quero [**eu** também ser falso moralista].
Not want.1SG I also be.INF false moralist
'I don't want it to be the case that I too am a false moralist.' [Sza09]:35

Italian

(253) Gianni ha deciso [di intervenire **lui**].
John has decided COMP intervene.INF he.NOM
'John has decided to intervene.' [Car99]

(254) Gianni$_1$ mi ha promesso [di farlo **lui**$_1$].
John CL has promised COMP do.INF he.NOM
'John$_1$ has promised me that he$_1$ will do it.' [Riz82]

(255) Maria mi ha chiesto [di parlare **io** con Gianni].
maria CL has asked COMP talk.INF I.NOM with Gianni
'Maria asked if I would talk to Gianni.' [Bel05]:21

(256) Il Rettore ha deciso [di aprire **lui** il convegno (***lui**)].
the Dean has decided COMP open he.NOM the conference he
'The Dean has has decided to open the conference.' [Car99]:79

Spanish

(257) El abogado propuso [interrogar **él** al testigo].
the lawyer proposed interrogate.INF he.NOM to-the witness
The lawyer proposed to interrogate the witness by himself. [LR95]:191

(258) Juan prometió a su profesor [hacer **él** los deberes].
John promised to his teacher do.INF he.NOM the homework
'John promised his teacher to do the homework by himself.' [Her11]:10

(259) Le prometí a Juan [escribir **yo** un libro (**yo**)]
CL promised.1SG to John write.INF I.NOM a book (I.NOM)
'I promised John to write a book.' [Her12]

(260) Julia prometió a Marta [encargarse **ella** del asunto].
Julia promised to Martha deal.INF-CL she.NOM of-the matter
'Julia promised Martha that she would deal with the matter.' [PG87]:161

(261) Julia quería [telefonear **ella**].
Julia wanted phone.INF she.NOM
'Julia wanted to be her who phones.' [PG87]:160

Appendix B. Proofs in CatLog2

Portuguese

a : $Nt(s(f))/CNs(f) : \iota$
a : $PPa/\exists a((Si\|Na)\sqcup(\langle\rangle Na\backslash Si)) : \lambda AA$
cedo : $Si\backslash Si : early$
chegar : $\langle\rangle\exists aNa\backslash Si : arrive$
decidiu : $\forall a((Na\backslash Sf)/((Si\|Na)\sqcup(\langle\rangle Na\backslash Si))) : \lambda A\lambda B((decided\ (A\ B))\ B)$
ele : $Nt(s(m))\|Nt(s(m)) : \lambda AA$
eles : $Nt(p(m))\|Nt(p(m)) : \lambda AA$
forçou : $((\exists gNt(s(g))\backslash Sf)/PPa)/\exists aNa : \lambda A\lambda B\lambda C(((force\ (B\ A))\ A)\ C)$
manifestantes : $CNp(m) : protesters$
o : $Nt(s(m))/CNs(m) : \iota$
os : $Nt(p(m))/CNp(m) : \iota$
Pedro : $Nt(s(m)) : p$
polícia : $CNs(f) : police$
problema : $CNs(m) : problem$
quer : $\forall a((Na\backslash Sf)/((Si\|Na)\sqcup(\langle\rangle Na\backslash Si))) : \lambda A\lambda B((want\ (A\ B))\ B)$
resolver : $(Si/\exists aNa)\div\langle\rangle\exists aNa : \lambda A\lambda B((solve\ B)\ A)$
sair : $\langle\rangle\exists aNa\backslash Si : go$

Italian

aprire : $(Si/\exists aNa)\div\langle\rangle\exists aNa : \lambda A\lambda B((open\ B)\ A)$
convegno : $CNs(m) : conference$
deciso : $(\exists aNa\backslash Sf)/(CPdi\sqcup CPche) : \lambda A\lambda B((decided\ (A\ B))\ B)$
di : $CPdi/\exists a((Si\|Na)\sqcup(Na\backslash Si)) : \lambda AA$
Gianni : $Ns(m) : j$
ha : $(\exists aNa\backslash Sf)/(\exists aNa\backslash Sf) : \lambda AA$
il : $Ns(m)/CNs(m) : \iota$
intervenire : $Si/\langle\rangle\exists aNa : intervene$
lui : $Ns(m)\|Ns(m) : \lambda AA$
rettore : $CNs(m) : dean$

Spanish

a : $PPa/\exists aNt(a) : \lambda AA$
deberes : $CNp(m) : homework$
el : $Nt(s(m))/CNt(s(m)) : \iota$
é**l** : $Nt(s(m))\|Nt(s(m)) : \lambda AA$
hacer : $(Si/\exists aNa)/\langle\rangle\exists aNa : \lambda A\lambda B((do\ B)\ A)$
Juan : $Nt(s(m)) : j$
los : $Nt(p(m))/CNp(m) : \iota$
profesor : $CNt(s(m)) : teacher$
prometió : $((\exists gNt(s(g))\backslash Sf)/\exists a((Si\|Na)\sqcup(Na\backslash Si)))/PPa : \lambda A\lambda B\lambda C(((promise\ (B\ C))\ A)\ C)$
su : $\forall g(Nt(s(g))|Nt(s(g)))/\exists gCNt(s(g)) : \lambda A\lambda B((of\ B)\ A)$

(0) **pedro+quer+chegar+cedo** : Sf

$Nt(s(m)) : p, \forall a((Na\backslash Sf)/((Si\|Na)\sqcup(\langle\rangle Na\backslash Si))) : \lambda A\lambda B((\mathit{want}\ (A\ B))\ B), \langle\rangle\exists aNa\backslash Si : \mathit{arrive}, Si\backslash Si : \mathit{early}\ \Rightarrow\ Sf$

$$
\begin{prooftree}
\AxiomC{}
\UnaryInfC{$Nt(s(m)) \Rightarrow Nt(s(m))$}
\RightLabel{$\exists R$}
\UnaryInfC{$Nt(s(m)) \Rightarrow \boxed{\exists aNa}$}
\RightLabel{$\langle\rangle R$}
\UnaryInfC{$[Nt(s(m))] \Rightarrow \boxed{\langle\rangle\exists aNa}$}
\AxiomC{}
\UnaryInfC{$\boxed{Si} \Rightarrow Si$}
\RightLabel{$\backslash L$}
\BinaryInfC{$[Nt(s(m))], \boxed{\langle\rangle\exists aNa\backslash Si} \Rightarrow Si$}
\AxiomC{}
\UnaryInfC{$\boxed{Si} \Rightarrow Si$}
\RightLabel{$\backslash L$}
\BinaryInfC{$[Nt(s(m))], \langle\rangle\exists aNa\backslash Si, \boxed{Si\backslash Si} \Rightarrow Si$}
\RightLabel{$\langle\rangle L$}
\UnaryInfC{$\langle\rangle Nt(s(m)), \langle\rangle\exists aNa\backslash Si, Si\backslash Si \Rightarrow Si$}
\RightLabel{$\backslash R$}
\UnaryInfC{$\langle\rangle\exists aNa\backslash Si, Si\backslash Si \Rightarrow \langle\rangle Nt(s(m))\backslash Si$}
\RightLabel{$\sqcup R$}
\UnaryInfC{$\langle\rangle\exists aNa\backslash Si, Si\backslash Si \Rightarrow \boxed{(Si\|Nt(s(m)))\sqcup(\langle\rangle Nt(s(m))\backslash Si)}$}
\AxiomC{}
\UnaryInfC{$Nt(s(m)) \Rightarrow Nt(s(m))$}
\AxiomC{}
\UnaryInfC{$\boxed{Sf} \Rightarrow Sf$}
\RightLabel{$\backslash L$}
\BinaryInfC{$Nt(s(m)), \boxed{Nt(s(m))\backslash Sf} \Rightarrow Sf$}
\RightLabel{$/L$}
\BinaryInfC{$Nt(s(m)), \boxed{(Nt(s(m))\backslash Sf)/((Si\|Nt(s(m)))\sqcup(\langle\rangle Nt(s(m))\backslash Si))}, \langle\rangle\exists aNa\backslash Si, Si\backslash Si \Rightarrow Sf$}
\RightLabel{$\forall L$}
\UnaryInfC{$Nt(s(m)), \boxed{\forall a((Na\backslash Sf)/((Si\|Na)\sqcup(\langle\rangle Na\backslash Si)))}, \langle\rangle\exists aNa\backslash Si, Si\backslash Si \Rightarrow Sf$}
\end{prooftree}
$$

$((\mathit{want}\ (\mathit{early}\ (\mathit{arrive}\ p)))\ p)$

(1) **pedro+quer+**[**ele**]**+chegar+cedo** : Sf

$Nt(s(m)) : p, \forall a((Na\backslash Sf)/((Si\|Na)\sqcup(\langle\rangle Na\backslash Si))) : \lambda A\lambda B((\mathit{want}\ (A\ B))\ B), [Nt(s(m))\|Nt(s(m)) : \lambda CC], \langle\rangle\exists aNa\backslash Si : \mathit{arrive}, Si\backslash Si : \mathit{early}\ \Rightarrow\ Sf$

$$
\begin{prooftree}
\AxiomC{}
\UnaryInfC{$Nt(s(m)) \Rightarrow Nt(s(m))$}
\RightLabel{$\exists R$}
\UnaryInfC{$Nt(s(m)) \Rightarrow \boxed{\exists aNa}$}
\RightLabel{$\langle\rangle R$}
\UnaryInfC{$[Nt(s(m))] \Rightarrow \boxed{\langle\rangle\exists aNa}$}
\AxiomC{}
\UnaryInfC{$\boxed{Si} \Rightarrow Si$}
\RightLabel{$\backslash L$}
\BinaryInfC{$[Nt(s(m))], \boxed{\langle\rangle\exists aNa\backslash Si} \Rightarrow Si$}
\AxiomC{}
\UnaryInfC{$\boxed{Si} \Rightarrow Si$}
\RightLabel{$\backslash L$}
\BinaryInfC{$[Nt(s(m))], \langle\rangle\exists aNa\backslash Si, \boxed{Si\backslash Si} \Rightarrow Si$}
\RightLabel{$\|R$}
\UnaryInfC{$[\boxed{Nt(s(m))\|Nt(s(m))}], \langle\rangle\exists aNa\backslash Si, Si\backslash Si \Rightarrow Si\|Nt(s(m))$}
\RightLabel{$\sqcup R$}
\UnaryInfC{$[Nt(s(m))\|Nt(s(m))], \langle\rangle\exists aNa\backslash Si, Si\backslash Si \Rightarrow \boxed{(Si\|Nt(s(m)))\sqcup(\langle\rangle Nt(s(m))\backslash Si)}$}
\AxiomC{}
\UnaryInfC{$Nt(s(m)) \Rightarrow Nt(s(m))$}
\AxiomC{}
\UnaryInfC{$\boxed{Sf} \Rightarrow Sf$}
\RightLabel{$\backslash L$}
\BinaryInfC{$Nt(s(m)), \boxed{Nt(s(m))\backslash Sf} \Rightarrow Sf$}
\RightLabel{$/L$}
\BinaryInfC{$Nt(s(m)), \boxed{(Nt(s(m))\backslash Sf)/((Si\|Nt(s(m)))\sqcup(\langle\rangle Nt(s(m))\backslash Si))}, [Nt(s(m))\|Nt(s(m))], \langle\rangle\exists aNa\backslash Si, Si\backslash Si \Rightarrow Sf$}
\RightLabel{$\forall L$}
\UnaryInfC{$Nt(s(m)), \boxed{\forall a((Na\backslash Sf)/((Si\|Na)\sqcup(\langle\rangle Na\backslash Si)))}, [Nt(s(m))\|Nt(s(m))], \langle\rangle\exists aNa\backslash Si, Si\backslash Si \Rightarrow Sf$}
\end{prooftree}
$$

$((\mathit{want}\ (\mathit{early}\ (\mathit{arrive}\ p)))\ p)$

(2) **pedro+decidiu+[ele]+resolver+o+problema** : Sf

$Nt(s(m)) : p, \forall a((Na\backslash Sf)/((Si\|Na)\sqcup(\langle\rangle Na\backslash Si))) : \lambda A\lambda B((decided\ (A\ B))\ B), [Nt(s(m))\|Nt(s(m)) : \lambda CC], (Si/\exists aNa)\div\langle\rangle\exists aNa : \lambda D\lambda E((solve\ E)\ D), Nt(s(m))/CNs(m) : \iota, CNs(m) : problem \ \Rightarrow\ Sf$

$$
\begin{prooftree}
\AxiomC{$Nt(s(m)) \Rightarrow Nt(s(m))$}
\RightLabel{$\exists R$}
\UnaryInfC{$Nt(s(m)) \Rightarrow \boxed{\exists aNa}$}
\RightLabel{$\langle\rangle R$}
\UnaryInfC{$[Nt(s(m))] \Rightarrow \boxed{\langle\rangle\exists aNa}$}
\AxiomC{$CNs(m) \Rightarrow CNs(m)$}
\AxiomC{$\boxed{Nt(s(m))} \Rightarrow Nt(s(m))$}
\RightLabel{$/L$}
\BinaryInfC{$\boxed{Nt(s(m))/CNs(m)}, CNs(m) \Rightarrow Nt(s(m))$}
\RightLabel{$\exists R$}
\UnaryInfC{$Nt(s(m))/CNs(m), CNs(m) \Rightarrow \boxed{\exists aNa}$}
\AxiomC{$\boxed{Si} \Rightarrow Si$}
\RightLabel{$/L$}
\BinaryInfC{$\boxed{Si/\exists aNa}, Nt(s(m))/CNs(m), CNs(m) \Rightarrow Si$}
\RightLabel{$\div L$}
\BinaryInfC{$[Nt(s(m))], \boxed{(Si/\exists aNa)\div\langle\rangle\exists aNa}, Nt(s(m))/CNs(m), CNs(m) \Rightarrow Si$}
\RightLabel{$\|R$}
\UnaryInfC{$[\boxed{Nt(s(m))\|Nt(s(m))}], (Si/\exists aNa)\div\langle\rangle\exists aNa, Nt(s(m))/CNs(m), CNs(m) \Rightarrow Si\|Nt(s(m))$}
\RightLabel{$\sqcup R$}
\UnaryInfC{$[Nt(s(m))\|Nt(s(m))], (Si/\exists aNa)\div\langle\rangle\exists aNa, Nt(s(m))/CNs(m), CNs(m) \Rightarrow \boxed{(Si\|Nt(s(m)))\sqcup(\langle\rangle Nt(s(m))\backslash Si)}$}
\AxiomC{$Nt(s(m)) \Rightarrow Nt(s(m))$}
\AxiomC{$\boxed{Sf} \Rightarrow Sf$}
\RightLabel{$\backslash L$}
\BinaryInfC{$Nt(s(m)), \boxed{Nt(s(m))\backslash Sf} \Rightarrow Sf$}
\RightLabel{$/L$}
\BinaryInfC{$Nt(s(m)), \boxed{(Nt(s(m))\backslash Sf)/((Si\|Nt(s(m)))\sqcup(\langle\rangle Nt(s(m))\backslash Si))}, [Nt(s(m))\|Nt(s(m))], (Si/\exists aNa)\div\langle\rangle\exists aNa, Nt(s(m))/CNs(m), CNs(m) \Rightarrow Sf$}
\RightLabel{$\forall L$}
\UnaryInfC{$Nt(s(m)), \boxed{\forall a((Na\backslash Sf)/((Si\|Na)\sqcup(\langle\rangle Na\backslash Si)))}, [Nt(s(m))\|Nt(s(m))], (Si/\exists aNa)\div\langle\rangle\exists aNa, Nt(s(m))/CNs(m), CNs(m) \Rightarrow Sf$}
\end{prooftree}
$$

$((decided\ ((solve\ (\iota\ problem))\ p))\ p)$

(3) **pedro+decidiu+resolver+[ele]+o+problema** : Sf

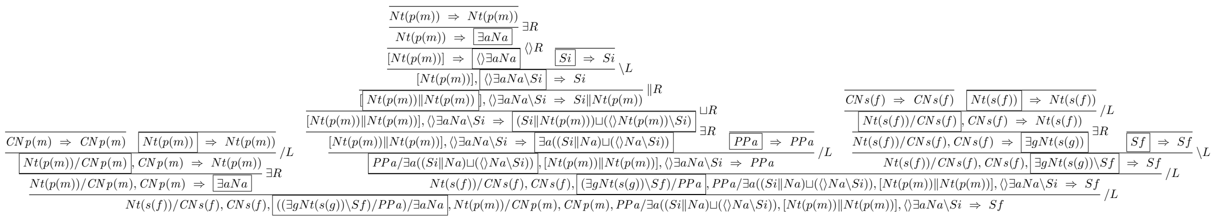
$Nt(s(m)) : p, \forall a((Na\backslash Sf)/((Si\|Na)\sqcup(\langle\rangle Na\backslash Si))) : \lambda A\lambda B((decided\ (A\ B))\ B), (Si/\exists aNa)\div\langle\rangle\exists aNa : \lambda C\lambda D((solve\ D)\ C), [Nt(s(m))\|Nt(s(m)) : \lambda EE], Nt(s(m))/CNs(m) : \iota, CNs(m) : problem \ \Rightarrow\ Sf$

$$
\begin{prooftree}
\AxiomC{$Nt(s(m)) \Rightarrow Nt(s(m))$}
\RightLabel{$\exists R$}
\UnaryInfC{$Nt(s(m)) \Rightarrow \boxed{\exists aNa}$}
\RightLabel{$\langle\rangle R$}
\UnaryInfC{$[Nt(s(m))] \Rightarrow \boxed{\langle\rangle\exists aNa}$}
\AxiomC{$CNs(m) \Rightarrow CNs(m)$}
\AxiomC{$\boxed{Nt(s(m))} \Rightarrow Nt(s(m))$}
\RightLabel{$/L$}
\BinaryInfC{$\boxed{Nt(s(m))/CNs(m)}, CNs(m) \Rightarrow Nt(s(m))$}
\RightLabel{$\exists R$}
\UnaryInfC{$Nt(s(m))/CNs(m), CNs(m) \Rightarrow \boxed{\exists aNa}$}
\AxiomC{$\boxed{Si} \Rightarrow Si$}
\RightLabel{$/L$}
\BinaryInfC{$\boxed{Si/\exists aNa}, Nt(s(m))/CNs(m), CNs(m) \Rightarrow Si$}
\RightLabel{$\div L$}
\BinaryInfC{$\boxed{(Si/\exists aNa)\div\langle\rangle\exists aNa}, [Nt(s(m))], Nt(s(m))/CNs(m), CNs(m) \Rightarrow Si$}
\RightLabel{$\|R$}
\UnaryInfC{$(Si/\exists aNa)\div\langle\rangle\exists aNa, [\boxed{Nt(s(m))\|Nt(s(m))}], Nt(s(m))/CNs(m), CNs(m) \Rightarrow Si\|Nt(s(m))$}
\RightLabel{$\sqcup R$}
\UnaryInfC{$(Si/\exists aNa)\div\langle\rangle\exists aNa, [Nt(s(m))\|Nt(s(m))], Nt(s(m))/CNs(m), CNs(m) \Rightarrow \boxed{(Si\|Nt(s(m)))\sqcup(\langle\rangle Nt(s(m))\backslash Si)}$}
\AxiomC{$Nt(s(m)) \Rightarrow Nt(s(m))$}
\AxiomC{$\boxed{Sf} \Rightarrow Sf$}
\RightLabel{$\backslash L$}
\BinaryInfC{$Nt(s(m)), \boxed{Nt(s(m))\backslash Sf} \Rightarrow Sf$}
\RightLabel{$/L$}
\BinaryInfC{$Nt(s(m)), \boxed{(Nt(s(m))\backslash Sf)/((Si\|Nt(s(m)))\sqcup(\langle\rangle Nt(s(m))\backslash Si))}, (Si/\exists aNa)\div\langle\rangle\exists aNa, [Nt(s(m))\|Nt(s(m))], Nt(s(m))/CNs(m), CNs(m) \Rightarrow Sf$}
\RightLabel{$\forall L$}
\UnaryInfC{$Nt(s(m)), \boxed{\forall a((Na\backslash Sf)/((Si\|Na)\sqcup(\langle\rangle Na\backslash Si)))}, (Si/\exists aNa)\div\langle\rangle\exists aNa, [Nt(s(m))\|Nt(s(m))], Nt(s(m))/CNs(m), CNs(m) \Rightarrow Sf$}
\end{prooftree}
$$

$((decided\ ((solve\ (\iota\ problem))\ p))\ p)$

(4) **a+policia+forcou+os+manifestantes+a+[eles]+sair** : Sf

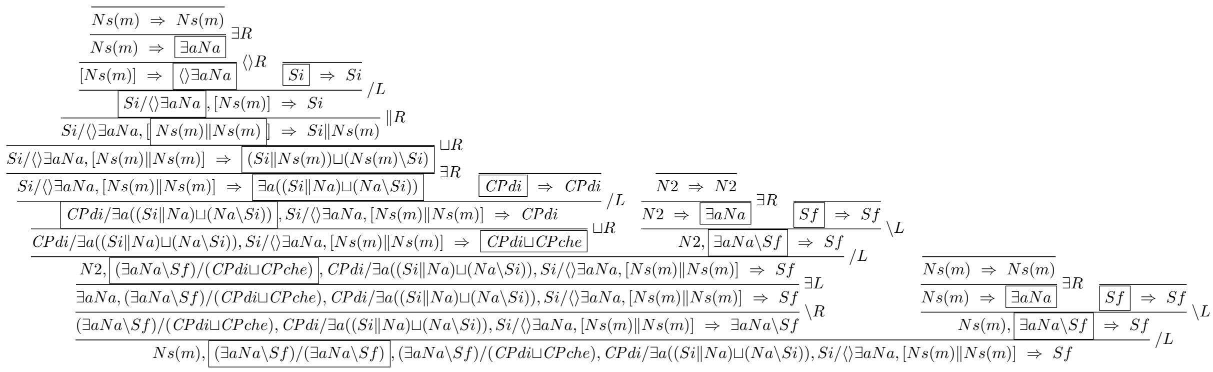
$Nt(s(f))/CNs(f) : \iota, CNs(f) : police, ((\exists gNt(s(g))\backslash Sf)/PPa)/\exists aNa : \lambda A\lambda B\lambda C(((force(B\ A))\ A)\ C), Nt(p(m))/CNp(m) : \iota, CNp(m) : protesters, PPa/\exists a((Si\|Na)\sqcup(\langle\rangle Na\backslash Si)) : \lambda DD, [Nt(p(m))\|Nt(p(m)) : \lambda EE], \langle\rangle\exists aNa\backslash Si : go \ \Rightarrow\ Sf$



$(((force\ (go\ (\iota\ protesters)))\ (\iota\ protesters))\ (\iota\ police))$

(5) **gianni+ha+deciso+di+intervenire+[lui]** : Sf

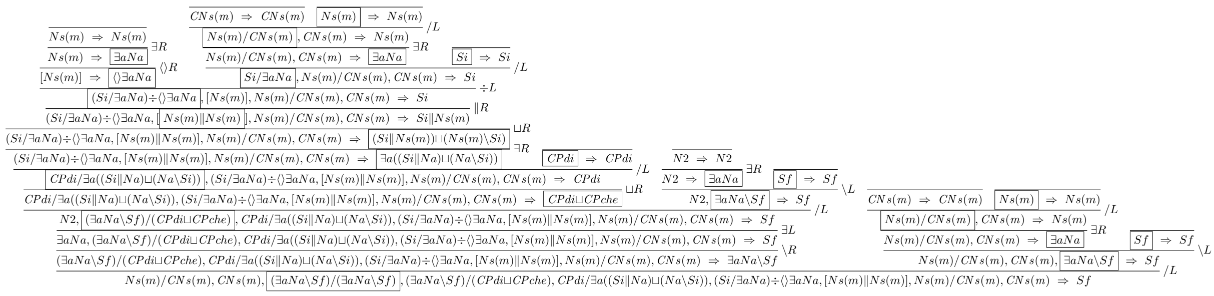
$Ns(m) : j, (\exists aNa\backslash Sf)/(\exists aNa\backslash Sf) : \lambda AA, (\exists aNa\backslash Sf)/(CPdi\sqcup CPche) : \lambda B\lambda C((decided\ (B\ C))\ C), CPdi/\exists a((Si\|Na)\sqcup(Na\backslash Si)) : \lambda DD, Si/\langle\rangle\exists aNa : intervene, [Ns(m)\|Ns(m) : \lambda EE] \ \Rightarrow\ Sf$



$((decided\ (intervene\ j))\ j)$

(6) **il+rettore+ha+deciso+di+aprire+[lui]+il+convegno** : Sf

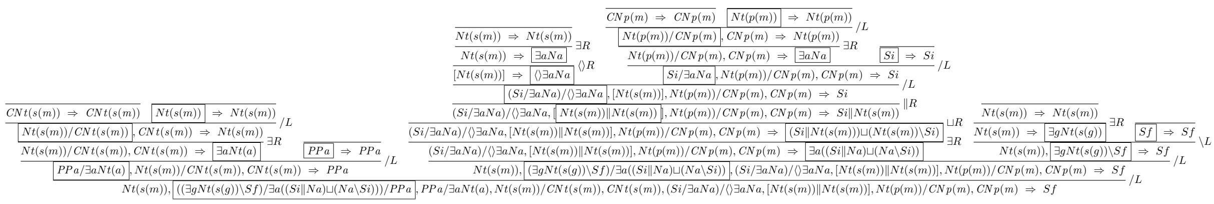
$Ns(m)/CNs(m) : \iota, CNs(m) : dean, (\exists aNa\backslash Sf)/(\exists aNa\backslash Sf) : \lambda AA,$
$(\exists aNa\backslash Sf)/(CPdi\sqcup CPche) : \lambda B\lambda C((decided\ (B\ C))\ C), CPdi/\exists a((Si\|Na)\sqcup(Na\backslash Si)) : \lambda DD,$
$(Si/\exists aNa)\div\langle\rangle\exists aNa\ :\ \lambda E\lambda F((open\ \ F)\ \ E), [Ns(m)\|Ns(m)\ :\ \lambda GG], Ns(m)/CNs(m)\ :\ \iota, CNs(m)\ :$
$conference\ \Rightarrow\ Sf$



$((decided\ ((open\ (\iota\ conference))\ (\iota\ dean)))\ (\iota\ dean))$

(7) **juan+prometio+a+el+profesor+hacer+[el]+los+deberes** : Sf

$Nt(s(m)) : j, ((\exists gNt(s(g))\backslash Sf)/\exists a((Si\|Na)\sqcup(Na\backslash Si)))/PPa :$
$\lambda A\lambda B\lambda C(((promise\ (B\ C))\ A)\ C), PPa/\exists aNt(a) : \lambda DD, Nt(s(m))/CNt(s(m)) : \iota, CNt(s(m)) : teacher,$
$(Si/\exists aNa)/\langle\rangle\exists aNa : \lambda E\lambda F((do\ \ F)\ \ E), [Nt(s(m))\|Nt(s(m)) : \lambda GG], Nt(p(m))/CNp(m) : \iota, CNp(m) :$
$homework\ \Rightarrow\ Sf$



$(((promise\ ((do\ (\iota\ homework))\ j))\ (\iota\ teacher))\ j)$

(8) **juan+prometio+a+su+profesor+hacer+[el]+los+deberes** : Sf

$Nt(s(m)) : j, ((\exists gNt(s(g))\backslash Sf)/\exists a((Si\|Na)\sqcup(Na\backslash Si)))/PPa :$
$\lambda A\lambda B\lambda C(((promise\ (B\ C))\ A)\ C), PPa/\exists aNt(a) : \lambda DD,$
$\forall g(Nt(s(g))\|Nt(s(g)))/\exists gCNt(s(g)) : \lambda E\lambda F((of\ F)\ E), CNt(s(m)) : teacher,$
$(Si/\exists aNa)/\langle\rangle\exists aNa : \lambda G\lambda H((do\ H)\ G), [Nt(s(m))\|Nt(s(m)) : \lambda II],$
$Nt(p(m))/CNp(m) : \iota, CNp(m) : homework\ \Rightarrow\ Sf$

$$
\begin{prooftree}
\AxiomC{}
\UnaryInfC{$CNt(s(m)) \Rightarrow CNt(s(m))$}
\RightLabel{$\exists R$}
\UnaryInfC{$CNt(s(m)) \Rightarrow \boxed{\exists gCNt(s(g))}$}
\AxiomC{}
\UnaryInfC{$Nt(s(m)) \Rightarrow Nt(s(m))$}
\AxiomC{}
\UnaryInfC{$Nt(s(m)) \Rightarrow Nt(s(m))$}
\RightLabel{$\exists R$}
\UnaryInfC{$Nt(s(m)) \Rightarrow \boxed{\exists aNt(a)}$}
\AxiomC{}
\UnaryInfC{$\boxed{PPa} \Rightarrow PPa$}
\RightLabel{$/L$}
\BinaryInfC{$\boxed{PPa/\exists aNt(a)}, Nt(s(m)) \Rightarrow PPa$}
\AxiomC{}
\UnaryInfC{$Nt(s(m)) \Rightarrow Nt(s(m))$}
\RightLabel{$\exists R$}
\UnaryInfC{$Nt(s(m)) \Rightarrow \boxed{\exists aNa}$}
\RightLabel{$\langle\rangle R$}
\UnaryInfC{$[Nt(s(m))] \Rightarrow \boxed{\langle\rangle\exists aNa}$}
\AxiomC{}
\UnaryInfC{$CNp(m) \Rightarrow CNp(m)$}
\AxiomC{}
\UnaryInfC{$\boxed{Nt(p(m))} \Rightarrow Nt(p(m))$}
\RightLabel{$/L$}
\BinaryInfC{$\boxed{Nt(p(m))/CNp(m)}, CNp(m) \Rightarrow Nt(p(m))$}
\RightLabel{$\exists R$}
\UnaryInfC{$Nt(p(m))/CNp(m), CNp(m) \Rightarrow \boxed{\exists aNa}$}
\AxiomC{}
\UnaryInfC{$\boxed{Si} \Rightarrow Si$}
\RightLabel{$/L$}
\BinaryInfC{$\boxed{Si/\exists aNa}, Nt(p(m))/CNp(m), CNp(m) \Rightarrow Si$}
\RightLabel{$/L$}
\BinaryInfC{$\boxed{(Si/\exists aNa)/\langle\rangle\exists aNa}, [Nt(s(m))], Nt(p(m))/CNp(m), CNp(m) \Rightarrow Si$}
\RightLabel{$\|R$}
\UnaryInfC{$(Si/\exists aNa)/\langle\rangle\exists aNa, [\boxed{Nt(s(m))\|Nt(s(m))}], Nt(p(m))/CNp(m), CNp(m) \Rightarrow Si\|Nt(s(m))$}
\RightLabel{$\sqcup R$}
\UnaryInfC{$(Si/\exists aNa)/\langle\rangle\exists aNa, [Nt(s(m))\|Nt(s(m))], Nt(p(m))/CNp(m), CNp(m) \Rightarrow \boxed{(Si\|Nt(s(m)))\sqcup(Nt(s(m))\backslash Si)}$}
\RightLabel{$\exists R$}
\UnaryInfC{$(Si/\exists aNa)/\langle\rangle\exists aNa, [Nt(s(m))\|Nt(s(m))], Nt(p(m))/CNp(m), CNp(m) \Rightarrow \boxed{\exists a((Si\|Na)\sqcup(Na\backslash Si))}$}
\AxiomC{}
\UnaryInfC{$Nt(s(m)) \Rightarrow Nt(s(m))$}
\RightLabel{$\exists R$}
\UnaryInfC{$Nt(s(m)) \Rightarrow \boxed{\exists gNt(s(g))}$}
\AxiomC{}
\UnaryInfC{$\boxed{Sf} \Rightarrow Sf$}
\RightLabel{$\backslash L$}
\BinaryInfC{$Nt(s(m)), \boxed{\exists gNt(s(g))\backslash Sf} \Rightarrow Sf$}
\RightLabel{$/L$}
\BinaryInfC{$Nt(s(m)), \boxed{(\exists gNt(s(g))\backslash Sf)/\exists a((Si\|Na)\sqcup(Na\backslash Si))}, (Si/\exists aNa)/\langle\rangle\exists aNa, [Nt(s(m))\|Nt(s(m))], Nt(p(m))/CNp(m), CNp(m) \Rightarrow Sf$}
\RightLabel{$/L$}
\BinaryInfC{$Nt(s(m)), \boxed{((\exists gNt(s(g))\backslash Sf)/\exists a((Si\|Na)\sqcup(Na\backslash Si)))/PPa}, PPa/\exists aNt(a), Nt(s(m)), (Si/\exists aNa)/\langle\rangle\exists aNa, [Nt(s(m))\|Nt(s(m))], Nt(p(m))/CNp(m), CNp(m) \Rightarrow Sf$}
\RightLabel{$\|L$}
\BinaryInfC{$Nt(s(m)), ((\exists gNt(s(g))\backslash Sf)/\exists a((Si\|Na)\sqcup(Na\backslash Si)))/PPa, PPa/\exists aNt(a), \boxed{Nt(s(m))\|Nt(s(m))}, (Si/\exists aNa)/\langle\rangle\exists aNa, [Nt(s(m))\|Nt(s(m))], Nt(p(m))/CNp(m), CNp(m) \Rightarrow Sf$}
\RightLabel{$\forall L$}
\UnaryInfC{$Nt(s(m)), ((\exists gNt(s(g))\backslash Sf)/\exists a((Si\|Na)\sqcup(Na\backslash Si)))/PPa, PPa/\exists aNt(a), \boxed{\forall g(Nt(s(g))\|Nt(s(g)))}, (Si/\exists aNa)/\langle\rangle\exists aNa, [Nt(s(m))\|Nt(s(m))], Nt(p(m))/CNp(m), CNp(m) \Rightarrow Sf$}
\RightLabel{$/L$}
\BinaryInfC{$Nt(s(m)), ((\exists gNt(s(g))\backslash Sf)/\exists a((Si\|Na)\sqcup(Na\backslash Si)))/PPa, PPa/\exists aNt(a), \boxed{\forall g(Nt(s(g))\|Nt(s(g)))/\exists gCNt(s(g))}, CNt(s(m)), (Si/\exists aNa)/\langle\rangle\exists aNa, [Nt(s(m))\|Nt(s(m))], Nt(p(m))/CNp(m), CNp(m) \Rightarrow Sf$}
\end{prooftree}
$$

$(((promise\ ((do\ (\iota\ homework))\ j))\ ((of\ j)\ teacher))\ j)$

(9) **juan+prometio+a+su+profesor+hacer+[el]+los+deberes** : $Sf\|Nt(s(g))$

$Nt(s(m)) : j, ((\exists gNt(s(g))\backslash Sf)/\exists a((Si\|Na)\sqcup(Na\backslash Si)))/PPa :$
$\lambda A\lambda B\lambda C(((promise\ (B\ C))\ A)\ C), PPa/\exists aNt(a) : \lambda DD,$
$\forall g(Nt(s(g))\|Nt(s(g)))/\exists gCNt(s(g)) : \lambda E\lambda F((of\ F)\ E), CNt(s(m)) : teacher,$
$(Si/\exists aNa)/\langle\rangle\exists aNa : \lambda G\lambda H((do\ H)\ G), [Nt(s(m))\|Nt(s(m)) : \lambda II],$
$Nt(p(m))/CNp(m) : \iota, CNp(m) : homework\ \Rightarrow\ Sf\|Nt(s(g))$

$$
\begin{prooftree}
\AxiomC{}
\UnaryInfC{$CNt(s(m)) \Rightarrow CNt(s(m))$}
\RightLabel{$\exists R$}
\UnaryInfC{$CNt(s(m)) \Rightarrow \boxed{\exists gCNt(s(g))}$}
\AxiomC{}
\UnaryInfC{$Nt(s(g)) \Rightarrow Nt(s(g))$}
\RightLabel{$\exists R$}
\UnaryInfC{$Nt(s(g)) \Rightarrow \boxed{\exists aNt(a)}$}
\AxiomC{}
\UnaryInfC{$\boxed{PPa} \Rightarrow PPa$}
\RightLabel{$/L$}
\BinaryInfC{$\boxed{PPa/\exists aNt(a)}, Nt(s(g)) \Rightarrow PPa$}
\AxiomC{}
\UnaryInfC{$Nt(s(m)) \Rightarrow Nt(s(m))$}
\RightLabel{$\exists R$}
\UnaryInfC{$Nt(s(m)) \Rightarrow \boxed{\exists aNa}$}
\RightLabel{$\langle\rangle R$}
\UnaryInfC{$[Nt(s(m))] \Rightarrow \boxed{\langle\rangle\exists aNa}$}
\AxiomC{}
\UnaryInfC{$CNp(m) \Rightarrow CNp(m)$}
\AxiomC{}
\UnaryInfC{$\boxed{Nt(p(m))} \Rightarrow Nt(p(m))$}
\RightLabel{$/L$}
\BinaryInfC{$\boxed{Nt(p(m))/CNp(m)}, CNp(m) \Rightarrow Nt(p(m))$}
\RightLabel{$\exists R$}
\UnaryInfC{$Nt(p(m))/CNp(m), CNp(m) \Rightarrow \boxed{\exists aNa}$}
\AxiomC{}
\UnaryInfC{$\boxed{Si} \Rightarrow Si$}
\RightLabel{$/L$}
\BinaryInfC{$\boxed{Si/\exists aNa}, Nt(p(m))/CNp(m), CNp(m) \Rightarrow Si$}
\RightLabel{$/L$}
\BinaryInfC{$\boxed{(Si/\exists aNa)/\langle\rangle\exists aNa}, [Nt(s(m))], Nt(p(m))/CNp(m), CNp(m) \Rightarrow Si$}
\RightLabel{$\|R$}
\UnaryInfC{$(Si/\exists aNa)/\langle\rangle\exists aNa, [\boxed{Nt(s(m))\|Nt(s(m))}], Nt(p(m))/CNp(m), CNp(m) \Rightarrow Si\|Nt(s(m))$}
\RightLabel{$\sqcup R$}
\UnaryInfC{$(Si/\exists aNa)/\langle\rangle\exists aNa, [Nt(s(m))\|Nt(s(m))], Nt(p(m))/CNp(m), CNp(m) \Rightarrow \boxed{(Si\|Nt(s(m)))\sqcup(Nt(s(m))\backslash Si)}$}
\RightLabel{$\exists R$}
\UnaryInfC{$(Si/\exists aNa)/\langle\rangle\exists aNa, [Nt(s(m))\|Nt(s(m))], Nt(p(m))/CNp(m), CNp(m) \Rightarrow \boxed{\exists a((Si\|Na)\sqcup(Na\backslash Si))}$}
\AxiomC{}
\UnaryInfC{$Nt(s(m)) \Rightarrow Nt(s(m))$}
\RightLabel{$\exists R$}
\UnaryInfC{$Nt(s(m)) \Rightarrow \boxed{\exists gNt(s(g))}$}
\AxiomC{}
\UnaryInfC{$\boxed{Sf} \Rightarrow Sf$}
\RightLabel{$\backslash L$}
\BinaryInfC{$Nt(s(m)), \boxed{\exists gNt(s(g))\backslash Sf} \Rightarrow Sf$}
\RightLabel{$/L$}
\BinaryInfC{$Nt(s(m)), \boxed{(\exists gNt(s(g))\backslash Sf)/\exists a((Si\|Na)\sqcup(Na\backslash Si))}, (Si/\exists aNa)/\langle\rangle\exists aNa, [Nt(s(m))\|Nt(s(m))], Nt(p(m))/CNp(m), CNp(m) \Rightarrow Sf$}
\RightLabel{$/L$}
\BinaryInfC{$Nt(s(m)), \boxed{((\exists gNt(s(g))\backslash Sf)/\exists a((Si\|Na)\sqcup(Na\backslash Si)))/PPa}, PPa/\exists aNt(a), Nt(s(g)), (Si/\exists aNa)/\langle\rangle\exists aNa, [Nt(s(m))\|Nt(s(m))], Nt(p(m))/CNp(m), CNp(m) \Rightarrow Sf$}
\RightLabel{$\|R$}
\UnaryInfC{$Nt(s(m)), ((\exists gNt(s(g))\backslash Sf)/\exists a((Si\|Na)\sqcup(Na\backslash Si)))/PPa, PPa/\exists aNt(a), \boxed{Nt(s(g))\|Nt(s(g))}, (Si/\exists aNa)/\langle\rangle\exists aNa, [Nt(s(m))\|Nt(s(m))], Nt(p(m))/CNp(m), CNp(m) \Rightarrow Sf\|Nt(s(g))$}
\RightLabel{$\forall L$}
\UnaryInfC{$Nt(s(m)), ((\exists gNt(s(g))\backslash Sf)/\exists a((Si\|Na)\sqcup(Na\backslash Si)))/PPa, PPa/\exists aNt(a), \boxed{\forall g(Nt(s(g))\|Nt(s(g)))}, (Si/\exists aNa)/\langle\rangle\exists aNa, [Nt(s(m))\|Nt(s(m))], Nt(p(m))/CNp(m), CNp(m) \Rightarrow Sf\|Nt(s(g))$}
\RightLabel{$/L$}
\BinaryInfC{$Nt(s(m)), ((\exists gNt(s(g))\backslash Sf)/\exists a((Si\|Na)\sqcup(Na\backslash Si)))/PPa, PPa/\exists aNt(a), \boxed{\forall g(Nt(s(g))\|Nt(s(g)))/\exists gCNt(s(g))}, CNt(s(m)), (Si/\exists aNa)/\langle\rangle\exists aNa, [Nt(s(m))\|Nt(s(m))], Nt(p(m))/CNp(m), CNp(m) \Rightarrow Sf\|Nt(s(g))$}
\end{prooftree}
$$

$\lambda A(((promise\ ((do\ (\iota\ homework))\ j))\ ((of\ A)\ teacher))\ j)$