Overtly Anaphoric Control
in Type Logical Grammar

María Inés Corbalán[1] and Glyn Morrill[2]

[1]Universidade Estadual de Campinas, São Paulo, Brazil
inescorbalan@yahoo.com.ar
[2]Universitat Politècnica de Catalunya, Barcelona, Spain
morrill@cs.upc.edu

Abstract. In this paper we analyse anaphoric pronouns in control sentences and we investigate the implications of these kinds of sentences in relation to the Propositional Theory versus Property Theory question. For these purposes, we invoke the categorial calculus with limited contraction, a conservative extension of Lambek calculus that builds contraction into the logical rules for a customized slash type-constructor.

Keywords: Jaeger's Calculus **LLC** · Obligatory Control · Portuguese · Pronominal Types · Pronouns · Type Logical Grammar

1   Introduction

In Type Logical Grammar (TLG) the analysis of an expression is a resource-conscious proof. Anaphora represents a particular challenge to this approach in that the antecedent resource is multiplied in the semantics. This duplication, which corresponds logically to the structural rule of Contraction, may be treated lexically or syntactically.

Prototypical cases of control, as exemplified in (1) and (2) below, involve a single overt nominal in the subject or object position of the main clause that appears to carry a subject semantic role in the embedded clause (cf. [9]).[1]

(1) The doctor condescended to examine John.

---

[1] This paper will deal only with what is called obligatory control (or exhaustive control in the nomenclature of [14]), cases where the controller must be, for each verb, a NP in a particular syntactic position in the sentence, not with types of arbitrary control (cf. [10], [25], [24]). An anonymous reviewer observes that our approach seems to provide resources to treat also cases of split control, where both the subject and the object matrix clause can jointly form the controlled embedded subject, as in the following Portuguese examples, where the inflection on the infinitive form (INFL) marks the plural predication:

    i. Eu convenci a Maria a/de viaja<u>rmos</u>.
       'I convince Mary to travel.INFL.'
    ii. João prometeu ao seu filho i<u>rem</u> ao cinema (juntos).
       'John promised his son to go.INFL to the cinema (together).'

(2) Barnett persuaded the doctor to examine Tilman.

Therefore, in the prototypical cases it seems that there is a mismatch between the syntactic and the semantic levels of representation: there is no overt nominal in the surface subject position of the embedded clause that carries the corresponding semantic role.

In most generative theories it is assumed that a deleted copy of the overt nominal or PRO occupies the embedded subject position in some non-phonological level of syntactic representation (cf. [23], [7,8], [11], [15]). Assuming this, most generative theories contend that the embedded clause in control structures denotes a proposition (but see [25], [15]). Hence, from these perspectives, the syntactic-semantic mismatch is resolved.

Since categorial grammar is a monostratal framework, the resort to a non-surface level of syntactic representation to avoid the mismatch is not available. Notwithstanding, prototypical obligatory control structures do not present a resource problem in a monostratal grammar such as TLG when there is no syntactic embedded subject controlled by a matrix constituent. From the Type Logical point of view, embedded subjectless clauses in control structures denote a property, not a proposition (cf. [6], [10]), and the lexical semantics of a control verb is multiple-bind. Hence, from a categorial perspective, the syntactic-semantic mismatch is resolved in this way. Treating the control complement as a property accounts for the sloppy reading in inferences from ellipsis of VP and quantification (cf. [6]), while treating it as a proposition does not.

Neverthelesss, if we assume the Property Theory the mismatch seems to reappear in a special kind of control structure in some pro-drop languages. The occurrence of overt, semantically controlled, pronouns in some pro-drop languages, as exemplified in (3)-(8) below, raises the question of reusing semantic resources in the context of control structures and raises the issue of the denotation of the controlled complement clause (cf. [4], [16], among others):[2]

(3) Pedro quer **ele** chegar (cedo). (BP)
    'Peter wants to arrive (early).'

There is not complete agreement in the literature as to whether split control is a type of obligatory control. In footnote 17 we will show how can we deal with split antecedents. As this reviewer notes such cases of control were discussed in the LFG Glue framework. Indeed, [2], following [5], observes that anaphoric control, but not functional control, allows split controlled antecedents. In future research we hope to compare our proposal for control with the one made in the related resource-sensitive formalism of LFG [2,3], which we unfortunately do not have space to discuss here.

[2] BP stands for Brazilian Portuguese, EP for European Portuguese, SPA for Spanish and IT for Italian. The infinitival subjects are highlighted. As is well known, Portuguese has an inflected form of infinitive (INFL), that is, an infinitive form that carries ending marks of agreement with their subject in both person and number. It is generally assumed that inflection is obligatory if there is an overt subject within the infinitive clause. And it is also generally assumed that the inflection must be deleted in cases where the reference of the (null) subject coincides with the reference of a matrix constituent. Thus, both rules confront each other in cases of control sentences with overt subjects.

(4) O João decidiu resolver **ele** o problema. (EP)
'John decided to solve the problem by himself.'

(5) A polícia forçou os manifestantes a **eles** saír(em). (BP)
'The police forced the protesters to leave.'

(6) María quería telefonear **ella**. (SPA)
'Mary wanted to phone.'

(7) Juan prometió a su profesor hacer **él** los deberes. (SPA)
'John promised his teacher to do the homework (personally).'

(8) Gianni me ha promesso di farlo **lui**. (IT)
'John promised me to do it (personally).'

Indeed, since there is an overt pronoun in the (pre- or post-verbal) subject position of the embedded clause, but the embedded clause denotes a property, we seem to have to assume that the denotation of the overt pronoun does not saturate the embedded predicate.

In this paper we analyse anaphoric pronouns in control sentences.[3] For these purposes, we invoke a categorial calculus with limited contraction, a conservative extension of Lambek calculus, that builds contraction into the logical rules for a customized slash type-constructor [13].

The structure of the paper is as follows. Section 2 presents Jaeger's system (in a Gentzen sequent format).[4] Section 3 considers an analysis following Jaeger's

---

[3] An anonymous reviewer indicates that the so-called Richard constructions, as exemplified below, have a number of suggestive parallel features with the variety of control which we deal with. Despite some similarities, it is important to note that Richard constructions are cases of copy raising; raising verbs, unlike control verbs, do not select for a *thematic subject* in the predicative complement, and copy raising verbs, unlike our control examples, *require* a pronominal bound copy in their complement clause:

    i. Richard seems like he is ill.
    ii. Richard seems like he is in trouble.

The phenomenon of copy raising is also attested in Portuguese. But two differences between typical Richard constructions and cases of copy raising in Portuguese must be pointed out: Firstly, the lexical copy is not obligatory in the Portuguese constructions; and secondly, the embedded copy in Portuguese can be not only a pronoun but also a lexical DP (cf. [4]):

    i. Acabou por ir **ele**/**o João** ao mercado.
      'It ended up being the case that he/John went to the market.'

For a treatment of copy raising in a resource-conscious framework or in a generative grammar we refer the interested reader to [1] and [21], respectively.

[4] A anonymous reviewer objects to the use of Gentzen format as "about as unfriendly as possible". Gentzen calculus, labelled and unlabelled natural deductions, proof

approach to pronouns and discusses some difficulties of it for the case of overt controlled pronouns. Section 4 presents an extension of Jaeger's system and develops our proposal. Section 5 concludes the paper. The Appendix contains a sample of the output generated by a version of the parser/theorem-prover CatLog2 (`www.cs.upc.edu/~droman/index.php`) for our final proposal.

2   LLC calculus

The Lambek calculus (**L**) with Limited Contraction (**LLC**) proposed by Jaeger [13] is a conservative extension of the Lambek-style core of TLG. In a nutshell, **LLC** extends **L** with a third kind of implication type-constructor, which compiles a limited version of the structural rule of Contraction into its left and right logical rules. Jaeger's calculus treats resource multiplication syntactically. Like Lambek calculus, Jaeger's calculus **LLC** is free of structural rules.[5]

Definition 1 (syntactic types of LLC). *Where P is a set of basic types, the set F of types of **LLC** is defined as follows:*

$$F ::= \ P \mid F\backslash F \mid F/F \mid F \bullet F \mid F|F$$

Definition 2 (semantic types). *The set T of semantic types is defined on the basis of a set δ of primitive semantic types by:*

$$T ::= \ \delta \mid T \,\&\, T \mid T \rightarrow T$$

As in **L**, the product type-constructor is semantically interpreted as Cartesian product and the implications, as function space formation. So, the category-to-type correspondence for **LLC** is given as follows:

Definition 3 (semantic type map for LLC). *The semantic type map for **LLC** is a mapping τ from syntactic types F to semantic types T such that:*

$$\tau\,(A \bullet B) = \tau\,(A) \,\&\, \tau\,(B)$$
$$\tau\,(A\backslash B) = \tau\,(B/A) = \tau\,(B|A) = \tau\,(A) \rightarrow \tau\,(B)$$

The sequent rules for the product and the slash connectives are as in Lambek calculus. The left and right rules for Jaeger's type slash | are as follows:

---

nets, categorical calculus, etc. are all of repute, all have their respective advantages and disadvantages, and are all notations for the same theory. We think that it is better to try to understand each notation than to censure one. The Gentzen format is really not so hard to read.

[5] Both systems also admit the Cut rule, i.e. adding the Cut rule does not give rise to any new theorems.

$$\frac{\Gamma \Rightarrow M : A \qquad \Delta\,(x : A; y : B) \Rightarrow N : C}{\Delta\,(\Gamma; z : B|A) \Rightarrow N\,[M/x]\,[(z\ M)/y)] : C}\ _{|L}$$

$$\frac{\Gamma\,(x_1 : C_1; \ldots; x_n : C_n) \Rightarrow M : B}{\Gamma\,(y_1 : C_1|A; \ldots; y_n : C_n|A) \Rightarrow \lambda z.M\,[(y_1\ z)/x_1]\ldots[(y_n\ z)/x_n] : B|A}\ _{|R}$$

Fig. 1. Left and Right rules for $|$

The **LLC** calculus is designed to treat different linguistic phenomena related to anaphora and thus to semantic resource multiplication. Jaeger uses his calculus to treat cases of personal pronouns bound by wh-operators and quantifiers, and reflexives and pronouns in ellipsis of VP, among other linguistic phenomena. In **LLC** anaphoric expressions are assigned a type $B|A$ and, in particular, (personal, possesive, reflexive) pronouns are assigned the syntactic category $n|n$. In semantic terms, a pronoun denotes the identity function $\lambda x.x$ over individuals; the reference of a pronoun is identical to the reference of its antecedent.[6]

As a basic example of application of **LLC**, consider the free and bound reading of the personal pronoun in the sentence in (9):

(9) John said he walks.

In one reading, the pronoun *he* is co-referential with the subject *John*; in the other, the pronoun remains free. In the first case, the category of the clause is s with the semantics $((say'\ (walk'\ j'))\ j')$; in the other case, the category is $s|n$ and the corresponding semantics is the function $\lambda x.((say'\ (walk'\ x))\ j')$. Figure 2 outlines these two derivations.

$$\frac{n \Rightarrow n \quad \genfrac{}{}{0pt}{}{\vdots}{n,(n\backslash s)\,/s,n,n\backslash s \Rightarrow s}}{n,(n\backslash s)\,/s,n|n,n\backslash s \Rightarrow s}\ _{|L} \qquad \frac{\genfrac{}{}{0pt}{}{\vdots}{n,(n\backslash s)\,/s,n,n\backslash s \Rightarrow s}}{n,(n\backslash s)\,/s,n|n,n\backslash s \Rightarrow s|n}\ _{|R}$$

Fig. 2. Derivations for *John said he walks*

The free and the bound readings for the pronoun can also be obtained when the matrix subject is a quantifier like *everyone*.[7]

Since we defend the Property Theory as the correct semantic analysis for the infinitive clause selected by a control verb, the syntactic category of this clause cannot be the simple type s despite the overt occurrence of a pronoun. Maintaining the Property Theory and the correct sloppy reading in ellipsis of VP, we test **LLC** in relation to anaphoric pronouns in control sentences.

---

[6] In this respect, Jaeger adopts Jacobson's proposal [12].

[7] Jaeger's proposal in itself does not capture either Binding Principle A (locality of anaphors) or Principle B (antilocality of personal pronouns). A categorial approach to locality of anaphors is given by modalities in [17] and a 'negation as failure' categorial approach including antilocality of personal pronouns is given in [20].

3   Control Structures and Property Theory

3.1   Applying LLC to Portuguese

Besides assuming Jaeger's proposal for pronouns, we use his slash to categorize the matrix control verb. The Spanish control sentence in (6) with a (post-verbal) overt pronoun can be derived assuming the following lexical assignments:

maría $: n : m'$
quería $: (n\backslash s) / (s|n) : \lambda x.\lambda y.\, ((wanted'\ (x\ y))\ y)$
telefonear $: s/n : \lambda x.\, (phone'\ x)$
ella $: n|n : \lambda x.x$

In words, the control verb *quería* 'wanted' is assigned a functional type that takes an unsaturated sentence with a pronominal gap as its complement. The lambda operator that binds two occurrences of the same variable guarantees the control relation between the matrix subject and the embedded pronoun.[8]

Figure 3 shows the derivation of the Spanish sentence in (6) and Figure 4 below shows the derivation of the control sentence in (3) containing a pre-verbal overt pronoun. Observe the nominal argument position in the type assigned to the infinitive embedded verb *chegar* 'to arrive'. Lexical assignments for the nominal, the pronoun and the matrix finite verb are as before.

$$\frac{\dfrac{\overline{n \Rightarrow n}\quad \overline{s \Rightarrow s}}{\dfrac{s/n, n \Rightarrow s}{\dfrac{s/n, n|n \Rightarrow s|n}{n, (n\backslash s)/(s|n), s/n, n|n \Rightarrow s}\, {}^{|R}}\, {}^{/L}}\quad \dfrac{\overline{n \Rightarrow n}\quad \overline{s \Rightarrow s}}{n, n\backslash s \Rightarrow s}\, {}^{\backslash L}}{n, (n\backslash s)/(s|n), s/n, n|n \Rightarrow s}\, {}^{/L}$$

Fig. 3. Derivation of *María quería telefonear **ella***

$$\frac{\dfrac{\overline{n \Rightarrow n}\quad \overline{s \Rightarrow s}}{\dfrac{n, n\backslash s \Rightarrow s}{\dfrac{n|n, n\backslash s \Rightarrow s|n}{n, (n\backslash s)/(s|n), n|n, n\backslash s \Rightarrow s}\, {}^{|R}}\, {}^{\backslash L}}\quad \dfrac{\overline{n \Rightarrow n}\quad \overline{s \Rightarrow s}}{n, n\backslash s \Rightarrow s}\, {}^{\backslash L}}{n, (n\backslash s)/(s|n), n|n, n\backslash s \Rightarrow s}\, {}^{/L}$$

Fig. 4. Derivation of *Pedro quer **ele** chegar*

---

[8] In SPA and IT subjects of (adverbial or subject) infinitive constructions necessarily occupy the post-verbal position. By contrast, in BP such subjects normally occupy the pre-verbal position. In EP subjects within infinitive clauses normally occur in the post-verbal position, but the pre-verbal position can also be admitted. Hence, the nominal argument position in the embedded verb in control structures is justified:

  i. Al sentir **él** los primeros síntomas de la gripe, Carlos se vacunó. (SPA)
     'When he feel.INF the flu symptoms, Carlos gets the vaccine.'
 ii. **Os meninos** sairem à noite preocupa suas mães. (BP/EP)
     'The boys go out.INFL at night worries his mothers.'
iii. Prima di morire **papà**, mama era felice. (IT)
     'Before die.INF dad, mom was happy.'

As we can see from Figure 5 below, the sequent $n, (n\backslash s)/(s|n), n|n, n\backslash s \Rightarrow s|n$ is not derivable. In words, in embedded sentences selected by control verbs the pronominal gap cannot be free and it has to be syntactically bound by a matrix nominal (if there is one).[9]

$$\cfrac{\cfrac{\quad}{n, n\backslash s \Rightarrow s|n}\ ^{*} \quad \cfrac{\vdots}{n, n\backslash s \Rightarrow s}}{\cfrac{n, (n\backslash s)/(s|n), n, n\backslash s \Rightarrow s}{n, (n\backslash s)/(s|n), n|n, n\backslash s \Rightarrow s|n}\ ^{|R}}\ ^{\backslash L}\ ^{/L}$$

Fig. 5. Illicit derivation for the type $s|n$

To sum up, it seems that applying Jaeger's proposal for pronouns is theoretically and empirically adequate for the analysis of overt pronouns occurring within control sentences in some pro-drop languages: we have used it to derive control sentences with overt pronominal subject warranting the control relation and also the Property Theory.

Notwithstanding, the previous proposal faces two adverse problems: over-generation and undergeneration. On the one hand, we cannot prove prototypical cases of control, as in (10), that do not contain an overt embedded subject.

(10) María quería telefonear.
      'Mary wanted to make a phone call.'

$$\cfrac{\cfrac{\quad}{n\backslash s \Rightarrow s|n}\ ^{*} \quad \cfrac{\cfrac{\quad}{n \Rightarrow n}\ \cfrac{\quad}{s \Rightarrow s}}{n, n\backslash s \Rightarrow s}\ ^{\backslash L}}{n, (n\backslash s)/(s|n), n\backslash s \Rightarrow s}\ ^{/L}$$

Fig. 6. Illicit derivation for *María quería telefonear*

On the other, we can derive several ungrammatical sentences, as exemplified in (11-13) below, containing a non-controlled subject expression within the complement clause selected by the subject control verb *quer* 'wants'.[10]

---

[9] If the controller were a pronoun, as in the example below, then the complex category $s|n$ can be derived, but this is in virtue of the matrix subject pronoun.

  i. Roberto, eu tentei **eu** enviar meu convite a você.
     'Robert, I tried to send my invitation to you.'

[10] As in other Romance languages, object pronouns in Portuguese take the clitic form: $(l)o/(l)a$. In Brazilian spoken language the third person (non-reflexive) clitics are not commonly used; instead, the (nominative) form $ele(s)/a(s)$ is usually used for accusative object:

  i. Visitei-o ontem. (EP)
  ii. Visitei ele ontem. (BP)
     '[I] visited him yesterday.'

(11) *Pedro$_1$ quer **ele**$_2$ ajudá-lo$_1$/ele$_1$.
'Peter wants for him to help him.'

(12) *Pedro$_1$ quer **Maria** ajudá-lo$_1$/ele$_1$.
'Peter wants Mary to date him.'

(13) *João$_1$ disse que Pedro$_2$ quer **ele**$_1$ ajudá-lo$_2$/ele$_2$.
'John said that Peter wants for him to help him.'

$$
\cfrac{
\cfrac{n \Rightarrow n \quad \cfrac{\overline{n \Rightarrow n} \quad \overline{s \Rightarrow s}}{n, n\backslash s \Rightarrow s}}{
\cfrac{n, (n\backslash s)/n, n \Rightarrow s}{\cfrac{n, (n\backslash s)/n, n|n \Rightarrow s|n}{} \, |R} \, /L} \, \backslash L
\quad
\cfrac{\overline{n \Rightarrow n} \quad \overline{s \Rightarrow s}}{n, n\backslash s \Rightarrow s} \, \backslash L
}{n, (n\backslash s)/(s|n), n, (n\backslash s)/n, n|n \Rightarrow s}
$$

Fig. 7. Derivation of **Pedro quer **Maria** ajudá-lo/ele*

In order to deal with these difficulties, we propose to extend **LLC**.

4   Proposal: Extending LLC

4.1   Semantically Inactive Disjunction Type

The first problem —undergeneration— can be easily tackled by adding the semantically inactive disjunction $\sqcup$ to **LLC** (cf. [18]).

$$
\cfrac{\Gamma\,(x : A) \Rightarrow M(x) : C \qquad \Gamma\,(y : B) \Rightarrow M(y) : C}{\Gamma\,(z : A \sqcup B) \Rightarrow M(z) : C} \, \sqcup L
$$

$$
\cfrac{\Gamma \Rightarrow M : A}{\Gamma \Rightarrow M : A \sqcup B} \, \sqcup R_1 \qquad \cfrac{\Gamma \Rightarrow N : B}{\Gamma \Rightarrow N : A \sqcup B} \, \sqcup R_2
$$

Fig. 8. Rules for semantically inactive disjunction type-constructor $\sqcup$

The optionality of the overt controlled pronoun can now be captured using such a disjunction type-constructor. We can deal with prototypical control sentences simply assigning a semantically inactive disjunction type to the complement argument of the control (transitive) verb:

quería : $(n\backslash s)/((s|n) \sqcup (n\backslash s))$

$$
\cfrac{
\cfrac{\cfrac{\vdots}{n\backslash s \Rightarrow n\backslash s} \, \backslash R}{n\backslash s \Rightarrow (s|n) \sqcup (n\backslash s)} \, \sqcup R
\quad
\cfrac{\overline{n \Rightarrow n} \quad \overline{s \Rightarrow s}}{n, n\backslash s \Rightarrow s} \, \backslash L
}{n, (n\backslash s)/((s|n) \sqcup (n\backslash s)), n\backslash s \Rightarrow s} \, /L
$$

Fig. 9. Derivation for *María quería telefonear*

4.2 Preliminary Proposal: Unlifted Pronominal Types

In order to deal with the second problem —overgeneration of pronoun distribution in control structures— it is important to note, in the first place, that even though the infinitive clauses contain a bound pronoun in (11-13) above, it does not appear as the subject, but the object of the infinitive verb *ajudar* 'to help'. The pronominal type proposed by Jaeger does not distinguish between subject and object pronouns, and so, there is no way to fix the Case of a pronoun.[11] Thus, in **LLC** subject and object (and also reflexive) pronouns are all of type $n|n$; therefore, the following clauses are both of type $s|n$ as they contain a free pronoun in some position.

(14) John saw him.

(15) He saw John.

In the second place, observe that in the three problematic examples, the embedded subject is not controlled by a matrix nominal: in (11) the subject is a free pronoun; in (12) a referential expression occupies the subject position; and in (13) the subject pronoun is bound by a higher nominal. Thus, despite the fact that the control clause contains a bound pronoun, a grammatical control relation is not exercised. While the argument type $(s|n) \sqcup (n\backslash s)$ of the control verb expresses that if there is a pronoun within the complement, it has to be bound, the type $s|n$ is not sufficient to adequately express the control conditions: i) the infinitive subject has to be a pronoun, and ii) it has to be bound (by a specific matrix nominal phrase). In other words, in control sentences it is necessary to ensure, first, that the embedded subject is a pronoun, and second, that the antecedent of the verbal argument type $s|n$ and that of *this* pronoun are the same. If the pronoun is the subject of an intransitive verb phrase, as in example (3) above, both conditions are correctly satisfied, but not when the complement contains the pronoun in the object position of a transitive complement.

In order to address both of these control conditions, we propose to extend **LLC** by adding a new type-constructor $||$ for proforms. The right and left rules for this new connective are to be the same as those of $|$. With this new syntactic type $B||A$ at hand, we can differentiate, in particular, between an expression containing an object pronoun $B|n$ and an expression containing a subject pronoun $B||n$.[12] Consequently, despite the fact that the sentences in (14) and (15)

---

[11] Jaeger's pronominal type does not distinguish between pre- and post-verbal position, this last difference being incorporated in the infinitive verb type in our previous proposal. As we have said before, Jaeger's proposal does not capture Principles A (locality) and B (antilocality) of the Binding Theory: that reflexive pronouns must be bound in their own clause and that accusative pronouns cannot take a c-commanding antecedent in their own clause. In order to take account of the pronominal *position*, we shall use lifted pronominal types, for example, $(s|n)/(n\backslash s)$ for pre-verbal subjects.

[12] An anonymous reviewer observes that this difference could be made by using features instead of introducing a new connective. Although we could have chosen that option,

both contain a free pronoun, they will have different types: $s|n$ and $s||n$, respectively. In words, $s|n$ is the type for a sentence that contains a pronominal free object, and type $s||n$ corresponds to sentences with a subject free pronoun.[13]

Thus the preliminary proposal, with the following lexical entries, derives control sentences without and with an overt controlled pronoun ((16) and (17-18), respectively).[14] And the derivation for a control sentence with an embedded referential subject and a bound object pronoun (as in (19)) is blocked:[15]

ajudar : $(n\backslash s)\,/n : \lambda x.\lambda y.\,((help'\ y)\ x)$
chegar : $n\backslash s : \lambda x.\,(arrive'\ x)$
ele (he): $n||n : \lambda x.x$
joão : $n : j'$
lo/ele (him): $n|n : \lambda x.x$
pedro : $n : p'$
quer : $(n\backslash s)\,/\,((s||n) \sqcup (n\backslash s)) : \lambda x.\lambda y.\,((want'\ (x\ y))\ y)$

(16) Pedro quer chegar.

(17) Pedro quer **ele** chegar.

(18) Pedro quer **ele** ajudá-lo/ele.

(19) *Pedro quer **João** ajudá-lo/ele.

$$
\cfrac{
  \cfrac{
    \cfrac{
      \cfrac{\vdots}{n, n\backslash s \Rightarrow s}\ \backslash L
    }{n||n, n\backslash s \Rightarrow s||n}\ ||R
  }{n||n, n\backslash s \Rightarrow (s||n) \sqcup (n\backslash s)}\ \sqcup R \qquad \cfrac{\vdots}{n, n\backslash s \Rightarrow s}
}{n, (n\backslash s)\,/\,((s||n) \sqcup (n\backslash s))\,, n||n, n\backslash s \Rightarrow s}\ /L
$$

Fig. 10. Derivation for *Pedro quer **ele** chegar*

---

we have preferred to extend Jaeger's proposal because using both proforms we can obtain the corresponding lifted types, and so, we can also distinguish between a pre- and a post-verbal pronoun.
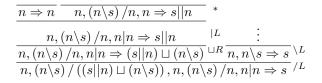
[13] Observe that the double free pronoun reading $\lambda x.\lambda y.((saw'\ y)\ x)$ for *He saw him*, which in Jaeger's system gets the category $(s|n)\,|n$, corresponds on our proposal to the type $(s||n)\,|n$ or $(s|n)||n$.

[14] Note that the bound reading for the object $n|n$ is obtained by using, not |R, but |L.

[15] There are two readings for *Pedro quer **ele** ajudar ele*: a reflexive bound reading $((wanted'\ ((help'\ p)\ p))\ p)$ and a free reading $\lambda x.\,((wanted'\ ((help'\ x)\ p))\ p)$. In the first case, the clause corresponds to the type s and in the second, to the type $s|n$. There is no derivation for the type $s||n$ with the subject free reading $\lambda x.\,((wanted'\ ((help'\ p)\ x))\ p)$. Considering that in BP the object pronoun can take the nominative form, it seems we have to take the sequence of types $n, (n\backslash s)\,/\,((s||n) \sqcup (n\backslash s))\,, n, (n\backslash s)\,/n, n||n$ into account for the sentence in (12). But, it must to be remembered that infinitive subjects are usually preverbal in BP.

$$\cfrac{\cfrac{\overline{n \Rightarrow n} \quad \overline{n,(n\backslash s)\,/n, n \Rightarrow s||n}}{\cfrac{n,(n\backslash s)\,/n, n|n \Rightarrow s||n}{n,(n\backslash s)\,/n, n|n \Rightarrow (s||n) \sqcup (n\backslash s)}\,{}^{\sqcup R} \quad \cfrac{\vdots}{n, n\backslash s \Rightarrow s}\,{}^{\backslash L}}{n,(n\backslash s)\,/\,((s||n)\sqcup(n\backslash s))\,, n,(n\backslash s)\,/n, n|n \Rightarrow s}}{\,}^{/L}$$

$$\overline{\qquad\qquad}\,{}^{*}\qquad {}^{|L}$$

Fig. 11. Illicit derivation for *Pedro quer **João** ajudá-lo/ele*

Observe that in embedded sentences selected by propositional verbs both the free and the bound reading for a pronominal subject are possible, as in Jaeger's system. Thus, an embedded pronoun can be bound by a higher quantifier or another nominal expression when it occurs within complement clauses of propositional verbs:[16]

(20) João disse que ele caminha.
      'John said he walks.'

These two readings result from the following derivations:

$$\cfrac{\cfrac{\vdots}{n,(n\backslash s)\,/s, n, n\backslash s \Rightarrow s}\,{}^{/L}}{n,(n\backslash s)\,/s, n||n, n\backslash s \Rightarrow s||n}\,{}^{||R} \qquad \cfrac{\overline{n \Rightarrow n} \quad \cfrac{\vdots}{n,(n\backslash s)\,/s, n, n\backslash s \Rightarrow s}\,{}^{/L}}{n,(n\backslash s)\,/s, n||n, n\backslash s \Rightarrow s}\,{}^{||L}$$

Fig. 12. Derivations for *João disse que ele caminha*

A control sentence with a prepositional control verb can also be derived assuming the anaphoric type $s||n$ as an argument of the selected preposition:[17]

---

[16] In EP, BP, SPA and IT such a free reading for a pronoun (and even a referential expression) is also allowed even if the embedded verb has (inflected or uninflected) *infinitive form*, when the complement clause is selected by a propositional verb (cf. [22], [16], among others):

    i. Eu penso/afirmo terem **os deputados** trabalhado pouco. (EP)
      'I think/affirm the congressmen have.INFL worked poorly.'
    ii. As italianas sabem serem **elas** encantadoras. (BP)
      'Italian girls know they are.INFL charming.'
    iii. Este documento prueba haber **tú** nacido en 1938. (SPA)
      'This document proves have.INF you was born in 1938.'
    iv. Credevo avere **egli** vinto. (IT)
      '[I] believed [that] he has.INF won.'

[17] To analyse split control as exemplified by the Portuguese sentence below, we suggest the following lexical entry for the prepositional control verb *convencer a/de* 'convince', where g groups individuals:

(21) A polícia forçou os manifestantes a **eles** saír(em). (BP)
      'The police forced the protesters to leave.'

(22) Acusa os colegas de **eles** ser(em) corruptos. (BP)
      '[S/He] accuses the partners of being rascals.'

forçou : $(n\backslash s) / (n \bullet (n\backslash s)) : \lambda x.\lambda y. (((force \ (\pi_2 x \ \pi_1 x)) \ \pi_1 x \ ) \ y)$
a : $(n\backslash s) / ((s||n) \sqcup (n\backslash s)) : \lambda x.x$
acusa : $(n\backslash s) / (n \bullet (n\backslash s)) : \lambda x.\lambda y. (((charge \ (\pi_2 x \ \pi_1 x)) \ \pi_1 x) \ y)$
de : $(n\backslash s) / ((s||n) \sqcup (n\backslash s)) : \lambda x.x$

$$
\cfrac{
\cfrac{
\cfrac{
\cfrac{\vdots}{n||n, n\backslash s \Rightarrow s||n} \ ^{||R}
}{n||n, n\backslash s \Rightarrow (s||n) \sqcup (n\backslash s)} \ ^{\sqcup R} \quad \cfrac{\vdots}{n\backslash s \Rightarrow n\backslash s}
}{
\cfrac{
\cfrac{(n\backslash s) / ((s||n) \sqcup (n\backslash s)), n||n, n\backslash s \Rightarrow n\backslash s \ ^{/L} \quad n/cn, cn \Rightarrow n \ ^{/L}}{n/cn, cn, (n\backslash s) / ((s||n) \sqcup (n\backslash s)), n||n, n\backslash s \Rightarrow n \bullet (n\backslash s)} \ ^{\bullet R} \quad \cfrac{\vdots}{n/cn, cn, n\backslash s \Rightarrow s} \ ^{/L}
}{n/cn, cn, (n\backslash s) / (n \bullet (n\backslash s)), n/cn, cn, (n\backslash s) / ((s||n) \sqcup (n\backslash s)), n||n, n\backslash s \Rightarrow s} \ ^{/L}
}{}
$$

Fig. 13. Derivation for *A polícia forçou os manifestantes a **eles** sair*

Although the previous proposal ensures that the subject embedded pronoun is controlled, the type $n||n$ does not block the derivation of ungrammatical sentences which contain a subject pronoun within a nominal phrase in the embedded subject position,[18] as exemplified below, neither captures the antilocality principle for the object type $n|n$:[19]

(23) *Pedro quer **o fato de que ele chegou** ser comemorado.
      'Peter wants the fact that he arrive be.INF celebrated.'

---

      i. Eu convenci a Maria a/de viajarmos.
      'I convince Mary to travel.INFL.'

      convenci : $(n\backslash s) / (n \bullet (n\backslash s)) : \lambda x.\lambda y. (((convinced \ (\pi_2 x \ \pi_1 x)) \ g(y, \pi_1 x)) \ y)$
      a/de : $(n\backslash s) / ((s||n) \sqcup (n\backslash s)) : \lambda x.x$

[18] Since the pronoun *ele* is assigned the type $n||n$ and the nominal phrase *os amigos d'ele* contains this pronoun it seems that we have to admit that the pronominal type $n||n$ is also assigned to it, and consequently, the sentence in (ii) could be derived:

      ii. João quer **os amigos d'ele** chegar(em).
      'John wants his friends to arrive(.INFL)'

Nevertheless, in this case *ele* is the complement of the preposition, and it is used as the third person possessive pronoun in order to avoid the ambiguity between the second and the third reading for the possessive *seu(s)* ('your'/'his'/'her'). Observe that the preposition *de* 'of' cannot take a first pronoun as its complement: **os amigos de mim/eu/nós*.

[19] Although for reasons of space we do not do so here, we believe our eventual, lifted pronoun type, proposal can prohibit this non-locality when it is semantically modalised; cf. the way non-locality for reflexives is blocked in [17].

4.3   Final Proposal: Lifted Pronominal Types

Like a type n, a pronominal type $n|n$ (and $n||n$) can also be lifted in **LLC**. In other terms, in addition to (24) and (25), the sequents (26) and (27) can also be derived in Jaeger's system:[20]

(24) $n \Rightarrow (s/n) \backslash s$

(25) $n \Rightarrow s/ (n \backslash s)$

(26) $n|n \Rightarrow (s/n) \backslash (s|n)$

(27) $n|n \Rightarrow (s|n) / (n \backslash s)$

Observe, in the first place, that lifted types are differentiated not only with respect to the position —left or right— of the argument, but also with respect to the type of the argument —s/n or $n \backslash s$. It seems clear that the lifted type $(s|n) / (n \backslash s)$ (or $(s||n) / (n \backslash s)$) in our proposal) could be used to categorize a pre-verbal subject pronoun, as it selects a verb phrase to the right.

For accusative pronouns we follow a strategy of lifting ('case as lifting') as well as has been done for nominative pronouns. There are two facts in particular which we capture. First, that accusative pronouns appear in non-subject positions,[21] and second, that they cannot take a subject antecedent in their own clause (antilocality). Following [20] we assign accusative pronouns types of the form $((s \uparrow n) - (J \bullet (n \backslash s))) \downarrow (s|n)$ where $\uparrow$ and $\downarrow$ and J are the extract and infix and discontinuous unit of the displacement calculus [19], and $-$ is difference [20]. The type says that the pronoun occupies a nominal position within a sentence where the position is not subject position $(-(J \bullet (n \backslash s)))$, and then seeks a nominal antecedent outside of the resulting sentence (and hence not the subject of the same sentence: antilocality).

Hence we arrive at an analysis illustrated by the mini-lexicon:

a : $Nt(s(f))/CNs(f) : \iota$
a : $(\exists a Na \backslash Saa)/\exists a((Si||Na) \sqcup (Na \backslash Si)) : \lambda A A$
acusa : $(\exists g Nt(s(g)) \backslash Sf)/\exists a(Na \bullet (Na \backslash Sde)) : \lambda A \lambda B(((charge\ (\pi_2 A\ \pi_1 A))\ \pi_1 A)\ B)$
ajudar : $(\exists a Na \backslash Si)/\exists a Na : help$
chegar : $\exists a Na \backslash Si : arrive$
colegas : $\forall g CNp(g) : partners$
corruptos : $CNp(m)/CNp(m) : corrupt$
de : $(\exists a Na \backslash Sde)/\exists a((Si||Na) \sqcup (Na \backslash Si)) : \lambda A A$
decidiu : $\forall a((Na \backslash Sf)/((Si||Na) \sqcup (Na \backslash Si))) : \lambda A \lambda B((decided\ (A\ B))\ B)$
disse : $(\exists g Nt(s(g)) \backslash Sf)/CPque : say$

---

[20] Note that the sequents are not derivable in the reverse direction. Hence, assigning lifted types preserves some but not all of the distribution of unlifted types.

[21] In the previous proposal for control sentences this fact is captured by assigning different pronominal types for the control argument —$s||n$— and the object pronoun —$n|n$. Notwistanding, as in Jaeger's proposal, there is no way to block the occurrence of an object pronoun in a subject position.

ela : $\forall v(((Sv{\uparrow}Nt(s(f)))-(J\bullet(Nt(s(f))\backslash Sv))){\downarrow}(Sv|Nt(s(f)))) : \lambda AA$
ela : $\forall v((Sv||Nt(s(f)))/(Nt(s(f))\backslash Sv)) : \lambda AA$
ele : $\forall v(((Sv{\uparrow}Nt(s(m)))-(J\bullet(Nt(s(m))\backslash Sv))){\downarrow}(Sv|Nt(s(m)))) : \lambda AA$
ele : $\forall v((Sv||Nt(s(m)))/(Nt(s(m))\backslash Sv)) : \lambda AA$
eles : $\forall v((Sv||Nt(p(m)))/(Nt(p(m))\backslash Sv)) : \lambda AA$
forçou : $(\exists gNt(s(g))\backslash Sf)/\exists a(Na\bullet(Na\backslash Saa)) : \lambda A\lambda B(((force\ (\pi_2 A\ \pi_1 A))\ \pi_1 A)\ B)$
joão : $Nt(s(m)) : j$
manifestantes : $CNp(m) : protesters$
maria : $Nt(s(f)) : m$
namora : $(\exists aNa\backslash Sf)/\exists aNa : love$
namorar : $(\exists aNa\backslash Si)/\exists aNa : love$
o : $Nt(s(m))/CNs(m) : \iota$
os : $Nt(p(m))/CNp(m) : \iota$
pedro : $Nt(s(m)) : p$
policia : $CNs(f) : police$
problema : $CNs(m) : problem$
que : $CPque/Sf : \lambda AA$
quer : $\forall a((Na\backslash Sf)/((Si||Na)\sqcup(Na\backslash Si))) : \lambda A\lambda B((want\ (A\ B))\ B)$
resolver : $(\exists aNa\backslash Si)/\exists aNa : solve$
sair : $\exists aNa\backslash Si : go$
ser : $(\exists aNa\backslash Si)/\exists a(CNa/CNa) : be$

In the Appendix we give some illustrative derivations from this lexicon generated by the parser/theorem prover CatLog2.[22]

5   Conclusions

In this paper we have analyzed overt pronouns in control sentences in Portuguese. Firstly, we have followed Jaeger's proposal for pronouns and we have exposed some problems resulting from adopting the same syntactic type for both subject and object pronouns: $n|n$. We have shown how can we extend Jaeger's system in order to guarantee the control relation with the subject embedded pronoun by distinguishing between subject $n||n$ and object $n|n$ pronominal types. But this strategy was still shown to be limited as we can derive some odd ungrammatical sentences. Finally, we have suggested adoption of lifted pronominal types to avoid these cases.

The analysis can be implemented in the parser/theorem prover CatLog2. In the Appendix we show analyses with lifted types.

---

[22] www.cs.upc.edu/~droman/index.php.

Appendix

(A) **pedro**+**quer**+**ele**+**chegar** : Sf

$Nt(s(m)) : p, \forall a((Na\backslash Sf)/((Si||Na)\sqcup(Na\backslash Si))) : \lambda A\lambda B((want\ (A\ B))\ B),$
$\forall v((Sv||Nt(s(m)))/(Nt(s(m))\backslash Sv)) : \lambda CC, \exists aNa\backslash Si : arrive\ \Rightarrow\ Sf$

Fig. 1: Derivation of (A)
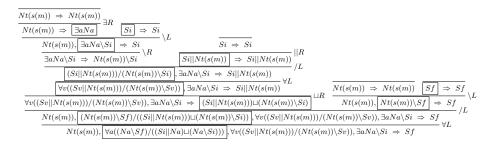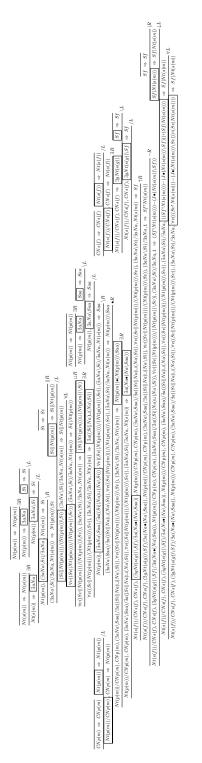
$((want\ (arrive\ p))\ p)$

(B) **a**+**policia**+**forçou**+**os**+**manifestantes**+**a**+**eles**+**ajudar**+**ele** : $Sf|Nt(s(m))$

$Nt(s(f))/CNs(f)\ :\ \iota, CNs(f)\ :\ police, (\exists gNt(s(g))\backslash Sf)/\exists a(Na\bullet(Na\backslash Saa))\ :$
$\lambda A\lambda B(((force\ (\pi_2 A\ \pi_1 A))\ \pi_1 A)\ B), Nt(p(m))/CNp(m) : \iota, CNp(m) : protesters,$
$(\exists aNa\backslash Saa)/\exists a((Si||Na)\sqcup(Na\backslash Si)) : \lambda CC, \forall v((Sv||Nt(p(m)))/(Nt(p(m))\backslash Sv))\ :$
$\lambda DD, (\exists aNa\backslash Si)/\exists aNa : help, \forall v(((Sv^{\uparrow}Nt(s(m)))-(J\bullet(Nt(s(m))\backslash Sv)))^{\downarrow}(Sv|Nt(s(m))))\ :$
$\lambda EE\ \Rightarrow\ Sf|Nt(s(m))$

$\lambda A(((force\ ((help\ A)\ (\iota\ protesters)))\ (\iota\ protesters))\ (\iota\ police))$

Fig. 2: Derivation of (B)

References

1. Asudeh, A.: Richard III. In Mary Andronis, Erin Debenport, Anne Pycha, and Keiko Yoshimura, eds., *CLS 38: The Main Session*, vol. 1, 31–46. Chicago, IL: Chicago Linguistic Society (2002).
2. Asudeh, A.: Control and Semantic Resource Sensitivity. *Journal of Linguistics* 41(3): 465–511 (2005).
3. Asudeh, A., Mortazavinia, M.: Obligatory Control in Persian: Implications for the Syntax–Semantics Interface. Handout from ICIL 4 Uppsala University (2011).
4. Barbosa, P.: Overt Subjects in Raising and Control Complements and the Null Subject Parameter. *LSA Annual Meeting Extended Abstracts* (2010).
5. Bresnan, J.: Control and complementation. *Linguistic Inquiry* 13:343–434 (1982).
6. Chierchia, G.: Anaphoric properties of infinitives and gerunds. In M. Cobler, S. MacKaye and M. Wescoat (eds). *Proceedings of the Third West Coast Conference on Formal Linguistics*. Stanford Linguistics Association, Stanford: CA. 28–39 (1984).
7. Chomsky, N.: *Lectures on Government and Binding*. Dordrecht: Kluwer (1981).
8. Chomsky, N.: *The minimalist program*. Cambridge: MIT Press (1995).
9. Davies, W. D., Dubinsky, S.: *The grammar of raising and control*. Blackwell Publishing (2004).
10. Dowty, D. R.: On recent analyses of the semantics of control. *Linguistics and Philosophy* 8(3): 291–331 (1985).
11. Hornstein, N.: Movement and control. *Linguistic Inquiry* 30(1): 69–96 (1999).
12. Jacobson, P.: Towards a variable-free semantics. Linguistics and Philosophy 22(2): 117–184 (1999).
13. Jaeger, G. *Anaphora and type logical grammar*. Springer: Dordrecht (2005).
14. Landau, I.: *Elements of control. Structure and meaning in infinitival constructions*. Kluwer: Dordrecht (2000).
15. Landau, I.: *A Two-Tiered on Control*. MIT Press (2015).
16. Mensching, G.: *Infinitive constructions with specified subjects*. Oxford University Press (2000).
17. Morrill, G.: Intensionality and Boundedness. *Linguistics and Philosophy*, 13(6):699–726 (1990).
18. Morrill, G. V.: *Type Logical Grammar: Categorial Logic of Signs*. Dordrecht: Kluwer Academic Publishers (1994).
19. Morrill, G., Valentín, O. and Fadda, M.: The Displacement Calculus. *Journal of Logic, Language and Information*, 20(1):1–48 (2011).
20. Morrill, G., Valentín, O.: Displacement logic for anaphora, *Journal of Computer and System Sciences*. 80(2): 390–409 (2014).
21. Polinsky, M., Potsdam, E.: Expanding the scope of control and raising. *Syntax* 9(2): 171–192 (2006).
22. Raposo, E.: Case theory and Infl-to-Comp: the inflected infinitive in European Portuguese. *Linguistic Inquiry* 18: 85–109 (1987).
23. Rosenbaum, P.: *The grammar of English predicate complement constructions*. Cambridge, MA: MIT Press (1967).
24. Stiebels, B.: Towards a typology of complement control. *ZAS Papers in Linguistics* 47: 1–80 (2007).
25. Wurmbrand, S.: Syntactic vc semantical control. Em C. J. Zwar e W. Abraham (eds). *Proceedings from the 15th Workshop on Comparative Germanic Syntax* (2002).