WPFG

The argument structure of deverbal nouns in Brazilian Portuguese

Roberto Gomes CAMACHO
Universidade Estadual Paulista, São José do Rio Preto, Brazil

The argument structure of deverbal nouns in Brazilian Portuguese

1. Introduction[1]

A considerable amount of work has been done on the argument structure of nominalizations in general and of deverbal nouns in particular, much of it within the functionalist perspective. In general, these studies are more concerned with syntax and semantics than with pragmatics, which is considered the most encompassing domain of explanation in FG (Dik 1997a). Moreover, little effort has been devoted to the analysis of real conversational data, which are highly relevant for a functionalist approach. My aim in this paper is to interpret the argument structure of deverbal nouns giving priority to the pragmatic domain on the basis of a corpus of spoken Portuguese.

As a derived type of term, deverbal nouns are generally assumed to keep the same argument structure as the input verb. Thus, deverbal nouns can be mono-, bi- or trivalent, depending on the valency of the input verb; similarly, deadjectival nouns can be mono- or bivalent, depending on the valency of the original lexical item. This is the position taken in Dik's (1985, 1997a) original proposal, according to which the arguments of the base verb or adjective are also arguments of the derived nouns. However, this interpretation of valency by natural inheritance is not uncontroversial because it is only very rarely that in constructions headed by deverbal nouns all the arguments of its underlying valency are expressed, as illustrated in (1a).[2]

(1)a
se eu	*fiz-er*		*est-e*	*gat-o*	*e*	*deix-asse*	
if I	make-SBJV.FUT.IRR.1SG		this-M	cat-M	and	leave-SBJV.PRF.IRR.1SG	

durante	*doze*	*mil*	*ano-s...*	*ele*	*v-ai*	*continu-ar*
during	twelve	thousand	year-PL	he	ir.AUX-IND.PRES.3SG	continue-INF

se-ndo	*um*	*gato*	*sem*	*valor....*	*não*	*t-em::*
be-PROG	a	cat	without	value	not	have-IND.PRS.3SG

nenhum-a..	*um*	*valor*	*artístico*	*est-a*	*represent-ação* (EF-SP-405)
any-F	a.M	value	artistic	this-F	represent-NMLZ

'if I made this cat and left it during twelve thousand years… he would going on being a cat without any value…(…) this <u>representation</u> does not have any artistic value'

b. $(e_i [f_i: representação_{N\ v} (f_i) (x_{i:}: eu (x_i)_{Ag} (x_j: gato(x_j))_{Pat}](e_i))$

In (1a) the two possible arguments present in the underlying structure of the noun *representação* 'representation', given in (1b), are not overtly expressed. This feature of deverbal nouns has led Mackenzie (1985, 1996, 2002) to argue that the prototypical nominalized predication is avalent. This, of course, raises the question of which hypothesis most appropriately and faithfully explains the empirical data, Dik's valency preservation hypothesis or Mackenzie's valency reduction hypothesis.

In this paper I argue in favour of the first position. Like Camacho & Santana (2004) and Keizer (2004b), I argue that deverbal nouns preserve the same argument structure as the corresponding verbal predicate. In her paper, Keizer (2004b) claims that the uniformity found between predications and terms by Rijkhoff (1992) is too restrictive and not descriptively adequate. Preferring a simplified term structure, she analyses all (syntactically) optional information within the clause as satellites and, in

addition, she considers some predicates, like basic relational predicates and most derived nominal predicates, as provided with one or more argument slots.

To support this position I will present quantitative empirical evidence on the basis of a sample extracted from the standard variety of spoken Brazilian Portuguese. Reconsidering, for instance, the instance contained in (1a), I will argue that there are functional reasons to treat the arguments expressed by zero anaphora in the internal structure of the NP headed by *representação* 'representation' as recovering the reference to the terms *eu* 'I' and *gato* 'cat' for the first and second argument, respectively. This involves a pragmatic process of reference identification widely known to characterize the valency structure of verbal predicates too. According to Keizer:

> It is this function of referent identification which explains why it is that the arguments of basic or derived nominal predicates can be omitted, while at predication level they must virtually always be expressed. Since within the term arguments typically serve to enable the addressee to identify the referent, they are only added when they are needed for identification; if not, they are left implied (Keizer 2004b:17).

I agree with the position that the identifying expressions at the term level are non-overt arguments for pragmatic reasons, but, as the analysis will show, there are also semantic reasons for not expressing them.

My analysis will also focus on the formal expression of argument structure, with the aim of identifying the principles motivating the choice of the possible terms and, therefore, the way the competition between the arguments for the possessor phrase slot (Dik 1985) is finally decided.

This paper is organized as follows: section 2 presents the existing proposals for the representation of argument structure of deverbal nouns: Dik's (1985, 1997a) valency preservation hypothesis, assumed more recently by Keizer (2004a), and Mackenzie's (1985, 1996) valency reduction hypothesis. Section 3 gives the most relevant criteria for selecting the sample of spoken language and a brief presentation of data organization. After these preliminary discussions, I will, in section 4, give some empirical evidence of semantic and pragmatic determination to overt and non-overt argument in favour of the valency preservation hypothesis. In section 5, I will finally present my own proposal for the interpretation of valency structure in the framework of Functional Discourse Grammar (henceforth FDG).

2. The treatment of deverbal nouns in FG
2.1. Dik's valency preservation hypothesis

Dik (1985; 1997b) formulates some general principles concerning the formation and expression of derived constructions in natural languages: the Principle of Formal Adjustment (henceforth PFA) and the Principle of Semantic Adjustment (henceforth PSA) govern formal expressions and semantic properties of the derived constructions. The PFA predicts that derived constructions have to adjust their formal expression to the prototypical model of expression for the non-derived constructions. The PSA predicts that, in addition, a derived construction also tends to adjust to the semantic properties of the Prototypical Expression Model. According to Dik, it is the PFA that explains how basically verbal predications acquire nominal properties: an embedded

predication acts as term in a higher level predication, and terms are defined as nominal expressions (see Fig. 1).
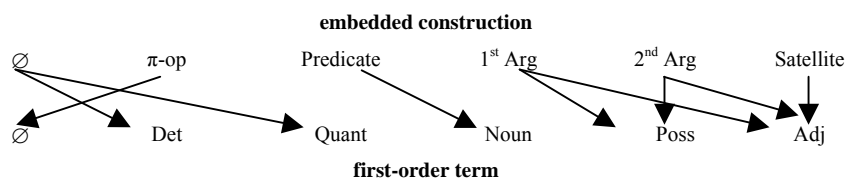
embedded construction

first-order term
Figure 1: Formal adjustments (Dik 1997b: 158)

Since deverbal nouns are embedded constructions, they can act as terms in a predication of a higher level; by virtue of this property, deverbal nouns are provided with an interesting grammatical feature: not only can they keep the argument structure of the input verbal predicate, preserving both valency and semantic functions of the arguments, but they can also themselves act as terms within the higher predication, displaying syntactic, semantic and pragmatic functions that cannot be assigned to verbs. The first and second arguments of the nominalization compete for only one structural slot, that of a possessor *of*-phrase, and this competition is typically won by the second argument of a transitive predication. According to the valency preservation hypothesis, the underlying structure of the nominalization will preserve the quantitative and qualitative valency of the input verbal predicate no matter whether one, or even both, of the two arguments are not overtly expressed. Example (2) represents a predicate formation rule for a noun like *destruição* ('destruction').

(2)   Formation rule for a deverbal noun formed by suffixing
        Input: pred [V] $(x_1)_{Agent}$ $(x_2)_{Goal}$
        Output: pred-*ção* [N] $(x_1)_{Agent}$ $(x_2)_{Goal}$ (cf. Dik 1997b:166)

The predicate frame representing the deverbal noun is specified separately in the lexicon, as illustrated in (3) and (4).

(3)   Destruir 'destroy' [V] $(x_1)_{Agent/Force}$ $(x_2)_{Goal}$

(4)   Destruição 'destruction' [N] $(x_1)_{Agent/Force}$ $(x_2)_{Goal}$

In spite of having a lexical entry distinct from that of the verb, the deverbal noun preserves the syntactic and semantic correspondence with the verbal term by inheritance. Since in FG the predicate frame is strongly associated with the corresponding lexical item, the formal difference between a verbal predicate and the derived nominal predicate is not located in the argument structure, but only in the category label, [V] or [N], assigned to the predicate.
        From this perspective, it is possible to argue that, since the underlying structure is abstract in nature, the argument structure of a predicate may be different from the structure actually used in conversation, in which some arguments are non-overtly expressed. This position can, however, only be inferred, since Dik (1985) only gives formal principles of adjustment without any reference to the functional motivations at work in actual use, which could block the formal insertion of all arguments. It is this weakness in the formulation of the valency preservation hypothesis that Mackenzie addresses in the following passage:

The drawback is that the adjustment rules (...) are stipulative, offering no answer to the question why the correspondences should be as they are presented. Furthermore, no indication is given of the circumstances under which a complement is adjusted or is not: the specific functionality of nominalizations (...) is not reflected in the formalism. (Mackenzie 1996: 11)

What I would like to add is that reducing the rules to the formal mechanisms of adjustment accounts only for the syntactic behaviour of the derived predicate and its arguments, thus ignoring the priority of pragmatic and semantic rules, the most important postulates of FG.

2.2 Mackenzie's valency reduction hypothesis

Mackenzie (especially 1996) provides an alternative to Dik's proposal, arguing that nominalized predicates, like the majority of full lexical nouns, are avalent. Since he admits that some arguments of the input predicate may be preserved in the deverbal noun, Mackenzie makes use of the notion of implied satellite (Dik 1997a:194).
Mackenzie (1985, 1996) uses this notion to advocate his valency reduction hypothesis, which treats the arguments of the input verbal predicate as satellites, not arguments of the derived noun. On the basis of the notion of implied satellite and of the representation suggested by Dik (1997a) for these two types of entities (x for arguments and y for satellites), Mackenzie (1996) adopts the following representation in his treatment of arguments and satellites: (x) for arguments, (y) for implied satellites and (z) for non-implied satellites. Y-satellites are as essential for dynamic and controlled predications as arguments. Z-satellites, on the other hand, are not automatically inherited from the verbal predicate; they can be optionally added to the predication. The possibility of adding arguments and satellites provides an interesting gradation from completely verbal constructions, like finite clauses, to completely nominal constructions, like productive nominalizations, as seen in (5a-d).

(5)a   That **my horse** (Arg_1) won **the race** (Arg_2) came as a great surprise. (finite clause)
  b   **My horse** (Arg_1) winning **the race** (Arg_2) came as a great surprise. (gerund)
  c   **My horse's** (y-satellite) winning **the race** (Arg2) came as a great surprise. (gerund-genitive)
  d   **My horse's** (z-satellite) winning **of the race** (y-satellite) came as a great surprise. (productive nominalization) (Mackenzie 1996: 2)

Mackenzie claims that while (5a) is a completely verbal predication and (5b) shows partial deverbalization, bearing full valency, (5c) is a case of valency reduction: while the second argument is preserved (*the race*), the first one can be omitted or expressed as a possessor satellite (*my horse's*), as stated by the rule in (6):

(6)   English gerund-genitive formation
       Input: $Pred_V$ Arg_1 Arg_2
       Output: $Pred_{VN}$ Arg_2 (Mackenzie 1996: 13)

According to Mackenzie, the productive nominalization in (5d) shows complete valency reduction, since both arguments can be either omitted or specified as Possessor satellites. This may be represented by the following rule:

(7)   English productive nominalization
       Input: $Pred_V$ Arg_1 Arg_2
       Output: $Pred_N$ (Mackenzie 1996: 13)

For the gerund-genitive in (5c) (*my horse's winning the race*) Mackenzie proposes to consider the Possessor (*my horse's*) as a y-satellite, since it is impossible to speak about *winning the race* without implying a possible agent. In nominalizations such as (5d), in which the input predicate (*win*) has the second argument as a Possessor (*my horse's winning of the race*), Mackenzie argues that the Possessor corresponding to the second argument of the input verbal predicate should be treated as a y-satellite and the first argument as a z-satellite. He justifies this claim by stating that " (...) Arg2-equivalents have a more intimate relation to nominalized predicates than Arg1-equivalents" (Mackenzie 1996:14). The second argument may appear either as a possessor in the post-head slot (*of-* phrase) or as a possessor in the pre-head slot (genitive). If an Arg-1 equivalent in the pre-head slot is focused, it is grammatically unacceptable to add an Arg-1 equivalent in the post-head slot, as illustrated by the differences between (8) and (9).

(8)   a. Caesar's (z) destruction of the city (y)
       b. CAESAR's destruction of the city

(9)   a. The city's (y) destruction by Caesar (z)
       b. *THE CITY'S destruction by Caesar   (Mackenzie 1996: 15)

Mackenzie's proposal predicts that, since the first argument corresponds to a z-satellite, it can be omitted, while the argument corresponding to a y-satellite is almost as essential to the predication as the arguments of a verb. A theoretical reason for representing optionally expressed participants in derived constructions as satellites is, according to Keizer that

> there is no need for a separate rule of term predicate formation rule (Mackenzie and Hannay 1982). This rule was introduced to make it possible for terms indicating e.g. location or possession to function as restrictors. For these terms to function as restrictors they would first need to be converted into predicates. If, on the other hand, they function as satellites, such conversion is no longer needed (Keizer 2004b: 11).

However, it needs to be mentioned that this problem no longer applies, since in the FDG framework separate variables are used to represent such semantic categories as location and time (Hengeveld & Mackenzie: 2006).

Another theoretical advantage of this proposal is that it predicts a formal correlation between the argument structures of verbs and nouns and the gradual nature of the deverbalization process of the input predicate: the valency of completely verbal constructions, such as (5a), is gradually reduced to zero in the completely nominal constructions, such as (5d). The process of valency reduction is a good parameter to confirm the gradual process of verbal decategorization and the consequent recategorization of the verbs as nouns.

According to Mackenzie (1996:15), the major advantage of this prosposal is that the predicate rules are no longer merely stipulative, as in Dik's valency preservation hypothesis. Moreover, Mackenzie's proposal allows us to reflect directly

the functionality of the nominalization, in terms of its compactness/efficiency, in the notion of valency reduction.

In spite of these attractions, however, the valency reduction hypothesis is far from unproblematic. It is possible to mention at least three points of criticism:

1. It is unclear how to deal, for instance, with the pragmatic salience of certain terms in discourse; that is, how can we explain that in some constructions a term consisting of new information needs to be expressed, even though the Addressee cannot be assumed to be able to identify the referent of the term? Could such terms be treated as y-satellites? Moreover, in a bivalent nominal predicate where the two terms contain information that is equally required for pragmatic reasons, could the first argument stop being a z-satellite and become a y-satellite?[3]
2. The valency reduction hypothesis seems to be even more problematic in terms of cognitive processing: it seems counterintuitive to reduce valency in the nominalization process and to add satellites referring to the potential arguments later. In other words, why reduce the arguments first to recover them later as satellites?[4]
3. Mackenzie (1996) employs the notion of z-satellite, as shorthand for any satellite that is not implied, to formulate the valency reduction hypothesis. However, the category of z-satellite was created by Dik (1997a) to contrast with the category of y-satellites, which were meant to cover the notion of 'implied satellite', i.e. a satellite that is always predicted in dynamic and controlled states of affairs. The point is that this addition of a new category runs counter to the methodological principle of economy, given the fact that it is possible to make an exhaustive analysis of the linguistic phenomena involved by means of the categories of argument and satellite already available. The same criticism may be extended to the valency reduction hypothesis.

I believe that these problems can be solved by analysing genuine examples of deverbal nouns from a sample of conversational data. The next task, then, is to proceed with data analysis and then reformulate the hypothesis of valency preservation in such a way that it becomes not only more efficient, but also more in line with the general principles of FDG.

3. The sample

The sample consists of 181 occurrences of deverbal nouns collected by Santana (2005), taken from three kinds of survey from the Standard Urban Norm Project (NURC/Brasil Project): Formal Elocutions (EF-377: Castilho & Preti: 1986); Dialogues between Informant and Interviewer (DID-237: Castilho & Preti: 1987); and Dialogues between two Informants (D2-360: Preti & Urbano: 1988). Since the first of these consists of a recorded lesson in classroom, the degree of interactional exchange between the participants is very limited as compared to the other two, which consist of typical conversations. Since the interviewees are all graduates, the three types of survey are representative of the standard variety of spoken Brazilian Portuguese.

The corpus was used to collect a sample of monovalent and bivalent deverbal nouns representing second-order entities. What these nouns have in common with the corresponding verbal predicates is that "they describe an entity that may be interpreted in terms of its temporal setting and in terms of its actuality status, the two

criterial features of states-of-affairs" (Hengeveld 2004a). Since it is relevant to observe the competition between two arguments to occupy the structural slot of a possessor phrase, the sample includes two kinds of input verbal predicates: those taking objects as their complements and those taking obliques as their complements. This methodological procedure was motivated by the fact that, unlike object complements, oblique complements already have a preposition; therefore, they impose different conditions in the competition for the post-nominal slot in the deverbal noun construction.

4. The expression of the argument structure

Before we will deal specifically with the relation between nominal valency and argument expression, it is important to recall that I here assume Dik's valency preservation hypothesis. According to this position, deverbal nouns and their respective input verbs make use of the same argument structure because of the strict correspondence between both categories, even if the nominal predicates rarely show the overt expression of the argument structure. This being so I will proceed to show the way the nominal valency is preserved with conversational data of spoken Brazilian Portuguese. In 4.1 I will show that there is a strong competition between the arguments for the possessor *de*-phrase slot, while 4.2 will focus on the nature of pragmatic and semantic motivation for each type of expression forms.

4.1. The competition for the adpositional phrase slot

Let us consider Table 1, which displays the relationship between potential valency and argument expression. The first row gives the types of verbal predicate that the deverbal nouns are derived from.

Table 1:Types of input verb and the potential valency structure of the derived noun

	Intransitive		Transitive		Oblique transitive		Total	
	N	%	N	%	N	%	N	%
A_1 overtly expressed	21	84.0	4	16.0	/	/	25	14.0
A_2 overtly expressed	/	/	33	92.0	3	8.0	36	20.0
A_1 and A_2 overtly expressed	/	/	/	/	5	100.0	5	3.0
Nothing overtly expressed	38	33.0	53	46.0	24	21.0	115	63.0
Total							181	

Note firstly the general results concerning the total amount of deverbal nouns with overt arguments inside the NP proper (last column to the right). The quantitative results show that the full valency of the corresponding base verb is rarely overtly expressed.

(i)   With only 14.0% (25/181) of the nouns is the first argument overtly expressed; the second argument with only 20.0% (36/181).

(ii)   In only 3.0% (5/181) of tokens we find first and second arguments simultaneously expressed.

(iii)   These scores mean that deverbal nouns with non-overt arguments amount to 63.0% (115/181), a number which represents the full majority of tokens.

These results appear to be favourable to the valency reduction hypothesis. Even so, I will argue, on the basis of more detailed analysis of the data, that there are semantic and pragmatic reasons to believe that non-overt arguments should be considered as being part of the argument structure of the noun.

Let us concentrate on the relation between the number of arguments expressed for each type of input verbs and the total number of overt arguments (see the rightmost column at table 1). We note that 84.0% (21/25) out of first arguments are expressed in nouns derived from intransitive verbs, while in nouns derived from transitive verbs the percentage of overtly expressed second arguments amounts to 92.0 % out of the total number of expressed arguments (33/36). It is interesting to note that all cases with overt expression of two arguments involve nouns derived from oblique transitive verbs.

In spite of the low quantitative significance of these results (n=5/181), the pattern that arises is to be expected, since the second argument inherits the grammatical marking of the oblique case of the input verb. This condition allows the first argument to be expressed as an adpositional phrase introduced by the preposition *de* ('of') (henceforth *de*-phrase).

The data also suggest that, in the case of bivalent deverbal structures, there is strong competition between the two arguments for expression in post-nominal position and that, in accordance with Dik's predictions, it is the second argument that wins out.[5] When only the second argument is expressed, there may be different reasons for not expressing the first argument. Thus, the first argument may be recoverable from the context, as in (10a-b), which may be regarded as cases of a zero anaphora, or the first argument is simply left unexpressed because its reference is undetermined, as in (10c).

(10)a *eles* *precis-am* *peg-ar* *pele* *para* *se* *esquent-ar...*
   they neeed-IND.PRES.3PL find–INF skin to REFL warm-INF

   e *t-er* *comida* *para* *com-er* *e* *se* *defend-er*
   and have-INF food to eat-INF and REFL defend-INF

   d-o-s *outr-o-s* *anima-is...* *então* *a-s*
   from-the.M-PL other-M-PL animal-PL then the.F-PL

   <u>*preocup-açõ-es*</u> *são* *MUITO... ahn::...***de**
   concern-NMLZ-PL be.IND.PRS.3PL much of

   tod-o (o) ***DIA:*** (EF-SP-405)
   every-M DAY

   'they need to catch skin to warm themselves... and to have food... to eat and to defend themselves from other animals... then the majority of <u>concerns</u> are... ahn::... **of every day**.'

 b *se* *eu* *quer-o* *cri-ar...* *um-a* <u>*réplic-a*</u>
   if I want-IND.PRS.1SG create-INF a-F copy-NMLZ

   d-a ***realidade...*** (EF-SP-405:54)
   of-F reality

   'if I want to create... a <u>copy</u> **of the reality**...'

 c *a* *validade* *é* *de* *dois* *ano-s*
   the validity be.IND.PRS.3SG of two year-PL

D-A		*public-ação* ...	*d-o-s*	**resultado-s**...
from-the.F		publish-NMLZ	of-the.M-PL	result-PL

d-a	*lista*	*de*	*aprov-ad-o-s*... (D2-SP-360:151)
of-F	list	of	approve-PAST.PTCP-M-PL

'the validity is two years from the <u>publication</u>... **of the results**... of the list of approved applicants'

It is important to note that the five cases with overt manifestation of two arguments consist of nouns which are all derived from bivalent verbs with an oblique complement, as shown in (11a) and (12a) for the deverbal nouns and in (11b) and (12b) for the respective verbal form.

(11)a
tod-a	*e*	*qualquer*	*manifest-ação (...)*	*v-ai*
every-F	and	each	manifest-NMLZ	go.AUX-IND.PRS.3SG

t-er	*que*	*est-ar*	*necessariamente*	*lig-ad-a*...
have-INF	that	be.DUR-INF	necessarily	link-PAST.PTCP-F

a	*est-a*	<u>*preocup-ação*</u>	*vital*	***d-o***.M	**homem**	**pré-histórico**
to	this-F	concern-NMLZ	vital	of-the.M	man	prehistorical

de...	***se***	***conserv-ar***	***viv-o***... (EF-SP-405:50)
of	REFL	keep-INF	alive-M

'every manifestation is going to be necessarily linked… to this vital <u>concern</u> **of prehistorical man of... keeping himself alive**...'

b
O	*homem*	*pré-histórico*	*se*	*preocup-a*	*vitalmente*
the.M	man	prehistorical	REFL	concern-IND.PRS.3SG	vitally

com	*conserv-ar*	*se*	*viv-o.*
with	keep-INF	REF.	alive-M

'The prehistorical man concerns vitally with keeping himself alive.'

(12)a
além	*d-a*	<u>*particip-ação*</u>	***d-o***	***artista***...	***n-o***	***filme***
besides	of-F	participate-NMLZ	of-M	artist	in-the.M	film

qua-is	*o-s*	*outr-o-s*	*elemento-s*	*importante-s*	*n-a*	*sua*
which.PL	the.M-PL	other-M-PL	element-PL	important-PL	in-the.F	your.F

opinião	*para*	*que*	*o*	*filme*	*sej-a*	*bem*	*suced-id-o*
opinion	in order	that	the.M	film	be-SBJ.IRR.3SG	well	succeed-PAST.PTCP-M

bem	*acei-t-o*	*pel-o*	*público?* (DID-SP-234:113)
well	receive-PAST.PTCP.M	by-the.M	public?

'in your opinion besides <u>the participation</u> **of the artist... in the film** which other important elements make the film successful and well received by the public?'

b
O	*artista*	*particip-a*	*n-o*	*filme*
the.M	artist	participate-IND.PRS.3.SG	in-the.M	film

'the artist participates (takes part of) in the film'

The two arguments in (11a) are both expressed as *de*-phrases, which is normally not allowed in Portuguese. Nevertheless, the speaker violates the rules of

grammar, inserting *de* 'of' into the prepositional slot for c*om* 'with', an insertion which seems to be motivated by the fact that the second argument is not a nominal phrase but a non-finite clause, which means that the competition for expression as *de*-phrase seems to be postponed.

There are some occurrences of first arguments expressed as adjectives,[6] as shown in (13).

(13)

tod-a	*a*	*evolu-ção*	**human-a...**	*não*	*deix-a*
every-F	the.F	evolve-NMLZ	human-F	not	let-IND.PRS.3SG

de	*s-er*	*exatamente*	*a*	*evolu-ção*	*d-o*
of	be-INF	exactly	the.F	evolve-NMLZ	of-the.M

domíni-o	*que*	*o*	*homem*	*t-em*	*sobre*
control-NMLZ	that	the.M	man	have-IND.PRS.3SG	over

a	*natureza...* (EF-SP-405)
the.F	nature

'every **human** evolution... is simply no more than the evolution of the control man has over nature...'

This kind of formal expression has, of course, little to do with the competition between first and second arguments for the possessor phrase slot; it does, however, play an extremely important role in recovering the NP *o homem pré-histórico* 'the prehistorical man'. There are only a small number of such constructions, which generally occur in the context of an underlying intransitive verb, as in (13), and, to a lesser degree, in the context of a transitive input verb, as in (14) and (15).

(14)

se	*ela*	*foi*	*cria-d-a...*	*para*	*um*	*FIM...*
If	she	be.AUX.IND.PRF.3SG	create-PAST.PTCP-F	for	an.M	end

Outr-o	*que*	*NÃO...*	*a*	*contempl-ação*	*estética...* (EF-SP-405:55)
Other-M	than	not	the.F	contemplate.NMLZ	aesthetic

if it was created... for an END... other than an **aesthetic** contemplation...'

(15)

depois	*de*	*um-a*	*represent-ação*	*teatral* (DID-SP-234:106)
after	of	a-F	represent-NMLZ	theatrical

after a **theatrical** representation

The reason why the second argument is selected for overt expression in (14) and (15) has to do with the nature of the other arguments: in both cases the first argument has undetermined reference. As such, there is no pragmatic need to express it. There is a semantic and a pragmatic motivation for an adjectival expression in the pre-head position when the post-head one is also available: the semantic motivation has to do with the generic meaning of the nominal predication while the pragmatic motivation is related to the textual recovering of nouns just mentioned before.

Like adjectives, possessive pronouns may be inserted into the pre-head position. The sample contains only one case of pronominal expression in the function of first argument, given in (16).

(16)

eu	*v-ou*	*trabalh-ar*	*com*	*barro*	*v-ou*
I	go.AUX.IND.PRS.1SG	work-INF	with	clay	go.AUX.IND.PRS.1SG

faz-er	***minh-a-s***	*cri-aç̃õ-es*		*ou*	*eu*	*v-ou*
make-INF	my-F-PL	create-NMLZ-PL		or	I	go.AUX-IND.PRS.1SG

pint-ar	*um*	*quadro...*	(EF-SP-405)
paint-INF	a.M	picture	

'either I̲ am going to work with clay I am going to make **my** creations or I am going to paint a picture...'

The pronominal references in (16) – both the personal pronoun *eu* 'I' and its anaphoric resumption as possessive pronoun *minhas* 'my' – have generic reference, meaning 'any prehistorical man ' or even 'any human being'; expression of the first argument of the nominal predicate *criação* 'creation' would normally be blocked by the presence of *eu* 'I' in the preceding sentence, as *criação* 'creation' shares this argument with the verb *trabalhar* 'to work'. Overt expression in the form of a possessive pronoun is meant to produce an aura of subjectivity, more particularly a kind of complicity between Speaker and Addressee.

A possible explanation for the high incidence of non-overt arguments is that from the point of view of discourse there seems to be no relevance to express the arguments in question: the argument slot for these arguments can already be filled on the basis of the preceding discourse. Thus the lack of overt expression does not mean argument reduction in the process of deverbalization; rather it means that there is competition not only between two arguments for the adpositional phrase slot, but also between formal expression as zero anaphora and adpositional phrase. In addition, the arguments, specially the first ones, can appear in pre-head position, either as a genitive (possessive pronouns) or as an adjective, both of which options are motivated by pragmatic reasons, such as the need for textual continuity of already mentioned information or emphasis. When the genitive expression, as *de*-phrase, appears in the post-head position the information it conveys is generally new.

4.2. The expression form of the arguments

Observe now the distribution of data in table 2, which displays the expression form of the arguments.

Table 2: expression form of arguments[7]

Argument expression →	A_1 expressed		A_1 not expressed		A^2 expressed		A^2 not expressed	
Expression form:	N	%	N	%	N	%	N	%
Possessor phrase	22	12.0			30	25.0		
Adjective	4	2.0			1	1.0		
Other expression form	4	2.0			10	8.0		
Total	**30**	**16.0**			**41**	**34.0**		
Semantically equivalent NPs			44	24.0			6	5.0
Zero anaphora			47	26.0			32	26.0
in the following text			4	2.0			4	3.0
Undetermined reference			26	15.0			17	14.0
Non-referential predicate			30	17.0			22	18.0
			151	**84.0**			**81**	**66.0**

Considering first the total scores, the number of second arguments expressed (34.0%) is more than twice the number of first arguments expressed (16.0%), a kind of incidence expected according to Dik's predictions (1985; 1997b). Recall that, while the preferred form of expression, that of the possessor phrase, is available for both arguments, the first argument is only expressed as a *de*-phrase when occurring with

monovalent nouns. If there are two available arguments in the formal argument structure of the deverbal noun, the second argument will be expressed as a *de*-phrase, while the first argument will take the form of an adjective or possessive pronoun.

Moreover, the data confirm the prediction that the overall preferred expression form is that of a *de*-phrase, since this form is assigned to 74.0% (22/30) out of the total number of first arguments expressed and 73.0% (30/41) out the total number of of second arguments expressed; these percentages show an impressively symmetrical relationship. Examples (17) and (18) provide instances of *de*-phrases expressing first and second arguments, respectively.

(17) *isso*   *signific-a*   *um*   <u>*aument-o*</u>
     this   mean-IND.PRS.3SG   a.M   raise-NMLZ

     de   ***vencimento-s***... (D2-SP-360)
     of' salary-PL

     'this means a <u>raise</u> **in salary'**

(18) *a*   *responsabilidade*   *n-a*   <u>*manut-enção*</u>   *d-a*
     the.F   responsability   in-the.F   maintain-NMLZ   of-the.F

     casa...   *v-êm*   *como*   *complemento*   *né?*   (D2-SP-360)
     home   get-IND.PRS.3SG   as   complement   doesn't?

     'the responsibility in the <u>maintenance</u> **of home**... gets as a complement, doesn't it?'

Considering now the form of the first and second arguments, it is possible to note that the column 'A$_1$ expressed' (meaning 'preceding or following the nominal predicate head'), contains a score of 30 cases. Column 'A$_1$ not expressed' (meaning 'not expressed before or after the nominal predicate head) indicates that there are 44 cases of zero anaphora expressing arguments semantically shared with the predicate of a matrix clause or a coordinated clause (cases of equi-NP deletion) and 47 cases of zero expression of arguments resuming some NP just mentioned earlier in the preceding text. In addition to this, there are 4 more cases where the first arguments are not expressed in the internal structure of the NP headed by a deverbal noun, but provided in the following context, by means a relative clause attached to the nominal head.

Considering these cases of anaphoric, the first type represents first arguments which are semantically shared with the main predicate, as illustrated in (19a).

(19)a *ajud-ar*   *um*   *pessoal*   *que*   *t-em*   *me*   *ped-ido*
     help-INF   a.M   people   who   have.AUX-IND.PRS.3SG   me   ask-PST.PTCP

     para   *faz-er::*   <u>*program-ação*</u>   *de*   *suco-s*
     to   make-INF   programm-NMLZ   of   juice-PL

     d-o   *Lanjal*   (DID-SP-234)
     of-M   Lanjal

     'to help some people who has asked me to make:: <u>programming</u> of Lanjal juices'

It should be noted that the same nominalization with the first argument overtly specified in postnominal position would not be a well formed construction in Portuguese:

(19)b *ajudar um pessoal que   tem **me** pedido para fazer:: <u>programação</u> de sucos do Lanjal (**por mim**)
'to help some people who has asked **me** to make:: <u>programming</u> of of Lanjal juices (**by me)'**

The second type of zero anaphora represents terms recovering some given entity that have just appeared in the precedent text not exactly in the matrix predication, as illustrated by *preocupação* 'concern' in (20a)

(20)a	*n-um-a*	*vida*	*d-ess-e*	*tipo...*	*a*	*<u>preocup-ação</u>*	*principal*
	in-a-F	life	of-this-M	kind	the.F	concern-NMLZ	main

	est-á		*centra-d-a*	*n-a*	*sobreviv-ência...* (EF-SP-405)
	be.DUR-IND.PRS.3SG		focus-PST.PTCP-F	in-the.F	survive-NMLZ

'in this kind of life... the main concern is focused on surviving'

Now note that the same nominalization with an overt first argument represented by *do homem* 'of man' would be a well-formed construction, meaning that the overt or non-overt expression is a real choice of the Speaker.

(20)b numa vida desse tipo... a <u>preocupação</u> principal **do homem** está centrada na sobrevivência... (EF-SP-405)
'in this kind of life... the main <u>concern</u> **of man** is focused on surviving

The third type of zero is not anaphoric but cataphoric. There is a limited number of cases where the non-overt arguments may be recovered not in the preceding text, but in the following text, as part of a relative clause serving as a modifier of the nominal head, as illustrated by *percepção* 'perception' in (21a)

(21)a	*é*	*MUIto*	*difícil (...)*	*a*	*gente*	*separ-ar*
	be.IND.PRS.3SG	VERY	difficult	the.F	people.1PL	separate-INF

	a	*<u>percepção</u>...*	*d-o*	*conceito*	***que***	***nós***
	the.F	**perceive-NMLZ**	**of-the.M**	**conception**	**that**	**we**

	faz-emos	*d-o*	*objeto..*	.(EF-SP-405:56)
	make-IND.PRS.1PL	of-the.M	object	

'it is very difficult for us to separate the <u>perception</u>... of the conception **that we make on the object**...'

Here again the insertion of a term in the first argument position makes the construction ill-formed, as seen in (21b)

(21)b. * *separar a <u>percepção</u> do objeto **por nós** do conceito que **nós** fazemos do objeto*
'to separate the <u>perception</u>... **by us** of... of the concept that **we** make on the object...'

If we add to these three kinds of argument manifestation those which are formally expressed as NP head modifiers, they make up 68.0% (125/181) of first arguments overtly expressed. The remaining 32.0% (56/181) consists of non-referential noun predicates (30 cases) and semantically undetermined subject arguments (26 cases), as illustrated by the underlined terms in (22) and (23a-c), respectively.

(22) *dentro*   *d-o*   *aument-o*   *de*   *vencimento-s*   *hav-eria...*
      inside   of-M   raise-NMLZ   of   salary-PL   there.be-IND.FUT.IRR.3SG

      um-a   *promo-ção*   **de**   **tod-o**   *o*   **pessoal...**   (D2-SP-360)
      a-F   promote-NMLZ   of   all-M   the.M   staff

      'together with the raise of salary there would be... a <u>promotion</u> **of all the staff**'

(23)a. *aquilo*   *é*   *um*   *complemento*   *quer dizer*
        that   be.IND.PRS.3SG   a.M   complement   I mean

        tudo   *que*   *v-em*   *é*   *ótimo*   (D2-SP-360)
        everything   that   come-IND.PRS.3SG   be.IND.PRS.3SG   good

        'that is a <u>complement</u> I mean everything coming is good'

    b. *eu*   *hav:-:ia*   *l-ido*   *um::*
        I   have.AUX-IND.IPFV.1SG   read-PTCP   a.M

        sobre   *um*   *nov-o*   *método*   *de*   *alfabetiz-ação*... (D2-SP-3600 '
        about   a.M   new-M   method   of   teach.reading-NMLZ

        I've:: been read a:: something about a new method of <u>teaching reading</u>...'

    c. *o*   *teatro*   *é*   *menos*   *acei-t-o*   *pel-o*
        the.M   theater   be.IND.PRS.3.SG   less   accept-PST.PTCP-M   by-the.M

        público   *por*   *falta*   *de*   *divulg-ação*   (DID-SP-234)
        public   for   lack   of   divulge-NMLZ

        'the theater is less accepted by the public because of some lack of <u>divulgation</u>, right?'

The promotion the speaker refers to in (22) must be carried through by the Ministry of Justice, which determines the career of the state attorney, but the speaker does not, at that point in the discourse, deem it relevant to mention the agent; a situation similar to that of passive voice constructions like (24a) or other predications containing an undetermined subject, like (24b).

(24)a. *O*   *pessoal*   *foi*   *promov-id-o*
        the.M   staff   be.IND.PRF.3SG   promote-PST.PTCP-M

        promov-eu   *se*   *o*   *pessoal*
        promote-IND.PRF.3SG   REFL   the.M   staff

        'The staff was promoted'

    b   *promov-eram*   *o*   *pessoal*.
        promote-IND.PRF.3PL   the.M   staff
        'One promoted the staff'

The deverbal constructions in (23a-c), on the other hand, contain non-referential nouns, functioning as a non-verbal predicate in (23a), a modifier PP in (23b), and the complement of a noun in (23c). This means that the noun assigns a property without referring, thus providing only a lexical description of a state-of-affairs. The noun gets close to the status of a zero-order entity, the less prototypical referential category: while referring to a zero-order expression, this kind of deverbal

noun can not refer to concrete entities, but only to the property it denotes (cf. Keizer 2004b).

If we apply the same procedure to second arguments, the results are surprisingly comparable: adding up the overtly expressed cases (n=41), the cases of non-expressed arguments for reason of semantic sharing (n=6), the cases of arguments pragmatically resumed by zero anaphora (n=26), and the cases of arguments expressed in the following context (n=4) amount to a total of 68.0% (77/122) of expressed arguments, compared to 18.0% (22/122) of undetermined referents and 14.0% (n=17) of non-referential nominal heads.

Example (23c), here repeated as (25), contains an instance of a semantically shared argument, where an overt expression of A_2 would render the sentence ill-formed.

(25)   *_o teatro é menos aceito pelo público por falta de divulgação **do teatro**_
        'the theater is less accepted by the public because of some lack of divulgation **of the theater**'

Example (26a) contains an instance of a nominal predication whose A_2 slot recovers some given entity that have just appeared in the precedent text. The insertion of an overt expression of this argument does not make it an ungrammatical construction, as illustrated in (26b).

(26)a.	*ele*	*perceb-eu*		*que*	*er-a*		*capaz*	*de*
	he	realize-IND.PRF.3SG		that	be-IND.IPFV.3SG	able	of

	CRI-AR::...		*e*	***cri-ar***	***um-a***	***imagem***..
	create-INF		and	create-INF	an-F	image

	(...) então::		*ele*	*v-ai*		*tent-ar*		*us-ar*
	(...) then		he	go.AUX-IND.PRS.3SG		try-INF		use-INF

	est-a	*cri-ação...*	*que*	*ele*	*é*		*capaz*	*de*
	this-F	create-NMLZ	that	he	be.IND.PRS.3SG		able	of

	faz-er...	*para*	*garant-ir*	*a*	*caça..*	(EF-SP-405)
	do-INF	to	guarantee-INF	the.F	hunted	

'he realized that he was capable TO CREATE::... and to **create an image**... (...) then:: he's going to try using this creation...he is able to do... to guarantee the hunted...'

b. *ele percebeu que era capaz de CRIAR::... e **criar uma imagem**...(...) então:: ele vai tentar usar esta criação **da imagem**... que ele é capaz de fazer... para garantir a caça....*

'he realized that he was capable TO CREATE::... e to **create an image**... (...) then:: he's going to try to use that creation **of the image**... he is able to do...to guarantee hunting...'

Example (27a) contains an instance of nominalization where a non-overt argument may be recovered in the following text by means of a relative clause. Here again insertion of a term in the second argument slot would make the construction ill-formed, as illustrated in (27b)

(27)a	*é*		*inCRÍvel*	*o que*	*aparec-e*		*lá*
	be.IND.PRS.3SG		incredible	what	happen-IND.PRS.3SG		there

o-s	*cort-e-s*	***que***	***eles***	***d-ão***	***n-a-s***
the.M-PL	cut-NMLZ-PL	that	they	make-IND.PRS.3PL	in-the.F-PL

cena-s	(DID-DP-234)
scene-PL	

'it is incredible what happens there the <u>cuts</u> **they make in the scenes**'

b.* *é inCRÍvel o que aparece lá os <u>cortes</u> <u>das cenas</u> que eles dão nas cenas é::*
   'it is incredible what happens there the <u>cut</u> of the scenes **that they make in the scenes**'

In (27a) reference to the arguments of *cortes* 'cuts' is made inside the relative clause that follows. The inappropriateness of an overt argument has much to do Economy Principle by Haiman (1983). Since this type of non-overt expression depends on semantic sharing of participants, FDG can house it very appropriately at the Representational Level.

The arguments of *criação* 'creation' in (26a), which are given zero anaphora expression in the nominalized term, are easily recoverable from the preceding text. The potential subject is *ele* 'he', that is, *o homem pré-histórico* 'the prehistorical man', and the potential object is *imagem* 'image'; unlike the arguments in (25) and (26), they could be expressed within the term, but the application of the Economy Principle leads to non-overt arguments in the term headed by the deverbal noun *criação* 'creation'.

Observe now (28a):

(28)a
na medida... em que		*acab-ava*		*a*	*caça*	*d-o*	*lugar*
as the same time that		be over-IND.IPFV.3SG		the.F	hunted	of-the.M	place

OU	*(que)*	*em virtude*	*d-a*	*época*	*d-o*		*ano*
OR	(that)	by virtue	of-the.F	season	of-the.M		year

n-o	*inverno*	*por*	*exemplo...*	*imigr-avam*		*para*	*lugar-es*
in-the.M	winter	for	example	migrate-IND.IPFV.3PL		to	place-PL

mais	*quente-s*	*eles*	*também*	*precis-avam*
more	warm-PL	they	also	need-IND. IPFV.3PL

acompanh-ar...	*a*	<u>*migr-ação*</u>	***d-a***	***caça*** (EF-SP-405:49)
follow-INF	the.F	migrate-NMLZ	of-the.F	hunted

'as the time that the hunted of the place was over OR (that) by virtue of the season in the winter for instance… they used to migrate to warmer places, they also needed to follow the… <u>migration</u> **of the hunted**'

The first argument of the deverbal noun *migração* 'migration' - *caça*, (literally 'hunted', meaning 'animals to be hunted') - does not represent secondary or backgrounded information; rather, by displaying a contrast with the parallel expressions given in (28b-c), the presence of the repeated PP *da caça* 'of the hunted' becomes absolutely necessary for the thematic continuity.

(28)b
migr-ação	*d-o*	*homem*
migrate-NMLZ	of-the.M	man

'man's migration'

16

c. *migr-ação*   *d-a*   *caça*
    migrate-NMLZ   of-the.F   hunted

    'the migration of the hunted'

Furthermore, although when first mentioned *a caça do lugar* 'the hunted of the place' consists of pragmatically new information, when mentioned again, it consists of given information; as such, one would expect this argument to be expressed as zero anaphora. Instead, however, it is represented by the full NP *a caça* 'the hunted' in order to not produce confusion between both types of migration, man's migration and hunted's migration; therefore, the overt expression of this argument is prompted by pragmatic and semantic motivations.

It is exactly the pragmatic role of verb-derived nouns in creating discourse continuity that determines the formal expression of the arguments. If the pragmatic function is the introduction of a new referent, the verb-derived noun should allow an overt NP corresponding to this argument; if, on the other hand, there is a discourse function of preserving text cohesion by recovering a preceding complete predication, the potential arguments are not overtly expressed.[8]

To sum up, there are two cases of zero anaphora: one triggered by semantic considerations and the other by pragmatic considerations. Both kinds of zero anaphora are governed by a pressure towards the maximal simplification of expression.

4.3. Some theoretical generalizations and implications

The data just analyzed seem to show that deverbal nouns preserve the argument structure of the corresponding input verb and that the overt or non-overt expression of their arguments within the term depends on a set of semantic and pragmatic factors, mainly related to the shared short-term information by the speech act participants. In the functionalist literature these motivations have been collectively referred to as the Economy Principle (Givón, 1980; 1990; Haiman 1983). Cristofaro (2003) describes this principle as follows:

> Syntagmatic economy is the tendency to reduce the length or complexity of any utterance, so that the most frequent expressions tend to be reduced phonetically and information that is redundant and/or recoverable from the context tends to be omitted (Cristofaro 2003: 248) .

In my own examples syntagmatic economy is strongly correlated to predetermination of semantic information in related states of affairs and non-overt expression of shared information for pragmatic reasons.

If, for example, the tense, aspect and mood of a dependent predication are predetermined by the semantic features of the main predication, they do not need to be specified in the verb of the dependent predication, as can be seen in (29a), an instance of a complex construction from Maricopa.

Maricopa (Northern Amerindian, Hokan)
(29)a   *[nyaa*   **m-mii-k***]*   *ʼii-m*
        I   2-cry-SUB   say-REAL
        'I said you cried' (Gordon 1986: 247 apud Cristofaro 2003: 66)

It is obvious that this type of non specification is not applied to nominalizations. But, similarly, if two predications share the same participants, reference to them may be omitted in the dependent construction, as illustrated in (29b)

Guugu Yimidhirr (Australian)

(29)b	*Ngayu*	*wawu-dgirr*	*[mayi*	*buda-nhu]*
	1SG.NOM	want-COM.ABS	food.ABS	eat-PURPV

'I want to eat food' (Haviland 1979: 135 apud Cristofaro 2003: 78)

This principle, also mentioned by Koptjevskaja-Tamm (1993), explains such cases of argument sharing in nominalizations as given in (29c).

(29)c	a	*caça*	*que*	*é*		*o que*	*oferec-e...*
	the.F	hunted	that	be.IND.PRS.3SG		what	offer-IND.PRS.3SG

um-a	*resist-ência* (EF-SP-450)
a-F	resist-NMLZ

'the hunted... that is the one that offers... some <u>resistance</u>

There are also cases in which the non-overtly argument is not semantically determined by the predicate of the main clause. When the participants are semantically predetermined, the speaker may omit reference to them, since the correspondent semantic information can be inferred from the subordination relation. However, when the participants are not predetermined and there is no overt reference to them, there is no semantic means of recovering the missing information. Example (30a) can be seen as an illustration of this type of non-overt argument.

(30)a	*a*	*preocup-ação*	*central...v-ai*		*s-er*
	the.F	concern-NMLZ	main	go.AUX-IND.PRS.3SG	be-INF

em	*torno*	*d-a*	*caça...* (EF-SP-405:51)
in	around	of-the.F	hunted

'the central <u>concern</u> ... will be around the hunted'

When we consider the preceding context, given in (11a), here repeated as (30b), we find that the first argument is the prehistorical man mentioned before.

(30)b	*tod-a*	*e*	*qualquer*	*manifest-ação (...)*		*v-ai*
	every-F	and	each	manifest-NMLZ		go.AUX-IND.PRS.3SG

t-er	*que*	*est-ar*	*necessariamente*	*lig-ad-a...*
have-INF	that	be.DUR-INF	necessarily	link-PAST.PTCP-F

a	*est-a*	*preocup-ação*	*vital*	*d-o*	*homem*	*pré-histórico*
to	this-F	concern-NMLZ	vital	of-M	man	prehistorical

de...	*se*	*conserv-ar*	*viv-o...* (EF-SP-405:50)
of	REFL	keep-INF	alive-M

'every manifestation will have to be necessarily linked… to this vital <u>concern</u> **of prehistorical man of... keeping himself alive**...'

Although expression of the same argument is morphosyntactically allowed, it is not overtly expressed because it consists of information pragmatically shared by the speech act participants. Therefore, the non-overt argument of the utterance in (30b) does not involve real loss of information, since the Addressee knows that, if the Speaker uses (30a), it is only because this expression shares the potential subject with that of (30b). This is a typical example of the Principle of Information Recoverability at work (Cristofaro 2003: 250-1); within FDG, this situation can be appropriately explained by motivations applying at the Interpersonal Level.

The Principle of Information Recoverability fits in well with the way FDG is organized. Thus, in FDG,

> each level of representation within the grammar feeds into the contextual component, enabling subsequent reference to various kinds of entity relevant at each level as soon as they are introduced in the discourse. The operation of formulation draws on this component so that that the availability of antecedents and visible referents may influence the composition of (subsequent) discourse acts (Hengeveld & Mackenzie 2006: 670-1).

In other words, there is considerable interaction between the Representational and Interpersonal Levels and the context. Thus, the choice between overt and non-overt expression of the arguments of deverbal nouns depends on the Speaker's assumptions concerning the Addressee's short-term memory, as represented by the Contextual Component.

Non-referential deverbal predicates, like the one in (31), and referentially undetermined arguments, exemplified in (32), fulfill the semantic and pragmatic conditions, respectively, which function as a signal to the Addressee not to expect referents to be recoverable from the context, a linguistic phenomenon not restricted to nouns.

(31) *bom*   *já*   *est-á*   *n-um-a::*   *idade*   *de*
      well   already   be.DUR-IND.PRS.3SG   in-an-F   age   of

      defin-ição   *quanto*   *a-o*   *segundo* *ciclo*   (D2-SP-360: 168)
      define-NMLZ   as   to-the.M   second   cycle

      'well she is already at an:: age of definition as to high school'

(32) *aument-a*   *um*   *pouco*   *mais*   *a*   *procur-a*
      raise-IND.PRS.3SG   a.M   litte bit more   the.F   search-NMLZ

      de   *engenheir-o*   *civil...* (D2-SP-360)
      of   engineer-M   civil

      'the search for civil engineers increases a little bit more...'

According to Hopper & Thompson (1984), whether or not we are dealing with a prototypical noun depends not so much on its semantic properties, but on its function in the discourse:

> Prototypicality in linguistic categories depends not only on independently verifiable semantic properties, but also - and perhaps more crucially - on the linguistic function in the discourse (Hopper & Thompson 1984: 708).

Thus, prototypical nouns introduce or refer to specific discourse entities. However, nouns do not always enable the speaker to construct or identify a discourse referent (Dik 1997a: 114). The noun *definição* 'definition' in (31) illustrates this situation very well.

In other words, whether or not a noun denotes a visible and concrete entity (which is not the case with second-order entities like *definition*) is not crucial to determine whether we are dealing with a prototypical noun. Instead, the decisive factor is the function that the noun performs in the communicative situation. Whereas the use of *definição* 'definition' in (31) does not enable the Addressee to construct or identify a discourse referent, the use of the same noun in (33) does allow the Addressee to identify a referent; this is why the noun is provided with its full set of possible arguments.[9]

(33) *Você*   *entend-eu*   *aquel-a*   *segund-a*   *defin-ição*
      you   understand-IND.PRF.3SG   that-F   second-F   define-NMLZ

      de   *morfema*   *que*   *o*   *professor*   *d-eu*
      of   morpheme   that   the.M   professor   give-IND.PRF.3SG

      n-a   *aula*   *de*   *ontem?*
      in-the.F   class   of   yesterday

      'Did you understand that second definition of morpheme the professor gave in class yesterday?'

In a similar way, semantic features are not enough to determine the prototypicality of verbs. To be considered a prototypical member of the verbal category, a verb form must be used to report a discourse event (Hopper and Thompson 1984: 726). Thus a non-finite form like *acertar* 'to shoot' in (34) should not be considered a prototypical instance; the finite verb in (35), however, could be well considered a prototypical instance, as it describes the discourse event of shooting.

(34) *Acert-ar*   *trave-s*   *exig-e*   *habilidade.*
      shoot-INF   post-PL   require-IND.PRS.3SG   ability

      'To shoot posts requires some ability.'

(35) *Ronaldo*   *acert-ou*   *a*   *bola*   *n-a*
      Ronaldo   shoot-IND.PRF.3SG   the.F   ball   in-the.F

      trave   *esquerd-a*   *para*   *prov-ar*   *sua*   *habilidade*.
      post   left-F   to   prove-INF   his.F   ability

      'Ronaldo shot the ball right in the left post to prove his ability.'

The discourse functions of referring to concrete entities or reporting discourse events thus form the main factor determining whether a given form is a central or a peripheral member of its category. This interpersonal motivation determines the degree of category membership.

Since deverbal nouns are modeled after regular non-derived nouns, they are characterized by the use of the possessor expression, which in the nominal domain forms the typical means of expressing arguments. But, while deverbal nouns make

reference to higher-order entities, the correspondence with the arguments of the verbal input should be reflected in the underlying structures of both classes of words.

As further evidence supporting the different behaviour of first and second-order nouns, we may consider cases of double expression of the possessor phrase, which occurs mainly in the genitive-possessive type of nominalization, like *my horse's winning of the race*. Although deverbal nouns adjust to the model of prototypical nouns, there is a clear difference between these deverbal nouns and first-order nouns in that the double expression of a possessor is allowed only for deverbal nouns. Compare (36a) and (36b).

(36)a   **o   livro   de   José   de   Maria*
the.M   book   of   José   of   Mary
'John's book of Mary'

   b   John's dismissing of Mary / John's dismissal of Mary

The only situation where a deverbal noun could be fully assimilated to a first-order noun is where the deverbal nouns do not describe states-of-affairs, but the result of action predicates, as seen in (37a) representing a result noun in opposition to (37b) representing an action noun.

(37)a   *Aquel-a <u>constru-ção</u>   d-o   alto   d-a   colina*
That-F   build-NMLZ   of-the.M   top   of-the.F   hill

         é   muito   sólid-a.
be-IND.PRS.3SG   very   steady-F

'That <u>building</u> on the top of the hill is very steady.'

   b   *A   <u>constru-ção</u>   d-a   casa   demor-ou   dois   ano-s.*
the.F   build-NMLZ   of-the.F   house   last-IND.PRF.3SG   two   year-PL
'The <u>building</u> of the house lasted two years.'

The underlined noun in (37a), which is derived from the verbal predicate *construir* 'to build', is not provided with an argument structure since it does not refer to a state-of-affairs, but to the result of an action; therefore, the prepositional phrase *do alto da colina* '[literally 'of the top of the hill (on the top of the hill)]' represents a possessor expression both formally and semantically, whose function is to locate the entity referred to. The same noun form in (37b) represents an embedded predication in subject position.

The category labels in (38) may be applied to both possible meanings of the noun in comparison with the verbal predicate.

(38)   construir   >   construção$_1$ >   construção$_2$
     [+ V – N]   [+V + N]   [-V + N]

Malchukov claims for a cline of deverbalization toward nominalization and toward deverbalization, establishing the relative ranking of features on verbal and nominal hierarchies; further he shows "the relevance of these hierarchies for constraining different types of possible transcategorial operations, such as nominalization" (Malchukov 2004: 56). In the cline of deverbalization, nominalizations only stop describing a state-of-affairs when they change into real first-order nouns. The more external verbal features and their respective layers are

more readily affected by the transcategorial operations than the more internal ones. Since valency is the innermost layer of the verbal lexeme, it is more likely to be preserved in the resulting noun than other features. Now consider the complementary processes of deverbalization and nominalization postulated by Malchukov, in (39).

(39)   *Generalized ScaleModel*
         [[[[[N] CL] NB] POS] DET] Case
          <-------------------------------- [[[[[[V] VAL] ASP] Tense] Mood] AGR] IF]
         nominalization   < --
                                                                          deverbalization
         (Malchukov 2004 : 27)

(At the nominal end, N stands for noun, CL for classifier, NB for number, POS for possessor, DET for determiner; at the verbal end, VAL stands for valency, ASP for Aspect, AGR for agreement and IF for Illocucionary Force.)

The form *construção₂* in (38) undergoes not only the process of decategorization from the verbal lexeme but also the process of nominalization into first-order entity and therefore it involves a real case of valency reduction, with the adoption of all the trappings of a prototypical noun. The form *construção₁* only undergoes the decategorization process, which implies that it keeps the status of a higher-order entity and, consequently, the preservation of the argument structure of the corresponding verbal predicate.

Malchukov does not use the notion of valency reduction in decategorization processes, but instead uses the concept of blocking effects. Thus, in gerund-genitive English forms like *My horse's winning the race*, the blocking effects reach the first argument, while second argument retains the same object marking as the verbal form. In fact, the blocking effects do not reduce the number of grammatical relations of the input verb; instead they are preserved, but with different kind of formal expression. Therefore, the use of prepositions *de* 'of' and *por* 'by' are formal devices of deverbal nouns to make visible the grammatical relations of Subject and Object in the Morphosyntactic Level, which, at least in Portuguese, are generally marked by word order and by person agreement.

This interpretation is supported by Keizer's (2004a) proposal of classification of English prepositions in the framework of FDG. She argues that prepositions can act as grammatical elements (operators) and as lexical elements (predicates). The grammatical function is confined to the prepositions *of* and *by* when introducing terms with those semantic functions which can be assigned to the first and second arguments of verbal and adjectival predicates. Such an approach is in line with Dik's hypothesis of valency preservation (1985, 1997b). Since this type of grammatical operator makes visible semantic functions, it leaves open the possibility of deverbal and deadjectival nominal heads being formally provided with argument structure, as will be clear from the following passage:

> In other words, the grammatical use of prepositions is restricted to those cases where the preposition expresses a semantic function which in the corresponding verbal or adjectival construction is not expressed by means of a preposition but by some (other) grammatical means, such as word order, agreement or case. It seems therefore reasonable to assume that when they appear in the derived construction the prepositions *of* and *by* do not have semantic content, but are indeed simply expressions of the semantic relations in question (Keizer 2004a: 27) .

In (37a), the preposition *de* 'of' establishes a modifying relation with the nominal head; as such it must be analyzed as a lexical preposition and the possession relationship specifies the location of the building. In (37b), on the other hand, *de* 'of' establishes a link between the nominal head and its inherited argument; as such it will be analyzed as a grammatical preposition.

This distinction between grammatical and lexical prepositions also seems to apply to the prepositions *por* 'by', *em* 'in' and *com* 'with' in examples (40a-b) and (41a-b), where the correlation between grammatical preposition for syntactic relations and lexical preposition for modifying relations can be clearly perceived.

(40)a   a   *destru-ição*   *d-a*   *cidade*   *[por epidemia]*
          the.F   destroy-NMLZ   of-the.F   city   [by epidemic]
          'destruction of the city [by epidemic]'

    b   *a*   *destru-ição*   *d-a*   *cidade*   *por*   *tropa-s*   *host-is*
        the.F   destroy-NMLZ   of-the.F   city   by   troop-PL   hostile-PL

        [n-o   *ano*   *passad-o]*
        [in-the.M   year   last-M]
        'the destruction of the city by hostile troops [in the last year]'

(41)a   *o*   *livro*   *[com*   *capa vermelh-a]*
          the.M   book   [with   cover   red-F]
          'the book [with red cover]'

    b   *a*   *preocup-ação*   *de*   *Maria*   *com*   *José*
        the.F   concern-NMLZ   of   Maria   with   José
        'the concern of Maria with José'

Modifiers in the nominal domain, indicated by brackets in examples (40) and (41), can express cause (40a), time (40b) or reference specification (41a); all these modifiers contain information that is not required by the nominal head, whereas information contained by the NPs introduced by *por* 'by' (40b) and *com* 'with'(41b) is required by the nominal head as part of its argument structure.

Note that this approach does not directly affect the interesting correlation between gradual deverbalization and valency reduction described by Mackenzie (1996). This correlation can also be accounted for in terms of Dik's PFA (1985; 1997b). At the more advanced stages of the nominalization process, the grammatical relation markings (word order, subject and object agreement and case) are taken over by other devices, such as the use of prepositions, genitive, possessive pronouns and adjectives, that is, the formal categories that normally follow the prototypical noun, as in (5a-d), rewritten here as (42a-d) for convenience, but with a different interpretation from Mackenzie's.

(42)a   That **my horse** (clausal Arg₁) won **the race** (sentential Arg₂) came as a great surprise. (finite verb)
    b **My horse** (sentential Arg₁) winning **the race** (sentential Arg₂) came as a great surprise. (gerund)
    c **My horse's** (genitive Arg₁) winning **the race** (sentential Arg₂) came as a great surprise. (gerund-genitive)
    d **My horse's** (genitive Arg₁) winning **of the race** (possessor phrase Arg₂) came as a great surprise. (productive nominalization)

This analysis preserves the idea that there is a gradual process of verbal decategorization; this process, however, does not necessarily mean valency reduction. Instead, it is reflected in the different kinds of grammatical relation markings used in

(42a-d). As the verbal predicate gradually gains nominal status, the clausal markings (agreement, word order and the like) give way to prepositional marking or some other device, like adjectival modification or the use of possessive pronouns which are perfectly appropriate to follow a nominal head.

5. A proposal for the analysis of the data

Given that my own interpretation of the data supports the valency preservation hypothesis, it will be clear that I would like to propose an analysis in which derived and non-derived nouns must already be selected from the lexicon with their own valency structure. Pragmatic and semantic motivations trigger the expression form at the Structural Level. As a first step in this analysis we need to consider the way deverbal nouns enter into the Formulation from lexical organization.

García Velasco & Hengeveld (2002) raise some problems that the notion of predicate frame poses with respect to standards of adequacy and the solution they propose is to split up predicate frames into predicates and frames by assigning the term *predication frame* for frame and *lexeme* for predicate. According to them, "the term predicate then becomes available for lexemes used as predicates in particular syntactic contexts. Predication frames specify the (underlying) configurations in which lexemes may occur" (García Velasco & Hengeveld 2002: 107).

Following these notions of predication frames and lexemes as separate theoretical units, I claim that noun lexemes combine with the same predication frame as the verb lexemes from which they are derived only if they are predicates, that is, deverbal nouns representing second-order entities such as *construção* 'construction' in (37b) above mentioned; when lexemes represent first-order entities, such as *construção* 'construction' in (37a), they should be really considered avalent because they are not predicates.

By way of illustration, consider the examples in (43a-b) and their respective predication frames in (44a-b). The quantitative valency is variable according to the nature of the lexeme chosen from the lexicon, which allows the expression of (43-b) in a similar way to that of the corresponding verbal predicate:

(43)a
A	*manifest-ação*	*d-o-s*	*grevista-s*
The.F	demonstrate-NMLZ	of-the.M-PL	striker-PL

caus-ou	*engarraf-amento.*
cause-IND.PRF.3SG	traffic.jam-NMLZ

'The demonstration of the strikers caused some traffic jam.'

b.
A	*destru-ição*	*de*	*Bagdá*	*pel-o*	*exército*	*americano*
the.F	destroy-NMLZ	of	Baghdad	by-the.M	army	american

caus-ou	*a*	*morte*	*de*	*civ-is.*
cause-IND.PRF.3SG	the.F	death	of	civilian-PL

'The destruction of Baghdad by the American Army caused the deaths of civilians.'

(44)a.   (R$_1$ [T   R$_1$] (R$_1$))
         (e$_1$: [(f$_1$: ♦ (f$_1$:)) (x$_1$)$_{Proc}$)] (e$_1$))

b (R$_1$ [T   R$_2$ R$_3$] (R$_1$))
   (e$_1$: [(f$_1$: ♦ (f$_1$:)) (x$_1$) $_{Ag}$ (x$_2$)$_{Go}$)] (e$_1$))

These frames mean that the lexemes *manifestação* 'demonstration' and *destruição* 'destruction' point to a relation (represented here by variable 'f') between two entities (represented by variable 'x'). The presence of these two variables in the representation of noun lexeme, according to García Velasco & Hengeveld (2002: 114), drives the process of linking towards a one-place predicate frame for (43a), and a two-place predicate frame for (43b).

Next, let's consider how to interpret the way the Interpersonal and the Representational levels trigger the expression form of the valency structure at the Structural Level. According to the valency preservation hypothesis, all the alternative forms are provided with full valency when they enter the process of Formulation, allowing the Structural Level to produce variants triggered by information specified at the Interpersonal and the Representational levels.

Mackenzie (2004), on the other hand, argues in favour of a valency reduction hypothesis, which permits distinct frames for intermediate categories at the Representational Level, not at the Interpersonal Level. One of the strongest arguments for his interpretation is that the choices do not depend on what is specified at the Interpersonal Level, since, as is well known, certain selection restrictions determined by the main predicate do not always allow all the alternative forms as embedded predicates. Consider the examples used by Mackenzie (2004), given in (45).

(45) a   I knew/said my horse won the race.
    b   *I knew/said my horse ('s) winning (of) the race. (Mackenzie 2004)

In fact, in a subordination relation the main and dependent predicates are marked by different degrees of interconnectiveness, which may be reflected in a sharing of tense, aspect and mood markings, as well as of participants, a relationship widely known in typological research as semantic integration (Givón 1980, 1990; Noonan 1985).[10]

This evidence of semantic integration seems to show that, in fact, the Representational Level plays an important role in the Formulation before any interference of the Interpersonal Level. It is important to recall that it is at the Representational Level that we can account for cases of zero anaphora motivated by semantic predetermination, a feature directly linked to degree of integration between main and embedded predicates. However, how can we explain cases of zero anaphora motivated by speaker's choice on the basis of referents presumably available in the Addressee's short-term memory, that is, the kind of choices motivated by information specified at the interpersonal level?

Recall that FDG is a multilevel model of grammar consisting of Interpersonal and Representational levels operating at the stage of formulation, and the Morphosyntactic and the Phonological levels operating at the codification. Given that FDG is a top-down model, its efficiency is directly proportional to the way it resembles language production; so, according to Hengeveld (2005), the idea of dynamic implementation calls for Depth First Principle and Maximal Depth Principle; both of them are meant to speed up the implementation of the grammar.

Figure 3 represents the pathways through the grammar. According to Hengeveld, the horizontal arrows concern the consultation of the sets of primitives by the various operations. The dynamic implementation is represented by vertical arrows.
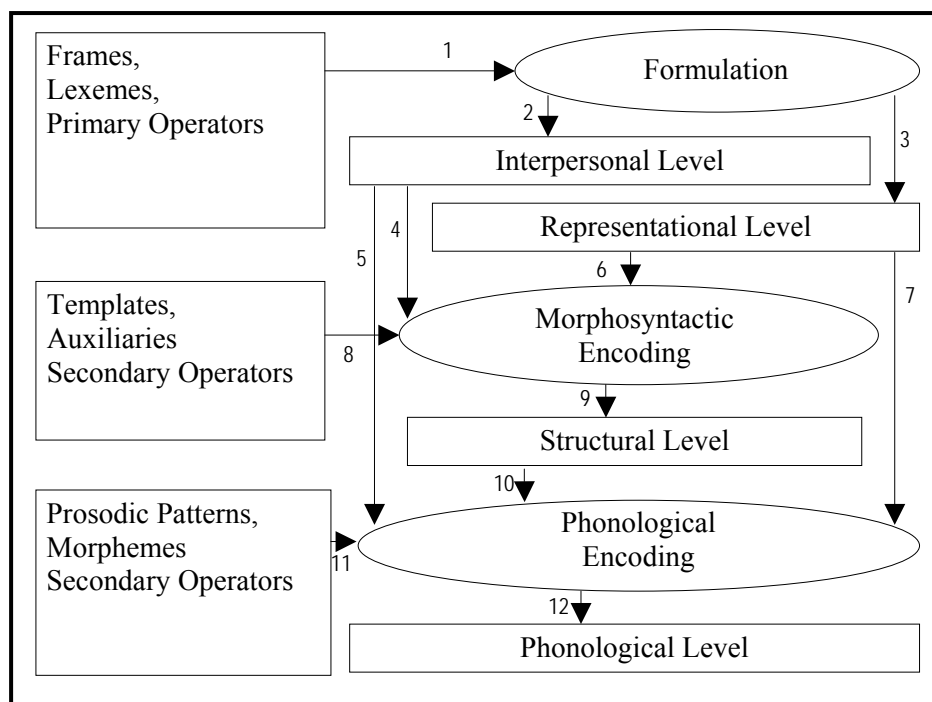
Figure 3: Pathways through the grammar (Hengeveld 2005: 75)

Since FDG represents a modular view of the grammar, pragmatics, semantics, morphosyntax and phonology are developed at independent but interrelated levels. The dynamic implementation provides a path through which short-term information may be accessed, while the relevant complementation choices at the Representational Level are being performed.

Thus, the semantic type of complement triggers the relevant choice of a given dependent construction; at the same time, the speaker consults the Contextual Component to check out which entities are already available to the Addressee in order to provide the Structural Level with the most suitable form. Although the first type of decisions draws on the Representational Level, and the second one draws on Interpersonal Level, the results may be exactly the same at the Structural level, namely, expression by zero anaphora.

The Interpersonal Level, the Representational Level and the Structural Level must be dynamically integrated in such a way that referential information stored in memory, which is contained in the Contextual Component, is readily accessible even after the relevant complementation choices at the Representational Level have been carried out. This way of processing is already predicted by the FDG framework, where the Interpersonal and the Representational levels are conceived of as operating independently from each other, while the Contextual Component may be accessed at any time, irrespective of whether an anaphoric pronoun is used to refer to speech acts, propositional contents, states of affairs or individuals, as well as to linguistic utterances from the previous discourse (Hengeveld 2004a: 3).

Thus, after the semantic type of the complement has triggered the relevant choice of a given dependent construction, the speaker needs to access the Contextual Component to check which entities are available to the Addressee in order to provide the Structural Level with the correct expression form. These decisions all draw on information specified at the Interpersonal Level. Let us resort to some examples of deverbal nouns, given in (46) and (47).

26

(46)a   *eles*   *consegu-em*   *cheg-ar...*   *a*   *é*
          they   succeed.in-IND.PRS.3PL   arrive-INF   to   be.IND.PRS.3.SG

          óbvio   *um-a*   *evolu-ção*   *[* **deles**]*   (EF-SP-405:57)
          obvious   an-F   evolve-NMLZ   [of theirs]

          'it is obvious that they succeed in arriving at... an <u>evolution</u> [*** of theirs**]'

(46)b.   *e eles conseguem chegar... a é óbvio a <u>evoluir</u> certo?*   (EF-SP-405:57)
          'and they can achieve... it is obvious to <u>evolve</u>, right?'

(47)a   *eles*   *tinh-am*   *que*   *acompanh-ar*
          they   have-IND.IPFV.3PL   that   follow-INF

          o   *<u>moviment-o</u>*   ∅   *[=d-o-s*   *anima-is]*   *também:* (EF-SP-405)
          the.M   move-NMLZ   ∅   [= of-the.M-PL   animal-PL]   too

          'they had to follow <u>moving</u> ∅ [= of the animals] 'too'

(47)b.   * *então eles tinham que acompanhar o <u>movimentar</u> também*
          'then they had to follow <u>moving</u> too'

     The construction in (46a) with a nominal predicate would admit an alternative dependent construction with a non-finite verb, as in (46b), but (47a) would not, since the resultant form in (47b) would look like strange, maybe excessively formal, though not exactly ungrammatical.

     After selecting the appropriate lexeme with its respective predicate frame at the Representational Level, zero anaphora is chosen for the expression of the arguments in (46a), whereas in (46b), zero anaphora expression is not semantically predetermined, but motivated by pragmatic reasons, that is, by the information status of referents in discourse. In this specific case, the Contextual Component needs to be checked in a principled way, and with the cooperation of both speech act participants. In this case, it is the evaluation availability of the referent at the Interpersonal Level that finally triggers the expression of zero anaphora to (46b) at the Structural Level.

     As Depth First Principle predicts, cases of semantically predetermined zero anaphora, as contained in (46a) are motivated by the following path through the grammar: 1 → 3 → 6 → 8 → 9 → 10; in this case the Interpersonal Level is circumvented. However, cases of pragmatically determined zero anaphora, such as those contained in (46b), are motivated by the following path through the grammar: 1→ 2 → 4 →8 →9 → 10; now it is the Representational Level that is circumvented.

     Consider now the example in (29c), repeated here as (48) for convenience.

(48)   a   *caça*   *que*   *é*   *o que*   *oferec-e...*
        the.F   hunted   that   be.IND.PRS.3SG   what   offer-IND.PRS.3SG

        uma   *<u>resist-ência</u>* (EF-SP-450)
        a-F   resist-NMLZ

        'the hunted... that is the one that offers... some <u>resistance</u>

     My claim would be that, during Formulation, the two levels work simultaneously to produce (48). However, given that these levels works independently from each other, the paths through grammar are exactly the same as those ones

followed in (46a) and (47a). In (48), on the other hand, they work simultaneously to produce both zero anaphora: the one referring to *caça* 'hunted' is motivated by semantic predetermination; the other, referring to *homens pré-históricos* 'prehistorical men', is pragmatically motivated.

As pointed out before, deverbal nouns have a hybrid nature, given that they may refer to second- or third order entities; in such cases, however, they have their structural expression modeled after prototypical nouns, that is, nouns referring to first-order entities. This raises the question of how they may be represented in the FDG framework.

Mackenzie (2004) argues that operators of identification and specification become more accessible as the nominal status of the embedded predication increases. He also adds that indefinite operators may act on the Referential Subact, as seen in *winning this battle is a winning of all battles* and *there won't be any winning of hearts and minds* (Mackenzie 2004). Thus, to do justice to the ambiguous character of the nominalizations, the most plausible explanation is to consider that, following Mackenzie (2004), their nominal half forces us to represent them as Referential Subacts and their verbal half to represent them as Ascriptive Subacts. By way of illustration, consider the two-place predicate *destruição* 'destruction' in (43b), repeated here as (49a) for convenience.

(49)a.
A	*destru-ição*	*de*	*Bagdá*	*pel-o*	*exército*	*americano*
the.F	destroy-NMLZ	of	Baghdad	by-the.M	army	American

caus-ou	*a*	*morte*	*de*	*civi-s.*
cause-IND.PRF.3SG	the.F	death	of	civilian-PL

'The destruction of Baghdad by the American Army caused the death of civilians.'

For the sake of simplification, let us take into account only the underlying representation of an embedded deverbal noun functioning as the subject of the main predication, as in (49a), in order to account for the mentioned hybrid categorial nature of deverbal nouns. Following Mackenzie (2004),[11] I would like to suggest that deverbal nouns represent a Subact of Ascription at the Interpersonal Level, embedded in a Subact of Reference, as represented in (49b); at the Representational Level they represent a relation between two referential entities (*Bagdá* and *exército americano*), as can be seen in (49c).

A destruição de Bagdá pelo exército americano
(49)b $(R_1 [(T_1)$ (R_2) (R_3) $]$ $(R_1))$
c $(e_1 [(f_1: destruição_N (f_1))$ $(x_1: exército americano(x_1))_{Ag}$ $(x_2: Bagdá (x_2))_{Pat}$ $]$ $(e_1))$

My interpretation differs from the one suggested by Mackenzie (1996) with regard to the question of whether the argument structure is preserved or not during the process of Formulation. I take the position that the way the arguments of embedded predications are expressed depends on pragmatic and semantic motivations as discussed earlier. These two motivations are directly reflected at the Structural Level, since they are responsible for triggering argument expression either as zero anaphora or as full lexemes.

Thus, to do justice to the hybrid character of deverbal nouns, I argue that the morphosyntactic expression must be represented as two layers, which I would call internal syntax and external syntax, following Haspelmath's (1995) representation of

inflectionally derived lexemes. Haspelmath refers to processes of derivation by means of transpositional inflection, which are very productive in a great number of languages. The way of representing this process of transpositional inflection is reflected in the hybrid character of the resultant forms, as illustrated by (50), which represents the German form *singende*.

(50) *der*   *im Wald*   *laut*   *sing$_V$ – ende$_{Adj}$*   *Wanderer*
      the   in:the forest   loud   sing-PTPC   hiker
      'the hiker (who is) singing loud in the forest'   (Haspelmath 1995: 44)

It seems useful to use the same strategy for other hybrid forms, too, which, like the participle form *singende*, which functions as an adjective, are produced by means of common processes of derivation. Considering that the layer called internal syntax refers to the verb form *sing* and the layer called external syntax refers to the participle form *singende*, I will propose the following morphosyntactic representation to the nominal expression *A destruição de Bagdá pelo exército americano* 'The destruction of Baghdad by the American Army':

(51)   [[d1[destru]$_V$ ição]$_N$ [de Bagdá]$_{SP}$] [pelo exército americano]PP] $_{NP}$

The frame in (52) is an attempt to represent the NP in accordance with the first three levels of organization predicted by FDG.

(52)   (R$_1$[(T$_1$)   (R$_2$)   (R$_3$)]   (R$_1$))
        (e$_1$ [(f$_1$)   (x$_1$)$_{Pat}$ (x$_2$)$_{Ag}$]   (e$_1$))
        [[d1[destru]$_V$ ição]$_N$ [de Bagdá]$_{PP}$ [pelo exército americano]$_{PP}$] $_{NP}$

This representation is also inspired by Hengeveld's (2004b) way of representing first-order deverbal nouns, like *teach-er*, as shown in (53). The hybrid character of this type of noun is clearly recognizable from the label Hengeveld assigns to the NP it heads: 'semi-verbal noun phrases':

(53)   a   teach-er
        INDEF   teacher-AG.NMLZ
        (1 x$_i$: (f$_i$: teach-$_V$ (f$_i$)) (x$_i$)$_\emptyset$) (cf. Hengeveld 2004b)

One of the discursive reasons to nominalize is the deverbal nouns' capacity of referring back to other predications already mentioned in the preceding text by means of lexical substitution, as observed in (54).

(54) ***cri-ar***   *um-a*   *pessoa*   *ou*   ***cri-ar***   *um-a*   *imagem*
      create-INF   a-F   person   or   create-INF   an-F   image

      é   *mais*   *ou*   *menos* *a*   *mesm-a* *coisa... n-o*
      be.IND.PRS.3SG   more   or   less   the.F   same-F   thing in-the.M

      sentido   *de*   *que*   *nós*   *est-amos*   ***cri-ando***
      sense   of   that   we   be.DUR-IND.PRS.1PL   create-PROG

      um-a   *coisa*   *nov-a...* *d-o*   *nada...*   *eu*   *não*   *tinh-a*
      a-F   thing   new-F   from-the.M   nothing   I   not   have-IND-IPFV.1SG

      nada   *aqui*   *pass-o*   *a*   *t-er*   *a*   *imagem*
      anything   here   pass-IND.PRS.1SG   to   have-INF   the.F   image

d-a	*minh-a*	*mão...*	*e*	*est-a*	*idéia*	*de*	***cri-a****ção*
of-the.F	my-F	hand…	and	this-F	idea	of	create-NMLZ

é		*que*	*ainda ( )*	*é*		*represent-ação...* (EF-SP-405)
be.IND.PRS.3SG		that	still ( )	be.IND.PRS.3SG		represent-NMLZ

'to **create** a person... or to **create** an image is more or less the same thing... in the sense that we are **creat**ing a new thing... from nothing... I had nothing here before and now I have the image of my hand ... and it is this idea of **creat**ion that still ( ) is representation...'

The underlined nouns constitute a referential chain based only on a common stem, irrespective of its category (verb or noun). This textual function seems to indicate that deverbal nouns can preserve the verbal nature of the input predicate partially, which justifies representing them in two layers at the Structural Level.

The same text relationship and, therefore, the same morphosyntactic representation may be applied to adjectives representing arguments in the valency structure of deverbal nouns, as it can be seen in (55) and (13), repeated here as (56) for convenience.

(55)
Nós	*v-amos*		*começ-ar*	*pel-a*		*Pré-História...[...] a_a*
we	go.AUX-IND.PRS.1PL		start-INF	by-the.F		pre-history… […]

a	*arte...*	*n-o*	*período*	*paleolítico*	*[...] a-s::*	*manifest-açõe-s*
the.F	art…	in-the.M	period	paleolithic	[…] the.F-PL::	manifest-NMLZ-PL

artístic-a-s	*começ-aram*		*a*	*aparec-er*
artistic-F-PL	start-IND.PRF.3PL		to	appear-INF

n-o	*paleolítico*	*superior* (EF-SP-405)
in-the.M	paleolithic	upper

'we are going to start from pre-history...[...] the art... in the paleolithic period [...]:: artistic manifestations started to appear in the upper paleolithic period'

(56)
tod-a	*a*	*evolu-ção*		***human-a****...*	*não*	*deix-a*
every-F	the.F	evolve-NMLZ		human-F	not	let-IND.PRS.3SG

de	*s-er*	*exatamente*	*a*	*evolu-ção*	*d-o*
of	be-INF	exactly	the.F	evolve-NMLZ	of-the.M

domíni-o	*que*	*o*	*homem*	*t-em*	*sobre*
control-NMLZ	that	the.M	man	have-IND.PRS.3SG	over

a	*natureza...* (EF-SP-405)
the.F	nature

'every **human** evolution... is simply no more than the evolution of the control man has over the nature...'

The expression of the first argument in the form of an adjective represents a cohesion relation between a preceding noun in (55) and a subsequent noun in (56); in addition to this kind of textual recovering, Keizer (2004b) mentions other kinds of resumption, illustrated in (57):

(57) In Paris Francine Stock examines the thinking behind the last minute *French* initiative what *they* hope do gain from it. <ICE-GB:S2B-010 #7:1:1> (Keizer 2004b: 12).

Although certain adjectives, like *French* in (57, can be used to indicate a participant in a state of affairs, Keizer argues that "they do so withhout taking argument position (...)"; despite the fact that "anaphoric reference to the implied agent is possible"(Keizer 2004b:12). In my opinion, the possibility of using both nouns ans adjectives to perform the same function of textual cohesion provides strong evidence that, in constructions like (57), adjectives, like nouns, do fill argument slots and thus do not perform the common function of modifying (first-order) nouns (as, for instance, the adjective *humana* in *ferramenta humana* 'human tool').

6. Conclusion

In answering the question of whether deverbal nouns are truly valent, as in Dik's valency preservation hypothesis or, whether they are avalent, as in Mackenzie's valency reduction hypothesis, I have argued in favour of the former position, providing empirical evidence of a quantitative nature based on spoken language. The main conclusions can be summarized as follows:

1. There are pragmatic and semantic motivations for not expressing overt arguments in the deverbal noun internal syntax; even so the referents of these arguments may be found in the different parts of the text, forming a referential chain.
2. This kind of evidence has led to the claim that full valency must be represented in the predicate frame, as advocated by the valency preservation hypothesis.
3. The arguments favouring the valency preservation hypothesis do not affect Mackenzie's main typological findings about the gradual process of nominalization.
4. These text-based data suggest that the linking between lexemes and predicate frames must be developed on the basis of internal and external syntax when representing deverbal nouns.
5. At the internal syntax level, a deverbal noun predicate occupies the structural slot of a verbal predicate only provisionally and the same is true for the adjectival argument in relation to a noun. At the external syntax level, they play the functions of nominal head and modifier respectively. Therefore, the formal adjustments referred to by Dik (1985; 1997b) apply only to the external syntax level.
6. From a theoretical point of view, the position assumed here does not go against any of the FDG postulates and, taking seriously into account scientific principles of economy and simplicity, does not allow for the proliferation of such theoretical concepts as y and z-satellites.
7. As a final advantage, by preserving the intuitive correspondence between derived nouns and input verbs, the valency preservation hypothesis provides a more direct link between theory and data.

Notes

1. I am very grateful to J. Lachlan Mackenzie, Kees Hengeveld and specially to Evelien Keizer for insightful comments on earlier draft versions of this paper. Of course, I am fully responsible for the shortcomings that still remain.
2. The tool I use to make a literal translation of the Brazilian Portuguese examples to English is Leipzig Glosses Rules. I will, however, skip the majority of zero representations especially nominal inflection of number, which will be restricted to plural marking. Note that since the examples are extracted from a sample of spoken language, they contain special signals like **::** for segment duration, **…** for hesitation, **(…)** for segments not clearly understood and so on. I will reserve the use of square brackets to indicate that some fragments were omitted for sake of simplicity.
3. Since the distinction between y-satellites and z-satellites may be considered purely semantic in nature, a satellite of either kind may be communicatively salient. In fact, in Mackenzie' s point of view, satellites are typically absent and introduced only if they are communicatively salient. But in my own approach, since noun valency is a propriety already predicted by the item chosen in the lexicon, it is the pragmatic status of New or Given Topic that applies to the argument structure; so, the notions of y- and z-satellites are completely precluded.
4. Since Mackenzie assumes an avalent frame for derived nouns in the Fund, his approach does not embody the operation of valency reduction preceding the attachment of satellites. Whether or not such a view is counterintuitive depends on the way we can see the lexicon; in my own approach there must be a semantic correspondence between the derived noun and the input verb and consequently a correspondent frame between both lexical categories. However, in an approach like Mackenzie's this question of processing would also depend on the number of cases in which arguments are expressed compared to the number of cases they are not. It is true that if in the majority of cases the arguments are not expressed¸ the cognitive effort involved in adding them as satellites would be smaller than the effort involved in processing the fact that certain argument positions are not being filled.
5. Observe that in Mackenzie's approach it is the y-satellite that 'wins' in such cases.
6. Of course we are assuming Dik's principle of formal adjustments, as represented in Fig 1. In the FDG framework, this kind of adjustments is addressed at the Structural Level, where they are regarded as being triggered by information about lexical cohesion specified at the Representational Level (see example (55)). According to Mackenzie, however, these elements are not arguments but adjectives, which can simply be added, as to any noun.
7. The terms 'expressed' and 'not expressed' for A_1 and A_2 mean that the first and the second arguments are either expressed or not expressed around the nominal head.
8. In Mackenzie's approach, the behaviour here described would be typically of **satellites,** not arguments.
9. Referential noun phrases are provided with the following features: (i) is headed by a noun; designates a first-order entity; (iii) designates through lexical means; (iv) is used referentially (Hengeveld 2004b).
10. According to Hengeveld (1999:8), the underlying structure of the complement clause depends semantically on the main predicate. Thus verbs denoting acts of speaking, such as *say*, take a subordinate clause as complement, verbs expressing propositional attitude, such as *believe*, take a proposition subordinate as complement and verbs, such as *bother*, which express the effect of a states-of-affairs on experiencers, take a subordinate predication as complement. This dependence relationship triggers strong predictions related to alternate possibilities to express TMA operators: the more outer the layer of a complement clause, the more probably it will be expressed as an independent verbal form.
11. Note that I insist on this view, despite the fact that I am aware that Mackenzie himself is no longer sure whether it would make sense to talk of Subacts embedded within other Subacts.
12. In Portuguese, there is no correspondence between the verb *iniciar* (more or less translated as 'initiate') and the noun *iniciativa* 'initiative', despite the fact that they are true cognates. Unlike English, Portuguese *iniciativa* can only be used in combination with a light verb like *tomar* 'to take'. In this sense the adjective could, however, still be regarded as an argument of the noun *iniciativa*.

Appendix: List of Abbreviations

1	first person	M	masculine
2	second person	NMIZ	nominalizer
3	third person	PL	plural
AUX	auxiliary	PRF	perfect
DUR	durative	PRS	present
F	feminine	PROG	progressive
FUT	future	PST	past
IND	indicative	PTCP	participle
INF	infinitive	REFL	reflexive
IPFV	imperfective	SBJV	subjunctive
IRR	irrealis	SG	singular

References

Camacho, R.G. & L. Santana (2004). Argument structure of deverbal nouns in Brazilian Portuguese. *Journal of Language and Linguistics* 3,2 :229-42.

Castilho, A. T. & Preti, D. (eds.) (1986). *A linguagem falada culta na cidade de São Paulo*: materiais para seu estudo. São Paulo: T. A. Queiroz, v. 1.

Castilho, A. T. & Preti, D. (eds.) (1987). *A linguagem falada culta na cidade de São Paulo*: materiais para seu estudo. São Paulo: T. A. Queiroz, v. 2.

Cristofaro, S. (2003). *Subordination*. Oxford: Oxford University Press.

Dik, S. C. (1985). Formal and semantic adjustment of derived constructions. In: A.M. Bolkstein et al. (eds). *Predicates and terms in Functional Grammar*. Dordrecht/Cinnaminson: Foris, p. 1-28.

Dik, S. C. (1997a). *The theory of Functional Grammar: Part I: The structure of the clause*, (2nd revised edition) ed Kees Hengeveld. Berlin and New York: Mouton de Gruyter.

Dik, S. C. (1997b). *The theory of Functional Grammar: Part II*.: *Complex and Derived Constructions*. (2nd revised edition) ed Kees Hengeveld.. Berlin and New York: Mouton de Gruyter.

García Velasco, D. & K. Hengeveld (2002). Do we need predicate frames? In: R. Mairal Usón & M.J. Pérez Quintero (eds). *New Perspectives on Argument Structure in Functional Grammar,* Berlin: Mouton de Gruyter, p. 95-123.

Givón, T. (1980). The binding hierarchy and the typology of complements. *Studies in Language* 4-3: 333-377.

Givón, T. (1990). *Syntax: a functional-typological introduction*. Amsterdam/Philadelphia: John Benjamins.

Haiman, J. (1983). Iconic and economic motivation. *Language* 59: 781-819.

Haspelmath, M. (1995). Word-class-changing inflection and morphological theory. *Yearbook of Morphology*. Kluwer Academic Publishers, p. 43-66.

Hengeveld, K. (1992). Non-verbal predication. Theory, typology, diachrony. Berlin: Mouton de Gruyter.

Hengeveld, K. (1999). Formalizing functionally. In: M. Darnell et al. (eds.) *Functionalism and formalism in linguistics*. Vol. 2: (Studies in Language Companion Series 42), 93-105. Amsterdam: Benjamins.

Hengeveld, K (2004a). *The architecture of a Functional Discourse Grammar*. In J. Lachlan Mackenzie & María Á. Gómez-González eds, *A new architecture for Functional Grammar* (Functional Grammar Series 24), 1-21. Berlin: Mouton de Gruyter Amsterdam.

Hengeveld, K (2004b). *Noun phrases in Functional Discourse Grammar*. Paper presented at 11th International Conference on Functional Grammar. Gijón, Spain.

Hengeveld, K. (2005). Dynamic Expression in Functional Grammar. In: de Groot, Casper, Hengeveld, K. (eds.) *Morphosyntactic Expression in Functional Grammar*. Berlin/New York: Mouton de Gruyter, p. 53-86).

Hengeveld, K. & J.L. Mackenzie (2006). Functional Discourse Grammar. In: Keith Brown ed., *Encyclopedia of Language and Linguistics,* 2nd Edition, Volume 4, 668-676. Oxford: Elsevier.

Hopper, P. J. & S.A. Thompson (1984). The discourse basis for lexical categories in universal grammar. *Language* 60: 703-752.

Keizer, E. (2004a). English Prepositions in FDG. Paper presented at *11th International Conference on Functional Grammar*. Gijón: Spain.

Keizer, E. (2004b). Term structure in FG: a modest proposal. *Working Papers in Functional Grammar* 78.

Koptjevskaja-Tamm, M. (1993). *Nominalizations*. London/New York: Routledge.

Leipzig Glossing Rules. Conventions for interlinear morpheme-by-morpheme glosses http://www.eva.mpg.de/lingua/pdf/LGR01Jan04.pdf

Mackenzie, J.L. (1985). Nominalization and valency reduction. In: A.M. Bolkestein et al. (eds). *Predicates and terms in Functional Grammar*. Dordrecht/Cinnaminson: Foris, p. 31-51.

Mackenzie, J.L. (1996). English nominalizations in the layered model of the sentence. In: B. Devriend, L. Goossens & J. Van der Auwera (eds). *Complex Structures: A Functionalist Perspective*. Berlim/New York: Mouton de Gruyter, p.325-55.

Mackenzie, J.L. (2002). The Predicate in Functional Grammar. In: R. Mairal Usón & M.J. Pérez Quintero (eds). *New perspectives on Argument Structure in Functional Grammar*. Berlin/New York: Mouton de Gruyter, p.1-38.

Mackenzie, J.L. (2004). FDG perspectives on nominalization. Paper presented at the *11th International Conference on Functional Grammar*. Gijón: Spain.

Malchukov, A. L. (2004). *Nominalization/verbalization: constraining a typology of transcategorial operations*. Lincom: Lincom Europa.

Noonan, M. (1985). Complementation. In: T. Shopen (ed.) *Language typology and syntactic description*. Vol II: Complex constructions. Cambridge: Cambridge University Press, p. 42-141.

Preti, D. & H. Urbano (eds.) (1988). *A linguagem falada culta na cidade de São Paulo*: V.3: materiais para seu estudo. São Paulo: T. A. Queiroz/FAPESP.

Rijkhoff, J. (1992). *The noun phrase. A typological of its form and structure*. University of Amsterdam: PhD dissertation**.**

Santana, L. (2005). *A expressão da estrutura argumental dos nomes derivados*. University of São Paulo State: M.A.dissertation.