# The Lexicon-Grammar of Predicate Nouns with *ser de* in Port4NooJ

Cristina Mota[1(✉)], Jorge Baptista[1,2], and Anabela Barreiro[1]

[1] L2F/INESC-ID, Lisbon, Portugal
`cmota@ist.utl.pt`, `anabela.barreiro@inesc-id.pt`
[2] Universidade do Algarve, Faro, Portugal
`jbaptis@ualg.pt`

**Abstract.** This paper provides continuity for previous efforts on the integration of complementary lexicon-grammars to expand the paraphrastic capabilities of Port4NooJ, the Portuguese module of NooJ (Silberztein 2016). We describe the integration of the lexicon-grammar of 2,085 predicate nouns, which co-occur in constructions with the support verb *ser de* 'be of' in European Portuguese, such as in *O Pedro é de uma coragem extraordinária* 'Peter is of an extraordinary courage', studied, classified and formalized by Baptista (2005b). This led to a 20% increase in the number of predicate nouns. We also extended previously created paraphrasing grammars, such as the grammars that paraphrase symmetric predicates, as well as the grammars that handle the substitution of the support verb by another support verb. Furthermore, we created new grammars to paraphrase negative constructions, appropriate noun constructions, adjectival constructions, and manner sub-clauses. The paraphrastic capabilities acquired have been integrated in the eSPERTo system.

## 1 Introduction

Previous research has shown that the distributional and transformational properties of adjectival and nominal predicates contained in lexicon-grammar tables can be used in paraphrasing tasks with successful results. Mota et al. (2016) describe the integration of the lexicon-grammar of human intransitive adjectives formalized by Carvalho (2007) and Mota et al. (2017) describe the integration of the lexicon-grammar of the predicate nouns co-occurring with the support verb *fazer* 'do' or 'make' formalized by Chacoto (2005).

This paper provides continuity for previous efforts on the integration of complementary lexicon-grammars to expand the paraphrastic capabilities of Port4NooJ, the Portuguese module of NooJ (Silberztein 2016). We describe the integration of the lexicon-grammar of 2,085 predicate nouns, which co-occur in constructions with the support verb *ser de* 'be of' in European Portuguese, such as in *O Pedro é de uma coragem extraordinária* 'Peter is of an extraordinary courage', studied, classified and formalized by Baptista (2005b). Many of these predicate nouns correspond to the nominalization of adjectival constructions,

so in those cases they are linked to the corresponding adjective in the lexicon-grammar table. This presents a major challenge in terms of integration into Port4NooJ, as in 55% of the cases where the predicate nouns have an equivalent adjectival construction, the adjective is homograph of a human intransitive adjective, already formalized in the lexicon-grammar of human intransitive adjectives. This means that one must find a way of harmonizing these entries so as not to have duplicates. However, for now, we did not tackle this challenge and will not discuss it further.

The predicate nouns that occur with the support verb *ser de* were classified into nine classes. The criteria to establish those classes are described in Sect. 3. The structural, distributional, and transformational properties represented in the lexicon-grammar table for each predicate noun led to the generation of a variety of paraphrases, which are achieved by different lexico-syntactic processes. One of these processes is the interchangeability with other elementary support verbs, such as *ter* 'have' or *haver* 'there be', and more rarely with *fazer* 'do' or 'make'. Some paraphrastic relations have also been established based on the (i) insertion of negation/negative prefixes (mainly *des-* and *in-*) on predicate nouns, (ii) complementary but analytical mechanism of negation involving the expression *falta de* 'lack of', (iii) NP restructuring, (iv) predicate nouns that can function as an adnominal adjunct of a common human noun, (v) predicate nouns most likely obtained from the reduction of complex sentence with a relative clause, and (vi) symmetric predicate noun constructions.

As in previous integration efforts, the resources gathered are to be incorporated into the eSPERTo paraphrasing system[1] and to be used in a wide variety of software applications, so as to allow rephrasing a sentence or text using different words. The evolution of eSPERTo's resources, starting from the original Port4NooJ to the current version, which integrates several lexicon-grammars, among other resources, is illustrated in Fig. 1. The combination of the different resources enables a large number of paraphrases that are increasing in volume and improving in quality. We plan to enlarge our growing database of pairs of paraphrastic units by converting more tables built (or to be built) for Portuguese or other languages at the same time as we evolve our integration of crowdsourcing and machine learning techniques to acquire raw data that require validation by linguists before integrating into the paraphrasing engine. Our methodology is based on grammar and other processes that we believe are the ones used by the human brain in processing language (Barreiro et al. 2011; Scott 2018).

## 2   Related Work

In our several scientific papers on the work towards increased integration of lexicon-grammar tables, we have referred to related published works on support verb constructions studied within the lexicon-grammar theory (Gross 1975,

---

[1] eSPERTo, which means 'smart' in Portuguese, stands for "**S**ystem of **P**araphrasing for **E**diting and **R**evision of **T**exts" ("**S**istema de **P**arafraseamento para **E**dição e **R**evisão de **T**exto").
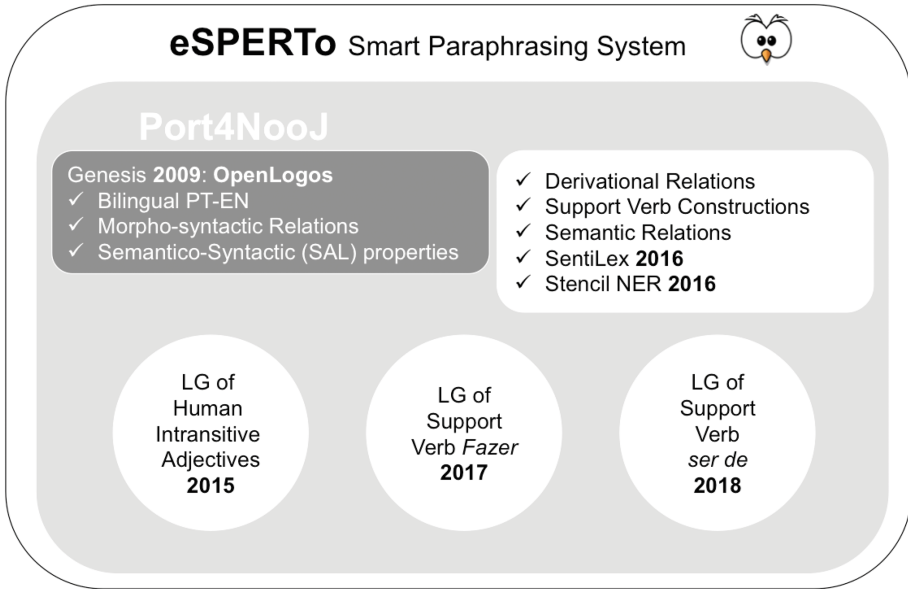
**Fig. 1.** Evolution of eSPERTo paraphrasing system

1996), for several Romance languages other than Portuguese (D'Agostino and Elia 1998; Laporte and Voyatzi 2008; Silberztein 1993), and contrastive studies aiming at machine translation between English and French (Salkoff 1990). Many of the studies within the lexicon-grammar framework focus on formalization of multiword units, most of them on the topic of support verb constructions. As for the Portuguese language, several detailed studies on support verb constructions have become an established theme in development cooperation, which have been presented at the last four International NooJ Conferences, addressing the integration of relations between support verbs and nominalizations (normally autonomous predicate nouns) or predicate adjective constructions. Each one of those papers also illustrates how the properties contained in the addressed lexicon-grammar tables are used in paraphrasing tasks.

The very first integration experiment by Mota et al. (2016) concerned the integration of the lexicon-grammar of human intransitive adjectives into the Port4NooJ module. After the success of this integration, we have started to integrate complementary lexicon-grammars, first the lexicon-grammar of predicate nouns, which co-occur with the support verb *fazer* (Mota et al. 2017), and now the lexicon-grammar of predicate nouns co-occurring with *ser de*. Several classes and subclasses have been defined for each one of these lexicon-grammars based on distributional and transformational properties of the predicate nouns with which each one of these verbs co-occurs. The main properties of the transformations of *ser de* that we represented in paraphrasing grammars are described

in Sect. 3. As far as we know, no other similar integration efforts have been done for any other language, at least within the NooJ framework[2].

## 3   Lexicon-Grammar of Predicate Nouns with $V_{sup}$ *ser de*

The lexicon-grammar of predicate nouns occurring in constructions with the support verb *ser de* 'be of' in European Portuguese, such as in *O Pedro foi de uma ajuda inestimável para a Ana* 'Pedro was of an invaluable help to Ana' was formalized by Baptista (2000, 2005b) after studying and classifying the structural, distributional, and transformational properties of 2,085 predicate nouns. The author identified seven classes of predicate nouns according to (i) the number or arguments (1 or 2) selected by the predicate noun, (ii) the syntactic (sentential/nominal) constraints, and (iii) the distributional (semantic) selection constraints on the nominal argument slots (human/non-human). Two special classes were established for: (i) nouns selecting a body-part noun as their subject, such as *Os músculos da Ana são de uma tonicidade impressionante* 'Ana's muscles are of an impressive tonicity', and (ii) constructions that have an equivalent symmetric construction, i.e., allow swapping the predicate noun's subject with its complement, such as *A Ana é de uma grande parecença com a irmã* 'A Ana is of a great resemblance to her sister'. Table 1 presents the breakdown of these nouns by classes and their basic syntactic structure. Of the many transformations formalized in this lexicon-grammar we chose a few to start creating paraphrasing grammars. In Sects. 3.1–3.7, we will briefly describe some of its properties.

### 3.1   Symmetry Restructuring

In symmetric predicates (see Baptista (2005a) for an overview in Portuguese), their two arguments have the same semantic role in relation to the predicate and, therefore, they can be swapped in their syntactic slots without changing the overall meaning of the sentence, e.g., *A aldeia é de uma proximidade à praia muito grande = A praia é de uma proximidade muito grande à vila* 'The village is of a great proximity to the beach' = 'The beach is of a great proximity to the beach'; or be coordinated in the same syntactic slot (the subject), e.g., *A vila é de uma grande proximidade à praia = A vila e a praia são de uma grande proximidade (uma da outra)* 'The village is of a great proximity to the beach' = 'The village and the beach are of a great proximity (to each other)'.

### 3.2   Support Verb Variants

Usually, the support verb occurring with a predicate noun can be replaced by lexical variants. However, that is not the case with *ser de* 'be of' which has no

---

[2] In Rassi et al. (2014, 2015), a proposal is presented for the integration of predicate noun constructions into STRING (Mamede et al. 2012), a hybrid, rule-based and statistical, natural language processing chain, specifically developed for the processing of Portuguese.

**Table 1.** Distribution of nominal predicates with support verb *ser de* by class attribute

| Class | Count | Percentage | Structure example |
|---|---|---|---|
| SdH1 | 388 | 19% | *(Nhum)$_0$ ser de N* |
| | | | O Zé é de uma alegria contagiante |
| | | | Zé is of a contagious joy/happiness |
| SdH2 | 54 | 3% | *(Nhum)$_0$ ser de N Prep N$_1$* |
| | | | O Zé é da confiança da Ana |
| | | | Zé is from Ana's trust |
| SdNH1 | 363 | 17% | *(Nnhum)$_0$ ser de N* |
| | | | Este molho é de uma acidez exagerada |
| | | | This sauce is of an exaggerated acidity |
| SdNH2 | 30 | 1% | *(Nnhum)$_0$ ser de N Prep N$_1$* |
| | | | Esta substância é de uma total indissolubilidade em água |
| | | | This substance is of a total indissolubility in water |
| SdNPC | 30 | 1% | *(Npc de Nhum)$_0$ ser de N* |
| | | | O rosto da Ana era de uma palidez doentia |
| | | | Ana's face was of a sick pallor |
| SdQ0 | 820 | 39% | *QueF$_0$ ser de N* |
| | | | Essa medida é de grande abrangência |
| | | | This measure is of broad coverage |
| SdQ1 | 308 | 15% | *QueF$_0$ ser de N Prep N$_1$* |
| | | | O Zé foi de uma agressividade desproporcionada para com a Ana |
| | | | Zé was disproportionately aggressive towards Ana |
| SdQ2 | 37 | 2% | *N$_0$ ser de N Prep QueF$_1$* |
| | | | O Zé é de uma grande habilidade para tratar das roseiras |
| | | | Ze is of great ability to treat rose bushes |
| SdSIM | 55 | 3% | *N$_0$ ser de N Prep N$_1$* [symmetry] |
| | | | O Zé e a Ana são de um companheirismo exemplar |
| | | | Zé and Ana are of exemplary companionship |
| Total | 2085 | | |

stylistic nor aspectual variants. Nonetheless, these predicate nouns may have equivalent constructions with other elementary support verbs, mostly *ter* 'have', *haver* 'there be', and, more rarely, with *fazer* 'do, make'[3] (see examples in Table 2).

---

[3] The construction with *fazer* is not strictly a paraphrase of the construction with *ser de*, as the former only characterizes the human gesture, but not the human quality expressed by the constructions with *ser de*.

**Table 2.** Paraphrasing *ser de* variants

| Type | Paraphrases |
|---|---|
| [Vsup=ter] | *O Pedro teve uma grande gentileza (para com a Maria)* |
| | *Pedro had a great kindness (towards Maria)* |
| [Vsup=fazer] | *O Pedro fez uma grande gentileza (à Maria)* |
| | *Pedro did a great kindness (to Maria)* |
| [Vsup=haver] | *Há no Pedro uma grande gentileza (para com a Maria)* |
| | *There is in Pedro a great kindness (with Maria)* |

### 3.3   Nominalizations

An important source for the analysis of paraphrastic relations among sentences is *nominalizations*, that is, equivalence relations between sentences with a predicate noun and a support verb, on the one hand, and a verb or an adjective (and its auxiliary verb), on the other hand. For the most part, the predicate nouns with support verb *ser de* correspond to the *nominalization* of adjectival constructions, e.g. *O Pedro foi de uma grande crueldade para com o João* 'Pedro was of a great cruelty towards João' = *O Pedro foi muito cruel para com o João* 'Pedro was very cruel to João'. More rarely, a verbal construction can be found: *O Pedro foi de uma grande compaixão para com o João* 'Pedro was of (=had) a great compassion for João' = *O Pedro compadeceu-se do João* 'Pedro took pity on João'. It should be noted, however, that the lexical-morphological relation between a predicate noun and a verb, or between a noun and an adjective, is necessary but insufficient to establish a nominalization, with a transformational status, in the sense of Gross (1981), Harris (1981). Not only the meaning of the sentences being related must be the same, but the distributional constraints of the predicate noun on its argument domain must be similar. Establishing such paraphrastic status, thus, requires highly granular, and systematic linguistic description.

### 3.4   Negation

In the lexicon-grammar of *ser de*, two types of negation were formalized: (i) prefixation, i.e., possibility of adding/removing a negative prefix like *des-* or *in-* to the predicate noun; and (ii) analytic negation involving the expression *falta de* 'lack of'[4]. Table 3 illustrates these two types of negation and the paraphrastic relation between them was represented in grammars displayed in Sect. 4.2. Whenever significant differences in meaning are found between the predicates starting with different negation prefixes, the prefixed and the base forms were treated as independent lexicon-grammar entries.

---

[4] Most of the predicates having this property express human qualities/attributes.

**Table 3.** Paraphrasing negation

| Type | Paraphrases |
|---|---|
| [in-N] | *O Pedro é de uma certa (in-)tolerância à lactose* |
| | *Pedro is of a certain (in-)tolerance to lactose* |
| [falta de N] (lack of N) | *O Pedro é de uma certa (falta de) tolerância à lactose* |
| | *Pedro is of a certain (lack of) tolerance to lactose* |

### 3.5   Appropriate Nouns and NP Restructuring

The subject of the sentences with *ser de* 'be of' is many times a complex noun phrase (NP) whose head is also a predicate noun. This head may occur with its arguments, particularly its semantic/notional 'subject' argument, in the form of a prepositional phrase (PP) introduced by the preposition *de* 'of': [*A disposição destes objetos*] *é de uma certa assimetria* '[The placement of these objects] is of a certain asymmetry'. An *appropriate* relation, in the sense of Gross (1981) [pp. 113–115], Harris (1976), usually exists between the predicate noun in the subject position and the sentence's main predicate, and a *NP restructuring* operation (Guillet and Leclère 1981) is then found, which splits the noun phrase into two distinct constituents: (i) the head of the PP becomes the sentence subject, while (ii) the predicate noun (formerly the head of the subject NP) is moved to a new PP, usually at the end of the sentence. For example, in [*Estes objetos*] *são de uma certa assimetria* [*na sua disposição*] '[These objects] are of a certain asymmetry [in their placement]', the reference of the possessive determiner in the PP (*sua* 'their', in the example) is constrained, and it has to refer to the subject of the predicate noun; obviously, this possessive can not be derived from a free PP, non-co-referent to the subject. The semantic roles of the elements are then kept exactly the same in spite of the formal changes the sentence undergoes.

### 3.6   Manner Sub-clauses Restructuring

An interesting distributional constraint regards the subject NPs with manner operator nouns (Gross 1975) *forma*, *maneira* and *modo* 'manner/way' (in Brazilian Portuguese, there is also the operator noun *jeito* '*idem*'). These operators can be construed with an *infinitive* sub-clause complement introduced by preposition *de* '*of*': *A forma/a maneira/o modo* **de** *o Pedro* **fazer** *isso é de uma arrogância impressionante* 'The way **of** Pedro **doing** this is of an impressive arrogance'; or a *pseudo-relative*, *finite* clause, introduced by the so-called interrogative adverb *como* 'how': *A forma/a maneira/o modo* **como** *o Pedro* **faz** *isso é de uma arrogância impressionante* 'The way **how** Pedro **does** this is of an impressive arrogance'. When the predicate noun accepts these operator nouns, these sentences qualify the way a process takes place or the manner in which an action is performed, rather than expressing the attributes of a person or an object. A similar *NP restructuring* as seen above (Sect. 3.5) operates on these manner constructions, splitting the complex NP, extracting the subject of the subordinate

clause to the subject of the predicate noun and leaving the operator noun as a PP manner complement: *O Pedro é de uma arrogância impressionante em_a forma/a maneira/o modo de fazer isso* 'Pedro is of an impressive arrogance in the way of doing that' *O Pedro é de uma arrogância impressionante em_a forma/a maneira/o modo como faz isso* 'Pedro is of an impressive arrogance in the way [he] does that' (Table 4).

**Table 4.** Paraphrasing manner sub-clauses

| Type | Paraphrases |
|---|---|
| [*RestNopQueF*] | *O modo de o Pedro agir é de uma certa teimosia* |
| | *Pedro's way of acting is of a certain stubbornness* |
| | *O Pedro é de uma certa teimosia no seu modo de agir* |
| | *Pedro is of a certain stubbornness in his way of acting* |
| [*RestNopQueF*] | *O modo como o Pedro age é de uma certa teimosia* |
| | *The way Pedro acts is of a certain stubbornness* |
| | *O Pedro é de uma certa teimosia no modo como age* |
| | *Pedro is of a certain stubbornness in the way he acts* |

### 3.7 Reduction of Finite Sub-clause to Infinitive and Restructuring

Other sub-clause transformations have also been represented in the lexicon-grammar, namely the possibility of an infinitive construction with operator-noun *facto* (fact), the reduction of a finite sub-clause to an infinitive and its restructuring in a way similar to the NP restructuring transformations seen above. Table 5 illustrates these phenomena.

## 4 Integration of Lexicon-Grammar Tables in Port4NooJ

As shown in Mota et al. (2017) and Mota et al. (2016), the integration of lexicon-grammars in Port4Nooj is a two-step process. First, one converts the lexicon-grammar tables into NooJ dictionary format, and then one builds grammars that use the linguistic knowledge encoded in the lexicon-grammar tables to identify relevant sentences or phrases and generate their corresponding paraphrases.

### 4.1 From Lexicon-Grammar Tables to NooJ Dictionaries

The procedure described in Mota et al. (2017) to convert the lexicon-grammar entries that occur with support *fazer* into a standalone dictionary was adopted,

**Table 5.** Paraphrasing of finite sub-clauses to infinitive

| Type | Paraphrases |
|---|---|
| [Nop facto] | *É de uma completa evidência que a Susana está a mentir* |
| | *It is of a complete evidence that Susana is lying* |
| | *É de uma completa evidência o facto de a Susana estar a mentir* |
| | *It is of a complete evidence the fact that Susana is lying* |
| [Vinf] | *É de uma certa teimosia que o Pedro aja assim* |
| | *It is of a certain stubbornness that Pedro acts in that way* |
| | *É de uma certa teimosia o Pedro agir assim* |
| | *Is of a certain stubbornness Pedro acting that way* |
| [Rest Vinf] | *O Pedro é de uma certa teimosia em agir assim* |
| | *Pedro is of a certain stubbornness in acting in this way* |

with minor adjustments, to convert the lexicon-grammar entries that occur with support verb *ser de*. The procedure is illustrated in Fig. 2.[5]

The minor adjustments were related to particularities of the lexicon-grammars properties or format that do not interfere with the overall procedure, but only with the way lexicon-grammar properties are converted into dictionary attributes. For example, the property PfxNeg can be filled in the lexicon-grammar entries with "-" when the predicate does not accept a negative prefix or with the value of the negative prefix, e.g., in-, des-.

The new standalone dictionary is comprised of 2,134 predicate noun entries that occur with Vsup *ser de*, corresponding to 1,376 different lemmas. Additional 797 entries await revision to be added to this version of the dictionary. They need revision of the inflectional codes, of derived adjectives or have problems with their format. The integration of this lexicon-grammar led to the creation of 450 new derivational paradigms, but there might be an overlap with paradigms created when integrating the lexicon-grammar of constructions with the support verb *fazer*.

50% of the predicate nouns in the lexicon-grammar already existed in the main dictionary of Port4Nooj. This corresponds to a 6% increase in nominal entries and 20% increase in predicate nouns. In 55% of the cases where the predicate nouns have an equivalent adjectival construction, the adjective was homograph of a human intransitive adjective, already formalized in the lexicon-grammar of human intransitive adjectives Carvalho (2007). This overlap implies harmonization of entries in order to eliminate duplicates.

It is also worth noting that 4% of the predicate noun lemmas (52) that occur with support verb *ser de* are homographs of predicate noun lemmas that occur

---

[5] As in Mota et al. (2017), the procedure allows to check whether words are compliant with the Portuguese Orthographic Agreement, but since all the words are pre-Agreement we are not using that feature and, hence, it is not described in the figure.
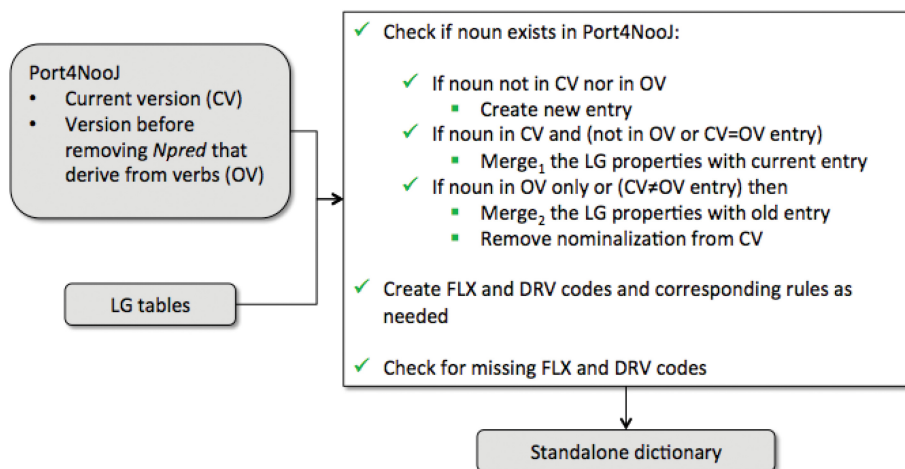
**Fig. 2.** Integration of LG entries in Port4NooJ

with support verb *fazer*, which had been previously integrated in Port4NooJ. Some predicates correspond to the same construction, like *O Zé é de uma patetice impressionante* 'Zé is of an impressive goofiness' seems a paraphrase of *O Zé fez uma patetice* 'Zé did a goofiness', which can be confirmed also by the fact that the entry for *patetice* accepts the Vsup *fazer* as a valid substitution for the Vsup *ser de*. Other predicate nouns, like *reserva* 'reservation' in *O Zé foi de uma grande reserva para com a Ana em relação à sua decisão* 'Ze was a of big reservation with Ana about her decision', and in *O Tó fez a reserva do bilhete* 'Tó made the ticket reservation', are not expressions of the same predicate. Further studies need to be made to see whether it is worth merging those entries together.

### 4.2    From Lexicon-Grammar Tables to NooJ Grammars

As we have already shown in Mota et al. (2016) and Mota et al. (2017), the process of integrating the lexicon-grammar with Port4NooJ dictionary entries is mostly automatic, but the process of creating the grammars that use the knowledge formalized in the lexicon-grammars is hand-crafted, hence, time-consuming. Some grammars can still be reused when they represent similar phenomena common to different lexicon-grammars although they usually need to be updated with the information of the newly integrated lexicon-grammar tables. The grammar that allows the substitution of the support verb by another support verb is one of such cases, as well as the grammar built to represent the equivalence between symmetric predicates, where the arguments have been swapped or coordinated. See Fig. 3 as an example of the latter. As illustrated in the figure, the equivalent adjectival constructions is also recognized and generated. In that case, the grammar guarantees the agreement in gender and in number with the subject (where the arguments have been coordinated or not) and the adjective.
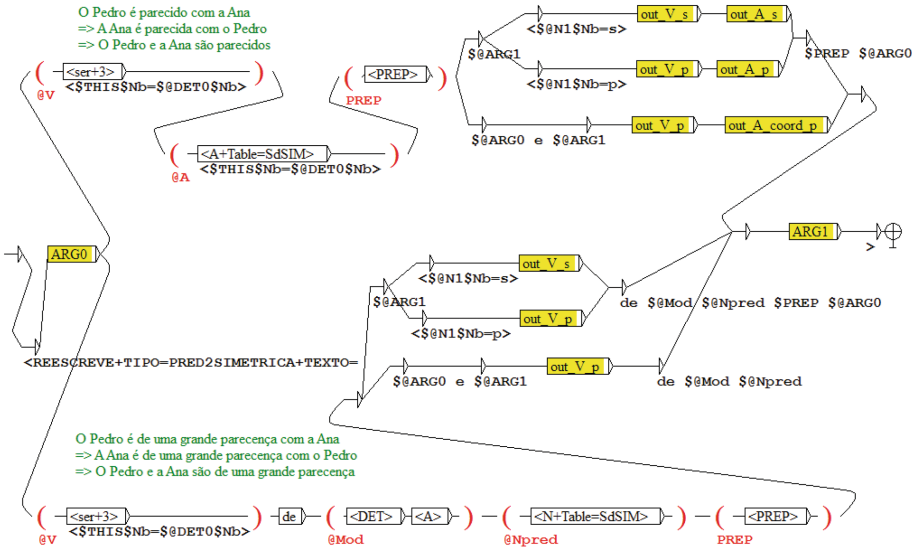
**Fig. 3.** Grammar for paraphrasing of symmetric constructions

Figure 4 shows the annotations that were added to the NooJ Text Annotation Structure after doing the linguistic analysis including the grammar that generates the paraphrasing of symmetric predicates illustrated in Fig. 3. Although the adjectival form *parecida* 'resembling' is feminine-singular, when swapping the argument of the adjective, it becomes masculine-singular to agree in gender-number with the noun *irmão* 'brother'; also, the adjective becomes masculine-plural to agree with the coordinated arguments, e.g. *a mulher e o irmão* 'the woman and the(=her) brother', since in Portuguese, whenever there is a masculine noun in coordinated noun phrases, the adjective is obligatorily inflected in the masculine form.

While developing new grammars, we encountered two new features that we had not encountered before: (1) we had to make use of more than one lexicon-grammar property to create the criteria to generate the paraphrase - in particular, to generate paraphrases of negative constructions, the predicate must have both attributes `PfxNeg` and `Negfaltade`, otherwise it is not enough to establish the paraphrastic relation between the two constructions; (ii) we established unidirectional paraphrases, i.e., we identify, for example, the construction A and generate the construction B; but we do not identify B and generate A - this happened when we need a larger context or a more complex analysis to be able to rephrase A given B, such as in the case of rephrasing a possessive with the appropriate noun phrase.

**Fig. 4.** Annotation and paraphrase generation of symmetric constructions

## 5  Conclusions and Future Work

This paper reported the progress of integration of existing lexicon-grammar tables in Port4NooJ, in continuity with previous integration efforts. This time, we have added 2,134 predicate nouns with Vsup *ser de* to our dictionary of predicate nouns, corresponding to 1,376 different noun lemmas. Half of the nouns already existed in the previous version of Port4NooJ (50%), which corresponds to a 6% increase in the number of nominal entries and a 20% increase in the number of predicate nouns for the current version of the Port4NooJ dictionary. Additional 797 entries await revision of the inflectional codes of the corresponding derived adjectives, or have format problems that require a fix prior to being added to this version. In addition, 450 new derivational paradigms were created, though some of them might overlap with paradigms already created, when integrating the predicate nouns used with the Vsup *fazer*. We have created some experimental grammars to generate production scale batches of paraphrases for different linguistic phenomena and exemplified paraphrasing with symmetric constructions in the NooJ syntactic parser. Our next steps will focus in four different axes: (i) consolidate and harmonize dictionaries, (ii) continue building lexicon-grammar-based paraphrasing grammars, (iii) review dictionaries and grammars, and (iv) integrate new lexicon-grammars, thus re-initiating the cycle. The new paraphrases can be immediately integrated in the eSPERTo system, because they are 100% precise and can serve the purposes of exploring language learning applications by using chatbots, a usage scenario in which our efforts are now engaged.

# References

Baptista, J.: Sintaxe dos Predicados Nominais construídos com o verbosuporte SER DE. Ph.D. thesis. Universidade do Algarve, Faro, Portugal (2000)

Baptista, J.: Construções simétricas: argumentos e complementos. In: Figueiredo, O., Rio-Torto, G., Silva, F. (eds.) Estudos de homenagem a Mário Vilela. Faculdade de Letras da Universidade do Porto, pp. 353–367 (2005a)

Baptista, J.: Sintaxe dos predicados nominais com 'ser de'. Fundação Calouste Gulbenkian, Fundação para a Ciência e a Tecnologia, Lisboa (2005b)

Barreiro, A., et al.: OpenLogos rule-based machine translation: philosophy, model, resources and customization. Mach. Transl. **25**(2), 107–126 (2011)

Carvalho, P.: Análise e Representação de Construções Adjectivais para Processamento Automático de Texto. Adjectivos Intransitivos Humanos. Ph.D. thesis. Universidade de Lisboa, Lisboa, Portugal (2007)

Chacoto, L.: O Verbo Fazer em Construções Nominais Predicativas. Ph.D. thesis. Universidade do Algarve, Faro, Portugal (2005)

D'Agostino, E., Elia, A.: Il significato delle frasi: un continuum dalle frasi semplici alle forme polirematiche. In: AA. VV, Ai limiti del linguaggio, Bari, Laterza, pp. 287–310 (1998)

Gross, M.: Méthodes en syntaxe: régime des constructions complétives. Actualités scientifiques et industrielles. Hermann (1975)

Gross, M.: Les bases empiriques de la notion de prédicat sémantique. Langages **15**(63), 7–52 (1981)

Gross, M.: Lexicon-Grammar. In: Brown, K., Miller, J. (eds.) Concise Encyclopedia of Syntactic Theories, pp. 244–259. Pergamon, Cambridge (1996)

Guillet, A., Leclère, C.: Restructuration deu group nominal. Langages **63**, 99–125 (1981)

Harris, Z.S.: The elementary transformations. In: Hiż, H. (ed.) Papers on Syntax. Synthese Language Library. Text and Studies in Linguistics and Philosophy, vol. 14, pp. 211–235. Springer, Dordrecht (1981). https://doi.org/10.1007/978-94-009-8467-7_9. ISBN 978-94-009-8467-7

Harris, Z.S.: Notes du Cours de Syntaxe. In: Gross, M. (ed.) Seuil, Paris (1976)

Laporte, E., Voyatzi, S.: An electronic dictionary of French multiword adverbs. In: Language Resources and Evaluation Conference. Workshop Towards a Shared Task for Multiword Expressions, pp. 31–34 (2008)

Mamede, N., et al.: STRING - a hybrid statistical and rule-based natural language processing chain for Portuguese. In: Computational Processing of the Portuguese Language (PROPOR 2012), vol. Demo Session, PROPOR, Coimbra, Portugal, PROPOR, s/p (2012)

Mota, C., Carvalho, P., Barreiro, A.: Port4NooJ v3.0: integrated linguistic resources for Portuguese NLP. In: Proceedings of the Tenth International Conference on Language Resources and Evaluation LREC 2016, Portorož, Slovenia, 23–28 May 2016

Mota, C., Chacoto, L., Barreiro, A.: Integrating the Lexicon-Grammar of predicate nouns with support verb *fazer* into Port4NooJ. In: Mbarki, S., Mourchid, M., Silberztein, M. (eds.) NooJ 2017. CCIS, vol. 811, pp. 29–39. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-73420-0_3

Rassi, A., et al.: The fuzzy boundaries of operator verb and support verb construc-
tions with dar "give" and ter "have" in Brazilian Portuguese. In: Proceedings of the
Workshop on Lexical and Grammatical Resources for Language Processing (LG-LP
2014), COLING 2014, Dublin, Ireland, 24 August 2014

Rassi, A.P., et al.: Integrating support verb constructions into a parser. In: Symposium
in Information and Human Language Technology (STIL 2015). SBC, Natal, Brasil
(2015)

Salkoff, M.: Automatic translation of support verb constructions. In: Proceedings of the
13th Conference on Computational Linguistics, COLING 1990, Helsinki, Finland,
vol. 3, pp. 243–246. ACL (1990)

Scott, B.: Translation, Brains and the Computer: A Neurolinguistic Solution to Ambi-
guity and Complexity in Machine Translation (2018)

Silberztein, M.: Les groupes nominaux productifs et les noms composés lexicalisés.
Lingvisticœ Investig. **17**(2), 405–425 (1993)

Silberztein, M.: Formalizing Natural Languages: The NooJ Approach, p. 346. Wiley,
Hoboken (2016)