

# Integrating the Lexicon-Grammar of Predicate Nouns with Support Verb *fazer* into Port4NooJ

Cristina Mota<sup>1</sup>, Lucília Chacoto<sup>2,3</sup>, and Anabela Barreiro<sup>1</sup>(✉)

<sup>1</sup> L2F/INESC-ID, Lisbon, Portugal

`cmota@ist.utl.pt`, `anabela.barreiro@inesc-id.pt`

<sup>2</sup> Universidade do Algarve, Faro, Portugal

`lchacoto@ualg.pt`

<sup>3</sup> CLUL, Lisbon, Portugal

**Abstract.** This paper describes the ongoing process of integrating approximately 3,000 predicate nouns into Port4NooJ, the Portuguese module for NooJ. The integration of these resources enables us to further extend the paraphrastic capabilities of eSPERTo paraphrasing system developed in the scope of a project with the same name. The integrated predicate nouns co-occur with the support verb *fazer* (*do* or *make*) and their syntactic and distributional properties are formalized in lexicon-grammar tables. These lexicon-grammar tables resulted in a standalone dictionary of predicate noun constructions and a few new grammars that can be used in paraphrase analysis and generation.

## 1 Introduction

Paraphrases are key linguistic assets in natural language interpretation and generation and play a very important role in NLP applications. Paraphrasing is a technique that implies the replacement or displacement of words, phrases or expressions in a sentence with semantically-equivalent linguistic structures so that the meaning of the output sentence remains the same. Support verb constructions represent an area of substantial interest in paraphrasing as they often have a corresponding verb associated with it. For example, the support verb construction *fazer uma sugestão* (*make a suggestion*), where the verb *fazer* has a weak semantic value and the predicate noun *sugestão* has a strong semantic value, is equivalent to the (strong) verb *sugerir* (*suggest*). Furthermore, support verb constructions are a rich source of paraphrases for reasons concerning stylistic variance allowed by the property of equivalence among different types of support verbs (elementary and non-elementary), among other reasons stated in early work (Barreiro 2009).

In this paper, we describe in detail the integration in the lexicon of Port4NooJ, the Portuguese module of NooJ (Silberztein 2016), of nearly 3,000 predicate nouns that co-occur with the support verb *fazer*, one of the most

frequent verbs in Portuguese<sup>1</sup>, thus complementing existing lexicon-grammar resources available for the Portuguese language. The resource integration work was developed within the eSPERTo<sup>2</sup> project whose main objective is to develop a smart, context-sensitive and linguistically enhanced paraphrasing system. At its current stage of development, eSPERTo comprises a paraphrase generator, a paraphrase acquisition module, and a web interactive application. eSPERTo recognizes semantico-syntactic, multiwords and other phrasal units, and transforms them into semantically equivalent phrases, expressions, or sentences. It uses local grammars to acquire linguistic knowledge that is applied in the identification/recognition and generation of different types of paraphrase, such as the support verb construction paraphrases described in Sect. 3. The new resources, which combine the dictionary of predicate nouns created by merging the information in lexicon-grammar tables with the entries in Port4NooJ plus the new transformational grammars contributed to over 5,000 entries before revision.<sup>3</sup> The utility of eSPERTo’s paraphrasing capabilities has been explored in a question-answering system to increase the linguistic knowledge of an intelligent conversational virtual agent and in a summarization tool to assist the paraphrasing task, but the resources described in this paper have not been tested in these applications. The broader application envisaged for the paraphrasing system is e-learning, namely in helping Portuguese language learners in editing and revising their texts. Precise paraphrases can also be helpful in professional translation, editing, and proofreading, among other tasks.

## 2 Related Work

Support verbs have been extensively and systematically studied within the lexicon-grammar theory proposed by Gross (1975), from the theoretical and methodological points of view, for many languages. Practical analytic formalization of multiword units (some including support verb constructions) in the Romance languages have been presented by several authors (D’Agostino and Elia 1998; Laporte and Voyatzi 2008; Silberstein 1993), among others. Support verb constructions have also been taken into account in contrastive studies for English and French (Salkoff 1990). With regards to Portuguese, most studies

---

<sup>1</sup> Support verbs constructions with the verb *fazer* are used both in written texts and spoken language, occurring extensively in Portuguese sentences and noun phrases. Sentences with support verb constructions are more frequent than sentences with their equivalent verbs. In fact, Barreiro (2009) points out that from a search on all sentences of the COMPARA parallel corpus (Frankenberg-Garcia and Santos 2003; Santos and Inácio 2006) where the infinitive form of *fazer* occurs with a noun or with a left modifier and a noun, 47% of the times the occurrence is in a support verb construction.

<sup>2</sup> eSPERTo, which means ‘smart’ in Portuguese, stands for “**S**ystem of **P**araphrasing for **E**ditng and **R**evison of **T**exts” (“**S**istema de **P**arafraseamento para **E**dição e **R**evisão de **T**exto”).

<sup>3</sup> Each entry in the lexicon-grammar can map to more than one existing Port4NooJ entry.

on support verb constructions outline the representation of argument mapping relations between a support verb and a nominalization (Baptista 2005; Chacoto 2005; Ranchhod 1990). Although support verbs often combine with autonomous predicate nouns, some studies focus on predicate adjective constructions (Carvalho 2007; Casteleiro 1981).

Mota et al. (2016) integrated the lexicon-grammar of human intransitive adjectives formalized by Carvalho (2007) into Port4NooJ, and showed how the properties contained in the lexicon-grammar tables can be used in paraphrasing tasks. The next step was to integrate complementary lexicon-grammars to expand the paraphrastic capabilities in Port4NooJ, building up on the previous work. Thus, the objective of the present work was to integrate the lexicon-grammar of predicate nouns which co-occur with the support verb *fazer* formalized by Chacoto (2005). The distributional and transformational properties of the predicate nouns addressed were used as the main criteria for establishing 19 classes and subclasses described in Sect. 3.

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

The systematic survey of predicate noun constructions with support verb *fazer* performed by Chacoto (2005) was carried out by assessing seven dictionaries and a subcorpus of five million words, i.e., Part 20 of the on-line 180 million word corpus CETEMPúblico.<sup>4</sup> The distributional and transformational properties of predicate nouns was used as the main criteria for establishing classes presented by these predicates, namely the syntactic structure, the number, and the type of constituents. Each class is represented in the form of a table, i.e., a binary matrix whose rows correspond to predicate nouns and whose columns represent the lexical and syntactic properties that these nouns have (+) or do not have (-). In total, 19 lexical and syntactic subclasses have been identified, as aforementioned.

The predicate nouns in the lexicon are, in general, everyday vocabulary (simple and compound nouns), with the exception of a group of predicate nouns of the sports and medical domains. The properties represented in the lexicon-grammar tables for each predicate noun lead to the generation of a variety of paraphrases, involving nominal group formation (active, passive and relative nominal groups), restructuring of the dative, shifting symmetrical nouns and symmetrical complements, conversion of the arguments, conversion into aspec-tual and stylistic variants of the support verb *fazer*, and transformation of the support verb, and relation between *fazer* and other support verbs. These properties establish semantic relationships between predicates allowing new features of paraphrases that have been described in NooJ grammars. For the sake of brevity, we illustrate only the case of paraphrasing based on the shifting of symmetrical nouns (Table 2), which are used with the predicates *aliar-se a* (*ally self with/to*), *casar com* (*marry with*), or *fazer um acordo com* (*make an agreement with*). The paraphrasing results from shifting the nouns from the position of subject to the position of indirect object and vice-versa.

<sup>4</sup> <http://www.linguateca.pt/acesso/corpus.php?corpus=CETEMPUBLICO>.

Table 1. Classes and distributional properties of predicate nouns with  $V_{sup}$  *fazer* ( $N_0$  *fazer* Det (Npred + C) W)

Class	$N_0$		C		W						Example		
	Det	NPred	E	V-n, ≠ Adj-n	≠ V-n, ≠ Adj-n	E	Prep	Prep	Prep N	Prep Que-F		de N	Prep N
	Nhum					=:a	=:de	=:/#a, de			=:/#a, de		
FC					+								O Zé fez voto de <b>ceibato</b>
FCAN					+	+							O Tó fez uma <b>serenata</b> à Ana
FCDN					+		+						O advogado fez as <b>alegações finais</b> do processo
FCPN					+			+					O avião fez uma <b>aterragem forçada</b> em Madrid
FCSI					+				+				O Zé fez um <b>acordo</b> com a Ana
FCQ					+					+			O Tó fez a <b>fineza</b> de convidar a Ana
FN				+									A instituição fez uma <b>angariação de fundos</b>
FNAN				+			+						O Zé fez um <b>aceno</b> ao Tó
FND		+											O Zé faz <b>ciclismo</b>
FNDN-hl				+				+					A Ana fez uma <b>classificação</b> dos dados
FNDNa				+			+						O rei fez a <b>abdicção</b> do trono
FNDNh				+			+						A Ana faz o <b>acolhimento</b> dos convidados
FNPn				+					+				O Zé fez um <b>acrescento</b> na prova tipográfica
FNSI				+					+				O Zé fez uma <b>aliança</b> com o Tó
FNQ				+						+			A Maria faz <b>tenção</b> de ter filhos
FNDNAN				+				+			+		O Zé fez uma <b>adaptação</b> do romance ao cinema
FNDNPN				+								+	O Tó fez o <b>câmbio</b> das pesetas em liras
FNDNPNSI				+								+	A Ana fez a <b>mistura</b> da farinha com o açúcar

**Table 2.** Paraphrasing based on symmetrical nouns

Symmetry	Paraphrases
[ $N_0$ V $N_1$ ]	<i>O Pedro aliou-se (com + a) a Maria</i>
	<i>Pedro joined/allied himself with/to Maria</i>
[ $N_1$ V $N_0$ ]	<i>A Maria aliou-se (com + a) o Pedro</i>
	<i>Maria joined/allied himself with/to Pedro</i>
[ $N_0$ Vcop Adj $N_1$ ]	<i>O Pedro é casado com a Maria</i>
	<i>Pedro is married to Maria</i>
[ $N_1$ Vcop Adj $N_0$ ]	<i>A Maria é casada com o Pedro</i>
	<i>Maria is married to Pedro</i>
[ $N_0$ Vsup N $N_1$ ]	<i>O Pedro fez um acordo com a Maria</i>
	<i>Pedro made an agreement with Maria</i>
[ $N_1$ Vsup N $N_0$ ]	<i>A Maria fez um acordo com o Pedro</i>
	<i>Maria made an agreement with Pedro</i>

## 4 Integration of Lexicon-Grammar Tables in Port4NooJ

The integration of the lexicon-grammar tables formalizing the properties of nominal predicates into Port4Nooj is a two-step process: first, one needs to create a dictionary from the lexicon-grammar and merge it with the Port4NooJ, and, then, it is necessary to create grammars that make use of the syntactic and distributional information described in the lexicon-grammar to identify and relate equivalent predicate constructions. The second step in our case aimed at expanding eSPERTo’s paraphrasing capabilities. Before describing the integration process, we start by mentioning our major challenges to date.

Given our prior experience integrating the lexicon-grammar tables of the human intransitive adjectives, we expected the integration of the lexicon-grammar tables for the predicate nouns with the support verb *fazer* to be straightforward. However, that was not what happened and we were faced with challenges at different levels.

The first one worth mentioning is that 65% of the predicate nouns already existed in Port4NooJ as nominalizations. These nominalizations derived from the corresponding verb, but did not contain detailed syntactic and distributional information as the one formalized in the lexicon-grammar tables. Nonetheless, these entries include relevant information that cannot be discarded, such as semantico-syntactic (SAL) information and English transfers. Another challenge faced in the integration process was that, in contrast with the lexicon-grammar tables of the human intransitive adjectives, which had the morphological equivalent verb and adjective explicitly mentioned in the table, the lexicon-grammar tables in discussion simply indicate whether the predicate noun has morphological equivalent predicates (the corresponding entry has a plus (+) sign), but neither the equivalent verb nor the equivalent adjective are explicit in the table.

Instead, they are listed co-occurring with additional information in one cross-reference table as an appendix, i.e., not having all information concentrated at a single site makes integration a more complex task. Additionally, there are 800 multiword predicate nouns comprised of a noun and an adjective, such as *transcrição integral* (*full transcript*) whose morphological equivalent is a multiword construction [verb + adverb], such as *transcrever integralmente* (*fully transcribe*). These cases need further revision, as the equivalent constructions need to be treated in a slightly different way. Finally, the spelling of a few predicate nouns in the lexicon-grammar were updated to comply with the Portuguese Orthographic Agreement<sup>5</sup>, but Port4NooJ does not yet comply with this Agreement.

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

We were compelled to adjust and create new scripts to overcome those challenges, even if the general process to create the equivalent standalone dictionary remained identical. For each entry in a lexicon-grammar table we had to (i) convert the corresponding lexicon-grammar properties into NooJ dictionary attributes; (ii) either create a new dictionary entry with those attributes or add those properties to an existing Port4NooJ dictionary entry or entries; and (iii) add the new entries or the old entries merged with the lexicon-grammar properties to the standalone dictionary. Each table was converted into a dictionary separately, but then all dictionaries were combined in a single standalone dictionary.

**Representation of Lexicon-Grammar Table Properties.** Given a lexicon-grammar entry, the new dictionary entry lemma is the predicate noun formalized in the lexicon-grammar, its POS tag is N, and its inflection code (the FLX attribute) is looked up in Port4NooJ or assigned automatically in cases where the word does not yet exist in the dictionary. The entry also receives the following attributes: `+Npred`, which indicates that it is a predicate noun, `+Vsup=fazer` which indicates that the support verb of the noun is *fazer*, and `+Table= <name_of_the_LG_table>` whose value is the name of the lexicon-grammar table where its distributional and syntactic properties are described.

Similar to what was done previously in Mota et al. (2016), for each different column in a lexicon-grammar table, a property `+<name_of_prop>` was created. If the noun row is marked with the value `+`, then that property was added to the noun entry, after removing or replacing special characters (e.g. `Vsupter` was generated from the LG property `Vsup=:ter`, and `NONnhum` was generated from `NO=:N-hum`), with the exception of the properties `V` and `Adj`, which indicate whether a noun has a morphologically-related verb and adjective, respectively. For these attributes, if the row has a `+` sign, instead of simply adding `+V` and `+Adj` attributes, a script looks up the noun in a file that lists all the morphologically-related verbs and adjectives to a noun and tries to obtain a derivation between the noun and verb or between the noun and

<sup>5</sup> <http://www.portaldalinguaportuguesa.org/main.html>.

the adjective. If the pair(s) noun/verb and noun/adjective exist, then a derivation paradigm is created and the following attributes are added to the entry: +DRV=N2V<drv\_code1>:<flx\_code1> and +DRV=N2A<drv\_code2>:<flx\_code2>, respectively. The *drv* code is determined and formalized automatically by finding the radical between the noun and the verb or adjective. For example, the noun *espuma* (*foam*) is associated with the corresponding verb *espumar* (*turn into foam*) and the corresponding adjective *espumoso* (*foamy*) through derivation rules (N2V2=r/V and A2V14=<B1>oso/A), which adds an *-r* to the noun to create the verb, and replaces the nominal ending *-a* with the adjectival ending *-oso*, respectively. As it happens with the inflection code of the lemma, the inflection of the derived word (*flx\_code1* or *flx\_code2*) is determined by consulting Port4NooJ (in this case, FLX=FALAR for the verb and FLX=ALTO for the adjective).

After all the attributes of a lexicon-grammar entry have been converted to NooJ format, we need to integrate the entry with the remaining Port4NooJ dictionaries.

**Integration with eSPERTO Dictionary Entries.** As previously mentioned, we had to use two dictionaries in order to properly integrate the lexicon-grammar entries into Port4NooJ: current version, i.e., the current version prior to integrating the lexicon-grammar entries, and old version, i.e., the dictionary version prior to the current version before removing the entries marked with +Npred that derive from verbs. In this way, even if the nominal predicate does not exist in the current version, it can still inherit the Port4NooJ properties as long as it makes part of the older version. There are 608 nominal predicates that only exist in the older version. If the nominal predicate exists in both dictionaries (there are 840 nominal predicates in this situation), and the linguistic information is the same in both dictionaries, the entry in the new standalone dictionary inherits that information, otherwise it only inherits the information that exists in the older dictionary as it means that the information in the current dictionary does not correspond to a nominal predicate - otherwise it would have been removed as well.

Given those two versions of the Portuguese dictionary, a new entry on the standalone dictionary of predicate nouns which co-occur with support verb *fazer* is created in one of the following situations.

Scenario 1: if the predicate noun or a predicate noun compliant with the pre-Ortographic Agreement<sup>6</sup> does not exist in Port4NooJ dictionary (neither current nor old version) then add the new entry created following the process described in the previous section as is. For example, the following entries did not exist in Port4NooJ and were created directly from the lexicon-grammar table:

<sup>6</sup> In order to properly look up for the nominal entries in Port4NooJ, we applied a simple rule that removes from words in Port4NooJ the first consonant in *cç*, *ct* and *pt* before comparing it to the nominal predicate in the lexicon-grammar. However, if there are derivation rules associated to the entry we still need to make sure that the rules do not introduce pre-agreement spelling.

batota,N+FLX=CASA+Npred+Vsup=fazer+Table=FN+NONhum+DetE+DetUMModif+Npred  
 +DRV=N2V2:FALAR+DRV=N2A6:ALTO+GNNdeNO  
 rodagem,N+FLX=ANO+Npred+Vsup=fazer+Table=FNAN+NONhum+DetO+Npred+Prepa  
 +N1Nnhum+VsuestarPrep+Vaspiniar+Vasproseguir+Vaspreparar  
 +DRV=N2V27:FALAR

Scenario 2: if noun (or the compliant noun) entry (i) exists in the current version but not in the old version or (ii) it is the same as in the old version, then merge the lexicon-grammar properties with the current entry, as in:

curva,N+FLX=ANO+Set=56+Subset=280+EN=curve+Npred+Vsup=fazer+Table=FN+NONhum  
 +NONhum+DetUMModif+DetO+Npred+Vsupdar+Vaspiniar+Vasproseguir  
 +Vaspreparar+DRV=N2V2:FALAR

All information up to +Npred already existed whereas after +Npred was obtained from the lexicon-grammar table.

Scenario 3: if noun (or the compliant noun) (i) exists in old version only or (ii) it exists in both, but current and old entries differ, then merge the lexicon-grammar properties with the old entry. For example, the noun *cruzamento* exists both in the current and old versions, but the entries differ. So, the lexicon-grammar attributes are only added to the old version, as the following 4 entries illustrate - the first three entries existed in the current version and the last entry only existed in the old verb, consequently that is the information to which the lexicon-grammar attributes were added:

cruzamento,N+FLX=ANO+AB+mot+EN=crossover  
 cruzamento,N+FLX=ANO+CO+recp+EN=frog  
 cruzamento,N+FLX=ANO+PL+nagcom+EN=crossings  
 cruzamento,N+FLX=ANO+PresPart+Set=68+Subset=551+EN=intersecting  
 +Npred+Vsup=fazer+Table=FNPN+NONhum+DetE+DetUMModif+DetO+Npred  
 +Preparar+N1Nnhum+DRV=N2V16:FALAR+GNNdeNOPrepN1

Additionally, remove previous Npred related properties, and then remove corresponding nominalizations from current version.

**Table 3.** Distribution of nominal predicates with support verb *fazer* by table attribute after integration into Port4NooJ

Class	# Lemmas	Class	# Lemmas	Class	# Lemmas
FNDN-hl	2145	FN	365	FNDNPN	120
FNDNa	1097	FNDNAN	313	FNSI	117
FNAN	560	FNDNPN	289	FND	92
FNPN	544	FCPN	246	FCSI	67
FNDNh	485	FCAN	180	FNQ	29
FC	406	FCDN	166	FCQ	27
Total: 6206					

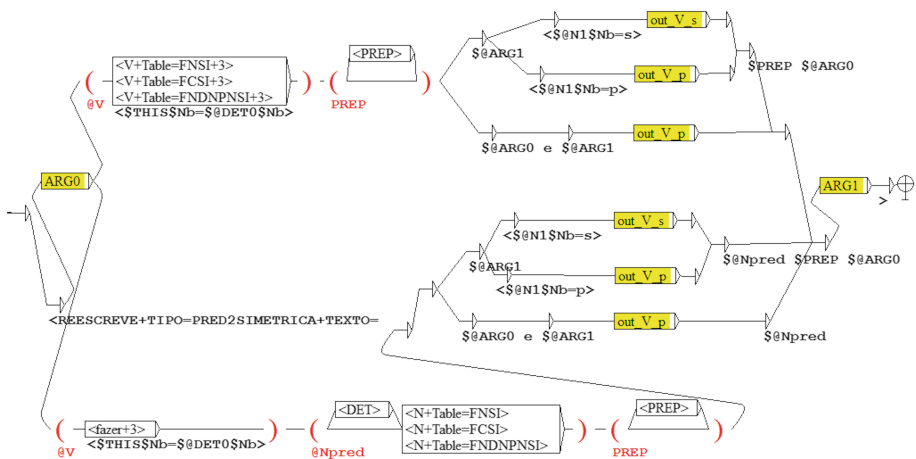


Finally, create inflectional (FLX) and derivational (DRV) codes and corresponding rules as needed, and check for missing inflectional (FLX) and derivational (DRV) codes.

We started by integrating 18 of the 19 tables<sup>7</sup>. The first attempt to convert the tables into a standalone Port4NooJ dictionary resulted in 6,205 nominal entries, corresponding to 1,918 different noun lemmas. Most entries already existed in Port4NooJ (55%), which corresponds to about a 7% increase in nominal entries in Port4NooJ from 11,719 different noun lemmas that existed in the previous version of Port4NooJ (i.e., before removing entries corresponding to nominal predicates that are nominalizations). In terms of predicate nouns only, there was an increase of 25%. In addition, 332 new derivational paradigms were automatically created. Table 3 shows the distribution of nominal predicates in the lemma dictionary by table attribute sorted by the most frequent table assigned.

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

Although we have a preliminary version of the standalone dictionary that still needs to be reviewed, we initiated the process of using the information in the lexicon-grammar tables of these constructions to do paraphrasing. In particular, we started by addressing the following issues: (i) Port4NooJ grammars already paraphrase support verb constructions (for example, paraphrasing the verbal construction with the nominal construction and vice-versa, or alternation between the support verb and other stylistic or aspectual verbs) - grammars are being updated with attributes from the new tables and are also being extended to take into account other paraphrases involving this type of constructions; (ii) the



**Fig. 1.** Grammar to identify and paraphrase constructions that allow symmetry

<sup>7</sup> For now, the table that formalizes medical technical terms was left out for a second stage.

grammars that paraphrase active and passive constructions were only paraphrasing sentences involving a verbal predicate - we are modifying these grammars so that they can also paraphrase nominal predicates in sentences and noun groups.

In addition, we also started developing new grammars to paraphrase equivalent constructions based on specific properties formalized in those lexicon-grammar tables. One such grammar paraphrases symmetric predicates, such as those illustrated in Table 2 and pictured in Fig. 1. The grammar in Fig. 1 applied to the sentences *O homem apostou com a mulher* (*The man betted with the woman*) and *O homem fez uma aposta com a mulher* (*The man made a bet with the woman*) results in the paraphrase *O homem e a mulher fizeram uma aposta* (*The man and the woman made a bet*).

## 5 Conclusions and Future Work

This paper described the ongoing process and preliminary results of integrating the lexicon-grammar that formalizes the properties of predicate noun constructions with the support verb *fazer*. This process is only complete after consolidation of linguistic information in Port4NooJ dictionaries and grammars. Although we already had scripts to generate a standalone dictionary from the lexicon-grammar of adjectives, these scripts had to be modified to generate a new standalone dictionary from the lexicon-grammar of predicate noun constructions as the challenges faced during this integration were different from the ones encountered during the integration of the adjectives. We will continue this integration work by creating all necessary grammars to process the constructions formalized in the lexicon grammar tables in question. After completion of this integration task, we intend to revise and evaluate the new resources in distinct applications.

In the near future, we plan to continue integrating and adapting additional lexicon grammars, such as the constructions with the support verb *ser de* (*be of*), such as in *ser de uma ajuda inestimável* (*be of invaluable help*) formalized by Baptista 2005, as these are a rich resource for paraphrasing in Portuguese.

**Acknowledgements.** This research was supported by Fundação para a Ciência e Tecnologia (FCT), under exploratory project eSPERTo (Ref. EXPL/MHC-LIN/2260/2013). Anabela Barreiro was also funded by FCT through post-doctoral grant SFRH/BPD/91446 /2012. The authors would like to thank Max Silberstein for his prompt support and guidance with all matters related to NooJ.

## References

- Baptista, J.: Sintaxe dos predicados nominais com ‘ser de’. Fundação Calouste Gulbenkian, Fundação para a Ciência e a Tecnologia, Lisboa (2005)
- Barreiro, A.: Make it simple with paraphrases: automated paraphrasing for authoring aids and machine translation, Porto, Portugal (2009)

- 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)
- Casteleiro, J.M.: *Sintaxe transformacional do adjetivo*. INIC, Lisboa (1981)
- 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*, pp. 287–310. Laterza, Bari (1998)
- Frankenberg-Garcia, A., Santos, D.: *Introducing COMPARA: the Portuguese - English parallel corpus*. In: Zanettin, F., Bernardini, S., Stewart, D. (eds.) *Corpora in Translator Education*, pp. 71–87. St. Jerome, Manchester (2003)
- Gross, M.: *Méthodes en syntaxe: régime des constructions complétives. Actualités scientifiques et industrielles*. Hermann, Paris (1975)
- 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)
- Mota, C., Carvalho, P., Raposo, F., Barreiro, A.: *Generating paraphrases of human intransitive adjective constructions with Port4NooJ*. In: Okrut, T., Hetsevich, Y., Silberztein, M., Stanislavenka, H. (eds.) *NooJ 2015. CCIS, vol. 607*, pp. 107–122. Springer, Cham (2016). [https://doi.org/10.1007/978-3-319-42471-2\\_10](https://doi.org/10.1007/978-3-319-42471-2_10)
- Ranchhod, E.: *Sintaxe dos predicados nominais com estar*. Linguística, INIC, Lisboa (1990)
- Salkoff, M.: *Automatic translation of support verb constructions*. In: *Proceedings of 13th Conference on Computational Linguistics, COLING 1990, vol. 3*, pp. 243–246. ACL, Helsinki (1990)
- Santos, D., Inácio, S.: *Annotating COMPARA, a grammar-aware parallel corpus*. In: Calzolari, N., et al. (eds.) *Proceedings of 5th LREC, Genoa, Italy*, pp. 1216–1221 (2006)
- Silberztein, M.: *Les groupes nominaux productifs et les noms composés lexicalisés*. *Linguisticæ Investigationes* **17**(2), 405–425 (1993)
- Silberztein, M.: *Formalizing Natural Languages: The NooJ Approach*, p. 346. Wiley, Hoboken (2016)