



HAL
open science

Description and classification of nominal predicates with support verb 'fazer' ('make'/'do') in Brazilian Portuguese

Cláudia Dias de Barros, Oto Araújo Vale, Jorge Baptista

► To cite this version:

Cláudia Dias de Barros, Oto Araújo Vale, Jorge Baptista. Description and classification of nominal predicates with support verb 'fazer' ('make'/'do') in Brazilian Portuguese. 32nd International Conference on Lexis and Grammar, Sep 2013, Faro, Portugal. pp.165-170. hal-04636129

HAL Id: hal-04636129

<https://hal.science/hal-04636129>

Submitted on 5 Jul 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Description and Classification of Nominal Predicates with the Support Verb *Fazer (Make/Do)* in Brazilian Portuguese*

Cláudia Dias de Barros¹, Oto Araújo Vale¹, Jorge Baptista²

¹Universidade Federal de São Carlos - UFSCar, ²Universidade do Algarve - UAlg

Abstract: This paper presents the study of description and classification of nominal predicates with the support verb *fazer (make/do)* in Brazilian Portuguese, under the perspective of Lexicon-Grammar Theory. We have established 30 formal properties (structural, distributional and transformational) in order to describe the nominal predicates, which were classified into 13 classes, according to arguments number and distribution and the possibility of symmetry.

Keyword: Support Verb, Lexicon-Grammar, Syntax, Predicative Noun, Semantic Role Labeling

1. Introduction

Since the 70's many authors have studied the nominal predicates under the perspective of Lexicon-Grammar (GROSS 1975), in many languages, like French (GROSS 1975, 1981), (GROSS 1989) (GIRY-SCHNEIDER 1978, 1987) and European Portuguese (EP) (RANCHHOD 1990), (BAPTISTA 2005), (CHACOTO 2005).

Nominal predicates are composed by a support verb (*SV*), like *dar (give)*, *estar (to be)*, *fazer (make/do)*, *ter (have)*, *ser de (to be of)*, etc., and a predicative noun (*Npred*), which is the predicator in a sentence and has arguments. The *SV* only provides to the *Npred*, through inflection, the aspect, mood and tense information. The main semantic content is provided by the noun, as shown in:

- (1) *Zé deu um beijo em Ana*
(Zé gave a kiss to Ann)¹
- (2) *Zé está em pânico*
(Zé is in a panic)
- (3) ***Zé fez um elogio a Ana***
(Zé made a compliment to Ann)
- (4) *Zé tem um comportamento estranho*
(Zé has a strange behavior)
- (5) *Zé é de uma grande coragem*
(Zé is very brave)

The correct identification of the predicator and the arguments of a sentence is important to the task of Semantic Role Labeling (SRL) (GILDEA and JURAFSKY 2002), because only with this identification the correct labels can be attributed (BARROS and VALE 2012), as is shown in (6):

- (6) *Zé_{Agente} faz_{SV} coleção_{Npred} de selos_{Tema}*
(Zé_{Agent} makes_{SV} a stamp_{Theme} collection_{Npred})

Besides EP and French, there are some researches studying nominal predicates in Brazilian Portuguese (BP) (RASSI *et al.* 2012) under the perspective of the Lexicon-Grammar (GROSS 1975) as well. These researches study nominal predicates with the *SV dar (give)* and *ter (have)*.

In this context, this paper describes the research about the nominal predicates with the *SV fazer (make/do)* in BP. This study is composed by approximately 2,400 predicative nouns, which were extracted from two sources: the study of Chacoto (2005) about predicative nouns with the *SV fazer (make/do)* of European Portuguese (EP), and the corpus PLN.Br (BRUCKSCHEN *et al.*, 2008), through the corpus processor Unitex (PAUMIER 2002).

This paper structure is: section 2 presents the main features of nominal predicates with the *SV fazer (make/do)*, including their description and classification; section 3 shows the data implementation proposal and section 4 presents the conclusion.

* Financially supported by Capes/Brazil.

¹ The English translations are just approximate from Brazilian Portuguese.

- (9) Zé fez a comparação de Pedro com João
(*Zé made a comparison between Peter and John*)
- (10) Zé fez a transferência do dinheiro do banco A para o banco B
(*Zé made the money transference from bank A to bank B*)
- (11) Zé fez uma homenagem a Ana
(*Zé made a tribute to Ann*)
- (12) Zé fez uma injustiça com Maria
(*Zé made an injustice to Mary*)
- (13) Zé fez a gravação de um disco
(*Zé recorded an album*)
- (14) Ana fez um acabamento no vestido
(*Ann did the finish in the dress*)
- (15) Zé fez a abertura do evento
(*Zé opened the event*)
- (16) Zé fez uma abordagem superficial do tema
(*Zé made a theme superficial approach*)
- (17) Zé fez adição de açúcar à receita
(*Zé added sugar to the recipe*)
- (18) Zé fez sua viagem
(*Zé made his trip*)
- (19) Zé fez uma advertência a Ana
(*Zé made a warning to Ann*)
- **Distributional:** the types of the arguments – human nouns (Nhum) (20), non human nouns (Nnhum) (21), plural nouns (Npl) (22), part-of-the-body nouns (Npc) (23) or a clause (QueF) (24);
- (20) Zé faz natação
(*Zé swims*)
- (21) A planta faz fotossíntese
(*The plant makes photosynthesis*)
- (22) As pessoas fizeram uma fila
(*People made a line*)
- (23) Zé fez uma contusão no pé
(*Zé made a foot injury*)
- (24) Zé fez a gentileza de convidar Ana
(*Zé was kind to invite Ann*)
- **Transformational:** syntactic changes on standard sentence, but no semantic alterations, as passive (25a) and (25b), symmetry (26a) and (26b), conversion (27a) and (27b), formation of nominal group from a relative clause reduction (28a) and (28b) and nominalization (29a) and (29b).
- (25a) Zé fez a revisão do carro
(*Zé made the car review*)
- (25b) A revisão do carro foi feita por Zé
(*The car revision was made by Zé*)
- (26a) Zé fez um acordo com Ana
(*Zé made a deal with Ann*)
- (26b) Ana fez um acordo com Zé
(*Ann made a deal with Zé*)
- (27a) Zé fez um elogio a Ana
(*Zé made a compliment to Ann*)
- (27b) Ana recebeu um elogio de Zé
(*Ann received a compliment from Zé*)
- (28a) Zé fez uma aposta com Ana
(*Zé made a bet with Ann*)
- (28b) A aposta de Zé com Ana...
(*Zé's bet with Ann...*)
- (29a) Zé apresentou seu trabalho publicamente
(*Zé presented his work publicly*)
- (29b) Zé fez a apresentação de seu trabalho publicamente
(*Zé made his work presentation publicly*)

We have marked all the formal properties in the binary table with a '+' if the structure presented the property and with a '-' in otherwise.

4. Nominal Predicates classification

After the nominal predicates description, as shown in subsection 2.1, we did the classification, taking into account some syntactic features, as the number of arguments and their distribution and the possibility of symmetry as well.

We made the first classification based on the number of arguments that the nominal predicates have:

- **PB-F1** – predicative nouns with just 1 argument, the subject (N0), as *natação* (*swimming*), in (30):
(30) Zé faz natação
(*Zé swims*)

This class presents 786 predicative nouns and was divided into three subclasses, according to the subject distribution:

- **PB-F1R** – the subject (N0) can be a human noun (Nhum) or non human noun (Nnhum), as *barulho* (*noise*), in (31):

- (31) (Zé + a máquina) faz barulho
(*Zé + the machine makes noise*)

- **PB-F1H** – the subject (N0) is a human noun (Nhum), as *basquete* (*basket*), as in (32):

- (32) Zé faz basquete
(*Zé does basket*)

- **PB-F1NH** – N0 is a non human noun (Nnhum), as *fotossíntese* (*photosynthesis*), in (33):

- (33) A planta faz fotossíntese
(*The plant does photosynthesis*)

- **PB-F2** – predicative nouns with two arguments: the subject (N0) and a complement (N1). This class presents 1,499 predicative nouns, so it has been divided into seven subclasses, according to subject and complements distribution and the possibility of symmetry.

- **PB-F2S** – symmetric nouns, as *fronteira* (*border*), as in (34):

- (34) Portugal faz fronteira com a Espanha
(*Portugal makes border with Spain*)

- **PB-F2Q** – N0 is a clause (QueF), as *falta* (*lack*), as in (35):

- (35) Que Zé não tenha vindo fez falta ao time
(*The Zé's absence lack the team*)

- **PB-F2Q1** – N1 is a clause (QueF), as *gentileza* (*kindness*), as in (36):

- (36) Zé fez a gentileza de receber Ana em sua casa
(*Zé was kind to receive Ann in his home*)

- **PB-F2HH** – N0 and N1 are human nouns (Nhum), as *elogio* (*compliment*), as in (37):

- (37) Zé fez um elogio a Ana
(*Zé made a compliment to Ann*)

- **PB-F2HNNH** – N0 is a human noun (Nhum) and the complement is a non human noun (Nnhum), as *diagnóstico* (*diagnosis*), as in (38):

- (38) O médico fez o diagnóstico da doença
(*The doctor made the disease diagnosis*)

- **PB-F2NHHH** – N0 and N1 are non human noun (Nnhum), as *polinização* (*pollination*), as in (39):

- (39) As abelhas fizeram a polinização das flores
(*Bees made the flower pollination*)

- **PB-F2HR** – N0 is a human noun (Nhum) and N1 can be a human noun or non human noun (Nnhum), as *protesto* (*protest*), as in (40):

- (40) Os manifestantes fizeram um protesto contra (o presidente + a guerra)
(*The protesters made a protest against the (president +war)*)

- **PB-F3** – predicative nouns with 3 arguments (N0, N1 and N2), as *comparação* (*comparison*), in (41):

- (41) Zé fez a comparação de Ana com Maria
(*Zé made the comparison of Ann with Mary*)

This class hasn't been divided yet, but it can be divided in a subclass, according to the possibility of symmetry in the complements, as *comparação* (*comparison*) in (42a) and (42b):

- (42a) Zé fez a comparação de Ana com Maria
(*Zé made the comparison of Ann with Mary*)

- (42b) Zé fez a comparação de Maria com Ana
(*Zé made the comparison of Mary with Ann*)

- **PB-F4** – predicative nouns with 4 arguments, as *transferência* (*transfer*), as in (43):
- (43) Zé fez a transferência do dinheiro do banco A para o banco B
(Zé transferred the money from bank A to bank B)
- This class won't be divided because it has only 6 nouns.

5. Data application proposal

The description and classification of nominal predicates with the SV *fazer* (*make/do*) in Brazilian Portuguese can be used as a lexical resource to improve some Natural Language Processing (NLP) applications, as automatic identification of nominal predicates and automatic semantic role labeling. The predicative nouns analyzed by the research described in this paper will be used also to the construction of a database called NomBank.Br, which will contain the predicate nouns and their argumental net.

NomBank.Br will be based on the cognate project to American English (AE) called NomBank (MEYERS 2004), developed in the University of New York, which can be defined as the task of labeling the argumental structure of the nouns present in *Penn TreeBank II* corpus (MARCUS *et al.* 1993). Figure 2 presents some examples of annotated sentences present in NomBank:

Figure 2: Example of NomBank annotation

In this example is possible to observe that NomBank has special labels to label the predicative noun (REL), and it's arguments: N0=ARG0, N1=ARG1, etc.

In NomBank.Br it will be used the same labels, and a sentence like (44) can be annotated as follows:

- (44) Zé fez um elogio a Ana
(Zé made a compliment to Ann)

ARG0 = Zé, SUPPORT= *fazer* (*make*), REL=elogio (*compliment*), ARG1= Ana (*Ann*)

The correct identification of the predicator in a sentence (in the case of (44), *elogio* (*compliment*), not the verb *fazer* (*make*)) is crucial to the task of (Automatic) Semantic Role Labeling (SRL), otherwise some mistakes can occur. For example, if the verb were considered as the predicator in (44) instead of *elogio* (*compliment*), this noun would be considered as another argument, what would be wrong.

NomBank.Br can be also used to the automatic identification of nominal predicates, because this database will contain nominal predicates with the SV *fazer* (*make/do*), and a system can use this nominal predicates as the base to automatically identify sentences with nominal predicates. The binary table built in the research described in this paper can also be used by these systems, as presented in (TOLONE AND CONSTANT 2010) and (TOLONE *et al.* 2013).

6. Conclusion

This paper presented the study of description and classification of approximately 2,400 nominal predicates with the support verb *fazer* (*make/do*) of Brazilian Portuguese, under the perspective of Lexicon-Grammar (GROSS 1975).

Nominal predicates were analyzed according 30 formal properties (structural, distributional and transformational) as the number of arguments, the type of prepositions that introduce the complements, the type of determiners between the SV and the predicative noun, the argument distribution, the possibility of passive, symmetry, conversion, etc.

After the linguistic description, the data were classified into 12 classes, according to the arguments number and distribution and the possibility of symmetry.

The data will be available to the construction of NomBank.Br, a database with predicative nouns and their argumental net. It can be used as a lexical resource to improve NLP systems that performs automatic semantic role labeling (GILDEA AND JURAFSKY 2002) and automatic nominal predicates identification.

References

- BAPTISTA, Jorge (2005). *Sintaxe dos Predicados Nominais com SER DE*. Lisboa: FCT/FCG.
- BARROS, Cláudia Dias de and VALE, Oto Araújo (2012). Brazilian Portuguese nominal predicates with ‘fazer’ (make/do): sports. In Radimsky, J., (ed.), *Actes du 31e Colloque International sur le Lexique et la Grammaire*, pp. 10–17, République Tchèque. Université de Bohême du Sud.
- BRUCKSCHEN, M., MUNIZ, F., SOUZA, J. G. C., FUCHS, J. T., INFANTE, K., MUNIZ, M., GONÇALVES, P. N., VIEIRA, R., e ALUÍSIO, S. M. (2008). Anotação linguística em xml do corpus PLN-Br. Nilc-TR-09-08, *Série de relatórios do NILC*.
- CHACOTO, Lucília (2005). *O verbo ‘fazer’ em construções nominais predicativas*. Dissertação de Doutoramento em Linguística (especialidade: Sintaxe). Universidade do Algarve, Faro, Portugal.
- GILDEA, Daniel AND JURAFSKY, Daniel (2002). Automatic Labeling of Semantic Role. *Computational Linguistics*, 28(3): pp. 245-288.
- GIRY-SCHNEIDER, Jacqueline (1978). *Les nominalisations en français : l’opérateur faire dans le lexique*. Librairie Droz, Genova.
- GIRY-SCHNEIDER, Jacqueline (1987). *Les prédicats nominaux en français: les phrases simples à verbes support*. Librairie Droz, Genova.
- GROSS, Gaston (1989). *Les construction converses du français*. Genève: Droz.
- GROSS, Maurice (1975). *Méthodes en Syntaxe*. Paris: Hermann.
- GROSS, Maurice. 1981. Les bases empiriques de la notion de prédicat sémantique. *Langages*, 63, 7–52.
- LECLÈRE, Christian (2002). Organization of the Lexicon-Grammar of French Verbs. *Linguisticae Investigationes*, 25-1, p. 29–48.
- MARCUS, M. P.; SANTORINI, B.; MARCINKIEWICZ, M. A (1993). Building a Large Annotated Corpus of English: The Penn TreeBank. *Computational Linguistics*, 19(2): 313-330.
- MEYERS, A.; REEVES, R.; MACLEOD, C.; SZEKELY, R.; ZIELINSKA, V.; YOUNG, B.; GRISHMAN, R (2004). Annotating Noun Argument Structure for NomBank. In: *Proceedings of LREC-2004*, Lisbon, Portugal.
- PAUMIER, SÉBASTIEN (2002). *Unitex: manuel d’utilisation*. Research report. University of Marne-la-Vallée, França.
- RANCHHOD, Elisabete (1990). *Sintaxe dos predicados nominais com ESTAR*. Lisboa: INIC.
- RASSI, Amanda Pontes (2008). Estatuto sintático-semântico do verbo ‘fazer’ no Português escrito do Brasil. Dissertação (Mestrado em Linguística e Língua Portuguesa). Faculdade de Letras, Universidade Federal de Goiás, Goiânia.
- RENOUF, A.; KEHOE, A.; BANERJEE, J. (2007). WebCorp: an integrated system for web text search. In: C. Nesselhauf; M. Hundt; C. Biewer (eds.), *Corpus Linguistics and the Web*. Amsterdam: Rodopi.
- ROCHA, Paulo and SANTOS, Diana (2000) (November). CETEMPúblico: Um corpus de grandes dimensões de linguagem jornalística portuguesa. *Actas do V Encontro para o processamento computacional da língua portuguesa escrita e falada, PROPOR’2000*, p. 131–140.
- TOLONE, E. and CONSTANT, M. (2010). A generic tool to generate a lexicon for NLP from Lexicon-Grammar tables. *Lingue D’Europa e del Mediterraneo, Grammatica Comparata*, 1:79-193.
- TOLONE, E; VOYATZI, S.; MARTINEAU, C. (2013). Utilisation des entrées adverbiales du DELA issues des tables du Lexique-Grammaire du Français. *Dialogar é preciso: Linguística para o processamento de línguas*, p. 243-258.