

VOLUME 8

ANO 2013

# LINGUÍSTICA

REVISTA DE ESTUDOS LINGUÍSTICOS  
DA UNIVERSIDADE DO PORTO

CENTRO DE LINGUÍSTICA  
DA UNIVERSIDADE DO  
PORTO

FACULDADE DE LETRAS  
DA UNIVERSIDADE DO  
PORTO

# LINGUÍSTICA

REVISTA DE ESTUDOS LINGUÍSTICOS  
DA UNIVERSIDADE DO PORTO

**Diretora:**

Ana Maria Brito (Universidade do Porto)

**Secretariado Editorial:**

João Veloso (Universidade do Porto)

**Conselho Científico:**

Alexandra Guedes Pinto (Universidade do Porto)  
Alexandra Soares Rodrigues (Instituto Politécnico de Bragança)  
Ana Cristina Macário Lopes (Universidade de Coimbra)  
Ana Maria Martins (Universidade de Lisboa)  
António Leal (Universidade do Porto)  
Belinda Maia (Universidade do Porto)  
Carmen Matzenauer (Universidade Católica de Pelotas, Rio Grande do Sul)  
Clarinda Azevedo Maia (Universidade de Coimbra)  
Claudia Brescancini (Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre)  
Cristina Martins (Universidade de Coimbra)  
Fátima Henriques da Silva (Universidade do Porto)  
Fátima Oliveira (Universidade do Porto)  
Fernanda Irene Fonseca (Universidade do Porto)  
Francisco Lacerda (Universidade de Estocolmo)  
Georges Kleiber (Universidade de Strasbourg)  
Gerd Wotjak (Universidade de Leipzig)  
Gianpaolo Salvi (Universidade Eötvös Loránd, Budapeste)  
Graça Rio-Torto (Universidade de Coimbra)  
Ignacio Bosque (Universidade Complutense de Madrid)  
Ildikó Szijj (Universidade Eötvös Loránd, Budapeste)  
Inês Duarte (Universidade de Lisboa)  
Isabel Galhano Rodrigues (Universidade do Porto)  
Isabel Margarida Duarte (Universidade do Porto)  
Isabel Pereira (Universidade de Coimbra)  
Jean-Pierre Angoujard (Université de Nantes)  
João Costa (Universidade Nova de Lisboa)  
Joaquim Brandão de Carvalho (Universidade de Paris 8)  
Krista Varantola (Universidade de Tampere)  
Lluís Payrató (Universidade de Barcelona)  
Luís Filipe Cunha (Universidade do Porto)  
Maria Clara Barros (Universidade do Porto)  
Maria da Graça Lisboa Castro Pinto (Universidade do Porto)  
Maria Helena Mateus (Universidade de Lisboa)  
Maria Helena Paiva (Universidade do Porto)  
Maria João Freitas (Universidade de Lisboa)  
Marina Vigário (Universidade de Lisboa)  
Norval Smith (Universidade de Amsterdão)  
Olívia Figueiredo (Universidade do Porto)  
Petra Sleeman (Universidade de Amsterdam)  
Rogélio Ponce de Léon Romeo (Universidade do Porto)  
Sónia Frota (Universidade de Lisboa)  
Thomas Hüsgen (Universidade do Porto)  
Tjerk Hagemeier (Universidade de Lisboa)  
Violeta Demonte (Universidade Autónoma de Madrid)

# LINGUÍSTICA

REVISTA DE ESTUDOS LINGUÍSTICOS  
DA UNIVERSIDADE DO PORTO

VOL. 8, 2013

## FICHA TÉCNICA

*Linguística*  
*Revista de Estudos Linguísticos da Universidade do Porto*  
Vol. 8, 2013

ISSN: 1646-6195

Periodicidade: Anual

Diretora:  
Ana Maria Brito

Secretariado Editorial:  
João Veloso

Editores:  
Faculdade de Letras da Universidade do Porto  
e Centro de Linguística da Universidade do Porto

Capa:  
José Osswald

Impressão e acabamentos:  
Invulgar - Artes Gráficas

Depósito Legal:  
248653/06

Tiragem:  
200 Exemplares

Os artigos publicados estão sujeitos a “peer review”.

Esta edição é integralmente financiada por:  
Centro de Linguística da Universidade do Porto

O Centro de Linguística da Universidade do Porto é uma Unidade de I&D financiada pela Fundação para a Ciência e a Tecnologia (PEst/LIN/UI0022/2011).

A Revista está registada no DOAJ e Latindex e está indexada na base de dados Fonte Académica.

<http://ler.letras.up.pt/site/default.aspx?qry=id04id191&sum=sim>

# SUMÁRIO

Espaço da direção.....	5
Foreword <i>Ana Maria Brito</i> .....	7
Artigos	
Nouns in apposition: Portuguese data <i>Graça Rio-Torto</i> .....	17
Noms épithètes dans les expressions binominales <i>Fernando Martinho</i> .....	39
Italian verb to noun conversion: the case of nouns in –a deriving from verbs of the 2 <sup>nd</sup> and 3 <sup>rd</sup> conjugation <i>Daniela Marzo</i> .....	69
Is conversion a syntactic or a lexical process of word formation? <i>Alexandra Soares Rodrigues</i> .....	89
Bare Morphology <i>Alina Villalva</i> .....	121
Deriving color adjectival nominalizations <i>Artemis Alexiadou</i> .....	143
Deadjectival human nouns: conversion, nominal ellipsis, or mixed category? <i>Petra Sleeman</i> .....	159
Les verbes à objets cognats en portugais européen: quelques réflexions sur la conversion et la génération de mots <i>Celda Choupina</i> .....	181

## Varia

Evidentials <i>dizque</i> and <i>que</i> in Spanish. Grammaticalization, parameters and the (fine) structure of Comp <i>Violeta Demonte, Olga Soriano</i> .....	211
--	-----

Óscar Lopes: recordar o linguista, o mestre e o amigo <i>Fátima Oliveira</i> .....	235
---	-----

## Recensões

R. do Carmo (org.) <i>Linguagem, Argumentação e Decisão Judiciária</i> <i>Belinda Maia</i> .....	243
---	-----

Fernão de Oliveira. <i>Gramática da Linguagem Portuguesa, Fac-simile, introdução e edição actualizada e anotada por José Eduardo Franco e João Paulo Silvestre</i> <i>Sónia Duarte</i> .....	249
---	-----

M. V. Navas Sánchez-Élez: <i>El barranquenho. Un modelo de lenguas en contacto</i> <i>Clara Araújo Barros</i> .....	255
--	-----

P. N. da Silva: <i>Tipologias Textuais - Como Classificar Textos e Sequências</i> <i>Isabel Margarida Duarte</i> .....	263
---	-----

Instructions to Authors.....	267
------------------------------	-----

## Espaço da Direção

Este é um número temático de *Linguística. Revista de Estudos Linguísticos da Universidade do Porto*. Com efeito, os textos publicados neste volume foram apresentados no workshop *Conversion and mixed categories* organizado pelo Centro de Linguística da Universidade do Porto (CLUP) e que teve lugar na Faculdade de Letras da Universidade do Porto, a 2 e 3 de Novembro de 2012. A direção da revista agradece a todos os autores, assim como aos membros da Comissão Científica que participaram na avaliação dos textos aqui apresentados.

O número contém ainda um texto de Violeta Demonte e Olga Soriano sobre enunciados em espanhol que contém um certo tipo de *que* ligado a discurso relatado. Seguem-se quatro resenhas de publicações recentes em *Linguística*.

O presente número de *Linguística* é publicado no mesmo ano da morte de três linguistas portugueses, Ernesto d'Andrade, da Faculdade de Letras da Universidade de Lisboa, Fernanda Menendez, da Universidade Nova de Lisboa, e Óscar Lopes, um cidadão e linguista a todos os níveis excepcional e que foi o fundador, em 1976, do Centro de Linguística da Universidade do Porto. A direção da revista lamenta profundamente a perda destes três docentes e investigadores e como forma de homenagear a contribuição que todos deram à *Linguística* portuguesa, dedica este volume à sua memória. Mas, porque a nossa dívida relativamente a Óscar Lopes é incomensurável, o volume contém ainda um texto de Fátima Oliveira sobre a sua obra em *Linguística*.





## Foreword

The texts published in this volume were presented at the international workshop *Conversion and Mixed categories* organized by the Centro de Linguística da Universidade do Porto (CLUP), which took place at the Faculdade de Letras da Universidade do Porto, on 2nd and 3rd november, 2012.

Grammatical categories have always been at the centre of linguistic reflexion. It is easy to classify the majority of words according to their syntactic category, based on phonological form and lexical meaning. As regards other words, however, phonological form and lexical meaning are insufficient for classification and it is crucial to take into account their syntactic behaviour. This is specially the case with words obtained by conversion. Conversion, sometimes referred to as “functional shift”, “functional change”, “transposition”, “zero derivation” and even “improper derivation” according to the approach adopted, is a process that has no similarities with canonical derivational processes because it does not involve affixes, being a process whereby a word’s category changes without any change of form.

English is a language rich in conversion phenomena, as in the noun / verb relation present in *bottle - to bottle* or in the adjective – verb relation present in *clear – to clear*. In this kind of examples, conversion is said to be total, because the converted verb gives rise to other nouns, such as *bottler*, and it can have morphological tense and person endings (e.g. the simple past *bottled*, the third person of the present *bottles*).

Deadjectival nouns referring to humans, such as *the Germans*, have also been analyzed as being the result of a (total) morphological conversion. By contrast, conversion is sometimes considered to be the result of a syntactic process of nominal ellipsis in an example such as *poor<sub>ADJ</sub>, the poor<sub>N</sub>* or as a case of partial conversion, because the noun is formed from an adjective by the presence of an article and it cannot have a plural ending, *\*the poors* being ungrammatical in English.

For a long time, grammarians described conversion as typical of English,

related to the gradual loss of inflection in this language. As Varela (2006: 174) clarifies, “the connection between conversion and inflection levelling largely depends on the value given to inflection (...) and to what units the concept may be applied, whether words or stems”. In the latter case, other languages are considered to have instances of conversion: in French *rose* (‘pink’) – *roser* (‘to pink’), in German *Regel* (‘rule’) – *Regeln* (‘to regulate’), in Spanish, Italian and Portuguese, in the so called “regressive” nouns, as in *brillo* (‘shine’) – *brillar* (‘to shine’) (cf. Varela, *loc. cit.*).

Conversion being a morphological process with syntactic repercussions, how and in which grammar component should this phenomenon be treated?

For many authors, there is a component responsible for word formation, Morphology, with its own rules and constraints, which are not dependent on syntactic principles, as argued by Jackendoff (2002) and, in this volume, by Graça Rio-Torto, Alexandra Rodrigues, Alina Villalva and Daniela Marzo. Ackema & Neeleman (2007), Williams (2007), developing Di Sciullo & Williams (1987), and many others, preserve the distinction between Morphology and Syntax, considering that Morphology is the system of words and Syntax is the system of phrases: they may share certain notions (such as merge, c-command, head, argument) but they also have specific operations and primitives that distinguish them.

A strong lexicalist view is also proposed by Hale & Keyser (1993). According to their framework, the Lexicon is the component whereby operations similar to movements may operate in the lexical structures. Starting with denominal verbs such as *bottle*, *saddle*, traditionally considered cases of conversion, Hale & Keyser (1993) proposed that this type of verbs is formed from nouns by a movement from an N position to a V position within the lexical structure. This proposal has had an important impact: one of its consequences being that intransitive verbs were analyzed by several linguists as «disguised» transitives; and such a way of looking to lexical structures influenced a lexicalist view of morphological processes.

At the other extreme, other authors consider that there is no distinction between Morphology and Syntax. Julien (2007) considers that the notion of word is even an epiphenomenon (see in particular pp. 212, 234), because it is based on syntactic conditions and because many words are derived from syntactic relations, such as head-head relation, specifier-head relation,

and others, among others. Svenonius (2007) defends that Morphology operates with the same types of structures and primitives as Syntax, and that there are movements responsible for certain linearizations. Moreover, for Embick & Noyer (2007), inspired by Halle & Marantz (1993), the founders of Distributed Morphology, the only generative component is Syntax, the words being introduced into terminal heads of syntactic structure. For Artemis Alexiadou and Petra Sleeman, in this volume, the category of a word may be considered the result of the movement of an acategorical root to functional categories that define its final category as a noun, or as a verb, for example.

The domain that mostly inspired Alexiadou (2001), Alexiadou *et al.* (2011) and Sleeman (2011a, 2011b) was mixed categories: participles (in their diversified uses), deverbal nominalizations, nominal and verbal gerundives in English and German (*John's performing of the sonata* versus *John's performing the sonata*), supines and infinitives in Roumanian (*citi(ul) contant al ziarelor*, 'the act of constantly reading newspapers', versus *constanta omitere(a) a unor informatii*, 'the constant omitting of information'), nominal and verbal infinitives in Spanish (*el murmurar la gente*, 'the fact that people (nominative) murmur', versus *el murmurar de las fuentes*, 'the murmuring of the fountains'). The choice of Distributed Morphology and the abandoning of both the Lexicalist Hypothesis and the traditional X' model seem justified because it is difficult to account for the intermediary steps that all these constructions may exhibit together with their hybrid properties.

From this brief presentation we see that morphological processes are at the centre of interesting discussions regarding the architecture of grammar. If it is true that conversion is classically considered a morphological process, as opposed to that which allows the creation of new words through affixes and different from syntactic constructions, it is also true that both conversion and the creation of words through affixes is an area in which the frontiers of grammar components must be carefully considered.

The texts here assembled analyse different phenomena and represent different theoretical approaches to the main topics of the workshop.

Graça Rio-Torto studies NN appositions such as *empresas-fantasma(s)* (lit. 'businesses-phantom(s)'), *palavras-chave(s)* (lit. 'words-key(s)'). Several questions have been raised about these constructions: are they compounds,

syntactic objects or lexicalizations of phrases? Are they phrasal compounds or phrasal nouns? Rio-Torto argues against a syntactic analysis, showing that they do not present any functional categories typical of nominal phrases, and argues that they are lexical phrases downgraded to lexical units.

A different approach to NN appositions (like *equipa maravilha* (équipe merveille, lit. 'team wonder') is taken by Fernando Martinho, who sees in this process, very common in Romance languages, the emergence of the second N as an adjectival modifier, very similar to relational adjectives, based on some syntactic behaviour.

Daniela Marzo studies another controversial issue, that represented by the relationship between the Italian nouns *cammino* ('walking, path'), *sosta* ('stop, rest, short break') with the verbs *camminare*, *sostare*. Are these nouns the result of conversion, zero-suffixation or overt suffixation? In her contribution these approaches are compared with respect to their explanatory power. While the suffixation account is ruled out for semantic and distributional reasons, it is shown that the conversion approach is the more plausible in being able to account for the two cases. In order to explain the formation and structure of a noun of the type of *basta* (verb *bastire*) it is proposed that it is the synchronic remnant of an Old Italian conversion type.

The differences between conversion, affix word formation and syntactic structures are highlighted by Alexandra Rodrigues, who analyses converted deverbal nouns such as *remendo* ('event of mending') and *curte* ('event of having fun'), and argues for their formation in the Lexicon, distinguishing them from purely syntactic types of nominalisation such as the one that occurs in *o estudar matemática traz-me vantagens* ('studying maths brings me benefits') or in *o remendar roupa é um recurso nesta época* ('mending clothes is a good resource nowadays').

Alina Villalva studies several cases of Portuguese words that share the same root string and argues that they represent different types of conversion. Such is the case with *firm*<sub>ADJ</sub> – *firmar*<sub>V</sub>; *abaix*<sub>ADV</sub> – *abaixar*<sub>V</sub>; *manch*<sub>N</sub> – *manchar*<sub>V</sub>; *afia*<sub>V</sub> – *afia*<sub>N</sub>. In order to set them apart one needs to take into account features from every grammar domain. Villalva concludes that this fact is one of the best arguments for considering conversion as a lexical word formation process (as opposed to morphological, syntactic or semantic notions).

Artemis Alexiadou examines two types of nominalizations related to colour adjectives in Greek, a suffixed one and a neutral one. The author shows that the two differ in that suffixed nominalizations denote stage level properties, while neuter nominalizations denote individual level properties. This difference is due to the fact that suffixed nominalizations are count nouns, while neuter nominalizations are mass nouns. A comparison between Greek, Dutch/German and English shows that different languages have different nominalization strategies: nominalization can take place on three layers: on the root level, on the nP level, and finally on the DP level. This explains the differences in distribution and interpretation among the different nominalization types across languages.

Sleeman analyses the human construction *de zieken* ('the sick people') in Dutch, which has mixed adjectival and nominal properties, and argues against the ellipsis analysis. An analysis within the framework of Distributed Morphology is proposed, which is a combination of the ellipsis analysis (without an empty noun) and the traditional derivational/conversion analysis.

Celda Choupina, comparing different approaches, studies cognate objects of verbs in Portuguese, *chorar um choro...* (lit. 'to cry a cry'), *cantar uma cantiga* ('to sing a song). Remember that Hale & Keyser (1993) proposed that English verbs such as *to saddle*, *to bottle* are formed from nouns by a movement from an N position to a V position within the lexical structure, according to a strong lexicalist view of morphological processes. According to another view, that supported by Distributed Morphology, there are no primitive word categories and, for instance, *dance*, in *the dance*, or *dance* in *to dance*, may be the result of the movement of an acategorical root to functional categories that define its final category as a noun or as a verb. At the same time, however, if there are cognate objects of certain verbs, why not consider that they are obtained by a movement of a root that leaves a copy, as argued by Haugen (2009), inspired as to the general idea of movement by copy of the Minimalist Program (Chomsky 1995)? We see that cognate objects are a central topic in the discussions regarding the form and function of grammar.

We sincerely hope that this volume will contribute to the discussion concerning the interface between Lexicon, Morphology, Syntax and Semantics not only amongst the linguists that were present at the workshop but also among the members of the linguistic community in general.

Ana Maria Brito  
July 2013

#### REFERENCES

- Ackema, P. & Neeleman, A. 2007. Morphology  $\neq$  Syntax. In Ramchand, G. & Reiss, C. 2007. (Eds.), 325-352.
- Alexiadou, A. 2001. *Functional Structure in Nominals: Nominalization and Ergativity*, Amsterdam: John Benjamins.
- Alexiadou, A., Iordăchioaia, G. & Schäfer, F. 2011. *Scaling the Variation in Romance and Germanic Nominalizations*. In Sleeman, Petra & Harry Perridon (Eds.) (2009) *The Noun Phrase in Romance and Germanic: structure, variation and change*. Amsterdam/Philadelphia: John Benjamins, 25-40.
- Chomsky, N. 1995. *The Minimalist Program*. Cambridge MA: The MIT Press.
- Di Sciullo, A. M. & Williams, E. 1987. *On the definition of word*. Cambridge, Mass.: The MIT Press.
- Embick, D. & Noyer, R. 2007. Distributed Morphology and the Syntax-Morphology Interface. In Ramchand, G. & Reiss, C. 2007. (Eds.), 289-324.
- Hale, K. & Keyser, S. J. (Eds.) *The View From Building 20: Essays in Linguistics in honor of Sylvain Bromberger*. Cambridge, Mass., MIT Press.
- Hale, K. & Keyser, S. J. 1993. On Argument Structure and the Lexical Expression of Syntactic Relations. In Hale, K. & Keyser, S. J. (Eds.), 53-109.
- Halle, M. & A. Marantz 1993. Distributed morphology and the pieces of inflection. *The View from Building 20*. In K. Hale & S. J. Keyser (Eds.), 111-176.
- Haugen, J. D. 2009. Hyponymous objects and Late Insertion. In *Lingua 119*, 242-262.
- Jackendoff, R. 2002. *Foundations of language. Meaning grammar and evolution*. Oxford: Oxford University Press.
- Julien, M. 2007. On the relation between Morphology and Syntax. In Ramchand, G. &

- Reiss, C. 2007. (Eds.), 209-238.
- Ljung, M. 1994. Conversion. In Asher, R. E. (ed.) *The Encyclopedia of Language and Linguistics*. Volume 2, Oxford: Pergamon Press, 758-759.
- Ramchand, G. & Reiss, C. 2007. (Eds.) *The Oxford Handbook of Linguistic Interfaces*, Oxford: Oxford University Press, 209-238.
- Sleeman, P. 2011a. The nominalised infinitive in French: structure and change. *Linguística. Revista de Estudos Linguísticos da Universidade do Porto*, volume 5, 1, 145-173.
- Sleeman, P. 2011b. Verbal and adjectival participles: Position and internal structure. *Lingua*, volume 121, 10, 1569-1587.
- Svenonius, P. 2007. 1...3-2. In Ramchand, G. & Reiss, C. (2007) (Eds.), 239-288.
- Varela, S. 2006. Conversion. In Keith Brown (Ed.) *Encyclopedia of Language and Linguistics*, 2<sup>nd</sup> edition, Amsterdam: Elsevier, 172-175.
- Williams, E. 2007. Dumping Lexicalism. In Ramchand, G. & Reiss, C. 2007. (Eds.), 353-381.





Artigos



# Nouns in apposition: Portuguese data

Graça Rio-Torto

gracart@fl.uc.pt

*Faculdade de Letras da Universidade de Coimbra (Portugal)*

*Centro de Linguística Geral e Aplicada*

**ABSTRACT.** The nature of N1N2 constructions, or nouns in apposition, is controversial: depending on the theoretical framework, they can be considered as compounds or as syntactic constructions.

Indeed, nouns in apposition function as a hybrid category, in a double way:

(i) the same lexical structure in apposition is viewed either as a coordinative construction, as a subordinative or as an attributive construction.

(ii) N2 functions as a modifier or as an attributive item of N1; in Portuguese, when plural is syntactically mandatory, N1 (the head) is systematically pluralized; N2 either rejects inflection or behaves as a predicator, allowing inflectional marks.

We claim that Romance NN behave as a specific type of compounds. This assumption is grounded on their behaviour by contrast with phrasal properties.

Portuguese compounds are characterized by a narrow relationship between internal structure, headness and inflectional patterns. In Portuguese, by default, the head of compound is inflected. NN related by an attributive semantic link are nowadays particularly unstable and problematic regarding inflection. Inflectional variation — widely attested — helps in determining the status of NN in apposition: as two inflectional patterns are available, we must verify if they correspond to two different constructions or to one structure with two readings. The analysis addressed is supported by empirical data of contemporary Portuguese language extracted from Brazilian and European databases, and requires the theoretical articulation of a double predicative class of N2 (holistic and partitive) with inflectional fluctuation of attributive N2 in the second situation: performing a continuum, double inflection is close to holistic predication and single inflection (of N1) is close to partitive predication; systematic double inflection is close to coordination and inflectional oscillation is close to attribution. The predicative power of nouns in apposition supports their partitive and their holistic reading. Inflectional fluctuation of N2 illustrate in an optimistic way their hybrid nature.

**KEY-WORDS.** lexicon, morphology, composition, apposition, Portuguese Language, Romance Languages

## 1. Introduction

Appositions are defined as «constructions in which either a noun phrase or a noun modifies the head of another noun phrase» (Kornfeld 2003: 212). This definition is neutral regarding the problems arising from such a concept, which poses interesting issues to the interface between Lexicon, Morphology and Syntax.

Nouns in apposition function as a hybrid category, in a double way:

- (i). N2 functions as a modifier or as an attributive item of N1; when plural is syntactically mandatory, Portuguese N1 is systematically pluralized; N2 either rejects inflection or allows inflectional marks;
- (ii). the same NN lexical structure is seen as a coordinative, as an attributive or even as a subordinate construction<sup>1</sup>. The variation in the inflectional patterns supports this oscillation.

The nature of NN constructions is controversial: they are considered either as compounds, mainly by lexicalist positions (Bisetto & Scalise 1999; Ribeiro 2010; Rio-Torto & Ribeiro 2012), or as syntactic constructions (nominal phrases), as supported by syntactic approaches (Kornfeld 2003 & Fábregas 2005). Di Sciullo & Williams (1987) consider compounding in English as a morphological process, and Lieber (1992) and Baker (1995) defend its syntactic nature.

Portuguese compounds are characterized by a narrow relationship between internal constituency, headness and inflectional patterns. In Romance languages NN compounds admit internal inflection (cf. plural morpheme *-s*), mandatory in N1 (*navios-escola* 'lit. ships-school') and sometimes also in N2: *couves-flores* 'cauliflowers' (cf. Cunha & Cintra 1985: 188). This internal inflection stresses their proximity with phrasal constructions.

In Portuguese, with the exception of NN structures, all types of compounds are regular according plural patterns: by default, the head of

---

<sup>1</sup> Spanish NN constructions illustrate this heterogeneity: *artículos-viaje* 'lit. articles-travel' are described as subordinate compound (Guevara 2012), *hombres-rana(s)* 'frogmen' as attributive compounds (Rainer 2012), *escuelas-museus* 'schools-museums', *bares restaurantes* 'bars restaurants' as coordinate compounds and *empresas-fantasma* 'ghost companies', *situaciones-limite* 'limit situations', *visitas-relámpago* 'lightning visits' as attributive appositions (Kornfeld 2003).

a compound is inflected (cf. section 4). On the contrary, NN related by an attributive link are nowadays unstable and problematic regarding inflection (cf. section 5.). Inflectional variation highlights the double nature of nouns in apposition: since two inflectional patterns are used, it is crucial to know if they correspond to two different constructions or to one structure with two readings. This goal entails the conciliation between (i) the nature of N2, when it is a predicator, and (ii) the no mandatory agreement (*visitas-relâmpago(s)*) between the nouns. This is the focus of section 6.

## 2. The state of the art: issues and assumptions

According to Val Alvaro 1999, there are three classes of nominal appositions: denominative (*rio Tamisa* 'river Thames'); classificative (*uma pintura Renascimento* 'a Renaissance painting'); and attributive (*homem-bomba* 'bomb-man', *situação-limite* 'limit situation', *amarelo-limão* 'lemon yellow', *café-concerto* 'café-concert').

We claim that nouns in apposition are compounds with morphological nature and not syntactic objects. Having a different structure and behavior, nouns in apposition are not prototypical phrases. As syntactic islands, appositions are similar to compounds (cf. section 3.). Two correlated questions must be addressed: is internal inflection compatible with their structural opacity? And is lexical integrity applicable to Romance compounds?

### 2.1. Theoretical statements

Syntacticist approaches to word formation (cf. Lieber 1992) claim that there are two types of words: (i) derivational or morphological words, built according to word syntax, and (ii) syntax words, generated by sentential syntax. According to this approach, compounds are syntax words. The mechanisms that govern morphological and syntactical objects are the same. Kornfeld (2003:223) states that «Spanish N+N compounds are syntactic words, that is, they derive from a syntactic (and not a morphological) means of word formation. This implies that, in Spanish, compounding is a derived phenomenon related to syntax and lexicon, and not a primitive notion of morphology». VN Romance compounds are perhaps the constructions in which the internal structure reproduces, with adaptations, the patterns of

prototypical phrases.

Lexicalist approaches state a different nature of lexical and syntactic units, rules and products. Within the lexicalist framework, compounds are plurilexematic units in which meaning refers to a delimited reality, with intensional and ontological identity. Each compound is used as a holistic denomination, and meaning does not always coincide with the parts' sum of meanings. Although plurilexematic, a compound behaves as a single lexical unit, exhibiting syntactic and lexical opacity. As "syntactic islands", compounds are marked by the inaccessibility of their constituents for syntactic rules. But the opacity to syntactic rules of NN compounds has been modulated, in order to account for the specific and scalar behavior of Romance compounds with regards to lexical integrity (Lieber & Scalise 2007).

Despite their nature, NN constructions can be inserted into syntactic structures and behave as listemes (Di Sciullo & Williams 1987), as linguistic expressions memorized and stored as listed objects. Section 3. describes the classes of Portuguese compound in order to situate the relevant features of nominal compounds in apposition.

## 2.2. Arguments against syntactic nature

This section presents crucial arguments in favor of the non-syntactic nature of nouns in apposition.

As their internal structure shows, NN appositions are not prototypical nominal phrases. As syntactic islands, appositions are similar to compounds. N2 in apposition is never preceded by a determiner. The ungrammaticality of such a sequence is systematic.

The canonical functional projections of nominal phrases are: [DP [QP [AgrP [NP]]]].

NN in apposition can agree in number with the nominal head they modify, but never admit QPs or DPs:

- (1) Este<sub>DP</sub> abre-latas vs. \*abre as<sub>DP</sub> latas, \*abre estas<sub>DP</sub> latas, \*abre muitas<sub>QP</sub> latas  
'This can-opener' vs. '\*opens the<sub>DP</sub> cans', '\*open these<sub>DP</sub> cans',  
'\*open many<sub>QP</sub> cans'

In opposition to phrases, compounds are opaque to internal syntactic operations: the constituents cannot be displaced (cf. 2-3), elided and anaphorically resetted (cf. 4), or coordinated (cf. 5):

- (2) *biocombustível* 'the biofuel' vs. \**o combustível-bio* '\*the fuel-bio'
- (3) *este saca-rolhas* 'lit. this opener stoppers' vs. *este rolhas-saca* 'lit. this stoppers opener'
- (4) \**este saca-rolhas é bom e saca-as bem* '\*this opens stoppers is good and opens them well'
- (5) \**este saca-rolhas e tampas*, lit. '\*this open stoppers and lids'

This kind of NN does not allow modifications (cf. 6) or complementations (cf. 7) with scope on one of the nouns. By contrast with (6), «*Este saca-rolhinhas*», with *-inh-*, a diminutive suffix, is accepted if interpreted as 'this little stopper opener'.

- (6) \**este saca-rolhinhas* '(-*inh-*: a diminutive suffix) \*lit. this open little stoppers'
- (7) *este saca-rolhas [de cortiça]<sub>pp</sub>* 'lit. this stopper opener [of cork]<sub>pp</sub>' (or cork stoppers, not plastic stoppers, for instance) vs. \**este saca [rolhas [de cortiça]<sub>pp</sub>]<sub>OD</sub>* 'lit. this open [stoppers of cork]

The structures analyzed are similar to other morphological constructions and to prototypical compounds in that they follow the 'Lexical Integrity Principle': the insertion of lexical items in the interior of this structure is mainly avoided, and any internal intrusion from syntactic procedures (elision, coordination, movement) is forbidden. [NN] in apposition (as well as [NA] or [N Prep. N]) have an internal structure which is not accessible to syntax, and that is clear by the impossibility of modifying only one of the nouns. By contrast with (6), «*Este saca-rolhinhas*», with *-inh-*, a diminutive suffix, is accepted if interpreted as 'this little cork opener'.

As we will observe (cf. section 6.1.), the fact that some NN compounds allow internal inflection does not deny the 'Lexical Integrity Principle' parameterized for Romance languages.

### 3. Classes of compounds

According to the morphological nature of their constituents, NN compounds can be divided into two groups: (i) single phonological words, also known as ‘morphological’ or neoclassical compounds; and (ii) double phonological words or ‘morphosyntactic’ compounds<sup>2</sup>. The first class includes roots, the other, words.

As far as the grammatical linking between the constituents of compounds is concerned, they are organized into three groups: subordinated, coordinated and attributive compounds.

The units of coordinate compounds are linked by a copulative operator, explicit (*whisky-and-soda*, *pepper-and-salt*) or not (*surdo-mudo* ‘deaf dumb’). The relationship between the two units, both heads of the construction, is symmetrical and additive. Both words share head-like characteristics.

We can distinguish appositional compounds, such as Eng. *maid-servant* or *trabalhador-estudante* ‘worker and student’, hyponyms of both *maid* and *servant* or ‘worker and student’ and copulative compounds (also called *dvanda* in the Sanskrit tradition), as *Boznia-Herzgovina*, (train) *Lisboa-Porto* ‘Lisbon-Oporto’. In these cases the compound is not an hyponym of either element, but the elements denote separate entities which combine to form a denomination of a new entity related to them.

Coordination, in compounding as well as in syntax, requires that both units belong to the same grammatical label. The major template is  $[NN]_N$ : *ator-autor* ‘actor-author’, *norte-sul* ‘north-south’, *trabalhador-estudante* ‘worker-student’.

In subordinate type the relation between the two constituents is a complement relation — a predicator-argument relation, a relation of implicit syntactic dependency —, namely a subordinative one, as  $[VN_{OD}]_N$  pt. *salva-vidas* ‘life-gard’, eng. *pickpocket*, eng. *taxi driver* show. The non-head (*vidas*, *pocket*, *taxi*), being a complement of the (de)verbal head (*salva*, *pick*, *driver*), is conceived, in LCS (Lexical Conceptual Structure), as a Thing (cf.

---

<sup>2</sup> As Rio-Torto & Ribeiro (2012) show, the duality ‘morphological’ vs. ‘morphosyntactic’ is not the most adequate to the compound data of Portuguese language. This dichotomy has been proposed by Di Sciullo & Williams (1997), to whom morphological objects include morphemes, roots, stems and compounds, and non morphological objects, such as VN, VA, VAdv, are phrasal structures. Structures like *database*, *weekend* would be syntactic products, reanalyzed and recategorized as words that can occupy a nuclear position (X<sup>0</sup>) in syntax. In fact, these two classes are better described as ‘single phonological words’ and ‘double phonological words’. Note that in Romance languages the absence of internal inflection would prove that VN nouns are compounds (not phrases) and that phonological words such as *database* or *weekend* might be also considered as compounds.



Jackendoff 1990).

In modificative or attributive type the constituents are not linked by a dependency relationship: the non-head modifies and delimits the intension of the head, or specifies an attribute of the head's body. The parts of the compound are linked by a modifier-modified or a modified-modifier relation. The non-head may be adjectival (*bomba atômica* 'atomic bomb'), nominal (*palavra-chave* 'key-word') or a prepositional phrase (*cartão de crédito* [lit. card of credit]). In LCS terms (Jackendoff 1990: 43-58) the modifier is interpreted as a Property and not as a Thing.

Compounds based on coordinate, subordinate and attributive internal relations are identified on both 'morphological' and 'morphosyntactic' compounds (cf. Table 1).

TABLE 1 – Compounds' internal constituency and grammatical relations

	Coordinate	Subordinate	Attributive
One phonological word 'Morphological' compound	[ <i>agridoce</i> ] <sub>A</sub> 'bittersweet'	[ <i>herbicida</i> ] <sub>N</sub> 'herbicide'	[ <i>neurocirurgia</i> ] <sub>A</sub> 'neurosurgery'
Two phonological words 'Morphosyntactic' compound	[ <i>poeta-pintor</i> ] <sub>N</sub> 'poet-painter'	[ <i>beija-mão</i> ] <sub>N</sub> (kiss hand) 'hand kissing'	[ <i>palavra-chave</i> ] <sub>N</sub> (word key) 'key-word'

#### 4. [NN]: inflectional patterns

In Romance languages the head and the grammatical relations between the constituents play a crucial role on compound inflection.

##### 4.1. Portuguese

In Portuguese there is an indelible relation between inflection and head. Nevertheless, this is not the only dimension that rules compound inflection. Other criteria are activated, such as stress structure, morphological nature of the constituents or grammatical relations between them.

Portuguese compounds may be divided into three classes, according to the presence of inflection at the right border of the hole compound (pattern 2), at the head (pattern 3), or its external occurrence (pattern 1). In

Portuguese, plural is primarily codified by the adjunction of -s to a nominal/adjectival form or to a determiner.

Pattern 1. [Det<sub>INFLECTION</sub> [XY]<sub>N<sub>N</sub></sub>]: exocentric compounds

Exocentric compounds in whose internal structure a V occurs do not allow internal inflection. It is the determiner that expresses the plural.

i) [VN]<sub>N</sub>: *o(s) lava-louça* (wash tablewear)'dishwasher'

ii) [VAdv]<sub>N/A</sub>: *o(s) fala-barato* (speak cheap)'gasbag(s)'

Pattern 2. [[X][Y<sub>INFLECTION</sub>]]<sub>N</sub>: inflection on the right constituent with scope on the whole

When the compound is a single phonological word (univerbation), inflection is mandatory at the end of it, as coordinate *sinojaponês(es)* 'sinojapanese', subordinate *herbicida(s)* 'herbicide' and attributive *hidromassagem(s)* 'hydromassage' illustrate.

This group also includes compounds formed by neoclassical roots, such as *anglófilo(s)* 'Anglophile(s)', *carnívoro(s)* 'carnivorous', and agglutinated compounds such as *aguardente(s)* 'brandy(ies)', *artimanha(s)* 'trick(s)'

Pattern 3. Head inflection: single manifestation [X<sub>INFLECTION</sub> Y] or double manifestation [X<sub>INFLECTION</sub> Y<sub>INFLECTION</sub>]

The head of the compound is pluralized. Three situations can occur.

(i) [NN]<sub>N</sub> in attributive/modifying compounds: *visita(s)-relâmpago* 'lightning-visit(s)'

(ii) [NN]<sub>N</sub> coordinate compounds: both heads are inflected: *médico(s)-escritor(es)* 'doctor(s)-writer(s)'; *surdo(s)-mudo(s)* 'deaf(s) and dumb(s)'

(iii) [NA]<sub>N</sub> or [AN]<sub>N</sub> head inflection and agreement: both constituents are inflected, due to a mandatory agreement between noun and adjective: an example such as *escola(s)<sub>N</sub> primária(s)<sub>A</sub>* (lit. school primary) 'primary school(s)' displays this kind of pattern, in that the head noun is inflected and the adjective agrees in number with it.

In sum, inflection is mainly governed by the head(s): two-inflectional features occur when the compound is double-headed (a scenery restricted to coordinative compounds). When the compound is  $[NA]_N$  or  $[AN]_{N'}$ , inflectional agreement is mandatory between N and A and these compounds also exhibit two inflectional features. NN patterns of inflection require a deep analysis. Table 2 summarizes the inflectional patterns of Portuguese compounds.

TABLE 2 – Internal constituency, grammatical relations and inflectional patterns

Inflectional patterns	Coordination	Attribution	Subordination
(i) exocentricity: [Det <sub>INFL</sub> [X Y]]	<i>o(s) pára-arranca</i> 'stop/start'	<i>o/a(s) fala-barato</i> (speak cheap) 'gasbag'	<i>o/a(s) sabe-tudo</i> 'know all'
(ii) morphological compound:[XY] <sub>INFL</sub>	<i>sinojapônês(es)</i> 'sino Japanese(s)'	<i>sul-africano(s)</i> 'south African(s)'	<i>herbicida(s)</i> 'herbicide(s)'
(iii) head inflection	<b>Coordination</b>	<b>Attribution</b>	
	head inflection	head inflection and agreement	
	<i>poeta(s) -pintor(es)</i> 'poet(s)-painter(s)  <i>surdo(s)-mudo(s)</i> 'deaf(s)and dumb(s)	<i>fim(s) ) de semana</i> (end(s) of week) 'weekend(s)' <i>palavra(s) chave</i> (word(s) key) 'key word(s)' <i>arma(s) branca(s)</i> (weapon(s) white(s)) 'sharp weapon(s)' <i>grande(s) superfície(s)</i> (big surface(s)) 'shopping center(s) '	

#### 4.2. Spanish

In Spanish the situation is not very different from that of Portuguese. Agreement in NN compounds varies according to the grammatical relations between constituents. The double marking of plural is typical of coordinate compounds (Guevara 2012). Kornfeld (2003: 213) claims that there is no number agreement between the head and N2 in 'classifier appositions' such as *hombres-pájaro* 'lit. men-bird', *carpetas-mochila* 'lit. folders backpack' and in 'attributive' ones (*situaciones-límite* 'lit. situations limit', *visitas-relámpago* 'lit. visits lightening'). In many NN compounds, the plural of N2 can be observed (cf. (s) signalizes plural optionality): *momentos*

*claves(s)* ‘moments-key’, *empresas fantasma(s)* ‘phantom companies’, *situaciones límite(s)* ‘limit situations’, *visitas relámpago(s)* ‘lightning visits’, *visitas sorpresa(s)* ‘surprise visits’. According to Rainer (2012: 378), this fact supports the conversion of N2 into an adjective.

Despite the similarities between nouns and adjectives, their nature (Bosque 1998, Wierzbicka 1986) and their syntactic way of functioning is very different. In the cases mentioned above, we believe that N2 is not converted into an adjective (cf. Martinho, in this book): it is a noun with a predicative functioning, a predicator.

#### 4.3. French

In French NN compounds, agreement varies according to the grammatical relations between the nouns. According to Villoing (2012: 53), «In coordinated and attributive compounds, plurality is marked on both constituents, while in subordinating NN compounds, it is only marked on the first element». Next table contains examples given by the author.

TABLE 3 – Patterns of NN plurality in French

N1 <sub>[+PL]</sub> N2 <sub>[-PL]</sub>	N1 <sub>[+PL]</sub> N2 <sub>[+PL]</sub>	
Subordination	Attribution	Coordination
bébés-éprouvette ‘test tube babies’ timbres-poste ‘postage stamps’ pauses-café ‘coffee breaks’	choux-fleurs ‘cauliflowers’ hommes-grenouilles ‘frogmen’ poissons-chats ‘catfish’	bars-tabacs ‘bars-tobacco’ boulangers-pâtisseries ‘bakers’

When the compound includes a color adjective, this is in principle invariable: *chemises rose-pâle* ‘pale pink shirts’ (\*roses-pâles), *cheveux brun-clair* ‘hair light brown’ (\*bruns-clairs). But, as Villoing also remarks (2012: 54), some variation is admitted: *des cuisines vert-olive* ‘kitchens olive green’ ou *vertes-olives*; *des robes bleues-ciel* ‘Blue Sky dresses’.

#### 4.4. Italian

In Italian, NN compounds such as *capi-stazione*, *navi-traghetto* reveal inflection on the head (cf. *-i-* plural's morpheme). Masini, Scalise (2012: 85) state that this type of compounds «[...] despite having two semantic heads, has just one formal head that is located on the canonical position of the head in Italian, i.e., on the left». Indeed, the formal head *nave* is responsible for the feminine gender of the whole (*traghetto* is masculine).

#### 5. Inflectional variation in attributive Portuguese NN

Cunha & Cintra 1985: 188, in their reference grammar of Portuguese, assume that the plural of NN compound occurs at least in N1 (*navios-escola* 'lit. ships-school') and sometimes also in N2 (*couves-flores* 'cawliflowers'). Recent data attest that NN plural variation is largest than in the past.

Data extracted from the written database CETEMPúblico ([www.linguateca.pt/cetempublico/](http://www.linguateca.pt/cetempublico/)), a large corpus of Portuguese newspaper language, show a relevant variation regarding plural of NN Portuguese compounds. Parallel data have been extracted from a Brazilian newspaper database — CETENFOLHA (Folha de São Paulo) —, also available at [www.linguateca.pt](http://www.linguateca.pt). The following table displays the number of tokens (by decreasing order) of each NN structure collected in both databases.

TABLE 4 – NN Pluralization in Portuguese and Brazilian databases

Single inflection: [N <sub>INFLECTION</sub> N]			Double inflection: [N <sub>INFLECTION</sub> N <sub>INFLECTION</sub> ]		
	CETM Público	Folha SP		CETM Público	Folha SP
(8 a) decretos-lei 'decree-laws'	179	3	(8 b) decretos-leis	60	9
(9 a) palavras-chave 'keywords'	179	23	(9 b) palavras-chaves	3	9
(10 a) projectos-piloto 'pilot projects'	89	2	(10 b) projectos-pilotos	3	0
(11 a) pombos-correio 'pigeons'	21	0	(11 b) pombos-correios	15	1
(12 a) homens-rã 'men-frog'	12	0	(12 b) homens-rãs	9	0
(13 a) cidades-dormitório 'dormitory towns'	12	0	(13 b) cidades-dormitórios	3	2
(14 a) cidades-fantasma 'ghost towns'	5	3	(14 b) cidades-fantasma	4	1
(15 a) empresas-fantasma 'ghost companies'	7	1	(15 b) empresas-fantasma	4	1
(16 a) fatos-macaco 'work overalls'	14	0	(16 b) fatos-macacos	7	0
(17 a) cafés-concerto 'café-concert'	17	0	(17 b) cafés-concertos	4	0
(18 a) mulheres-objecto 'women-object'	2	0	(18 b) mulheres-objects	0	0
(19 a) visitas-relâmpago 'visits-lightning'	5	5	(19 b) visitas-relâmpagos	0	0
(20 a) governos-sombra 'shadow governments'	4	0	(20 b) governos-sombras	0	0
(21 a) cães-polícia 'police dogs'	23	0	(21 b) cães-polícias	33	0
(22 a) couves-flor 'cauliflowers'	1	0	(22 b) couves-flores	5	0

The data presented by decreasing values show that in Portugal as well as in Brazil the old pattern (Single inflection), with N1 marked by plural, is best represented than the new one (Double inflection).

The exceptions are the following data from Portugal: *cães-polícia* 'police dogs' (23 tokens) vs. *cães-polícias* (33 tokens) and *couves-flor* 'cauliflowers' (1 token) vs. *couves-flores* (5 tokens) and *palavras-chaves* in Brazil.

Nevertheless, the plural of N2 increased in the last decades and promotes a changing stage of Portuguese, Italian, Spanish languages.

In order to determine if single or double plural pattern is correlated with syntactic context, we analyzed the phrasal and sentential constructions selected by NN, and we conclude that there are no syntactic constraints to any inflectional pattern. Data illustrate (cf. next table) that single and double inflection takes place in the same syntactic conditions: variation remains independently of them.

TABLE 5 – Syntactic contexts associated with NN constructions

<b>Single inflection:</b> [X <sub>INFL</sub> Y]	<b>Double inflection:</b> [X <sub>INFL</sub> Y <sub>INFL</sub> ]
<b>(i) relative subordinate sentence:</b>	
<i>empresas-fantasma</i> que recebem capitais ‘ghost firms that receive capital’  <i>empresas-fantasma</i> que se dedicam à subcontratação ‘ghost companies that are engaged in outsourcing’	<i>empresas-fantasma</i> s que se dedicam ao tráfico ‘ghost companies that that are engaged in trafficking’
<b>(ii) prepositional phrase with modifier function:</b>	
<i>empresas-fantasma</i> para burlas financeiras ‘ghost companies for financial scams’	<i>empresas-fantasma</i> s de comunicação social ‘ghost companies of social media’

The following table illustrates the three solutions adopted by Portuguese, Spanish and French for cognate NN. The same shading pattern is adopted for Romance equivalent NN.

TABLE 6 – NN Pluralization in Portuguese, Spanish and French

<b>Pluralization</b>	<b>Portuguese</b>	<b>Spanish</b>	<b>French</b>
<b>I.</b> N <sub>[+PL]</sub> - N <sub>[±PL]</sub>	visitas-relâmpago	visitas-relâmpago	visites-éclair
	empresas-fantasma	empresas-fantasma, situaciones-límite	
	palavras-chave		notions/mots-clé
<b>II.</b> N <sub>[+PL]</sub> - N <sub>[±PL]</sub>	palavras-chave(s)	momentos claves(s) visitas relâmpago(s)	mots-clés (web) moments-clés (web) notions-clés (web)
	empresas-fantasma(s)	empresa fantasma(s), situaciones límite(s)	
	homens-rã(s)	hombres-rana(s)	
<b>III.</b> N <sub>[+PL]</sub> - N <sub>[+PL]</sub>			hommes-grenouilles

This table displays three situations: in III. both nouns are systematically marked by plural. In I. only N1 is pluralized, which means that N1 is seen as the head and N2 as a noun in apposition: an adjective would typically agree in number with the head of the compound/phrase. In II. N2 is either marked or not marked by plural: this situation corresponds to a fluctuation in which N2 is taken by the speakers as a noun, traditionally not pluralized (cf. table 4), or as a nominal predicator, and then marked by plurality.

There is a relevant asymmetry between French data collected outside and inside the web. The newest pattern  $N_{PL} - N_{(\pm PL)}$  is well attested in several web sites.

### 5.1. Debating the data

Coordinate compounds behave as follows: both nouns are marked by plural, because a double head is present. NN coordinate compounds are defined as ‘an N that is both N1 and N2’, that is to say, a holistic merge of both constituents occurs. This is not the case of fr. *hommes-grenouilles*, pt. *homens-rã(s)*, sp. *hombres-rana(s)* ‘frogmen’, because the human beings of this class are not frogs. How to explain that in Portuguese and in Spanish the plural of *homens-rã(s)* and *hombres-rana(s)* oscillates between Pattern I (N[+pl])-N[-pl]) and Pattern III (N[+pl])-N[±pl])? The conjunction of N1 and N2 into a unique being is referentially mandatory in case of coordination, and it is unavailable with the literal meanings of both nouns. Semantic coindexation is possible, if undergrounded by a figurative transfer.

NN such as *empresas-fantasma* ‘ghost companies’, *visitas-relâmpago* ‘lit. visits-lightning’, *palavras-chave* ‘keywords’, *situações-limite* ‘limit-situations’ can be described as attributive compounds: they are left headed endocentric constructions (not double headed constructions) whose meaning can be paraphrased as an N ‘that is N1 and also an N2, with a metaphoric interpretation of N2’. The gender of NN is deduced from the head:

(25) as	<i>empresas-</i>	<i>fantasma</i>
(FEM.PL)	(FEM.PL)	(MASC.SING)
(26) os	<i>homens-</i>	<i>rã</i>
(MASC.PL)	(MASC.PL)	(FEM.SING)



Semantically, the NN attributive classification, characterized by a relation of modification between the head and the non-head constituent, is more adequate than a coordinative classification.

However, a great rate of variation can be observed between single and double plural strategy of N2 (cf. tables 4, 6). We claim that the plural of N2 does not transform N2 into an adjective, it is due to its predicative functioning. According to Fábregas (2005: 206), «Predication in an appositional construction is due to the presence of a relational structure that makes it possible that the noun that acts as a predicator selects the other element as its subject».

What features of N2 allow it to function as a predicator? Why doesn't the language select an adjective? Some N2 denote beings marked by animacy (*fantasmas* 'ghosts', *rãs* 'frogs', *macacos* 'monkeys', *polícias* 'police') which combine with N1 marked by the same feature. Other N2 denote inanimate objects (*chaves* 'keys') and combine in an open way with N1 denoting diverse classes of entities.

Some denominal adjectives cognate of N2 are not available, probably also because the mental lexicon of the speakers contains information about the predicative values of these nouns:

- (25) *empresas-lídere(s)*: what would the denominal adjective derived from *líder* be?
- (26) *situações limite(s)*: what would the denominal adjective derived from *limite* be?
- (27) *momentos chaves(s)*: what would the denominal adjective derived from *chave* be?
- (28) *visitas relâmpago(s)*: what would the denominal adjective derived from *relâmpago* be?

The adjunction of derivational suffixes *-ic-* and *-al*, the most productive in the formation of relational adjectives in contemporary Portuguese, is grammatical, but indeed they are not used with these nouns. In other cases the adjectives correlated with N2 are blocked by a different meaning:

- (29) *eleições farsa(s)*: [N *farsantes*] cannot replace [N *farsa(s)*], because *farsantes* has an agentive meaning that *farsa* lacks
- (30) *empresas-lídere(s)* ‘lider-organization(s)’ vs. *liderantes* ‘leading’
- (31) *situações-limite* ‘limit situations’ vs. *limitantes* or *limitativas* ‘limiting or restricting’
- (32) *visitas-surpresa(s)* are visits unexpected, pleasant or not; the adjective *surpreendentes* means ‘surprising; amazing; fabulous’
- (33) the adjective *milagrosas*, in *terapias milagrosas* ‘miraculous, fabulous therapies’, can codify a causative (‘that products miracles’) and a figurative (‘fabulous’) meaning that is not carried by *milagre* (*terapias milagre(s)* ‘miracle therapies’).

The predicative impact of the noun is greater than the one of the adjective. As the construction with a nominal predicator is less prototypical, it is open to a broad variety of readings. Typically, a noun denotes a kind of an entity, rather than a single property (Wierzbicka 1986). But these nouns behave as predicators and, by consequence, at the same time they can perform two functions: they delimit a recognizable category of entities on the base of a property: *chave* ‘crucial, core’, *fantasma* ‘inexistent, false’, *limite* ‘limit, extreme’, *relâmpago* ‘short and unexpected’, *surpresa* ‘unexpected’.

When the cognate adjective is available, as in the following case, the adjective is not equivalent to the noun:

- (34) a. [*crianças*]<sub>N</sub> [*prodígio(s)*]<sub>N</sub> ‘lit. children prodigy(s)’  
b. [*crianças*]<sub>N</sub> [*prodigiosas*]<sub>A</sub> ‘prodigious children’.

The adjective (cf. 34 b.) assigns a stable feature to the entity denoted by N1. Apart from its categorizing frame, N2 modifies N1 (cf. 34 a.) through its predicative features (a *prodigy* ‘something extraordinary’). N1 is then characterized by a noun that is a predicator: a *prodigy*. In this way, N1 is not defined or definitively characterized by the property associated to a state adjective (cf. 34 b.).

Summarizing: the noun is a more optimal and impressive choice than the cognate adjective, which is often unavailable. The internal cohesion of the construction is strengthened with a noun. An adjective can be displaced

(NA or AN) and modifiable, which does not happen with N2.

The next section aims at articulating the duality of plural patterns with the predicative classes of N2. Before, we remember that the Portuguese language is characterized by a rich contextual inflection: N1 is the head of the structure, and the modifiers must agree with it. Agreement is systematic when the modifier is an adjective, and unsystematic when the modifier is a nominal predicative.

#### 6. NN: grammar and predication

Regarding non-coordinative NN constructions, two situations must be distinguished:

- (i) a subclassification type: N2 denotes a subclass, mainly a telic one, of N1: *artigos-viagem* 'lit. articles-travel', *seguro-saúde* 'health insurance'. The construction refers to a class of entities (N1) that is underspecified by the features of N2. In this case, as in other subordinate constructions, only the head is marked by plurality.
- (ii) a qualitative type: N2 qualifies N1, with a set of properties typical of N2: the whole corresponds to 'N1 with some qualitative properties of N2': *desporto-rei* 'lit. sport-king', *situação-limite* 'limit situation', *empresa-líder* 'lit. company-leader': a stereotypical feature associated to an entity (*líder* 'leader', *fantoche* 'puppet', *piloto* 'pilot') is attributed, but the word that codifies this feature is not an adjective, although it denotes a property (*chave* 'crucial, core', *limite* 'limit, extreme', *rei* 'major'). These N2 behave either as nouns, rejecting plural, or as adjectives, being pluralized together with the noun (N1) they modify.

As the first subtype illustrates, plural variation of N2 is not unrestricted: when N2 is uncountable (cf. *saúde* 'health') or when N2 is used in a generic meaning and with a telic relation with N1 (cf. *artigos-viagem* 'articles for travel'), singular of N2 is the preferential solution. Henceforth, we are no longer focused on these subordinate compounds.

6.1. NN qualitative type and the ‘Lexical Integrity Principle’ applied to Romance Languages

The NN structures analyzed are similar to prototypical compounds since they obey the ‘Lexical Integrity Principle’: the insertion of lexical and grammatical items in the interior of this structure is mainly avoided, and any internal intrusion from syntactic procedures (elision, coordination, movement) is forbidden. NN in apposition (as well as NA) have an internal structure which is not accessible to syntax, and that is visible by the impossibility of modifying only one of the nouns.

As many authors (Scalise & Bisetto 1999, Fábregas 2005, Lieber & Scalise 2007) assume, internal inflection is not an adequate and applicable criterion for Romance languages; internal and double plural does not deny the ‘Integrity Principle’ of Romance compounds. Agreement is a syntactic device with scope on the whole structure; sometimes agreement is visible in some of the constituents and other times it is visible in all the constituents. Nevertheless, inflectional patterns are crucial for the delimitation of compound classes.

6.2. Classes of predication

The solutions presented for NN constructions are not disjunctive: NN compounds are to be taken as appositions with syntactic atomicity listed in the lexicon. It is well known that phrases can be ‘downgraded’ to words by a process of reanalysis. Within a syntactically oriented morphology, Fábregas (2005) states that NN are a specific type of phrases with a specific type of a projection. Next table presents the three classes of appositions proposed by this author.

TABLE 7 – Classes of NN constructions (Fábregas 2005: 202-203)

(i) PREDICATIVE APPPOSITIONS	Partitive: <i>pantalones campana</i> ‘trousers-bell, trousers with the shape of a bell’
	Holistic: <i>poeta pintor</i> ‘poet painter’, ‘a poet that is also a painter’
(ii) KIND APPPOSITIONS: <i>el presidente Zapatero</i> ‘the president Zapatero’	
(iii) PROPER NAME / SURNAME: <i>Noam Chomsky</i>	

We assume that some NN constructions admit two ways of projection: as plurilexematic words, they can display a more or a less phrasal nature, with different patterns of inflection. As Fábregas (2005: 206) says, these «two classes of predicative appositions are the manifestation of the same type of structure, and they are only different in terms of the functional projections merged with that structure».

NN containing a predicative N2 are in the border of morphological and phrasal constructions. Within an attributive NN construction, as *festas surpresa* ‘surprise parties’, N2 is a noun, not an adjective, which explains that its plural is not systematic<sup>3</sup>; but N2 presents an attributive value: the noun *surpresa* ‘surprise’ (*festas surpresa* ‘surprise parties’) adds to its categorizing frame a predicative value; in *festas surpreendentes* ‘surprising parties’ the cognate adjective *surpreendentes* has a predicative meaning and lacks a categorizing power. In these NN constructions, there is a full predication of N1 combined with a partitive predication of N2 (Fábregas 2005)<sup>4</sup>.

This double nature of this kind of N2 undergrounds the holistic or the partitive predication allowed and explains the oscillations regarding plural as well as regarding the grammatical relation ( $\pm$  attributive) between the nouns.

In the last decades N2 plural tends to increase (cf. *empresas-âncora(s)* ‘lit. companies anchor(s)’, *empresas-membro(s)* ‘lit. companies member(s)’, *escolas-depósito(s)* ‘lit. schools deposit(s)’, *escolas-aviário(s)* ‘lit. schools aviary(s)’). As mentioned above (section 4.), inflectional oscillation is close to attribution and systematic double inflection is close to coordination.

The next table intends to illustrate that partitive and holistic predication are two poles of a scale and that the gradience between single and ((a)systematic) double inflection is correlated with these two kinds of predication. The data are distributed in a continuum on the base of these two dimensions. The left column contains NN constructions with N2 [-Plural] and in the other extreme of the scale both nouns of the construction are pluralized. In the central position of this *continuum* there are NN constructions with a hybrid

<sup>3</sup> Remember that when the construction contains an adjective, the compound NA has to be double inflected.

<sup>4</sup> A NN construction with a partitive predication is defined as follows by Fábregas 2006, p. 114: «en ella el núcleo es un sintagma predicación (Bowers 1993, 2001), una proyección relacional que selecciona como complemento un predicado y como especificador el argumento del predicado».

status: semantically they are close to attribution, inflectionnaly they behave either like attributive compounds or like coordinate compounds, and the predication they perform is then between a partitive and a holistic one.

TABLE 8 – Classes of predicative NN constructions

Partitive predication		Holistic predication	
N2 [ $\pm$ Plural] (Inflectional oscillation )		N2 [+ Plural]	
single inflection	double inflection	(Systematic) double inflection	
<i>empresas-chave</i> 'lit. companies-key'	<i>empresas-chaves</i> 'lit. companies-keys'	<i>escolas-empresas</i> 'lit. schools-companies'	
<i>empresas-líder</i> 'lit. companies-leader'	<i>empresas-líderes</i> 'lit. companies-leaders'	<i>igrejas-fortalezas</i> 'lit. churches-fortresses'	
Attribution Single inflection	<-----> <----->	Coordination Double inflection	

Plural oscillation of N2 (cf. left vs. central columns) depends on the way in which N2 is processed by the speakers: (i) as a noun in apposition, and therefore not pluralized, or (ii) as a nominal predicator, and then marked by plurality. Systematic double inflection is close to coordination and inflectional oscillation is close to attribution. Single inflection is close to partitive predication and double inflection is close to holistic predication.

### 6.3. Final remarks

Only binary and non-interactive conceptions of language domains reject hybrid and boundary categories: it is not impossible that a NN structure can be seen as a construction in the frontier of morphology and syntax, with two readings. The borderline nature of a noun codifying a class by means of predicative features explains the fuzziness of the construction. A renewed insight as this one overcomes the compartmentalized views of the architecture of the language.

Nouns in apposition whose N2 have predicative power are borderline constructions since they support a partitive and a holistic reading. Inflectional fluctuation of N2 illustrates in an optimistic way their hybrid nature.

REFERENCES

- Ackema, P., Neeleman, A. 2005. Word-Formation in Optimality Theory. In: P. Stekauer; R. Lieber, *The Handbook of Word-Formation*. Dordrecht: Springer, 285-313
- Arnaud, P. 2003. *Les noms composés timbre-poste*. Lyon: Presses Universitaires de Lyon.
- Bauer, L. 2001. Compounding. In: M. Haspelmath et al. (Eds.), *Language Typology and Language Universals* vol. I, Berlin: Walter de Gruyter, 695-707.
- Bisetto, A., Scalise, S. 1999. Compounding: morphology and/or syntax? In: L. Mereu (Ed.), *Boundaries of morphology and syntax*. Amsterdam, Philadelphia: John Benjamins, 31-48.
- Bisetto, A., Scalise, S. 2006. The classification of compounds. *Lingue e Linguaggio*. IV (2): 319-332.
- Bosque, I. 1998. *Las Categorías Gramaticales*. Madrid: Editorial Sintesis.
- Brito, A. M. 2003. Categorías sintácticas. In: M. H. M. Mateus et al., *Gramática da Língua Portuguesa*. Lisboa: Caminho, 323-432.
- Cunha, C., Cintra, L. 1998. *Nova gramática do português contemporâneo*. Lisboa: Sá da Costa Editores.
- Di Sciullo, A. M., Williams, E.. 1987. *On the Definition of Word*. Cambridge, MA: The MIT Press.
- Fábregas, A. 2005. *The definition of the grammatical category in a syntactically oriented morphology*. <http://uit.academia.edu/AntonioFábregas/Papers/525958/>
- Fábregas, A. 2006. La relación entre morfología y sintaxis: compuestos de dos sustantivos y nombres de color. *Verba*. 33: 100-122.
- Guevara, E. 2012. *Spanish compounds*. *Probus*. 24: 175-195.
- Guevara, E., Scalise, S. 2009. Searching for universals in compounding. In S. Scalise; E. Magni & A. Bisetto (Eds.), *Universals of Language Today*, Berlin: Springer, 101-128.
- Jackendoff, R. 1990. *Semantic Structures*. Cambridge, MA: The MIT Press.
- Jackendoff, R. 2009. Compounding in the Parallel Architecture and Conceptual Semantics. In: R. Lieber, P. Štekauer (Eds.), *The Oxford Handbook of Compounding*. Oxford: Oxford University Press, 105-129.
- Kornfeld, L. 2003. Compounds N+N as formally lexicalized appositions in Spanish. In: G. Booij, J. De Cesaris, A. Ralli & S. Scalise (Eds.), *Topics in morphology. Selected papers from the Third Mediterranean Morphology Meeting*. Barcelona: Universitat Pompeu Fabra, 211-225.
- Kornfeld, L. 2009. La composición en español. In: R. Lieber; P. Stekauer, *The Oxford Handbook of Compounding*. Oxford: Oxford University Press, 436-452.

- Lieber, R. 1992. *Deconstructing Morphology*. Chicago: University of Chicago Press.
- Lieber, R. 2009. A lexical Semantic Approach to Compounding. In: R. Lieber; P. Štekauer (eds.), *The Oxford Handbook of Compounding*. Oxford: Oxford University Press, 78-104.
- Lieber, R., Scalise, S. Lieber, R. & S. Scalise 2007. The Lexical Integrity Hypothesis in a new theoretical universe. In: G. Boij et al. (Eds.), *On-line Proceedings of the Fifth Mediterranean Morphology Meeting* (Fréjus, 15-18 September 2005). University of Bologna. <http://www.morbo.lingue.unibo-it/> (accessed 24 October 2010).
- Martinho, F. (in this book), Noms épithètes dans les expressions binomiales
- Piera, C., Varela, S. 1999. Relaciones entre morfología y sintaxis. In: Bosque, I., V. Demonte (Eds.) *Gramática Descriptiva de la Lengua Española*. Madrid: Espasa, vol. 3, 4367-4422.
- Rainer, F. (2012), Escarceos sobre la conversión sustantivo → adjetivo en español. In: A. Fábregas; E. Felú; J. Martín & José Pazó (Eds.), *Los límites de la morfología. Estudios ofrecidos a Soledad Varela Ortega*. Madrid: UA Ediciones, 369-382.
- Ribeiro, S. 2010. *Compostos nominais em português. As estruturas VN, NN, NprepN e NA*. München: LINCOM.
- Ribeiro, S. & Rio-Torto, G. (2010), *Denominações compositivas de estrutura VN, NN, NprepN e NA: nexos intralexicaís*. In: Iliescu, M. et al. (Eds.), *Actes du XXV Congrès International de Philologie et de Linguistique Romanes*. Tome VII. Berlin : Mouton de Gruyter, 477-487.
- Rio-Torto, G. 2006. O Léxico: semântica e gramática das unidades lexicais. In Francisca Athayde (coord.), *Estudos sobre léxico e gramática*, 11–34. Coimbra: Cadernos do CIEG 23.
- Rio-Torto, G. & Ribeiro, S. 2012. Portuguese compounds. *Probus*. 24: 119-145.
- Scalise, S. & Bisetto, A. 2009. *La struttura delle parole*. Bologna: Il Mulino.
- Val Alvaro, J. F. 1999. La composición. In: I. Bosque; V. Demonte (Eds.) *Gramática Descriptiva de la Lengua Española*. Madrid: Espasa, vol. 3, 4757-4841
- Villalva, A. 2003. Formação de palavras: composição. In: M. H. M. Mateus et al., *Gramática da Língua Portuguesa*. Lisboa: Caminho, 969–983.
- Villoing, F. 2012. French compounds. *Probus*. 24: 29-60.
- Wierzbicka, A. 1986. What' s in a noun? (or: how do nouns differ in meaning from adjectives?). *Studies in Language*. 10 (2): 353-389.



# Noms épithètes dans les expressions binominales \*

Fernando Martinho

fmart@ua.pt

*Departamento de Línguas e Culturas - Universidade Aveiro*

**ABSTRACT.** Adjectives have the exclusive property of being attributive nominal modifiers (or ‘epithets’). This exclusivity allows them to integrate the DP as modifiers of the lexical head N, while sharing, in Romance, its inflectional features. Some conversion models, however, allow the projection of a DP with two nominal heads in a modifying relation. Departing from previous proposals in Portuguese (Rio Torto 1998, Villalva 2003, Ranchhod 2003), our analysis implies that some N1N2 endocentric sequences, like «*criança modelo*» (*child model*) or “*império fantasma*” (*empire ghost*), in fact correspond to a modified structure, where N2 is an “attributive noun” (Martinho 2007). In these N1N2 “binominal modified constructions”, usually classified as cases of “adjunction”, we will consider, based on a corpus of cases from the CETEM database, that N2 appear to be a mixed category. Since N2 checks some basic adjectival properties (eg., N2 accepts multiple modification, predicative position and typical adjectival prefixation), we argue that this mixed category is equivalent to an adjectival modifier. Moreover, since N2 is only partially meeting these same properties (eg., N2 has no degree), we conclude that, in this kind of modification common to all Romance languages (Fabre 1996, Villoing 2002, Savary 2004, Montermini 2006), the N2 nominal epithet is converted into a relational adjective (Demonte 1999).

**KEY-WORDS.** Syntax, Lexicon, adjectival modifiers, attributive adjectives, relational adjectives, mixed categories, conversion, composition.

**RÉSUMÉ.** Les adjectifs ont la propriété exclusive d’être des modifieurs attributifs (ou épithètes) du nom. Cette exclusivité leur permet d’intégrer le DP en tant que modifieurs de la tête lexicale N, tout en partageant ses traits flexionnels. Certains modèles de conversion permettent cependant la projection d’un DP pourvu de deux têtes nominales en relation de modification. Se dissociant de propositions précédentes en portugais (Rio Torto 1998, Villalva 2003, Ranchhod 2003), notre analyse implique que certaines séquences endocentriques N1N2, comme « *criança modelo* » ou « *império fantasma* », correspondent en fait à une structure modifiée, dans laquelle N2 serait un « nom attributif » (Martinho 2007).

---

\* Une version préliminaire de ce travail a été présentée lors du *Workshop on Conversion and Mixed Categories* (Faculté des Lettres de Porto, 2 et 3 novembre 2012). Je tiens à remercier le public du séminaire pour sa réceptivité et ses suggestions. Je remercie aussi les deux réviseurs anonymes de ce texte, dont les remarques et les commentaires ont grandement contribué à en corriger les défauts essentiels et à en assurer la pertinence scientifique.

Dans ces « constructions binominales modifiées » N1N2, habituellement classées comme des cas d' « adjonction », nous considérerons, en nous fondant sur un corpus constitué d'extraits de la base de données informatisée CETEM, que N2 correspond à une catégorie mixte. Étant donné que N2 vérifie certaines propriétés de base des adjectifs (par exemple, N2 accepte la modification multiple, la position prédicative et la préfixation adjectivale), nous préciserons que cette catégorie mixte équivaut à un modifieur adjectival. Par ailleurs, puisque N2 ne satisfait que partiellement à ces mêmes propriétés (par exemple, N2 n'a pas de degré), nous concluons que, dans ce processus de modification commun à toutes les langues romanes (Fabre 1996, Villoing 2002, Savary 2004, Montermini 2006), l'épithète nominale N2 est convertie en adjectif relationnel (Demonte 1999).

MOTS-CLÉ. Syntaxe, lexique, modification adjectivale, adjectifs attributifs, adjectifs relationnels, catégories mixtes, conversion, composition.

## Introduction

Si l'idée de l'existence de classes de mots fondées sur leur signification est aussi vieille que la grammaire, l'étude de la diversité des langues nous force à admettre que, pour bien des items lexicaux, il y a conflit entre nature et usage. Nous nous proposons ici de réfléchir à certains items de la classe des noms utilisés comme modifieurs adjectivaux en portugais, en partant d'un corpus informatisé annoté, le *CETEM Público*. En tant que modifieurs prototypiques du nom (attributifs ou épithètes), les adjectifs sont classiquement, dans le DP (Determiner Phrase), des unités lexicales superficiellement adjacentes d'une tête lexicale N, partageant avec cette tête ses traits flexionnels et modifiant ses traits sémantiques. Cependant, il existe, dans la plupart des langues, et aussi en portugais, de nombreux cas (plus d'une centaine dans la base *CETEM*) résultant apparemment de la rencontre entre deux N, pour lesquels, à la différence de ce qui se produit dans les cas de composition, l'un des deux N est réinterprété comme épithète de l'autre.<sup>1</sup>

La piste que nous proposons ici concerne un sous-ensemble de séquences binominales N + N, en marge des processus de composition classiques

<sup>1</sup> Rappelons qu' 'épithète' dérive du grec *epithetos*, « ajouté ». Les ouvrages de référence associent habituellement la fonction épithète à un ensemble de distinctions lexico-sémantiques, dont l'une (« les épithètes de nature ») semble relativement pertinente: ainsi, dans « un ciel immense », *immense* serait une épithète de nature, exprimant un caractère intrinsèque et durable (par opposition aux « épithètes de circonstance »). Cette interprétation typique et générique attribuée par la langue classique à certaines épithètes est à rapprocher de notre analyse prototypique des séquences N<sub>1</sub>N<sub>2</sub> ci-dessous. Cf. *Dictionnaire de l'Académie française*, <http://www.academie-francaise.fr/>.

(qui, eux, sont de fréquence très élevée), qui correspondent en fait à une structure modifiée de type N + A, dans laquelle le deuxième N (ou N<sub>2</sub>) est une catégorie mixte, identifiée ici, faute de mieux, comme « nom épithète ». Dans ces « constructions binominales », nous verrons, pour certaines d'entre elles, (« les constructions binominales modifiées »),<sup>2</sup> que N<sub>2</sub> est recyclé en catégorie adjectivale et résulte sans doute d'une conversion catégorielle du trait [+V].

## 1. Les expressions binominales et les procédés de composition

Nous commencerons par présenter le corpus informatisé consulté et reprendre sommairement quelques analyses préalables des expressions binominales.

### 1.1. Corpus de travail

Définissons le corpus de travail, qui illustre le type de séquences dont il va être question. Les 112 exemples suivants en (1.a) et (1.b) –divisés ici en deux groupe en fonction de leur graphie–, correspondent tous à des expressions binominales NN, et résultent de concordances extraites de la base de données CETEM<sup>3</sup>, une fois expurgées des faux positifs :

(1.)Liste de concordances N<sub>1</sub>N<sub>2</sub> (*syntaxe de recherche: [pos="N»]*  
*[pos="N»]*)

a. Expressions N<sub>1</sub>N<sub>2</sub> avec trait d'union (*orthographe originale*):

Projecto-piloto, contrato-promessa, contrato-programa, projecto-lei, medida-padrão, clube-satélite, estado-membro, episódio-chave, tema-tabu, saltada-relâmpago, vencimento-base, homem-aranha, saco-cama, governo-sombra, palavra-chave, ministro-presidente, posto-chave, versão-base, estado-fantoches, lei-quadro, trabalhador-estudante, notícia-acusação, linha-mestra, queixa-crime, edifício-montra, biografia-testemunho, cidade-dormitório, ano-luz, carta-repto, futebol-espectáculo, jogo-treino, estado-nação, negócio-

<sup>2</sup> Cf. « ATAP (attributive / appositive) compounds », Scalise & Bisetto (2009).

<sup>3</sup> Adresse Internet de la base CETEM: <http://www.linguateca.pt/cetempublico/>

base, empresa-fantasma, salário-base, programa-base, cão-polícia, público-alvo, material-base, período-chave, região-piloto, carro-patrolha, ordenado-base, estado-providência, retrato-tipo, edifício-sede, objecto-fetichê, data-limite, valor-limite, convidado-mistério, processo-crime, efeito-surpresa, prazo-limite, programa-quadro, mulher-polícia, jogo-fantasma, homem-golo, taxa-base, ideia-choque, plano-sequência, jogo-miragem, café-concerto, bloco-notas, poupança-reforma, caminhão-cisterna, estúdio-auditório, cidade-luz, equipamento-base, engenheiro-chefe, quadro-legenda, zona-tampão, retrato-robot, objecto-fetichê

b. Expressions  $N_1N_2$  sans trait d'union (*orthographe originale*):

Empresa propriedade, argumento base, espelho retrovisor, linha mestra, zona sul, andar modelo, zona centro, sector automóvel, versão cliente, cinema paixão, medicamento anti-acidez, satélite espião, detecção radar, avião pirata, circulação automóvel, déficit recorde, sala estúdio, tecnologia laser, império fantasma, verão assassino, estado membro, emissor pirata, equipa sensação, cópia vídeo, andar modelo, equipa maravilha, tradutor intérprete, vigilância vídeo, plano anti-crime, disciplina base, caixa multibanco, divisão anti-sequestro, rainha mãe, momento chave, país anfitrião, encontro teste, protesto anti-aborto, caixa serviço, programa tipo, vencimento base, polícia anti-motim, fabricante líder, mulher empresária

1.2. Modèles de composition  $N_1N_2$

Dans le domaine des études morphosyntaxiques, différents auteurs ont abordé, dans le domaine du portugais, les structures endocentriques  $N_1N_2$  et les ont considérées, généralement, comme des cas de composition.

Commençons par ce que Villalva (2003) qualifie, au sens large, de « compositions morphosyntaxiques ». Celles-ci (qui englobent entre autres la structure  $N_1N_2$ ) correspondent bien à des catégories lexicales et se divisent, selon l'auteur, en deux grandes familles, les structures d'adjonction en (2.) – tête nominale typique composée de deux éléments, dont le 1<sup>er</sup> est un nom

et le 2<sup>ème</sup> un « modifieur nominal »<sup>4</sup> qui ne fléchit pas –, et les structures de conjonction en (3.) – dans lesquelles les constituants partagent leur flexion mais sont opaques à la réanalyse lexicale, et dont on ne peut identifier avec précision la tête de la structure.

(2.) *Composition morphosyntaxique - structures d'adjonction*<sup>5</sup>

Homem-aranha, governo-sombra, peixe-espada

(3.) *Composition morphosyntaxique - structures de conjonction*

Autor-compositor, conta-poupança, café-bar-restaurant

Une autre analyse de  $N_1N_2$  est celle de Rio-Torto (1998), qui les définit comme des unités composées de « deux unités lexicales dotées de capacité référentielle » et en distingue, en conséquence, deux types : (i) la composition coordinative, en (4.), qui permet de rendre compte naturellement de la lecture de dénotation par intersection typique de ces composés (un *compositeur interprète* est à la fois l'un et l'autre) ; (ii) la composition subordonative, en (5.), qui, elle, n'implique pas de dénotation intersective, et repose plutôt sur des modèles dérivationnels du type « Déterminé - déterminant », dans lequel l'élément déterminant est, dans certains cas, un nom.

(4.) *Composition coordinative*<sup>6</sup>

Porco-espinho, compositor-intérprete

(5.) *Composition subordonative*

Turma-piloto, visita-relâmpago, casa-mãe

Selon Ranchhod 2003, enfin, les « noms composés »  $N_1N_2$  seraient un cas particulier des « expressions figées hétérogènes », illustrées en (6.) formées à partir de mots simples au moyen de règles combinatoires préalables. Ces composés occupent les positions syntaxiques typiques des noms et ont, en principe, une interprétation compositionnelle. Leur forme graphique

---

<sup>4</sup> L'expression originale est "modificador nominal". Cf. Villalva (2003 :980).

<sup>5</sup> Exemples de Villalva (2003).

<sup>6</sup> Exemples de Rio-Torto (1998).

habituelle avec trait d'union semble provenir, d'après l'auteur, du fait que la lexicologie les considère comme des composés, et les traite donc comme des entrées normales du dictionnaire.

(6.) Noms composés<sup>7</sup>

braço-de-ferro, cara-metade, raio laser, café-concerto

Les auteurs cités classent indifféremment les séquences NN précédentes comme des compositions, bien que de forme assez diverse. Parmi les tests proposés pour détecter les propriétés de ces « noms composés », les auteurs suggèrent qu'on peut s'intéresser à leur morphologie interne, ainsi qu'à leur compositionnalité (par exemple, leur dénotation supporte ou non une réduction par dissociation du composé). Il est aussi noté que certaines de ces compositions résistent plutôt bien à des opérations de stress lexical, comme la suppression néologique d'un segment syllabique intermédiaire ou le *blending*.<sup>8</sup> Par ailleurs, comme le note Villalva, la flexion de N<sub>2</sub> dans les structures d'adjonction semble être assez inconstante.

Il nous semble que les analyses précédentes, toutes pertinentes dans leur volonté de classer ces entrées lexicales ambiguës, soulèvent cependant plusieurs difficultés, plutôt, pensons nous, en termes de frontière entre lexique et syntaxe qu'entre composition lexicale ou morphosyntaxique. Ces analyses semblent aussi associer de façon transversale des cas de nature et d'interprétation très différentes et ne donnent finalement qu'une faible attention aux séquences N<sub>1</sub>N<sub>2</sub>, qui seules nous intéressent<sup>9</sup>, face aux autres types de composés (V-N, A-N...).<sup>10</sup>

<sup>7</sup> Exemples de Ranchhod (2003)

<sup>8</sup> Rio-Torto (1998) précise ainsi que les composés analysés se caractérisent par divers procédés "réducteurs, abstracter, supprimeurs et abrégiateurs" et que leur résistance à ces procédés doit permettre de les classer objectivement. Les composés subordinatifs, par exemple, sont sensibles à la suppression d'un segment médial (« televoto, petrodólar, moto-serra, cineclub, aparthotel »); les composés coordinatifs, eux, réagissent plutôt au *blending* (« diciopédia, cantriz »). Ces exemples, observe l'auteur, correspondent souvent à des « modes éphémères », dont certains deviennent éventuellement des néologismes.

<sup>9</sup> Les expressions binominales N<sub>1</sub>N<sub>2</sub> semblent bien être le « parent pauvre » des analyses citées, probablement pour deux raisons différentes: (i) il s'agit de séquences effectivement rares et éphémères, vis-à-vis des autres types; (ii) leur statut grammatical de « composés » est plus délicat à établir que celui de séquences reposant sur des catégories lexicales distinctes (N-A, V-N, etc.).

<sup>10</sup> Des travaux plus récents existent dans le domaine de la composition. Citons ceux de Ribeiro (2010) et de Rio-Torto & Ribeiro (2012) (je remercie l'un des réviseurs pour ces références, que je n'ai malheureusement pu consulter à temps).

Notre position consiste à considérer que les expressions binaires  $N_1N_2$ , quelle que soit la catégorie grammaticale à laquelle elles seraient apparentées, correspondent à un ensemble de structures hétérogènes et fragiles, dont une partie n'a en fait rien à voir, pensons-nous, avec la composition. Nous allons proposer, en effet, qu'une expression comme *visita-relâmpago* (*visite éclair*), contrairement à *porco-espinho* (*porc-épic*), par exemple, n'est pas le résultat d'un processus de composition, mais bien une expression nominale modifiée, dans laquelle  $N_2$  serait, donc, une épithète.

## 2. Expressions binominales

Les expressions de type  $N_1N_2$  représentent une caractéristique originale de la langue véhiculaire moderne, ainsi qu'une source d'enrichissement et de création néologique de son vocabulaire, et sont, de plus, relativement communes, car elles permettent une extension du lexique fondée sur des lexèmes déjà acquis et des opérations préexistantes. Du point de vue formel, la rencontre entre deux noms dans une relation d'intersection dénotative crée un terme spécifique, fondé principalement sur des relations de sémantique lexicale –en particulier l'hyponymie–, et caractérise en cela la façon dont le lexique reflète l'organisation conceptuelle de certains domaines du savoir. Les publications scientifiques contiennent en effet, comme le souligne Fabre (1996) –dans ce cas, en anglais–, un nombre croissant de ces expressions composées néologiques (comme *noun phrase*, *light year* ou *milk disease*).

Nous allons distinguer en portugais, dans les structures  $N_1N_2$  extraites de la base de données en ligne *CETEM*, entre 'expressions binominales composées', qui concernent effectivement les cas de composition, et que nous aborderons rapidement, et 'expressions binominales modifiées', comparativement beaucoup plus rares, qui seules ici nous intéressent.<sup>11</sup> Nous n'attribuerons aucune valeur particulière à leur graphie, qui, étant donné la relative nouveauté de ces formes, est assez inconstante.

---

<sup>11</sup> La distinction entre ces deux catégories d'expressions n'est pas toujours évidente, ainsi que l'ont fait remarquer deux réviseurs anonymes. Ce travail est, de fait, une tentative de définition et de distinction fondée sur un certain nombre de critères formels, appliqués à des données linguistiques par ailleurs assez flottantes et fragiles. Voir aussi les notes 14 et 19.

## 2.1. Expressions binominales composées (EBC)

Les expressions binominales composées impliquent que, à côté de  $N_1$ ,  $N_2$  est bien un nom et que nous sommes donc devant des cas de compositions comme ceux définis initialement (voir paragraphe 1.2). Il serait, de même, possible de distinguer entre binominaux composés intersectifs, en (7.) (pour lesquels l'adjacence entre deux N équivaut, par intersection, à une nouvelle expression référentielle, une coordination à considérer sur le plan sémantique comme hyponyme de  $N_1$ ) et binominaux composés argumentaux, en (10.) (dans ce cas, le prédicat  $N_1$  possède une grille argumentale et attribue un rôle thématique à  $N_2$ ). Voyons les EBC intersectives en (7.):

### (7.) EBC intersectives (*source: CETEM*)

trabalhador-estudante, saco-cama, ministro-presidente, cidade-dormitório, edifício-sede, cão-pastor, cardeal-patriarca, cine-estúdio, actor-encenador, emissor-receptor, sobrinho-neto, juiz-conselheiro, artesão-empresário, casa-museu, cirurgião-dentista, cantora-compositora, director-tesoureiro, médico-veterinário, sapadores bombeiros, cidadãos eleitores, estado membro, peixe espada, área escola, jantar recepção, sala estúdio, espelho retrovisor, opinião destaque, macaco vampiro, curso formação, mulheres padres, salas-gabinete

Observons que, provenant d'une situation d'intersection dénotative, ces cas devraient accepter différents tests morphologiques (comme la pluralisation, en (8.)) et syntaxiques (comme l'inversion  $N_2N_1$ , en (9.)), ce qui se produit effectivement, malgré le poids de leur lexicalisation :

### (8.) EBC intersectives: flexion du pluriel

juízes-conselheiros, artesãos-empresários, cirurgiões-dentistas, cantoras-compositoras, directores-tesoureiros, médicos-veterinários

### (9.) EBC intersectives: inversion $N_2N_1$

- a. trabalhador-estudante / estudante-trabalhador
- b. mulher padre / padre mulher
- c. poupança-reforma / ? reforma-poupança
- d. contrato-promessa / ? promessa-contrato



- e. sapadores bombeiros / ? bombeiros sapadores
- f. mulher-polícia / ? polícia-mulher

Quant aux composés argumentaux, ou subordonnés,  $N_1N_2$ , ils sont relativement rares. On peut cependant trouver des exemples dans lesquels  $N_2$  est interprété comme Thème ou But de  $N_1$  (10.). A noter que, dans ce cas, (voir (11.)), l'inversion  $N_2N_1$  est totalement bloquée, puisque si  $N_2$  est l'argument de  $N_1$ , il ne peut quitter sa position thématique.<sup>12</sup>

(10.) EBC argumentales (source: CETEM)

medicamento anti-acidez, cópia vídeo, filme catástrofe, plano anti-droga, desfile pró-aborto, jogo vídeo, propaganda anti-propinas, viagens intra-empresas, análise custo-benefício, queixa crime, processo-crime<sup>13</sup>

(11.) EBC argumentales: inversion  $N_2N_1$

a. cópia vídeo / \* vídeo cópia,

b. medicamento anti-acidez/ \*anti-acidez medicamento

## 2.2. Expressions binominales modifiées (EBM)

### 2.2.1. Modifieurs métaphoriques

Intéressons-nous maintenant aux binominaux modifiés. Du point de vue sémantique, ces expressions reposent sur des procédés dénotatifs stéréotypés, à la limite du connotatif et de la norme grammaticale, dans lesquels le second terme ( $N_2$ ) sert souvent de métaphore (*governo-sombra*) ou de métonymie (*lado lâmina*) pour qualifier le premier ( $N_1$ ). Dans ces expressions,  $N_2$  passe par un processus d'allègement sémantique, qui isole ou extrait un trait spécifique de son concept.

Ainsi, une expression comme *avião pirata* (*avion pirate*) est-elle une expression référentielle dans laquelle l'extension du terme « avion » est

<sup>12</sup> Dans ces cas,  $N_2$  se caractérise parfois par une préfixation (« anti, intra, pro ») qui équivaut, sur le plan syntaxique, à une construction prépositionnelle.

<sup>13</sup> Comme l'a fait remarquer un réviseur, certains exemples de EBC intersectives pourraient être réinterprétés comme des structures prédicatives, comme, par exemple, « poupança-reforma » (*économie-retraite*), à condition toutefois de considérer  $N_2$  comme le But de  $N_1$ . Notons, de plus, que certaines expressions, comme « filme catástrofe », pourraient, selon la lecture qu'on leur attribue, osciller entre le statut de EBC et de EBM : dans le 1<sup>er</sup> cas, on parlerait d'une catastrophe (par exemple naturelle) qui est l'objet d'un film ; dans le 2<sup>ème</sup> cas (plus naturel), on parlerait d'un film mettant en scène les conséquences dramatiques d'une catastrophe. L'acceptation naturelle de « filme catástrofe » en fait donc une EBM.

manipulée et réinterprétée par celle du second terme, qui contribue à en identifier le référent. L'absence éventuelle de  $N_2$  n'affecte pas directement la dénotation de  $N_1$  (*un avion pirate est un avion*).

Recourir à la modification métaphorique et analogique se traduit sur le plan structurel par une situation d'apparente adjacence entre deux têtes nominales. Il s'agit ici, évidemment, d'expressions dénotatives qui ne proviennent pas d'une opération compositionnelle sur des entités, mais bien d'un procédé par lequel le second terme fonctionne comme modifieur et non comme expression référentielle. C'est la nature fossilisée d'une partie de ces expressions qui explique que certaines, comme *governo-sombra* (*gouvernement ombre*), acquièrent une stabilité lexicale suffisante pour devenir de nouvelles lexies, figurant dans le dictionnaire à côté d'EBC comme *processo-crime* (*procès-crime*).

### 2.2.2. Analyse des EBM

Notre première hypothèse est donc que  $N_2$ , dans les expressions binominales modifiées  $N_1N_2$ , correspond à une épithète. Il s'agit bien d'une relation de type nom-modifieur, à la différence des EBC, et, à ce titre, sont exclus tous les cas de dénotation intersective ou de composition argumentale, comme ceux vus en (7.) et en (10.). Voici en (12.) une liste étendue de EBM, extraite des concordances de la base CETEM :

#### (12.) EBM – $N_1N_2$ (source : CETEM)

projecto-piloto, medida-padrão, clube-satélite, zona-tampão, retrato-robot, linha mestra, zona sul, andar modelo, zona centro, cinema paixão, défice recorde, tecnologia laser, império fantasma, emissor pirata, negócio-base, valor-guia, desporto-rei, região-piloto, frase-chave, rali-maratona, prova vedeta, astro-rei, cidade-berço, cidade martírio, factor novidade, mulher paixão, homem orquestra, cidade património, fator risco, grupo piloto, factor casa, documento síntese, programa candidato, torre montanha, músico anfitrião, efeito porto, organização fantoche, projecto síntese, data limite, fator tempo, papéis vedetas, circuito satélite, pai herói, linhas mestras, elementos chave, questões base, lado lâmina,

discos piratas, equipa maravilha.<sup>14</sup>

Ces EBM sont, dans leur majorité, purement fortuites, c'est-à-dire forgées de toutes pièces selon les besoins inhérents à une situation de communication : à part quelques expressions institutionnalisées par l'usage (*cidade-berço* – *ville berceau*), il s'agit de formes lexicalement fragiles, qui prolifèrent probablement sur la base d'anglicismes, causant un impact sur l'auditeur par la substitution inattendue d'un adjectif par un nom. S'il ne s'agit pas de lexies stables, elles ne sont pas non plus opaques, et ont en commun le fait d'être des entrées récentes dans les dictionnaires – lorsque cela se produit. Elles ne proviennent pas non plus, contrairement aux EBC, d'une opération de juxtaposition ou de concaténation. Une fois institutionnalisée, l'EBM devient un item lexical familier, plus tard pétrifié en lexème (*desporto-rei*, *sport-roi*).

En tant que constructions de modification, ces EBM correspondent à la projection de têtes lexicales et à des opérations d'édification de structure syntaxique, et proviennent de principes d'organisation endocentriques dans lesquels le NP dérivé (en portugais) doit être considéré comme de tête initiale. Notre description de ces  $N_1N_2$  part donc de l'hypothèse qu'il s'agit de structures syntaxiques dérivées sur la base des catégories lexicales N et A. Comme  $N_2$  n'est visiblement pas un A mais un N, nous considérerons cette 'catégorie mixte' comme un cas de conversion catégorielle.

Notons qu'un ensemble ouvert de noms prototypiques<sup>15</sup> (voir (13.)), capables de dénoter ou connoter des modèles consensuels (« modelo, padrão, tipo »), des termes de comparaison à valeur analogique (« base, vedeta, chave »), des unités de mesure non discrètes (« recorde, relâmpago »), des approximations sémantiques (« pirata, maravilha »), etc., sont utilisés systématiquement comme  $N_2$  dans les EBM :

---

<sup>14</sup> Voici le contexte original de trois de ces EBM, dont l'interprétation est particulièrement ambiguë: (1) "Sempre com a Pilar, minha *mulher paixão*."; (2) "Debaixo da «*sikhara*», a *torre montanha*, está o santuário, escondido como se estivesse numa gruta."; (3) "No primeiro caso, o *efeito porto* não existe, mas no segundo é por demais importante." (Source: CETEM). Voir note 21 pour l'EBM "cinema *paixão*".

<sup>15</sup> Nous utilisons ici le concept de « prototype » au sens de E. Rosch, qui le définit comme le membre le plus central ou le plus représentatif d'une catégorie, fonctionnant comme un point de référence cognitif. Cf. Kleiber (1990)

(13.) N<sub>2</sub> prototypiques (*source: CETEM*)

base, modelo, tipo, padrão, limite, mestre, piloto, chave, líder, pirata, fantasma, relâmpago, mãe, satélite, vedeta, recorde, maratona, maravilha, casa, fanteche ...

Les N<sub>2</sub> prototypiques en (13.) sont de fréquence très inégale. Voici dans le tableau ((14.), ci-dessous, les fréquences pour chaque N2, classées par ordre décroissant :<sup>16</sup>

(14.) Fréquence des N<sub>2</sub> prototypiques (*source: CETEM*)

N2	Fréq. totale	Fréq. pluriel
base	1514	6
recorde	1315	55
limite	1108	45
tipo	1108	0
chave	626	26
líder	452	65
piloto	425	23
fantasma	376	82
modelo	365	10
satélite	331	79
pirata	296	157
mãe	249	5
casa	172	0
relâmpago	156	2
vedeta	108	38
padrão	105	1
mestre	73	4
fanteche	72	15
síntese	59	0
maratona	26	0
maravilha	13	0
TOTAL	8777	613(7%)

<sup>16</sup> Exemple de syntaxe de recherche pour N<sub>2</sub> = « base » : [pos = »N«] lema = [«base»]

<sup>17</sup> Le N<sub>2</sub> « casa » n'a été détecté qu'en association avec le N<sub>1</sub> « factor », produisant ainsi l'EBM « factor casa »

Ces termes semblent, du fait de leur interprétation métaphorique, avoir perdu leur capacité référentielle, tout en affectant, en retour, la capacité référentielle de  $N_1$ , assumant en conséquence une fonction de modificateurs génériques à faible degré de représentativité. Cet allègement sémantique des  $N_2$  se manifeste par des propriétés typiques (« base, tipo »), idéalisées (« modelo, síntese ») et éventuellement connotatives (« pirata, fantoche »), propriétés qu'il attribue alors à  $N_1$ . Il n'est plus question d'entités référentielles qui, par intersection, manipulent principalement la dénotation de  $N_1$ , mais bien d'un procédé dans lequel  $N_2$  manifeste des propriétés reproductibles de modificateur de  $N_1$ . Nous pouvons donc définir la classe des  $N_2$  prototypiques sur la base d'une combinaison de différents traits sémantiques rendant compte de ces observations, parmi lesquels pourraient figurer les traits [+générique] et [-référentiel].<sup>18</sup>

On peut par ailleurs prévoir que le procédé ici analysé n'est en aucune façon épuisé par les cas répertoriés, et peut, au contraire de ce qui se produit dans le domaine de la composition, donner de nouvelles formes originales et aléatoires, dont on peut varier les combinaisons à l'infini. Ainsi, bien que non identifiées dans la base de données CETEM, il serait tout à fait possible de détecter des formes comme les suivantes (en 15.):

(15.) EBM –  $N_1N_2$  (*non présentes dans la base CETEM*)

- a. homem chave, batalha chave, região chave, zona chave
- b. clube modelo, projecto modelo, região modelo, cidade modelo
- c. clube piloto, região piloto, cidade piloto, família fonte
- d. discurso marco, pena exemplo, decisão limite

La possibilité de combiner librement  $N_1$  avec une gamme limitée de  $N_2$  détermine donc une classe ouverte d'expressions, dont le procédé de notation orthographique cesse d'ailleurs d'être rigide, ce qui renforce d'autant l'autonomie relative des éléments combinés. Contrairement aux cas

---

<sup>18</sup> Si l'existence de  $N_2$  prototypiques comme « pirata, limite, fantasma, maravilha » est pacifique, en revanche, le fait de considérer de même d'autres  $N_2$ , comme « novidade, património, montanha, anfitrião », semble plus discutable. Ces termes obéissent pourtant au même principe qui identifie tous les  $N_2$  prototypiques: à cause de (grâce à) leur lecture métaphorique, ils ont bel et bien été allégés de leur capacité référentielle et il ne leur reste plus qu'à assumer une fonction prototypique d'épithètes [+génériques] et [-référentielles]. Ainsi, une « torre montanha » (*tour montagne*) est-elle une tour qui possède les attributs d'une montagne ; une « cidade património » (*ville patrimoine*), une ville d'une grande richesse historique et architecturale, etc.

(7.), dont la nature lexicale de composés binominaux laisse peu de doute, nous sommes donc avec (12.) ou (15.), devant des structures syntaxiques originales, qui ne sont que très peu lexicalisées. Ces observations renforcent notre proposition de considérer  $N_2$  comme l'équivalent d'un adjectif.  $N_2$  émerge, dans ces EBM, en position d'épithète. Nous avons, cependant, à déterminer de quel type d'adjectif il s'agit, dans la mesure où la nature de  $N_2$  affecte probablement sa sélection par  $N_1$ .

### 2.3. Quelques tests sur les EBM

En partant des critères définis par Bhat (1994) pour identifier les adjectifs, évaluons la nature des  $N_2$  présents dans ces EBM. Notre intention est de présenter, à la suite, dix tests tendant à suggérer que  $N_2$  vérifie les propriétés typiques des adjectifs relationnels.

Pour commencer, notons que  $N_2$  diffère singulièrement des adjectifs pour trois raisons essentielles : (i) les consultations de la base CETEM indiquent que  $N_2$  épithète ne fléchit pas<sup>19</sup> ou très peu (16.); (ii)  $N_2$  n'accepte pas d'intégrer la tête des petites propositions ou 'small clauses' (17.); (iii)  $N_2$  n'accepte pas de variation de degré, et n'est donc pas graduable (18.):

- (16.) EBM –  $N_1N_2$  Flexion de  $N_2$
- |  |                             |
|--|-----------------------------|
| a. as emissões <b>pirata?(s)</b> proibidas         | ( <i>emissão pirata</i> )   |
| b. as empresas <b>fantasma?(s)</b> investigadas    | ( <i>empresa fantasma</i> ) |
| c. visitámos ontem os andares <b>modelo?(s)</b>    | ( <i>andar modelo</i> )     |
| d. ministérios anunciam défices <b>recorde?(s)</b> | ( <i>défiçe recorde</i> )   |
| e. mau tempo sobre as zonas <b>centro*(s)</b>      | ( <i>zona centro</i> )      |

- (17.) EBM –  $N_1N_2 N_2$  Tête de "petite proposition" (small clause)
- |   |                            |
|---|----------------------------|
| a. Quero estas emissões <b>proibidas</b> imediatamente    |                            |
| a'. * Quero estas emissões <b>pirata(s)</b> imediatamente | ( <i>emissão pirata</i> )  |
| b. Os Americanos consideram a cidade <b>deserta</b>       |                            |
| b'. * Os Americanos consideram a cidade <b>martírio</b>   | ( <i>cidade martírio</i> ) |

<sup>19</sup> La flexion se manifeste avec beaucoup d'indétermination: l'apparition de formes fléchies est très minoritaire, en moyenne 10 à 15% selon les EBM. Voir le tableau (14.). Pour une analyse plus approfondie des problèmes de flexion de  $N_2$ , cf. Rio-Torto & Ribeiro (2012).

- (18.) EBM –  $N_1N_2$                       Degré de  $N_2$
- a. \* Esta prova é mais **vedeta** que a anterior                      (*prova vedeta*)
- b. \* Guimarães é a cidade mais **berço** de Portugal                      (*cidade-berço*)
- c. \* Este acontecimento não é tão **limite** quanto aquele (*acontecimento limite*)

Ces trois tests, indispensables pour déterminer la nature adjectivale d'une unité lexicale, montrent donc que nous ne sommes pas devant des épithètes adjectivales ordinaires. Le fait que, dans les bases de données, la majorité des cas impliquent des formes non fléchies n'est pas ici dépourvu d'intérêt, quoique non décisif. Cependant, ces tests sont insuffisants pour écarter la catégorie adjectivale, puisque les adjectifs relationnels ne passent pas, eux non plus, les tests (17.).

En second lieu, d'autres tests appliqués aux EBM montrent que  $N_2$  peut être remplacé par un adjectif explicite (19.) et qu'il ne peut avoir de position pré- $N_1$  (20.) (exigeant donc la position post- $N_1$ ).  $N_2$  correspond donc, en portugais, à une position exclusivement post-nominale :

- (19.) EBM –  $N_1N_2 \rightarrow N_1 A$
- a. projecto-piloto —> projecto inovador                      (*projecto-piloto*)
- b. clube-satélite —> clube dependente                      (*clube-satélite*)
- c. engenheiro-chefe —> engenheiro principal                      (*engenheiro-chefe*)
- d. emissor pirata —> emissor clandestino                      (*emissor pirata*)

- (20.) EBM –  $N_1N_2 \rightarrow N_2N_1$
- a. \*robot-retrato, \*cliente versão, \*base-negócio, \*satélite-clube
- b. \*fetiche-objecto, \*fantasma império, \*laser tecnologia

En troisième lieu, les EBM acceptent la modification attributive classique pré ou post nominale (21.), de même qu'ils acceptent la modification relationnelle ( $A_{rel}$ ) post-nominale (22.):

- (21.) EBM – (A) N<sub>1</sub>N<sub>2</sub> (A) (A: *modifieurs attributifs*)  
 a. clube-satélite **pequeno** (*clube-satélite*)  
 b. **estranho** objecto-fetichê (*objecto-fetichê*)  
 c. retrato-robot **fiel** (*retrato-robot*)  
 d. **novo** andar modelo (*andar modelo*)  
 e. **arrepiaante** império fantasma (*império fantasma*)
- (22.) EBM – N<sub>1</sub>N<sub>2</sub> (A<sub>rel</sub>) (A<sub>rel</sub>: *adjectif relationnel*)  
 a. clube-satélite **desportivo** (*clube-satélite*)  
 b. império fantasma **galáctico** (*império fantasma*)  
 c. retrato-robot **policial** (*retrato-robot*)  
 d. banda-culto **japonesa** (*banda-culto*)

En quatrième lieu, le couple N<sub>1</sub>N<sub>2</sub> ne peut pas être séparé *via* modification attributive (c.a.d. insertion de A entre N<sub>1</sub> et N<sub>2</sub> – voir (23.)), sauf si A est un adjectif relationnel, qui lui peut séparer N<sub>1</sub> et N<sub>2</sub> (voir (24.)):

- (23.) EBM – N<sub>1</sub> (A) N<sub>2</sub> / (A) N<sub>1</sub>N<sub>2</sub> (A) (A: *modifieur attributif*)  
 a. \*projecto **novo** piloto/ **novo** projecto piloto  
 (*projecto piloto*)  
 b. \*emissão **habitual** pirata / **habitual** emissão pirata  
 (*emissão pirata*)  
 c. \*cinema **comovente** paixão / cinema paixão **comovente**  
 (*cinema paixão*)<sup>20</sup>  
 d. \*zona **estreita** tampão / **estreita** zona-tampão.  
 (*zona-tampão*)
- (24.) EBM – N<sub>1</sub> (A<sub>rel</sub>) N<sub>2</sub> (A<sub>rel</sub>: *adjectif relationnel*)  
 a. clube **desportivo** satélite (*clube satélite*)  
 b. império **galáctico** fantasma (*império fantasma*)  
 c. habitação **social** modelo (*habitação modelo*)

<sup>20</sup> Signalons, au sujet de “cinema paixão”, que l’interprétation attribuée à cette EBM en portugais est “cinéma qui est comme une passion”, ou « celui qui a la passion du cinéma », et s’applique donc métaphoriquement aux amateurs du 7<sup>ème</sup> art. L’ambiguïté de l’interprétation est probablement due à la nouveauté de l’expression.



Si l'on parcourt les cas précédents, on note que les adjectifs relationnels n'affectent pas, par leur présence, l'intégrité de l'EBM originale. Plus précisément, un relationnel est capable de s'insérer entre  $N_1$  et  $N_2$ , ce qui est impossible pour un modifieur attributif (23.).

En cinquième lieu,  $N_2$  ne semble pas compatible avec une position prédicative avec copule, bien que  $N_2$  puisse, dans certains cas, être paraphrasé à l'aide d'un relationnel prédicatif (25.):

- (25.) EBM –  $N_1$  é  $N_2$  /  $N_1$  é  $A_{rel}$  ( $A_{rel}$ : *adjectif relationnel*)
- a. \* A equipa é **sensação** / é sensacional (*equipa sensação*)
  - b. ?A versão é **cliente** (*versão cliente*)
  - c. ?A decisão é **limite** (*decisão limite*)
  - e. \*A zona é **centro** /é central (*zona centro*)
  - f. \* A disciplina é **base** / é básica (*disciplina-base*)

En sixième lieu,  $N_2$  accepte des préfixes adjectivaux modifieurs (26.), même si  $N_2$  n'accepte lui-même ni degré ni modifieur (27.).

- (26.) EBM –  $N_1$  préfixe- $N_2$
- a. programa **pré**-candidato (*programa candidato*)
  - b. cidade **quase**-património (*cidade património*)
  - c. emissão **semi**-pirata (*emissão pirata*)
  - e. projecto **ante**-piloto (*projecto piloto*)

- (27.) EBM –  $N_1$  Q- $N_2$  (Q – *quantifieur*)
- a. \*momento **muito** chave (*momento chave*)
  - b. \*fabricante **bastante** líder (*fabricante líder*)
  - c. \*emissor **pouco** pirata (*emissor pirata*)
  - d. \*empresa **incrivelmente** fantasma (*empresa fantasma*)

En septième lieu,  $N_2$  peut être coordonné à un adjectif relationnel complémentaire, y compris lorsqu'il ne partage pas sa flexion (28.):

- (28.) EBM – N<sub>1</sub>N<sub>2</sub> - Coord - A<sub>rel</sub> (A<sub>rel</sub>: *adjectif relationnel*)
- a. as disciplinas **base**?(s) e **específicas** (*disciplina base*)
  - b. as tecnologias **laser**\*(s) e **nucleares** (*tecnologia laser*)
  - c. as zonas **centro**\*(s) e **costeiras** (*zona centro*)
  - d. as emissões **pirata**(s) e **ilegais** (*emissão pirata*)
  - e. as empresas **fantasma**(s) e **fraudulentas** (*empresa fantasma*)

Cette même coordination est cependant bloquée si l'adjectif en question n'est pas un relationnel (29.):

- (29.) EBM – N<sub>1</sub>N<sub>2</sub> - Coord - A (A: *modifieur attributif*)
- a. \*as disciplinas **base** e **fáceis** (*disciplina base*)
  - b. \* uma tecnologia **laser** e **perigosa** (*tecnologia laser*)
  - c. \* as zonas **centro** e **húmidas** (*zona centro*)
  - d. \* as emissões **pirata** e **antigas** (*emissão pirata*)
  - e. \* uma empresa **fantasma** e **recente** (*empresa fantasma*)

En huitième lieu, N<sub>2</sub> n'admet pas de complément, au contraire de N (30.):

- (30.) EBM – N<sub>1</sub>N<sub>2</sub> - Cmpl/ N - Cmpl
- a. \* emissoras **pirata da rádio** / piratas da rádio  
(*emissora pirata*)
  - b. \* equipa **sensação de vitória** / sensação de vitória  
(*equipa sensação*)
  - c. \* cidade **berço para dormir** / berço para dormir  
(*cidade berço*)
  - d. \* versão **cliente de confiança** / cliente de confiança  
(*versão cliente*)

En neuvième lieu, les EBM acceptent le phénomène d' «incrustation» (31.) signalé par Demonte (1999), selon lequel «quand, dans un syntagme nominal, coexistent plusieurs adjectifs relationnels, ceux-ci s'incruster les uns dans les autres, et l'adjectif situé le plus à droite modifie toujours l'unité formée par le nom et le précédent adjectif relationnel». Ces incrustations

admettent aussi que  $N_2$  soit mobile (32.):

- (31) EBM –  $N_1$  -  $A_{rel}$  -  $A_{rel}$  -  $N_2$  ( $A_{rel}$ : *adjectif relationnel*)
- a. um [[clube desportivo] juvenil] **satélite**  
(*clube satélite*)
  - b. as [[viagens marítimas] portuguesas] **chave**  
(*viagens chave*)
  - c. as [[provas desportivas] nortenhas] **vedetas**  
(*prova vedeta*)
  - d. um [[projecto universitário] português] **piloto**  
(*projecto piloto*)

- (32.) EBM –  $N_1$ -  $N_2$ -  $A_{rel}$  -  $A_{rel}$  ( $A_{rel}$ : *adjectif relationnel*)
- a. um [[projecto **piloto** universitário]] português  
(*projecto piloto*)
  - b. as [[viagens marítimas] **chave**] portuguesas  
(*viagem chave*)
  - c. um [[clube **satélite**] juvenil] desportivo  
(*clube satélite*)

En dixième et dernier lieu, le modifieur  $N_2$  ne vérifie aucune relation de polarité ni de complémentarité et ne possède pas non plus d'antonymes (33.):

- (33.) EBM –  $N_1$ -  $N_2$  / antonymie - polarité
- a. clube **satélite** / \* clube planeta
  - b. versão **cliente** / \* versão vendedor
  - c. equipa **sensação** / \* equipa não-sensação
  - d. viagem **relâmpago** / \* viagem não-relâmpago

Les dix points précédents convergent donc fortement vers la conclusion selon laquelle  $N_2$  vérifie toutes les propriétés typiques des adjectifs relationnels, dont  $N_2$  ne se distingue ponctuellement qu'en termes d'accès limité à la flexion.

La proximité entre les noms et les adjectifs relationnels a été suggérée

par différents auteurs, comme Uehara (1998), Demonte (1999) ou Anward (2000). Demonte note ainsi que, de même que certains noms, «c'est une caractéristique des adjectifs relationnels de se faire accompagner de préfixes prépositionnels tel que *ante-*, des préfixes adverbiaux comme *pré-* ou *anti-*, de même que des préfixoïdes greco-latins du genre *néo-* ou *paléo-*». Ces mêmes préfixes caractérisent  $N_2$  (voir (26.)). La diversité sémantique des adjectifs relationnels nous incite cependant à approfondir la syntaxe de  $N_2$ .

#### 2.4. Syntaxe de $N_2$ relationnel

Du point de vue de la structure des constituants, deux tests supplémentaires peuvent nous aider à déterminer de façon objective la nature de  $N_2$  : les constructions elliptiques et les constructions prédicatives.

##### 2.4.1. $N_2$ elliptique

Il existe des cas où  $N_2$  révèle nettement sa nature de modifieur adjectival. Du point de vue des tests de consistance associés à la manipulation de constituants effacés ou ellipsés, on peut remarquer qu'il existe une différence de sensibilité, relativement aux ellipses, entre les EBM et les EBC. En effet, comme l'ont noté Bosque & Picallo (1996), la présence d'un adjectif relationnel permet d'identifier un nom ellipsé ou sous-entendu, comme en (34.a), au contraire de ce qui se passe avec les compositions classiques (34.b), qui n'autorisent pas cette opération d'ellipse du nom :

(34.) Ellipse de N suivi de  $A_{rel}$  ( $A_{rel}$ : *adjectif relationnel*)

a. A sociedade industrial e a [-] **cibernética** marcam as grandes etapas do séc. XX

b. \* Vi los hombre-rana y los (-) anuncio

(Espagnol: *hombre-rana, hombre-anuncio* (exemple de Bosque & Picallo))

On rencontre précisément la même symétrie que (34.a) avec les EBM, où  $N_2$  identifie, dans ce type d'ellipse,  $N_1$  :

(35.) EBM – Ellipse de  $N_1$  suivi de  $N_2$

a. São mais frequentes as viagens longas que as [-] relâmpago  
*(viagem relâmpago)*

- b. A versão base e a [-] cliente têm preços diferentes  
(*versão cliente*)
- c. A zona centro e a [-] norte não serão afectadas  
(*zona norte*)

L'ellipse est donc possible puisque la structure des constituants de ces constructions symétriques est capable d'identifier  $N_1$  sur la base de l'interprétation de  $N_2$  comme constituant capable de légitimer la position elliptique, c'est-à-dire comme un adjectif (Cf. Martinho 1998).<sup>21</sup>

#### 2.4.2. N2 prédicatif

Il existe, comme on sait, une restriction forte quant à l'accès des adjectifs relationnels à une position prédicative. Globalement, comme noté par Demonte (1999), cette position est ouverte à une sous-classe de relationnels, les « relationnels non-argumentaux » (36.), mais est d'accès difficile pour les « relationnels argumentaux » (37.).

##### (36.) Relationnels non argumentaux

- a. revista **mensal** / a revista é **mensal**
- b. comédia **musical** / a comédia é **musical**

##### (37.) Relationnels argumentaux

- a. engenheiro **eléctrico** / ?\* O engenheiro é **eléctrico**
- b. câmara **municipal** / ?\* a câmara é **municipal**

Dans les EBM, certains  $N_2$  semblent vérifier effectivement les propriétés typiques des relationnels argumentaux, puisqu'ils sont interdits d'accès à la position prédicative, alors que d'autres ont, au contraire, accès libre à cette même position. Ceci nous invite à conclure que la classe des  $N_2$  est une classe relationnelle elle-même divisée en deux sous-ensembles, celui des  $N_2$  non argumentaux (38.), potentiellement prédicatifs, et celui des  $N_2$  non prédicatifs, strictement arguments de  $N_1$  (39.).

---

<sup>21</sup> Rappelons que, comme l'illustre l'exemple (34.b), l'ellipse est bloquée par les procédés de composition.

- (38.) EBM – N<sub>2</sub> relationnel non argumental
- a. O nosso clube é associado, não é **satélite**  
 (*clube satélite*)
  - b. Esta nova tecnologia não é óptica, é **laser**  
 (*tecnologia laser*)
  - c. Esta prova é **rainha**  
 (*prova rainha*)
  - d. A Rádio-Gaga nunca foi declarada, sempre foi **pirata**  
 (*rádio pirata*)
- (39.) EBM – N<sub>2</sub> relationnel argumental
- a. \* A cidade de Guimarães não é recente, é **berço**  
 (*cidade berço*)
  - b. ? O retrato do criminoso é muito fiel, não é **robot**  
 (*retrato robot*)
  - c. \* Paris não é cidade qualquer, é **luz**  
 (*cidade luz*)
  - d. \* Estas zonas são periféricas, não são **centro**  
 (*zona centro*)

Il semble donc légitime de proposer que, dans une minorité de EBM (**cidade berço**), les deux membres maintiennent une relation de type prédicat / argument, et que, dans les autres cas (**clube satélite**), N<sub>1</sub> et N<sub>2</sub> ont bien une simple relation de modification. Ainsi, des contrastes comme (38.a) et (39.a), par exemple, trouveraient une explication indépendante basée sur la classification interne des adjectifs relationnels. Il s'agit là, évidemment, d'une preuve forte en faveur de la nature relationnelle de N<sub>2</sub>. Notons aussi, au passage, que les EBM impliquant un N<sub>2</sub> argumental semblent rejeter la flexion de N<sub>2</sub> (voir 39.d), ce qui, là encore, procède d'une explication indépendante : cela tiendrait au fait qu'un argument interne n'a pas à partager ses traits flexionnels avec le noyau prédicatif.

Ces exemples nous invitent aussi à considérer qu'une distinction supplémentaire devrait être introduite entre expressions binominales modifiées par un relationnel non argumental et expressions binominales saturées par un relationnel argumental ; nous n'explorerons pas cette voie plus avant ici.

## 2.5. Les EBM dans d'autres langues

Dans d'autres langues, les données extraites du portugais renforcent notre analyse, d'autant plus que les EBM y sont aussi fréquentes. Commençons par le basque, langue fertile en expressions binominales, où l'on observe en (40.) des EBM de type  $N_2N_1$ , dans lesquelles  $N_2$  précède et modifie  $N_1$  (cf. Coyos (2004) :

### (40.) EBM – $N_2N_1$ (Basque)

- a. aho<sub>N<sub>2</sub></sub> literatura<sub>N<sub>1</sub></sub> (lit. bouche littérature - '*littérature orale*')
- b. esku<sub>N<sub>2</sub></sub>-lan<sub>N<sub>1</sub></sub> (lit. main travail - '*travail manuel*')
- c. ama<sub>N<sub>2</sub></sub> eskola<sub>N<sub>1</sub></sub> (lit. mère école - '*école primaire*')

En basque, à l'image des langues de type *Déterminant-Déterminé*, l'ordre est  $N_2N_1$  et ces expressions y sont donc de tête finale. L'ordre  $N_2N_1$  y correspond bien à des cas d'EBM où  $N_2$  est un modifieur épithète de  $N_1$ . Notons aussi que  $N_2$  est interprété systématiquement comme un adjectif relationnel (« main / manuel »). A la différence des autres langues, les EBM semblent être en basque un processus lexical systématique et majeur.<sup>22</sup>

Dans les langues romanes, les EBM sont, de même, très fréquentes, bien qu'à une échelle moindre et avec une diversité lexicale variable (Fabre 1996). Monterini (2006) décrit en Italien (41.) l'existence de formes  $N_1N_2$  dans lesquelles  $N_2$  manifeste les propriétés prototypiques précédemment identifiées en portugais:

### (41.) EBM – $N_1N_2$ (Italien)

- a. viaggio-lampo, reunione-lampo (lit. *voyage éclair, réunion éclair*)
- b. telefonata-fiume (lit. *appel fleuve – 'appel téléphonique interminable'*)
- c. capoazienda (lit. *chef entreprise – 'siège social'*)<sup>23</sup>
- d. casa madre

---

<sup>22</sup> Pour d'autres exemples en Basque, cf. Martinho (2007). Etant donné l'extrême diversité de ces cas en basque – où certains  $N_2$  ne sont pas forcément des N prototypiques-, ceux-ci ne peuvent être circonscrits aux cas classiques d'EBM sur lesquels porte notre étude.

<sup>23</sup> Dans ce cas, il s'agit d'une EBM de type  $N_2N_1$  (voir note 25).

En français, les EBM en (42.) sont, de même, très fréquentes et ont été répertoriées par Villoing (2002), qui cite de nombreux cas de ce qu'il appelle les « composés par modification », dans lesquels (selon nous)  $N_2$  modifie métaphoriquement  $N_1$  :

(42.) EBM –  $N_1N_2$  (Français)

- a. roman fleuve ('roman interminable')
- b. espace jeunes ('espace destiné aux jeunes')
- c. guerre éclair ('guerre à durée limitée')
- d. mot valise (mot résultant de *blending*)
- e. invité surprise ('invité non prévu')<sup>24</sup>

En espagnol, les mêmes structures EBM apparaissent en (43.), de même que  $N_2$  métaphorique :

(43.) EBM –  $N_1N_2$  (Espagnol)

- a. casa piloto (*maison pilote*)
- b. escuela tipo (*école type*)
- c. empresa modelo (*entreprise modèle*)
- d. guerra relámpago (*guerre éclair*)

Il est intéressant de noter qu'en espagnol, les adjectifs relationnels 'classiques' sont souvent, dans le contexte de la traduction, en compétition avec  $N_2$  relationnel (voir 44.), ce qui suggère une équivalence traductologique

---

<sup>24</sup> L'un des réviseurs, que je remercie, observe qu'on peut aussi trouver l'ordre  $N_2N_1$ , comme en français dans l'exemple "candidat-vaccin" (cf. "Un nouveau *candidat-vaccin* contre le virus d'Ebola", titre du magazine *Science & Avenir*, 06-06-2005). La même expression se retrouvant en anglais (« candidate vaccine »), nous en concluons qu'il s'agit effectivement d'une variante EBM, de type  $N_2N_1$ . Après avoir élargi la recherche, il s'est avéré que la forme « vaccine candidate » ( $N_1N_2$ ) est aussi attestée en anglais (« Malaria Vaccine Candidate Gives Disappointing Results » - New York Times, 9-11-2012), de même que « vaccin candidat » en français (« Medicago produit avec succès un *vaccin candidat* à PPV contre la souche H7N9 » - Association des Médecins Vétérinaires Praticiens du Québec, 8-05-2013). Si nous avons ici affaire indéniablement à des expressions binominales modifiées, il semble que l'ordre interne du couple Modifiant – Modifié ne soit pas (encore) l'objet d'une forme lexicale stable, ce qui, comme nous l'avons noté précédemment, est une caractéristique des EBM. De fait, loin d'être opaque, comme c'est le cas des compositions classiques, nous sommes ici en présence d'expressions ouvertes qui vérifient les propriétés d'instabilité graphique (avec ou sans trait d'union) et lexicale notées plus haut, ce qu'atteste bien l'autonomie relative des éléments combinés ici mentionnés. Le terme « candidate / candidat » est donc un  $N_2$ , qui, après avoir perdu sa capacité référentielle, assume une fonction de modification, ce qui peut expliquer que, en tant que modifieur, il puisse occuper une position pré ou post-nominale : *vaccin candidat / candidat vaccin*.



entre les séquences  $N_1N_2$  et  $N-A_{rel}$  :

(44.) EBM –  $N_1N_2$  à  $N - A_{rel}$  ( $A_{rel}$ : *adjectif relationnel*)

- a. droit canon  $N_2$  / derecho **canónico**  $A_{rel}$
- b. séance marathon  $N_2$  /sesión **maratoniana**  $A_{rel}$

## 2.6. Le 'nom relationnel' $N_2$

On peut donc conclure, à la lumière des données précédentes, que les expressions binominales modifiées remplissent une double fonction : dénotative (pour le nom  $N_1$ ) et restrictive (pour l'épithète post-nominale  $N_2$ ). Il est clair que  $N_2$  apparaît comme un élément restrictif dans les EBM. Celles-ci sont constituées sur la base d'une paire de noms dont le second modifie la dénotation du premier et appartient, en principe, à un ensemble semi-fermé de lexèmes. Les différents tests présentés précédemment suggèrent que  $N_2$  équivaut à un adjectif relationnel: il accepte une position prédicative non argumentale, il est exclusivement post-nominal, il refuse le degré adjectival, il dénote (prototypiquement) des entités, il s'incruste avec les relationnels classiques, il se coordonne avec des adjectifs relationnels mais pas avec des attributifs et il accepte des préfixes exclusivement relationnels (*macro-, infra-, pré-, néo-*). Ces  $N_2$  seraient donc des adjectifs défectifs, des noms qui n'auraient pas réussi totalement leur conversion en modifieurs et qui, en particulier, n'auraient accès ni à la flexion ni au degré adjectival, c'est-à-dire à la morphologie, et seraient donc restés sur leur base lexicale nominale, ceci de façon transversale aux langues consultées.

Est-il légitime de voir en  $N_2$  l'équivalent d'un adjectif relationnel? La proximité entre formes relationnelle et nominale est reconnue. Comme le note Uehara (1998), lorsqu'un nom apparaît (en japonais) comme modifieur, il est construit comme un terme relationnel et continue d'autant moins à maintenir une lecture nominale que sa syntaxe se rapproche de celle de l'adjectif. Et Demonte (1999), évoquant des séquences du type (44.), suggère que « si, dans les séquences NA, les relationnels ont des caractéristiques nominales, ils pourraient en fait former des séquences NN ». <sup>25</sup>

---

<sup>25</sup> Par ailleurs, certains adjectifs relationnels (par exemple, les adjectifs de nationalité) reçoivent un rôle thématique lorsqu'ils sont liés à une tête nominale. Cf. Pysz (2007).

La possibilité que  $N_2$  soit un relationnel est encore renforcée par sa nature lexicale:  $N_2$  correspond en portugais, comme nous l'avons vérifié par consultation de la base CETEM, à un ensemble de noms à usage métaphorique, desquels dérivent, pour la plupart, des adjectifs relationnels classiques :

(45.) EBM –  $N_2$  à A (*Dérivation lexicale*)

- a. base → básico, exemplo → exemplar, tipo → típico
- b. centro → central, marco → marcante, herói → heróico
- c. maravilha → maravilhoso, modelo → modelar ...
- d. verdade → verdadeiro, ficção → fictício,

Lorsque la métaphore ne se traduit pas par un procédé de dérivation, elle est alimentée par des procédés dénotatifs ou connotatifs fondés sur la sélection de traits sémantiques génériques, comme par exemple les équivalences sémantiques en (46.):

(46.) EBM –  $N_2$  / A (*Equivalence sémantique*)

- a. limite / excessivo, forçado
- b. mestre / principal, central
- c. piloto / pioneiro, inovador
- d. chave / essencial, determinante
- e. pirata / clandestino, ilegal
- f. fantasma / misterioso, invisível
- g. relâmpago / rápido, instantâneo
- h. vedeta / célebre, famoso
- i. recorde / excessivo
- j. satélite / próximo, complementar

En somme, les EBM analysées confirment que  $N_2$  peut être considéré comme un adjectif. Ainsi, lors de l'édification de la structure syntaxique du DP, la rencontre entre un N référentiel et un N prototypique détermine la projection d'une tête N modifiée par un relationnel. Cette projection résulte de restrictions générales sur la structure des constituants, bannissant, par exemple, l'adjacence redondante de traits catégoriaux et référentiels de

deux entités qui ont normalement une fonction argumentale. De façon à encadrer les EBM dans le modèle générativiste, on pourrait proposer, à titre d'hypothèse, que le système computationnel du langage humain convertit  $N_2$  en A en lui attribuant un trait [+V] en syntaxe implicite, au niveau de la Forme Logique.<sup>26</sup>

### 3. Conclusion

L'existence de ces séquences EBM, du genre *equipa maravilha* (*équipe merveille*), dans lesquelles  $N_2$  est un objet lexical et syntaxique ambigu, soulève évidemment de nombreux doutes, en particulier sur la façon dont serait menée cette conversion catégorielle. Celle-ci va effectivement à l'encontre de l'information lexicale initiale, qui stipule que  $N_2$  est un constituant [+N][-V] et est donc, par définition, incompatible avec une position de modifieur. Comme nous l'avons vu, les EBM manifestent diverses restrictions de sélection et de proximité entre  $N_1$  et  $N_2$ . Ainsi, par exemple,  $N_1$  et  $N_2$  doivent toujours être adjacents,  $N_1$  est toujours (en portugais) la tête du NP, et  $N_2$  peut être absent. Tout cela correspond bel et bien à la syntaxe de l'adjectif, ce qui, face au paradigme théorique en vigueur, est tout de même contradictoire et met partiellement en cause le modèle de projection des catégories lexicales en syntaxe.

L'existence de ce genre de lexies instables, contenant des éléments mixtes, mi-noms mi-adjectifs, à mi chemin entre lexique et syntaxe, représente un défi important pour l'actuel modèle lexical de la grammaire générative, suggère l'existence de zones d'ombre quant à son application et demande donc la formulation de modèles lexicaux alternatifs.

On peut voir dans les théories du « recyclage » une réponse partielle à ces problèmes. Le recyclage (Croft 1990, Anward 2000) consiste à reconnaître l'existence d'une orientation générale dans laquelle les langues tendent à optimiser leurs ressources lexicales: au lieu de multiplier les items lexicaux distincts pour chaque combinaison de concepts et de fonctions, elles « recyclent » certains items en différentes fonctions. Ces items lexicaux seraient en position d'équilibre délicat entre la définition rigoureuse issue des classes traditionnelles du lexique et le besoin de flexibilité catégorielle

---

<sup>26</sup> Cela équivaut à proposer que  $N_2$  serait bien un nom en Syntaxe, mais fonctionnerait et serait interprété comme un adjectif.

manifesté par certains domaines lexicaux. Ce serait le cas de ces « noms épithètes » ici abordés. Ces épithètes, qui n'ont pas besoin d'identification fonctionnelle, se contentent en effet d'une position d'adjacence au terme modifié et sont bien, comme cela est prévu pour les catégories recyclées, dépouillées de leur flexion.

Cette mode des binominaux modifiés s'inscrirait alors dans un modèle plus large, dans lequel les catégories lexicales sont utilisées de façon aussi transversale que possible, selon le principe de la maximalisation des signifiés et de la minimalisation des ressources.

#### RÉFÉRENCES

- Anward, J. 2000. A dynamic model of part-of-speech differentiation., in *Approaches to the Typology of Word Classes*, Vogel, Petra M. & Comrie, Bernard (Eds.), Mouton de Gruyer.
- Bhat, D.N.S. 1994. *The adjectival category: Criteria for differentiation and identification*. (Studies in Language Companion Series 24.) Amsterdam: Benjamins.
- Bosque, I. & Picallo C. 1996. Postnominal adjectives in Spanish DPs, *Journal of Linguistics*, 32, 349-385.
- Coyos, J.-B. 2004. Basque, in *Le nom composé: Données sur 16 langues*, Arnaud, Pierre J.L. (Ed.), Lyon: Presses Universitaires de Lyon, 47-70.
- Croft, W. 1990. Conceptual framework for grammatical categories (or, a taxonomy of propositional acts), *Journal of Semantics*, 7, 245-279.
- Demonte, V. 1999. A Minimal Account of Spanish Adjective Position and Interpretation. In Franco J., Landa A. & Martín J. (Eds.), *Grammatical Analyses in Basque and Romance Linguistics*, Amsterdam, John Benjamins.
- Fabre, C. 1996. *Interprétation automatique des séquences binominales en anglais et en français. Application à la recherche d'informations*, thèse de doctorat, Université de Rennes.
- Kleiber, G. 1990. *La sémantique du prototype*, Paris: PUF.
- Martinho, F. 1998. *A Elipse Nominal em Português e Francês*, dissertation de Maîtrise, Universidade do Porto.
- 2007. *Sintaxe e semântica dos adjetivos graduáveis em Português*, dissertation de

- doctorat, Universidade de Aveiro.
- Montermini, F. 2006. La Composition en Italien dans un Cadre de Morphologie Lexématique, communication présentée à la Journée d'études "Composition", Arras, 24 mars 2006.
- Pysz, A. 2007. The structural location of adnominal adjectives: Prospects for Old English, *Journal of Theoretical Linguistics*, n°5, 59-85.
- Ranchhod, E. 2003. O lugar das expressões 'fixas' na gramática do Português, in *Razões e Emoção, Miscelânea de estudos em homenagem a Maria Helena Mira Mateus, Castro, I. & Duarte, I. (Eds.)*, Vol. II, Imprensa Nacional-casa da Moeda, 239-254.
- Ribeiro, S. 2010. *Compostos nominais em português. As estruturas VN, NN, NprepN e NA*. München: LINCOM (LINCOM Studies in Romance Linguistics 64).
- Rio-Torto, G.M. 1998. *Morfologia Derivacional. Teoria e Aplicação ao Português*, Coleção Linguística 12, Porto Editora.
- Rio-Torto, G.M. & Ribeiro, S. 2012. Portuguese Compounds. *Probus* 24: 119-145.
- Savary, A. 2004. *Recensement et description des mots composés – méthodes et applications*, Thèse de doctorat en Informatique Fondamentale, Laboratoire d'Automatique Documentaire et Linguistique, Université Paris 7.
- Scalise, S. & Bisetto, A. 2009. Classification of compounds. In R. Lieber & P. Stekauer (Eds.), *The Oxford Handbook of Compounding*, Oxford: Oxford University Press, 49-82.
- Uehara, S. 1998. *Syntactic categories in Japanese: A typological and cognitive introduction*. Tokyo: Kurosio.
- Villalva, A. 2003. Formação de palavras: composição, in Mateus, M. H., Brito, A.-M., Duarte, I., Frota, S., Matos, G., Oliveira, F., Vigário, M. & Villalva, A., 2003, *Gramática da Língua Portuguesa*, série Linguística, Caminho.
- Villoing, F. 2002. *Les mots composés [VN] N/A du français: réflexions épistémologiques et propositions d'analyse*, thèse de doctorat, Paris-X Nanterre.
- <http://www.linguateca.pt/cetempublico/> (last access: 31-05-2013).



# Italian verb to noun conversion: the case of nouns in –a deriving from verbs of the 2<sup>nd</sup> and 3<sup>rd</sup> conjugation<sup>1</sup>

Daniela Marzo  
daniela.marzo@ling.uni-stuttgart.de  
University of Stuttgart (Germany)

ABSTRACT. The formation and structure of the following Italian deverbal nouns is an open question in Italian word-formation research: (1a) *cammino*, (1b) *ritegno*, (1c) *riparto*, (2a) *sosta*, (2b) *classifica*, (3a) *tema*, (3b) *basta*, (3c) *perquisita*. While instances such as in (1) and (2) have, depending on the theoretical background, been analysed either as *root conversion* (cf. (1)) and *conversion of the infinitival stem* (cf. (2)) or as *suffixation* (cf. (1) and (2)), the spectrum of accounts for the instances in (3) includes, in addition, *root conversion with analogical gender assignment* (cf. (3a) and (3b)) as well as *analogical backformation* (cf. (3c)).

In the present contribution it is shown that the current conversion approaches as well as the suffixation approach are insufficient to explain the formation and structure of the nouns in (1) to (3). While the suffixation account is ruled out mainly for semantic and distributional reasons, it is shown that the existing conversion accounts are, in principle, more plausible to account for the nouns in (1) and (2). However, as for the analogy-based conversion approaches, it is shown that only (3c) is a proper analogy-driven formation. In contrast, type (3b) will be shown to be an instance of root conversion remodelled by blocking. Moreover, in order to explain the formation and structure of the noun type instantiated by (3a), an additional type of conversion has to be distinguished: (3a) is shown – by the help of data from Old Italian – to be a synchronic remnant of an Old Italian conversion type: conversion of stem 2 of the Italian verbal stem space (terminology from Giraudo, Montermini & Pirrelli 2009), i.e. the stem on which also some Italian inflectional forms, as e.g. some of the present subjunctive forms are constructed.

KEY-WORDS. Conversion, suffixation, verbal stem space, blocking, analogic word-formation, networks.

---

<sup>1</sup> I would like to thank Birgit Umbreit for the fruitful discussions of earlier versions of this paper.

## 1. Introduction

The formation and structure of Italian deverbal nouns such as in (1) to (3) is an open and controversial question in Italian word-formation research.

- (1) a. *cammino* (N) ‘path, walking’ – *camminare* (v) ‘to walk’  
 b. *ritegno* (N) ‘reservation’ – *ritenere* (v) ‘to retain’  
 c. *riparto* (N) ‘partitioning’ – *ripartire* (v) ‘to divide, to portion’
- (2) a. *sosta* (N) ‘stop, rest, stopping’ – *sostare* (v) ‘to rest, to stop’  
 b. *classifica* (N) ‘classification’ – *classificare* (v) ‘to classify’
- (3) a. *tema* (N) ‘fear, fearing’ – *temere* (v) ‘to fear’  
 b. *basta* (N) ‘stitch, stitching’ – *bastire* (v) ‘to stitch’  
 c. *perquisa* (N) ‘search, searching’ – *perquisire* (v) ‘to search’

While instances such as in (1) and (2) have, depending on the theoretical background, mainly been analysed either as *root conversion* and *conversion of the infinitival stem* (cf. Thornton 2004: 516-520 for (1) and (2a) respectively) or as *suffixation* (for *overt* suffixation cf. Scalise 1994: 271-275; for *zero* suffixation cf. Dardano 1988: 56), the spectrum of accounts for the instances in (3) includes, in addition, *root conversion with analogical gender assignment* (cf. Thornton 2004: 518 for (3a) and (3b)) as well as *analogical backformation* (cf. Thornton 2004: 520 for (3c)).

In the present contribution two important conversion and suffixation accounts will be compared with respect to their explanatory power of the formal and the semantic aspects of nouns such as in (1) to (3) as well as with regard to their relation to the respective verbs (cf. section 2.). While the suffixation account is ruled out mainly for semantic and distributional reasons, it is shown that the existing conversion accounts are, in principle, more plausible to account for the nouns in (1) and (2). However, in order to explain the formation and structure of all instances in (3), different subgroups of conversion have to be distinguished (cf. section 3.) Type (3a) will be analysed as a synchronic remnant of an Old Italian conversion type (cf. 3.1), i.e. conversion of stem 2 of the Italian verbal stem space (terminology from Giraudo, Montermini & Pirrelli 2009), a stem on which also the Italian 1<sup>st</sup> to



3<sup>rd</sup> singular present subjunctive and the 1<sup>st</sup> and 3<sup>rd</sup> plural present indicative forms are constructed. In contrast, type (3b) will be shown to be an instance of root conversion remodelled by blocking (cf. 3.2). Type (3c) will be discussed as a product of analogy to existing word-family and network relations such as *classifica* (N) 'classification' – *classificare* (v) 'classify' – *classificazione* (N) 'classification' (cf. 3.3). Section 4 will summarize the findings of this paper and give a brief outlook on open questions and further research.

## 2. The data groups and existing accounts

### 2.1. Some background information on the data groups

The deverbal nouns in (1) to (3) belong to two different declensional classes, marked by either (prototypically feminine) *-a* (cf. (2) and (3)) or (prototypically masculine) *-o* (cf. (1)) in the singular and *-e* and *-i* in the respective plural forms, whereas their verbal counterparts belong to three different conjugational classes that are traditionally distinguished mainly by the theme vowel of their infinitival forms (*-a-*, *-e-* and *-i-* respectively; cf. (1a) and (2) in contrast to (1b) and (3a) as well as (1c), (3b) and (3c)).

As for the semantics of the nouns, it can be observed that independently of the declensional class they belong to, they usually are event nominals and often display, in addition, readings obtained by semantic shifts, such as locative readings (cf. (1a) *cammino*) or e.g. resultant state readings (cf. (3a) *tema*). On a more general level, this means that their readings are connected to the readings of the respective verbs either by a relation of conceptual identity or one of conceptual contiguity (cf. Koch 2001, Blank 1998). In sum, the semantic patterns are quite similar to those of overt deverbal nominalizations (for an overview of the observed semantic patterns cf. Thornton 2004: 516-520; Tollemache 1954: 17).

By far, the biggest subgroup of the nouns in (1) to (3) is, according to Thornton (2004: 516), type (1) (nominal *-o* + infinitival theme vowel *-a-/e-/i-*), that has approximately 800 members. Group (2) and (3) (nominal *-a* + infinitival *-a-/e-/i-*) are, in contrast, much smaller (together not even 300 instances). Type (2) has, in addition, to be split up into two subgroups (Thornton 2004: 517-520): (i) proper deverbal nouns such as (2a) *sosta* (based on the counts in Thornton, about 60% of the deverbal nouns in *-a*) and (ii) nouns that are in a direct relation to their verbs only from a

synchronic point of view, but that are denominal backformations from a diachronic perspective<sup>2</sup>, such as (2b) *classifica* (about 40% of the instances) that is a backformation from *classificazione* ‘classification’.<sup>3</sup> Subgroup (3) (nominal –a + infinitival –e/–i–) is by far the smallest group and concerns only a handful of nouns as Thornton (2004: 518) puts it (some more instances can be found in Tollemache 1954: 157-164 and Rohlfs 1969: 473).

Though there seem to be some neologisms in the 20<sup>th</sup> century, the noun patterns under inspection are not very productive nowadays (less than 30 neologisms for type (1a); cf. Thornton 2004: 517; only 2 for type (2) and 1 for type (3), i.e. (3d); cf. Thornton 2004: 518, 520).

## 2.2. Type (1) to (3): Conversion or suffixation?

In what follows, two important existing accounts for the instances in (1) to (3), more precisely *root conversion* and *conversion of the infinitival stem* on the one (cf. 2.2.1) and *suffixation* (cf. 2.2.2) on the other hand will be presented and compared with respect to their benefits and shortcomings in explaining the data. As for (3a) to (3c), alternative hypotheses will be introduced in section 3.

### 2.2.1. Conversion

According to Thornton (2004: 516-520) the types instantiated by (1) to (3) are all instances of conversion (cf. also Tollemache 1954). More precisely, she analyses type (1), characterized by nominal –o and verbal –a/–e/–i–, such as (1a) *cammino* – *camminare* as *root conversion*. Her argumentation goes as follows: The verbal root *cammin–* (v) is converted into a nominal element *cammin–* (n). In order for *cammin–* (n) to be employed as a prototypical Italian noun, it has (i) to get a vocalic ending, (ii) to receive grammatical gender and (iii) to be integrated into a noun class. In the absence of a vowel ending that could be reinterpreted as gender and noun class marker, and the default gender for Italian converted nouns being masculine (cf. e.g. Thornton 2003), *cammin–* (v) is assigned the most

<sup>2</sup> According to Thornton (2004: 518-519) denominal backformation is, in these cases, more plausible than deverbal conversion because the suffixed noun can be found in text corpora much earlier than the short noun and because they are usually perfectly homonymous; however, cf. section 2.2.1 and footnote 5.

<sup>3</sup> According to Thornton (2004: 518-520) this pattern was productive in Italian bureaucratic language of the 18<sup>th</sup> and the 19<sup>th</sup> centuries. Scalise (1994: 41), too, confines cases like these to special and sectorial languages.

productive masculine inflectional class (cf. Dressler/Thornton 1996), i.e. –o (SG)/–i (PL) (cf. (4)).

(4) *cammin–* (v) → *cammin–* (N) + [–o (MASC.SG)/–i (MASC.PL)]

In contrast to type (1), type (2) is, in Thornton's analysis (2004: 518), an instance of conversion of the stem of the infinitive, such as *sosta–* (v) (cf. (2a)). As, in contrast to the above-mentioned *cammin–* (v), *sosta–* (v) already has a vocalic ending, it can, in principle, be used as a prototypical Italian noun without its form being altered. As –a prototypically represents feminine gender<sup>4</sup> the most economic solution is chosen and the vocalic ending of the stem is reinterpreted as feminine gender and assigned to the inflectional class –a (SG)/–e (PL), (cf. (5)).

(5) *sosta* (v) → *sosta* (N.FEM.SG) / *soste* (N.FEM.PL)

As pointed out in section 2.1, type (2) is, from a diachronic perspective, to be split up into two subtypes (cf. Thornton 2004: 517-520). From this perspective the analysis in (5) can be applied to (2b) only from a synchronic perspective.<sup>5</sup> From the diachronic point of view, (2b) *classifica* is an instance of backformation from *classificazione*. The suffix –zione is cancelled and the final –a of the remaining short form is reinterpreted as feminine singular (cf. (6)).

(6) *classificazione* (N.FEM.SG) → *classificazione* (N.FEM.SG) / *classifiche* (N.FEM.PL)

The instances in (3) that are characterized by nominal –a and verbal –e/–i– such as (3a) *tema – temere*, are, in turn, according to Thornton (2004: 518), just like type (1) cases of root conversion that have, in contrast to type (1), been integrated into the –a/–e noun class by analogy with type (2). Though this is certainly a plausible scenario from a descriptive perspective, Thornton leaves open the important question of why type (3) has been assigned feminine gender and been integrated into the –a (SG)/–e

<sup>4</sup> It also can be assigned masculine gender, but only under specific circumstances (see Thornton 2003).

<sup>5</sup> This is also perfectly justified from the naïve native speaker's point of view, cf. Marzo (2013: 47); Thornton (2004: 519) for the treatment of these cases in synchronic lexicography; Umbreit (2011) for motivational networks; Gaeta (2010) for synchronic collision of diachronically different word formation patterns.

(PL) class in analogy to type (2), if the most productive noun class is the prototypically masculine one marked by  $-o_{(SG)}/-i_{(PL)}$  (cf. section 2.1).

### 2.2.2. Suffixation

In contrast to Thornton, Scalise (1994: 274-275) views all instances in (1) to (3) as suffixation. In his stem-based word-formation theory, all morphologically constructed words are constructed on stems, more precisely on a root plus theme vowel combination. The theme vowel is always the one of the infinitival form, i.e.  $-a-/e-/i-$ . In order to account for the noun types exemplified by (1) to (3) he posits that the theme vowels of the stems are cancelled by a phonological vowel cancellation rule when the suffixes  $-a$  or  $-o$  are added, cf. (7):

- (7) a. *camminā* (v) +  $-o$  (N) → *cammino* (N) cf. type (1)  
 b. *sostā* (v) +  $-a$  (N) → *sosta* (N) cf. type (2)  
 c. *temā* (v) +  $-a$  (N) → *tema* (N) cf. type (3)

As for type (2), Scalise (1994: 275) is aware of the fact that a stem conversion analysis is, in principle, also possible, but for reasons of analogy he deliberately opts for the suffixation solution. While this choice certainly makes sense from within his framework, Scalise does not offer any description of the semantics and the distribution of the supposed suffixes. Such a description would, however, be necessary, to corroborate the status of  $-a$  and  $-o$  as suffixes, though it is certainly not an easy task in view of the coinciding semantic structures of the words in question (cf. section 2.1, but also Scalise 1994: 275).

### 2.2.3. Summary: Conversion or suffixation?

While Thornton's analysis (root conversion, conversion of the infinitival stem, analogy) cannot convincingly explain all examples in (3), Scalise's approach would in principle also account for the instances in (3), if the semantics and distribution of the suffixes could be described more properly, which is far from being a self-evident task (cf. e.g. Scalise 1994: 275). Interestingly, Scalise himself does not seem to bother about words of type

(3) at all, as he states that the posited suffixes only apply to verbs of the 1<sup>st</sup> conjugation (Scalise 1994: 275).

In view of these completely different approaches and in comparison to the situation in other Romance languages, where the proposed analyses of the (roughly) corresponding noun types either clearly speak in favour of conversion (for French Meinschaefer 2005: 212-219; but especially Tribout 2010 and 2012; for Portuguese Rodrigues 2009) or in favour of suffixation (for Spanish e.g. Meinschaefer 2005: 212-219 and Rainer 1993: 382-383; 457-458; 620-622), the formation and structure of the Italian nouns in (1) to (3) still lack a satisfying explanation.

### 3. Three subtypes of Italian verb to noun conversion: *tema*, *basta* and *perquisa*

#### 3.1 *Tema*: A synchronic remnant of an Old Italian conversion type

In what follows I will argue that in Old Italian there existed, in addition to the conversion types proposed by Thornton (cf. 2.2.1 on *root conversion* and *conversion of the infinitival stem*), another type of stem conversion forming feminine nouns in –a: conversion of the verbal stem 2 (numbering taken from Girauo et al. 2009: 3), i.e. the same stem on which in inflectional morphology the 1<sup>st</sup> person singular and 3<sup>rd</sup> person plural present indicative forms as well as the 1<sup>st</sup> to 3<sup>rd</sup> singular and 3<sup>rd</sup> plural present subjunctive forms are constructed (for examples see table 1 below).

While the potential role of the totality of the verbal stem space for word formation has been extensively studied for some languages (cf. for French Bonami, Boyé & Kerleroux 2009, Fradin 2009, Tribout 2010, 2012, Villoing 2010; for German Nolda 2012),<sup>6</sup> studies of the Italian verbal stem space have so far been concentrating on inflectional phenomena (cf. Pirrelli & Battista 2000, Maiden 2003, Girauo et al. 2009; for French inflection see Boyé & Bonami 2002 and 2003, Bonami, Boyé, Girauo & Voga 2008). However, there is reason to assume that in Italian, too, other verbal stems than just the infinitival stem (= stem 12 in the terminology of Girauo et al. 2009: 3) are relevant for word-formation.

Several reasons speak in favour of stem 2 being an input form to Italian

---

<sup>6</sup> For French, e.g., it has even been shown that in the verbal stem space there can be stems that are hidden to inflection and only visible to word-formation processes (cf. e.g. Tribout 2010, 2012 for conversion and Bonami et al. 2009 for derivation).

verb to noun conversion<sup>7</sup>: (i) from a synchronic perspective stem 2 and Italian nouns of the type exemplified by (1a) *tema* as well as the stem of the above mentioned inflectional forms are perfectly homonymous (cf. examples in table 1).

Table 1 – Homonymy of type (3a) nouns with stem 2 and present subjunctive forms<sup>8</sup>

Modern Noun Form (FEM.SG)	1 <sup>st</sup> Occurrence (Zingarelli Online)	Modern Verbal Stem 2*	Present Subjunctive (1 <sup>st</sup> to 3 <sup>rd</sup> singular)	Infinitive
<i>beva</i> 'drink, taste of wine'	1625	<i>beva</i>	<i>beva</i>	<i>bere</i> 'to drink'
<i>cerna</i> 'selection'	1321	<i>cerna</i>	<i>cerna</i>	<i>cernere</i> 'to select'
<i>doglia</i> 'pain'	1249	<i>dolga</i>	<i>dolga</i> and <i>doglia</i>	<i>dolere</i> 'to hurt'
<i>piova</i> 'rain'	1266	<i>piova</i>	<i>piova</i>	<i>piovere</i> 'to rain'
<i>possa</i> 'power'	1306	<i>possa</i>	<i>possa</i>	<i>potere</i> 'to be able to'
<i>tema</i> 'fear'	1294	<i>tema</i>	<i>tema</i>	<i>temere</i> 'to fear'
<i>stringa</i> 'string, lace'	1437	<i>stringa</i>	<i>stringa</i>	<i>stringere</i> 'to string'
<i>vaglia</i> 'worth, value'	1350	<i>valga</i>	<i>valga</i> , but † <i>vaglia</i>	<i>valere</i> 'to be worth'
<i>voglia</i> 'longing, want'	1250	<i>voglia</i>	<i>voglia</i>	<i>volere</i> 'to want'

\*According to Girauo et al. (2009: 3), but cf. footnote 10.

(ii), if there is no homonymy between the modern noun and the verbal stem on which modern standard subjunctive forms are constructed, there is homonymy between the modern noun and an ancient one (cf. *vaglia* in table 1; cf. e.g. Rohlfs 1968: 297 for the Old Italian subjunctive form) or a modern alternative verbal stem form (cf. *doglia* in table 1; for the alternative

<sup>7</sup> Verb plus noun compounds instead, that are another controversial issue in Italian word-formation research (for an overview of the competing approaches see e.g. Bisetto 2004), seems to be fed by stem 3 of the Italian verbal stem space (for the numbering see Girauo et al. 2009: 3; for an analysis of French composition see Villouing 2010), which is the same stem on which e.g. the Italian imperative (2<sup>nd</sup> person singular) is construed: cf. *lanciafiamme* 'flame-thrower' (1<sup>st</sup> conjugation), *rompicapo* 'problem' (2<sup>nd</sup> conjugation) and *apriscatole* 'tin-opener' (3<sup>rd</sup> conjugation). This analysis is in line with propositions according to which the verbal element only has the form, but not the meaning of an imperative (e.g. Rainer 2001: 389).

<sup>8</sup> All words in this table apart from *doglia*, *vaglia* and *voglia* are also listed in Thornton (2004: 518), Rohlfs (1969: 473) and Tollemache (1954: 163).

verbal inflectional forms, cf. e.g. Zingarelli Online, *dolere*).<sup>9</sup>

TABLE 2 – Homonymy of Old Italian nouns with stem 2 and present subjunctive forms

Old Italian Noun	Modern Italian Noun	Modern Standard Verbal Stem 2	Present Subjunctive Form (1 <sup>st</sup> to 3 <sup>rd</sup> singular)	Infinitive
† <i>buglia</i> 'tumult, confusion' (cf. Tollemache 1954: 163)	<i>subbuglio</i>	<i>bolla</i>	<i>bolla</i> , † <i>buglia</i> (cf. e.g. (8))	<i>bollire</i> 'to boil' † <i>bullire</i>
† <i>convegna</i> 'gathering, meeting' (cf. Tollemache 1954: 163)	<i>convegno</i>	<i>convenga</i>	<i>convenga</i> , † <i>convegna</i> (cf. Rohlfs 1968: 297 on <i>venire</i> and OVI)	<i>convenire</i> 'to get together'
† <i>dorma</i> 'sleeping, sleep' (cf. Tollemache 1954: 163)	<i>sonno</i>	<i>dorma</i>	<i>dorma</i>	<i>dormire</i>
† <i>fugga</i> 'escape, flight' (cf. Zingarelli Online <i>fuga</i> )	<i>fuga</i>	<i>fugga</i>	<i>fugga</i>	<i>fuggire</i> 'to flee'
† <i>ritegna</i> 'reservation'	<i>ritegno</i>	<i>ritenga</i>	<i>ritenga</i> , † <i>ritegna</i> (cf. Rohlfs 1968: 297 on <i>tenere</i> and OVI)	<i>ritenere</i> 'to retain'
† <i>sieda</i> 'chair' (cf. Zingarelli Online <i>sedia</i> )	<i>sedia</i>	<i>sieda</i>	<i>sieda</i>	<i>sedere</i> 'to sit'

(iii), nouns such as (3a) *tema* were formed in very old stages of the Italian language (cf. table 1). In those stages, there were other nouns of the same type that have since undergone important changes, such as †*fugga* and †*sieda* (cf. table 2) or that have since been replaced by other words, such as †*buglia* (cf. table 2), but all of them correspond, in their ancient form,

<sup>9</sup> In their description and definition of the Italian verbal stem space Pirrelli & Battista (2000) also account for doublets in Italian standard inflection, as e.g. for the *devo/debbo* (1<sup>st</sup> SG.IND.PRES.; 'I must') alternation (cf. e.g. 2000: 328-329) by establishing two stem spaces for the same verb. However, in their "overall distribution schema" (2000: 337) these parallel stems are no longer included, neither are they in Giraudo et al. (2009: 3), whose stem numbering is used throughout this paper. Moreover, neither Pirrelli & Battista (2000) nor Giraudo et al. (2009) take into account stems on which Old Italian and modern Italian dialectal forms are constructed. Consequently, as there is no complete description of the verbal stem spaces of other varieties than standard Italian available at the moment, and as even the above-mentioned stem numbering system abstracts away from parallel (and potentially differently organized) standard stem spaces, the numbering I use here might turn out to be incompatible with data from Old Italian and modern dialects once their stem spaces are fully described.

either to the still existing stem 2 and subjunctive form (*fugga, sieda*) or to an ancient subjunctive form, such as *buglia* that can be found in the OVI corpus (= *Corpus OVI dell'italiano antico*), as in (8):

(8) Hit number 4 in OVI for: form = *buglia* (*Ricette mediche bolognesi*  
14th cent.)

[...] *One cosa buglia in lo vino dolce e*  
*every thing boils in the wine sweet and*  
(3RD PERS.SG.PRES.SUBJ.)

*façasse* *l'impiastro [...]*  
*makes.itself the ointment*  
(3RD PERS.SG.PRES.SUBJ.)

'everything has to boil in sweet wine and the ointment has to emulsify'

The data presented in tables 1 and 2 clearly speak in favour of the stem 2 conversion analysis. Even if from a synchronic point of view the group (3a) *tema* consists only of very few members (cf. table 1), it is quite safe to assume that there are, at least in Old Italian, more nouns of type (3a) *tema* (cf. table 2) than commonly cited (e.g. by Thornton 2004: 518, Rohlfs 1969: 473 and Tollemache 1954: 163). In order to fully evaluate the degree of productivity of the stem 2 conversion pattern in Old Italian, it will, however, be necessary to complement the thorough description of the verbal stem space of modern standard Italian presented in Pirrelli & Battista (2000) and Giraudo et al. (2009) by the description of the verbal stem spaces of Old Italian varieties. In addition, an extensive and systematic corpus search in an Old Italian corpus (e.g. OVI) should be realized in order to increase the data group of type (1a). Only these steps would allow us to ultimately verify the stem 2 conversion hypothesis presented in this section.

### 3.2. *Basta*: Root conversion remodelled by blocking?

In contrast to (3a) *tema*, (3b) *basta* cannot be analysed as conversion of stem 2, as stem 2 of the verbal stem space of *bastire* is *bastisca* (and consequently, the noun should be *\*bastisca* (FEM.SG), which does not exist). It cannot be an instance of conversion of the infinitival stem (= stem 12)



either, because in this case we should expect the hypothetical nominal form \**basti* (FEM.SG) (cf. Thornton’s analysis of (2a) *sosta* in section 2.2.1). Nor can (3b) *basta* be a simple instance of root conversion as, in this case, it should display a masculine default –o (cf. Thornton’s analysis of type (1a) *cammino*). A plausible hypothesis is, instead, that (3b) *basta* is an instance of root conversion that does not have the default ending –o because the ending was blocked by an older homonym ending in –o. An observation that pleads in favour of this analysis is that at the time of the formation of words such as (1b) *basta* there were already corresponding forms in –o in the Italian lexicon (cf. table 3, *basta*, *falla*, *scherma*).

TABLE 3 – Blocking of nouns in –o by older homonyms?

Root conversion + integration into –a class	Verb	Modern Italian Stem 2	1 <sup>st</sup> Noun Occurrence (Zingarelli Online)	Root conversion + integration into –o class	1 <sup>st</sup> Occurrence (Zingarelli Online)
<i>basta</i> 'stitch, hem'	<i>bastire</i> 'to stitch'	<i>bastisca</i>	1612	<i>basto</i> 'burden, weight'	1305
<i>falla</i> 'leak'	<i>fallire</i> 'to fail, to make an error' <i>fallare</i> 'to make an error, to fail'	<i>fallisca</i>	1612	<i>fallo</i> 'failure, error'	1294
<i>scherma</i> 'fencing'	<i>schermire</i> 'to fence'	<i>schermisca</i>	1601	<i>schermo</i> 'shield'	1266
<i>scherna</i> 'mock, mockery'	<i>schernire</i> 'to mock, to deride'	<i>schernisca</i>	1348	<i>scherno</i> 'mock, mockery'	1336

As the concepts designated by the newer words are not identical (and in some cases not even related) to those designated by the older words ending in –o, the difference in meaning might have been signalled by a difference in form. An economic way to realize this conceptual non-correspondence linguistically is to integrate the new noun into a different noun class, in

the case of (1b) *basta* the prototypically feminine –a class instead of the prototypically masculine –o class (for the role of gender alternation in Italian word formation see e.g. Koch 2001).

This could, in principle, be a plausible explanation for the examples in table 2 not ending in the expected default –o, although the data do not allow, for the time being, for sound generalisations: on the one hand, the size of the (3b) group is simply too small and has to be complemented by more data; on the other hand, the case of *falla* in table 3 might be even more complicated because of the concurring base verbs that appeared at about the same time in Old Italian (cf. Zingarelli Online *fallare* and *fallire*). Moreover, if more cases like *scherna* were found, they might falsify the hypothesis altogether: in this case, not only the masculine and the feminine word appeared at about the same time in written texts (cf. Zingarelli Online), but they are even synonymous.<sup>10</sup> Besides, even if it was totally sure that the integration into the –o noun class of type (1b) *basta* has been blocked by existing older homonyms, we would have to explain why root conversion has, in the first place, been preferred to conversion of stem 2 (\**bastisca*). There are several plausible explanations: (i) at the time of the formation of type (3b) words (probably in the 16th/17th century, cf. the first occurrences in table 3), conversion of stem 2 was no longer a productive pattern in Italian; (ii), as existing forms attract the formation of formally similar lexical neighbours (cf. e.g. Rainer 1997) even if they are not semantically related, the already existing –o words might have facilitated the formation of root converts with subsequent integration into the –a class because of the above-mentioned need of differentiation; (iii) there might have been a need to avoid suffixes containing /sk/ (present in the respective stem 2 endings in table 3) that is usually associated with certain diminutive forms (cf. e.g. *pulviscolo* ‘fine dust’; cf. e.g. Merlini Barbaresi 2004) because of a certain risk of ambiguity.

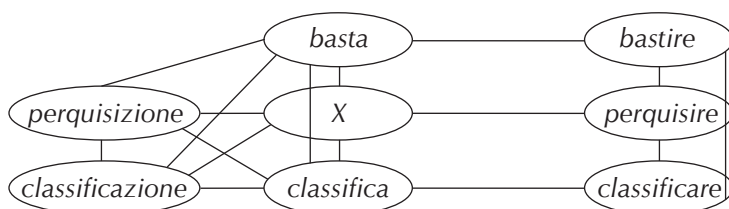
### 3.3 *Perquisa*: A product of analogy to existing word-family network relations

According to Thornton (2004: 518) (3c) *perquisa* is an instance of back-formation from *perquisizione* created in analogy to instances such as

<sup>10</sup> A closer look for other gender doublets and at their semantics and distribution is left for future research, cf. section 4.

revoca (2b), that has been coined in administrative language (cf. sections 2.1 and 2.2.1). Her main argument is that the word (3c) *perquisita* originated in prisoner slang in which the administrative language is well-known and frequent (Thornton 2004: 520). However, from a synchronic cognitive perspective there is no difference between *sosta* (2a) and *classifica* (2b): both are directly related to the verb (Marzo 2013: 47; Umbreit 2011; cf. 2.1 and 2.2.1). Moreover, native speakers do not necessarily know ancient word-formation rules of their language that have, in addition, produced only very few words (cf. 2.1 for the group size). It is thus highly questionable that (3c) *perquisita* really has been formed by analogy to type (2b) *classifica*. However, *perquisita* (3c) might still be an analogy-driven formation, but the analogy is more likely to have been drawn to synchronic word-family and network relations than to a diachronic word formation process.

FIGURE 1 – An empty slot X in a network for type (1c) *perquisita*



As both the series *basta* – *bastire* and *classificazione* – *classifica* – *classificare* exist in the speaker’s mind and are connected by paradigmatic relations, the formation of *perquisita* might have been possible thanks to – so to speak – an empty slot X in a network, such as in figure 1.

Interestingly, *perquisita* is not the only instance of type (3c) neither from a diachronic, nor from a synchronic perspective. Italian *dorma* ‘sleeping, sleep’ that is in disuse today according to Tollemache (1954: 163) and *dormizione* that has, according to Zingarelli Online, in addition to a religious meaning also the archaic meaning ‘sleeping, sleep’ are related by exactly the same relations to their infinitive *dormire* ‘to sleep’ as (3c) *perquisita* and *perquisizione* are – the only difference being that *dorma* can

be analysed as an instance of stem 2 conversion such as type (3a) *tema*, while *perquisita* cannot. From a synchronic point of view, too, *perquisita* is not alone, as an explorative search on google for short forms of nouns in *-izione* having an infinitive in *-ire* has shown (cf. table 4). An open question for further research is thus whether the increased (and increasing?) amount of exemplars of nouns of type (1c) *perquisita* that are characterized by nominal *-a*, infinitival *-i-* and the existence of a synonymous noun in *-izione* is leading or already has led to a new schema in the sense of Bybee (2010) or Langacker (1987, 1991).

TABLE 4 – Neologisms of type (1c)

Neologisms	Google hits	Infinitives	Suffixed Nouns
<i>acquisa</i> 'acquisition'	3 (04/23/13)	<i>acquisire</i> 'to acquire'	<i>acquisizione</i> 'acquisition'
<i>inquisita</i> 'inquisition'	7 (04/23/13)	<i>inquisire</i> 'to inquire'	<i>inquisizione</i> 'inquisition'
<i>disquisita</i> 'too detailed and long discussion'	36 (04/23/13)	<i>disquisire</i> 'to discuss in detail'	<i>disquisizione</i> 'too detailed and long discussion'

Results of a search for *la* (DET.ART.) + *Xizione*

#### 4. Conclusion and outlook

This paper focused on a special Italian noun type characterized by nominal *-a* and infinitival *-e-/-i-* as in (3a) *tema*, (3b) *basta* and (3c) *perquisita*, a noun type that is traditionally considered as marginal (cf. the few examples in e.g. Tollemache 1954, Rohlf's 1969, Thornton 2004) with respect to its counterparts in (1) and (2) (nominal *-o* plus infinitival *-a-/-e-/-i-* and nominal *-a* plus infinitival *-a-* respectively). In section 2 it was shown that traditional accounts (conversion of the infinitival stem, root conversion and suffixation) all fail in explaining the totality of the data types in (1) to (3) convincingly. Especially the analyses of and comments on type (3) turned out to be deceiving. An alternative analysis was then proposed in section 3: Type (3) was divided into three subgroups. Type (3a) (cf. section 3.1) was shown to be a remnant of an Old Italian conversion type, more precisely conversion of stem 2 of the Italian verbal stem space. The instances in group (3b) (cf. section 3.2) were hypothesized to be cases of root conversion that

have been integrated into the –a (SG)/–e (PL) noun class because of blocking of the –o (SG)/–i (PL) class by older homonyms. As for (3c) (cf. section 3.3), Thornton's (2004) analogy-driven back-formation analysis was questioned and a network-based (but still analogy-driven) alternative was presented. All in all, it has been shown that the noun type (3) is not that marginal if we take into account data from other Italian varieties than just modern standard data.

While in this paper the main focus was on the verbal stem space of modern standard and Old Italian data, future research concerned with the role of the verbal stem space for Italian word-formation in general and verb to noun conversion in particular could additionally take into account modern dialectal data, because the homonymy of nouns with alternative verbal stems suggests (cf. *doglia* in table 1, section 3.1) that other synchronic varieties might hide conversion bases. In addition, formal irregularities of type (1) such as the stem alternation in (1b) *ritegno* 'reservation' – *ritenere* 'to retain' ([ɲ] – [n]) (cf. Thornton 2004: 516) that seem inexplicable from the perspective of the modern standard Italian verbal stem space, might be better understood when systematically compared against the modern dialectal verbal stem space that contains [ɲ] as well as [n], as e.g. in (9a). Interestingly, the same phenomenon can, again, be observed in Old Italian dialects, cf. (9b). More precisely, [ɲ] seems to appear systematically in those inflectional forms that are constructed – again – on the verbal stem 2 (1st to 3rd sing. subj. ind. and 1st sg. and 3rd pl. pres. ind.).

- (9) a. Sicilian, Yahooogroup *Linguasiciliana*, message n° 4311, 23.04.13:

**Ritegnu** ca nn'avemu nu prubbrema quannu nu pòpulu nun  
think (1ST.SG.PRES.IND.) that we have a problem if a people not  
*si po lèggiri la so littiratura.*  
itself can read the its literature.  
'I think we have a problem if a people cannot read its own literature'

- b. Old Tuscan: Hit number 2 in OVI for form = *ritegna* (*Legg. G. di Procida*, 13th cent.)

*miseri scacciati de regno di Cicilia e di Puglia che non*  
poor chased away from kingdom of Sicily and of Apulia who not

*truovano albergo che li ritegna, [...]*

find inn that them keeps (3RD SG.PRES.SUBJ.)

‘the poor guys chased away from the kingdom of Sicily and Apulia  
who don’t find an inn that allows them to stay’

All in all, the following remaining questions will best be tackled by a systematic and extensive corpus research of different Italian varieties: first, the exact impact of the verbal stem space of Old Italian and modern Italian dialects (more data are needed) on verb to noun conversion and other word-formation phenomena (cf. section 3.1); second, the blocking hypothesis for type (3b) *basta* (cf. section 3.2); third, the distribution of –o and –a and therefore also of gender doublets in Old Italian, where, in addition to gender doublets for type (3b) *basta* (cf. *basto*) doublets can be found across the borders of other noun types, such as e.g. (1b) *ritegno* and (3a) *ritegna* (cf. table 2); last but not least, the existence of such noun pairs raises serious doubts about whether Thornton’s root conversion hypothesis can be held up for type (1) in general.

#### REFERENCES

- Bisetto, A. 2004. La composizione. In: M. Grossmann; F. Rainer (Eds.). *La formazione delle parole in italiano*. Tübingen: Niemeyer, 33-51.
- Blank, A. 1998. Kognitive italienische Wortbildungslehre. *Italienische Studien*. 19: 5-27.
- Bonami, O. & Boyé, G. 2002. Suppletion and dependency in inflectional morphology. In: F. van Eynde; L. Hellan; D. Beerman (Eds.). *Proceedings of the HPSG’01 Conference*. Stanford: Stanford CSLI publications, 51-70.
- Bonami, O. & Boyé, G. 2003. Supplétion et classes flexionnelles dans la conjugaison du français. *Langages*. 15: 102-126.
- Bonami, O., Boyé, G., Giraudo, H. & Voga, M. 2008. Quels verbes sont réguliers en français?. In: J. Durand; B. Habert; B. Laks (Eds.). *Actes du Congrès Mondial de Linguistique Française*. Paris: Institut de Linguistique Française, 1511-1523.
- Bonami, O., Boyé, G. & Kerleroux, F. 2009. L’allomorphie radicale et la relation flexion-construction. In: B. Fradin; F. Kerleroux; M. Plénat (Eds.). *Aperçus de morphologie du*

- français*. Saint-Denis: Presses Universitaires de Vincennes, 103-125.
- Bybee, J. 2010. *Language, Usage and Cognition*. Cambridge: Cambridge UP.
- Dardano, M. 1988. Formazione delle parole. In: G. Holtus; M. Metzeltin; C. Schmitt (Eds.). *Lexikon der romanistischen Linguistik*. Vol. IV. Tübingen: Niemeyer, 51-63.
- Dressler, W.U. & Thornton, A.M. 1996. Italian nominal inflection. *Wiener linguistische Gazette*. 57-59: 1-26.
- Fradin, B. 2009. Romance: French. In: R. Lieber; P. Štekauer (Eds.). *Oxford Handbook on Compounding*. Oxford: Oxford University Press, 417-435.
- Gaeta, L. 2010. On the viability of cognitive morphology for explaining language change. In: A. Onysko; S. Michel (Eds.). *Cognitive Perspectives on Word-Formation*. Berlin; New York: De Gruyter, 75-95.
- Giraud, H., Montermini, F. & Pirrelli, V. 2009. Processi cognitivi nell'analisi delle classi verbali dell'italiano: un approccio sperimentale. In: P.M. Bertinetto; V. Bambini; I. Ricci (Eds.). *Linguaggio e cervello / Semantica. Atti del XLII Congresso della Società di Linguistica Italiana. Pisa: Scuola Normale Superiore 25-27 settembre 2008*. Vol. 2. Roma: Bulzoni, CD-Rom.
- Koch, P. 2001. Lexical typology from a cognitive and linguistic point of view. In: M. Haspelmath; E. König; W. Oesterreicher; W. Raible (Eds.). *Language Typology and Language Universals. An International Handbook*. Vol. 2. Berlin; New York: De Gruyter, 1142-1178.
- Langacker, R. 1987-1991. *Foundations of Cognitive Grammar*. 2 vols.. Stanford: Stanford University Press.
- Maiden, M. 2003. Il verbo italo-romanzo: verso una storia autenticamente morfologica. In: A. Rocchetti; M. Giacomo-Marcellesi (Eds.). *Il verbo italiano. Studi diacronici, sincronici, contrastivi, didattici*. Roma: Bulzoni, 3-21.
- Marzo, D. 2013. *Polysemie als Verfahren lexikalischer Motivation. Theorie und Empirie am Beispiel von Metonymie und Metapher im Französischen und Italienischen*. Tübingen: Narr.
- Meinschaefer, J. 2005. *Deverbale Nominalisierungen im Französischen und Spanischen*. Konstanz: Universität Konstanz, habilitation thesis.
- Merlini Barbaresi, L. 2004. L'alterazione. In: M. Grossmann; F. Rainer (Eds.). *La formazione delle parole in italiano*. Tübingen: Niemeyer, 286-293.
- Nolda, A. 2012. *Konversion im Deutschen – Muster und Beschränkungen. Mit einem Grundriss einer allgemeinen Theorie der Wortbildung*. Berlin: Philosophische Fakultät der Humboldt-Universität zu Berlin, habilitation thesis.

- OVI = *Corpus OVI dell'italiano antico*. Retrieved June 19, 2013, from the World Wide Web: [gattoweb.oivi.cnr.it](http://gattoweb.oivi.cnr.it).
- Pirrelli, V. & Battista, M. 2000. The paradigmatic dimension of stem allomorphy in Italian verb inflection. *Rivista di Linguistica*. 12(2): 307-380.
- Rainer, F. 1993. *Spanische Wortbildungslehre*. Tübingen: Niemeyer.
- Rainer, F. 1997. I meccanismi di diffusione di allomorfi: il caso della rianalisi del tipo latino *man-u-alis* in italiano. In: P. Benincà; A. Mioni; L. Vanelli (Eds.). *Fonologia e morfologia dell'italiano e dei dialetti d'Italia. Atti del XXXI congresso della Società di Linguistica italiana (Padova 25-27 settembre 1997)*. Roma: Bulzoni, 423-452.
- Rainer, F. 2001. Compositionality and paradigmatically determined allomorphy in Italian wordformation. In: C. Schaner-Wolles; J. Renniso; F. Neubarth (Eds.). *Naturally! Linguistic Studies in Honour of Wolfgang Ulrich Dressler Presented on the Occasion of His 60th Birthday*. Torino: Rosenberg & Sellier, 383-392.
- Rodrigues, A. 2009. Portuguese converted deverbal nouns: constraints on their bases. *Word Structure*. 2(1): 69-107.
- Rohlf, G. 1968-1969. *Grammatica storica della lingua italiana e dei suoi dialetti*. Vol. 2-3. Roma: Einaudi.
- Scalise, S. 1994. *Morfologia*. Bologna: Il Mulino.
- Thornton, A.-M. 2004. Conversione. In: M. Grossmann; F. Rainer (Eds.). *La formazione delle parole in italiano*. Tübingen: Niemeyer, 500-553.
- Thornton, A.-M. 2003. L'assegnazione del genere in italiano. In: F. Sanchez Miret (Ed.), *Actas del XXIII Congreso Internacional de Lingüística y Filología Románica*. Tübingen: Niemeyer, 467-481.
- Tollemache, F. 1954. *I deverbali italiani*. Firenze: Sansoni.
- Tribout, D. 2010. *Les conversions de nom à verbe et de verbe à nom en français*. Paris: Université Denis Diderot (Paris VII), thèse de doctorat.
- Tribout, D. 2012. Verbal stem space and verb to noun conversion in French. *Word Structure*. 5(1): 109-128.
- Umbreit, B. 2011. Motivational networks. An empirically supported cognitive phenomenon. In: K.-U. Panther & G. Radden (Eds.). *Motivation in Grammar and the Lexicon*. Amsterdam; Philadelphia: John Benjamins, 269-286.
- Villoing, F. 2010. Les mots composés VM. In: B. Fradin; F. Kerleroux; M. Plénat, (Eds.) *Aperçus de morphologie du français*. Saint-Denis: Presses Universitaires de Vincennes, 175-197.



Zingarelli Online = *Lo Zingarelli 2013. Vocabolario della lingua italiana*, Bologna, Zanichelli.



# Is conversion a syntactic or a lexical process of word formation?

Alexandra Soares Rodrigues

afs@ipb.pt

*Escola Superior de Educação de Bragança – Instituto Politécnico de Bragança*

*CELGA – Universidade de Coimbra (Portugal)*

**ABSTRACT.** Conversion is sometimes described as a syntactic phenomenon by which a lexical item changes its lexical category according to the syntactic environment where it is inserted. This syntactic-ordered approach comes from theoretical fields that conceive the lexicon as the domain of irregularity, whilst regular patterns are treated in syntax (Chomsky 1995).

However, Portuguese converted deverbal nouns (*remendo* ‘event of mending’, *curte* ‘event of having fun’, *trinca* ‘event of biting’) manifest a structural behaviour that permits us to situate their formation in the lexicon instead of in the syntax. According to the theoretical allusion we made above, this would characterise converted deverbal nouns as lacking a regular pattern. However, what we mean is that the lexicon is not the field of irregularity. Apart from the irregular material that must be stored in long term memory as to be used by speakers, such as inherited lexemes (e.g. *rato* ‘mouse’, *cão* ‘dog’, *rir* ‘to laugh’, *verde* ‘green’, etc.), the lexicon is the domain of word formation, which is constraint-based (Rodrigues 2008, 2009). This means the word formation part of the lexicon is constrained by regular patterns that are neither directional in principle, nor syntactic in nature.

We follow Jackendoff (2002) conception on the lexicon, conceiving it an interface of syntax, phonology and semantics.

Converted deverbal nouns formation seems to agree with this conception, since it depends on phonological, semantic and syntactical constraints (Rodrigues 2004, 2009). Portuguese verb-into-noun conversion is not a simple case of syntactic environment. This is specially visible when we confront this lexical conversion with a purely syntactic type of nominalisation (Kerleroux 1996), such as the one that occurs in *O estudar matemática traz-me vantagens*. ‘Studying maths brings me advantages’ or *O remendar roupa é um recurso nesta época*. ‘Mending cloths is a good resource nowadays’. Apart from the syntactic constraints that converted deverbal nouns must follow (e.g. as real nouns they cannot directly theta-mark, whilst syntactic nominalisations can), these nouns display phonological constraints in their formation and a semantic variability that is not observable in syntactic nominalisations. This semantic variability is not exclusively explainable by the syntactic environment. Moreover, the formation of converted deverbal nouns obeys

to structural constraints that do not act upon syntactic nominalisation. All this means that conversion is not syntactic in nature. Yet it is not irregular either. This evidences that the mainstream generative grammar conception on the lexicon should be abandoned.

KEY-WORDS. lexical conversion, nominalisation, Portuguese word formation, deverbal nouns.

## 1. Introduction

We intend to demonstrate that V to N conversion, following the process formation of [*abraç*]<sub>V</sub>[*o*]<sub>N</sub>, is not a syntactic process. For that, we will distinguish lexical conversion from syntactic nominalisations.

We follow the parallel architecture of Jackendoff (2002). According to this, the lexicon is an interface between syntax, phonology and semantics. We consider word formation a lexical domain with its own rules, which are basically constraints, which are not dependent on syntactic principles.

We believe the lexicon is not the domain of irregularity, as Chomsky proposes. According to Chomsky (1981), grammar is divided into the core and the periphery. According to Chomsky, only regular matters are the object of linguistics. However, this is to say that many linguistic phenomena are put aside. The problem is that every language contains irregular material, and people acquire and use that material as proficiently as they use the regular one.

Moreover, following Chomsky's proposal of focusing on the core of the grammar and leaving aside the periphery brings another risk: whenever the theoretical model fails to explain some phenomenon, it is very tempting to leave it to the periphery. Nevertheless, one should not decide of the 'peripheral character' of the object without having tested the effectiveness of the model.

What have these theoretical and philosophical considerations to do with conversion?

Conversion is sometimes characterised as an irregular domain. We think it is not so irregular as that, we would say, since it obeys to structural constraints, belonging to phonological, syntactic and semantic domains (Rodrigues 2004; Rodrigues 2009). This means that conversion does neither generate every hypothetical word, nor it is subdued to overgeneration.

Since word formation is not syntax, or it is not reducible to syntax, we need to observe that object with appropriate lens. Word formation is only periphery if analysed by the wrong lens. Otherwise, it will be explainable.

Thus, we have three steps to track:

1<sup>st</sup>: is conversion a syntactic process? Or does it involve other structures? Do syntactic principles explain every part of conversion phenomenon? No. It is not explainable by syntax principles.

2<sup>nd</sup>: if it is not syntax, is it irregular, that is, periphery, following Chomsky's label? No. It is not syntax, but, in spite of it, it is not irregular. On the contrary, converted deverbal nouns (CDN) formation has a regular pattern.

3<sup>rd</sup>: So, what is it? It is word formation at a lexical level, with its own rules, with its own regularity.

## 2. Defining lexical categories

One problem that emerges when dealing with conversion is the definition of lexical categories. For this, we will follow Baker (2003: 17). Since we are concerned with the distinction between verbs and nouns, we make only reference to these two categories. Baker (2003: 20) defines verbs as a lexical category that has a specifier («Only verbs are true predicates, with the power to license a specifier, which they typically theta-mark.»).

In what concerns nouns, these are defined by Baker (2003: 96) as having a referential index. Since Baker proposal is to define lexical categories according to their differences, we have here the criteria that permit us to distinguish verbs from nouns: «No syntactic node can both license a specifier and bear a referential index.» (Baker 2003: 96).

Bearing a referential index is dependent on the criterion of identity. According to Geach (1962), only a noun can occur in the frame "X is the same \_\_\_ as Y".

Because of this identity criterion, nouns are the only category that may occur with determiners, bear plurality and quantification, and bear a relationship of coreference, such as anaphora.

We will evaluate nominalisations using those different syntactic structures that depend on the identity criterion and see what they will evidence about them.

### 3. Comparing syntactic nominalisation with lexical conversion

#### 3.1. The identity criterion

According to Baker (2003), only nouns have the criterion of identity. This means that only nouns may occur in sentences such as “X é o mesmo \_\_\_\_\_ que Y” (‘X is the same \_\_\_\_\_ as Y’).

This is true for non-constructed nouns such as *lápiz* ‘pencil’ or *carro* ‘car’, as examples (1) show. (Every sentence was attested with 60 naive native speakers.)

- (1) a. *Este lápis é o mesmo que vimos ontem.*  
‘This pencil is the same we saw yesterday.’
- b. *Este carro é o mesmo que vimos ontem.*  
‘This car is the same we saw yesterday.’

In what concerns event deverbal nouns, the same construction is possible if the *denotatum* has ontological conditions that fit it, for instance, if it means a type and not an instance. This is because events have ontological limits that coarct their durability in time (example 2).

- (2) \**Este é a mesma plantação de árvores a que assistimos ontem.*  
‘This is the same planting of trees we have attended yesterday.’

Event nouns may only appear in this structure if they designate an event with no interruption along time. Thus, we consider that this is not a good criterion to measure nominality.

#### 3.1.1. Determiners

Because of having a referential index, nouns are syntactically related with determiners and quantifiers.

Thus, according to Baker, the occurrence of determiners is only bearable

by nouns. This is true for deverbal nouns and for infinitive nominalisations (examples 3).

- (3) a. *O apodrecimento da fruta.*  
'The rot of the fruit.'
- b. *O apodrecer a fruta/ o apodrecer da fruta.*  
The rot-INF the fruit/ the rot-INF of.the fruit.  
'Rotting the fruit./ The rotting of the fruit.'
- c. *A apanha da fruta.*  
'The gathering of the fruit.'

### 3.1.2. Plurality

Plurality is another criterion that encompasses nouns. Notice that when verbs have number morphemes, these are a syntactic consequence of agreement with a noun projection (Baker 2003: 107). Let us examine what happens with pluralising verbs (4a, 5b, 6a), comparing them with deverbal nouns (examples 4b, 5 a, 6c).

- (4) a. *\*Os remendares são cuidadosos.*  
The-PL mend-INF-PL are careful  
'The mendings are careful.'
- b. *Os remendos a que assistimos são cuidadosos.*  
The-PL mend-CDN-PL to which attend-1<sup>st</sup> PL are careful  
'The mendings we have attended are careful.'
- c. *O remendar a que assistimos é cuidado.*  
The mend-INF to which attend-1<sup>st</sup>PL is careful.  
'The mending we have attended is careful.'
- (5) a. *As reconstruções da cidade foram trabalhosas.*  
'The reconstructions of the city were arduous.'

- b. \**Os reconstruíres da cidade foram trabalhosos.*  
The-PL reconstruct-INF PL of.the city were arduous  
'\*The reconstructings of the city were arduous.'
- (6) a. \**Estamos a assistir aos saltares em série dos atletas.*  
Be-1stPLto attend to.the.PL jump-INF PL in series of.the-PL  
athletes  
'\*We are attending to the jumps in series of the athletes.'
- b. *Estamos a assistir aos saltos em série dos atletas.*  
Be-1st PL to attend to.the-PL jumps-PL in series of.the athletes  
'We are attending to the jumps in series of the athletes.'
- c. *Estamos a assistir ao saltar do atleta.*  
Be-1stPL to attend to.the jump-INF of.the athlete  
'We are attending to the athlete jumping.'

Examples (4-6) show that, although there had been a nominalisation, verbs cannot undergo pluralisation, contrarily to CDNs and suffixed deverbal nouns.

This seems to evidence that when changes occur at the morphological level, they are not accepted, whilst if changes are merely syntactic, they are accepted, in infinitive nominalisations. What is inside the scope of the word in itself is not accepted, only what is outside. This make way to the hypothesis that nominalisation is not the same as conversion and, more than that, conversion is at a different level than nominalisation – it occurs in the lexicon.

### 3.1.3. Quantifiers

Quantifiers evidence the criterion of identity focused on by Baker in relation to nouns.

In what concerns syntactic nominalisations, quantifiers are restricted, since pluralisation also is.

- (7) a. \**Os dois saltares a que assistimos foram consecutivos.*  
The-PLtwo jump-INF.PL to which attend-1stPL be-3rdPL



### consecutive-PL

‘\*The two jumpings we attended to were consecutive.’

b. *Os dois saltos a que assistimos foram consecutivos.*

The.PL jumps-N PL to which attend-1stPL be-3rdPL consecutive  
‘The two jumps we attended to were consecutive.’

(8) a. \**Os dois abraçares foram sinceros.*  
The-PL two hug-INF PL be-3rdPL sincere  
‘The two huggings were sincere’.

b. *Os dois abraços foram sinceros.*  
The-PL two hugs-N PL be-3rdPL sincere  
‘The two hugs were sincere.’

#### 3.1.4. Anaphora

Another criterion that provides nominality is anaphora (Baker 2003: 127-128). This is true because «the antecedent of a pronoun must be a referential-index-bearing noun projection.» (p. 128). See examples (9-11) below. From these examples we may see that infinitive nominalisations do not accept anaphora.

(9) *Isto é o bolo. O João fê-lo ontem.*  
‘This is the cake. John made it yesterday.’

Anaphora is possible with event nouns (10a, 11c) but it is not grammatical with infinitive nominalisations (10b; 11a,b):

(10) a. *Isto é a comemoração. O João organizou-a ontem.*  
‘This is the commemoration. John organised it yesterday.’  
b. \**Isto é o comemorar. O João organizou-o ontem.*  
This is the commemorating-INF. The John organised it yesterday  
‘\*This is the commemorating. John organised it yesterday.’

- (11) a. *\*Isto é o apanhar a fruta. O João organiza-o bem.*  
 This is the gather-INF the fruit. The John organises-it well  
 ‘\*This is the gathering the fruit. John organises it well.’
- b. *\*Isto é o apanhar da fruta. O João organiza-o bem.*  
 This is the gather-INF of.the fruit. The John organises-it well  
 ‘This is the gathering of the fruit. John organises it well.’
- c. *Isto é a apanha da fruta. O João organiza-a bem.*  
 This is the gather-CDN of.the fruit. The John organises-it well.  
 ‘This is the gathering of the fruit. John organises it well.’

3.1.5. Conclusions on the criterion of identity and its syntactic consecutions

Table 1 summarises data exposed till now.

	Concrete nouns	Suffixed Deverbal nouns	Converted deverbal nouns	Syntactic Nominalisation
X is the same ___ as Y	+	--	--	-- depends on the semantics
determiners	+	+	+	+
quantifiers	+	+	+	-
pluralisation	+	+	+	-
anaphora	+	+	+	-

Table 1. Different nouns and nominalisations and its acceptability towards syntactic environments of nouns

Table 1. synthesises that syntactic nominalisations only bear determiners. They fail the criterion of identity with the structure “X is the same \_\_\_ as Y”, since this depends more on semantic typology. Infinitive nominalisations also fail plurality, quantification and anaphora.

These results evidence that a lexical category is not a simple question of syntactic environment.

Now we will add more syntactic occurrences that characterise verbs in general and specifically in Romance languages (Kerleroux 1996:70).

### 3.2. Clitics

Verbs have their complements actualised by clitics (examples 12a, b; 13a, b). Nouns do not (12 c, d, e, f; 13c, d).

(12) a. *O derrubar árvores é uma atitude insensata.*  
The cut.down-INF trees is a attitude senseless  
'Cutting down trees is a senseless attitude.'

b. *O derrubá-las é uma atitude insensata.*  
The cutt.down-INF-it-FEM PL be-1<sup>st</sup> SG a attitude senseless  
'Cutting down them is senseless.'

c. *O derrube de árvores é insensato.*  
The cut.down-CDN of trees is senseless  
'The cutting down of trees is senseless.'

d. *\*O derrube-as é insensato.*  
The cut.down-CDN-them is senseless  
'\*The cutting down them is senseless.'

e. *O derrubamento de árvore.*  
The cut.down-SUFDEVN of trees  
'The cutting down of trees.'

f. *\*O derrubamento-as.*  
The cut.down-SUFDEVN-them  
'\*The cutting down them.'

(13) a. *O desejar livros é uma atitude sensata.*  
The desire-INF books is a attitude sensible  
'Desiring books is a sensible attitude.'

b. *O desejá-los é uma atitude sensata.*  
The desire-INF-them is a attitude sensible  
'Desiring them is a sensible attitude.'

c. O *desejo* *de livros é sensato.*  
The *desire-CDN* *of books is sensible*  
'The desire for books is sensible.'

d. \*O *desejo-* *los é sensato.*  
The *desire-CDN-* *them is sensible*  
'\*The desire them is sensible.'

Examples (12) and (13) show that clitics may only occur with verbs, even if they are nominalised.

### 3.3. Theta-mark

Verbs may theta-mark directly (examples 14a, b), whilst nouns need a preposition (14d, e, f).

(14) a. O *João derrubou uma árvore.*  
'John cut down a tree.'

b. O *derrubar* *árvores é insensato.*  
The *cut.down-INF* *trees is senseless*  
'Cutting trees is senseless.'

c. O *derrubar* *de árvores é insensato.*  
The *cut.down-INF* *of trees is senseless*  
'The cutting down of trees is senseless.'

d. O *derrube* *de árvores é insensato.*  
The *cut.down-CDN* *of trees is senseless*  
'The cutting down of trees is senseless.'

e. \*O *derrube* *árvores é insensato.*  
The *cut.down-CDN* *trees is senseless*  
'\*The cutting down trees is senseless.'

f. \*O *derrubamento*                      árvores                      é *insensato*.  
 The cut.down-SUFDEVN                      trees                      is senseless  
 ‘\*The cutting down trees is senseless.’

### 3.4 Modification

Verbs are modified by adverbs (examples 15a, b), whilst nouns are modified by adjectives (example 15c).

(15) a. O [*saber*                      *pouco*] *apoquenta-a*.  
 The know-INF                      little                      scares- her  
 ‘Knowing little scares her.’

b. O [*passear ocasionalmente*] *entorpece as pernas*.  
 The walk-INF occasionally                      numbs the legs  
 ‘Walking occasionally numbs the legs.’

c. \*O [*passeio ocasionalmente*] *entorpece as pernas*.  
 The walk-CDN occasionally                      numbs the legs  
 ‘\*The walk occasionally numbs the legs.’

### 3.5 Sum up on verb vs. noun criteria

We have now conditions to add more details about verb vs. noun behaviour in table 2.

	Concrete nouns	Suffixed Deverbal nouns	Converted deverbal nouns	Syntactic Nominalisation [VP] <sub>DP</sub>
adverbs	-	-	-	+
direct theta-mark	-	-	-	+
complements as clitic	-	-	-	+

Table 2. Syntactic behaviour of verbs in Romance languages in nominalisations and nouns.

Table 2 shows syntactic occurrences of verbs in Romance languages. As we can see, only [VP]<sub>DP</sub> syntactic nominalisations admit those occurrences such as adverbs, complements as clitics and a direct theta-mark. Concrete

nouns, suffixed deverbal nouns and CDNs do not. We use determiner phrase following Brito (1993), instead of noun phrase. This is so because we defend there is no real noun, but a nominalization which is built by the determiner. Otherwise nominalisations would accept every criteria applied to nouns.

Comparing data in table 1 and 2, we may conclude that syntactic nominalisation detaches from other nominal occurrences. The last ones display a typical noun syntactic behaviour, whilst nominalised [VP] do not. Nominalised [VP] shows a typical syntactic verbal behaviour, whilst CDNs do not. As we will see in section 4, there is a kind of nominalisation that is in-between both poles, but it does not follow every N criteria.

#### 4. Are all syntactic nominalisations the same?

We want to demonstrate that conversion is not a syntactic nominalisation. However, first we need to observe if syntactic nominalisations of verbs are all the same (See Brito 2012 for different kinds of nominalisations.).

##### 4.1 CDNs

Converted deverbal nouns correspond to real nouns. They are not a syntactic nominalisation. As we will see in section 5, there are structural marks that point out this. Example 16 provides an example of a CDN.

- (16) *O João fez uma compra fabulosa.*  
'John made a fabulous purchase.'

##### 4.2. [VP]<sub>DP</sub>

Nominalisation of a verbal phrase (example 17) constitutes a kind of syntactic nominalisation.

- (17) *O [comprar livros] faz-me bem.* [VP]<sub>DP</sub>  
The buy-INF books do-3<sup>rd</sup>SG-me well  
'Buying books does me well.'

Syntactic behaviour that evidence that [VP]<sub>DP</sub> corresponds to a syntactic nominalisation operated by the determiner are focused on as following:

- i) It does not admit pluralisation (example 18):

- (18) \*Os *comprares livros fazem-me bem.*  
The-PL buy-INF PL books do-PL-me well  
'\*The buys books do me well.'

ii) It admits clitics (examples 19-20):

- (19) O *[abrir-se a porta] foi efeito do vento.*  
The openINF-itself-ANTCAUS the door was effect of.the  
wind  
'The opening of the door was the wind effect.'

- (20) O *[barbear-se o João] foi efeito do noivado.*  
The shave-INF-himself the John was effect of.the  
engagement  
'Shaving himself by John was the engagement effect.'

Notice that, when there is a clitic, a preposition to theta-mark is excluded (19-20). Clitics and preposition exclude each other. The clitic is verbal and the occurrence of Preposition, which serves the noun because this does not theta-mark directly, is nominal.

This shows that what is nominalised is a phrase [VP] and not a lexical item [V]. Notice also that a real lexicalised nominal infinitive excludes the clitic (see section 4.5).

#### 4.3. Metalinguistic nominalisation

Metalinguistic nominalisation is represented in examples (21a, b), which correspond to citations.

- (21) a. *Aquele 'não tenho tempo para ler' é falso.* [S]<sub>DP</sub>  
'That 'I have no time to read' is false.'
- b. O *'as armas e barões assinalados' está na memória de toda a gente.* [s]<sub>DP</sub>  
'The 'as armas e os barões assinalados' is in the memory of everybody.'

Metalinguistic nominalisations do not take obligatory lexicalisation. If so, we would always have to presuppose a noun for every word, phrase and sentence (Kerleroux 1996: 101).

#### 4.4. [V]<sub>DP</sub>

The nominalisation of the verb as a lexical item is exemplified in (22a, b, c, d, g, h).

- (22) a. *Aquele [caminhar] lento.* [V]<sub>DP</sub>  
That walk-INF slow  
'That slow walking.'
- b. *Aquele [observar] de aves.* [V]<sub>DP</sub>  
That watch-INF of birds  
'That watching of birds.'
- c. \**O [observar] lento aves.*  
The watch-INF slow birds  
'\*That slow watching birds.'
- d. *O [observar] lento de aves.* [V]<sub>DP</sub>  
The watch-INF slow of birds  
'The slow watching of birds.'
- e. *O [observar lentamente aves].* [VP]<sub>DP</sub>  
The watch-INF slowly birds  
'The slowly watching birds.'
- f. \**O [observar lentamente de aves].*  
The watch-INF slowly of birds  
'\*The slowly watching of birds.'
- g. \**O observá-las lento.*  
The watch-INF-them slow  
'\*The slow watching them.'



- h. \*Os observares de aves.  
The-PL watch-INF of birds  
'\*The watchings of birds.'

When there is a VP nominalisation, it may occur with adverb (22e) and not with adjective (22c) and it excludes preposition to theta-mark (22e, f). When there is a V nominalisation, it cannot occur with an adverb (22f), but with an adjective (22d) and it does not directly theta-mark (22c), but via preposition (22d). It does neither admit clitics (22g) nor pluralisation (22h).

#### 4.5. Lexicalised nominalisations

There are in the lexicon some nominalisations that have been lexicalised. This means there is an infinitive verbal form that, by means of its usage frequency, has undergone a process of blockage as a noun. This means that this form is not lexically coincident with the inflectional verb form; instead, it corresponds to a different lexical item, with its own conceptual and syntactic structures (Kerleroux 1996: 75). Examples of lexicalised nominalisations are provided in examples (23-27).

- (23) *Os olhares do João.*  
'John's glances.'

- (24) *O pôr do sol*  
'The sunset'

- (25) *O comer está na mesa.*  
'The meal is on the table.'

- (26) *O jantar está na mesa.*  
'Dinner is on the table.'

- (27) *O cantar de amigo.*  
'A kind of medieval song'

Lexicalised nominalisations are characterised by the following syntactic behaviours (i-vi):

i) it does not take clitics (examples 28a, c), contrarily to verbs (examples 28b, d):

(28) a. \**O pôr-se do sol.*  
The set-itself of.the sun  
'\*The sun-itself-set.'

b. *O sol pôs-se às 20:11.*  
The sun set-1st SING-itself at.the-PL 20:11  
'The sun set at 20:11.'

c. \**O jantar-se está na mesa.*  
The dinner-INF-NOM CL is on.the table  
'\*One having dinner is on the table.'

d. *Jantar-se às 21:00 é tarde?*  
Dinner-INF-NOM CL at.the-PL 21:00 is late  
'Is it late to have dinner at 21:00?'

ii) it takes pluralisation (*os olhares, os jantares*);

iii) it does not occur with adverbs (examples 29b), but with adjectives (examples 29a)

(29) a. *O jantar delicioso*  
'The delicious dinner'

b. \**O jantar deliciosamente*  
'\*The deliciously dinner.'

iv) it may be the base of a compound following the paradigm V+N (*guarda-chuva* literally 'save<sub>V</sub>-rain<sub>NN</sub>' 'umbrella', *corta-vento* literally 'cut<sub>V</sub>-wind<sub>NN</sub>' 'windcheater', *guarda-rios* literally 'save<sub>V</sub>-rivers<sub>NN</sub>' 'kingfisher'), such as demonstrated in the examples (30).

(30) a. *O guarda-jantares*

The keep-dinners

b. *O guarda-cantares*

The keep-songs

c. *O resguarda-olhares*

The keep.away-glances

Notice that constructions V + INFINITIVE such as the ones exemplified in (31) are not accepted.

(31) a. \**O guarda- passeares*

The keep- walk-INFINITIVE.PL

b. \**O guarda- pintares*

The keep- paintINFINITIVE.PL

c. \**O guarda- trepares*

The keep- climbINFINITIVE.PL

Examples (31) show that the infinitive form must be in the lexicon as such to function as a base in this form. Otherwise they will not integrate word-formation bases (Corbin 1987). Contrasting examples (30) with (31), we may check that only lexicalised nominalisations have occurrences equal to the ones of CDNs, in what refers to compounds formation. CDNs may constitute the base of compounds following the paradigm V + N (examples 32).

(32) a. *para- roncós*

stop- snores-CDN.PL

b. *para- choques*

stop- collisions- CDN.PL

'bumper'

#### 4.6. Summing up CDNs vs. nominalisations

Table 3. summarises the main conclusions on the differences between conversion (CDNs) and syntactic nominalisations.

	pluralisation	Theta-mark via Prep	Adjectival modification	Base of other words	anaphora	Adverbial modification	clitics
CDNs	+	+	+	+	+	-	-
Lexicalised $[M]_N$	+	+	+	+	+	-	-
$[M]_{DP}$	-	+	+	-	-	-	-
$[VP]_{DP}$	-	-	-	-	-	+	+

Table. 3 Syntactic environment of conversion and nominalisation

From the analysis of table 3, we may observe that CDNs and lexicalised Vs are in a pole which is typical of nouns. They undergo pluralisation, they are only able to theta-mark via preposition, they admit adjectival modification, and they may function as bases for other words. They reject adverbial modification and clitics.  $[VP]_{DP}$  are in the opposite pole obeying to verb environment. In spite of occurring with determiner,  $[VP]_{DP}$  do not pluralise, theta-mark directly instead of via preposition, and they do not function as bases for other words. On the contrary, they accept adverbial modification and clitics. In between these two poles are  $[M]_{DP}$ . Behaving like verbs, they can neither pluralise nor be the bases of other words, in the infinitive form. Behaving like nouns, they theta-mark via preposition and not directly, they take adjectives as modifiers and reject adverbs for it, and they reject clitics.

#### 5. Why conversion is not a simple question of syntactic mechanisms

Until now we have observed that syntactic nominalisations are not at the same level of nouns. Syntactically speaking, on one hand, nominalisations cannot undergo the syntactic environment every noun can bear. On the other hand, syntactic nominalisations undergo syntactic environment verbs can bear. In between nominalisation  $[M]_{DP}$  does not provide a total nominal behaviour to those items.

It is now the time to search for evidence about conversion at the lexical level. Is the production of items such as *arranque* 'starting-CDN', *mordo* 'bite-CDN', *apanha* 'gathering-CDN', *abate* 'cut(trees)'; 'slaughter (animals)-CDN', *derrube* 'nocking down-CDN', *descanso* 'rest-CDN' the same as nominalisations such as *o arrancar* '(the) starting-INFINITIVE', *o morder* '(the biting-INFINITIVE)', *o apanhar* '(the) gathering'-CDN? Until now, we have concluded that the first kind of nouns seems to be real nouns, since they obey to every criteria identified by Baker (2003) that a noun must have. Nominalised infinitives still obey to verbal criteria, despite of the occurrence of the determiner. Even in between nominalisations fail some N behaviour.

We need to understand if conversion is a real word formation process or if it is simply a syntactic one. Let us compare deverbal derivations in the next sentence (examples 33 to 37).

(33) a. *O encaixar a madeira leva muito tempo.*

The fit-INF the wood takes a.lot.of time  
'Fitting the wood takes a lot of time.'

b. *O encaixe da madeira leva muito tempo.*

The fit-CDN of.the wood takes a.lot.of time  
'The fitting of the wood takes a lot of time.'

c. *O encaixe que escolhi é de metal.*

The fit-CDN-CONC that choose-1st SING is of metal  
'The dowel that I have chosen is made of metal.'

(34) a. *O abrigar dos animais por Noé foi rápido.*

The shelter-INF of.the-PL animals by Noah was quick  
'The sheltering of the animals by Noah was quick.'

b. *O abrigo dos animais por Noé foi rápido.*

The shelter-CDN of.the-PL animals by Noah was quick  
'The sheltering of the animals by Noah was quick.'

c. O abrigo era de madeira.  
The shelter-CDN-CONC was of wood  
'The shelter was made of wood.'

(35) a. O derrubar sobreiros é proibido por lei.  
The cut.down-INF cork.oak is forbidden by law  
'Cutting down cork oaks is forbidden by law.'

b. O derrube de sobreiros é proibido por lei.  
The cut.down-CDN of cork.oak is forbidden by law  
'The cutting down of cork oaks is forbidden by law.'

c. \*O derrube de sobreiros que vi ontem é um mau produto.  
The cut.down-CDN of cork.oak that see-1stSING  
yesterday is a bad product  
'\*The cutting down of cork oaks I saw yesterday is a bad product.'

(36) a. O remendar roupa é uma boa solução.  
The mend-INF cloth is a good solution  
'Mending cloth is a good solution.'

b. O remendo da roupa é uma boa solução.  
The men-CDN of.the cloth is a good solution  
'The mending of the cloth is a good solution.'

c. O remendo que vi ontem é um bom produto (feito à mão).  
The mend-CDN.CONC that see-1stSING yesterday is a  
good product (made by.the hand)  
'The mend that I saw yesterday is a good product (handmade).'

- (37) a. *O atinar a resposta é ótimo.*  
 The find-INF the answer is great  
 'Finding the answer is great.'
- b. *O atino da resposta é ótimo.*  
 The find-CDN of.the answer is great  
 'The finding of the answer is great.'
- c. *\*O atino que vi ontem é um bom produto.*  
 The CDN that see-1stSING yesterday is a good product  
 'The finding that I saw yesterday is a good product.'

In the examples (33) to (37), sentences (a.) represent examples of transposition of verbs into a nominal syntactic behaviour. Sentences (b.) represent examples of a converted deverbal noun meaning 'event'. This is the most regular meaning of this kind of derived words. Sentences (c.) represent the actualisation of converted deverbal nouns with a concrete meaning - 'product', which is one of the most concrete meanings of these words (different levels of concreteness can be observed in Rodrigues 2001: 139-145).

As we can see in sentences (a.) every verb may be integrated in a sentence with a determiner, which gives the verbal phrase (VP) a determiner phrasal character. This is in accordance with data provided in § 3.1.1. We may check if this syntactic mechanism works for every verb (examples 38a):

(38) a

<i>o fumar</i> the smoke-INF
<i>o beber</i> the drink-INF
<i>o analisar muitas frases</i> the analyse-INF many sentences
<i>o escrever muitos artigos</i> the write-INF many papers
<i>o conduzir muitos carros</i> the drive-INF many cars
<i>o haver pouco dinheiro</i> the there.be-INF few money
<i>o ser boa pessoa</i> the be-INF a good person

<i>o estar em Londres</i> the be-INF in London
<i>o viver alegremente</i> the live-INF joyously
<i>o morrer sem dor</i> the die-INF without pain
<i>o nascer de parto natural</i> the be.born-INF of natural childbirth
<i>o permanecer em casa</i> the remain-INF at home
<i>o desenvolver as teorias</i> the develop-INF the theories
<i>o arborizar muitos terrenos</i> the plant.trees.in-INF many plots
<i>o refletir sobre os assuntos</i> the reflect-INF on the subjects
<i>o dar alimentos para a ajuda de África</i> the give-INF of food in aid of Africa
<i>o ler histórias aos meninos</i> the read-INF stories to the children
<i>o rir muito faz bem à saúde</i> the laugh-INF a.lot is good for health

In (38a) there are unaccusative verbs (*morrer* ‘to die’, *nascer* ‘to be born’, *permanecer* ‘to remain’), transitive verbs (*arborizar* ‘to plant trees in a plot’, *dar* ‘to give’, *ler* ‘to read’), unergative verbs (*rir* ‘to laugh’), verbs with an erudite morphological structure (*conduzir* ‘to drive’, *arborizar* ‘to plant trees in a plot’), and with non-erudite morphological structure (*ler* ‘to read’, *dar* ‘to give’), verbs with just one syllable in the infinitive form (*ler*, *dar*, *rir*). We could add more classes, as the reader can see.

Going back to examples (33) to (37), sentences (b.) show converted deverbal nouns from the verbs *encaixar* ‘to fit’, *abrigar* ‘to shelter’, *derrubar* ‘to cut down’, *remendar* ‘to mend’ and *atinar* ‘to find’. However, four verbs are not sufficient to decide if every verb may generate converted deverbal nouns. Taking into account the sentences we have presented above, we may conclude that none of those verbs whose VPs have been nominalised ([[VP]]<sub>DP</sub>) are now able to create a CDN shown in (38 b):

## (38) b

* <i>o fumo</i> the smoke-CDN
* <i>o bebo</i> the drink-CDN
* <i>o analiso de muitas frases</i> the analyse-CDN of many sentences
* <i>o escrever de muitos artigos</i> the write-CDN of many papers
* <i>o conduzo muitos carros</i> the drive-CDN of many cars
* <i>o haver??? pouco dinheiro</i> the there.be-CDN of few money



*o ser??? boa pessoa the be-CDN a good person
*o estar??? em Londres the be-CDN in London
*o vivo alegremente the live-CDN joyously
*o morro sem dor the die-CDN without pain
*o nasço de parto natural the be.born-CDN of natural childbirth
*o permaneço em casa the remain-CDN at home
*o desenvolvo das teorias the develop-CDN of the theories
*o arborizo de muitos terrenos the plant.trees.in-CDN of many plots
*o refleto?? sobre os assuntos the reflect-CDN on the subjects
*o/a do/da/de de alimentos para a ajuda de África the give-CDN of food in aid of Africa
*o/a le/lallo histórias aos meninos the read-INF stories to the children
* o/a ri/rie/ria faz bem à saúde the laugh-INF is good for health

Syntactic environment has been changed from (38a) to (38b) to create an environment adequate to nouns.

Notice that some of the constructions in (38 b) are acceptable, if and only if they would correspond to a citation role in speech. These would be [o [vivo alegremente]<sub>VP</sub>]<sub>DP</sub> (and not [o vivo alegre], since a verb would not occur with adjective; this would correspond to a NP), [o [morro sem dor]<sub>VP</sub>]<sub>DP</sub>, [o [permaneço em casa]<sub>VP</sub>]<sub>DP</sub>, which corresponds to a speech situation like the one represented in :

Marta says to João: -*Eu vivo alegremente.*

And João tells António: - *I don't agree with 'o [vivo alegremente]' da Marta.*

So far, verbs seem to be syntactically nominalised in [VP]<sub>DP</sub> (38 a), but not every verb is lexically convertible in a noun [N] (38b).

Sentences (33-37 c) show something more interesting: even those verbs that may generate CDNs are not able to produce concrete meanings of those nouns. Sentences (33-37) are simply a random sample. But if sentences (c) are even coerced in a universe of five objects, it is easy to imagine that its number among language is fewer than sentences (b), which, from the examples in (38), are less than (a).

This evidences that CDNs are not syntactically derived. Why not? Some reasons may be adjudged here:

- a. CDNs are not producible from every verb, contrarily to syntactic nominalisations;
- b. As every word formation mechanism, deverbal conversion is constraint-based. These constraints are of word formation type and not of syntactic type (cf. Rodrigues 2009; Rodrigues 2004 for those constraints).
- c. CDNs have a semantic structure that is much more complex than a simple syntactic nominalisation would admit. Do syntactic nominalisations in sentences (33) to (37) admit a concrete reading?

*Atino* does not provide a concrete reading. We should not expect nominalisations to behave in a way that not even lexical nouns do. So let us choose verbs whose CDNs display a concrete meaning (examples 39 and 40). CDNs are presented in bold in examples 39; correspondent verbs are presented in bold in examples 40.

- (39) a. O **enxerto** da pereira vê-se no nó.  
'The graft of the pear tree is seen at the knot.'
- b. O **remendo** das calças é azul.  
'The mend of the trousers is blue.'
- c. O **respiro** da baleia é maior do que o respiro do forno.  
'The blowhole of the wale is bigger than the air hole of the oven.'
- d. O **saltarico** traz calças verdes.  
'The hopper wears green trousers.'
- e. O **agasalho** é azul.  
'The garment is blue.'
- f. O **abordo** é de madeira.  
'The wharf is made of wood.'
- (40) a. \*O **enxertar** da pereira vê-se no nó.  
'\*The graft-INF of the pear tree is seen at the knot.'

b. \*O **remendar** das calças é azul.

'The mend-INF of the trousers is blue.'

c. \*O **respirar** da baleia é maior do que o respirar do forno.

'The breathe-INF of the wale is bigger than the breathe-INF of the oven.'

d. \*O **saltaricar** traz calças verdes.

'The hop-INF wears green trousers.'

e. \*O **agasalhar** é azul.

'The muffle.up-INF is blue.'

f. \*O **abordar** é de madeira.

'The board-INF is made of wood.'

From examples (39) and (40) we conclude that a concrete meaning is not available in syntactic nominalisations, even when the correspondent CDN provides a concrete meaning. Thus, CDNs are not a simple matter of syntactic environment as it is syntactic nominalisation.

Besides, not every converted deverbal noun displays any concrete meaning.

If the meaning of the word depended only on the context, any converted noun would have any meaning. But this does not happen. This could lead the followers of mainstream generative grammar to believe that this kind of word is not constructed by any regular mechanism, and thus, that CDNs have to be stored like words such as *gato* 'cat' and *mesa* 'table'. However, the multiplicity and the non-transversality of meanings in CDNs are not good reasons to consider CDNs as irregular, belonging to the periphery of language. Non-transversality is a clue on regularity, not syntactic regularity, but word formation regularity. Consider this: if every CDN had every meaning, it would be aleatory, belonging to periphery. There would be overgeneration in the lexicon. What is regular in syntax may be irregular in the lexicon and vice versa. Take the contrast: if every verb could originate 'agent' meaning nominalisations, these would occur in the syntax.

This does not happen with CDNs. We have to check if some kind of

verbs generates a specific kind of meaning.

In the examples (39) above, there are three classes of concrete meanings: 'product', 'agent', 'place/part of the body'. *Enxerto* 'graft', *remendo* 'mend', *agasalho* 'garment' have a 'product' meaning; *saltarico* 'hopper' has an 'agent' meaning; and *abordo* 'wharf' and *respiro* 'air hole; blowhole' have a 'place /part of the body' meaning. Notice that *enxertar* 'to graft', *remendar* 'to mend', *agasalhar* 'to muffle' are accomplishment verbs; and that the 'product' meaning corresponds to the projection of the [point of arrival] and the [telic] features of the verbs (Rodrigues 2008). *Abordar* 'to board' and *respirar* 'to breathe' imply a [path] in their semantic structures, which is projected in the concrete meanings of the CDNs.

Unergative verbs of emission of light and substance (Levin & Rappaport Hovav: 1995) generate CDNs with 'product' meaning, apart from the 'event' meaning. However, [V]<sub>DP</sub> of the correspondent verb do not display those meanings (examples 41-42).

(41) a. \*O *mijar* é excessivamente amarelo.  
The piss-INF is excessively yellow  
'\*Pissing is excessively yellow.'

b. O *mijo* é excessivamente amarelo.  
The piss-CDN is excessively yellow  
'The piss is excessively yellow.'

(42) a. \*O *brilhar* é amarelo.  
The glitter-INF is yellow  
'\*Glittering is yellow.'

b. O *brilho* é amarelo.  
The glitter-CDN is yellow  
'The glitter is yellow.'

Notwithstanding, unergative verbs of emission of substance or light do not generate CDNs with 'agent' meaning, whilst unergative verbs of behaviour involving motion do (examples 43 and 44). Once again, notice

that [V]<sub>DP</sub> do not display the 'agent' meaning (examples 43-44b).

(43) a. *O João é um saltarilho.*

'John is a hopper.'

b. \**O João é um saltarilhar.*

'\*John is a hopping.'

(44) a. *O João é um estardalho.*

'John is a rowdy.'

b. \**O João é um estardalhar.*

'\*John is a behave noisily.'

These few examples support the hypothesis that CDNs generation belongs to a specific word formation domain, not to syntax, since their formation depends on lexical (among others) features of the verbal bases.

### 5.1. Constraints

We are now going to summarise structural constraints that rule CDNs formation. We will go fast through these, because they have been studied in (Rodrigues 2009).

Following Rodrigues (2009), these constraints belong to different structures and are in interface with each other.

1- Prosodic constraint in connection with processing conditions: the verbal stem must obey the minimal prosodic word requirement so that the obtained noun is lexically identified with its base;

2 - Morphological constraint in connection with pragmatic and processing conditions: the verbal stem must not contain morphological constituents etymologically classified as [+Latin] and simultaneously prototypical of a cultivated usage;

3 - Lexical-conceptual structure and argument structure constraints: the verbal base must have in its basic construction a) a [+concrete] meaning; b) an external argument that semantically corresponds to a Proto-Agent role.

These are not syntactic constraints.

These constraints do not apply to infinitive nominalisation. There seems to be no constraints of this kind ruling syntactic nominalisation.

## 5.2. Other differences between CDNs and infinitive nominalisations

Until now we have observed constraints that rule the formation of CDNs. These constraints do not apply to the infinitive nominalisations, and this shows that the latter are generated in the syntax, whilst the former are generated in the lexicon. However, there are more structural differences between them that evidence a distinct origin.

### 5.2.1 Morphological clues

CDNs may bear the three theme vowels of Portuguese. These theme vowels have no relation with the theme vowel of the verb. This is a lexical choice, not a syntactic one (table 4).

Verbal Theme vowel	CDN with -a	CDN with -e	CDN with -o
-a	<i>limpa</i>	<i>trote</i>	<i>choro</i>
-e	<i>colha</i>	<i>combate</i>	<i>mordo</i>
-i	<i>zurza</i>	<i>cuspe</i>	<i>cuspo</i>

Table 4. Verbal Theme Vowels and nominal theme vowels

Even the same stem may undergo the three theme vowels (table 5). This has no relation with syntax.

Verbal Theme vowel	CDN -a	CDN -e	CDN -o
-a	descasca - desmama - encarna - encrava enleia ensaca entalha espalha pinga queima quita - requeima	descasque desgaste desmame encalhe encarne encoste enclave - ensaque entalhe - pingue - quite rasgue requeime	- desgasto - encalho - encosto encravo enleio - entalho espalho pingo queimo - rasgo requeimo
-e	recolha	-	recolho
-i	-	cuspe	cuspo

Table 5. CDNs stems that go with different theme vowels.

### 5.2.2. Gender

Apart from the theme vowel variance, CDNs may bear masculine or feminine gender. This is seen by the determiner before the N (examples 45).

- (45). a. O                      *aguço*                      do              *lápis*.  
 The-MASC              sharpen-CDN              of.the      pencil  
 'The sharpening of the pencil.'
- b. A                      *apanha da*                      azeitona.  
 The-FEM                      gather-CDN              of.the      olive  
 'The gathering of the olive.'
- c. A                      *encrava*                      do              *mecanismo*.  
 The-FEM                      jam-CDN                      of.the      device  
 'The jamming of the device.'
- d. O                      *encravo*                      do              *mecanismo*.  
 The-MASC                      jam-CDN                      of.the      device  
 'The jamming of the device.'

e. O                    *encrave*                    *do*                    *mecanismo.*  
The-MASC           jam-CDN            of.the            device  
'The jamming of the device.'

f. A                    *poda*                    *da*                    *árvore.*  
The-FEM            prune-CDN           of.the            tree  
'The pruning of the tree.'

g. A                    *curte*                    *da*                    *música.*  
The-FEM            enjoy-CDN           of.the            music  
'The enjoyment of the music.'

On the contrary, infinitive nominalisations only bear masculine determiner, which corresponds to the unmarked gender in Portuguese. This seems to show that gender is dependent on the lexicon and not on the syntax, even if it is (only) observable in syntax (Of course we are not talking about syntactic agreement.). If dependent on the syntax, the gender will always be masculine. This is the same for French, Spanish and Italian (see examples 46).

(46) a. *El mirarte a los ojos.* Vs. \**La mirarte a los ojos.*

b. *Le lancer du disque.* Vs. \**La lancer du disque.*

c. *Il cantare aiuta la respirazione.* Vs. \**La cantare aiuta la respirazione.*

## 6. Conclusions

CDNs construction occurs in the lexicon. CDNs have meaning variables not dependent on syntax. Their construction obeys to structural constraints not explainable by syntax. They bear the determiners *o* and *a*. Nominalisations have no multiple meaning and do not undergo those constraints. Nominalisations only take the determiner *o*. It functions as a kind of referenciation.

Every nominalisation occurs in the syntax. This does not mean that the



lexicon is the field of irregularity. The lexicon is generative, without being syntax. Constraints in interface show that.

[V]<sub>DP</sub> are not in the lexicon as nouns, they are in the syntax. Thus, it takes just a determiner to make a nominalisation; but it takes many more things to make a noun. Pluralisation, anaphora, and the possibility of generating words following other nouns paradigms are crucial.

Of course, with our perspective we do not mean that the syntactic study of lexical nominalisations is not rightful. Of course it is, apart from bringing evidence to the distinction we have made here between syntactic nominalisation and lexical conversion, a syntactic study should bring light to the syntactic behaviour of these lexemes. However, understanding their syntactic behaviour is not the same as explaining their formation process.

#### REFERENCES

- Baker, M. 2003. *Lexical categories. Verbs, nouns, and adjectives*. Cambridge: Cambridge University Press.
- Bhat, D. 1994. *The adjectival category: criteria for differentiation and identification*. Amsterdam: John Benjamins.
- Brito, A. M. 1993. Aspects de la syntaxe du SN en Portugais et en Français. *Revista da Faculdade de Letras do Porto, Série Línguas e Literaturas*. X, 25-53.
- Brito, A. M. 2003. Categorias sintácticas. In: Mateus, M. H. et al. *Gramática da língua portuguesa*. Lisboa: Editorial Caminho, 5.<sup>a</sup> ed., 323-432.
- Brito, A. M. 2012. A nominalização do infinitivo em português europeu. *Textos seleccionados, XXVII Encontro Nacional da Associação Portuguesa de Linguística*. Lisboa: APL, 98-120.
- Chomsky, N. 1970. Remarks on nominalization. In: R. Jacobs; P. Rosenbaum (Eds.). *Readings in English Transformational Grammar*. Waltham: Ginn, 184-221.
- Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, N. 1995. *The Minimalist Program*. Massachusetts: The MIT Press.
- Chomsky, N & Lasnik, H. 1993. The Theory of Principles and Parameters. In: N. Chomsky. *The minimalist program*. Massachusetts: The MIT Press, 13-127.
- Culicover, P. W. & Jackendoff, R. 2005. *Simpler syntax*. Oxford: Oxford University Press.

- Croft, W. 1991. *Syntactic categories and grammatical relations*. Chicago: University of Chicago Press.
- Geach, P. 1962. *Reference and generality*. Ithaca, NY: Cornell University Press.
- Grimshaw, J. 1990. *Argument structure*. Massachusetts: The MIT Press.
- Gupta, A. 1980. *The logic of common nouns*. New Haven, Connecticut: Yale University Press.
- Hengeveld, K. 1992. *Non-verbal predication: theory, typology, diachrony*. Berlin: Mouton de Gruyter.
- Jackendoff, R. 2002. *Foundations of language. Meaning grammar and evolution*. Oxford: Oxford University Press.
- Kerleroux, F. 1996. *La coupure invisible. Études de syntaxe et de morphologie*. Villeneuve d'Ascq. Presses Universitaires du Septentrion.
- Rappaport Hovav, M. & Levin, B. 1995. *Unaccusativity. At the syntax-lexical semantics interface*. Cambridge; Massachusetts: The MIT Press.
- Rodrigues, A. S. 2001. *A construção de postverbais em português*. Porto: Granito Editores.
- Rodrigues, A. S. 2004. Condições de formação de nomes postverbais em português. In: Graça Rio-Torto et al. *Verbos e nomes em português*. Coimbra: Livraria Almedina, 129-185.
- Rodrigues, A. S. 2008. *Formação de substantivos deverbais sufixados em português*. München: Lincom.
- Rodrigues, A. S. 2009. Portuguese converted deverbal nouns: constraints on their bases. *Word Structure*. 2,1, 69-107.
- Rodrigues, A. S. 2012 a. What affixes reveal about interfaces in word formation. In: M. Bloch-Trojnar; A. Bloch-Rozmej. *Modules and interfaces*. Lublin: Wydawnictwo. (Studies in Linguistics and Methodology; 4), 255-270.
- Rodrigues, A. S. 2012 b. *Jackendoff e a arquitetura paralela. Apresentação e discussão de um modelo de linguagem*. München: Lincom.

# Bare morphology

Alina Villalva

alina\_villalva@yahoo.com

*Faculdade de Letras da Universidade de Lisboa (Portugal)*

*CLUL, Laboratório de Psicolinguística*

ABSTRACT. Several cases of Portuguese words that share the same root string are studied, representing different types of conversion. Such is the case with *firm*<sub>ADJ</sub> – *firmar*<sub>V</sub>; *abaix*<sub>ADV</sub> – *abaixar*<sub>V</sub>; *manch*<sub>N</sub> – *manchar*<sub>V</sub>; *afia*<sub>V</sub> – *afia*<sub>N</sub>. In order to set them apart one needs to take into account features from every grammar domain. It may be concluded that this fact is one of the best arguments for considering conversion as a lexical word formation process (as opposed to morphological, syntactic or semantic notions).

KEY-WORDS. conversion, derivational process, morphology.

Conversion is usually defined as a derivational process linking lexemes of the same form but belonging to different word classes. Virtually all of this has been questioned at one point or another, and yet the concept of conversion remains in use, very much as the conventional system of word-classes does in languages for which it is theoretically inadequate.

Bauer & Varela (2005: 8)

Eight years later, this opinion by Bauer & Varela (2005) still remains the state-of-the-art on conversion: it is still not difficult to accept that conversion may not be a derivational process, that it may not link lexemes, that, whichever entities are involved, they may not have the same form and that they may not even belong to different word classes. Nevertheless, at some point of the discussion on word formation, it is necessary to consider these cases that involve neither affixation nor compounding. This kind of (negative) definition allows treating as equal what has traditionally been considered as different. Portuguese grammars generally mention two sorts of ‘bare’ word formation processes, both classified under the label of derivation: ‘*derivação regressiva*’, for an allegedly subtractive deverbal

noun formation<sup>1</sup>, and ‘*derivação imprópria*’<sup>2</sup>, for the use of a word in a non-canonical syntactic position. Yet, ahead of his time, Said Ali (1923-1931, 1965: 230-231) questioned this view by arguing that “*a mudança de sentido e de função que sofrem as palavras [...] deverá denominar-se semântica e não derivação*”. So, much remains to be clarified in this domain. In this paper, some of the basic questionings will be confronted with data from Portuguese, from the viewpoint of word grammar<sup>3</sup>. The relevant data includes a wide variety of situations, some of which may be seen below:

- (1) i. a. *Não gosto de bebidas **doces***<sub>adj</sub>  
‘I don’t like sweet drinks’  
b. *Vou fazer um **doce***<sub>n</sub>*de amora*  
‘I’m going to prepare a blackberry jam’
- ii. a. *A palha **seca***<sub>adj</sub>*arde depressa*  
‘Dry hay burns quickly’  
b. *A **seca***<sub>n</sub>*destruiu as colheitas*  
‘The drought destroyed the crops’  
c. *Temos de pôr o barco em **seco***<sub>n</sub>  
‘We have to drydock the boat’  
d. ***Seca***<sub>v</sub>*as batatas antes de fritá-las*  
‘Dry the potatoes before frying them’
- iii. a. *Estas flores estão **murchas***<sub>adj</sub>  
‘These flowers are wilted’  
b. *Estas flores **murcharam***  
‘These flowers wilted’

<sup>1</sup> Cunha & Cintra (1984: 104) describe the ‘*derivação regressiva*’ as “*um processo de criação vocabular [...] que consiste na redução da palavra derivante por uma falsa análise da sua estrutura*”. Even in this kind of literature, some evolution has taken place. In a recent edition of the *Moderna Gramática Portuguesa*, Bechara (2009: 370) claims that “*intimamente relacionada com a derivação temos a formação regressiva ou deverbal, que consiste em criar palavras por analogia, pela subtração de algum sufixo, dando a falsa impressão de serem vocábulos derivantes: de atrasar tiramos atraso, de embarcar, embarque; de pescar, pesca; de gritar, grito.*”

<sup>2</sup> Cunha & Cintra (1984: 105) call ‘*derivação imprópria*’ the “*processo de enriquecimento vocabular pela mudança de classe de palavras*”, “*sem sofrer modificação na forma*”. In this case, Bechara (2009: 372) offer as an example the nominal use of a negation adverb (*terrível palavra é um não*) and he calls it a case of conversion, but he still feels the need to make an observation about the previous label: “*Os casos de conversão recebem o nome de derivação imprópria*. The terminological turn is unconvincingly justified by Bechara for the fact that “*a conversão não repercute na estrutura do significante de base [...]*.”

<sup>3</sup> The background analysis of words for the description in this paper is based on Villalva (2000; 2008; 2011).

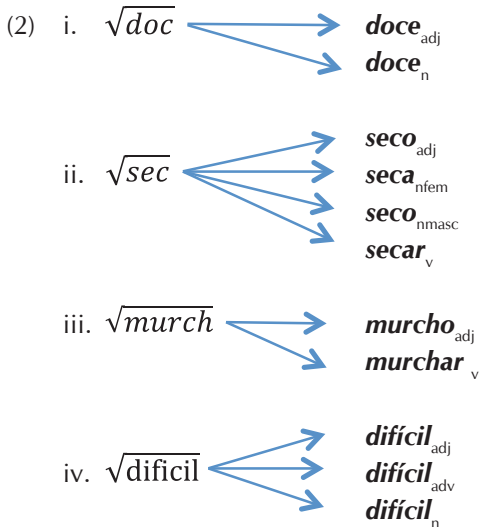
- iv. a. *Não me faça perguntas **difíceis***<sub>adj</sub>  
 ‘Don’t ask me difficult questions’
- b. *Ele adora falar **difícil***<sub>adv</sub>  
 ‘He likes to speak in a complicated manner’
- c. *O **difícil***<sub>n</sub> *não é chegar a horas*  
 ‘The hard part is not arriving on time’
- v. a. *Este poço é muito **fundo***<sub>adj</sub>  
 ‘This is a very deep well’
- b. *Ele respira **fundo***<sub>adv</sub>  
 ‘He breathes deeply’
- c. *Quem não nada vai ao **fundo***<sub>n</sub>  
 ‘Those that can’t swim will keel over’
- d. *Eles **fundaram***<sub>v</sub> *um clube*  
 ‘They founded a club’
- vi. a. *Ele é de **firmes***<sub>adj</sub> *convicções*  
 ‘He has firm convictions’
- b. *Vai **firme***<sub>adv</sub>  
 ‘Go steady’
- c. *Com esta fotografia **firmaremos***<sub>v</sub> *este momento tão particular*  
 ‘With this picture, we will document this peculiar special moment’
- vii. a. *Já fiz **imensas***<sub>adj</sub> *tentativas*  
 ‘I’ve already made lots of attempts’
- b. *Gosto **imenso***<sub>adv</sub> *de queijo*  
 ‘I like cheese a lot’
- viii. a. *O professor mora **longe***<sub>adv</sub>  
 ‘The teacher lives far away’
- b. *Não vejo nada bem ao **longe***<sub>n</sub>  
 ‘I can’t see well at a distance’
- ix. a. *Passa **adiante***<sub>adv</sub>  
 ‘Move forward’
- b. *Vou **adiantar***<sub>v</sub> *este projecto*  
 ‘I’m going to keep on working on this project’

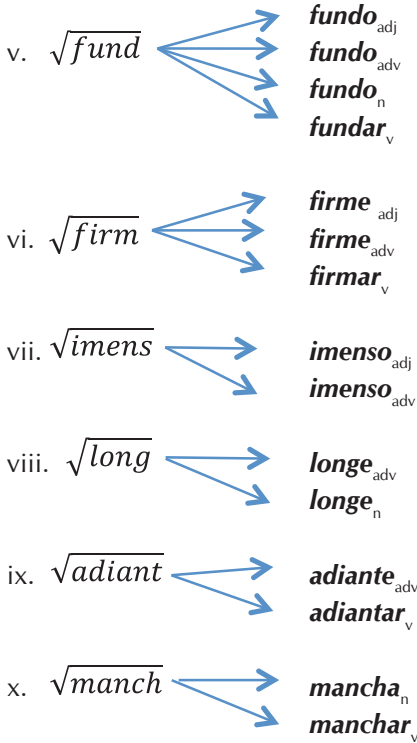
- x. a. Há **manchas**<sub>n</sub> de humidade na parede  
'There are damp patches on the wall'  
b. A humidade **manchou**<sub>v</sub> a parede  
'Damp stained the wall'

The basic issue under discussion in this paper concerns the common vs different nature of all the situations in (1). If they are the outcome of a single process, we could call it conversion, but what kind of process is this and how does it operate? If, instead, they involve different word formation processes, then it is relevant to ask if there is any of them that is a word formation process, or if it is a lexical, a morphological or a syntactic matter, or even, as Said Ali suggested, if it is all about semantics.

1. The underspecified root hypothesis and its reverse

Hypothetically, one could admit that all the words in (1) are projections of the following set of roots that would, in each case, unfold into a variable number of words:





The problem with this hypothesis stems from the fact that the words formed by this set of rules bear a number of inherent and selectional properties that are idiosyncratic. Features such as gender or thematic class for nouns (eg. *fundo* and *seco* are *-o* ending masculine nouns, whereas *mancha* and *seca* are *a-* ending feminine nouns), variability for adjectives (eg. *fundo/funda* is a variable adjective, but *doceis* invariable), the thematic class of adverbs (eg. *fundo* is an *-o* ending adverb, but *firme* is an *e-* ending adverb) and the conjugation for verbs (*fundar* is 1<sup>st</sup> conjugation and *fugir* is 3<sup>rd</sup> verb), are not derivable – they must be listed somewhere. Notice that the similarity of words involved in conversion is self-evident if we consider citation forms, but less convincing if we take into account full lexical paradigms or their syntactic and semantic behaviour:

$\sqrt{\text{fund}}$	adj	adv	n	v		
citation form	<i>fund</i>	<i>fund</i>	<i>fund</i>	<i>fundar</i>		
other forms	<i>fundos</i> <i>funda</i> <i>fundas</i>		<i>fundos</i>	<i>fund</i> <i>fundas</i> <i>funda</i> <i>fundamos</i> <i>fundais</i> <i>fundam</i> <i>fundei</i> <i>fundaste</i> <i>fundou</i> <i>fundámos</i> <i>fundastes</i> <i>fundaram</i> <i>fundava</i> <i>fundavas</i> <i>fundávamos</i> <i>fundáveis</i> <i>fundavam</i>	<i>fundara</i> <i>fundaras</i> <i>fundara</i> <i>fundáramos</i> <i>fundáreis</i> <i>fundei</i> <i>fundes</i> <i>fundemos</i> <i>fundeis</i> <i>fundem</i> <i>fundasse</i> <i>fundassem</i> <i>fundássemos</i> <i>fundásseis</i> <i>fundassem</i> <i>fundares</i> <i>fundarmos</i>	<i>fundardes</i> <i>fundarem</i> <i>fundarei</i> <i>fundarás</i> <i>fundará</i> <i>fundaremos</i> <i>fundareis</i> <i>fundarão</i> <i>fundaria</i> <i>fundarias</i> <i>fundaria</i> <i>fundarias</i> <i>fundaríamos</i> <i>fundarieis</i> <i>fundariam</i> <i>fundado</i> <i>fundando</i>

There is a relatively broad consensus that idiosyncratic properties are lexically established, so, if the roots that are inscribed in the lexicon are underspecified roots, where does this idiosyncratic information come from?

The reverse of the underspecified root hypothesis is the one root for each lexical paradigm hypothesis. Again, if we consider the data in (1), the output of this new hypothesis is as follows:

- (4) i.  $\frac{adj \sqrt{doc}}{n \sqrt{doc}}$
- ii.  $\frac{adj \sqrt{sec}}{n \sqrt{sec}}$   
 $\frac{n \sqrt{sec}}{v \sqrt{sec}}$
- iii.  $\frac{adj \sqrt{murch}}{v \sqrt{murch}}$
- iv.  $\frac{n \sqrt{difícil}}{adj \sqrt{difícil}}$   
 $\frac{adv \sqrt{difícil}}$



- v.  $\begin{array}{l} \sqrt[n]{fund} \\ \sqrt[adj]{fund} \\ \sqrt[v]{fund} \end{array}$
- vi.  $\begin{array}{l} \sqrt[adj]{firm} \\ \sqrt[adv]{firm} \\ \sqrt[v]{firm} \end{array}$
- vii.  $\begin{array}{l} \sqrt[adj]{imens} \\ \sqrt[adv]{imens} \end{array}$
- viii.  $\begin{array}{l} \sqrt[n]{long} \\ \sqrt[adv]{long} \end{array}$
- ix.  $\begin{array}{l} \sqrt[adv]{adiant} \\ \sqrt[v]{adiant} \end{array}$
- x.  $\begin{array}{l} \sqrt[n]{manch} \\ \sqrt[v]{manch} \end{array}$

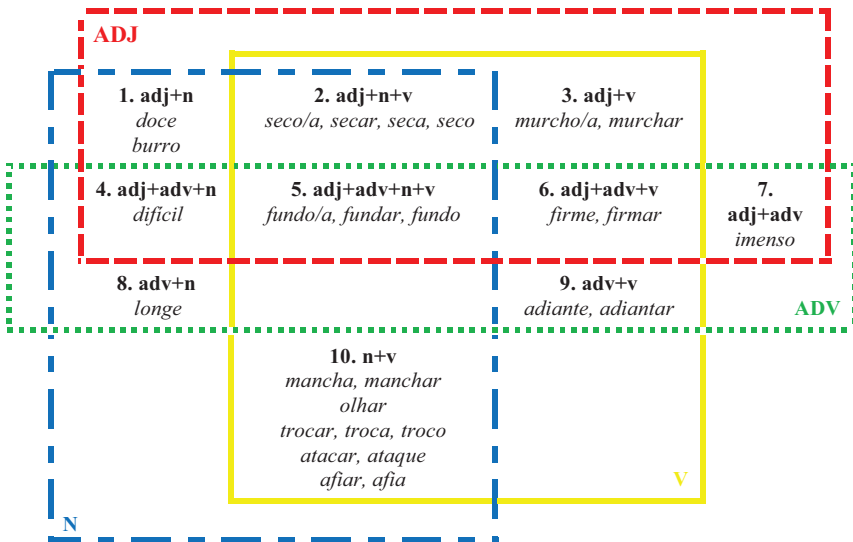
Aside from being blatantly uneconomical, the main problem of this counter-hypothesis is that it fails to acknowledge that probably not all the roots from (4i) to (4x) have the same lexical status. Even intuitively, it is easily demonstrated that the use of *difícil* as an adjective or as a noun does not require a huge effort, from the point of view of the knowledge of these words. However, the use of pairs such as *mancha<sub>N</sub>* / *manchar<sub>V</sub>* and *trocar<sub>V</sub>* / *troca<sub>N</sub>*, for instance, requires a more significant effort, since in one of these cases, the verb is a sort of causative involving the noun, and, in the other case, the noun is a sort of deverbal action noun. This is exactly what can be found in typical morphological word formation processes:

- (5) a. *mancha<sub>N</sub>* '(the) damp patch'    *manchar<sub>V</sub>* 'to stain'  
       *plano<sub>N</sub>* '(the) plan'                *planear<sub>V</sub>* 'to make plans'  
       *cristal<sub>N</sub>* '(the) cristal'            *cristalizar<sub>V</sub>* 'to cristalize'  
       *fruto<sub>N</sub>* '(the) fruit'                *frutificar<sub>V</sub>* 'to fructify'

b. <i>trocar</i> <sub>V</sub>	'to change'	<i>troca</i> <sub>N</sub>	'(the) change'
<i>apresentar</i> <sub>V</sub>	'to present'	<i>apresentação</i> <sub>N</sub>	'(the) presentation'
<i>esquecer</i> <sub>V</sub>	'to forget'	<i>esquecimento</i> <sub>N</sub>	'(the) forgetting'
<i>tolerar</i> <sub>V</sub>	'to tolerate'	<i>tolerância</i> <sub>N</sub>	'(the) tolerance'

Strictly speaking, conversion is not a morphological word formation process, since no affixes or compounding strategies are involved, but it is also quite obvious that conversion is a word formation process of a different nature, or, most probably, even more than one word formation process. In the following sections, I will try to identify different subtypes of conversion in Portuguese. At the same time, I will attempt to consolidate a more principled analysis of the type of data in (1), considering an expanded set of words, shown in (6):

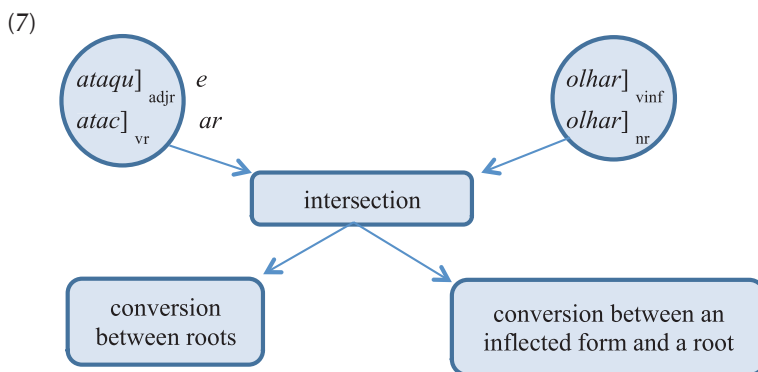
(6)



We will first consider some morphological clues, enabling us to identify the morphological nature of the input and of the output of conversion and, eventually, to uncover the directionality of conversion.

## 2. Morphological category of conversion forms

The idea of treating pairs of words such as *atacar<sub>V</sub>* / *ataque<sub>N</sub>*, as well as *olhar<sub>V</sub>* / *olhar<sub>N</sub>*, as conversion products, was an interesting insight. However, this approach requires identification of the morphological category of the intersection string, in order to be able to obtain some descriptive adequacy. In fact, in *atacar<sub>V</sub>* / *ataque<sub>N</sub>* conversion targets the root, where as in *olhar<sub>V</sub>* / *olhar<sub>N</sub>* conversion involves an inflected form of a root:



From this observation, we can surmise that the identification of the morphological category of conversion members is required. However, probably, the most relevant outcome is related to the fact that the output of a conversion operation is a root, as this pushes (this kind of) conversion into the lexicon, pulling it away from syntax and from semantics. From this point of view, conversion is not very different from regular derivation that also yields roots as its output:

- (8) *ole<sub>nr</sub>*                      *ole-os<sub>adjr</sub>* (*o/a*)                      *ole-os-idad<sub>nr</sub>* (*e*)  
 'oil'                              'oily'                              'oiliness'

## 3. Assessment of the directionality of conversion

Looking at the derivational history of complex words provides an interesting tool for assessing the directionality of conversion. Words such as *simpatizante* 'fan, supporter', which can surface both as nouns and adjectives are probably not the best choice. In fact, since *-nt(e)* suffixation is a vestige

of the Latin present participle, which had a predominantly adjectival nature, we could admit that deverbal *-nt(e)* words are generated as adjectives, and later converted into nouns. But although words such as *atraente* ‘attractive’, *exigente* ‘demanding’ *oufluyente* ‘fluent’ only occur as adjectives, the vast majority of *-nt(e)* words have a hybrid categorical behavior, occurring as either adjectives or nouns.

A much more convincing case concerns pairs such as *embarcar<sub>v</sub>* / *embarque<sub>N</sub>*. *Embarcar* is a parasynthetic verb, based on a noun root (ie. *barc-*). Parasynthetic verb formation is a very productive resource in Portuguese (as in other Romance languages). Since parasynthesis is typically a verb formation process, the directionality puzzle is solved in this case:

- (9) *braç<sub>nr</sub>* ‘arm’ → *abraç<sub>vr</sub>* ‘to embrace’ → *abraç<sub>nr</sub>* ‘hug’  
*barc<sub>nr</sub>* ‘boat’ → *embarc<sub>vr</sub>* ‘to embark’ → *embarqu<sub>nr</sub>* ‘boarding’  
*purg<sub>nr</sub>* ‘purge’ → *expurg<sub>vr</sub>* ‘to expunge’ → *expurg<sub>nr</sub>* ‘purge’  
*bord<sub>nr</sub>* ‘board’ → *transbord<sub>vr</sub>* ‘to transship’ → *transbord<sub>nr</sub>* ‘transshipment’

The same can be said about pairs such as *bloquear<sub>v</sub>* / *bloqueio<sub>N</sub>*. The verb is formed by suffixation, so it is the noun that must be obtained by conversion:

- (10) *bloc<sub>nr</sub>* ‘block’ → *bloque<sub>vr</sub>* ‘to block’ → *bloquei<sub>nr</sub>* ‘blockade’  
*cabec<sub>nr</sub>* ‘head’ → *cabece<sub>vr</sub>* ‘to nod’ → *cabecei<sub>nr</sub>* ‘nod’

These previous examples also allow us to notice that conversion competes morpho-semantically with morphological word formation processes, such as action noun formation. In some cases, the output of conversion even competes with morphological derived words (cf. 11c):

- (11) a. *Essa empresa abate gado*  
‘This company slaughters cattle’  
*O abate de gado por essa empresa ...*  
‘The slaughter of cattle by this company...’

- b. *Esta escola **forma** professores*  
'This school trains teachers'  
*A **formação** de professores por esta escola ...*  
'The training of teachers by this school ...'  
*Este conductor **atropelou** duas pessoas*  
'This driver ran over two people'  
*O **atropelamento** de duas pessoas ...*  
'The running over of two people ...'
- c. *acumular* 'accumulate'      *acúmulo, acumulação* 'accumulation'  
*atropelar* 'runn over'      *atropelo, atropelamento* 'trampling'

A morphologically controlled paraphrase of the conversion words is also a useful tool for assessing directionality. In the case of adjective-noun pairs such as *doce*, we can see that the adjective participates in the paraphrase of the noun, but the reverse is not possible. Consequently, we may hypothesize that the noun is the output form<sup>4</sup>:

- (12) *Vou beber um vinho **doce***  
'I am going to prepare a sweet wine'

*Vou fazer um **doce** (= algo doce)*  
'I am going to prepare a sweet (= something sweet)

Finally, we will consider etymology. This is a criterion that needs to be used with care. Speakers seldom acquire the knowledge of etymology, and word etymons are often unknown to philologists<sup>5</sup>. The following example is of a conversion of an adjective (*AMICUS*) into a noun (*AMICUS*) that took place in Latin. The existence of an adjective (i.e. *amigo* 'friendly') and a noun (i.e. *amigo* 'friend') in Portuguese is probably due to the fact that both existed in

---

<sup>4</sup> Notice that this claim does not imply that the nominal surfacing of *doce* demands the ellipsis of a noun. It's up to the syntactic analysis to decide upon that matter.

<sup>5</sup> There is a huge contingent of words that have an unknown origin (eg. *palerma* 'fool'). In some other cases, the origin of the word is controversial. The Portuguese word *burro*, that may occur as a noun (meaning 'donkey' or 'fool') or an adjective (meaning 'stupid'), is related to a Latin adjective (i.e. *BURICUS*), which meant 'those that come from Buri and have peculiar horses', but eventually the Latin adjective is not the ancestor of the Portuguese word. Originally, *BURI* is a Germanic toponym.

Latin and both evolved into Portuguese. There is no reason to believe that any kind of conversion took place in Portuguese in this instance:

$$(13) \text{ in Latin:} \qquad \qquad \qquad \text{in Portuguese:}$$

$$\begin{array}{ccc} \sqrt[v]{am} & \rightarrow & \sqrt[adj]{amic} & > & \sqrt[adj]{amig} \\ & & \sqrt[n]{amic} & > & \sqrt[n]{amig} \end{array}$$

In the case of *melhor*, historical information clearly shows that its predecessor (i.e. *MELIOR*) was created in Latin, as the superlative of the adjective *BONUS*, that also entered the Portuguese lexicon, as *bom*. The verbal and the adverbial valences of  $\sqrt{melhor}$  are Portuguese innovations, so these may have been obtained through conversion:

(14) *Este é o meu melhor casaco* (adj)  
'This is my best coat'

*Vê lá se melhora dessa gripe* (v)  
'Get better from that flu'

*Nesse caso, o melhor é ir embora* (n)  
'In that case, we had better be leaving'

*Um chapéu branco protege melhor do sol* (adv)  
'A white hat protects better from the sun'

In this section, I endeavored to assemble a set of tools that can serve to determine the directionality of conversion, which amounts to saying that there is no single way of determining this – we need to gather information from various sources. In the following two sections, we will consider two different cases of conversion as word formation processes.

4. Case study #1: deverbal nominalization

There is a large number of pairs of verbs and nouns. In this section, we will see that some of them can be sorted into three subtypes. Subtype I includes action nouns formed from a verb root. As mentioned before, the presence of derived and parasynthetic bases attests that conversion generates deverbal nouns and not the opposite.

(15)

derivation		conversion	
noun root		verbroot	→ action noun
<i>barc-</i> <i>caix-</i> <i>cas-</i>	parasynthetic	<i>embarc-</i> <i>encaix-</i> <i>descasc-</i>	<i>embarque</i> <i>encaixe</i> <i>descasque</i>
<i>baf-</i> <i>boc-</i> <i>far-</i>	<i>ej-</i>	<i>bafej-</i> <i>bocej-</i> <i>farej-</i>	<i>bafejo</i> <i>bocejo</i> <i>farejo</i>

The semantics of these action nouns indicates that they are the result of a word formation process:

- (16) *ataque* ‘ação de atacar’  
 ‘attack’ ‘the act of attacking’

The main problem with the description of these nouns as lexically formed words is the lack of a pattern for their thematic class membership and gender value. Apparently, everything is possible and unpredictable (cf. 17a). We can even find more than one noun formed upon a single verb root (cf. 17b) and we can also find variation between European Portuguese and Brazilian Portuguese:

(17) a.

	base verb	vr → nr	deverbal noun
1 <sup>st</sup> C	<i>avisar</i> ‘to warn’ <i>derramar</i> ‘to spill’ <i>pescar</i> ‘to fish’	<i>avis-</i> <i>derram-</i> <i> pesc-</i>	<i>aviso</i> (-o masc) ‘warning’ <i>derrame</i> (-e masc) ‘leakage’ <i>pesca</i> (-a fem) ‘fishing’
2 <sup>nd</sup> C	<i>socorrer</i> ‘to help’ <i>abater</i> ‘to slaughter’ <i>escolher</i> ‘to choose’	<i>socorr-</i> <i>abat-</i> <i>escolh-</i>	<i>socorro</i> (-o masc) ‘help’ <i>abate</i> (-e masc) ‘slaughter’ <i>escolha</i> (-a fem) ‘choice’
3 <sup>rd</sup> C	<i>resumir</i> ‘to summarize’ <i>curtir</i> ‘to enjoy’ <i>fugir</i> ‘to run away’	<i>resum-</i> <i>curt-</i> <i>fug-</i>	<i>resumo</i> (-o masc) ‘summary’ <i>curte</i> (-e fem) ‘enjoyment’ <i>fuga</i> (-a fem) ‘escape’

b.

<i>trocar</i> ‘to change’	<i>troca</i> ‘changement’ <i>troco</i> ‘change’
---------------------------	--

c.

<i>cuspir</i> ‘to spit’	<i>cuspo</i> (PE) ‘spit’ <i>cuspe</i> (PB) ‘spit’
-------------------------	--

Based on this kind of evidence, the nature of this word formation process seems unquestionably difficult to capture and the lexicon seems the only place capable of hosting it.

Subtype II concerns nouns formed from an impersonal infinitive form (eg. *olhar*). Virtually, all verbs share this possibility of nominalizing an infinitive and all of them, in spite of the conjugation class of the base verb, are masculine nouns belonging to the Ø thematic class:

(18)

	infinitive	noun root	masculine noun
1st C	<i>latejar</i> ‘to throb’	<i>latejar-</i>	<i>latejar</i> ‘throbbing’
	<i>poupar</i> ‘to spare’	<i>poupar-</i>	<i>poupar</i> ‘sparing’
	<i>fumar</i> ‘to smoke’	<i>fumar-</i>	<i>fumar</i> ‘smoking’
2nd C	<i>beber</i> ‘to drink’	<i>beber-</i>	<i>beber</i> ‘drinking’
	<i>conhecer</i> ‘to know’	<i>conhecer-</i>	<i>conhecer</i> ‘knowing’
	<i>derreter</i> ‘to melt’	<i>derreter-</i>	<i>derreter</i> ‘melting’
3rd C	<i>sentir</i> ‘to feel’	<i>sentir-</i>	<i>sentir</i> ‘feeling’
	<i>vestir</i> ‘to wear’	<i>vestir-</i>	<i>vestir</i> ‘wearing’
	<i>dormir</i> ‘to sleep’	<i>dormir-</i>	<i>dormir</i> ‘sleeping’



These infinitive nouns (traditionally called verbal substantives) may compete with a 'cognate' derived noun (cf. 19a-c) or even with another deverbal noun formed by root conversion (cf. 19d). These coincidences, available mostly if there is a semantic specificity for each of these nouns<sup>6</sup>, enable them to act as members of the same class of action nouns:

- (19) a. *Gosto de assistir ao **nascer** do sol*  
 'I like to watch the sunrise'  
*Gosto de assistir ao **nascimento** de uma criança*  
 'I like watching the birth of a child'
- b. *Incomoda-me muito este **padecer** do João*  
 'John's suffering bothers me a lot'  
*Incomoda-me muito este **padecimento** do João*  
 'John's suffering bothers me a lot'
- c. *O **poupar** é muito importante*  
 'Saving is very important'  
*A **poupança** é muito importante*  
 'Saving is very important'
- d. *O **pertencer** a essa família só lhe trouxe preocupações*  
 'Belonging to this family only brought him concerns'  
*A **pertença** a essa família só lhe trouxe preocupações*  
 'Belonging to this family only brought him worries'

---

<sup>6</sup> The existence of a derived action noun often blocks the use of an infinitive noun or, at least, it renders it quite marginal:

*A comemoração<sub>N</sub> do teu aniversário foi muito divertida.*

*O 'comemorar<sub>N</sub> do teu aniversário foi muito divertido.*

'The celebration of your birthday was very nice'

*A apanha<sub>N</sub> da fruta começa em Junho.*

*O 'apanhar<sub>N</sub> da fruta começa em Junho.*

'The harvest of the fruit begins in June'

Nevertheless, there are occasions that seem to point in the opposite direction, particularly when we consider collocations or idioms:

*Até ao lavar dos cestos é vindima*

\*[Até à lavagem dos cestos é vindima]

'The grape harvest isn't over if the baskets haven't been washed'

Like other (derived) action nouns, some of these infinitive nouns get a secondary meaning, which, independently of its precise meaning, implicates countability, evident in the fact that these nouns can pluralize, whereas the previous ones can't:

- (9) a. *É preciso ter cuidado com o prazo de **acabamento** dos trabalhos*  
 'It's necessary to be careful with the deadline for completion of work'  
*acabamento* 'completion' = o acto de acabar  
*Neste apartamento há muitos problemas de **acabamentos***  
 'In this apartment there are many problems with the finishing'  
*acabamentos* 'finishings' = remates

b. *andar* 'walk'

*Por este **andar**, não acabamos hoje o trabalho*

'At this pace, we won't finish this job today'

*por este andar* 'at this pace' = a este ritmo

*Ela tem dois **andares** em Lisboa*

'She has two apartments in Lisbon'

*andares* 'apartments' = apartamentos

*dever* 'duty'

*Ele tem o **dever** de cuidar bem dos filhos*

'He has the obligation to take good care of the children'

*dever* 'obligation' = obrigação

*Vais brincar quando acabares de fazer os **deveres***

'You'll play when you're done with the homework'

*deveres* 'homework' = trabalhos de casa

*jantar* 'dining'

*Às vezes os **jantares** de família são maçadores*

'Sometimes family dinners are boring'

*jantares* 'dinners' = refeições nocturnas

*parecer* 'looking'  
*Quem escreveu estes pareceres sobre a tese?*  
 Who wrote these opinions about the thesis?  
*pareceres* 'opinions' = opiniões

*ser* 'essence' = essência  
*O autor inventou uns seres muito estranhos*  
 'The author has invented some very strange beings'  
*seres* 'beings' = criaturas

Once again, based on this kind of evidence, this word formation process seems to be a lexical matter.

Subtype III deals with subject-nouns<sup>7</sup> formed upon present indicative, 3<sup>rd</sup> person singular verb forms. The identification of the input form is decided on the evidence of 3<sup>rd</sup> conjugation verbs and, particularly, irregular verbs:

(9)

	INFINITIVE	3 <sup>RD</sup> SING. PRES. IND.	NOUN
1 <sup>st</sup> C	<i>empatar</i> 'to hinder'	<i>empata</i>	<i>empata</i> 'hinderer'
	<i>penetrar</i> 'to penetrate'	<i>penetra</i>	<i>penetra</i> 'gatecrasher'
	<i>afiar</i> 'to sharpen'	<i>afia</i>	<i>afia</i> 'sharpener'
2 <sup>nd</sup> C	<i>pertencer</i> 'to belong'	<i>pertence</i>	<i>pertence</i> 'belonging'
	<i>comer</i> 'to eat'	<i>come</i>	<i>come</i> 'eater' <i>come e dorme</i> 'lazy person'
3 <sup>rd</sup> C	<i>dormir</i> 'to sleep'	<i>dorme</i>	<i>dorme</i> 'sleeper'      'lazy person'
	<i>desentupir</i> 'to unclog'	<i>desentope</i>	<i>desentope</i> 'toilet plunger'
Irregular verbs	<i>fazer</i> 'to do'	<i>faz</i>	<i>faz(-tudo)</i> 'handyman'
	<i>ir</i> 'to go'	<i>vai</i>	<i>vaivém</i> 'shuttle'
	<i>vir</i> 'to come'	<i>vem</i>	

<sup>7</sup> Booij (1986: 507) suggests that agent nouns are more adequately labeled as subject-nouns since the word formation process that forms them links the thematic role of the base verb to the derived word. Thus, deverbal subject nouns can be agent nouns (which is the most frequent case), but they can also be patient nouns or experiencer nouns (cf. *sofredor* 'sufferer', *merecedor* 'deserver').

Notice that these subject-nouns, just like morphologically derived subject-nouns<sup>8</sup>, may have a [+human] or a [-animate] interpretation:

(22) a. [+human]

*Ele/a é um/a **empata**.*

‘He/she is a hinderer’

*Ele/a é um/a **penetra**.*

‘He/she is a gatecrasher’

*Ele é um **conquistador**.*

‘He’s a conqueror’

*Ele é um **vindimadeiro**.*

‘He is a grape harvester’

b. [-animate]

*Preciso de comprar um **afia**.*

‘I need to buy a sharpener’

*O **escorrega** precisa de ser **arranjado**.*

‘The slide needs to be fixed’

## 5. Case study #2: verb-forming conversion

This case is about nouns such as *albergue* or *mancha*, that are also related to verbs (e.g. *albergar* ‘to host’, *manchar* ‘to stain’), by sharing the same root. But, unlike the previous cases, they are not action nouns or agent nouns. These are usually concrete, countable nouns, such as *albergue* ‘hostel’. The verbs, however, exhibit some regularities: they are always 1<sup>st</sup> conjugation verbs and they always have a *latusensu* causative meaning:

(23)

noun	nr → vr	verb	paraphrase
<i>albergue</i> ‘hostel’	<i>alberg-</i>	<i>albergar</i> ‘to host’	<i>dar albergue</i>
<i>alerta</i> ‘alert’	<i>alert-</i>	<i>alertar</i> ‘to alert’	<i>pôr alerta</i>
<i>mancha</i> ‘stain’	<i>manch-</i>	<i>manchar</i> ‘to stain’	<i>pôr mancha(s)</i>

<sup>8</sup> Dressler (1986) established a hierarchy for the conceptual category of Agent that captures the observation that agent noun forming resources are also available for instrument noun and locative or source nouns.

These denominal verbs have properties that can be observed in verbs similarly based on adjective and adverb roots. The input form is different, but the output verb resembles the denominal instance: like the previous cases, they are always 1<sup>st</sup> conjugation verbs and they always have a *latusensu* causative meaning.

(24)

adjective	adjr→vr	verb	paraphrase
<i>firme</i> ‘firm’	<i>firm-</i>	<i>firmar</i> ‘to firm’	<i>tornar firme</i>
<i>murcho/a</i> ‘wilted’	<i>murch-</i>	<i>murchar</i> ‘to wilt’	<i>ficar murcho</i>
<i>melhor</i> ‘better’	<i>melhor-</i>	<i>melhorar</i> ‘to better’	<i>tornar/ficar melhor</i>

(25)

adverb	advr→vr	verb	paraphrase
<i>abaixo</i> ‘below’	<i>abaix-</i>	<i>abaixar</i> ‘lower’	<i>pôr abaixo</i>
<i>adiante</i> ‘further’	<i>adiant-</i>	<i>adiantar</i> ‘advance’	<i>pôr adiante</i>

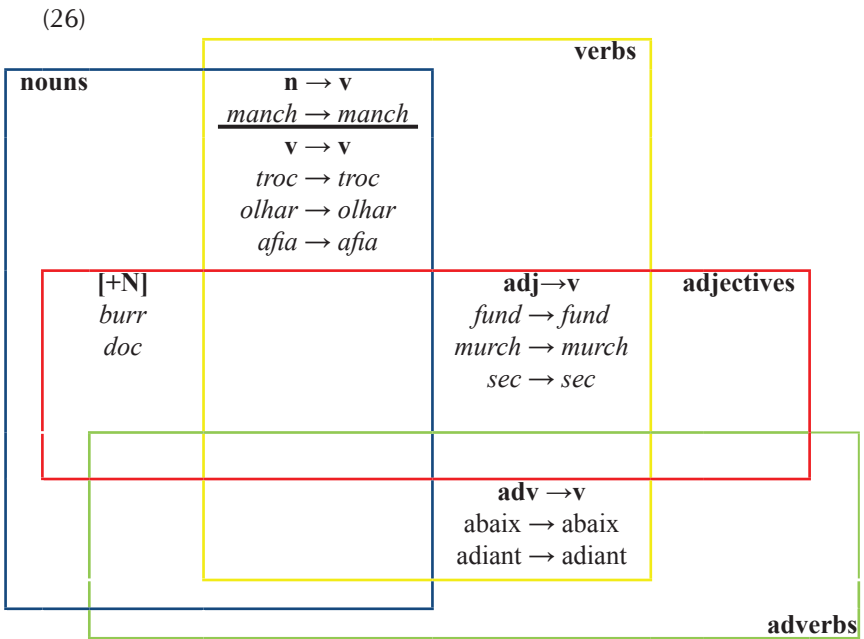
Etymology can provide valuable clues. The Portuguese noun *alarme*, or the English version *alarm*, are borrowings from the Italian noun *all'arme*, which was based on the interjective prepositional phrase *all'arme*, used to call the soldiers for a defensive intervention. In Portuguese, both the noun *alarme* (eventually under the form *alarma*) and the verb *alarmar* ‘to cause alarm’ are not registered prior to the 19<sup>th</sup> century, but, according to the authors of *Le Trésor de la Langue Française Informatisé* ([atilf.atilf.fr/dendien/scripts/tlfiv4/showps.exe?p=combi.htm;java=no](http://atilf.atilf.fr/dendien/scripts/tlfiv4/showps.exe?p=combi.htm;java=no;)), the French word *alarme* has been used since the beginning of the 14<sup>th</sup> century and the verb *alarmer*, is acknowledged as existing from the 17<sup>th</sup> century onwards. The authors of the *Tesoro della Lingua Italiana delle Origine* ([tlio.ovi.cnr.it/TLIO/](http://tlio.ovi.cnr.it/TLIO/)) consider that the first documented use of the word occurs in a 13<sup>th</sup> century text (the verb *allarmare* is not registered at all in this lexicographic source). So, historical information, namely in Italian where the noun first appeared, clearly favours the hypothesis that the use of the noun is available prior to use of the verb in several languages.

This is, so far, the most regular case of word formation by conversion, since any noun root, adjective root or adverb root prone to representing a culmination state may be converted into a 1<sup>st</sup> conjugation causative verb root.

6. Closing remarks

Our initial question about the different or common nature of the data in (1) found an obvious response: not all the instances of root sharing are alike. If conversion covers all of them, then we must admit that there are different types of conversion. We also concluded that in order to set them apart we need to take into account features from every domain of the grammar, which is probably the best argument for considering conversion as a lexical word formation process rather than a morphological, syntactic or semantic one.

In practical terms, the data in (1) that have been scrutinized in this paper can be arranged as in table (26):



## REFERENCES

- Bauer, L. & Varela, S. 2005. Conversion or zero-derivation: an introduction. In Bauer, L. & Varela, S. (Eds.), 7-18.
- Bauer, L. & Varela, S. 2005 (Eds.). *Approaches to Conversion / Zero-derivation*. Münster: Waxmann.
- Bechara, E. 2009. *Moderna Gramática Portuguesa*. 37ª edição. Rio de Janeiro: Nova Fronteira.
- Booij, G. E. 1986. Form and meaning in morphology: the case of Dutch 'agent nouns'. *Linguistics* 24: 503-517.
- Cunha, C. & Cintra, L. 1984. *Nova Gramática do Português Contemporâneo*. Lisboa: João Sá da Costa.
- Don, J. 1993. *Morphological Conversion*. Utrecht: Research Institute for Language and Speech.
- Don, J., Trommelen, M. & Zonneveld, W. 2000. Conversion and category indeterminacy. In *An International Handbook on Inflection and Word-formation* 1, 943-952.
- Villalva, A. 2000. *Estruturas Morfológicas. Unidades e Hieraquias nas Palavras do Português*. Lisbon: FCG and FCT.
- Villalva, A. 2008. *Morfologia do Português*. Lisbon: Universidade Aberta.
- Villalva, A. 2011. Os radicais no léxico do Português contemporâneo. *Atas do ICIEL* (in print). Salvador da Baía, Brasil.





# Deriving color adjectival nominalizations

Artemis Alexiadou

artemis@ifla.uni-stuttgart.de

*Universität Stuttgart (Germany)*

**ABSTRACT.** In this paper I examine two types of nominalizations related to color adjectives in Greek, a suffixed one and a neutral one, which I will compare to their English and Dutch (and German) counterparts. I show that the two differ in that suffixed nominalizations denote stage level properties, while neuter nominalizations denote individual level properties. This difference is due to the fact that suffixed nominalizations are count nouns, while neuter nominalizations are mass nouns. A comparison between Greek, and Dutch/German and English shows that languages have different nominalization strategies: nominalization can take place at three layers: at the root level, at the nP level, and finally at the DP level. This explains the differences in distribution and interpretation among the different nominalization types across languages.

**KEY WORDS.** Color adjectives, suffixed nominalization, neuter nominalization, stage level, individual level.

## 1. Introduction

In the recent literature, the extent to which natural languages exploit the distinction between (category preserving) inflection and (category changing) derivation to create subtle nuances of meaning that enrich the palette of reference to abstract objects has emerged as a new focus on the study of the interface between morpho-syntactic representations and their interpretation, see e.g. McNally & de Swart (2011), Villalba (2009) among others.

In this paper, I contribute to this discussion by looking at two types of nominalizations related to color adjectives in Greek, which I will compare to their English and Dutch counterparts. The first type, labeled here suffixed nominalization, is mostly derived via the addition of an affix to an adjectival stem, e.g. *-ada* (2a), or *-ila* (2b), and bears feminine gender. The second type, labeled here neuter nominalization,<sup>1</sup> simply surfaces in what looks like the neuter form of the adjective in question (2c) vs. (1):

---

<sup>1</sup> Here I refer to (2c) as neuter (adjectival) nominalization, following Villalba (2009).

- (1) to aftokinito ine prasino  
the car-nom is green-neut  
the car is green
- (2) a. i prasinada ton mation tis *suffixed nominalization*  
the green-ness the eyes hers  
the greenness of her eyes
- b. o tihos ehi endoni asprila  
the wall has intense whiteness  
there is an intense whiteness on the wall
- c. to prasino tu milu ine endono *neuter nominalization*  
the red the apple-gen is intense  
the red of this apple is intense

Dutch, as described in McNally & de Swart (2011), shows a three way distinction. Next to the inflected form (3b), corresponding to the neuter form in (2c), and the suffixed form (3c), there is also a non-inflected form of the adjective that can appear in nominal distribution, see e.g. (3a). A similar partition is found in German, see Alexiadou, Iordachioaia, Marzo & Umbreit (2012).

- (3) a. Misschien kun je het **rood** van de aardbeien nog een beetje roder maken?  
maybe can you the red of the strawberries yet a bit redder make  
'Maybe you can change the red (shade) of the strawberries so that it is a bit redder?'
- b. Het **rode** van de aardbeien, het witte van de mascarpone en de the red[+ -e] of the strawberries, the white[+ -e] of the mascarpone and the slagroom en het groene mintblaadje kleurden prachtig bij elkaar.  
cream and the green mint-leaf.dim colored beautifully with each-other

'The red (aspect) of the strawberries, the white (aspect) of the mascarpone and the cream, and the green mint leaf are a wonderful color combination.'

c. De **roodheid** van de huid kan achterliggende oorzaken hebben.  
the redness of the skin can deeper causes have  
'The redness of the skin can have deeper causes

In contrast, English lacks the inflected form, and only shows the contrast between two forms: *the red* and *the redness*.

The questions that I will deal with in this paper are:

- What are the differences between the various forms?
- How do these relate to their morpho-syntactic representations?
- How does the two way distinction in Greek relate to the two way distinction in English, and how do these relate to the Dutch three way distinction?

## 2. The morpho-syntax of de-adjectival nominalizations

### 2.1. The neuter nominalization

There are two possible analyses for examples such as the ones in (2c): either these are cases of nominalization of an adjective or these could be analyzed as an instance of nominal ellipsis, as in (4):

- (4) esi tha agorasis to prasino aftokinito ki ego to kokino ~~aftokinito~~  
you fut buy.2sg the green care and I the red one  
You will buy the read care and I the red one

At first sight, an analysis according to which such strings contain an elided noun seems plausible, since in Greek the noun for color, *hroma*, is also neutral. If, however, these cases involved ellipsis, we would expect the remnant to behave like an adjective, contrary to fact. Importantly, strings of the type in (2c) do not tolerate gradation (*\*to pio ble tu uranu* 'the more blue of the sky').

Importantly, there seems to be some evidence in favor of the

nominalization analysis. First of all, they allow adjectival modification, in contrast to their Dutch counterparts, see the contrast in (5) vs. (6). If this were a case of ellipsis, we would expect the color term to behave like an adjective and only allow adverbial modification, which is not the case:

- (5) to endono / \*endona ble tu uranu  
the intense(NEUT)/intensely blue the sky-gen
- (6) het intens/\*intense rode van de ondergaande zon  
the intense (ADV)/intense (ADJ) red of the setting sun

Again in contrast to their Dutch counterparts, they tolerate other determiners.

- (7) afto to kitrino ginete grigora roz/poli kitrino epese  
this the yellow becomes soon pink/much yellow is around
- (8) \*een/dit/veel rode  
a/this/much red

Unlike other types of neuter nominalizations, color nominals can only take an argument in the genitive and not in the PP form:

- (9) to ble tu uranu/\*me ton urano  
the blue the sky.Gen/at the sky
- (10) to kalo me ti Maria  
the good at the Mary

In addition, there is a difference in meaning between (9) and (10): while (9) entails the predication of the adjective about the genitive DP, the neuter nominal in (10) does not. The latter rather establishes a partitive relation with the genitive DP, which is missing in the former:

- (11) to kalo me ti Maria = > I Maria ehi kati kalo  
the good at the Mary Mary has something good
- (12) to ble tu uranu = > O uranos ine ble  
the blue the sky-gen the sky is blue

Further examples illustrating the presence of a partitive reading with non-color nominalizations are given in (13), from Alexiadou, Iordachioaia, Marzo & Umbreit (2012):

- (13) a. To kalo tu Jani ine i ipomoni tu.  
the good the John-gen is the patience his  
The good thing about John is his patience.
- b. To kalo me ta vilia tu Larson den ine i dedalodis plokas.  
the good with the books the Larson-gen neg are the daedalian plots  
The good thing about Larson's books is not the daedalian plots.

The situation in (13) seems similar to what has been observed for Spanish LO- nominalizations, see Villabba (2009), where the data come from, building on Bosque & Moreno (1990):

- (14) a. Lo interesante del libro es el primer capítulo.  
LO interesting of-the book is the first chapter  
The interesting part of the book is the first chapter.
- b. Me asusta lo peligroso de la empresa.  
to.me frightens LO dangerous of the.FEM enterprise  
It frightens me how risky the enterprise is.

Importantly, such a partitive reading is absent from color nominals. There is a further difference between color nominalizations and other neuter nominalizations concerning productivity: the neuter nominalization of non-color adjectives, is idiosyncratic, in the sense that it is possible with some adjectives but not with all:

(15) a.	perifanos proud	perifan-i-a pride-fem	*to perifano the proud	
	b.	ilikrinis honest	ilikrin-i-a honesty-fem	*to ilikrines the honest
	c.	mikro small-petty	mikro-tit-a pettiness	*to mikro the petty

In contrast, all color adjectives produce nominals of the type in (2c). This, in addition to the interpretational differences outlined above seems to suggest that the neuter nominal in (13) is a case of substantivization, i.e. nominalization out of a root, as proposed in Giannakidou & Stavrou (1999), while color terms seem to involve a case of de-adjectival nominalization. The question then arises, what about the suffixed nominalization.

## 2.2. The suffixed nominalization

The suffixed nominalizations in *-ada* or *-ila* are clearly nouns. The following observations support this claim. First, they take adjectival modification:

- (16) i endoni kokinila  
the intense redness

Second, they combine with a variety of determiners:

- (17) tin/afti i /poli asprila  
the/this the/much whiteness

Third, they can pluralize, in contrast to the neuter nominalization:

- (18) a. i apriles  
the whitenesses
- b. \*? ta aspra  
the white-pl

This suggests that the suffixed nominalization is a count noun, while the neuter nominalization behaves like a mass noun.

There is a further difference between the two types of nominalization, and this relates to their temporal structure. While the suffixed nominalization can occur as the subject of a predicate such as *last*, this is not possible with the neuter nominalization. If, as it is standardly assumed, predicates such as *last* require eventualities as their subject, the contrast below suggests that suffixed nominalizations have some temporal structure (cf. Martin 2010):

(19) *i asprila kratise deka meres*  
the whiteness lasted 10 days

(20) *\*to aspro kratise 10 meres*  
the white lasted 10 days

A second piece of evidence in favor of this view comes from the observation that suffixed nominalization can be spatio-temporally restricted:

(21) a. *i kokinila sto heri mu me anisihise*  
the redness in the hand my me worried

b. *\*to kokino sto heri mu me anisihise*  
the red in the hand my me worried

c. *i kokinila ton proigumenon imeron eksafanistike*  
the redness the past days disappeared

d. *\*to kokino ton proigumenon imeron eksafanistike*  
the red the past days disappeared

In conclusion, suffixed nominalization behave like stage-level predicates (see also Roy 2010), while neuter nominalizations behave like individual level predicates. While the former are count nouns, the latter are mass nouns.

Note that both suffixed nominalizations, i.e. *-ada* and *ila* ones, can

have such stage-level readings. In fact, most of the time, the two forms are used interchangeably. Occasionally, suffixed nominalizations can take idiosyncratic interpretations, e.g. *prasinada* 'green-ness' can also mean *grass*:

- (22) a. i asprila tu paniu                      i asprada tu paniu  
          the whiteness the cloth-gen        the whiteness the cloth
- b. to derma mu ehi kokinadas/kokiniles/aspriles/asprades  
              my skin has rednesses/whitenesses

Finally, as far as the behavior of the accompanying genitive is concerned, I note that this again can appear only in the genitive form, and it does not differ in interpretation from the genitive that co-occurs with the neuter nominalization, suggesting that also in this case the genitive is an argument of the embedded adjective.

### 3. Towards an analysis

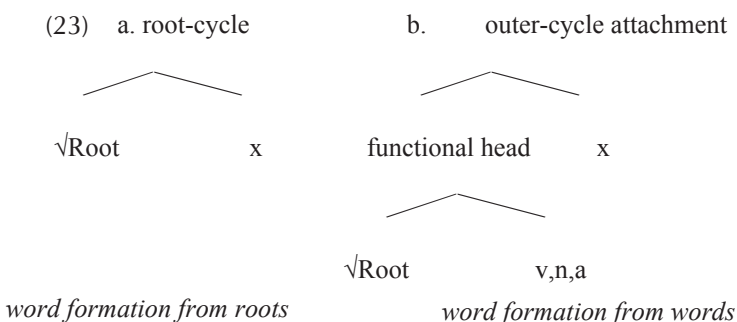
#### 3.1. Background

I will assume a view on word formation couched within the framework of Distributed Morphology (see Arad 2005, Marantz 2001, Embick 2010). From this perspective, the following pieces constitute the building blocks of word formation:

1. Language has atomic, non-decomposable, elements, called roots.
2. Roots combine with the functional vocabulary and build larger elements.
3. Roots are category neutral. They are then categorized by combining with category defining functional heads.

There are two cycles for word-formation (Marantz 2001/to appear), i.e. two levels at which categorizing affix can appear: the root cycle and the outer-cycle. Affixation at the root cycle leads to word formation out of roots, while affixation at the level that includes already a categorizing affix involves word formation out for words:





The two processes have very different properties. To begin with, roots are assigned interpretation at cycle (23a), i.e. the constraint in (24) holds:

(24) Locality constraint on the interpretation of roots/Cyclic generalizations:

Roots are assigned an interpretation in the environment of the first category-assigning head with which they are merged. Once this interpretation is assigned, it is carried along throughout the derivation. (Arad 2005, Embick 2010).

Moreover, merger with root implies:

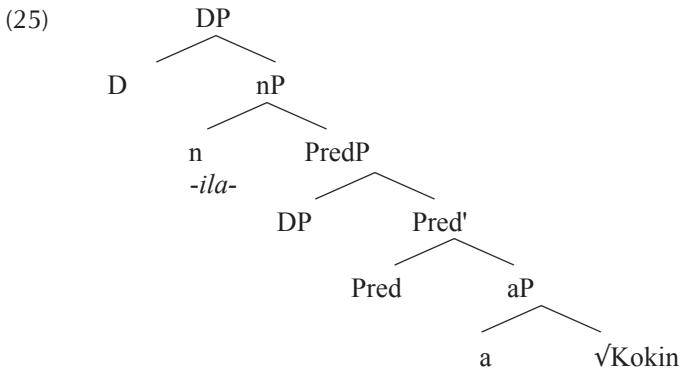
1. negotiated (apparently idiosyncratic) meaning of root in context of morpheme
2. apparent semi-productivity (better with some roots than others)
3. meaning of construction cannot be an operation on “argument structure” but must depend on root semantics independent of argument structure
4. corollary of the above: cannot involve the “external argument”

In contrast, merger above a category-determining morpheme implies:

1. compositional meaning predicted from meaning of stem
2. apparent complete productivity
3. meaning of structure can involve apparent operation on argument-structure
4. can involve the external argument. (Marantz 2001/to appear).

### 3.2. Deriving the two types of adjectival nominalizations

We have seen that both types of adjectival nominalization are actually nominal. Both suffixed and neuter nominalizations are productive and inherit the argument of the adjective. From the perspective of the framework introduced in section 3.1, this means that the root must be first categorized by *a*, an adjectivizer, and then by *n* which will host the nominalizing suffix, as in (25). I follow Roy (2010) and assume that on top of the *aP*, the nominalization also includes a *PredP* (see also Bowers 1993), which hosts the argument of the adjective inherited by the nominal itself.

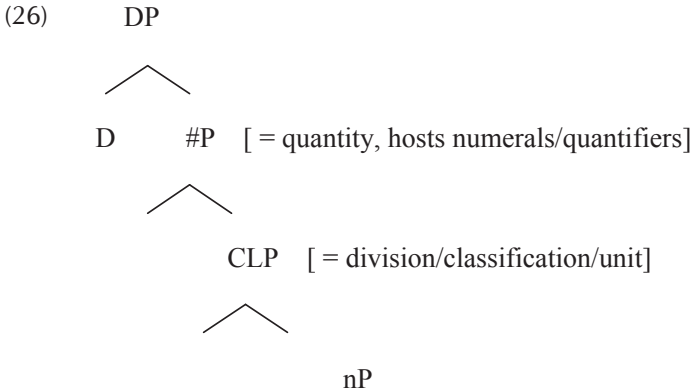


In view of the fact that the neuter nominalization realizes a genitive argument of which the property of the adjective is predicated, it should be similar in structure with the suffixed nominalizations, with the difference that they do not have a suffix.

The question that arises is then the following: if (25) is the structure for both the suffixed and the neuter nominalization, how can we account for the differences between the two types in terms of mass vs. count interpretation? Second, how is the stage-level/individual level distinction grammatically realized?

In order to deal with these issues, I will build on Husband (2006), who argues that these two properties are related. To begin with, recent work on the nominal structure of noun phrase argues that count and quantity interpretation arise from the presence of a specific functional projection within the extended projection of the nominal, see Borer (2005) for detailed

discussion. Building on Borer (2005), different functional projections between the DP and the NP are assumed: (i) the quantity phrase (#P in her system; similar but not equal to NumberP) and the Classifier Phrase (CLP in her system). All nouns enter the derivation as mass, and become count in the syntax, via ClassP.



In view of the fact that suffixed nominalizations can be pluralized, they must contain CLP. That is, suffixed nominalizations contain the layers #P and CLP between nP and DP, while neuter nominalizations lack these layers, see (27). Crucially, following Husband (2006), a stage level interpretation in nominals is related to the presence of the quantity phrase. As Husband argues, a quantity structure in the stative domain is interpreted always as stage-level.



#### 4. Crosslinguistic variation

In this final section, I will turn to issues concerning cross-linguistic variation. I mentioned in section 1 that languages seem to cut the domain of reference to abstract objects differently. First, there are languages like Greek, which have two types of nominalizations: the neuter one and the suffixed one, respectively. Second, there are languages like Dutch and German that show a three way distinction: next to the two types of nominalizations found

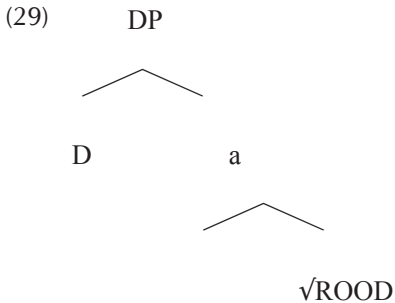
in Greek, they also have a non-inflected form. Third, there are languages like English which show a two-way distinction but different from the one found in Greek: they only have suffixed and non-inflected nominalizations.

McNally & de Swart (2013) address this issue and identify three point of variation: i) the presence of number inflection in the nominal system, ii) the semantics and productivity of derivational morphology that forms nouns from adjectives, and iii) the existence of specifically neuter definite articles.

Let me begin by contrasting the inflected adjective found in Germanic to its Greek counterpart. As McNally & de Swart (2011), the inflected adjective in Dutch does not behave like a noun. It can only appear with a specific determiner, namely *het*, (28):

- (28) \**een/dit/veel rode*  
one/this/many red

This then suggests, as McNally & de Swart (2011) also argue, that the inflected forms are actually adjectives that become nominal in the context of a determiner:



The difference between Greek neuter nominalizations and their Germanic counterparts thus is captured in the model outlined in 3.1 by means of the presence of nP. The categorizer n nominalizes a root or another category, but importantly it brings in a full internal nominal syntax, i.e., gender features and declension information. This means that it will force the resulting structure to fully behave like a noun, i.e., only adjectives will be allowed as modifiers (adjectives usually agree with the noun), and all types

of determines should be licit.

However, languages seem to also nominalize structures by means of a D alone (lordachioaia to appear). This has been argued to be the case, among others, with the verbal gerund in English, where the possessive is the only nominal marking, everything else indicates a verbal structure (see Alexiadou 2001, Alexiadou, lordachioaia & Schäfer 2011):

(30) John's **constantly**/\***constant** reading this novel

In this respect, Dutch inflected nominalizations are similar to the verbal gerund in English, because they lack an nP and are nominalizations by D alone. Importantly, D introduces an external nominal syntax, but not an internal one; it thus basically accommodates a structure of a different category into a nominal context. The crucial evidence in favor of this view is the fact that the two nominalizations sharply contrast with respect to the selection of determiners. While neuter nominalizations in Greek are compatible with just any determiner, Dutch inflected forms are restricted to the definite determiner. I assume here that the un-inflected form in Dutch (and German), and the suffixed nominalization found in these languages behave similarly to the Greek neuter and suffixed nominalization respectively.

Turning now to English, Kennedy & McNally (2010) point out that color nouns are routinely used to refer to colors, and as exemplified in (31), they are mass nouns. This is shown by their ability to appear in bare singular form and their compatibility with mass quantifiers.

- (31) a. Green was a surprising choice for the color of the dining room.  
b. There was a little/not much green in the carpet.

McNally (2011) argued that color nouns differ from suffixed nominalizations in that only the latter can pick out proxy properties, color terms cannot. Taking as point of departure the framework put forth in Kennedy & McNally (2010), color adjectives are argued to be ambiguous between a true color reading and a so-called proxy reading, where proxy is understood as having the property denoted by the color term:

- (32) a. ...if there is blueness around the mouth....  
b. ...if there is blue around the mouth...

Since color adjectives can have both proxy and true color readings, McNally argues that the contrast above shows that a) *redness* involves nominalization of an adjective, and b) the color noun is more basic than the color adjective. In turn, from the perspective of the framework in 3.1, this suggests that *red* involves nominalization of a root.

On the basis of these observations, we can now summarize the analysis of the different types of color nominalizations. Suffixed nominalizations in all languages are nominalizations of an adjective. Neuter nominalizations in Greek are similar to their suffixed counterparts, the difference between the two relating to the presence of quantity structure in the latter but not in the former. Inflected nominalizations in Dutch and German are nominalizations of an adjective, which, however, acquires nominal properties via the presence of a determiner. Non-inflected nominalizations in Dutch and German are again similar to their suffixed counterparts. In contrast, bare nominalizations in English are nominalizations of a root.

## 5. Conclusions

In this paper, I examined two types of nominalizations related to color adjectives in Greek, a suffixed one and a neutral one, which I compared to their English and Dutch (and German) counterparts. I showed that the two differ in that suffixed nominalizations denote stage level properties, while neuter nominalizations denote individual level properties. This difference is attributed to the fact that suffixed nominalizations are count nouns, while neuter nominalizations are mass nouns. A comparison between Greek, and Dutch/German and English showed that languages have different nominalization strategies: nominalization can take place at three layers: at the root level, at the nP level, and finally at the DP level. This explains the differences in distribution and interpretation among the different nominalization types across languages.

REFERENCES

- Alexiadou, A. 2001. *Functional structure in nominals: nominalization and ergativity*. Amsterdam: John Benjamins.
- Alexiadou, A., G. Iordăchioaia & F. Schäfer. 2011. Scaling the variation in Romance and Germanic nominalizations. In: P. Sleeman & H. Peridon (Eds.). *The noun phrase in Romance and Germanic*. Amsterdam: John Benjamins, 25-40.
- Alexiadou, A. & Iordăchioaia, D. Marzo & B. Umbreit. 2012. Comparing lexicalist and syntactic insights on deadjectival nominalizations. Paper presented at BCGL 7, Brussels, December 2012.
- Arad, M. 2005. *Roots and patterns*. Springer: Dordrecht.
- Borer, H. 2005. Structuring sense vol. I: *In name only*. Oxford: Oxford University Press.
- Bowers, J. 1993. The syntax of predication. *Linguistic Inquiry*. 24 (4): 591-656.
- Embick, D. 2010. *Localism vs. Globalism in Morphology and Phonology*. Cambridge, Mass.: MIT Press.
- Giannakidou, A. & M. Stavrou. 1999. Nominalization and ellipsis in the Greek DP. *The Linguistic Review*. 16: 295-331.
- Husband, E. 2006. Stage-level/individual-level predicates and aspect. Ms. Michigan State University.
- Iordăchioaia, G. to appear. The interaction between NP and DP in nominalizations. Proceedings of NELS 43.
- Kennedy, C. & L. McNally. 2010. Color, context and compositionality. *Synthese* 174: 79-98.
- Marantz, A. 2001/To appear. Words and Things. Ms., MIT & NYU.
- Martin, F. 2010. Stage-level and individual level readings of quality nouns: a study in aspectual morpho-semantics. Paper presented at Going Romance 23, Leiden, December 2010.
- McNally, L. 2011. Color terms: a case study in natural language ontology. Paper presented at the Workshop on the syntax and semantics of nounhood and adjectivehood. Barcelona, March 2011.
- McNally, L. & H. de Swart. 2011. Inflection and derivation: how adjectives and nouns refer to abstract objects. Proceedings of the *18th Amsterdam Colloquium*: 425-434.
- McNally, L. & H. de Swart. 2013. Inflectional and derivational strategies for referring to abstract objects. Paper presented at the Workshop on adjectives and their nominalizations, Stuttgart, February 2013.

- Roy, I. 2010. Deadjectival nominalizations and the structure of the adjective. In: A. Alexiadou; M. Rathert (Eds.). *The syntax of nominalizations across languages and frameworks*. Berlin; Mouton de Gryuter, 1291-58.
- Sleeman, P. 1996. *Licensing empty nouns in French*. Ph.D. Dissertation. HIL.
- Villalba, X. 2009. Definite adjective nominalizations in Spanish. In: M.T. Espinal; M. Leonetti; L. McNally (Eds.). *Proceedings of the IV Nereus International Workshop Definiteness and DP Structure in Romance Languages*.
- Villalba, X. & Bartra-Kaufmann, A. 2009. Predicate focus fronting in the Spanish determiner phrase. *Lingua* 120: 819-849.



# Deadjectival human nouns: conversion, nominal ellipsis, or mixed category?

Petra Sleeman

p.sleeman@uva.nl

*ACLCL, University of Amsterdam (The Netherlands)*

ABSTRACT. Whereas deadjectival nouns referring to humans such as *the Germans* have been analyzed as the result of morphological conversion, the human construction *the rich* in English has been analyzed as a special case of nominal ellipsis. In this paper counterarguments are presented against the ellipsis analysis, mainly focusing on the human construction in Dutch, which has mixed adjectival and nominal properties. Traditionally, deadjectival human nouns ending in the suffix *-e* are analyzed as the result of morphological derivation. In the ellipsis analysis the suffix *-e* is analyzed as an inflectional suffix rather than a derivational one, licensing an empty noun. The plural suffix *-n* and the determiner would provide the human interpretation. In this paper an analysis in the framework of Distributed Morphology is proposed, which is a combination of the ellipsis analysis (without an empty noun) and the traditional derivational/conversion analysis.

KEY-WORDS. Deadjectival noun, human construction, ellipsis, mixed category, Dutch.

## 1. Introduction

In many languages human nouns can be derived from an adjective by means of an overt suffix, meaning 'N that has the property A'. This is exemplified in (1):

- |        |  |         |
|--------|--|---------|
| (1) a. | <i>strange</i> 'strange' + <i>er</i> → stranger                    | ENGLISH |
|        | b. <i>riche</i> 'rich' + <i>ard</i> → <i>richard</i> 'rich person' | FRENCH  |
|        | c. <i>dik</i> 'fat' + <i>erd</i> → <i>dikkerd</i> 'fat person'     | DUTCH   |

However, adjectives can also be used to refer to a person without a nominalizing suffix. This is exemplified in (2)-(4) for English:

- (2) Of all these musicians, John is the best.
- (3) the Germans and the Russians
- (4) the rich and the poor

Kester (1996a,b) argues that the analysis of the relevant noun phrases in (2)-(4) is not the same. According to Kester, the noun phrase *the best* in (2) can best be analyzed as a case of nominal ellipsis: the noun is missing in the noun phrase and gets its semantic content from an antecedent noun in the previous discourse:

- (5) Of all these musicians, John is the best (musician).

For (3) she argues that the noun phrases contain a noun that is the result of conversion: the adjectives have been lexically transformed into nouns, as witnessed by the fact that they can be morphologically pluralized.

Although the noun phrases in (4) are semantically plural, they do not bear a plural morpheme. Therefore Kester argues that they do not contain a nominalized adjective, as in (3), but that, just as in (2), the noun phrases in (4) are cases of ellipsis: there is an empty noun and *rich* and *poor* are adjectives.<sup>1</sup>

Kester claims that the Dutch equivalents of the noun phrases in (4) do not contain a nominalized adjective either, but do also contain an empty noun, in spite of the fact that they can be pluralized, which is expressed by the plural morpheme *-n* in (6). Kester analyzes the schwa in the noun phrases in (6) not as a nominalizing affix, but as an inflectional affix on the adjective preceding the empty noun, represented as [e] in (7).

- (6) *de rijken en de armen*  
the rich and the poor  
'the rich and the poor'

---

<sup>1</sup> A similar analysis of (2)-(4) is put forth by Borer & Roy (2010).

- (7) *de rijk-e-n*            *[e] en de arm-e-n [e]*  
the rich-ADJ.INF-PL      and the poor.ADJ.INF-PL  
'the rich and the poor'

The inflectional properties of *rijken* and *armen* in (6) make them mixed categories: the schwa is an adjectival property, whereas the plural is a nominal property.<sup>2</sup>

Within the model of the Lexicalist Hypothesis (Chomsky 1970), morphological operations such as derivation or conversion took place in the Lexicon. Forms such as those in (2) and (3) were the result of the morphological conversion of an adjective into a noun. In this model, the schwa on the adjective in (7) could not be interpreted as an inflectional suffix expressing syntactic concord, as in Kester's ellipsis analysis, but was analyzed as a derivational suffix, converting the adjective into a noun, just like the suffixes in (1). More recent research, couched within the framework of Distributed Morphology (Halle & Marantz 1993, 1994; Harley & Noyer 1999; Embick & Noyer 2006), makes use of Syntax to account for the properties of mixed categories (e.g., Borsley & Kornfilt 2000; Alexiadou 2001; Embick 2004). In the light of syntactic analyses of mixed categories that have been proposed, the analysis of the noun phrases in (2)-(4) is re-examined in this paper.

The paper is organized as follows. In §2, an analysis of mixed categories within a Distributed Morphology approach is illustrated. In §3, an analysis of deadjectival nominalizations in terms of ellipsis is discussed. In §4, an analysis of deadjectival nominalizations within a Distributed Morphology approach is proposed. A conclusion is presented in §5.

## 2. Mixed categories and conversion

Due to their category-shifting nature, mixed categories can present properties both of the original base and of the resulting category. In the Principles and Parameter framework of the Generative model (Chomsky 1981) mixed categories posed a theoretical problem. The X'-structure of phrases made a category switch within syntax theoretically impossible. This

---

<sup>2</sup> In Dutch, attributive adjectives can take an inflectional schwa. Nouns, but not adjectives, can take a plural morpheme in the plural.

is illustrated by nominalized infinitives in Dutch.

Nominalized infinitives are verbs used as nouns and they can present properties of both categories. In the literature, the more verbal types are generally called verbal infinitives and the nominal types are called nominal infinitives (e.g., Plann 1981, Alexiadou, Iordăchioaia & Schäfer 2011), a distinction that has also been made by Chomsky (1970) for English gerunds. Verbal properties are the combination with a subject, direct complementation, i.e. the combination with direct objects, the combination with auxiliaries, and the combination with adverbs. Nominal properties are the use of a determiner (article, possessive or demonstrative pronoun), modification by an adjective instead of an adverb, and the combination with genitives instead of a subject or a direct object, gender distinctions, and pluralization. In its most verbal use, the nominalized infinitive is used without a determiner, but occurs in argument position. In its most nominal use, the nominalized infinitive functions in all respects as a noun. Verbal infinitives and nominal infinitives are situated on a scale between these two extremes. The middle of the scale contains nominalized infinitives in which verbal and nominal properties are mixed.

The following examples illustrate the ambiguity of the Dutch nominalized infinitive. In (8), taken from Sleeman (2001), the infinitive is purely verbal: there is no determiner and the direct object precedes the infinitive (Dutch is an SOV language). In (9), also taken from Sleeman (2001), the infinitive is purely nominal. Within the Lexicalist model, its nominal category is the result of a conversion rule in the Lexicon, changing a verb into a noun.<sup>3</sup>

(8) *Alcohol drinken kan schadelijk zijn.*  
alcohol drinking can harmful be  
'Drinking alcohol can be harmful.'

(9) *Heb je je drinken al op?*  
have you your drinking already finished  
'Have you already finished your drink?'

---

<sup>3</sup> In the case of "relisting", a rule simply changes a category into another without suffixation. Another rule that has been proposed to account for the category shift, zero-derivation, makes use of a null suffix: [V [Ø]<sub>N</sub>]<sub>N</sub> (see, e.g. Don 2005).

Sentences (10)-(12) have been taken from Ackema & Neeleman (2004):

*Deze zanger is vervolgd voor...*

This singer has-been prosecuted for...

(10) ... *dat stiekem succesvolle liedjes jatten*

that sneaky.ADV successful songs pinch

'that sneaky pinching of successful songs'

(11) ... *dat stiekeme succesvolle liedjes jatten*

that sneaky.ADJ successful songs pinch

'that sneaky pinching of successful songs'

(12) ... *dat stiekeme jatten van succesvolle liedjes*

that sneaky.ADJ pinch of successful songs

'that sneaky pinching of successful songs'

In (10), the infinitive is modified by an adverb and is preceded by a direct object. These are verbal properties. Differently from (8), the infinitive in (10) is introduced by a determiner, which is a nominal property. In (11), the infinitive is preceded by its direct object (verbal property), but is introduced by a determiner and is modified by an adjective (nominal properties). In (12), instead of a prenominal direct object, there is a prepositional phrase, which is a nominal property. In these five examples, the nominal infinitive changes thus from purely verbal (8) into purely nominal (9), with three intermediary steps (10)-(12).

Both within the Lexicalist Hypothesis and in the traditional X' model, it is difficult to account for these intermediary steps. They show that the infinitive has verbal and nominal properties at the same time. The X' model does not allow the insertion of a category with verbal properties, i.e. a verb, under, e.g., a nominal head, accounting for the mixed behavior.

The Distributed Morphology model (Halle & Marantz 1993, 1994) and comparable models offered a solution. Category-neutral roots are dominated by, e.g., verbal and nominal functional projections. In this way, the inner verbal behavior (lower verbal functional projections) and the outer nominal

behavior (higher nominal functional projections) can be accounted for.

Alexiadou, Iordăchioaia & Schäfer (2011) analyze verbal and nominal non-derived nominalizations (infinitive, supine, gerund) in several European languages (see also Sleeman 2010). In the verbal type, the verbal properties dominate, in the nominal type the nominal properties dominate. On the basis of the presence of the subject of the infinitive in the Spanish example (13), Alexiadou et al. (2011) analyze the verbal nominalized infinitive in Spanish as the most verbal type. In its structure, it has TP as its highest verbal functional projection (14).

- (13) *el cantar yo la Traviata*  
 the sing.INF I.NOM the Traviata  
 ‘me singing the Traviata’

- (14) [DP [TP [Aspect [VoiceP [vP [Root]]]]]]

For nominal non-derived nominalizations a structure as in (15) is proposed by Alexiadou et al.:

- (15) [DP [(NumberP) [ClassP [nP [AspP [VoiceP [vP [Root]]]]]]]]

Languages differ in the presence or nature of functional projections in the structure of nominal non-derived nominalizations. Languages like English and Romanian have, e.g., ClassP [+ count], because the nominal non-derived nominalizations in these languages can be pluralized, as illustrated for the English gerund and the Romanian infinitive:

- (16) a. *the repeated killings of unarmed civilians*  
 b. *demolările frecvente ale cartierelor vechi*  
 demolish.INF-PL frequent-PL of quarters.GEN old  
 ‘the frequent demolitions of old quarters’

Mixed categories have mixed properties, expressed by the presence of both verbal and nominal functional projections in structures (14)-(15). The

purely verbal infinitive in (8) would only have verbal functional projections in its structure, whereas the purely nominal one in (9) – the result of conversion within a Lexicalist Hypothesis – would only have nominal functional projections in its structure.

In the next section I will turn to the presentation of the subject of this paper, non-derived deadjectival human nouns in Dutch, and Kester's (1996) analysis.

### 3. An ellipsis analysis of nominalizations

Traditionally, deadjectival human nouns as *de zieke* 'the sick person' in Dutch are analyzed as the result of derivation, through which the derivational suffix *-e*, just like a null suffix in a conversion analysis for *the poor* in (4) in English (see fn. 3), converts the adjectival base into a noun (e.g., Geerts et al 1984):

(17) [A [e] ]<sub>N</sub> ]<sub>N</sub>

Kester (1996a,b) argues, however, that they should be analyzed as cases of nominal ellipsis, on a par with cases of genuine nominal ellipsis, as (18), in which [e] represents the omitted noun:

(18) *Jan kocht de rode auto en de groene [e].*  
John bought the red car and the green  
'John bought the red car and the green one.'

Kester presents several arguments against a nominalization, i.e. derivational, analysis. First, human nouns in Dutch can be preceded by degree modifiers and adverbs, as in (19)-(20), they can appear in comparative and superlative forms, as in (21), and they can be preceded by adjuncts and complements, as in (22):

(19) *de zeer rijken*  
the very rich  
'the very rich'

- (20) *de vandaag gearriveerden*  
the today arrived  
'the people arrived today'
- (21) a. *de nog ziekeren*  
the yet sicker  
'the people that are (even) sicker'
- b. *de allerbesten*  
the all-best  
'the best of all'
- (22) a. *de naar onze maatstaven zeer vermogenden*  
the to our standards very wealthy  
'the people that are very wealthy to our standards'
- b. *de hiervan afhankelijk*  
the here-on dependent  
'the people dependent on this'

Second, the complement precedes the adjective in (22b), which suggests that the adjective is in prenominal position, cf. (23). When the adjective is in predicative position, the complement can precede or follow it (24):

- (23) a. *de hiervan afhankelijke vluchtelingen*  
the here-on dependent refugees  
'the refugees dependent on this'
- b. \**de afhankelijk hiervan*  
the dependent here-on  
'the people dependent on this'
- c. \**de afhankelijke hiervan vluchtelingen*  
the dependent here-on refugees  
'the refugees dependent on this'



(24) a. *De vluchtelingen zijn hiervan afhankelijk.*

the refugees are here-on dependent

'The refugees are dependent on this.'

b. *De vluchtelingen zijn afhankelijk hiervan.*

the refugees are dependent here-on

'the refugees are dependent on this'

Third, the human nouns do not allow diminutive formation:

(25) \**een zieketje*

a sick.DIM

'a small sick person'

Fourth, the human construction has an irregular plural. Contrary to most nouns ending in schwa in Dutch, they form the plural by means of the affix [-n] instead of [-s]:

(26) a. *de zieken*

the sick.pl

'the sick'

b. \**de ziekes*

the sick.pl

'the sick'

Instead of the traditional nominalization analysis, Kester proposes a syntactic analysis, viz. an ellipsis analysis, just as for (18). She adopts Lobeck's (1995) analysis of nominal ellipsis, according to which the adjective in the ellipsis construction has to formally license and identify small *pro* (Rizzi 1986), the head of the noun phrase. The assumption that the human construction is an adjective modifying an elided noun in a syntactic ellipsis construction, explains, according to Kester, why modification by degree modifiers and adverbs (19-20) and comparative and superlative formation (21) are possible, why the complement precedes the human construction

(22), and why diminutive suffixation is not allowed (25).

As for the plural suffix, which in the standard case does not show up on adjectives in Dutch, Kester proposes that it identifies *pro* endowed with the inherent person features [+human, +plural]. As an argument in favor of her analysis, she presents the data in (27):

- (27) *Ik heb er enkele(n) in Amsterdam gezien.*  
 I have of-it some in Amsterdam seen  
 'I have seen some of them in Amsterdam.'

She states that the use of the plural ending *-n* is optional in this example, but that its presence leads to the [+human] interpretation.

Giannakidou & Stavrou (1999) argue against an ellipsis analysis and in favor of a nominalization analysis for the human noun in Greek. One of their most important arguments is that the human construction has a kind interpretation. Since the Greek human construction does not have the same properties as the Dutch human construction – degree modifiers, adverbs, comparatives and superlatives are not possible – Giannakidou & Stavrou remain agnostic as to whether their analysis of the Greek human construction can be extended to languages such as Dutch. In what follows I will also for Dutch argue against an ellipsis analysis.

The first problem with Kester's analysis concerns the properties of the empty noun. Kester adopts Lobeck's (1995) analysis of the empty noun as a small *pro*, and argues that in Dutch in all ellipsis cases *pro* is formally licensed by adjectival inflection. In the standard ellipsis case, *pro* does not have features, but is identified by its antecedent. In (18), *pro* is interpreted as [–human, –plural, –generic] on the basis of the context. In the human construction, *pro* is identified as [+human, +plural, ±generic] by the suffix *-n* on the adjective, and as [+human, –plural, ±generic] by the determiner if there is no *-n* on the adjective (*de rijke* 'the rich person').<sup>4</sup> The analysis

---

<sup>4</sup> To account for the uninflected forms *the rich* and *the poor* in the English example (4) Kester assumes that [+human], [+plural], [+generic] are the default features of *pro*, which is non-referential in this case and does not have to be formally licensed by adjectival inflection. In this way she accounts for the presence in English of the human construction with a generic interpretation only. However, if adjectival inflection is not needed to license the default features of *pro*, the question arises why adjectival inflection is needed in the plural generic reading in Dutch:

(i) \**de rijk\*(e) en de arm\*(e)*  
 the rich and the poor  
 'the rich and the poor'

of the empty nominal either as *pro* or as an empty noun in the Lexicon (Panagiotidis 2003) has been abandoned in more recent accounts of nominal ellipsis, in favor of analyses in which the noun is either deleted at PF under identity with an antecedent (see, e.g., Depiante & Masullo 2004), or is not inserted as a phonological matrix at the level of PF (see, e.g., Kornfeld & Saab 2002). In both analyses, identity with the antecedent plays a crucial role. However, in Kester's analysis of the human construction, there is no antecedent. In another recent account, focalization or a contrastive interpretation plays an important role (see, e.g., Corver & van Koppen 2009, Eguren 2009). In the analysis of the human construction presented above the adjective is not focalized or contrasted.

A second problem is Kester's analysis of the schwa, which she analyzes as an inflectional morpheme. In canonical DPs the inflectional schwa is used on adjectives in Dutch if the determiner is definite or plural or, with an indefinite singular article, if the noun is a common noun. This means that in the human construction it would have to be assumed that *pro* also has the feature [+common], which is, however, a morphological feature and not a semantic feature as [human], [plural], and [generic]. Furthermore, human nouns that would have been "omitted" in this context, are neuter nouns like *mens* 'human' or *persoon* 'person', which would not require an inflected adjective in an indefinite, singular context (28). However, in an indefinite singular noun phrase the adjective in the human construction bears the suffix –e (29):

(28) a. *een ziek persoon*

a sick person.N

'a sick person'

b. *een ziek mens*

a sick human.N

'a sick human being'

---

Following Lobeck (1995), Kester assumes that, because of the absence of adjectival inflection in English, nominal ellipsis is not licensed in other contexts than in the human construction. Sleeman (1996), however, shows that nominal ellipsis in which the antecedent is recovered from the context is not excluded in English. She proposes that partitivity and not adjectival inflection is the licensing factor.

- (29) *een zieke*  
a sick  
'a sick person'

Participial forms ending in *-en* cannot be inflected in their attributive use. However, in the human construction they bear the suffixal *-e*:<sup>5</sup>

- (30) *de verlaten(\*e) echtgenote*  
the abandoned-ADJ.INF spouse  
'the abandoned spouse'

- (31) *de verlatene*  
the abandoned-NOM.  
'the abandoned person'

A third problem is the plural *-n*. In genuine elliptical constructions, plural indefinite quantifiers in Dutch bear the suffix *-e* (32). In a human reading (32b) *-n* may be added (33):

- (32) a. *Ik heb veel boeken. Sommige heb ik gelezen.*  
I have many books. Some have I read  
'I have many books, some of which I have read.'

- b. *Ik heb veel vrienden. Sommige zie ik bijna nooit.*  
I have many friends. Some see I almost never  
'I have many friends, some of whom I almost never see.'

- (33) *Ik heb veel vrienden. Sommigen zie ik bijna nooit.*  
I have many friends. Some see I almost never  
'I have many friends, some of whom I almost never see.'

However, in my intuition, in combination with the quantificational pronoun *er* adding *-n* in (27) is incorrect, even in the human interpretation

---

<sup>5</sup> Participles used in the human construction can be combined with an agent, a complement or an adjunct, as (20) shows.

(34). If *-n* is added, *er*, which is obligatory in combination with an indefinite elliptical DP in object position, cannot be used (35):<sup>6</sup>

(34) *Ik heb er enkele(\*n) in Amsterdam gezien.*

I have of-it some in Amsterdam seen

'I have seen some of them in Amsterdam.'

(35) *Ik heb enkelen in Amsterdam gezien.*

I have some in Amsterdam seen

'I have seen some in Amsterdam.'

This would mean that only the grammatical variant of (34) is a case of ellipsis, the quantitative pronoun replacing the noun in object position. Since there is no *er* in (35), this cannot be a case of ellipsis, contrary to what Kester claims.

Because of these problems with Kester's ellipsis analysis of the human construction in Dutch, I will propose another analysis in the next section.

#### 4. A root-analysis of nominalizations

In the previous section it was shown that a nominalization analysis, i.e. a morphological derivation analysis, of deadjectival human nouns in Dutch such as *de zieke* 'the sick person' raises problems, because they can, e.g., be combined with adverbs and complements, and allow comparative and superlative formation. This suggests that they are still adjectives and have not been converted into nouns in the Lexicon. It was, however, also shown that a syntactic ellipsis analysis, in which the empty noun has the form of a small *pro* with the features [+human], [±plural], and [±generic] is also not without problems, which concern the interpretation of the human construction, the features on *pro*, the schwa, and the plural *-n*. Because of these problems, I will propose another analysis in this section, which is a mix of the nominalization and the ellipsis analysis.

In section 2, I showed that in a Distributed Morphology approach

---

<sup>6</sup> The Dutch grammar Geert et al. (1984) does not give a rule or an example confirming or infirming my intuition. Van den Toorn (2008:179), however, gives an example containing *er* and *enkelen*, which infirms my intuition. In Google, in examples such as (27) involving a human interpretation, however, *-n* is used in half of the cases. I attribute the use of *-n* here to the confusion with the use in (33) and (35).

morphological operations like conversion and derivation can be accounted for in Syntax. For mixed categories this syntactic approach to morphology is an ideal way of accounting for their mixed properties. The functional projections dominating category-neutral roots determine to what degree properties are mixed. If a root is only dominated by nominal functional projections, it is purely nominal, if there are, e.g., verbal and nominal projections, we are dealing with a deverbal nominalization. In Alexiadou's (2001) analysis of nominalizations (ending in the suffix *-ion* in English), the number of functional verbal projections determines if the nominalization is a process noun or a result noun (Alexiadou 2001).

Since the human nouns discussed in this paper are deadjectival, we must assume that they are roots dominated by at least an adjectival functional projection: *aP*.<sup>7</sup> We saw that, in Dutch, adjectival properties such as complementation, degree modification, and comparative and superlative formation can also be present. This means that the head of *aP* must have the property to license a complement, and that *aP* can be dominated by the adjectival functional projection *DegP*.

In Kester's analysis, the schwa morpheme on the adjective in the human construction is analyzed as an inflectional morpheme. However, it has been pointed out in the previous section that this analysis raises a problem: attributival participles ending in *-en* cannot get an inflectional schwa, but in a nominalized form they do end in a schwa morpheme. I therefore assume, just as in traditional analyses (Geerts et al. 1984), that in the human construction *-e* is a derivational suffix, and not an inflectional suffix. In Distributed Morphology it is assumed that derivational suffixes (or their features) are inserted in the head of the functional nominal projection *nP*. Since I analyze the schwa in the human construction as a derivational suffix, I propose that it is inserted in the head of *nP*. Its features are [+human] and [+common]: it creates human nouns that combine with the singular definite determiner *de* in Dutch.

Since, in this analysis, the human construction is a (syntactic)

---

<sup>7</sup> In the head of *aP* (features bundles of) adjectival suffixes can be inserted. Suffixed adjectives can also be used in the human construction:

(i) *de aandacht-ig-e-n*  
the attention.suff.-e-n  
'those paying attention'

nominalization and not a case of ellipsis, pluralization is simply the consequence of the presence of *nP* in the functional domain dominating the root.<sup>8</sup> Borer (2005) proposes that the distinction between mass nouns and count nouns is made in Syntax by means of the functional projection Classifier Phrase. Count nouns are roots dominated by ClassP. In section 2, we saw that, in Alexiadou et al. (2011) analysis of non-derived deverbal nominalizations, the possibility to function as a count noun is expressed as the feature [+count] in the head of ClassP, accounting for the possibility of pluralization. Finally, the presence of the functional projection DP accounts for the presence of a determiner in the human construction.<sup>9</sup>

The structure that results from the combination of the proposed functional projections dominating the root in the human construction is the following:

$$(36) [_{DP} [_{FP} [_{NumP} [_{ClassP[+count]} [_{nP} [_{DegP} [_{aP} [root]]]]]]]]]]$$

Based on (36), structure (37b) represents the human construction in (37a), in which for concreteness the morphemes have a phonological shape and are not simply feature bundles:<sup>10</sup>

(37) *a.de volkomen afhankelijk*  
 the completely dependent  
 ‘the completely dependent persons’

$$b.[_{DP} [de] [_{NumP} [_{ClassP[+count]} [n] [_{nP} [e] [_{DegP} volkomen [_{aP} [afhankelijk]]]]]]]]]]$$

The presence of functional projections in the structure is determined by the combinatorial properties of the human construction. A genuine “lexicalized” deadjectival nominalization like *de blanke* ‘the white person’,

<sup>8</sup> Although in Kester’s analysis, the human construction is analyzed as a case of ellipsis, it does not have an antecedent. This is the same in the DM analysis.

<sup>9</sup> The human construction can be modified by an adjective, which is inserted in FP:

(i) *een opgewekte zieke*  
 a good-humoured sick  
 ‘a good-humoured sick person’

<sup>10</sup> If a complement is licensed by the head of *aP*, the order “complement – degree phrase” in (i) is unexpected. Siononius (1992: 112) suggests that this ordering suggests that the complement has been preposed:

(i) *de hiervan volkomen afhankelijken*.  
 the of.this completely dependent  
 ‘the persons who are completely dependent on this.’

which does not allow for modification by a degree adverb, does not have a DegP in its structure and has an 'intransitive head of aP', because it cannot be combined with a complement.

So far we have seen that a syntactic root analysis can account for the adjectival properties of the human construction (the combination with modifiers, comparative and superlative formation, and complementation), just like the ellipsis analysis. The root analysis differs from the ellipsis analysis in that the schwa is analyzed as a derivational morpheme and not as an inflectional morpheme. The possibility of pluralization is related to the presence of *nP* and *ClassP* in the structure, which accounts for the absence of the quantificational pronoun *er* with *enkelen* in (35), *er* being also analyzed as *nP* by Kranendonk (2010). In the ellipsis analysis, on the other hand, *-e* and *-n* serve to formally license resp. identify *pro*, which has the features [+human], [+plural], [ $\pm$ generic]. In the singular, the common definite determiner *de* instead of *-n* identifies *pro* as being human.<sup>11</sup> However, I argued that the omitted noun in the human construction rather seems to be a neuter noun like *het mens* 'the human being' or *het persoon* 'the person', which is an argument against the ellipsis analysis. Furthermore I follow Giannakidou & Stavrou (1999) in assuming that the human construction results from the transformation of a property into a kind (and eventually into an individual). The root analysis proposed in this section has the advantages of the ellipsis analysis, but crucially does not involve ellipsis, but nominalization. Therefore the human construction can have a kind reading, and *-e*, *-n* and *de* do not have a formal licensing resp. identificational function, but simply express nominalization.

Although the root analysis seems to account quite nicely for the mixed properties of the human construction in Dutch, it raises some questions. First, as shown in the previous section, one of the arguments used by Kester in favor of her ellipsis analysis is the fact that the plural morpheme is *-n* and not *-s*. She notices that other deadjectival nouns ending in *-e* take the plural morpheme *-s* and not *-n*:

---

<sup>11</sup> Kester (1996a,b) also analyzes abstract deadjectival nouns as elliptical constructions. In the abstract construction, *pro* is again formally licensed by the schwa, which Kester assumes also here to be inflectional, and is identified by the neuter definite determiner *het*: *het goede* 'the good thing'.



(38) a. *rond*  
'round'

b. *de ronde*  
'the round (in a competition)'

c. *de rondes*  
'the rounds'

Kester tries to account for the exceptional behavior of human nouns like *de zieken* 'the sick people' by analyzing *-n* in this case as a morpheme that identifies *pro* as [+human], [+plural]. However, the human construction is not exceptional in taking *-n* as the plural morpheme. As noticed by Van der Hulst & Kooij (1997), *-(e)n* is the unmarked plural affix in Dutch. The affix *-s* is marked. It occurs in loan words, but also in some native words, such as *wapens* 'weapons' or *lengtes* 'lengths' (next to *lengten*). Although Kester takes the plural morpheme *-n* in the human construction to be exceptional, the choice of the *-* unmarked *-* plural suffix *-n* in this deadjectival syntactic construction is, on the contrary, rather as expected.

Second, as observed in the previous section, the human construction does not take the diminutive form *-tje*, whereas other nouns ending in *-e* do so:

(39) \**een zieketje*  
a sick.DIM  
'a small sick person'

(40) *een rondetje*  
a round.DIM  
'a small round (in a competition)'

The impossibility of diminutive formation might be related to the construction's deadjectival nature. As we saw, Giannakidou & Stavrou (1999) follow Chierchia (1998) in assuming that in the case of nominalization a property is changed into a kind and eventually into an individual.

De Belder (2011) assigns to the Dutch word *chocolade* 'chocolate' a kind reading and a unit reading. According to De Belder, kind nouns in Dutch can be countable:

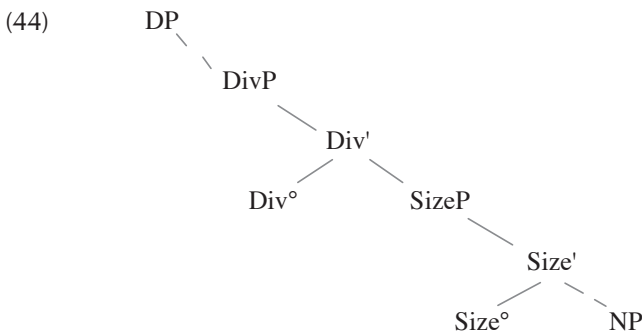
- (41) *Ik proefde chocolades*  
I tasted chocolate-PL  
'I tasted different kinds of chocolate.'  
# 'I tasted pieces of chocolate.'

However, only in the unit reading a diminutive suffix can be added:

- (42) *Ik proefde een chocola-tje.*  
I tasted a chocolate-DIM  
'I tasted a piece of chocolate.'  
# 'I tasted a certain kind of chocolate.'

- (43) *Ik proefde chocola-tje-s.*  
I tasted chocolate-DIM-PL  
'I tasted pieces of chocolate.'  
# 'I tasted different kinds of chocolate.'

For the unit interpretation De Belder proposes structure (44), in which DivP corresponds to ClassP in the structure of the nominalized infinitives discussed in section 2. Both notions are used by Borer (2005). In the kind reading there is no SizeP. With a mass noun DivP would also be absent.



I suggest that the human construction lacks the Size Phrase projection, because the human construction basically expresses a property turned into a kind (*een zieke* 'someone having the property of being sick') rather than a unit ('a sick person'). This would explain why pluralization but not diminutive formation is possible. For English generic plural human nouns such as *the rich*, as in (4), I suggest that both SizeP and DivP are missing, with plurality expressed in NumP.

## 5. Conclusion

In this paper, I have argued against a syntactic ellipsis analysis of the human construction in Dutch. Following Kester, I have recognized that a lexical nominalization analysis cannot account either for the mixed properties of the human construction. I have proposed, instead, an analysis that is a mix of a syntactic ellipsis analysis and a nominalization analysis.

In a Distributed Morphology approach, I have proposed that the mixed character of the human construction in Dutch is expressed in the diversity of the functional projections dominating the category-neutral, non-elliptical, root. The adjectival functional projections account for the combination with adverbs and for the possibility of complementation, whereas the nominal functional projections make the analysis of the schwa as a derivational morpheme inserted in the head of nP and the availability of pluralization, expressed by the unmarked suffix *-n*, possible. The absence of diminutive formation has been related to the basic kind interpretation of the human construction.

## REFERENCES

- Ackema, P. & Neeleman, A. 2004. *Beyond Morphology. Interface Conditions on Word Formation*. Oxford & New York: Oxford University Press.
- Alexiadou, A. 2001. *Functional Structure in Nominals: nominalization, and ergativity*. Amsterdam & Philadelphia: John Benjamins.
- Alexiadou, A., Iordăchioaia, G. & Schäfer, F. 2011. Scaling the variation in Romance and Germanic nominalizations. In: P. Sleeman & H. Perridon (Eds.). *The Noun Phrase*

- in Romance and Germanic: Structure, variation and change*, [Linguistik Aktuell/ Linguistics Today, volume 171]. Amsterdam & Philadelphia: John Benjamins, 25-40.
- Belder, M. De. 2011. A morphosyntactic decomposition of countability in Germanic. *The Journal of Comparative Germanic Linguistics* 14(3): 173-202
- Borer, H. 2005. *In Name Only*. [Vol. 1 of *Structuring Sense*]. Oxford: Oxford University Press.
- Borer, H. & Roy, I. 2010. The name of the adjective. In: P. Cabredo-Hofherr; O. Matushansky (Eds.). *Adjectives: Formal analyses in syntax and semantics*. [Linguistik Aktuell/Linguistics Today 153]. Amsterdam & Philadelphia: John Benjamins.
- Borsley, R.D. & Kornfilt, J. 2000. Mixed extended projections. In: R.D. Borsley (Ed.). *The Nature and Function of Syntactic Categories*. New York: Academic Press, 101-131.
- Chierchia, G. 1998. Reference to kinds across languages. *Natural Language Semantics* 6(4): 339-405.
- Chomsky, N. 1970. Remarks on nominalization. In: R. Jacobs; Rosenbaum, P. (Eds.). *Readings in English Transformational Grammar*. Waltham, MA: Blaisdell, 184-221.
- Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- Corver, N. & Koppen, M. van. 2009. Let's focus on NP-ellipsis. *GAGL (Groninger Arbeiten zur germanistischen Linguistik)* 48: 3-26.
- Depiante, M. & Masullo, P. 2004. Gender is in the lexicon, number is in the syntax: evidence from nominal ellipsis in Spanish. Paper presented at GLOW 27, Thessaloniki.
- Don, J. 2005. On conversion, relisting and zero-derivation. A comment on Rochelle Lieber: English word-formation processes. *SKASE Journal of Theoretical Linguistics* 2(2): 2-16.
- Eguren, L. 2009. Adjectives and deleted nominals in Spanish. In: E. Aboh, E. van der Linden, J. Quer & P. Sleeman (Eds.). *Romance Languages and Linguistic Theory. Selected Papers from Going Romance 2007*. Amsterdam & Philadelphia: John Benjamins, 67-86.
- Embick, D. 2004. On the structure of resultative participles in English. *Linguistic Inquiry* 35(3): 355-392.
- Embick, D. & Noyer, R. 2006. Distributed Morphology and the Syntax/Morphology interface. In: G. Ramchand & C. Reiss (Eds.). *Oxford Handbook of Linguistic Interfaces*. Oxford: Oxford University Press, 289-324.
- Geerts, G., Haeseryn, W., De Rooij, J. & van den Toorn, M. 1984. *Algemene Nederlandse Spraakkunst*. Groningen: Wolters-Noordhoff.
- Giannakidou, A. & Stavrou, M. 1999. Nominalization and ellipsis in the Greek DP.

- Linguistic Review* 16: 295-331.
- Halle, M. & Marantz, A. 1993. Distributed Morphology and the pieces of inflection. In: K. Hale; S. J. Keyser (Eds.). *The View from Building 20*. Cambridge, MA: MIT Press, 111-176.
- Halle, M. & Marantz, A. 1994. Some key features of Distributed Morphology. *MIT Working Papers in Linguistics* 21: 275-288.
- Harley, H. & Noyer, R. 1999. Distributed Morphology. *Glott International* 4(4): 3-9.
- Hulst, H. van der & Kooij, J.G. 1997. Prosodische keuzes bij meervoudsvorming in het Nederlands. In: A. van Santen & M. van der Wal (Eds.), *Taal in Tijd en Ruimte. Voor Cor van Bree bij zijn afscheid als hoogleraar Historische Taalkunde en Taalvariatie aan de Vakgroep Nederlands van de Rijksuniversiteit Leiden*. Leiden: SNL, 365-373.
- Kester, E.-P. 1996a. *The Nature of Adjectival Inflection*. PhD dissertation, Utrecht University.
- Kester, E.-P. 1996b. Adjectival inflection and the licensing of empty categories in DP. *Journal of Linguistics* 32(1): 57-78.
- Kornfeld, L. & Saab, A. 2002. Nominal ellipsis and morphological structure in Spanish. In: R. Bok-Bennema, B. Hollebrandse, B. Kampers-Manhe & P. Sleeman (Eds.), *Romance Languages and Linguistic Theory 2002. Selected Papers from Going Romance*. Amsterdam & Philadelphia: John Benjamins, 183-198.
- Kranendonk, H. 2010. *Quantificational Constructions in the Nominal Domain: Facets of Dutch Microvariation*. PhD dissertation, Utrecht University.
- Lobeck, A. 1995. *Ellipsis, Functional Heads, Licensing and Identification*. Oxford: Oxford University Press.
- Panagiotidis, P. 2001. Empty nouns. *Natural Language and Linguistic Theory* 21: 381-432.
- Plann, S. 1981. The Two *el* + infinitive Constructions in Spanish. *Linguistic Analysis* 7: 203-240.
- Rizzi, L. 1986. Null objects of Italian and the theory of *pro*. *Linguistic Inquiry* 17: 501-557.
- Sleeman, P. 1996. *Licensing Empty Nouns in French*. The Hague: Holland Academic Graphics.
- Sleeman, P. 2001. Deverbale procedés in het Nederlands en het Frans. In: L. Beheydt & P. Godin (Eds.). *Contrastief Onderzoek Nederlands-Frans*. Bibliothèque des Cahiers de l'Institut de Linguistique de Louvain 108, 195-209.
- Sleeman, P. 2010. The nominalized infinitive in French: structure and change. *Linguística*.

*Revista de Estudos Linguísticos da Universidade do Porto* 5: 145-173.

Svenonius, P. 1992. The extended projection of N: identifying the head of the noun phrase. *Working Papers in Scandinavian Syntax* 49: 95-121.

Van den Toorn, M.C., 2008, 9<sup>th</sup> edition. *Nederlandse Grammatica*, [www.dbnl.org/tekst/toor004nede01\\_01/toor004nede01\\_01.pdf](http://www.dbnl.org/tekst/toor004nede01_01/toor004nede01_01.pdf) (first edition: 1984. Groningen: Wolters-Noordhoff).

# Les verbes à objets cognats en portugais européen: quelques réflexions sur la conversion et la génération de mots\*

Celda Morgado Choupina

celda@ese.ipp.pt

*Instituto Politécnico do Porto (Portugal)*

*Centro de Linguística da Universidade do Porto*

**ABSTRACT.** The text analyzes the verbs with cognate objects in European Portuguese and provides an analysis of this phenomenon in the framework of Distributed Morphology, and a few ideas of the Minimalist Program. In this paper we only develop the analysis of the constructions of the type of *chorar um choro...* [cry a cry...] and *cantar uma cantiga* [sing a song]. Since a lexicalist theory as the one developed by Hale & Keyser (1993; 2002) can't solve all the issues raised by the various types of constructions in EP, we will try an analysis that combines the contributions of Haugen (2009), on cognate objects, Marantz (1993) and Alexiadou (2001, 2006), on verbal functional heads, and (Chomsky 1995), on the idea of movement by copy.

**KEY-WORDS.** cognate objects, Distributed Morphology, Minimalist Program, movement by copy, functional head.

**RÉSUMÉ.** Le texte analyse les verbes à objets *cognats* en Portugais Européen et propose une analyse de ce phénomène dans le cadre de la Morphologie Distribuée et de quelques idées du Programme Minimaliste. Dans cet article nous développons uniquement l'analyse syntaxique des constructions du type *chorar um choro...* [pleurer un pleur...] et *cantar uma cantiga* [chanter une chanson]. Puisqu'une théorie lexicaliste comme celle développée par Hale & Keyser (1993; 2002) ne parvient pas à résoudre toutes les questions soulevées par les différents types de constructions en PE, nous essayerons une analyse qui concilie les contributions de Haugen (2009), à propos des objets cognats, Marantz (1993) et Alexiadou (2001, 2006), sur l'existence de têtes fonctionnelles verbales, et (Chomsky 1995), sur l'idée de mouvement par copie.

**MOTS-CLÉS.** objets cognats, Morphologie Distribuée, Programme Minimaliste, mouvement par copie, tête verbale fonctionnelle.

---

\* En tant que membre du CLUP cette recherche a été appuyée par FEDER / POCTI U0022/2003.

Nous remercions Ana Maria Brito par ses observations sur une version préalable de ce texte. Nous remercions aussi les évaluateurs anonymes de cet article et les participants du *Workshop on conversion and mixed categories*, les 2 et 3 novembre 2012, à la Faculté des Lettres de l'Université de Porto, pour leurs commentaires. Nous remercions également notre collègue José António Costa pour la première traduction de ce texte en français.

## 1. Introduction

Il y a, dans des langues telles que l'anglais, un large éventail d'éléments lexicaux qui appartiennent à la fois à la catégorie des noms et des verbes, comme c'est le cas de *dance*, *laugh*, *bottle* et *saddle*, ce qui a conduit Hale & Keyser (1993) à proposer l'analyse des verbes dénominaux par incorporation / mouvement de la base de N vers V. Suivant ce point de vue, la conversion est conçue comme un processus lexical, impliquant un mouvement à l'intérieur de la structure lexicale elle-même. Ces mêmes verbes peuvent apparaître avec des objets cognats de types différents, les uns des cognats (OC), *to sleep a sleep...*, les autres des objets hyponymes (OH), *to dance a jig* (Hale & Keyser 2002).

Dans une perspective totalement différente, non lexicaliste et basée sur la Morphologie Distribuée (MD), il est possible d'analyser ces phénomènes à partir de la notion de racine acatégorielle, qui acquiert sa nature nominale ou verbale en fonction des morphèmes fonctionnels avec lesquels elle fusionne (Haugen 2009; Alexiadou 2001 et 2006). Ces idées, liées surtout au processus de mouvement par copie du Programme Minimaliste (PM), (Chomsky 1995), suggèrent des alternatives intéressantes pour l'analyse des verbes à OC.

Dans cet article, et développant Choupina (2013), nous nous proposons de discuter, du point de vue théorique et de l'apport descriptif, la contribution de ces auteurs et de présenter une analyse du phénomène en PE.

Ainsi, nous présenterons, dans la section 2, une synthèse des analyses de Hale & Keyser (1993; 2002) et de Haugen (2009) pour les OC en anglais. Dans la section 3.1., nous présenterons l'analyse des restrictions sémantiques et syntaxiques de ce que nous allons appeler des vrais OC et des OC apparentés. Au §3.2., nous développerons une analyse syntaxique de ces deux types d'OC en essayant de concilier les idées fondamentales de la MD avec l'idée de mouvement par copie et l'importance des têtes verbales fonctionnelles. Finalement, au §4 nous esquisserons nos conclusions.

## 2. Deux analyses de référence des constructions à objets cognats

### 2.1 Analyse lexicaliste (Hale & Keyser 1993; 2002)

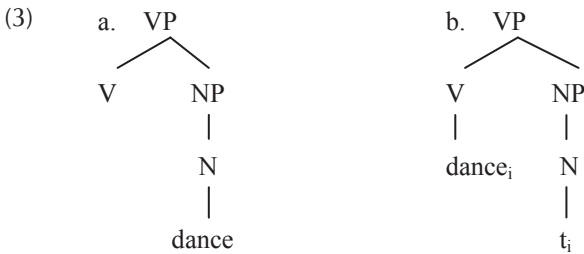
En anglais il y a des mots qui appartiennent à la fois à la catégorie des noms et à la catégorie des verbes, tels que *dance* [*la danse/danser*] et *laugh*



[*le rire/rire*]. Les verbes *to dance* et *to laugh*, traditionnellement considérés intransitifs, peuvent intégrer des constructions transitives, ce qui déclenche la cooccurrence du verbe et du nom morphologiquement relié, comme on le voit en (1) et (2).

- (1) He danced a dance.
- (2) He laughed his last laugh. (Hale & Keyser 2002: 71)

Pour analyser ces constructions avec des OC, Hale et Keyser (1993) proposent que les verbes inergatifs dénominaux en anglais soient considérés comme des « transitifs » ou des « transitifs cachés » à un niveau sous-jacent, formés par un processus lexical (abstrait) («l-syntactic») d'incorporation, c'est-à-dire, un déplacement de tête, comme illustré dans la dérivation de a. vers b. en (3)<sup>1</sup>.



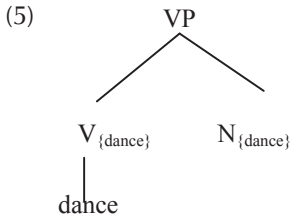
De ce point de vue, la conversion N > V est conçue comme un processus lexical, qui implique un mouvement à l'intérieur de la structure lexicale elle-même.

L'existence d'une possible structure transitive explicite mène les auteurs à réadapter un peu leur proposition initiale (Hale & Keyser 2002). Dans ces constructions, l'argument sujet est introduit au niveau de la syntaxe, étant donc externe, et, en principe, le verbe ne permettrait pas la réalisation du nom *dance* en position de complément, ou plutôt, il empêcherait l'épellation (*spell out*) de la structure transitive (4a), tout en permettant un hyponyme (4b).

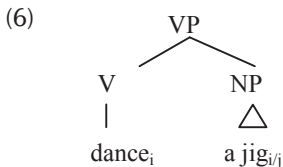
<sup>1</sup> Dans la structure syntaxique nous adopterons les sigles en anglais.

- (4) a. He danced a dance.  
b. He danced a jig.

Considérée d’abord comme le résultat d’une opération de déplacement d’une tête nominale vers la position de la tête verbale, comme décrit en (3), la notion d’incorporation est améliorée par la suite, de façon à capturer l’existence d’objets cognats et d’hyponymes et elle désigne donc un processus de copie; les auteurs introduisent la notion de «conflation» en tant que « the process of copying the p-signature of the complement into the p-signature of the head, where the latter is “defective”» (Hale & Keyser 2002: 63 et 93). Voir (5).



Certains auteurs (entre autres, Ramchand 2008; Haugen 2009 et Gallego 2012), ont remarqué que certains verbes ne pourraient pas résulter d’une incorporation nominale dans une racine verbale phonologiquement déficiente, puisqu’en (4b) cette position est occupée par le vrai complément – l’objet hyponyme (OH) –, celui-ci ne pouvant pas être justifié par une identité morphologique (étant donné qu’il ne présente pas la même racine que le verbe préalablement formé). La seule façon de « sauver » la proposition présentée par Hale et Keyser est d’admettre que l’OH ne peut être introduit que dans la syntaxe à travers une *Insertion Tardive*, opération qui n’est autorisée que dans un cadre non lexicaliste (voir (6)).



Il faut observer aussi que les verbes ayant des OC du type *to sleep* a *sleep* n'admettent jamais (ou très difficilement) la construction avec OH, comme *dance a jig*. On a, donc, au moins deux constructions avec OC: celles qui admettent un OC ou un OH et celles n'admettant qu'un OC (que nous allons considérer comme de vrais cognats). En effet, OC et OH sont distincts en anglais, parce que les verbes avec de vrais OC ne permettent pas de constructions avec des OH. Comparons (7), qui est grammatical, avec (8), qui est agrammatical.

(7) *to sleep* + OC

She slept *the sleep of the just*.

(8) \**to sleep* + OH

\* She slept *her last nap/a long winter slumber*.

Ces verbes refusent d'ailleurs la pronominalisation, comme on peut le constater par la mauvaise formation de constructions avec OC (9), tandis que les constructions avec OH l'acceptent (10) (Hale & Keyser 2002: 71):

(9) \**to sleep* + OC / pronominalisation

\* John slept *the sleep of the just* and Bill slept *it too*.

(10) *to dance* + OH / pronominalisation

John danced *the tango* and Bill danced *it too*.

En ce sens, dans des langues comme l'anglais, les OC du type *to sleep* a *sleep* et les OC du type *to dance* a *dance* ne pourraient pas être analysés par la même opération, incorporation / *conflation*, comme le soutiennent Hale & Keyser (2002). Pour le portugais, ce qui rend difficile d'accepter la proposition de Hale et Keyser c'est aussi l'impossibilité d'expliquer, à travers la même opération lexicale, la formation des verbes admettant les deux constructions, c'est-à-dire, des verbes qui n'admettent que des OC, tels que *dormir um sono* [dormir un sommeil], et des verbes qui admettent des OC et des OH, tels que *dançar uma dança/ um tango*, [danser une dance/un tango]; *cantar uma cantiga/ um fado* [chanter une chanson/un fado], comme on le verra dans le § 3.

Ce constat nous oblige à repenser ce cadre et à rechercher des modèles

théoriques capables d'expliquer avec plus de rigueur les différentes constructions existant dans les langues romanes et en anglais. Ce besoin s'en fait sentir aussi en raison des réflexions suivantes:

- d'une part, par l'association de la *conflation* à la sélection d'arguments, les auteurs affaiblissent la composante du lexique et par conséquent leur approche, car ils utilisent des concepts et des opérations nettement syntaxiques;
- d'autre part, dans les langues romanes, il n'existe pas, pour la plupart des cas, d'homonymie parfaite entre la racine verbale et le OC, contrairement à ce qui se passe en anglais, langue qui présente tant de verbes dénominaux.

## 2.2 Proposition selon la Morphologie Distribuée (Haugen 2008; 2009)

Dans le cas particulier des OC, Haugen (2009) adopte la théorie de la MD (développée, entre autres, par Halle & Marantz 1993; Harley & Noyer 1999), étant donné que, selon lui, c'est la structure qui détermine la formation des mots et la plupart du sens des constructions. Il s'agit donc d'un modèle opposé au modèle lexicaliste de Hale et Keyser, comme on l'a vu auparavant. Selon Haugen (2009), la révision de la notion de *conflation* de Hale & Keyser (2002) n'est pas complètement satisfaisante, surtout parce qu'une complémentation stricte limitée par *conflation* n'exclut pas la possibilité d'un «résidu» (par exemple, un adjectif) qui modifie le nom.

Les propriétés les plus importantes de la MD pour le projet de Haugen (2009) sont, d'une part, l'utilisation de l'insertion tardive et, d'autre part, la non primauté des catégories désignées comme noms et verbes (les racines sont sous-spécifiées, donc dépourvues de traits de catégorie). Ainsi, la MD place la structure syntaxique hiérarchique tout au long du parcours, y compris des projections fonctionnelles (*n* ou *v*) et des racines à interpréter.

Haugen suit aussi la Théorie de la Copie de Chomsky (1995), où les traits sont considérés comme des copies totales d'un constituant déplacé, étant co-indexés par une chaîne de mouvements. Ces copies sont des faisceaux de traits plutôt que du matériel lexical morphologiquement spécifié. Cependant, il assume que ces traits ne sont pas supprimés après leurs copies.

À partir de ces idées, la thèse centrale de Haugen (2009: 248) est que nous pouvons résoudre le problème des constructions avec des OH

en admettant l'insertion (par Insertion Tardive) de racines nominales non cognates dans des copies hautes et basses, après l'application de l'opération de déplacement par copie. Le nom le plus bas sera interprété comme un hyponyme, de façon à ce que n'importe quel nom radical soit inséré dans la tête de la chaîne de déplacements / copies.

Selon la Théorie de la Copie traditionnelle, la syntaxe rend visible une copie du complément dans la position incorporée (c'est-à-dire, attaché à *v*), comme si elle y était générée, tout en laissant des traces dans la position de base du complément. Haugen estime que "The complex verb then moves higher up the tree for tense, aspect and mood inflection, as per standard incorporation" (Haugen 2009: 249).

L'auteur accepte, par conséquent, la possibilité de *spell out* des deux copies à la fois: la copie basse et la copie haute attachée à *v*. Remarquons que Haugen voit l'opération d'incorporation non seulement en tant que mouvement du nominal vers la position verbale (suivant Hale & Keyser 1993), mais aussi selon la Théorie de la Copie (Chomsky 1995), l'opération consistant à copier ('copying') au lieu de déplacer au sens strict.

En permettant l'insertion de matériau lexical non cognat dans la position du radical de *v* en position de copie basse, Haugen résout d'une façon élégante la problématique des OH. De plus, selon l'auteur, on pourrait accepter que les conditions pragmatiques spécifiques des langues (les maximes de quantité de Grice, par exemple) légitiment l'insertion de certaines racines et leur *spell-out* dans certaines positions de la chaîne de mouvements (Haugen 2009: 253 et suivantes). L'auteur ne présente pourtant aucune structure syntaxique en anglais<sup>2</sup>, ni avec des OC, ni avec des OH. Il ne se prononce pas non plus sur les nominalisations déverbales (*Ele dançou magnificamente o tango* vs. *A dança do tango foi magnífica*. [Il a dansé le tango magnifiquement vs. La danse du tango a été magnifique.]) et il ne distingue pas les constructions qui n'admettant que des OC de celles présentant des OC et des OH. Comme on le verra, le Portugais Européen (PE) et les langues romanes en général soulèvent de nouvelles questions dans le domaine des verbes avec des OC et avec des OH, ce qui justifie l'exploration d'une analyse de quelques-unes des constructions discutées dans ce texte.

---

<sup>2</sup> Les structures que Haugen (2009) présente proviennent de l'Hopi (Uto-Aztecan), langue agglutinante présentant une syntaxe différente de celle de l'anglais.

## 3. Les vrais objets cognats et les constructions apparentées en PE

Les exemples (11) à (13) présentent ce que, dans la littérature, on nomme des objets cognats (Hale & Keyser 1993) ou des arguments ombre (Pustejovsky 1998).

(11) a. Amália *cantou* uma cantiga portuguesa.

Amalia a chanté une chanson portugaise

‘Amalia a chanté une chanson portugaise’

b. Amália *cantou* um fado.

Amalia a chanté un fado

‘Amalia a chanté un fado’

(12) a. Ele *chorou* um choro silencioso.

il a pleuré un pleur silencieux

‘il a pleuré des pleurs silencieux’

b. *Dormimos* um sono reparador.

[-] avons dormi un sommeil réparateur

‘nous avons dormi d’un sommeil réparateur’

(13) a. Ele *morreu* de uma morte lenta.

il est mort de une mort lente

‘il est mort d’une mort lente’

b. *Caminhavam* por um caminho difícil

[-] ont marché par un chemin dur

‘ils ont marché par un chemin dur’

Ces exemples illustrent des constructions différentes d’objets cognats et montrent que ces constructions ne sont pas une classe homogène - le nom et le verbe n’ont pas les mêmes relations morphologiques dans tous les exemples, et les constructions ne sont donc pas soumises aux mêmes restrictions syntaxiques et sémantiques. En (14) nous résumons une possible typologie de ces constructions pour le PE (cf. Choupina 2013):

(14) Typologie des constructions à OC en PE:

- (i) **Les vrais OC: sous-type a.** le nom est morphologiquement relié au V; ce sont des structures qui présentent forcément un indéfini et un modificateur restrictif, *chorar-choro...*; **sous-type b.** le nom, bien que n'étant pas morphologiquement lié au verbe, maintient une stricte dépendance sémantique avec lui et il ne peut être remplacé par un OC morphologiquement relié au V; ce sont aussi des structures présentant toujours un indéfini et un modificateur restrictif, *dormir-sono...* [dormir-sommeil]. Dans les deux types a. et b. le nom représente un évènement.
- (ii) **Les OC apparentés:** le nom est morphologiquement relié au verbe et, donc, ces compléments semblent des cognats; mais, parce que ce sont des constructions qui n'ont pas les mêmes contraintes syntaxiques (la présence des indéfinis et du modificateur restrictif n'est pas obligatoire), ils ne peuvent pas être considérés comme de vrais OC: *cantar-cantiga* [chanter – chanson].
- (iii) **Les OH:** le nom exprime un sous-type de l'entité désignée par le nom morphologiquement relié au V ou sémantiquement dépendant de lui; l'indéfinitude et la modification restrictive ne sont pas obligatoires: *cantar-fado* [chanter - fado].
- (iv) **Les OCP:** ce sont des SPREP contenant un nom morphologiquement relié au V: *morrer de uma morte...* [mourir d'une mort ...]; *caminhar por um caminho...* [marcher par un chemin...]<sup>3</sup>.

Pour des questions d'espace, dans cet article nous analysons les types (i), (ii) et (iii).

### 3.1 Restrictions syntaxiques et sémantiques sur les vrais OC et des OC apparentés

---

<sup>3</sup> Par leur comportement linguistique, les OCP peuvent être reportés à la classe des vrais OC (Choupina 2013), mais nous ne développons pas ici cette idée.

- (1) a. he died [of] a slow death [Anglais]  
b. Jorge morreu uma morte trágica [Portugais du Brésil]  
c. ele morreu de uma morte lenta [Portugais Européen] [Il est mort d'une mort lente]
- (2) a. they walked [on] a difficult way [Anglais]  
b. caminharam um caminho difícil [Portugais du Brésil]  
c. caminhavam por um caminho difícil [Portugais Européen] [Ils marchaient par un chemin dur].

Les exemples (11) et (12), ci-dessus, incluent, respectivement, des OC apparentés et de vrais OC. Voyons d'abord les exemples (11a) et (12a), repris en (15) et (16):

(15) Amália *cantou* uma *cantiga portuguesa*. (OC apparenté)  
'Amalia a chanté une chanson portugaise'

(16) Ele *chorou* um *choro silencioso*. (vrai OC)  
'il a pleuré un pleur silencieux'

Les constituants *uma cantiga portuguesa* [une chanson portugaise] et *um choro silencioso* [un pleur silencieux] sont des objets directs des verbes. Ils contiennent un nom cognat, morphologiquement relié au verbe (*cantar-cantiga* [chanter-chanson]; *chorar-choro* [pleurer-pleur]). Le SN des exemples (15) et (16) comprend un déterminant indéfini et un adjectif en position post-nominale (cf. Choupina 2013; Silva 2010)

Cependant, l'indéfinitude et la modification n'ont pas le même statut dans les deux exemples. En (15), la définitude et la modification sont optionnelles, surtout en fonction du discours préalable<sup>4</sup>, comme les exemples (17) à (19) le montrent, tandis que, en (16), les deux propriétés sont nécessaires. Remarquons l'agrammaticalité de (20) et (21), ce qui montre que la construction *chorar-choro* [pleurer-pleur] est beaucoup plus restrictive que *cantar-cantiga* [chanter-chanson].

(17) Amália *cantou* a *cantiga portuguesa*.  
Amalia a chanté la chanson portugaise  
'Amalia a chanté la chanson portugaise'

(18) Amália *cantou* uma *cantiga*.  
Amalia a chanté une chanson  
'Amalia a chanté une chanson'

<sup>4</sup> Les exemples deviennent plus acceptables dans des constructions du type (i) et (ii).

- (i) Amália *cantou* a *cantiga portuguesa por excelência, o fado*.  
Amalia a chanté la chanson portugaise par excellence, le fado.  
'Amalia a chanté la chanson portugaise par excellence, le fado.'
- (ii) Amália *cantou* a *cantiga e o Marceneiro cantou o fado*.  
Amalia a chanté la chanson et le Marceneiro a chanté le fado  
'Amalia a chanté la chanson et Marceneiro a chanté le fado.'



- (19) Amália cantou a cantiga.  
Amalia a chanté la chanson  
'Amalia a chanté la chanson'
- (20) \*Ele chorou o choro silencioso.  
il a pleuré le pleur silencieux  
'il a pleuré le pleur silencieux'
- (21) \*Ele chorou um choro.  
il a pleuré un pleur  
'il a pleuré un pleur'

Pour cette raison, nous faisons une distinction entre les vrais OC (voir la paire *chorar-choro*) et les OC apparentés (pour *cantar-cantiga*) (cf. Choupina 2013). De même Silva (2010: 17, 92 ss), pour le portugais du Brésil, présente les propriétés suivantes comme des caractéristiques des OC prototypiques: «verbe principal intransitif transitivisé par l'instanciation postverbale d'un Syntagme Nominal interprété, dans le prédicat, comme un objet direct, qui contient un déterminant (faible, non référentiel), une tête (nom apparenté du verbe intransitif) et un modificateur (SA / SP / SC)» (Silva 2010: 17)<sup>5</sup>.

Parfois, le critère d'indéfinitude n'est pas considérée comme ayant un comportement homogène dans les différentes langues où l'on retrouve ces constructions, et certains linguistes (cf. Real-Puigdollers 2008), en présentant des exemples littéraires ou des expressions idiomatiques (comme *reir la risa de un niño* [rire d'un rire d'enfant] et *pleurer toutes les larmes de son corps*), ont même mis en cause ce critère dans plusieurs langues. Nous nous éloignons de cette critique et nous acceptons le critère d'indéfinitude comme une exigence fondamentale pour l'existence de vrais OC, associé au critère de modification.

La possibilité de remplacer un OC par un OH renforce aussi la distinction entre deux grands types de construction. La construction (15) a un OC apparenté qui peut être remplacé par un OH, tandis que la construction avec un vrai OC, en (16), produit une agrammaticalité, ne permettant pas d'alternance, comme (22) et (23) l'illustrent.

---

<sup>5</sup> Notre traduction.

- (22) Amália cantou um fado lisboeta/ o fado lisboeta/ um/o fado.  
Amalia a chanté un fado de Lisbonne / le fado de Lisbonne / un /le fado  
'Amalia a chanté un fado lisboète/ le fado de Lisbonne/ un /le fado'
- (23) Ele chorou um choro sufocante / \*o choro sufocante /\*o choro.  
il a pleuré un pleur étouffant/ le pleur étouffant/ le pleur  
'il a pleuré un pleur étouffant / le pleur étouffant /le pleur'

En (22), le verbe *cantar* [chanter] peut être suivi d'un OH, soit avec un indéfini *um fado lisboeta* [un fado de Lisbonne], soit avec le défini (*o fado* [le fado]) et l'absence de modificateur restrictif; par contre, *chorar* [pleurer] (22) ne fournit aucun hyponyme et exige le nom, l'article indéfini et un modificateur restrictif.

Gallego (2012: 103) suit cette même ligne de distinction, nommant les constructions avec les verbes *cantar* [chanter], *comer* [manger] et *beber* [boire] (en espagnol standard) des hyponymes alternants. De même, Silva (2010: 17), pour le PB, qualifie les constructions avec des verbes d'alternance de verbes à objet cognat non prototypique, par opposition aux verbes à cognat prototypique, qui ne se produisent qu'avec des verbes traditionnellement considérés intransitifs.

En développant cette ligne de raisonnement, nous considérons comme des verbes propices à la construction avec des OC apparentés les verbes d'ingestion (*comer* [manger], *beber* [boire], *fumar* [fumer], etc.), les verbes de mouvement (*dançar* [danser], *caminhar* [marcher], *correr* [courir], etc.) et quelques autres portant sur des processus intellectuels (*estudar* [étudier], *ler* [lire], *pensar* [penser]) (cf. aussi Bosque & Gutiérrez-Rexach 2009: 361). Ce sont des verbes qui peuvent toujours être transitifs.

Un autre critère de distinction de ces constructions est la possibilité / l'impossibilité d'ajouter une expression du type «*mas não sei o quê*» [mais je ne sais pas quoi] (Bosque & Gutiérrez-Rexach 2009: 360). Voyons (24) et (25).

- (24) a. Amália cantou, mas não sei o quê.  
Amalia a chanté, mais [-] ne sais le quoi  
'Amalia a chanté, mais je ne sais pas quoi'
- b. Amália comeu/ bebeu/dançou, mas não sei o quê.  
Amalia a mangé/a bu/ a dansé, mais [-] ne sais le quoi  
'Amalia a mangé/a bu/a dansé, mais je ne sais pas quoi'
- (25) a. \*Ele chorou, mas não sei o quê.  
Il a pleuré, mais [-] ne sais le quoi  
'il a pleuré, mais je ne sais pas quoi'
- b. \*Ele dormiu/espирrou, mas não sei o quê.  
il a dormi/a éternué, mais [-] ne sais le quoi  
'il a dormi/ a éternué, mais je ne sais pas quoi'

Examinons, maintenant, le critère de pronominalisation, souvent utilisé dans la littérature pour faire la distinction entre types d'objets cognats. Prenons l'exemple (26) et reprenons les exemples (11a) et (12a), ici numérotés (27) et (28):

- (26) Ele chorou lágrimas de dor.  
il a pleuré larmes de douleur  
'il a pleuré des larmes de douleur'
- (27) Amália cantou um fado.  
Amalia a chanté un fado  
'Amalia a chanté un fado'
- (28) Ele chorou um choro silencioso.  
il a pleuré un pleur silencieux  
'il a pleuré un pleur silencieux'

Voyons maintenant les constructions (29) à (32) avec pronominalisation.

- (29) ele chorou- **as** (-as = lágrimas)  
il a pleuré les (les = larmes)  
'il les (= larmes) a pleurées'
- (30) \*Ele chorou- **o**. (-o = choro silencioso)  
il a pleuré le (le = pleur silencieux)  
'il l' (= pleur silencieux) a pleuré'
- (31) Amália cantou- **o**. (-o = um fado)  
Amalia a chanté le (le = un fado)  
'Amalia l' (= un fado) a chanté'
- (32) Amália cantou- **a**. (-a = a cantiga portuguesa)  
Amalia a chanté la (la = la chanson portugais)  
'Amalia l' (= une chanson portugaise) a chantée'

Les OD (objet direct) (29), les OH (31) et les OC apparentés (32) permettent la construction avec pronominalisation, ce qui contribue à considérer leur statut argumental: ce sont de vrais arguments des verbes qui les accueillent et ils présentent le cas accusatif et le rôle d'objet/thème. En revanche, l'agrammaticalité de la construction avec de vrais OC, en (30), peut suggérer le statut non-argumental de ces objets (cf. Horita 1996; Kitahara 2006). En outre, *um choro* est un évènement.

Toutefois, au sein de la MD, la distinction entre arguments et non-arguments, autrement dit, la discussion du statut argumental des OC, est une dimension syntaxique considérée comme dérivée structurellement, car les racines sont, dans la base, non catégorielles et dénuées de grille argumentale et de grille thématique. Les racines, les morphèmes lexicaux, sont dotés seulement de traits minimaux de sens ou d'informations aspectuelles (les primitifs sémantiques) et non de grille argumentale ou thématique.

En ce qui concerne le verbe *dormir*, qui en PE permet la cooccurrence avec le nom *sono* [sommeil] mais pas le nom cognat morphologique du verbe, \**dormir uma dormida*, (33), il peut aussi être considéré comme un vrai OC (type b.); et cela pour trois raisons: tout d'abord, (i) la langue ne fournit pas le cognat morphologique du verbe (33); (ii) le nom autorisé n'est

que sémantiquement cognat (34); (iii) il y a des langues qui permettent le cognat morphologique, comme l'anglais (35).

(33) \**Dormimos* uma dormida reparadora.  
 nous avons dormi une dormition<sup>6</sup> réparatrice  
 'nous avons dormi d'un sommeil réparateur'

(34) *Dormimos* um sono reparador.  
 nous avons dormi un sommeil réparateur  
 'nous avons dormi d'un sommeil réparateur'

(35) *We slept* a refreshing sleep.  
 nous avons dormi un réparateur sommeil  
 'nous avons dormi un sommeil réparateur'

Pour ces raisons, nous envisageons aussi ces objets comme de vrais OC, du type b. (mots cognats sémantiques). Leur comportement relativement à l'indéfinitude et à la modification renforce cette proposition. Ces tests avec le verbe *dormir* soulèvent bien évidemment plusieurs questions que les limites de ce texte ne permettent pas d'analyser en détail.

Comme synthèse de ce que nous avons vu, nous présentons ci-dessous (Tableau 1) un résumé des similitudes et des différences entre les sous-classes d'OC analysés, dans cet article, en PE.

Tableau 1 – Synthèse des critères/propriétés contrastifs des vrais OC et des OC apparentés

Critères/propriétés	vrai OC		OC apparenté
	sous-type a.	sous-type b.	
Catégorie SN	+	+	+
Adjacence à V	+	+	+
Relation morphologique V-N	+	-	+
Modification	+	+	+/-
Indéfinitude	+	+	+/-
Substitution par OH	-	-	+
Expansion par "mais je ne sais pas quoi"	-	-	+

<sup>6</sup> En français, les dérivés de *dormir* se construisent autrement, d'où l'impossibilité d'une traduction exacte de l'expression portugaise *dormir um sono* vs. *dormir uma dormida*. Ce serait à peu près *dormir un sommeil* vs. *dormir une dormition*. Le non *dormition* signifie la mort de la Vierge, dans la religion chrétienne.

### 3.2 Analyse syntaxique: une proposition

Étant donné la présentation des propriétés fondamentales des constructions à objet cognat, dans ce paragraphe nous allons esquisser une analyse syntaxique des OC.

À la suite des propositions de la MD, notamment l'analyse de Haugen (2009) pour les OC et aussi la notion de têtes verbales fonctionnelles de Marantz (1993) et Alexiadou (2001, 2006), mais sans négliger le PM, nous allons maintenant examiner la structure des constructions avec de vrais OC et des OC apparentés.

La grammaire traditionnelle explique la formation de la paire *chorar-choro* [pleurer-pleur] à partir de la dérivation régressive du verbe *chorar*, par soustraction du marqueur d'infinitif du verbe et échange de la voyelle thématique (VT) / indice thématique (IT). Cela veut dire que, dans cette perspective, la direction dérivationnelle a lieu du verbe *chorar* [pleurer] vers le nom *choro* [pleur].

Rodrigues (2001) soutient l'hypothèse lexicale/ morphologique de la conversion, en cherchant à préciser la façon dont le reclassement du radical verbal en nominal se produit sans avoir recours à des suffixes dérivatifs.

À partir de Hale & Keyser (1993), comme nous l'avons vu au paragraphe 2., les verbes inergatifs sont qualifiés de dénominaux, car le verbe est formé par incorporation d'une tête nominale dans la position de la tête verbale, dans la composante lexicale, ce qui expliquerait l'homonymie parfaite entre, par exemple, *to cry - a cry*; *to laugh - a laugh*; *to sneeze - a sneeze*.

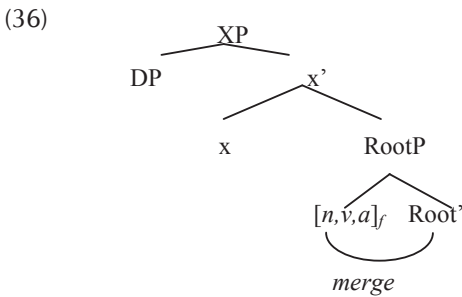
Ce type d'analyse, selon laquelle il y a une composante lexicale génératrice de mots – la Morphologie – avant la syntaxe, qui comprend tous les mots déjà formés, ne nous permet pas d'analyser pleinement toutes les constructions d'objets cognats dans des langues comme le portugais; pour cette raison nous allons surmonter le problème de la direction dans la formation des paires V-N, en utilisant une idée chère à la MD, déjà présentée au paragraphe 2. Selon ce modèle, les racines intégrées dans la syntaxe sont considérées comme non catégorielles et la même racine peut épeler deux ou plusieurs copies; dans ces conditions, une même racine peut produire deux éléments de vocabulaire, même de catégories différentes. Cela veut dire que, selon cette perspective, il y a, dans les cas des vrais OC, création de mots à partir d'une même racine, non catégorielle, et non pas une conversion des

racines ou des mots déjà formés, et encore moins une dérivation régressive. En syntaxe se construisent tous les éléments de vocabulaire nécessaires, par fusion de la racine avec les différents morphèmes fonctionnels (des morphèmes fonctionnels de catégorie et des morphèmes qui projettent les arguments, obligatoires ou facultatifs) (cf. aussi Marantz (1993), Pylkkänen (2002, 2008), Alexiadou (2001, 2006), Siddiqi (2009)).

Adoptant l'architecture de la grammaire proposée par la MD, la morphologie se trouve distribuée dans la structure syntaxique, soit en termes de formation de mots, soit dans un syntagme ou une phrase, ce qui exclut un composant génératif – la Morphologie – qui puisse précéder la syntaxe.

Il n'y a aucun doute que l'acceptation des vrais OC (*chorar um choro...*, *sonhar um sonho...*, *dormir um sono...*) est limitée syntaxiquement, comme indiqué dans la section précédente, par la présence d'un déterminant indéfini et d'un modificateur restrictif. Cependant, ces conditions sont facultatives aussi bien pour les OC apparentés (*cantar uma/a cantiga*, *dançar uma/a dança*) que pour les OH (*cantar um/o fado*; *dançar um/o tango*).

En supposant donc que la syntaxe est le seul composant génératif du système, les racines (l-morphème) abstraites et non catégorielles sont insérées dans la tête terminale, associées par fusion (*merge*) à un morphème fonctionnel (les *f*-morphèmes, qui peuvent être *n*, *v*, *a*), comme illustré en (36).

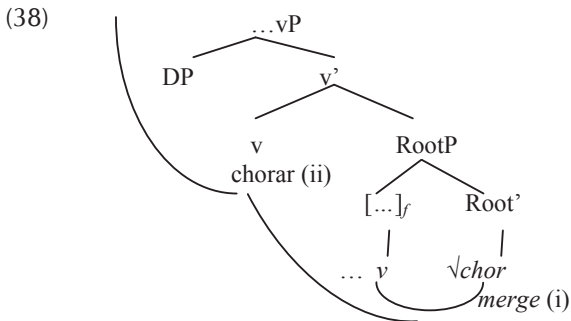


Cette opération permet la réalisation phonologique de la racine et forme l'élément de vocabulaire. Dans le cas particulier de nos constructions, les racines vont être associées à un *f*-morphème de type *v*.

Considérons d'abord le cas des constructions avec de vrais OC du type *a*, comme dans la phrase (14a), répétée ici en (37a).

- (37) a. Ele *chorou* um choro silencioso (vrai OC)  
b. Ele *chorou*. (OC caché)

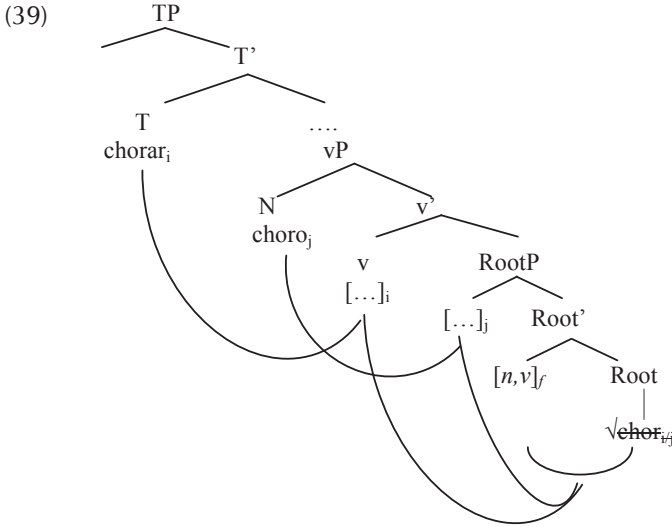
En suivant les idées précédentes de la MD et aussi le modèle de Alexiadou (2001), qui propose que la syntaxe verbale ait au moins deux niveaux (RootP et vP), nous proposons en (38) la structure partielle pour un verbe du type *chorar* [pleurer].



L'élément de vocabulaire résultant de la fusion de la racine *chor-* avec le *f-* morphème verbal (étape (i)) est, par le processus de mouvement par copie, copié vers *v* (étape (ii)), une catégorie fonctionnelle verbale responsable de la transitivité de la construction; après la montée, la copie laisse des traces phonologiques dans la position de la racine (la copie basse), qui peut ou non être visible dans le module Encyclopédie<sup>7</sup>. Cette dernière option permet la construction (37a) ou la construction (37b). En (37a) la tête, dans une seule *épellation*, fait deux copies: une copie haute (tête de [*v*, *v*']), qui monte ensuite afin de vérifier les traces de Voix, Aspect et Temps, spécifiques des formes verbales flexionnelles, et la copie basse (spécificateur de RootP), qui se déplace dans une position de spécificateur de vP pour vérifier les traces de l'accusatif, comme illustré en (39).

<sup>7</sup> Le module de l'Encyclopédie est l'un des trois éléments qui composent l'architecture de la grammaire en MD, contenant des entrées qui fusionnent avec les items de vocabulaire et qui peuvent accueillir des informations extralinguistiques liées à l'interprétation sémantique des éléments déjà formés grammaticalement.

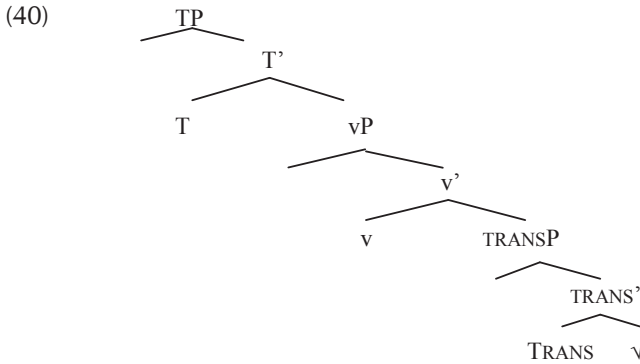




De cette façon, la construction à OC *chorar - choro* [pleurer-pleur] est produite.

Cependant, deux problèmes se posent: d'abord, il est nécessaire que la tête *choro* soit insérée dans un DP, en intégrant le déterminant indéfini et le modificateur, deux conditions de grammaticalité en PE; deuxièmement, la position dans laquelle la copie nominale (N) est projetée est une position de spécificateur de vP, laquelle, pour certains auteurs, est réservée à l'insertion de l'argument «externe», avant de monter vers la position de spécificateur de TP.

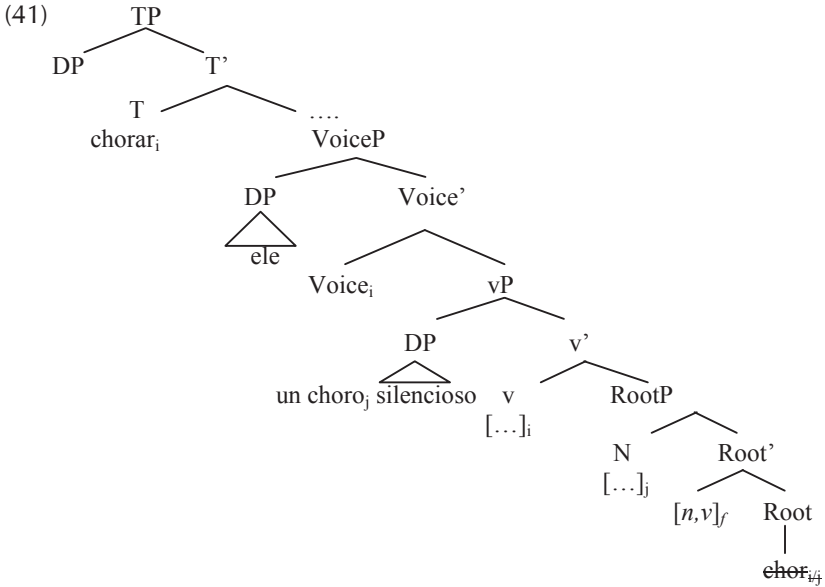
Selon Siddiqi (2009), un moyen de résoudre la deuxième question serait d'admettre que, entre la tête RootP et v', il doit y avoir une autre tête fonctionnelle, responsable de la visibilité du DP, qui contient le nom cognat et lui assigne l'accusatif. Pour Siddiqi (2009: 75), cette tête est source de l'argument thème, qui est projeté en position du spécificateur de TRANSP, comme illustré en (40).



La tête fonctionnelle TRANS fusionne avec la racine et permet l'insertion des arguments et l'attribution des rôles thématiques. Dans cette perspective, il y aurait autant de têtes fonctionnelles que d'arguments nécessaires à la structure syntaxique; les rôles thématiques sont plutôt conçus comme des informations attribuées par les têtes fonctionnelles dans la syntaxe et non comme des propriétés des mots, déterminées dans la composante lexicale. Admettre la structure de Siddiqi (2009) ne suppose pas défendre que les OC soient de vrais arguments, puisque les mêmes têtes fonctionnelles peuvent projeter différents types d'arguments, soit ceux qui sont exigés et requis sémantiquement par les traits idiosyncratiques de la racine, soit ceux qui sont requis après la fusion des traits particuliers de la racine avec les traces des têtes fonctionnelles. De ce point de vue, la transitivité est considérée comme fonctionnelle et syntaxique, étant responsable des arguments syntaxiques et sémantiques présents dans les structures, dans la forme phonologique.

Alexiadou (2006), à la suite de Kratzer (1996), résout le problème d'une façon plus intéressante que celle de Siddiqi (2009). Supposons une catégorie fonctionnelle verbale au-dessus de vP, la catégorie VoiceP, qui permet l'insertion de l'argument «externe», le sujet syntaxique, dans la position de spécificateur et libère la position de spécificateur de vP, laquelle, à partir de la suggestion ci-dessus, sert à accueillir la montée de la copie basse et les mécanismes d'attribution de cas accusatif.

Par conséquent, pour la construction (37a), nous proposons la structure en (41), qui résout quelques problèmes soulevés par la structure (39).



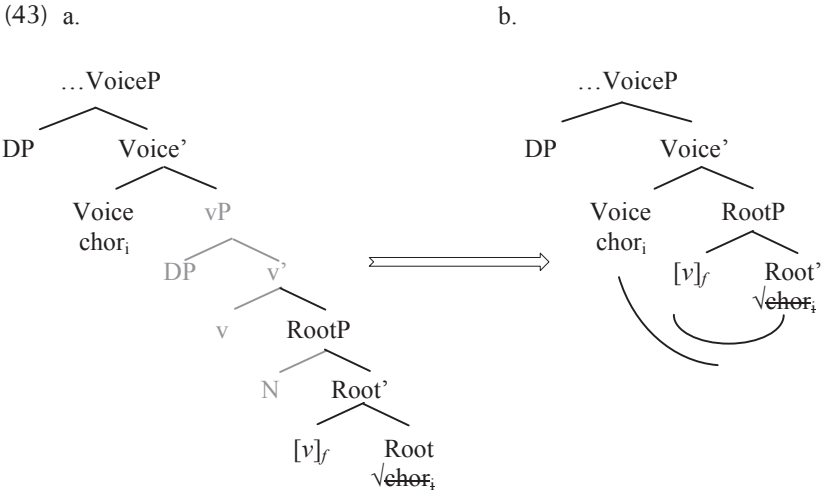
Cette approche ressemble à l'idée des têtes applicatives de Marantz (1993), développée pour analyser des constructions de Double Objet en anglais, et, en général, les constructions applicatives, dans les langues bantu. Ainsi, les catégories VoiceP et vP (ou TRANS P dans le modèle de Siddiqi, 2009) sont des têtes fonctionnelles verbales et, de ce point de vue, sont des têtes applicatives, qui permettent la construction de la grille argumentale et de la grille thématique des prédicats. Les têtes fonctionnelles sont cachées dans toutes les structures, mais elles ne sont activées que si des traits de la racine sont compatibles avec les traits des têtes fonctionnelles, ce qui permet d'expliquer, au moins apparemment, l'impossibilité de quelques verbes d'une langue particulière de se construire avec des OC, comme *chegar* [arriver] et *nascer* [naître] en PE, comme on peut le voir en (42).

(42) a. \*Ele chegou uma chegada inesperada.  
 'Il est arrivé une arrivée inattendue'

b. \*O bebé nasceu (de) um nascimento difícil.  
 'Le bébé est né (de) une naissance difficile'

Dans ces exemples, l'introduction d'arguments est bloquée par l'incompatibilité des traces intrinsèques de la racine et des traits spécifiques des têtes fonctionnelles.

Revenons aux exemples en (37). Si la construction ne porte pas d'OC, comme en (37b), un seul exemplaire de la copie sera conçu, la copie haute, celle qui contient la forme phonologique de *chorar*. Cependant, les traces de la copie basse ne sont pas supprimées, ce qui justifie, selon Haugen, qu'elles puissent être récupérées discursivement. Dans cette construction, les catégories fonctionnelles vP, v' et le spécificateur de RootP sont absentes (ainsi que le gris et le pointillé dans la structure en (43a) servent à indiquer), ce qui justifie l'intransitivité, comme décrit en (43b).



Notre proposition n'explique pas totalement le fait de ne pas y avoir une homonymie parfaite entre toutes les paires V-N dans les constructions avec de vrais OC, comme le démontre l'exemple (44).

- (44) \*dormir uma *dormida*  
 dormir une dormition  
 'dormir une dormition [dormir un sommeil]'

Une explication de l'agrammaticalité de (44) peut être fondée sur le fait que le PE fait un usage considérable de verbes légers (*light verbs*) (*ter* [avoir], *dar* [donner] et *fazer* [faire]) dans des paraphrases de ces constructions avec des OC: *fazer uma dormida* [faire une dormition].

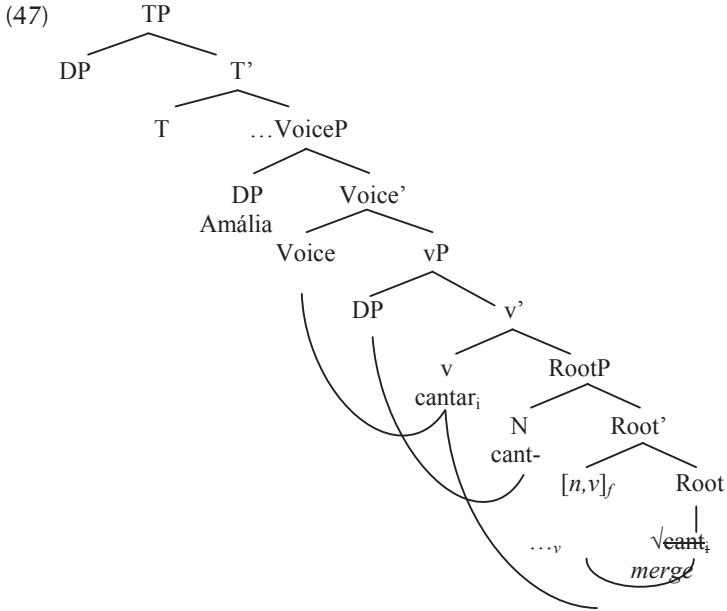
Considérons maintenant les exemples de constructions avec des OC apparentés (45) et avec des OH (46).

(45) Amália cantou *uma cantiga portuguesa*. (OC apparenté)

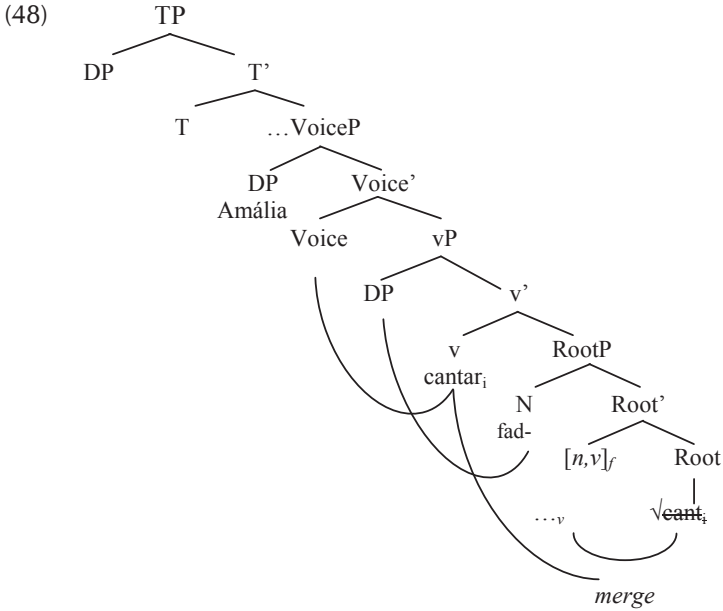
(46) Amália cantou *um fado*. (OH)

Les deux premières étapes de la dérivation sont identiques à celles des vrais OC: la racine abstraite est associée, par fusion, au *f*-morphème verbalisateur et par déplacement et copie elle se déplace vers la tête de la catégorie fonctionnelle [v, v'], ce qui laisse des traces dans la position d'origine. Cependant, il y a une différence importante à noter: seule la copie haute est phonologiquement visible; dans la position de la copie basse [spécificateur de RootP] est inséré, par insertion tardive, du matériel lexical nouveau, qui peut être cognat ou non cognat.

Dans le cas des OC apparentés (*cantar uma cantiga portuguesa* [chanter une chanson portugaise]), la seconde racine insérée est morphologiquement identique à la première (maintenant déplacée vers [v, v'] et ensuite vers [Voice, Voice']), comme décrit en (47).



Dans le cas des OH (*cantar um fado* [chanter un fado]), le matériel inséré est non cognat, et les traits de cette racine et de la copie basse (qui était cachée dans la même position, spécificateur de RootP) sont partiellement identiques, établissant une relation d’hypo / hypéronymie. La structure avec le OH est présentée en (48).



Il est important de souligner que la nouvelle racine (*cognate* pour les OC apparentés et non *cognate* pour les OH) est insérée directement dans la position de spécificateur de RootP et non dans le nœud terminal; la racine nominale insérée, soit celle de l'OC apparenté soit de l' OH, et la racine verbale sont projetées à différents moments de la formation de la structure hiérarchique; il y a donc deux épellations dans chaque construction. Ces propriétés sont le point qui nous permet de différencier la cognation apparentée et la cognation avec de vrais OC, où les deux copies (V et N) sont de la même racine et résultent de la même épellation (*spell out*), ayant beaucoup plus de contraintes sémantiques et syntaxiques.

#### 4. Quelques conclusions

Bien que l'anglais présente des cas d'homonymie parfaite entre des mots de classes différentes (*to sleep/ a sleep; to laugh/ a laugh...*), ce qui a justifié un traitement de ce phénomène basé sur le mécanisme de l'incorporation nominale, la distinction entre les OC et les OH présuppose une analyse plus complète et plus diversifiée que celle de Hale & Keyser (1993; 2002).

Quelques idées issues de la MD, en particulier celles de Haugen (2009), articulées avec la notion de mouvement par copie du PM et avec des approches fondées sur la notion de tête verbale applicative (Marantz 1993, Sidiqqi 2009, Alexiadou 2001, 2006), nous ont permis de faire une analyse des vrais OC, des OC apparentés et des OH.

En ce qui concerne les données pour le PE, cet article vise à développer une typologie des OC déjà présentée dans des travaux précédents (Choupina 2013). En effet, il y a une distinction claire entre les vrais OC et les OC apparentés, en raison de quelques critères / propriétés: relation morphologique V-N; indéfinitude; modification; substitution par OH; expansion avec «mais je ne sais pas quoi» et pronominalisation.

À la suite de Haugen (2009) et Alexiadou (2001, 2006), on a pu esquisser une analyse syntaxique des constructions étudiées, selon laquelle les racines sont non catégorielles, et c'est au niveau de la syntaxe que ces racines rejoignent des morphèmes fonctionnels et acquièrent leur structure argumentale et événementielle. De cette façon, le problème de la directionnalité de la dérivation N-V est résolu.

#### RÉFÉRENCES BIBLIOGRAPHIQUES:

- Alexiadou, A. 2001. *Functional Structure in Nominals: Nominalization and Ergativity*, Amsterdam: John Benjamins.
- Alexiadou, A. 2006. On the morphosyntax of (anti-)causative verbs. Disponible dans: <http://ifla.uni-stuttgart.de/files/artemis-anitaworkpap-revised3.pdf> (28/05/2013).
- Bosque, I. & Gutierrez-Rexach, J. 2009. *Fundamentos de Sintaxis Formal*, Madrid: Akal.
- Choupina, C. 2013. Contributos para uma análise sintática dos objetos cognatos em PE. In *Studia Romanica Posnaniensia*, Adam Mickiewicz University Press: Poznan, vol. XL/1, 2013, 59-79. ISBN 978-83-232-2542-3.
- Chomsky, N. 1995. *The Minimalist Program*. Cambridge MA: The MIT Press.
- Gallego, A. 2012. A note on cognate objects: cognation as doubling. In Bentzen, K. & Fábregas, A. (Eds.) *Nordlyd* 39.1: 95-112, University of Tromsø. Disponible dans: <http://www.ub.uit.no/baser/nordlyd/> (11/05/2012).
- Hale, K. & Keyser, S. J. 1993. On Argument Structure and the Lexical Expression of



- Syntactic Relations. In Hale, K. & Keyser, S.J. (Eds.) *The View From Building 20: Essays in Linguistics in honor of Sylvain Bromberger*. Cambridge, Mass., MIT Press, 53-109.
- Hale, K. & Keyser, S.J. 2002. *Prolegomenon to a theory of argument structure*. Cambridge, Mass., MIT Press.
- Halle, M. & Marantz, A. 1993. Distributed morphology and the pieces of inflection. *The View from Building 20*. In K. Hale & S.J. Keyser (Eds.), 111–176. Cambridge, Mass.: MIT Press.
- Harley, H. & Noyer, R. 1999. State-of-the-Article: Distributed Morphology. *Glott International* 4.4, pp 3-9.
- Haugen, J. D. 2009. Hyponymous objects and Late Insertion. In *Lingua* 119, 242-262.
- Horita, Yuku 1996. English Cognate Object Constructions and Their Transitivity. In *English Linguistics* 13, 221-247.
- Kitahara, Ken-ichi 2006. On the Form and Meaning of Cognate Object Constructions in English: A Construction Grammar Approach. In *Tsukuba English Studies*, vol. 25, 125-126. Disponible dans: <http://www.tulips.tsukuba.ac.jp/mylimesio/dl/page.do?issuclid=867952&tocid=100066195&page=125-126> (27/05/2012)
- Marantz, A. 1993. Implications of Asymmetries in Double Object Constructions. In Mchombo, S. A. (Eds.) *Theoretical aspects of Bantu Grammar*, vol. 1. Stanford: CSLI Publications, 113-150.
- Pustejovsky, J. 1998. *The Generative Lexicon*. Cambridge, Mass: The MIT Press.
- Pylkkänen, L. 2002. *Introducing Arguments*. Ph. D. Dissertation, MIT.
- Pylkkänen, L. 2008. *Introducing Arguments*. Cambridge: MIT Press.
- Ramchand, G. 2008. *Verb Meaning and the Lexicon: a first phase syntax*, Cambridge University Press.
- Real-Puigdollers, C. 2008. The Nature of Cognate Objects. A Syntactic Approach. In *Proceedings ConSOLE XVI*, pp. 157–178. Disponible dans: <http://media.leidenuniv.nl/legacy/console16-real-puigdollers.pdf> (27/05/2011).
- Rodrigues, A. S. 2001. *A construção de postverbais em português*. Porto: Granito Editores e Livreros.
- Siddiqi, D. 2009. *Syntax within the Word. Economy, allomorphy, and argument selection in distributed morphology*, Linguistik aktuell/Linguistics Today: JB.
- Silva, M. L. 2010. *As construções com objeto cognato em Português – análise e esboço de modelagem sociocognitiva*. Disponible dans: [http://marcelolopesdasilva.com.br/index.php/textosdoautor/doc\\_details/12-2010b-as-construcoes-com-objeto-cognato-em-portugues-analise-e-esboco-de-modelagem-sociocognitiva.html](http://marcelolopesdasilva.com.br/index.php/textosdoautor/doc_details/12-2010b-as-construcoes-com-objeto-cognato-em-portugues-analise-e-esboco-de-modelagem-sociocognitiva.html) (27/05/2012).



Varia



# Evidentials *dizque* and *que* in Spanish. Grammaticalization, parameters and the (fine) structure of Comp\*

Violeta Demonte

violeta.demonte@cchs.csic.es

*Instituto de Lengua, Literatura y Antropología - Spanish National  
Research Council*

Olga Fernández-Soriano

olga.fernandez@uam.es

*Universidad Autónoma de Madrid (Spain)*

ABSTRACT. In this paper we study one type of Spanish *que* (the equivalent to the complementizer “that”) that can be shown to belong into the crosslinguistically restricted list of evidentials. In particular, we will claim that it encodes the (most basic) marks of non-first-hand or *indirect* (reported) evidence. Our point of departure is certain (apparently) independent clauses of Spanish headed by an overt complementizer (*que*). Some tests will be presented that support the idea that that one type of *que* introducing a well specified subset of root sentences shares most of the properties that have been claimed to characterize reportative evidentials in languages such as Quechua (Faller 2002, 2006). As for the properties of reportative *que*, it will be further shown that it does not encode any features related to epistemic modality (reliability or (im)probability) and we will propose that it is better analyzed as an illocutionary operator, affecting the illocutionary force (in line with Faller 2002 among others) and not as an epistemic modal (Izvorsky 1997 among others). In order to determine the nature of this reportative element and its origin, we contrast it with an old Spanish form, *dizque*, which exists nowadays in certain modern American varieties. This particle also has the properties of an evidential but behaves as an epistemic modal. In the last sections, we will propose that both evidential particles (*que* and *dizque*) are the result of a process of grammaticalization (i.e. ‘upward reanalysis’, or categorial change, of functional material, in the sense of Roberts and Roussou 2003) of the complex structure headed by a communication verb, *dicen que* “they say that”. We will tentatively describe such process and introduce a hypothesis as to the nature and role of the parameter involved in the claimed reanalysis.

---

\* Previous versions of this paper have been presented at the XLII Symposium of the *Sociedad Española de Lingüística*, and at the 61 *Seminário CEL* (USP; Brazil). We thank both audiences for helpful comments and suggestions. We deeply thank Heloisa Salles for her interesting suggestions and observations. Thanks also to Elena Castroviejo, and to M. Jesús Torrens for crucial help with one text in Old Spanish. The research behind this paper has benefited from the research projects FFI2012-32886 and FFI2011-23829.

KEY-WORDS. Evidentials, complementizers, grammaticalization, epistemic modals, illocutionary operators.

## 1. Introduction

In this paper we will study one type of Spanish *que* (complementizer “that”) and propose that has to be incorporated into the crosslinguistically restricted list of ‘evidentials’. Our point of departure is certain (apparently) independent clauses of Spanish which are headed by an overt complementizer, as those in (1).

- (1) a. ...de repente se oye detrás: oye, **que** Manolo puso la bandera.  
 (CREA, ORAL Spain)  
 ...suddenly it is heard from behind, listen *that* Manolo put the flag
- b. (Oye), **que** ha dimitido el decano.  
 Listen, *that* the dean has resigned

Those in (1) are root clauses which can be discourse initial (pronounced out-of-the blue, as can be seen in (1a)). The semantic contribution of *que* is to incorporate reference to a speech event heard (and reported) by the speaker. So the sentence in (1b) is used in a context where the dean has resigned and someone (different from the speaker and the hearer) actually *said* it. We will hypothesize that *que* in this particular construction is an (indirect) ‘reportative’ grammatical evidential, similar to the ones that have been identified for American Indian languages, as well as for languages like Turkish, Balkan languages, Tibetan, Japanese, Korean, etc. (Chafe & Nichols 1986, Aikhenvald 2006).

This paper is organized as follows. In section 2 we define evidentiality and we briefly summarize the three semantic views on the conceptual nature of evidentials. In section 3 we present and describe the Spanish evidential *dizque*. We trace its origins from Old Spanish to some present forms in American Spanish and we also justify the meanings and functions of this form, which, as an evidential, is better characterized as an epistemic modal. In section 4, mainly based on Demonte & Fernández-Soriano (in

press), we introduce a set of diagnoses that allow us to analyze the Spanish *que* appearing in well specified Spanish root sentences as a reportative (indirect) evidential. In 6 we develop a syntactic and semantic-pragmatic formal analysis of this form which we define as an illocutionary operator (as opposed to the form *dizque*). Finally, in section 6 we introduce the view of grammaticalization in formal grammar and we hypothesize that there is a path from *dicen que* “they say that” to *dizque* and *que* and that it meets the main features of a case of grammaticalization (‘upward reanalysis’). We finally present a conjecture as to the parameter that could be involved in this process.

## 2. Evidentiality. Three views on evidentials

Evidentiality is a linguistic category encoding speaker-oriented qualifications of propositions in terms of the evidence they are based on. All languages have a way of expressing the speaker’s source of information by lexical devices such as adverbs like *allegedly*, *reportedly*, etc., verbal constructions such as *it is said*, *I heard*, and other lexical means. These are called ‘evidentiality strategies’ (Aikhenvald 2004). But only some languages grammaticize evidentiality and encode it in their (inflectional) morphology or in their particle system (complementizers, for instance) (cf. Willett 1988, Palmer 1986). Evidentials are thus generally morphological (verbal) markers or particles, some derived from verbs like *see*, *hear* and *say* (Gordon 1986, Aikhenvald 2006). As Speas (2004: 255) puts it “some languages have evidential morphemes which mark the Speaker’s source for the information being reported in the utterance”. As an example we will take evidentials in Quechua, as described in Faller (2002). Quechua has three types of evidentials (which are always enclitic): direct, reportative and conjectural.

### (2) Para-sha-n-mi/-si/-ch´a.

rain-prog-3-bpg/rep/conj

‘It is raining.’

-**mi**<sub>EV</sub>: s sees that it is raining.

-**si**<sub>EV</sub>: s was told that it is raining.

-**ch´a**<sub>EV</sub>: s conjectures that it is raining. (From Faller, 2006)

We will be concerned with what have been called ‘indirect’ evidentials. Izvorsky (1997) examines the meaning of indirect evidentials such as the perfect of evidentiality of languages like Bulgarian, Turkish and Norwegian and compares it to English adverbs like *apparently*. Essentially, this author analyzes evidentials as “epistemic modals with a universal modal force and a more restricted domain of quantification than that of ‘ordinary’ epistemic operators” (1997:225. See also Chung 2007, Matthewson *et al.* 2007, Rullmann *et al.* 2008). Basically evidentials would mark the speaker’s degree of certainty and/or the necessity/possibility of the truth of the propositional content.

In an alternative view, evidentials are not considered as epistemic modals but have been analyzed as encoding illocutionary modifiers (e.g., Faller 2002, 2006, 2007) which affect the illocutionary force, including the illocutionary points and sincerity conditions. According to Faller, evidentials “add to or modify the sincerity conditions of the act they apply to” (2002: 231), they are functions from speech acts to speech acts. Conceptually the marking of the speaker’s degree of certainty and/or the necessity/possibility of the truth of the propositional content would be clearly distinct from the kind of evidence a speaker has, although the later may of course often determine the former. The view of evidentiality as a conceptual category distinct from epistemic modality does however not preclude the possibility (which is well attested in the world’s languages, see for example Willett (1988), Chafe and Nichols (1986), Aikhenvald and Dixon (2003), and Aikhenvald (2004)) that specific linguistic markers may combine both. There is also an intermediate position (e.g., Garrett 2001), according to which evidentiality can be encoded as epistemic modals in some morphemes and as illocutionary operators in others (See Lim 2010).

Going back to the examples in (1), we have to note that these sentences do not involve any kind of modality, in the sense that the notion of doubt is not present in their meaning, not even as a pragmatic inference. The sentence introduced by what we will prove to be an evidential *que* is presented as a (true) assertion that has been previously heard by the speaker. This suggests that evidential *que* is an illocutionary operator.

In order to justify the previous claim and at the same time clarify the status of the *que* in (1) we will first compare it to an old form, which remains



in some modern varieties of (American) Spanish. In particular, we will now analyze the form *dizque*, derived from *dicen que* “they say that” which, being an evidential is not identical to our *que*, in the sense that its main properties are clearly those of an epistemic modal.

### 3. The form *dizque* as an evidential

#### 3.1 Historical Data

As Kany (1944) points out, Old Spanish form *diz que* stood for *dicen que* “they say that” or *se dice que* “it is said that”. It was common in the old language and it began to decline shortly after 1500. The author also notes that

[...] nevertheless *diz que* did not become obsolete; it became dialectal, provincial, or rustic. At the beginning of the seventeenth century Covarrubias (*Tesoro*, 1611, p. 324) registered it as “palabra aldeana, que no se deve usar en Corte.” It lingered on in regional literature and speech into the nineteenth century. Even today it is occasionally heard in restricted areas of Spain, but only as an archaism in familiar or jocose style. (Kany 1944: 168).

As can be seen in the following example of a narrative the meaning of this form is that of source of information, that is usually attributed to reportative evidentials.

(3) ... vase a la comedia, que **diz que** estaba cuajá de señorío prencipal . . . fuéronse a ver a la enferma, que **diz que** paecia un sol. (Pereda, *Obras completas*, VI, Madrid, 1897, p. 140, 409, from Kany 1944, fn. 1).

‘(He) goes to the comedy that *was said* to be crowded with gentry... they went to see the sick woman who *they say* looked like a sun.’

There are many other cases of *dizque*<sup>1</sup> with a reportative evidential value. In CORDE, we found 84 cases out of 54 Spanish documents. Some are given below:

---

<sup>1</sup> This form has many variants in Latin America: *izque*, *isque*, *i que!* y *que*, *es que*, *quizque*, *quesque* among others.

- (4) E Cartagena le respondió, que con el mejor marinero de la nao le habia salvado, y que quizá otro dia le salvaria con un page. Y **dizque** dende, en tres dias el dicho Cartagena no lo tornó á saludar. [1521, López de Recalde, J.: *Carta del contador Juan López de Recalde al Obispo de Burgos*]  
 ‘And Cartagena answered him that he had saved him with the best sailor in the ship and that maybe another day he would save him with a page. And *it is said* that since then the above mentioned Cartagena did not say hello to him for three days’.
- (5) & a Nós es fecha relación que [...] los dichos alcaldes de la tierra no visitan commo deuen, ni exsecutan en ella, la nuestra justicia, ni oyen los querellosos commo deúan; antes, **dizque** ponen su<sup>2</sup> tenientes en las dichas alcaldías ombres legos, escuderos & otras personas no vsadas de tener judgado, los quales **dizque** fazen injusticias & estorsiones a los vezinos. [1492, Anónimo: *Ordenanzas reales de la ciudad de Sevilla*]  
 ‘And to us was reported that ... the mentioned mayors of the land do not visit as they ought to, nor do they execute in it, our justice, nor do they listen to the complainants as they should; on the contrary, *it is said that* they put as their lieutenants in the above mentioned [...] mayoralities, laymen, squires and other people not used to having judged, who *it is said that* make unfair things and extortion to the neighbours’.
- (6) ¡Con induztria **dizque** le engolda! [1617, Suárez de Figueroa, C.: *El pasajero*]  
 ‘With industry *they say that* he feeds him’.

It is important to notice that in all these cases the *dizque* has the added meaning of “doubt” or uncertainty about the proposition it introduces. Crucially, as Kany observes, “... in many instances their force is reduced to that of an adverb of doubt.” (Kany op. cit. p. 171). In this sense it contrasts with the reportative *que* in (1) that we will analyze in section 3. First we

<sup>2</sup> For a better understanding and a relatively different and clearer version of this *Ordenanza*, see I. Carrasco Cantos & P. Carrasco Cantos, *Estudios lingüísticos de las Ordenanzas de Sevilla de 1492*. Málaga, Universidad de Málaga, where it reads:

(i)[...] antes/ diz que ponen **por** sus tenjentes en las dichas alcaldias onbres/ legos, escuderos & otras personas non vsadas de tener juzgados los/ quales diz que hasen ynjusticias & extorsiones a los vezinos de la tierra [...]

Nothing crucial for our purposes hinges upon the text chosen.

will concentrate on the properties of the form *dizque* in modern American Spanish.

## 2.2 *Dizque* in American Spanish

Laprade (1976) analyzes the form *dizque* (as well as the use of certain verbal forms) as expressions of evidentiality (non-first hand information) in the area of La Paz (see also Hardman 1986, Klee & Ocampo 1995) and relates this use in the Andean variety of Spanish zone to contact with Quechua and Aymara, which have evidentiality systems. Travis (2006) also analyzes this evidential form in Spanish spoken in Colombia (where there is no contact with languages with evidentiality systems). The important fact for our purposes is that, as Travis notes:

The range of use of *dizque* extends from functioning as a purely evidential marker, encoding reported speech and hearsay with a notion of doubt implied in some contexts, to a marker of epistemic modality, encoding extensions of the notion of doubt implied in its evidential use and nothing about source of information" (Travis 2006: 1269).

Travis distinguishes thus a true reportative evidential encoding source of information, which can carry an added value of doubt. That value, according to Travis "is available as a pragmatic inference according to the context, but is not inherent in the semantics of *dizque* itself" (2006: 1272). In a second case *dizque*, a kind of adverbial form as can be seen, for instance, from its position within the sentences, is only a marker of epistemic stance, "source of information is no longer encoded, and the notion of doubt has been conventionalized and has even been extended to one of nonvolitionality." (Travis 2006: 1272). Examples are given below of some of the different values carried by *dizque*:<sup>3 4</sup>

---

<sup>3</sup> Travis (2006) describes other related values, such as "labeling". See the cited work for details.

<sup>4</sup> Treviño (2008) reports another use of *que* in Spanish spoken in México, whose distribution seems to be similar to *dizque*. This *que* is subject to prosodic restrictions. The author provides the following examples:

- (i) a. Lo ponen así **que** para ver los nervios.  
They put it like that *que* to see the nerves  
b. ¿Dónde lo habrá conseguido? —Lo compró **que** en Sanborns.  
Where did he get it? — He bought it *que* in Sanborns  
c. (Reading the label of a product)  
- **Que** hay que guardarlo **que** en el congelador.  
Que it has to be stored *que* in the freezer

*Reportative value* [with an added value of doubt]:

- (7) Por ejemplo, el a – aquí el alcalde, Todo lo que ha hecho, Y...y ahorita, **dizque** ya lo están investigando.

'For example, the mayor here, all that he's done, and now, *dizque* he's under investigation.' (Travis 2006: ex. 12)

*Dubitative/mirative value*:

- (8) . . . se abrió **dizque** a comerciar. Comerciar era llegar a los almacenes, pedir, no pagar, y exigir dinero de vuelta. . . les sacaban dinero **dizque** de vuelta.

' . . . she began *dizque* to do business. Do business was to go into the shops, make an order, not pay, and demand change. . . they would get money out of them *dizque* in change.' (Travis 2006, ex. 24 *apud* Castro Caycedo 1994: 100–101)

The diachrony of *dizque* appears thus to be the following. Basically *dizque* started as a pure evidential, which encodes source of information (the most extended value); from there it extended its meaning to mark the information as doubtful (or even false). In this case, it can take first person subjects, as opposed to the pure reportative evidential use (see below). This is a more restricted use. From this notion of doubt it evolved into a pure epistemic modality marker. Travis (2006: 1270) notes:

[...] the semantic change from reported speech to doubt is a regular crosslinguistic pattern that has been widely commented on in grammaticized evidential systems. [...] This is related to the nature of reported speech: attributing an utterance to someone else allows the speaker to distance him/herself from the material being presented, and thus such an utterance can take on overtones of speaker doubt about the veracity of the information.

In the following section we will turn to sentences in (1) in order to complete our view of the Spanish set of evidentials.

#### 4. Non subordinate *que* in Spanish. A reportative evidential

##### 4.1 Two types of non subordinate *que*

In Demonte & Fernández-Soriano (in press) sentences like (1) are analyzed. Other similar examples are provided in (9):

- (9) a. **Que** el Barça ha ganado la Champions. [Etxepare 2007: 25-26]  
That the Barça has won the Champions
- b. (Bueno, pues estaba una mañana en Interview y me llamó mi well, so I.was one morning in Interview and to-me called my hermano y me dice:) Oye, **que** ha sido depuesta Benazir. brother and to.me tells: listen that has been deposed Benazir  
'Well, one morning I was in Interview and my brother calls and says: hey, Benazir has been deposed.' (CREA, oral, Spain)
- c. En la cena nos llama un compañero, oye **que** está nevando en el at the dinner us calls a companion, listen that is snowing at the campo de vuelo. field of flying  
'At dinner a colleague calls and says, hey it is snowing at the flying field'. (Taken from the Internet)

These sentences, as opposed to the corresponding versions without *que*, refer to a speech event reported by the speaker. Another property of these structures is that they are matrix clauses which can (and most often are) discourse initial (pronounced out of the blue). Our claim is that *que* in these cases is a reportative evidential. In this two aspects they are different from sentences of the type in (10) that we would like to briefly present just to clarify the data:

- (10) a. Moment A:  
- Viene el autobús  
Comes the bus  
Moment B:  
- **Que** viene el autobús (¿no me oyes?)  
that comes the bus (can't you hear me)

- b. Sí/naturalmente **que** me iré.  
Yes/naturally that I.will.leave
- c. Speaker A:  
- He votado al PP / María es estupenda.  
I.have voted to.the PP/ Mary is great  
Speaker B (scornfully / angrily):  
- ¡**Que** has votado al PP!! / ¡**Que** María es estupenda?  
That you-have voted to-the PP/ That Mary is great
- d. Y él, **que** llegábamos tarde, **que** no se podía salir con nosotros...  
And he that we-arrived late that not one-can get-out with us  
'And he kept saying that we were late, that you cannot meet up with us'
- e. (Vamos/ vaya) **que** no aguantas más  
(In-sum...) that not you -bear anymore  
'In sum, you cannot bear it anymore, right?'

In Demonte & Fernández-Soriano (in press) we show that these are echoic structures, where: "the speaker does not report a particular state of affairs but reproduces or refers to another utterance or thought to show her reaction" (Wilson 2006). Different from the sentences with an evidential *que* that we will analyze in the following subsection the sentences in (10) are instances of root sentences introduced by *que*, similar to other 'in subordinate' sentences<sup>5</sup> (Demonte & Fernández Soriano in press), some of them depending on a silent verb and where *que* had the added echoic value. The structure proposed for those cases is the one in (11):

(11) (V) [ForceP [**que**... [TP...]]]

In what follows we will come back to the *que* in (1) and (9) and show that is a reportative evidential similar to the ones in Quechua, which appears above the left periphery.

<sup>5</sup> Some of these cases have been reported in some varieties of spoken English. Radford (2013) provides data such as the following, taken from English broadcast media:

(i) a. *Obviously that* the Achilles was giving him a bit of a problem (Ian Chappell, BBC Radio 5)  
b. *Inevitably that* there'll be some temptation there for cricketers (Gus Fraser, BBC Radio 5 Sports Extra)

#### 4.2 Properties of reportative *que*

a) Reportative evidentials never report a speaker's or a hearer's saying. Reportative *que* also shows this property. This is, in fact, a very clear restriction: the president of a nation, for example, cannot report his own war declaration (which may well be a speech event) headed by this *que*.

(12) #Ciudadanos, **que** {se ha/ hemos} declarado la guerra.

Citizens that it has been/ we have declared the war

b) Another property of reportative evidentials is that they are restricted to declarative sentences. More specifically, Aikhenvald (2004:242) notes that declarative sentences are the most natural environment for evidentials to occur cross-linguistically. Evidentials can also occur in questions, at least in a subset of the languages which possess an evidential system. They seem quite infrequent in commands (Aikhenvald, 2004:250). Aikhenvald does not comment on other minor clause types, such as exclamatives. As can be seen in (13), all these clause types are incompatible with reportative *que*:

(13) a. #Oye, **que** ¡qué bonito día hace!

Listen, that what nice day it is

b. #Oye, ¿**Que** hemos ganado la liga?<sup>6</sup>

listen, that we have won the league?

c. ??Oye, **que** !andando/ a comer / que nos vayamos!

Listen that walking / to eat / that we go

c) Faller 2002 shows that reportative evidentials do not allow for the speech eventuality they imply to be accessed by linguistic operations bearing on propositional truth, such as negation/dissent. Negation/dissent can only access the proposition introduced by the evidential, not the "source of the information" it refers to. Once again, reportative *que* behaves on a par with (indirect) evidentials in this respect: the source of information introduced by *que* cannot be affected by negation/dissent. This is why

---

<sup>6</sup> Faller (2006) and other authors show that some evidentials can be under the scope of interrogatives, giving rise to particular evidence "shifts" (see Lim (2010)). Davis, Potts and Speas (2007) argue that this contribution is genuinely pragmatic.

the second part of the following dialogue cannot mean that the speaker did not actually hear that Madrid has won the Champions league but only that the proposition itself (that Madrid has won) is false, as the impossibility of (14b) indicates:<sup>7</sup>

- (14) - Oye, **que** el Madrid ha ganado la Champions.  
 Listen, that Madrid has won the Champions  
 - No, qué va, no pueden haber ganado.  
 No, no way, they cannot have won.  
 - #No, hombre, no has escuchado eso en ninguna parte.  
 No, man, you have not heard that anywhere.

d) It has been noted that indirect/reportative evidentials are common in folklore tales (see Lim 2010). As Aikhenvald (2006:324) notes: “the genre of the text may determine the choice of an evidential. Traditional stories are typically cast in reported evidentials”. In this sense, it is interesting that, in the (orally transmitted) Spanish literature, one can trace some examples of reportative *que* like the following:

- (15) **Que** de noche lo mataron / al caballero. La gala de Medina, la flor de Olmedo.  
 That of night CL they-killed / To-the knight / The jewel of M the flower of O  
 ‘For at night they killed/ That noble soul/ The jewel of Medina / The flower of Olmedo’ (*El Caballero de Olmedo*, Lope de Vega)

e) First person effect. Aikhenvald (2006) notes that reportative evidentials carry additional meanings if the proposition they heard has a first person subject. These are basically new information, ‘unprepared mind’ and surprise. In other words, when the subject of the prejacent is first person, the sentence carries an additional implication that the speaker is not aware of her act, or does not believe what she is asserting (Lim 2010:60-63). This kind of effect is also found in the Spanish *que* under analysis, as shown by

<sup>7</sup> In this sense they are different to epistemic modals. But see Faller (2006) for a further observation suggesting that what has to be taken into consideration is also “external scrutability”.



the following examples:

(16) a. Scenario: Listening to the lottery results, someone suddenly hears his number:

(Oye,) **que** he ganado la lotería.

Listen that I-have won the lottery [Surprise]

b. Scenario: Someone receives a letter saying that she has been nominated Dean:

(Oye,) **que** soy la nueva decana.

Listen that I-am the new dean [Unawareness]

c. Scenario: There is a party, the bell rings, a neighbor complains about the noise:

(Oye,) **que** somos muy ruidosos y tenemos que irnos.

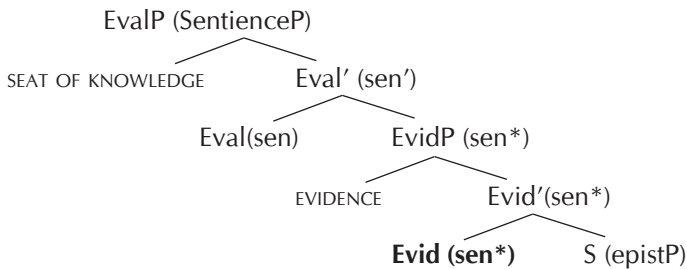
Listen that we-are very noisy and we-have to leave

[Surprise, disagreement]

5. The syntactic and semantic/pragmatic analysis of evidential *que*

In this section we will propose a syntactic structure for the constructions headed by the just described evidential *que*. Our point of departure is the approach to discourse markers proposed by Speas & Tenny (2003). Within this framework, evidentials head an Evidentiality Phrase inside a Speech Act Phrase, as depicted in (17). It is our contention that *que* is the head of this EvPhrase.

(17) [Speech Act Phrase [ SPEECH ACT [CP ...]]]



As for the semantics of evidential *que*, we have already suggested that it is better characterized as an illocutionary operator, which modifies the illocutionary force of the sentence. We thus adopt Faller's analysis in (18):<sup>8</sup>

$$\begin{array}{ccc}
 (18) & \text{ASSERT}(p) & \text{PRESENT}(p) \\
 & \text{que} & \longrightarrow \\
 & \text{SINC} = \{Bel(s,p)\} & \text{SINC} = \{\exists s_2[\text{Assert}(s_2, p) \wedge s_2 \notin \{h,s\}]\}
 \end{array}$$

(Faller 2002:200)

(18) corresponds to Faller's representation for Quechuan indirect evidential *-si*. The illocutionary act in (18), that Faller calls Presentation, would be defined by the sincerity condition (SINC) that someone distinct from the speaker and the hearer has asserted the content of the report. No correlation exists between the presence of reportative *que* and the degree of certainty with which the speaker believes (*Bel*) the embedded proposition, however. Its use is entirely compatible with a situation in which a speaker is convinced that the prejacent proposition is true, as well as with a situation in which the speaker is convinced that it is false (see the first person effect cases).

This would explain its initial position, always restricted to sentential scope. Another fact that would be accounted for under the analysis proposed is the incompatibility of *que* with other evidentials (and not with epistemic modals such as *posiblemente*, as in (19a)), as expected. Example (19b) shows that reportative *que* cannot coappear with conditional mood, which behaves as a reportative evidential in journalist style:

- (19) a. Oye, **que** *posiblemente* el Banco Central ha sido asaltado esta mañana por unos adolescentes.  
 Listen, that probably the Bank Central has been assaulted this morning by some teenagers
- b. #Oye, **que** el Banco Central *habría sido asaltado* esta mañana por unos adolescentes

<sup>8</sup> As Faller (2006) notes, the illocutionary operator nature of evidential *que* should indicate that it could occupy a position higher than Speech Act rather than below it. We will not pursue this issue here.

Listen, that the Bank Central would-have been assaulted this morning by some teenagers  
Intended meaning: The speaker reports what the news give as reported information.

As a final remark, let us say that since evidential *que* is restricted to root clauses and only takes declarative structures, the possibility of embedding under conditionals is precluded, as seen in (20):

- (20) \*Oye, si **que** el decano ha dimitido, no hay que preocuparse.  
Listen, if that the Dean has resigned there is no anything to be worried about

Faller claims this to be a property characteristic of illocutionary operators, which crucially distinguishes them from epistemic modals. The reasons why illocutionary operators cannot be under the scope of conditionals are however more complex (see Faller 2006 for a detailed explanation). The behavior of *que* with respect to negation also indicates that it behaves like illocutionary operators. The same is true for the assent/dissent test.

#### 6. Variation and change: from *dicen que* “they say that” to reportative *que*. A case of grammaticalization

As said above, it is an extended idea that *dizque* derives from the construction ‘verb *decir* ‘say’ plus complementizer *que*.’ In this section we will set this descriptive generalization into Robert & Roussou (2003) theory of grammaticalization and we will suggest a process whose final step is the bare evidential *que* that we have just identified.

##### 6.1 The notion of grammaticalization

According to standard definitions, grammaticalization is seen as the creation of new functional material through the reanalysis of other existing either functional or lexical material (Lehmann 1985). A typical example of grammaticalization is that of the evolution of English main verbs to auxiliaries; another relevant case would be the evolution from Latin functional demonstrative *ILLE* to Romance article *il/le/el*. In functionalist

approaches, grammaticalization is conceived as a specific path of change. In formal grammar, the most conspicuous approach to linguistic change and grammaticalization is Roberts & Roussou (2003) (henceforth R&R). These authors argue for the strong hypothesis that grammaticalization is a regular case of parameter change rather than an isolated or specific type of change. Thus in this view grammaticalization is seen as an epiphenomenon of other more general processes. We will adopt this general view in our attempt to explain the evolution from *dice que* to *dizque* and possibly to *que*.

R&R (2003) give two main properties of grammaticalization:

- a) "Grammaticalization is *upward* reanalysis due to parameter resetting."
- b) "Grammaticalization is reanalysis [that] gives rise to a new exponent for a *higher* functional head X." (R&R 2003: 300) (our italics)

These properties locate the main features of grammaticalization processes, mostly accepted by all approaches, inside the frame of a parametric theory of linguistic change. We will refer to this approach as 'formal grammaticalization'. In general terms, grammaticalization goes along with:

- a) (Syntactic) Reanalysis: (Upwards) categorial change of lexical or functional material,
- b) (morpho) phonological reduction (from *ILLE* to *le*, for instance),
- c) semantic bleaching (cf. Hopper & Traugott 1993: 87): loss of a semantic property (demonstrative property, in the case of articles) or "change in the meaning of the reanalysed element." (R&R 2003: 219).

More specifically, reanalysis and grammaticalization consist in the formation of new functional material that carries a category change and a structural simplification (R&R 2003:2). Crucial to this theory of change is then the idea that functional heads (Comp, *v*, T, Det, Ev, etc.) are present in all languages, although they may not be realized morphophonologically, and that this P(honetic) F(orm) realization may be achieved either by (external) Merge or by Move (material from other node in the clause

structure is moved to the functional head). Things being that way, language change is the result of change in the PF realization of functional heads; syntactic change is caused by phonological change and semantic change. Change occurs, according to these authors, when the trigger experience for a parameter setting (e.g., whether a functional head is realized by Merge or Move) is ambiguous or obscure.<sup>9</sup>

In the next subsections, we will present in a very schematic and tentative way our proposal that the evolution from *dicen que* to reportative *que* fits into this general frame and is thus a case of formal grammaticalization. We are aware that complete development of this hypothesis needs a thorough empirical study as well as a more careful approach to the question of parameter resetting in Spanish.

## 6.2 Our case: from *dicen que* to *dizque / que*

To sustain our hypothesis we rely on Rizzi's (1977) theory of the structure of sentence left-periphery (see below the structure in (23)) extended with Speas and Tenny's proposal about encoding of pragmatic features in syntactic structures (see above (17)). We also accept the extended idea that functional categories bear syntactic features which trigger syntactic operations. With these provisos, we postulate a process in the terms that follow.

Whatever its order in the whole story might be, a crucial step is the loss of features of the functional category Comp. Roussou (2000) has shown that the Force head of Rizzi's (1997) ForceP splits into two (sub) heads: "the clause-typing one, which is essentially an Operator (Op) head [...], and an even higher C head that has the properties of a 'subordinator' (in the sense that it functions as a clause-linking element)" [R&R 2003: 78]. We believe that it is the subordinator feature that is lost in the evolution under study. Although this assumption is in need of further testing and clarification, it appears to us that in a certain period there is ambiguity between the two meanings and uses (we also contend that together with the loss of the categorial feature in Comp the meaning of *decir* is blurred). Recall the example in (4) that we repeat below as (21); in this sentence *dizque* may be understood as the complex V+Comp or as a report of the speaker uncertainty in front of the reported sentence (i.e., as pure evidential).

---

<sup>9</sup> "A change occurs when the trigger experience for a parameter setting provided by the input has become obscure or ambiguous. This can happen in a variety of ways, for example through language contact, morphophonological erosion, etc." (R&R 2003: 12)

(21) E Cartagena le respondió, que con el mejor marinero de la nao le habia salvado, y que quizá otro día le salvaría con un page. Y **dizque** dende, en tres días el dicho Cartagena no lo tornó á saludar. [1521, López de Recalde, J.: *Carta del contador Juan López de Recalde al Obispo de Burgos*]

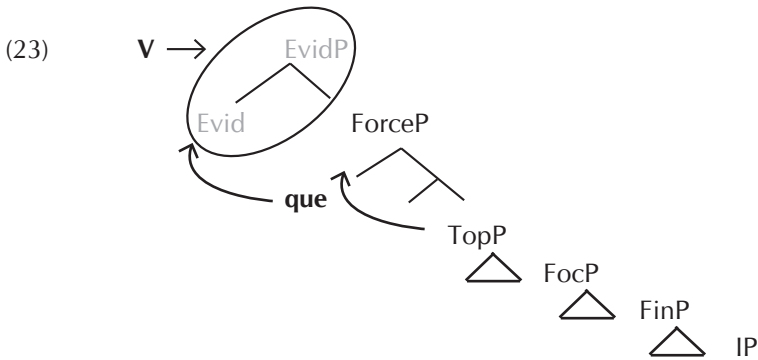
'And Cartagena answered him that he had saved him with the best sailor in the ship and that maybe another day he would save him with a page. And *it is said* that since then the above mentioned Cartagena did not say hello to him for three days'.

In (22), from Travis (2006: 1286) *dicen que* and *dizque* coexist in the same text but "while *dicen que* 'they say that' presents objectively what is put forward by theories of witchcraft ("it is said that things change their properties"), *dizque* implies that the material it introduces is a belief that is more open to question ("supposedly they take on magic properties").

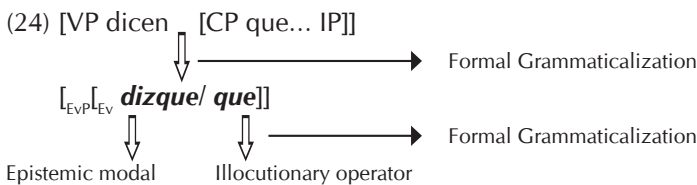
(22) En brujería dicen que todas esas cosas cambian sus propiedades por los rezos y las alumbradas que se les hacen y **dizque** se convierten en sustancias mágicas, ¿me entiendes?

'In witchcraft they say that all these things change their properties through the devotions and illuminations that are done to them and *dizque* they become magic substances, you see?' (Castro Caycedo 1994: 118, from Travis 2006, ex. (16))

The provisional syntactic analysis which correlates with the semantic bleaching and semantic ambiguity is that the *que* with an operator feature moves upwards to the matrix verb, now with no lexical content. The fusion of two heads gives rise to another functional node [Evid], situated in EvidP (above ForceP). The result is a root sentence with an activated Ev node (whose head is occupied by *dizque* or, alternatively, *que*) in its periphery. This is shown in the following diagram, based on Rizzi's (1997) "left periphery":



To summarize, the cause of grammaticalization would be the reduction of a feature in the head of ForceP (ForceP = Sub > Op → Op). This reduction involves formation of new functional material (*dizque / que*) and a categorial change. This change implies a ‘structural simplification’, which appears to be a central property of grammaticalization processes (R&R 2003:2). A structural representation is simpler (and therefore less marked) than another iff it contains “fewer formal feature syncretism” (Lash 2012: 197), where syncretism, generally speaking, is resolution of features in conflict. That is, grammaticalization of an element implies that the representation associated with it is simpler. In this framework, “simplicity” refers to feature content (including being negatively marked for a feature). This would be the case of the evolution under study, where the subordinate feature of Force is lost together with the verbal features associated to *decir*. But structural simplification also refers to the choice between (external) merge vs. move (internal merge). Merge ([X\*merge] in R&R formalization) is simpler (less marked) than move ([X\*move]), since movement is featurely more complex. We will come back to this concept in the next section. The following diagram intends to summarize the steps of the whole process and be a graphic recapitulation of the change we have analyzed:



The examples presented up to now show, though, that the distributional behavior of *dizque* and *que* is not identical. *Que* appears only in the initial position of independent sentences, *dizque* appears more frequently in subordinate structures (or taking scope over the VO when it is simply an adverb). As we have seen, *dizque* (as well as the form *que* present in Mexican Spanish analyzed by Treviño (2008) (see fn. 4)) has the properties of an evidential which is also an epistemic modal. The epistemic modal value, moreover, is the only value which is present in the adverbial form *dizque*, which can modify any sentence constituent. The standard Spanish form *que* has no features related to modality but presents the properties of an evidential which behaves as an illocutionary operator, this would explain its being restricted to root contexts.

### 6.3. Parameter resetting?

In this section we will give a rather tentative explanation of how and why parameter resetting has taken place in the evolution from *dicen que* to (*diz*)*que* in Spanish. Our basic claim is that the change to a [Evid] might be related to the value of the parameter which determines the movement of the head of Fin to the head of Force. Or, put it in more formal terms, to the (simpler) properties of the category Force, which would change from  $[C^*_{\text{move}}]$  to  $[C^*_{\text{merge}}]$  (from “attract” to “(external) merge”). Let us go step by step. First, it is usually assumed that the heads Fin and Force conflate together in languages like English (Rizzi 1997):  $C [+Fin_{\text{move (to Force)}}]$ . It is our contention that this process does not take place in Spanish ( $C [-Fin_{\text{move}}]$ ) and therefore both Force and Fin can be licensed by (independent) Merge (see the representation in (23)).

Second, as for the nature of Fin, Pesestky & Torrego (2000) show that, as a result of movement of T to C, the complementizer *that* in English is not a real complementizer but the spell out of I(nf). In Demonte & Fernández-Soriano (2005, 2009) we show that in Spanish this movement does not take place and that *que* (in Fin) is, in this sense, a real, “independent” complementizer.<sup>10</sup> Therefore, movement of Fin to Force (with the apparent “single Comp” structure) does not take place either, so there is also a

<sup>10</sup> We refer the reader to the mentioned work for further details.



possibility for the complementizer to be independently generated in Force.<sup>11</sup> A piece of evidence which supports this claim is that Spanish *que* can in certain cases materialize both in Fin *and* in Force, giving rise to structures such as (25):

- (25) Prefiero [<sub>SForce</sub> **que**, [<sub>STop</sub> ese chico, [<sub>SFin</sub> **que** [se vaya]]]]  
I.prefer that that boy that leaves

At a certain point, the parameter might have changed from Force[+move] (attraction of Fin by Force) to Force[+merge]. The trigger, as is usually the case for grammaticalization (see R&R 2003), may have been the appearance of ambiguity, in this case between (*dice* +) *que*<sub>Fin</sub> and *que*<sub>Force</sub>. This might have also paved the way for the appearance of recomplementation (as a reinforcement of Force (Villa García 2012)), as in cases like (26) (see Demonte and Fernández Soriano (2009) for details):

- (26) a. Nos regaña mucho porque dice **que** desde que somos ricos,  
desde que tenemos piscina, **que** ya no vamos al pueblo tanto...  
S/he scalds us a lot because s/he says that since we are rich,  
since we have a swimming pool that we do not go to the village  
as much as we used to
- b. ... y me dijo **que** cuando se tomaba coñac **que** creías que eras  
inmortal.  
... and s/he told me that when you drank cognac that you  
thought you were immortal (CREA Oral, España)

Another hallmark of grammaticalization as a result of parameter resetting is that it does not affect all the instances of the category which undergoes the process, only those cases that appear in the relevant contexts. This is certainly the case of Spanish *dicen que* to *dizque* and *que*, which did not affect relative *que*, for instance, nor all the cases of bisentential *decir+que* cases. So we can claim to have a genuine case of grammaticalization in R&R senses, that is, upward reanalysis of features along the clausal spine (the left periphery in our case), associated with parameter change.

<sup>11</sup> As Ledgeway (2005) shows, in certain Italian dialects movement of Fin to Force may leave (visible), intermediate copies. See Demonte & Fernández-Soriano (2013) for the contrast of these data with Spanish cases.

## REFERENCES

- Aikhenvald, A. Y. 2004. *Evidentiality*. Oxford: Oxford University Press.
- Aikhenvald, A. Y. 2006. Evidentiality in grammar. In Keith Brown (Ed.), *Encyclopedia of Language and Linguistics*, 2nd edition, vol. 4. Oxford: Elsevier, 320–325.
- Aikhenvald, A.Y. & R. M. W. Dixon. (2003) *Studies in Evidentiality*, Amsterdam: John Benjamins.
- Chafe, W. L., & J. Nichols (Eds.). 1986. *Evidentiality: the Linguistic Coding of Epistemology*. Norwood, NJ: Ablex Publishing.
- Chung, K-S. 2007. Spatial deictic tense and evidentials in Korean. *Natural Language Semantics*. 15(3): 187–219.
- Davis, C., C. Potts & P. Speas. 2008. The Pragmatic Values of Evidential Sentences. *SALT*.
- Demonte, V. & O. Fernández Soriano. 2005. Features in Comp and syntactic variation: The case of “(De)queísmo” in Spanish. *Lingua*. 115: 1063–1082.
- Demonte, V. & O. Fernández-Soriano. 2009. Force and finiteness in the Spanish complementizer system. *Probus*. 21(1): 23–49.
- Demonte, V. & O. Fernández-Soriano. 2013. El ‘que’ citativo en español y otros elementos de la periferia oracional. Variación inter e intralingüística. In D. Jacob & K. Ploog (Eds). *Autour de ‘que’ El entorno de que*, Frankfurt am Main: Peter Lang. 43 – 63.
- Demonte, V. & O. Fernández-Soriano. In press. Evidentiality and illocutionary force. Spanish matrix *que* at the syntax-pragmatics interface. In A. Dufter & A. Octavio, de Toledo (Eds.), *Left Sentence Peripheries in Spanish: Diachronic, Variationist, and Typological Perspectives*. [Linguistics Today Series] Amsterdam: John Benjamins.
- Etxepare, R. 2007. Some aspects of the quotative construction in Iberian Spanish. *Anuario del Seminario de Filología Vasca Julio de Urquijo: International Journal of Basque Linguistics and Philology*. 41(2): 25-58.
- Etxepare, R. 2010. From hearsay evidentiality to samesaying relations. *Lingua* 120(3): 604–627.
- Faller, M. 2002. *Semantics and Pragmatics of Evidentials in Cuzco Quechua*. PhD dissertation, Stanford University.
- Faller, M. 2006. Evidentiality and epistemic modality at the semantics/pragmatics interface. Talk presented at the *Workshop on Philosophy and Linguistics*, November 3–5, University of Michigan at Ann Arbor.
- Faller, M. 2007. The Cusco Quechua Reportative evidential and rhetorical relations. P. K. Austin & A. Simpson *Linguistische Berichte, Sonderheft 14*.

- Garrett, Edward. 2001. *Evidentiality and Assertion in Tibetan*. PhD dissertation, UCLA.
- Gordon, L. 1986. The development of evidentials in Maricopa. In Chafe & Nichols (Eds.), 75–88.
- Hardman, M. 1986. Data source marking in the Jaqi languages. In Chafe & Nichols (Eds.), 113–116.
- Hopper, P. & E. Traugott. 1993. *Grammaticalization*, Cambridge: Cambridge University Press.
- Izvorski, R. 1997. The present perfect as an epistemic modal. In Aaron Lawson (Ed.), *Proceedings of SALT VIII*. Ithaca, NY: Cornell University Press. 222–239.
- Jakobson, R. 1957. *Shifters, Verbal Categories and the Russian Verb*. Cambridge, MA: Harvard Univ., Dept. of Slavic Languages and Literatures.
- Kany, C. E. 1944. Impersonal *dizque* and its variants in American Spanish. *Hispanic Review* .12, 168–177.
- Klee, C. A. & Ocampo, A. M. 1995. The expression of past reference in Spanish narratives of Spanish-Quechua bilingual speakers. In Carmen Silva Corvalán (Ed.), *Spanish in Four Continents: Studies in Language Contact and Bilingualism*, Washington, DC: Georgetown University Press. 52–70.
- Laprade, R. A. 1976. Some Salient Dialectal Features of La Paz Spanish. MA thesis, University of Florida.
- Lash, E. 2012. The diachronic development of the Irish comparative particle. In Ch. Galves, S. Cyrino, R. Lopes, F. Sandalo & J. Avelar (Eds.) *Parameter theory and Linguistic change*. Oxford: OUP. 196-214.
- Ledgeway, A. 2005. Moving through the left periphery: the dual complementiser system in the dialects of southern Italy. *Transactions of the Philological Society*. 103(3): 339–396.
- Lehmann, C. 1985. Grammaticalization: synchronic variation and diachronic change, *Lingua e Stile* 20:303–318.
- Lim, D. 2010. *Evidentials and Interrogatives: A Case Study from Korean*. PhD dissertation, University of Southern California.
- Matthewson, L., H. Davis & H. Rullmann. 2007. Evidentials as epistemic modals: evidence from St’át’imcets. *Linguistic Variation Yearbook* 7: 201–254.
- Palmer, F. R. 1986. *Mood and Modality*. Cambridge: Cambridge University Press.
- Pesetsky, D. & Torrego, E. 2000. T-to-C Movement: Causes and Consequences. In M. Kenstowicz (Ed.) *Ken Hale: a Life in Language*. Cambridge: MIT Press. 355-426.
- Radford, A. 2013. The complementiser system in spoken English. Evidence from broadcast

- media. In Camacho-Taboada, V., A. L. Jiménez-Fernández, J. Martín-González & M. Reyes-Tejedor (Eds.), *Information Structure and Agreement*. Amsterdam: John Benjamins. 11-54.
- Rizzi, L. 1997. The fine structure of the left periphery. In Liliane Haegeman (Ed.), *Elements of Grammar. Handbook of Generative Syntax*. Dordrecht: Kluwer. 281-337.
- Roberts, I. & A. Roussou. 2003. *Syntactic change. A minimalist approach to grammaticalization*. Cambridge: Cambridge University Press.
- Roussou, A. 2000. On the left periphery: modal particles and complementisers. *Journal of Greek Linguistics*. 1: 65-94.
- Rullmann, H., L. Matthewson & H. Davis. 2008. Modals as distributive indefinites. *Natural Language Semantics*. 16(4): 271-295.
- Speas, M. 2004. Evidentiality, logophoricity and the syntactic representation of pragmatic features. *Lingua*. 114(3): 255-276.
- Speas, P. & C. Tenny. 2003. Configurational properties of point of view roles. In A. M. Di Sciullo (Ed.), *Asymmetry in Grammar*. Volume 1: *Syntax and Semantics* [Linguistik Aktuell/ Linguistics Today 57]. Amsterdam: John Benjamins. 315-344.
- Travis, C. E. 2006. *Dizque*: a Colombian evidentiality strategy. *Linguistics*. 44(6): 1269-1297.
- Treviño, E. 2008. *Que*, una marca de evidencialidad en español, paper presented at X *Encuentro Internacional de Lingüística en el Noroeste*, UNISON, Hermosillo.
- Villa-García, J. 2012. Re complementation and locality of movement in Spanish. *Probus*. 24(2): 257-314.
- Willett, T. 1988. A cross-linguistic survey of the grammaticization of evidentiality. *Studies in Language*. 12(1): 51-97.
- Wilson, D. 2006. The pragmatics of verbal irony: Echo or pretence? *Lingua*. 116(10): 1722-1743.

## Óscar Lopes: recordar o linguista, o mestre e o amigo

Fátima Oliveira  
foliveir@netcabo.pt

*Faculdade de Letras da Universidade do Porto  
Centro de Linguística da Universidade do Porto*

A 22 de março, dia da Universidade do Porto, Óscar Lopes deixou-nos, mas a sua investigação e a memória do seu extraordinário saber e da sua enorme sensibilidade continuarão sempre a acompanhar-nos.

Muitos conhecem as suas obras, muitos lhe podem chamar mestre e alguns tiveram o privilégio de o conhecer também e chamar amigo, mas lembrar Óscar Lopes é correr um enorme risco pois dificilmente seremos fiéis ao essencial.

Por isso me ocorrem estas suas palavras: “Não confio em qualquer título de autorreconhecimento, porque tanto as nossas imagens a um espelho polido como as nossas imagens que os olhos alheios devolvem estão, não apenas erradas na sua simetria axial, mas medusadas pelo reflexo inverso do nosso próprio olhar que fita, e que fixa, essas nossas imagens”. (Lopes, 1986:19-20)

Óscar Lopes é muito mais conhecido pelos seus trabalhos de crítica, ensaio e história literários, mas a sua extraordinária faceta de linguista, embora menos conhecida ou até acessível, está perfeitamente de acordo com a sua personalidade. Foi, por certo, uma forma de resistir à proibição de ensinar Literatura a partir de 1953, mas a Linguística foi também para Óscar Lopes uma das formas de tentar encontrar respostas às muitas perguntas que fazia sobre o homem e sobre a vida. A busca de sentido e o sentido que a vida faz levaram-no desde muito cedo a interessar-se por Semântica, em particular Semântica formal, e por uma certa Pragmática.

É certo que escreveu: “Não faço linguística: trata-se apenas de, com a mais rigorosa metodologia disponível, refletir sobre certos gestos do nosso mais espontâneo modo de falar, gestos que têm que ver com relações especiais de

tempo, de atitude e de referência na comunicação social possível” (Lopes, 1986:20). No entanto, a sua investigação marcou o início em Portugal dos estudos formais em Linguística e em particular da Semântica e influenciou uma geração de linguistas.

Para Óscar Lopes, a Linguística está sempre presente de uma certa forma porque a língua, materialidade da comunicação e do pensamento nos envolve e nos constitui. E por isso nos diz: “Não é apenas apaixonante, em si mesmo, tentar apreender a álgebra da comunicação verbal, sobretudo daquelas palavras, ou operadores, graças aos quais todas as mais diversas maneiras de pensar e agir, mesmo as que são cientificamente mais sofisticadas, se conjugam com os dados mais radicais da comunicação (o *eu* e o *tu*, o *agora* ou não, o *aqui* ou não, o *assim* ou de outro modo), – mas suponho ainda que, para a apreensão mais fina da poesia, não será indiferente o facto de, por exemplo, a palavra *isto*, nos seus mais óbvios usos em português, exigir que o contexto desambigue entre os 76 principais tipos diferentes de localização espacial que comporta” (Lopes, 1986: 21).

A *Gramática Simbólica do Português – um esboço* representa o culminar de uma investigação extraordinária longamente amadurecida e testada com alunos do então Ciclo Preparatório, cujas perspetivas teóricas o interessavam desde pelo menos os anos 50 e que estão patentes em artigos de 1958 como “Lógica Gramatical e Lógica Simbólica”, ou “Gramática escolar - reduto de um anacronismo”. E numa comunicação ao VI Congresso do Ensino Liceal, em 1971, “Perspetivas de uma Gramática Simbólica”, o Professor Óscar Lopes diz o seguinte: “O cálculo proposicional bivalente elementar supre, com enorme vantagem gramatical e estilística, a taxinomia usual das orações, conjunções e complementos circunstanciais delas sintaticamente deriváveis, substituindo a didática classificatória, fechada, por uma didática de análise aberta a todos os matizes contextuais e a todas as interpretações alternativas, o que fomenta a autonomia judicativa e a criatividade do educando, munindo-os ao mesmo tempo de uma disciplina interna”.

Esta citação, entre muitas outras possíveis, serve para ilustrar a sua conceção extremamente inovadora de Gramática, associada a instrumentos de Matemática e de Lógica. No entanto, a preocupação com a formalização em Linguística surge só no século XX, devendo contudo assinalar-se que é só a partir de meados dos anos 50 que surgem os primeiros textos fundadores

e é principalmente a partir dos anos 60 que se assiste a um impulso da Semântica Formal, sobretudo com os trabalhos de Montague. Mas estes textos criadores de novas perspetivas sobre a análise linguística não são então conhecidos em Portugal, quer pelo isolamento em que vivíamos, quer porque alguns deles só mais tarde foram publicados.

Na *Gramática Simbólica*, utilizando instrumentos teóricos do cálculo proposicional, da lógica de predicados, da teoria dos conjuntos, de topologia, de relações matemáticas e lógicas de vários tipos e suas propriedades, Óscar Lopes trabalha um número considerável de questões linguísticas que vão desde a distinção semântica entre nomes próprios e comuns, questões de quantificação, definidos e indefinidos, demonstrativos, «coordenação» e «subordinação» de vários tipos, até questões de tempo, aspeto e modo. É, aliás, interessante notar, a título exemplificativo, como Óscar Lopes, já em 1971, recorre, para o estudo do tempo, à noção de intervalo como conjunto de pontos unilinearmente ordenados, quando o texto clássico e fundador desta abordagem em semântica linguística formal, escrito por Bennet e Partee, data de 1972, mas só é publicado em 1978.

Como em qualquer domínio científico, a linguística mudou muito e as teorias semânticas desenvolveram-se bastante nos últimos trinta anos, o que o entusiasma bastante estudar, mas este trabalho excepcional do Professor Óscar Lopes, depois de mais de 40 anos, ainda causa espanto e admiração, é fonte de inspiração e dá um enorme prazer ler.

Depois desta obra fundamental, Óscar Lopes continuou a desenvolver investigação em Linguística, abarcando assuntos desde a semântica dos nomes massivos, usando a difícil formalização de sistemas de lógica intensional, fez estudos sobre contrastivas, condicionais, concessivas e causais, passando por estudos sobre o Presente e a dêixis, e interessando-se também por partículas discursivas como *pois*. Alguns destes textos foram reunidos e publicados em 2005.

A necessidade de ultrapassar a superfície deste contínuo sonoro que é a língua e a consciência da sua complexidade levam-no a tratar com igual rigor e atenção frases como '*Os homens trazem a mobília*', '*O Zé chega sempre às sete*', '*Isto é maravilhoso*', '*O João sai, ainda que chova*' ou

textos literários.<sup>1</sup>

Foi também um extraordinário professor de Linguística quando, depois de o 25 de Abril lhe ter tornado possível ensinar na Universidade, iniciou o curso de Linguística Matemática, o que ao tempo foi também totalmente inovador. Mas também lecionou outras disciplinas como Semântica, História da Língua, Introdução aos Estudos Linguísticos ou Pragmática tanto a estudantes de licenciatura como de mestrado.

A sua capacidade de relação com assuntos tratados por outras áreas em Linguística e com outros domínios científicos, como, por exemplo, a Biologia, a Física e a Astrofísica, ou ainda com a Filosofia, a Antropologia ou a História fez as delícias de todos os que com ele privaram em aulas, conversas ou em leituras.

De facto, de modo direto ou indireto influenciou muitos linguistas e pessoas interessadas no estudo de questões linguísticas, que procuram em diferentes modos de formalização instrumentos adequados que tornem possível encontrar algumas respostas que vão para além de um descritivismo pré-científico. Tentou também ensinar-nos que as possíveis respostas constituem uma plataforma para novas perguntas e que a necessária redução do campo de análise na investigação, em qualquer domínio científico, não deve fazer-nos esquecer os dados nem a própria reflexão crítica. Por isso defendeu a formalização como meio de tornar mais claras e passíveis de infirmação a descrição e a explicação em Linguística.

Porém, o seu desejo de captar a ordem na desordem aparente, de a enfrentar e a dominar através de teorias altamente formalizadas não o impediram de escrever “Não há ciência, não há prática, não há política, não há nada de rigoroso ou eficaz sem fantasia” (Lopes, 1983: XX).

As palavras, essa espécie de gestos falados, sempre o intrigaram e fascinaram porque de uma certa forma estava sempre à procura de uma resposta e, de preferência, diferente da sua, pois, para ele, era nesse diálogo que se cria um olhar novo. Um diálogo que sempre desejou e ao mesmo tempo o perturbava pela presença humana que o fazia sentir que “há um recado pessoal que não sei dar ou receber” (Lopes, 1986: 24).

---

<sup>1</sup> Veja-se o que Óscar Lopes diz na entrega do Prémio Vida Literária: “Raramente escrevo no embalo de duas ou três ideias fáceis, e nunca pratico a improvisação” in *Óscar Lopes – um homem maior do que o seu tempo, na comemoração dos seus noventa anos*, p. 191.



Por isso, o seu monólogo mais solitário sempre foi, de uma certa forma, um diálogo, uma polémica. Soube afinal de contas colocar-se na encruzilhada correta, soube ouvir o som, a princípio ténue, do novo. Soube olhar para longe, explorar o futuro e não deixou de sonhar.

Por tudo isto e por tudo quanto fica por dizer, estou convicta de que o Professor Óscar Lopes teria tido, no domínio da Linguística, uma projeção internacional muito maior, se tivesse desenvolvido a sua investigação num país e num tempo diferentes, nas várias leituras semânticas da palavra *diferente*.

Mas citando Óscar Lopes, que parafraseava Pessoa, *ser diferente é ter a oportunidade de ser superior*. Por isso, é urgente ler ou reler os seus textos, ouvir a sua voz e refletir, criando, se possível, novos olhares e formulando novas perguntas pois é isso que espera de nós.

#### REFERÊNCIAS

- Bennet, Michael & Barbara Partee. 1972. *Toward the Logic of Tense and Aspect in English*. Report, System Development Corporation, Santa Monica. Bloomington: Indiana University Club, 1978.
- Lopes, Óscar. 1971. *Gramática Simbólica do Português (um esboço)*. Lisboa: Fundação Calouste Gulbenkian – Instituto Gulbenkian de Ciência, 2ª edição corrigida, 1972.
- Lopes, Óscar. 1983. Como Eu Gosto deste Livro. Prefácio ao livro de Haroldo Maranhão *A Porta Mágica*, Coimbra: Vértice.
- Lopes, Óscar. 1986. Palavras de Auto-Explicação, proferidas na sessão de entrega do Prémio da Crítica da Associação Portuguesa de Escritores em 1984, publicado em *Uma Arte da Música e outros Ensaios*, Porto: Oficina Musical, 19-25.
- Lopes, Óscar. 1996. Texto lido na entrega do prémio vida literária. In *Óscar Lopes – um homem maior do que o seu tempo, na comemoração dos seus noventa anos, 2007*, Câmara Municipal de Matosinhos, 187-193.
- Lopes, Óscar. 2005. *Entre a Palavra e o Discurso*. Estudos de Linguística 1977-1993. Coordenação da edição de Oliveira F. e A. M. Brito. Porto: Campo das Letras.



Recensões



Rui CARMO (org.) *Linguagem, Argumentação e  
Decisão Judiciária*. Coimbra: Coimbra Editora. 2012,  
180 pp.  
ISBN 978-972-32-2123-7

Belinda Maia  
bhsmaia@gmail.com  
*Faculdade de Letras da Universidade do Porto (Portugal)*  
*Centro de Linguística da Universidade do Porto*

O livro em apreço resulta de um seminário no âmbito do programa *Justiça XXI*, da responsabilidade da Associação Sindical dos Juizes Portugueses, do Centro de Estudos Sociais da Universidade de Coimbra, do Observatório Permanente da Justiça e do Sindicato dos Magistrados do Ministério Público.

Este livro é um contributo valioso para a discussão da importância de uma comunicação mais estreita entre aqueles que estudam o Direito e a Linguística e a ligação que existe entre estas áreas. O volume tem também contributos de outras áreas de saber, como as ciências exatas, e a filosofia da linguagem.

Na minha recensão, vou comentar a parte mais ligada ao estudo da linguagem ou à Linguística, dado o meu interesse, não só pelo estudo da linguagem como objecto de fascínio, mas, e essencialmente, pelas funções que a linguagem tem na sociedade humana.

A colecção de artigos começa com uma reflexão por Carlos Fiolhais sobre 'As Leis Naturais e as Leis Humanas', que nos oferece uma pequena viagem pela história da ciência e a formulação das leis que regem a mundo físico à nossa volta, e que servem de base obrigatória para a investigação científica – até chegar uma nova lei que ou constrói sobre as leis existentes, ou muda radicalmente a perspectiva da ciência. Estas leis podem ser expressas em fórmulas como a de Einstein –  $E=mc^2$ , e há aqui a ligação com os 'Positivistas Lógicos' do círculo de Viena, representado por Wittgenstein. Mas Wittgenstein, depois de procurar este ideal, chegou à conclusão de

que a linguagem não poderia ser reduzida a fórmulas. Por muito que se tente, o Direito também não está isento de influência da linguagem como fenómeno social; o autor acaba com uma reflexão sobre o modo como as novas descobertas da própria Ciência oferecem desafios ao Direito e à Ética.

O segundo artigo, de Manuel Simas Santos, começa por contrastar os ‘factos científicos’ com ‘os factos de um processo judiciário’ e apresenta uma reflexão sobre como os ‘factos’ estão sempre sujeitos a alguma subjectividade, embora inconsciente, e conclui que “as reacções *intuitivas, emocionais e sentimentais* do julgador face ao caso concreto” desempenham um papel importante num julgamento.

O dilema do juiz é o tema do terceiro artigo, de Conceição Carapinha, que apresenta reflexões sobre o discurso judiciário no tribunal, do ponto da vista do leigo, que dificilmente compreende o rigor do discurso, tal como é encarado pelos representantes da Lei. Utilizando as perspetivas da Linguística que focam o ‘contexto’, a ‘história’ e até a linguagem não verbal do sujeito, apela para a aparente falta de interesse por parte da Lei em relação ao contexto do réu ou da testemunha.

Os dois artigos seguintes discutem o problema da compreensibilidade do discurso jurídico, o movimento da ‘Linguagem Clara’, derivado do ‘Plain English’, e a sua relação com a democracia e a transparência do discurso jurídico para o público em geral. Como muito bem observam, o assunto não é de maneira nenhuma ‘claro’. Aliás, o problema debate-se em várias áreas onde o discurso entre pares a alto nível é incompreensível para o público em geral. Mas na era de informação para todos a preocupação com a inteligibilidade, ou não, desta informação para o público em geral tornou-se uma questão candente. Para um técnico, a explicação de uma máquina para o entendimento do público tornou-se uma questão de segurança (eventualmente a ser julgado em tribunal, caso instruções deficientes conduzam à morte ou provoquem danos do comprador). Para o Direito, a compreensão dos meandros da Lei tornou-se uma questão de democracia.

Rui do Carmo argumenta que a compreensibilidade na linguagem jurídica “é indissociável de outras três características essenciais” – a concisão, a segurança e o rigor jurídico.

Maria Manuel Leitão Marques pergunta com razão se ‘Simplificar’ é ‘Uma missão possível’, mas continua o artigo elaborando sobre o que é

entendido como ‘Linguagem Clara’, e, referindo os passos a serem adotados para este fim, demonstra que certos textos podem e devem manter-se para a informação entre pares, mas que é bem possível adaptar textos desnecessariamente elaborados para a informação do público em geral.

Giovanni Damele apresenta um artigo sobre ‘Verdade e comunicação – Notas sobre argumentação e decisão Judiciária’, que há de agradar a quem gosta destas questões vistas do ponto de vista filosófico.

Joana Aguiar e Silva continua esta reflexão sobre a ‘verdade’ com ‘As Narrativas do Direito e a Verdade judicial’, num artigo focando as ‘chamadas teorias narrativistas do direito’. Explora como a ‘história’ de um caso evolui, perguntando ‘O que é a verdade jurídica senão fruto de convenções jurídicas?’ e afirmando que ‘o Direito cria as suas próprias verdades que, nessa medida, são sempre relativas’.

Os últimos artigos focam a forma como se constroem os textos judiciais. No primeiro, ‘Decisão Judiciária: processo de elaboração e fundamentação’, José Augusto Ferreira da Silva continua a reflexão sobre o efeito da democracia em relação à transparência de informação; neste caso a obrigação, por parte dos juizes, de fundamentar as suas decisões claramente e com base em cada caso apresentado. Documenta um processo muito interessante: a luta pelos advogados para obrigar os juizes a fundamentar as suas decisões de uma maneira ‘expressa, clara, coerente e suficiente’ para ‘convencer o destinatário das suas razões’.

O artigo ‘A Decisão Judicial: O processo de Fundamentação’ de José Mouraz Lopes, mostra os critérios que devem reger o processo de criação da fundamentação da decisão judicial: “generalidade, indisponibilidade, completude, publicidade e duplo grau de jurisdição”. Defende a ‘suficiência da fundamentação’, ‘uma fundamentação coerente’, e ‘uma fundamentação razoável’. Acaba reflectindo sobre a dificuldade de tornar o discurso ‘cerrado, opaco, excessivo e genericamente pouco compreensível’ acessível ao público em geral, convidando os seus colegas a ‘cultivar a capacidade de síntese’ e ‘um estilo de linguagem escurreita, clara, sintética, mas simultaneamente rigorosa’.

No último artigo, ‘Apontamentos sobre a elaboração e fundamentação de algumas decisões judiciais do Ministério Público em Processo Penal’, Vítor Paiva debruça-se sobre ‘a elaboração e fundamentação das decisões de

não instauração de inquérito, arquivamento e acusação', apelando – como outros nesta colecção – a uma escrita objetiva, sintética e clara. Oferece vários conselhos a quem escreve estes documentos de modo a apresentar os factos de uma maneira objectiva. Importante também para a Linguística Forense é a sua referência à necessidade de haver uma documentação clara e objectiva desde o começo de todo o processo, começando com a que cabe aos órgãos da polícia criminal.

O foco de interesse neste volume é essencialmente monolíngue e relacionado com a situação portuguesa, num contexto judicial. Porém, existem diversas outras áreas onde a Linguística pode dar um importante contributo em contextos legais. Esta área, baseada em abordagens que levam em consideração a interação entre a linguagem e o Direito, e que se designa genericamente por Linguística Forense, permitirá, porventura, incentivar um trabalho conjunto entre os profissionais das ciências jurídicas e das ciências da linguagem com vista a complementar as competências daqueles que se relacionam quotidianamente com o processo judicial. Para além do contexto judicial, em justo destaque neste volume, são ainda áreas da Linguística Forense o apoio à investigação policial, não só em termos investigativos, mas também processuais, e a investigação em análise de autoria. A primeira consiste em apoiar os investigadores policiais, no sentido de analisar os procedimentos policiais de natureza linguística, e recomendar boas práticas que sejam mais transparentes, quer para as pessoas envolvidas na investigação, quer para os próprios investigadores. A segunda consiste em contribuir, através da especialização em Ciências da Linguagem, para análises aprofundadas de casos nos quais esteja em causa a verdadeira autoria dos textos, sejam eles escritos ou orais.

Há, para além destas vertentes, uma gama de problemas que se levantam e que permitem encontrar pontos de ligação com outros sistemas de Direito e as culturas com eles relacionadas. Alguns destes problemas são discutidos num relatório da Comissão Europeia – *Language and Translation in International Law and EU Law (2012)* - mas o debate sobre a tradução de textos jurídicos tem uma longa história. O facto de os sistemas, as leis e os processos legais nos diferentes países de Europa terem origens, histórias e culturas diferentes, reflete-se na terminologia e fraseologia utilizadas nos documentos usados. Estas diferenças dificultam a intercompreensão



em casos em que mais que uma língua e cultura são alvos de processos judiciais. Gostaria de apelar para a necessidade de estudar os problemas que se levantam para o sistema judicial quando confrontado com as políticas europeias de multilinguismo e multiculturalismo. A Diretiva 2010/64/EU sobre o Direito à Tradução e Interpretação no processo judicial nos tribunais entra em vigor em 2014 e implica a formação de pessoas devidamente preparadas. Outra questão, relacionada com esta, é a questão dos direitos linguísticos que afectam a vida de cada um de nós, enquanto cidadãos.

Estas são áreas que podem ser desenvolvidos com a colaboração entre a Linguística e o Direito, sobretudo em termos de investigação.



Fernão de OLIVEIRA. *Gramática da Linguagem Portuguesa, Fac-simile, introdução e edição actualizada e anotada por: José Eduardo Franco e João Paulo Silvestre*, Lisboa: Fundação Calouste Gulbenkian. 2012. 247 pgs. ISBN 978-972-31-1447-8 (edição em capa dura)

Sónia Duarte

duarte.sonia@sapo.pt

*Centro de Linguística da Universidade do Porto (Portugal)*

Como evidencia Rogelio Ponce de León Romeo (2010: 553-554) num artigo dedicado à atividade de crítica textual no âmbito gramaticográfico, apesar das insuficiências, “tem havido um aumento da publicação - em forma de edições críticas e facsimiladas - de textos metagramaticais”. Entre estas, o mesmo autor (Ponce de León 2010: 554) destaca, no quadro peninsular e nos vinte e cinco anos precedentes, o trabalho crítico-textual sobre os textos de Antonio de Nebrija e de Fernão de Oliveira. Curiosamente, reforçando esta apreciação, nos últimos dois anos, saiu à luz uma edição crítica da *Gramática Castellana* (Madrid 1492) de Nebrija e uma edição atualizada e anotada da *Grammatica da lingoagem portuguesa* (Lisboa 1536) de Oliveira. A primeira é da responsabilidade de Carmen Lozano (2011); a segunda, limitada a 750 exemplares, é o objeto desta recensão.

Os editores desta nova edição do texto quinhentista dedicam-se a diferentes áreas de investigação: José Eduardo Franco é especialista em história da cultura e João Paulo Silvestre em historiografia linguística, lexicologia e edição de corpora e ambos publicaram já sobre Oliveira (Franco 2009; Silvestre 2009) num mesmo volume coletivo. Neste trabalho que agora empreendem em comum, os referidos investigadores dão sentido à articulação das suas diferentes áreas de especialidade sublinhando a abrangência do valor documental do texto, tanto relativamente à história da língua como relativamente à história da cultura e das ideias, descrevendo a presente edição como “proposta de uma leitura sincrética, que respeita

esse amplo valor testemunhal” (p. 8). Nesse sentido, como se explica na introdução, João Paulo Silvestre encarregou-se da fixação do texto e dos índices lexicais e José Eduardo Franco assumiu a responsabilidade dos comentários históricos e culturais.

A leitura que aqui se apresenta foi ainda recentemente precedida de outras duas sobre o mesmo texto: a edição crítica de Torres & Assunção (2000), reeditada em 2007, e a edição diplomática de Paiva (2002), comentadas ambas por José Eduardo Franco e João Paulo Silvestre, aquando da explicitação dos critérios da sua própria edição (39-43), e contabilizadas na listagem de edições impressas do texto, incluída entre a bibliografia final (p. 232). Relativamente às anteriores edições modernas de Oliveira e no tocante especificamente à fixação do texto, como se expõe na introdução, o principal objetivo dos editores, é o de “proporcionar uma leitura actualizada que observa e mantém a informação metalinguística, e que oferece o confronto com o fac-simile da impressão de 1536”<sup>1</sup> (p. 8). Com efeito, é considerando estes aspetos que a edição aqui visada se posiciona em especial relativamente à única leitura atualizada que os precede (Buescu 1975), na qual, segundo os editores do volume recenseado, a atualização dos exemplos sem indicação da forma original deu lugar a “incoerências no enunciado gramatical” (p. 40). Os editores salientam ainda divergências, relativamente à supracitada edição de Buescu, quanto à não manutenção de determinadas formas arcaizantes com continuidade no registo literário. Como se deduz da passagem acima transcrita, afigura-se-lhes ainda como uma limitação o facto de a edição de 1975 não ir acompanhada de fac-simile, o qual, por seu turno, foi editado pela primeira vez em 1981, por iniciativa da Biblioteca Nacional. Da mesma forma, relativamente à edição crítica de Torres & Assunção (2000), que os editores assumem como leitura de referência (p. 43), a edição de 2012 propõe-se igualmente reconsiderar a uniformização de certas opções, assegurando que a resolução de variantes se submete a critérios específicos que não contrariem a teoria metalinguística exposta no texto de Oliveira.

No que concerne à estrutura, esta está organizada em três partes: uma que corresponde ao texto introdutório (pp. 7- 43), outra dedicada ao texto

---

<sup>1</sup> Nas citações, foi respeitada a opção ortográfica dos autores, que não seguem o acordo ortográfico de 1990.

metagramatical (pp. 45-195) e outra ainda, aos índices e bibliografia (pp. 197-247). Na introdução, após algumas breves considerações sobre o carácter multifacetado da vida e obra de Oliveira, sobre a amplitude do valor da sua produção metagramatical e sobre os propósitos e especificidades desta edição, o discurso incide sobre determinadas dimensões da obra, que se articulam em torno dos seguintes eixos: i) a língua e o texto metalinguístico no quadro de um projeto político nacional de afirmação e expansão; ii) a exposição da teoria gramatical de Oliveira e o seu enquadramento na tradição metalinguística portuguesa<sup>2</sup>. Assim, dentro do primeiro ponto, evidencia-se em que medida o texto de Oliveira é um exemplo original de afirmação linguística no quadro teórico da questão da língua em Portugal<sup>3</sup>, mas também a confirmação de que o processo de gramatização tem implicações extralinguísticas<sup>4</sup> que põem em relevo a pertinência da aliança entre história e filologia, assumida pelos editores. No âmbito do segundo ponto, destacam-se os seguintes aspetos: o papel da observação enquanto marca do espírito humanista visível no tratamento dos fenómenos linguísticos por Oliveira; a reivindicação do vernáculo e a legitimação da sua singularidade; a consideração ou não da obra dentro do paradigma textual de *gramática*<sup>5</sup>; a heterodoxia das estratégias de construção da autoridade subjacente às propostas gramaticais, nomeadamente a opção pelo discurso na primeira pessoa e a valorização de dados empíricos; a opção por uma terminologia

---

<sup>2</sup> Como advertem os próprios autores, esta parte da introdução corresponde a uma significativa ampliação do trabalho anteriormente citado de Silvestre (2009).

<sup>3</sup> Utilizo este conceito tal como descrito por Buescu, no seguimento de Picchio (1959), no âmbito da tensão bipolarizada, por um lado, face ao latim e, por outro, face ao castelhano:

“Ora o binómio português/castelhano, aparentemente adversativo do binómio latim/português traz, afinal, a neutralização deste, na medida em que a posição em relação ao castelhano releva duma *praxis*: apresenta a iminência de um risco que os humanistas pressentem. O do predomínio de uma língua competitiva, forma de expressão de uma nação de algum modo rival –em termos objetivos mais poderosa – no plano político interno e também no plano duma política expansionista e imperial.

Assim a reaproximação com o Latim representa o estreitamento de um vínculo que, sendo tutelar, é também libertador” (Buescu 1983: 225).

<sup>4</sup> Uma referência incontornável sobre essas implicações no âmbito específico da obra de Oliveira é o conhecido trabalho de Eugenio Asensio (1974[1960]: 11-13), onde se aproxima a dimensão imperial do texto de Oliveira à sustentada por Nebrija.

<sup>5</sup> Os editores parecem decantar-se pela segunda opção, não o descrevendo como tal, mas caracterizando-o alternativamente como “primeiro ensaio de uma codificação da língua portuguesa” (p. 7), “breve tratado” (p. 7), ou, apropriando-se das palavras de Oliveira, como “primeira anotação” (p. 24). Na referência ao debate em torno desta questão, os autores (p. 10. n. 10) citam Buescu (1984[1983]: 17), que relativiza a aplicação à obra de Oliveira da designação de *gramática* “o título de gramática contido no rosto, corresponde efectivamente, a um título genérico, talvez imposto pelo editor, prosseguindo na tradição medieval de chamar *gramática* a todas as obras de conteúdo linguístico”. Mais recentemente, esta posição foi ainda reforçada por Schäfer-Prieß (no prelo: 1.1.6).

diferenciada da utilizada na tradição gramatical de descrição do latim; o reduzido impacto do texto na tradição gramaticográfica subsequente.

Após a introdução, na parte dedicada ao texto metagramatical, há lugar à explicitação dos critérios editoriais, durante a qual se tecem ainda algumas considerações sobre o panorama precedente na edição da obra de Oliveira, bem como sobre os recursos lexicográficos utilizados no esclarecimento de dúvidas pontuais e que, como explicitam os editores, correspondem “preferencialmente ao testemunho da tradição lexicográfica antiga<sup>6</sup> (p. 43). Seguidamente, a leitura e fac-símile do texto de Oliveira surgem em páginas alternadas, de forma a reproduzir a distribuição de páginas original, encontrando-se numeradas de forma sequencial a partir da folha de rosto. Nas notas de rodapé, para além do esclarecimento de léxico e de referências histórico-culturais – como, de resto ocorria já, nas notas finais, na edição de Buescu (1975: 127-137) – os editores investiram ainda, coerentemente com os propósitos definidos, em deixar transparecer, quer o texto original, quer os desvios à já referida leitura crítica de Torres & Assunção.

No que diz respeito aos índices e bibliografia, constitui um contributo a destacar a integração de uma ferramenta de grande utilidade como é um índice de exemplos (pp. 197-205) e um índice metalinguístico (pp. 207-229), assim como a atualização dos dados bibliográficos desde a última edição do texto (Torres & Assunção 2007).

Como comentário final, cabe sublinhar a coerência do trabalho de edição com os objetivos propostos e o preenchimento oportuno de um espaço no contexto do trabalho editorial em torno da obra metagramatical de Oliveira, e que, como salientam os editores na introdução (p. 9), seria útil que fosse ampliado com a edição das obras completas de Fernão de Oliveira.

#### REFERÊNCIAS

Asensio, E. 1960. La lengua compañera del imperio. Historia de una idea de Nebrija en España y Portugal. *Revista de Filología Española*. XLIII: 399-413.

---

<sup>6</sup> Esta aparece identificada pelos editores com as obras de Bento Pereira, Rafael Bluteau e Jerónimo Cardoso.

- Buescu, M. L. C. (Ed.). 1975. *A Gramática da Linguagem Portuguesa*. Introdução, leitura atualizada e notas. Lisboa: Imprensa Nacional – Casa da Moeda.
- Buescu, M. L. C. 1983. *Babel ou a ruptura do signo. A gramática e os gramáticos portugueses do Século XVI*. Lisboa: Imprensa Nacional - Casa da Moeda.
- Franco, J. E. 2009. Fernando Oliveira, ideógrafo da proto-nacionalidade portuguesa. In: C. Morais (Coord.). *Fernando de Oliveira. Um humanista genial*. Aveiro: Universidade de Aveiro – Centro de Línguas e Culturas, 547-580.
- Nebrija, A. de. 2011[1492]. *Gramática sobre la lengua castellana*. Edición estudio y notas de Carmen Lozano. Barcelona: Real Academia Española, Círculo de Lectores, Galaxia Gutenberg.
- Oliveira, F. de. 1981. *Gramática da Linguagem Portuguesa*. Edição fac-similada. Lisboa: Biblioteca Nacional (reeditada em 1988).
- Paiva, M. H. 2002. *Os gramáticos portugueses quincentistas e a fixação do padrão linguístico: contribuição da Informática para o estudo das relações entre funcionamento, variação e mudança*. Porto. Tese de doutoramento em Linguística Portuguesa apresentada à Faculdade de Letras da Universidade do Porto.
- Ponce de León, R. 2010. "Ars grammatica" e "ars critica": breves considerações sobre a crítica textual aplicada aos textos metagramaticais no âmbito da Península Ibérica. In F. Topa & M. J. Reynaud (Eds.) *Crítica Textual & Crítica Genética em Diálogo; Colóquio Internacional: Porto, 18-20 de Outubro de 2007: actas – I*, 553-563.
- Schäfer-Prieß, B. No prelo. *A Gramaticografia Portuguesa de 1540 até 1822: Condições da sua génese e critérios de categorização, no âmbito da tradição latina, espanhola e francesa*. Traducción de Jaime Ferreira da Silva, revisada y actualizada por la autora.
- Silvestre, J. P. 2009. A autoridade heterodoxa da *Grammatica* de Fernão de Oliveira. In C. Morais (Coord.). *Fernando de Oliveira. Um humanista genial*. Aveiro: Universidade de Aveiro – Centro de Línguas e Culturas, 233-241.
- Stegagno-Picchio, L. (Ed.). 1959. "La questione della lingua in Portogallo." In: João de Barros. *Diálogo em louvor da nossa Linguagem*. Modena: Soc. Tipográfica Modonese.
- Torres, A. & Assunção, C. (Eds.). 2000. *Gramática da Linguagem Portuguesa (1536) / Fernão de Oliveira*. Edição crítica, semidiplomática e anastática. Lisboa: Academia das Ciências de Lisboa (reeditada em 2007).





Maria Victoria Navas SÁNCHEZ-ÉLEZ. *El barranquenho. Un modelo de lenguas en contacto*.  
Madrid: Editorial Complutense. 2011. 319 pp,  
ISBN: 978-84-9938-099-5

Clara Araújo Barros  
mbarros@letras.up.pt  
*Faculdade de Letras da Universidade do Porto (Portugal)*  
*Centro de Linguística da Universidade do Porto*

Na origem deste trabalho sobre o dialeto barranquenho está a investigação desenvolvida no âmbito do *Projeto Diacronia e Sincronia: linguagens fronteiriças* da responsabilidade da Autora que o criou e coordenou no Centro de Linguística da Universidade de Lisboa.

A obra em análise apresenta, de forma exaustiva e sistemática, os resultados da investigação elaborada durante um longo período de recolha e de interpretação de dados, aprofundando um tema sobre o qual existiam anteriormente muito poucos estudos. O livro é constituído por uma introdução, cinco capítulos e seis anexos.

Na *Introdução* (pp.11-19), a A. expõe os objetivos do trabalho, a sua relevância, os principais aspetos da metodologia utilizada e esclarece as condições que originam a sua investigação, que surge já em 1987, por sugestão do professor Lindley Cintra. Os materiais para a pesquisa foram recolhidos durante a estadia da autora na região e na comunidade de Barrancos, de 1988 a 1990, seguida de diversas estadias que foram mais breves. Foram ainda utilizadas entrevistas efetuadas pelo professor Lindley Cintra em 1965. São abordadas questões metodológicas e aspetos teóricos concretos que definem um caminho original de investigação sobre esta *língua* minoritária. A metodologia utilizada combina o método antropológico e a sociolinguística variacional iniciada por William Labov, cruzando dados da dialetologia, da sociolinguística e da estatística matemática. As entrevistas, que analisam diferentes situações de comunicação e contemplam trinta

informadores, têm como ponto de partida um questionário desenhado especificamente para este estudo, que está descrito no Anexo I.

O barranquenho, situado entre a variedade portuguesa alentejana e as variedades espanholas estremenas e andaluzes é um dialeto meridional peninsular que revela afinidades com os dialetos portugueses centro-meridionais do interior, mas apresenta traços particularmente arcaizantes e marcas de uma forte influência do espanhol. Tais características explicam-se pela especificidade da sua localização e da sua história – marcadas pelo isolamento no contexto português e pela continuidade de povoamento espanhol. Este trabalho parte da análise das realizações da sibilante implosiva /-s/ – traço selecionado como específico, mas trata outros traços fonéticos, aspetos morfossintáticos e lexicais. Recolhe também e analisa textos da literatura oral / tradicional. Nas páginas finais da introdução figuram uma lista dos informadores, de ambos os sexos, agrupados em quatro gerações e em três grupos construídos com base no nível da sua formação escolar e uma lista dos símbolos utilizados.

No capítulo I – *El espácio geográfico y su historia* (pp. 21-47) – são abordados três aspetos fundamentais: a continuidade do espaço geográfico e da atividade económica nesta área meridional da península, o isolamento da região de Barrancos no contexto português e as diversas vagas de povoamento espanhol nessa região.

A A. começa por descrever as condições históricas e geográficas que favoreceram o aparecimento da fala barranquenha. Salienta a ausência de diversidade do espaço físico entre o Alentejo e as províncias de Huelva e Badajoz, regiões que delimitam a área de Barrancos. A fronteira política foi estabelecida nos tratados de Badajoz (1267) e de Alcañices (1297), mas a continuidade das comunidades humanas parece ter prosseguido imperturbada, sem uma fratura, desde o período medieval.

Houve sempre em Barrancos uma grande tradição de povoamento espanhol de estremenos ou andaluzes. Em finais do século XVII (1674-1704) 20% da população é espanhola. Dados de 1877-1894 revelam que um terço da população de Barrancos é espanhola e, dos restantes dois terços, 80% é de ascendência espanhola. No período das invasões francesas, famílias ricas de agricultores e comerciantes andaluzes instalaram-se e permaneceram na região e durante a guerra civil espanhola houve também intenso movimento

de espanhóis em direção a Barrancos.

O isolamento foi um fator de manutenção do dialeto: as ligações eram más em todas as direções, mas sobretudo precária a comunicação em relação ao lado português, sendo recente a ligação a Moura (a 50 km) e às restantes povoações portuguesas (a uma distância sempre superior a 20 km). Os contactos culturais e humanos com a Espanha, a escassos 9 Km, foram sempre mantidos. Daí a presença de aspetos culturais de nítida influência espanhola como os touros de morte, o folclore andaluz, as romarias, e os festejos natalícios com canções espanholas.

O capítulo II – *Descripción del barranquenho* (pp.49-169) – é central na obra. Nele se efetiva o inventário de características linguísticas do dialeto. As diversas propostas aqui apresentadas mostram-se inovadoras pelo tratamento dos fenómenos que ilustram a fala barranquenha e pela explicitação de razões que explicam as suas características. A análise sociolinguística variacional foi aplicada ao estudo das realizações da sibilante em posição implosiva.

A análise da situação linguística na comunidade ocupa a parte inicial do capítulo. Segundo a A. falam-se e ouvem-se três línguas na região, embora os falantes de Barrancos não sejam todos trilingues como tem sido afirmado. Poucas pessoas se expressam sempre em português – apenas os professores e outros funcionários. É também falado em português tudo o que respeita à religião: missas, novenas e rosários. Em situações formais, usam por vezes o português os que frequentaram o ensino secundário e superior em cidades portuguesas. Apesar disso, cerca de metade da população ainda tem pouco contacto com o português e existe um número significativo de pessoas que nunca usa o português. As famílias de origem espanhola, sobretudo nas gerações mais velhas, falam um espanhol regional, com arcaísmos, influências diversas e hipercorreções. Finalmente, falam barranquenho as pessoas ali nascidas, com pais da região. Falam-no sobretudo em família. As línguas mais ouvidas são o barranquenho e o espanhol. Os falantes trilingues são raros – só os que falam português no trabalho, barranquenho em família e espanhol com os espanhóis.

O tema central do capítulo é a descrição geral do dialeto que evidencia traços da variedade portuguesa alentejana e das variedades espanholas andaluzas e estremenhas, sendo marcado também por arcaísmos, alguma

influência leonesa, e arabismos. A complexidade da situação linguística exige uma descrição que contemple as características do português e do espanhol e que tome em consideração dados da sua variação sincrónica e diacrónica. A descrição fonológica começa pelo sistema vocálico, propondo interessantes hipóteses históricas sobre a não-elevação das vogais átonas e sobre a monotongação dos ditongos nasais – explicadas por uma influência castelhana ou por conservadorismo, porque correspondem também às realizações do português antigo. No barranquenho as sibilantes são pré-dorso-dentais como no português central e meridional, mas por vezes apicoalveolares – como no espanhol. A sibilante final pode não ser articulada, quer no sintagma nominal – no artigo, no nome, no adjectivo –, quer na 1ª pessoa do plural dos verbos. Este fenómeno, a articulação da sibilante implosiva /-s/, revela variação sociolinguística, sendo a sua realização condicionada por variáveis como o estilo e o nível de escolarização: em 67% dos casos é articulada a sibilante, em 16% dos casos é aspirada /-h/, e em 17% dos casos, não é articulada/-0/. Quanto mais formal for o estilo e maior o grau de escolarização, mais a sibilante é pronunciada. A variação corresponde também ao contexto linguístico sendo a elisão mais frequente nos determinantes: artigos, demonstrativos, indefinidos e possessivos. A descrição contempla outros fonemas como /// e /r/, que a A. confronta com realizações de outras variedades meridionais, do português do Brasil e do espanhol. O /v/ realizado como bilabial é de provável influência castelhana, porque a isófona horizontal portuguesa está muito afastada a norte. São também descritos aspetos da morfossintaxe em que se observam diferenças em relação à norma do português para as quais a A. apresenta dois tipos de explicação: o isolamento, que é propício à conservação de traços arcaicos, e a grande proximidade do espanhol. São referidos o sistema verbal, partículas diversas, categorias como o artigo, o género e o número, com formas observáveis quer em português arcaico quer em espanhol atual. Por motivos fonológicos (a elisão da sibilante implosiva /-s/), a marca de número pode não estar presente quando a consoante não é pronunciada. É assinalada ainda a presença do diminutivo arcaizante *-ito*, como em *canito*, *granito* em que surge o *n* intervocálico arcaizante, típico dos dialetos moçárabes meridionais peninsulares. No sistema pronominal é sublinhada a presença de pronomes como *vocedes* e

com nós, presentes também em galego e noutras variedades de português. O pronome complemento apresenta traços arcaizantes: *la vendia, le disse, lo chamam, sas levou*.

Em relação ao léxico, algumas formas espanholas estão mais integradas no dialeto, outras são superficiais e tendem a desaparecer nos falantes que dominam o português. Na antroponímia e na toponímia encontram-se nomes tipicamente espanhóis. Na segunda parte do capítulo é sistematizada em quadros e tabelas a distribuição da sibilante e é estudado o efeito dos diferentes contextos linguísticos – na presença da sibilante, na articulação aspirada e na elisão. Figuram igualmente quadros com os resultados da comparação com as realizações das vilas espanholas mais próximas e com outros dados dialetais de Portugal. A A. estabelece ainda uma comparação do barranquenho com os dialetos portugueses do Uruguai, em que se observa a elisão de /-s/ na 1ª p. dos verbos, fenómeno com diferente distribuição em relação ao espanhol sul-americano. E faz também um confronto entre a distribuição de /-s/ em barranquenho, onde a presença da sibilante é maioritária, e alguns dados do espanhol das Caraíbas, revelando que neste dialeto hispânico a presença da sibilante é minoritária. A A. demonstra que é possível observar momentos diferentes de evolução para este fenómeno muito difundido em variedades peninsulares meridionais e perspectiva em números globais e percentuais a variação da sibilante, que apresenta parâmetros complexos. A comunidade barranquenha revela um uso heterogéneo, podendo ocorrer as três articulações num mesmo falante. A variação é analisada nos diferentes contextos com graus de formalidade variável e a análise dos dados estatísticos explicita a correlação de fatores linguísticos e sociais. Todas as variáveis analisadas neste capítulo estão expostas no Anexo II. A interpretação dos resultados revelou que, no barranquenho, em 82% dos casos a sibilante está presente, sendo 67% a norma e 33% desvio – a articulação aspirada. A presença da sibilante verifica-se mais em posição interna e a elisão mais em posição final. A A. analisa em seguida as consequências do fenómeno – presença ou ausência do morfema de número, no sintagma nominal e na terminação verbal da 1ª pessoa do verbo. A parte final do capítulo inclui as entrevistas de Lindley Cintra em Barrancos e um estudo comparativo do léxico do barranquenho, do português-padrão, do espanhol de Encinasola e do espanhol-padrão para

observar a influência das áreas portuguesa e espanhola em contacto.

No capítulo III – *Las lenguas fronterizas* (pp.171-200) – a A. procede à análise dos diversos casos de existência de línguas fronteiriças entre a Espanha e Portugal, marcadas por bilinguismo e por mútua influência linguística. No seu conjunto, os aspetos reunidos neste capítulo perseguem dois objetivos fundamentais: a apresentação do conceito de línguas em contacto e a sugestão da utilidade da aplicação deste conceito a casos concretos exemplificados pela tipologia do barranquenho e da sua comparação com o “fronterizo” sul-americano.

A abordagem das línguas fronteiriças em contacto é fonética, mas também morfossintática e há uma integração dos fenómenos observados nos dados conhecidos de outras línguas românicas. São apresentados, em quadros e em tabelas, os resultados do inquérito realizado – resultados de /r/, /ll/ e da alternância entre /r/ e /ll/ em posição interior; da lábio-dental /v/ e da bilabial /b/; e da elisão de /-d-/ em posição intervocálica. Todos estes traços são comuns nas falas meridionais da península. A A. estabelece uma comparação/contraste entre o barranquenho e o ‘fronterizo’ falado na região do Uruguai em fronteira com o Brasil. Expõe o contraste fonológico, o contraste morfossintático e alguns exemplos lexicais do *fronterizo* e do barranquenho, em quadros comparativos. As condições de contacto de línguas análogas – o português e o espanhol – nestes dois casos, reveste-se de traços semelhantes, tanto mais que as variedades das duas línguas ibéricas que se encontram em contacto na região sul-americana são também meridionais. Estes dois dialetos em confronto apresentam dados socioeconómicos diferentes, mas com uma mesma determinação e atitude positiva dos falantes em relação à sua língua. Na parte final do capítulo III, a A. faz ainda uma análise detalhada de fenómenos de mútua influência do português e do castelhano e estabelece paralelos com a situação do galego e do leonês.

No Capítulo IV – *Literatura oral y tradicional. Testimonios de música popular* (pp. 201-263) – a A. abre o âmbito do estudo a dimensões literárias, antropológicas e etnológicas sem nunca se afastar de um propósito eminentemente linguístico, reunindo um precioso acervo de aforismos, refrões, anedotas, contos, canções e romances tradicionais. Destaca-se a recolha e análise de composições da música popular tradicional desta

região, nomeadamente de textos das canções dos “quintos de Barrancos”, relacionadas com a partida dos jovens para o serviço militar. Por vezes, o tema das composições é mais geral e poder-se-iam incluir nas “cantigas de ronda” conhecidas na Espanha meridional e no Alentejo. A recolha dos textos é seguida de uma análise detalhada das características fonéticas tipicamente meridionais e de um levantamento tipológico de traços morfológicos e lexicais a partir do estudo de diversas versões dos textos orais/tradicionais. A perspetiva comparatista sublinha a importância relativa do contexto português ou espanhol.

De referir ainda a extensa informação bibliográfica que abarca um domínio multidisciplinar, com predomínio da linguística da variação, comportando estudos de dialetologia e de sociolinguística, de linguística românica, portuguesa e espanhola e também de literatura oral/tradicional, folclore, antropologia e história.

Esta obra está direccionada para a reflexão sobre a importância do estudo das línguas em contacto e a A. defende que o processo de diferenciação é complexo, fazendo apelo à necessidade de incorporar a dimensão social à determinação dos grupos a inquirir. É de sublinhar o facto de serem abordados conjuntamente os aspetos teóricos da questão das línguas em contacto e os problemas metodológicos concretos inerentes à identificação, descrição e explicação das características de uma variedade dialetal. Para além da importância dos contributos gerais que as implicações teóricas e práticas da metodologia preconizada nesta obra trazem à teoria linguística, ela interessa de um modo particular aos investigadores que estudam temas relacionados com dialetos em situação de bilinguismo, nomeadamente na península ibérica e na América Latina. É sobretudo nestes domínios que as diversas propostas teóricas e metodológicas aqui apresentadas se mostram inovadoras, pelo tratamento pluridimensional dos dados de um inquérito e pela conjugação de uma tradição dialetológica de investigação de variedades linguísticas com correntes mais recentes da ciência linguística. Trata-se da única obra de fôlego publicada sobre o tema e resulta de um período muito longo de investigação. Durante as duas décadas em que decorreu a sua investigação, a A. publicou um grande número de trabalhos parcelares relacionados com o estudo do dialeto barranquenho. Mas sentia-se sem dúvida a falta desta obra de síntese que apresenta de modo muito

completo e sistemático os fecundos resultados dessa investigação. Pena é, talvez, que exista um lapso temporal grande entre esta publicação e a fase inicial do trabalho, mas o prolongado período de investigação permitiu o desenvolvimento da análise e o amadurecimento das conclusões. Trata-se de um estudo multifacetado e inovador do ponto de vista teórico e metodológico: descreve as características do dialeto barranquenho, enquadra-o num cenário de variação e propõe análises que suscitam conhecimentos da dialetologia do espanhol e do português como idiomas em contacto e da evolução histórica das duas línguas. Situando-se, assim, na interseção de parâmetros de variação sincrónica e diacrónica, esta obra constitui um valioso contributo para o estudo da linguística da variação.



Paulo Nunes da SILVA. *Tipologias Textuais - Como Classificar Textos e Sequências*, Coimbra: Editora Almedina. Coleção CELGA, 2012. 220 pp.  
ISBN 9789724048789

Isabel Margarida Duarte  
iduarte@letras.up.pt  
*Faculdade de Letras da Universidade do Porto (Portugal)*  
*Centro de Linguística da Universidade do Porto*

Paulo Nunes da Silva doutorou-se em 2006 com uma dissertação intitulada *O tempo no texto*. É docente do Departamento de Humanidades da Universidade Aberta, membro do Celga (Centro de Estudos de Linguística Geral e Aplicada) da Universidade de Coimbra, e trabalha, preferencialmente, nas áreas de Semântica, Pragmática e Linguística Textual. As suas publicações recobrem, sobretudo, os seguintes tópicos, que referimos por estarem todos eles implicados no livro em apreço: tempo, aspeto e tipologias textuais (géneros discursivos, sequências textuais).

Esta obra, de publicação financiada pela Fundação para a Ciência e a Tecnologia, faz parte de uma edição conjunta Almedina / CELGA, incluindo-se numa Série coordenada pelas docentes Ana R. Luís e Joana Vieira Santos, ambas da Faculdade de Letras da Universidade de Coimbra, série cujo aparecimento se saúda e através de cujas publicações se pretende dar a conhecer “investigação proveniente da pesquisa e do ensino em Linguística e Ciências da Linguagem, junto de especialistas, professores e alunos universitários” (da badana).

O livro que agora se aprecia é de indiscutível interesse para docentes de Português dos Ensinos Básico e Secundário, mas também se revela de grande utilidade para estudantes universitários que frequentem Unidades Curriculares do âmbito do texto e do discurso, sejam de Línguas ou de outras áreas, como Jornalismo e Ciências da Comunicação, tradução, etc., para estudantes de Linguística Geral e até de Didática do Português.

O autor sistematiza, de forma muito clara, organizada e bem fundamentada teoricamente, os principais contributos científicos sobre texto em geral e tipologias textuais, em particular, com a vantagem de ir fornecendo frequentes exemplos, muito bem selecionados, que permitem ao leitor ir percebendo e aplicando, na prática, a teoria explanada. Não esquece também, além da abundante bibliografia de referência estrangeira, sobretudo francesa, o que foi escrito em Portugal na área em causa (sobretudo por Maria Antónia Coutinho e Ana Cristina Macário Lopes). Cita os autores mais óbvios deste campo, baseando muito do que escreve nos ensinamentos de Adam (1992), mas utiliza fontes teóricas que vão muito para além do núcleo central do âmbito em estudo. A bibliografia é atualizada e completa e o livro inclui, antes dela, um Índice temático que facilita a consulta pontual.

A escrita de Paulo Nunes da Silva é muito clara, enxuta e elegante. Essa clareza facilita o acesso a teorias por vezes complexas, que o autor não teve, felizmente, a tentação de simplificar para divulgação. Assim, não abdica do rigor científico, mas torna a teoria acessível a um público alargado, por ser organizado, não usar jargão metalinguístico excessivo e utilizar de forma muito didática os exemplos selecionados. Também os inúmeros esquemas, quadros e mapas conceptuais bem como a feliz apresentação gráfica do livro favorecem a leitura.

A parte final, em que são apresentados exercícios com soluções pormenorizadas e bem justificadas, dá um cariz ainda mais prático ao manual, permitindo que o leitor treine os conhecimentos teóricos que foi adquirindo e favorecendo atitudes de autoformação.

A obra é, assim, totalmente adequada aos eventuais destinatários, sobretudo aos professores de Português dos Ensinos Básico e Secundário que bem necessitados estão dela, porque devem abordar, com os respetivos alunos, tendo em conta documentos orientadores como o Programa de Português do Ensino Básico (2009) e o *Dicionário Terminológico*, questões teóricas do âmbito das tipologias textuais que, muitas vezes, não fizeram parte das matérias estudadas na respetiva formação inicial universitária. No que concerne aos estudantes universitários, o facto de terem agora ao dispor uma boa obra de síntese, em português, didaticamente adequada, é de inegável utilidade, porquanto muitos deles não têm acesso aos textos

originais, por exemplo, escritos em francês.

O livro divide-se em 4 capítulos. No primeiro, o autor define o conceito de texto, sublinhando a sua omnipresença por ser o produto das nossas constantes interações verbais e mostra-nos o seu caráter de objeto pluridimensional. Apesar da enorme variedade dos textos com que lidamos, há regularidades e constantes que os definem, às quais o autor dedica atenção neste capítulo. No segundo, problematiza de forma muito interessante e aprofundada as questões que dizem respeito à própria atividade de classificação de textos de que se vai ocupar no capítulo seguinte. Assim, neste segundo capítulo, explica em que consiste a atividade de classificação de textos, refere as vantagens de a levar a cabo e explicita alguns princípios metodológicos a ter em conta aquando da realização de tal atividade. Com efeito, conseguir classificar textos é facilitador não só das tarefas de receção e leitura, mas também das de produção textual. O capítulo terceiro, o mais extenso de longe, ocupa um lugar central na obra, pois é nele que se desenvolve a questão da classificação de textos, assim desdobrada: (1) tipos de textos; (2) géneros discursivos; (3) tipos de textos e conceito de sequência textual; e, por fim, (4) tipos de sequências textuais. Neste último ponto, o autor debruça-se sobre as características linguísticas e discursivas das sequências narrativas, descritivas, argumentativas, explicativas e dialogais.

O facto de o autor possuir uma sólida formação também na área da Semântica permite-lhe convocar conhecimentos desse âmbito que vêm reforçar e credibilizar os critérios utilizados para a classificação textual (veja-se, por exemplo, a utilização da teoria de Moens & Steedman (1988), em vários momentos da obra). As explanações constantes do capítulo 3 permitem ao leitor, sobretudo ao que for docente de Português, perceber a complexidade dos fenómenos em causa e evitar simplificações abusivas que distorçam a teoria, quando se tratar de didatizar e operacionalizar os modelos apresentados para os trabalhar com alunos mais novos. No último capítulo, como se disse, o autor apresenta um conjunto de exercícios bem pensado e respetivas propostas de soluções, assim promovendo uma desejável atitude de autoformação.

Outra mais-valia deste livro prende-se com a escolha dos exemplos utilizados. O autor, livre de preconceitos e fora de polémicas estéreis em torno de falsos binómios como literário / não literário, literatura / língua,

etc., socorre-se de um conjunto muito bem selecionado de textos, sejam eles dos *media*, do cânone literário escolar, ou de obras literárias menos canónicas. Assim, se o leitor encontra, eventualmente com satisfação, no caso de ser professor de Português, exemplos de textos que ele próprio utiliza (ou pode vir a utilizar) com os seus alunos, ficando com a sensação de que a leitura do livro é imediatamente útil para o seu trabalho docente, também se depara com textos que previsivelmente não conhece e cuja leitura, à luz das teorias que exemplificam, constituirá uma abertura para a mudança e a novidade.

#### REFERÊNCIAS

- Adam, Jean-Michel (1992). *Les textes: types et prototypes*. Paris: Éditions Nathan.
- Moens, Marc & Steedman, Mark (1988). Temporal ontology and temporal reference. *Computational linguistics* vol. 14 (nº 2), 15-28.

## Instructions to authors

### **Editorial policy**

*Linguística. Revista de Estudos Linguísticos da Universidade do Porto* accepts papers on any linguistic topic. Papers from either fundamental or applied research will equally be considered for publication, no matter the theoretical background of the submitted studies.

### **Submission and acceptance**

Prospective authors are encouraged to submit manuscripts within the scope of *Revista de Estudos Linguísticos da Universidade do Porto*. Submitted papers must be previously unpublished – though revised versions or translations of previously published texts can be exceptionally admitted as candidates for publication – and must not be under consideration for publication elsewhere.

All submissions are transmitted to at least two anonymous referees.

For initial submission an email copy should be sent to the Editor, without any authors' identification, accompanied by a separate title page, giving each author's name, affiliation, address (to which proofs and editorial correspondence can be sent), phone, fax and e-mail address. For co-authored papers, first author's contacts will be used in further correspondence.

After a paper is accepted for publication, authors must present a final version as specified below and in full accordance with the instructions provided by this style sheet.

*Word processors and other software:* MS Word for Windows (DOC, DOCX or RTF file). For papers with specific symbols or figures (phonetic transcription, OT tableaux, syntax trees a.s.o.), a PDF version (in addition to a DOC or RTF version) is required. Moreover, clear indication of the used software must be given to the Editor. For phonetic transcription, please use the IPA-samd Uclphon1 SILDoulos font.

## Style

Final versions of accepted papers will observe the following specifications. The Editor reserves the right to return the manuscript to the authors for any corrections when these norms are not respected.

As the style specifications globally corresponds to the MIT Press style sheet, for further information please consult <http://mitpress.mit.edu/journals/LING/li-style.pdf>.

## Length

*Articles:* The preferred length of articles is 20 A4 pages approximately, Times New Roman 12 points, double spacing, 2.5 cm margins on all sides. Full-length articles should deal with original topics or research.

*Research notes:* The preferred length of research notes is 8 A4 pages approximately, Times New Roman 12 points, double spacing, 2.5 cm margins on all sides. Research notes could include brief accounts of research or report important work in advance of a more comprehensive paper.

*Book/Software/Webpage reviews:* The preferred length of reviews is 4 A4 pages approximately, Times New Roman 12 points, double spacing, 2.5 cm margins on all sides. Reviews will normally be commissioned by the Editor; nevertheless, offers to review recent books, software or webpages are welcomed. Please get in touch with the Editor if you wish to publish a review. Each book review should specify full bibliographic details of the reviewed book (title, author(s)/editor(s), place and year of publication, publisher, number of pages, edition, hardback/paperback, ISBN). Software reviews should specify full authorial and technical details (commercial designation, authors, copyright owner, version number, required computer operating system). Webpage reviews should give all necessary details regarding the web host, page creators, http address and date of retrieval.

## Languages

Papers must be written in English, Portuguese or any widespread language. Contributions in English may use either British or American spelling, provided it is used consistently. Do not hyphenate English words. Contributions in Portuguese may use either Portuguese or Brazilian spelling, provided it is used consistently.

## Layout

*Margins:* 2.5 cm on all sides. Use A4 format for the printed copies.

*Font:* Times New Roman, 12 pt. For long quotations and captions: 11 pt (see below).

*Line-spacing:* Double-spacing, except for abstracts and key-words, tables and figures, long quotations and reference list (where single-spacing should be used).

*Page numbers:* Page numbers at page bottom, centred.

*Phonetic symbols:* For phonetic transcription, please use the IPA-samd Uclphon1 SILDoulos font.

*Title, authors' identification, abstract and key-words:*

✓ Top of first page: Title of the paper. Times New Roman, 18 pt, bold, centred, normal capitalisation.

✓ Empty line (18 pt)

✓ Author(s)' name(s) (name(s) and surname(s)). Times New Roman, 16 pt, regular, centred, normal capitalisation. One author per line. Together with each name, in a separate line underneath the author's name, give an e-mail address (Times New Roman, 12 pt, regular, centred). In the following line, indicate author's affiliation (institution, country, with the country's name in brackets). Times New Roman, 14 pt, italics, centred, normal capitalisation.

✓ Empty line (14 pt)

✓ Abstract in English (if original language different from English). Times New Roman, 11 pt, regular, justified. Up to 500 words approximately. Heading (11 pt, first line of the abstract text): ABSTRACT.

✓ Empty line (11 pt)

✓ Key-words in English (if original language different from English). Times New Roman, 11 pt, regular, justified. Up to 6 key-words in the paper's language. Heading (11 pt, first line of the abstract text): KEY-WORDS.

✓ Abstract in the paper's language. Times New Roman, 11 pt, regular, justified. Up to 500 words approximately. Heading (11 pt, first line of the abstract text): Empty line (11 pt)

✓ Key-words. Times New Roman, 11 pt, regular, justified. Up to 6 key-words in the paper's language. Heading (11 pt, first line of the abstract text).

- ✓ Empty line (11 pt)
- ✓ 2 empty lines (12 pt)
- ✓ Text

*Notes and acknowledgements:* Footnotes in the text should be identified by superscript numbers and listed consecutively at each page bottom. Acknowledgements should be made in a first note, marked with an asterisk (this note should be introduced immediately after the title's last word).

*Section headings:* All sections and subsections should have a heading. Section headings should be numbered as in the following:

1. Section title
- 1.1 Subsection title level 1
- 1.1.1 Subsection title level 2

*Examples, tables, figures, etc.:* Examples, tables and figures should be inserted in the text and numbered consecutively with Arabic numerals. Each table and figure should have a title, at its top (Times New Roman, 12 pt, single-spacing, left-aligned) according to the following examples.

TABLE 1 – Title of table

FIGURE 1 – Title of figure.

Captions must occur at the figure or table bottom: Times New Roman, 11 pt, single-spacing.

In the full printed version and in the electronic copy, tables and figures must be included in their intended locations. On separate sheets and separate files (DOC/RTF and PDF), additional copies of tables and figures should be provided (1 figure or table per A4 page). These additional versions may fit camera-ready quality (clear black print, laser or high quality ink-jet printer). Their lettering should be large enough to be legible after reduction. Only black and white tables and figures can be accepted for final publication.

*Italicisation:* Do not underline examples or emphasised terms; these should be italicised. Bold type or small capitals can also be used.

*Examples:* All examples included in paragraphs should be given in italics (except when representing phrase/sentence structure).

*Numbering of examples:* examples which are not included in the text of a paragraph should be numbered, with the number placed in parentheses. Tabs (not spaces) should be used to align the examples. Sets of related items



may be listed together, and numbered with lowercase letters, as in the example below:

- (1) a. a slice of bread, a glass of water
  - b. a pile of books, a row of houses
- (2) strawberry, raspberry, blueberry

Examples should be numbered consecutively throughout the whole text. List numbering may be used automatically. This allows adding and updating cross-references to examples. To enable it, choose Format > Bullets and Numbering > Numbered, or right-click and choose Bullets and Numbering > Numbered.

The preceding and the following text should be separated from the example(s) by one blank line.

References to numbered examples should take the following form: “as in (1b) and (2)”.

If the paper includes examples from a language that does not coincide with the one of the paper, glosses and translations must be given in the language of the paper; use SMALL CAPITALS to gloss a grammatical category or grammatical category morpheme in a linguistic example, like in the following:

- (1) Paolo li                      ha già       letti.  
 Paolo them (MASC.PL) has already read (MASC.PL)  
 ‘Paolo has already read them.’
- (2) Kodomo ga 3-nin waratta.  
 kids NOM       3-CL laughed  
 ‘Three kids laughed.’

*Quotations:* Short quotations are included in the text, enclosed in quotation marks (Times New Roman, 12 pt). Longer quotations should begin a new line and be indented, in Times New Roman, 11 pt, single space, without any quotation marks. After each long quotation, its source must be indicated (right-aligned, Times New Roman, 11 pt, single-space), following the bibliographical references style (see below). Inside a quotation, a suppression of any original passage should be marked with [...].

*Experimental data:* Authors should supply sufficient information to enable replication of investigations. Statistical results must be clearly indicated, following the norms of the American Psychological Association. Give subjects' chronological ages in years, years:months or years:months.days (when appropriate).

*References in the text:* Reference in the text should be to author's name and date. When appropriate, indicate relevant chapter/section or, preferably, page numbers (see following examples).

According to Kuhn (1962: 44), ...

The links between emotion, language and behaviour are taken into consideration by several authors (e. g.: Cross, Blake, Turnbridge & Gill 2001: 228 ff.).

For co-authored papers, include '&' before the last author's surname (see example above). For papers with three or more co-authors, indicate all co-authors' names in the first mention; thereafter, indicate first author's name, '*et al.*' (italicised) and date of reference:

Cross *et al.* 2001

All personal communications should be identified as 'p.c.' after the source name and given a date (if possible) (e. g.: Matthews p.c. 2004).

#### *List of references*

References should be listed alphabetically by author at the end of the article. Please type **REFERENCES** (Times New Roman, 12 pt, bold, small capitals, left-aligned) before the first reference. An empty line (12 pt) should be kept immediately above and underneath this heading. All references in Times New Roman 12 pt, single-spacing, indented, as in the following examples. For references with more than one author, use a comma; and '&' to separate the last co-authors' name; or *et al.* (for three or above three authors) in the reference list.

Use the references style proposed at:

<http://mitpress.mit.edu/journals/LING/li-style.pdf>

Here there are some examples:

- Books:

Kuhn, T. 1962. *The structure of scientific revolutions*. Chicago: The University of Chicago Press.

- Papers in journals:

Jusczyk, P. W., Goodman, M. B. & Baumann, A. 1999. Nine-Month-Olds' Attention to Sound Similarities in Syllables. *Journal of Memory and Language*. 40(1): 62-82.

- Chapters in books:

Goodluck, H. 1986. Language acquisition and linguistic theory. In: P. Fletcher & M. Garman (Eds.). *Language Acquisition*. Cambridge: Cambridge University Press, 49-68.

Do not include the edited volume as a separate entry of the reference list, unless it is explicitly referred to as such in the text. In this latter case, proceed as follows:

Fletcher, P. & Garman, M. (Eds.). 1986. *Language Acquisition*. Cambridge: Cambridge University Press.

- Documents retrieved from the Internet:

Zeichner, K. M. 1983. Alternative paradigms of teacher education. *Journal of Teacher Education*. 34 (3): 3-13. Retrieved January 25, 1996, from the World Wide Web: <http://www.apa.org/journals/zeichner.html>.

For documents not available as printed publications:

Skehan, P. 2002. *Individual differences in second and foreign language learning*. Retrieved April 19, 2005, from the World Wide Web: <http://www.lang.ltsn.ac.uk/resources/goodpractice.aspx?resourceid=91>.

- Unpublished material and other sources:

Give as many details as you can. For unpublished manuscripts or

mimeographs, consider them as books and indicate 'ms' instead of publisher's identification. For submitted or forthcoming papers, treat them as papers and supply information such as 'forthcoming', 'in press' or 'in preparation'.

### **Appendices**

When absolutely essential, a final section of appendices can be included after the reference list. This section may contain experimental items, corpora or iconic materials relevant for the illustration of the authors' points of view or for the demonstration of experimental results. Appendices are ordered consecutively with capital letters (Appendix A, B, C...). The Editor reserves the right to judge any appendix irrelevant and therefore to suggest its suppression from the final publication. The inclusion of a section of appendices should be regarded as exceptional.

### **Proofreading**

Once a paper is reformulated on the basis of the referees' suggestions and its final version is accepted, no substantial modifications will be allowed. Normally, all proofreading will be carried out by the Editorial Committee. Nonetheless, the Editor can ask the author to review a set of page proofs. No alterations other than of printer's errors will be admitted at this stage.

















# LINGUÍSTICA

REVISTA DE ESTUDOS LINGUÍSTICOS  
DA UNIVERSIDADE DO PORTO

Foreword, *Ana Maria Brito*

Artigos

Nouns in apposition: Portuguese data, *Graça Rio-Torto*

Noms épithètes dans les expressions binominales, *Fernando Martinho*

Italian verb to noun conversion: the case of nouns in –a deriving from verbs of the 2<sup>nd</sup> and 3<sup>rd</sup> conjugation, *Daniela Marzo*

Is conversion a syntactic or a lexical process of word formation? *Alexandra Soares Rodrigues*

Bare Morphology, *Alina Villalva*

Deriving color adjectival nominalizations, *Artemis Alexiadou*

Deadjectival human nouns: conversion, nominal ellipsis, or mixed category?  
*Petra Sleeman*

Les verbes à objets cognats en portugais européen: quelques réflexions  
sur la conversion et la génération de mots, *Celda Choupina*

Varia

Evidentials *dizque* and *que* in Spanish.

Grammaticalization, parameters and the (fine) structure of Comp, *Violeta Demonte*,  
*Olga Soriano*

Óscar Lopes: recordar o linguista, o mestre e o amigo, *Fátima Oliveira*

Recensões

R. do Carmo (org.) *Linguagem, Argumentação e Decisão Judiciária*, *Belinda Maia*

Fernão de Oliveira. *Gramática da Linguagem Portuguesa*, Fac-simile, introdução e edição  
atualizada e anotada por José Eduardo Franco e João Paulo Silvestre, *Sónia Duarte*

M. V. Navas Sánchez-Élez: El barranquenho. Un modelo de lenguas en contacto,  
*Clara Araújo Barros*

P. N. da Silva: Tipologias Textuais - Como Classificar Textos e Sequências, *Isabel Margarida Duarte*

VOLUME 8

ANO 2013