

27/1

revista do centro de estudos humanísticos
série ciências da linguagem
2013

diacrítica

LEIUS



Universidade do Minho
Centro de Estudos Humanísticos

27/1

revista do centro de estudos humanísticos
série ciências da linguagem
2013

diacrítica

CEH



Universidade do Minho
Centro de Estudos Humanísticos

Título: DIACRÍTICA (N.º 27/1 – 2013)
Série Ciências da Linguagem

Diretora: Ana Gabriela Macedo

Diretores-Adjuntos: Carlos Mendes de Sousa; Vítor Moura

Coordenadora: Maria do Pilar Barbosa

Comissão Editorial: Pilar Barbosa (U. Minho), Cristina Moreira Flores (U. Minho), José de Sousa Teixeira (U. Minho)

Comissão Redatorial: Aldina Marques (U. Minho), Ana Cristina Macário Lopes (U. Coimbra), Ana Maria Brito (U. Porto), Anabela Barros (U. Minho), Anabela Gonçalves (U. Lisboa), Antónia Coutinho (U. Nova de Lisboa), Augusto Silva (U. Católica Portuguesa), Conceição Paiva (U. Federal do Rio de Janeiro), Denise Kluge (U. Federal do Paraná), Eugénia Duarte (U. Federal do Rio de Janeiro), Graça Rio-Torto (U. Coimbra), Henrique Barroso (U. Minho), Isabel Margarida Duarte (U. Porto), Ivo Castro (U. Lisboa), João Veloso (U. Porto), José Teixeira (U. Minho), Maria Cristina Figueiredo (U. Federal de Santa Catarina), Maria João Freitas (U. Lisboa), Maria Lobo (U. Nova de Lisboa), Nélia Alexandre (U. Lisboa), Rolf Kemmler (U. Trás-os-Montes), Tjierk Hagemeier (U. Lisboa)

Comissão Científica: Jorge Morais Barbosa (U. Coimbra); António Branco (U. Lisboa); Ana Brito (U. Porto); Ivo Castro (U. Lisboa); Antónia Coutinho (U. Nova de Lisboa); Maria João Freitas (U. Lisboa); Jürgen M. Meisel (U. Hamburgo / U. Calgary); José Luís Cifuentes Honrubia (U. Alicante); Mary Kato (U. Campinas); Rui Marques (U. Lisboa); Fátima Oliveira (U. Porto); Graça Rio-Torto (U. Coimbra); José Luís Rodrigues (U. Santiago de Compostela); Eduardo Paiva Raposo (U. da Califórnia, Sta. Bárbara); Conceição Paiva (Universidade Federal do Rio de Janeiro); Augusto Soares da Silva (U. Católica Portuguesa)

Edição: Centro de Estudos Humanísticos da Universidade do Minho em colaboração com Edições Húmus – V. N. Famalicão. *E-mail:* humus@humus.com.pt

Publicação subsidiada por
FCT – Fundação para a Ciência e a Tecnologia

ISSN: 0807-8967

Depósito Legal: 18084/87

Composição e impressão: Papelmunde – V. N. Famalicão

ÍNDICE

- 5 ***Se faire + Vinf: un outil au service de la construction d'une diathèse 'maléfactive' de l'objet (in)direct***
Sílvia Lima Gonçalves Araújo
- 39 ***A descriptive study of Brazilian offensive phrases***
Ana C. Bastos-Gee
- 69 ***Menço ou minto? Regularização de paradigmas verbais***
Maria João Colaço e Esperança Cardeira
- 95 ***"A beached whale posing in lingerie" Conflict talk, disagreement and impoliteness in online newspaper commentary***
Isabel Ermida
- 131 ***Nota de morfologia histórica do português: sufixo -ó***
Przemysław Dębowskiak
- 153 ***Revisitando a variação entre ditongos nasais finais átonos e vogais orais na comunidade de fala do Rio de Janeiro***
Christina Abreu Gomes, Cássia Mesquita e Taís da Silva Fagundes
- 175 ***Mecanismos de textualização e construção textual: para uma abordagem sócio-discursiva do *cartoon****
Audria Leal e Ana Caldes
- 195 ***Gramaticalização e especialização funcional: o caso do conector *pois****
Maria da Conceição de Paiva e Maria Luiza Braga
- 217 ***SNS como rótulos em livros didáticos de história do Brasil: simples ou complexos?***
Vera Lúcia Paredes Pereira da Silva e Gabrieli Pereira Bezerra
- 239 ***"Componente" como substantivo uniforme de dois géneros***
Iva Svobodová

QUESTÕES DE ÉTICA E POLÍTICA LINGUÍSTICAS

- 267 **A situação atual do ensino da língua portuguesa na Galiza**
Xurxo Fernández Carballido
- 287 **How can sociolinguistic data be used?**
Celeste Rodrigues e Deolinda Simões

RECENSÕES

- 311 **Pinto, Paulo Feytor (2010), *O Essencial sobre Política de Língua.***
Henrique Barroso
- 317 **Ruth Amossy, (2010). *La présentation de soi – Ethos et identité verbale,***
Micaela Aguiar
- 321 **Sánchez Rei, Xosé Manuel (2011), *Lingua galega e variación dialectal,***
Maria Aldina Marques
- 325 **Normas de apresentação**

SE FAIRE + VINf: UN OUTIL AU SERVICE DE LA CONSTRUCTION D'UNE DIATHÈSE 'MALÉFACTIVE' DE L'OBJET (IN)DIRECT

Sílvia Lima Gonçalves Araújo*
saraujo@ilch.uminho.pt

Résumé

Nous entendons proposer ici une réflexion sur l'utilité d'un corpus monolingue dans le domaine touchant à l'étude du langage sur la base d'un exemple concret, celui des constructions en *se faire*+*Vinf*. À partir d'un corpus littéraire et journalistique que nous examinons dans le sens de déterminer les types de verbes préférentiellement sélectionnés par *se faire*, nous cherchons essentiellement à rendre compte du mode de fonctionnement de ce marqueur diathétique dans des contextes où le sujet est construit avec une absence de télicité sur ce qui lui arrive. Comme nous le verrons, *se faire* s'associe bien volontiers à des verbes détritimaux appartenant à différents registres de langue (notamment non standard) pour construire une diathèse maléfactive dérivée d'un accusatif ou d'un datif.

Mots-clés: diathèse (maléfactive), *se faire*, argot, passif, objet (direct ou indirect)

Abstract

This paper is intended to establish a space for reflection on the usefulness of a monolingual corpus in fields concerning the study of language, using for that purpose a concrete example, i.e. that of *se faire*+*Vinf* constructions. Based on a literary and journalistic corpus, which we queried to determine the types of verbs more often selected by *se faire*, we essentially aim to describe the linguistic mechanics behind this diathetic marker in contexts where the subject lacks telicity over the events that overcome him. As we will attempt to demonstrate, *se faire* willingly associates with negative verbs belonging to different language registers (particularly non-standard) in order to build a detrimental diathesis derived from the direct or indirect object.

Keywords: (malefactive) diathesis, *se faire*, slang, passive, (direct or indirect) object

* Universidade do Minho, Departamento de Estudos Românicos, Braga, Portugal.

1. L'introduction

La plupart des études signalent que la signification de *se faire-Inf* «semble aller aussi bien du côté de l'actif que du côté du passif, au gré des contextes, ce qui paraît à la fois étonnant et agaçant» (Blanche-Benveniste, 2007: 155). Comme le note, à juste titre, Blanche-Benveniste (2007: 164) à qui nous empruntons les quatre exemples qui suivent, les effets d'activité ou de passivité que l'on peut obtenir par le biais de *se faire* n'ont rien à voir avec la syntaxe de *se faire* mais dépendent entièrement du lexique mis en jeu et du contexte socio-culturel. «La tendance naturelle est de prendre pour actifs les exemples dans lesquels l'action semble bénéfique au sujet [exs. *je me suis fait ouvrir la porte; elle s'est fait teindre les cheveux*] et de prendre pour passifs ceux dont l'action lui est néfaste et pour laquelle il est peu probable qu'il ait mis en œuvre sa volonté [*je me suis fait fermer la porte au nez; elle s'est fait griffer le visage*]» (Blanche-Benveniste, 2007: 164).

C'est très certainement cette richesse diathétique (Kokutani, 2005) qui amène les linguistes à définir de manière aussi variée cette construction. Comme le remarque, à juste titre, Novakova (2009), pour certains auteurs (Spang-Hanssen, 1967; Roggero, 1984; Riegel *et al.*, 1993, Labelle, 2002, entre autres), il s'agit d'une forme de passif. D'autres (Tasmowski & Van Oevelen, 1987) proposent un traitement unitaire: malgré des valeurs très similaires à la construction passive, le tour reste causatif (le passif est un sous-cas du causatif pronominal). D'autres encore (Kupferman, 1995) renoncent au traitement unitaire au profit d'une analyse binaire⁽¹⁾ puisqu'ils divisent la construction, selon que l'action est positive (volontaire) ou négative (involontaire) pour le sujet, en deux interprétations: l'une qui relève

(1) Tasmowski & van Oevelen (1987) considèrent que *se faire + Vinf* a une valeur propre réfléchie dont le passif est issu. Les données diachroniques semblent aller à l'encontre de cette thèse. En effet, comme l'indique Creissels (2003-2004: chap. XVII, p. 9), à partir d'une valeur causative, «on serait passé à une possibilité d'interprétation passive par l'intermédiaire d'une réflexivisation de la construction causative sans marque morphologique». Cette évolution en trois étapes est schématisée par Novakova 2009 de la façon suivante: **causatif** (*X a fait assassiner Y*) → **réflexivisation** (*Y s'est fait assassiner*) → **passif** (*Y a été assassiné*). Kupferman (1995: 76) postule, au contraire, l'existence de deux constructions «hétérogènes» en *se faire* (une «causative réfléchie» et une «passive»); selon lui (1995: 57), «la ressemblance morphologique entre deux formes syntaxiques ne signifie pas nécessairement qu'elles soient typologiquement apparentées». Ce serait, pour lui, la construction passive en *se faire* qui est «devenue homonyme» et «sémantiquement indépendante de la construction causative réfléchie». L'auteur précise que la phrase qui suit, par exemple, est ambiguë, en ce qu'elle contient soit la construction en *se faire-Inf* causative, qui se glose en anglais par (ii), soit la construction en *se faire-Inf* passive, qui se glose en anglais par (iii): (i) *il s'est fait tuer dans la 404*; (ii) *he had himself killed in his car* (lecture causative); (iii) *he got killed in his car* (lecture passive).

de la construction causative, «où le sujet est un 'instigateur', c'est-à-dire un agent» (Kupferman, 1995: 60) et l'autre qui porte un sens passif «où le sujet est seulement un patient dans un procès dont l'agent peut être explicite ou non» (*idem*). Selon la valeur des verbes auxquels s'associe *se faire*, les exemples semblent donc se diviser en deux catégories principales: les actions bénéfactives et les actions maléfactives, dont les dernières correspondent à ce que les linguistes nomment couramment «actions désagréables ou violentes» (cf. Spang-Hanssen, 1967; Gaatone, 1983).

Dans les pages qui suivent, nous nous occuperons davantage des exemples qui décrivent un acte désavantageux pour le sujet car c'est précisément cet emploi détrimental de *se faire* qui distingue le plus nettement *se faire* de son homologue portugais (*fazer-se Inf*) ou espagnol (*hacerse Inf*). En effet, ces derniers ne semblent pas avoir développé de sens non agentif. Ils révèlent donc peu d'affinités avec les verbes d'appréciation négative (Araújo, 2008), car il n'est pas du tout logique d'être responsable d'une action néfaste qui retombe sur soi-même:

- (1a) En novembre 2008, le rugbyman Shane Geoghegan *s'est fait descendre* devant sa maison. Une des quinze balles tirées par l'assassin l'a touché mortellement à la tête.
- (1b) Em Novembro de 2008, Shane Geoghegan, jogador de rugby, *foi assassinado* à porta de sua casa. Uma das quinze balas disparadas pelo assassino atingiu-o mortalmente na cabeça.
- (1c) En noviembre del 2008, el jugador de rugby Shane Geoghegan *fue abatido* a balazos a las puertas de su casa. Una de las quince balas que disparó el asesino le dio en la cabeza.

Ce qu'il faut retenir, en effet, c'est que les exemples avec *se faire* qui désignent des actions négatives pour le sujet ont tendance à être traduites par le passif périphrastique en portugais et en espagnol. Parmi les 88 exemples (puisés dans le corpus *Europarl – Opus*⁽²⁾) qui sont négatifs pour le sujet, 56 d'entre eux sont traduits, en portugais, par le passif (soit en 63,6% des cas). En espagnol, dans ce même corpus, sur ces 88 exemples, 33 d'entre eux ont été, eux aussi, traduits par le passif classique, soit 37,5%. Ces traductions suivent la division en deux interprétations faite ci-dessus, où les actions négatives ont un sens passif. Dans les exemples suivants, le verbe *hacerse volar* ou *fazer-se ir pelos ares* qui décrit un attentat suicide appartient au

(2) <http://opus.lingfil.uu.se/bin/opusqcp.pl?corpus=Europarl3>

domaine du désagréable, et pourtant l'emploi de *hacerse/fazer-se* inf est de mise:

- (2a) Notre ville a vu des auteurs d'attentat-suicide, des gens prêts à *se faire sauter* eux-mêmes, à semer la destruction par des attentats terroristes.
- (2b) Hemos tenido terroristas suicidas en nuestra ciudad, gente dispuesta a *hacerse volar* a sí misma, para destruir mediante actos terroristas.
- (2c) Tivemos bombistas suicidas à solta na nossa cidade, gente pronta a *fazer-se ir pelos ares* e a espalhar a destruição com ataques terroristas.

S'il est vrai que la lecture intentionnelle de *hacerse/fazer-se* est plus probable avec les procès bénéfactifs, elle n'en est pour autant exclue avec les procès détrimentaux à condition que le cotexte l'explícite. En effet, des exemples du type de (2a)-(2b) montrent que le détrimental n'implique pas nécessairement l'antitéléonomie. Nous n'arrivons pas à imaginer un plus haut degré d'intentionnalité et de contrôle que lorsque le sujet se force, comme ici, à faire quelque chose qui porte si brutalement atteinte à son intégrité physique.

2. Conditions d'emploi de *se faire INF*

Le corpus littéraire (CL dans le tableau 1 qui suit) auquel nous avons fait appel pour le français correspond à la base de données textuelles disponible en ligne: *FRANTEXT*. Quant au corpus représentant le registre journalistique (CJ dans le tableau 1), il a été construit grâce aux archives du journal *Le Monde Diplomatique*⁽³⁾ (ci-après, LMD) disponible sur CD-Rom⁽⁴⁾.

(3) *Le Monde Diplomatique* - LMD constitue, à présent, l'un des corpora du Corpus multilingue *Per-Fide*, partiellement financé par le projet PTDC/CLE-LLI/108948/2008 de la Fondation pour la Science et Technologie. Le corpus *Per-Fide* (Araújo et al., 2010) est composé d'un ensemble de sous-corpus couvrant plusieurs domaines (religieux, littéraire, juridique, journalistique et technique), sur un total de sept langues (Português, Español, Russian, Français, Italiano, Deutsch, English). Les textes insérés jusqu'à présent dans la base de données peuvent être consultés librement à partir d'un concordancier, encore à l'état expérimental (<http://www.per-fide.ilch.uminho.pt/query>), qui permet d'entrer des critères de recherche non seulement pour la langue source mais aussi pour la langue cible ou bien pour les deux à la fois et de visualiser rapidement l'élément recherché dans différents corpora simultanément.

(4) Le fait d'avoir deux registres (littéraire et journalistique) permet d'obtenir une analyse à plus large couverture pour le comportement de la construction en question. Comme nous le verrons, cette analyse permettra également de contraster certains emplois d'après leur registre d'appartenance.

2.1. Classement des occurrences de *se faire INF* dans la base textuelle FRANTEXT et celle de LMD

La recherche de toutes les occurrences de [N1 *se faire** INF (N2) (par N3)]⁽⁵⁾ dans la partie catégorisée de la base FRANTEXT entre 1985 et 2006 délivre 519 occurrences. Cette même recherche effectuée sur le corpus journalistique entre 1980 et 2000 en délivre 1279⁽⁶⁾.

2.1.1. Les verbes communs les plus fréquents dans les deux corpora

Les verbes les plus fréquents qui se recoupent dans les deux corpora sont regroupés dans le tableau qui suit:

Tableau 1: verbes dont le nombre d'occurrences est supérieur ou égal à 10 dans CJ et CL

SE FAIRE INF	CJ (901 occurrences de <i>se faire INF</i>)		CL (341 occurrences de <i>se faire INF</i>)		Corpus (verbes communs) (1242 occurrences de <i>se faire INF</i>)	
	Nombre d'occur- rences	Pourcen- tage %	Nombre d'occur- rences	Pourcen- tage %	Nombre d'occur- rences	Pourcen- tage %
entendre	165	18,31%	28	8,21%	193	15,53%
sentir	188	20,87%	13	3,81%	201	16,18%
remarquer	22	2,44%	12	3,52%	34	2,73%
passer	20	2,22%	11	3,23%	31	2,49%
tuer	14	1,55%	23	6,74%	37	2,98%
appeler	13	1,44%	10	2,93%	23	1,85%
Total	422	46,83%	97	28,44%	519	41,76%

Les données ont révélé que les fréquences les plus élevées de *se faire+ Vinf* dans les deux corpus correspondent à des expressions plus ou moins lexicalisées, qui fonctionnent comme une unité lexicale: *se faire entendre* (193

(5) La requête inclut toutes les formes fléchies de *faire* compatibles avec le pronom *se*.

(6) Le corpus général (littéraire et journalistique) comprend donc 1798 phrases employant *se faire INF*.

occurrences dans le CJ et CL), *se faire sentir* (201 occurrences), *se faire remarquer* (34), *se faire passer* (31), etc. Il est intéressant de constater que les verbes les plus fréquents dans les deux bases de données appartiennent à la classe des verbes de perception, bien qu'ils se répartissent très variablement au sein des deux corpora.

2.1.2. *Se faire-Inf* et les verbes de perception

Le verbe *sentir* est davantage représenté dans le CJ avec 188 occurrences (comparativement à 165 occurrences de *entendre*). À l'inverse, le CL utilise deux fois plus le verbe *entendre* avec 28 occurrences (comparativement à 13 occurrences de *sentir*). Bien que ces deux verbes de perception (indiquant soit une faculté olfactive ou tactile soit une faculté auditive) soient nettement plus représentés dans le CJ que dans le CL, il est bien clair que les deux corpora montrent le même attrait pour la correspondance entre la construction *se faire INF* et ces deux verbes décrivant une situation de nature sensorielle ou perceptive.

Les linguistes semblent d'accord pour dire que la construction en *se faire-Inf* exige un sujet animé, exception faite pour les infinitifs *entendre*, *sentir* (et *attendre*), qui acceptent un sujet inanimé. Il suffit pour s'en convaincre de consulter ces deux entrées verbales dans le *Trésor de la Langue Française Informatisé*⁽⁷⁾ (TLFi) disponible en ligne:

(7) <http://atilf.atilf.fr/>

Tableau 2: *se faire entendre/sentir* dans le TLFi

SENTIR, verbe trans.	ENTENDRE, verbe trans.
<i>Empl. factitif. Se faire sentir.</i> Se manifester, devenir sensible.	<i>Se faire entendre.</i> (Quasi-)synon. de <i>être entendu</i> (mais plus expressif, avec insistance sur l'effort ou sur l'événement qui se produit).
citations proposées dans le TLFi	
<i>La douleur, la faim, la soif se fait sentir; action, nécessité qui se fait sentir. Le découragement commençait à se faire sentir dans toutes les sphères de l'armée et même au grand quartier général (Joffre, Mém., t. 1, 1931, p. 319).</i>	<i>Enfin, un mouvement de pas se fit entendre en dehors du cachot (Hugo, Han d'Isl., 1823, p. 546).</i> <i>Mais vient un moment où un chant trop uniforme ne se fait plus entendre et où il faut un cri pour attirer l'attention (Mounier, Traité caract., 1946, p. 732).</i>
<i>En partic.</i> [Le suj. désigne un agent atmosphérique]	
Citations proposées dans le TLFi	
<i>Le froid se fait sentir. La bonne chaleur du soleil (...) avait commencé à se faire sentir dès le mois de mars (Ramuz, Gde peur mont., 1926, p. 14).</i> <i>Malgré l'approche du mauvais temps dont les premiers effets se faisaient déjà violemment sentir au débouquer, matelots et marchands faisaient cercle autour de l'unique mât (Cendrars, Bourlinguer, 1948, p. 11).</i>	

Se faire sentir est associé à des sujets qui peuvent décrire: soit une sensation (de soif ou de douleur) qui se manifeste, soit à un élément naturel/atmosphérique qui est présenté dans le TLFi comme une cause intervenant directement dans la survenue du stimulus perceptif:

- (3) A partir de l'année 1065, les eaux du Nil ne s'étant pas élevées à une hauteur suffisante, la famine commença à *se faire sentir*. (LMD, août 1988, page 6)

Quant à *se faire entendre*, il est intéressant de voir qu'il est mis en parallèle avec la construction passive canonique *être entendu* (cf. *supra*, tableau 2), ce

qui montre bien la valeur passive sous-jacente à ces constructions en *se faire* qui requièrent l'accompagnement de deux arguments (un expérienceur non spécifié dans l'énoncé et un stimulus de type sonore en position de sujet syntaxique): les citations proposées à titre d'exemple montrent que l'objet (= la cause) de la perception est un bruit dont on identifie plus ou moins bien l'origine (ex. *en dehors du cachot*) et la nature (ex. *un chant, ...*). On notera, tout de même, qu'il s'agit souvent de SN inanimés (*un mouvement de pas, un cri, une voix ...*), qui, par un emploi métaphorique ou métonymique, visent quand même des animés, ce qui semble corroborer la préférence de *se faire* pour les sujets à référent humain:

- (4) Le déclic se fit entendre et je raccrochai. (FRANTEXT, Aventin C./*Le cœur en poche*/1988, page 202)

Il est d'ailleurs étonnant que le TLFi n'ait proposé aucun exemple où *se faire entendre* est clairement associé à un sujet animé humain. Dans nos corpora, cette combinaison est, nous le verrons, fort fréquente:

- (5) Nicolas, assis à l'arrière, trouvait difficile de *se faire entendre* à cause du bruit de la soufflerie, poussée au maximum pour désembuer les vitres (FRANTEXT, Carrère, E./*La classe de neige*/1995, Page 7)

Outre les verbes *entendre* et *sentir*, on trouve, dans le CJ et le CL, d'autres verbes de perception. Le verbe *voir* qui relève, quant à lui, du domaine de la vision surgit, en effet, dans nos deux corpora mais il y est nettement moins représenté (avec 4 occurrences dans le CJ et 8 occurrences dans le CL) que les deux autres verbes de perception mentionnés plus haut. Selon Willems (2000: 172), il est possible de ranger tous les autres verbes de perception visuelle soit dans le sous-champ de *voir*, soit dans celui de *regarder*:

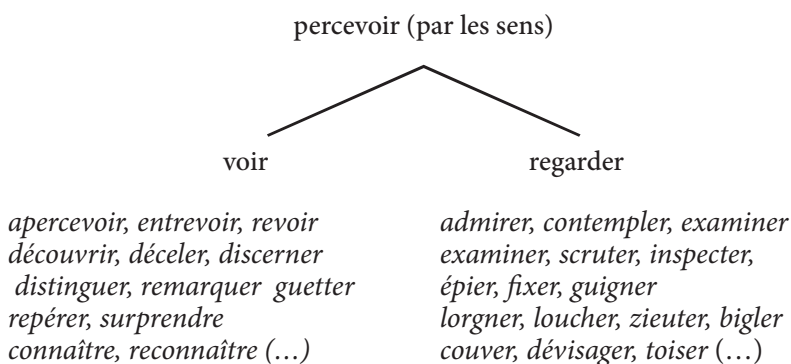


Figure 1: verbes de perception visuelle rangés dans le sous-champ de voir ou celui de regarder

Dans les deux corpora monolingues que nous avons consultés, il n'y a aucune attestation de verbes placés sous la tutelle de *regarder*. En revanche, *se faire* semble montrer une préférence assez marquée pour certains verbes appartenant au domaine de *voir*: le verbe *reconnaître* est davantage représenté dans le CJ avec 38 occurrences (comparativement à 22 occurrences de *remarquer*). À l'inverse, le CL utilise presque trois fois plus le verbe *remarquer* avec 12 occurrences (comparativement à 3 occurrences de *reconnaître*). *Reconnaître* apparaît ainsi au cinquième rang des verbes les plus employés dans le CJ, alors que dans le CL, le quatrième rang est occupé par le verbe *remarquer* que le TLFi présente comme suit:

Tableau 3: *se faire remarquer* dans le TLFi

REMARQUER, verbe trans.
Se faire remarquer. Se distinguer, attirer l'attention sur soi, en bien ou en mal.
Exemples proposés dans le TLFi
<i>Se faire remarquer par son courage, son dévouement, ses excentricités; le désir de se faire remarquer. Ces deux voitures à roues avant motrices, se sont fait remarquer notamment par leur carrosserie</i> (Tinard, <i>Automob.</i> , 1951, p. 369) À l'école préparatoire, je m'étais fait remarquer surtout par une expression perpétuelle de surprise, qui ne passe pas, à tort ou à raison, pour une marque de grande intelligence et me faisait juger un peu simple... A. France, <i>Pt Pierre</i> , 1918, p. 267

En fait, on s'aperçoit que *se faire remarquer* est interprété comme indéterminé du point de vue de l'assignation d'intentionnalité. Seul le contexte

nous permet d'interpréter le procès décrit comme intentionnel (s'il y a réellement le désir de se faire remarquer) ou comme non intentionnel (tel est visiblement le cas dans le deuxième exemple (cf. *supra*, tableau 3) qui met en scène un enfant qui attire involontairement l'attention sur lui). On notera, d'ores et déjà, que ce verbe semble parfaitement compatible avec des sujets à référent inanimé (ex. *ces deux voitures à roues avant motrice*) qui, de par leurs caractéristiques plus ou moins spéciales (ex. *par leur carrosserie*), parviennent à attirer le regard d'un observateur attentif. Nous n'avons trouvé qu'une seule occurrence de ce type dans le CJ:

- (6a) Ce dernier film *s'est fait remarquer* par l'intensité de son style réaliste. (LMD, mars 1985, page 24)

Dans nos deux corpora, *se faire remarquer* est majoritairement associé à des sujets humains:

- (6b) Voilà longtemps que Bourbogne essayait de *se faire remarquer*. (FRANTEXT, Ormesson, J. D'/Le bonheur à San Miniato/1987, page 94)
- (6c) ... La porte qui claque: il est à la traîne - quel con - et il trouve moyen de *se faire* encore *remarquer* en trébuchant à l'entrée! (FRANTEXT, Bayon/Le lycéen/1987, page 209)

La distinction humain/non humain distingue l'agent de la cause mais on observe un continuum dans la force d'agentivité exercée par l'humain selon qu'il déclenche plus ou moins volontairement le processus: l'exemple (6b) prend une valeur clairement active (car le sujet agit intentionnellement pour attirer le regard de son entourage), l'exemple (6c) semble, au contraire, proche d'un passif: le fait que le référent du sujet, déjà très en retard, trébuché (involontairement) à l'entrée fait qu'il est remarqué par le professeur et ses camarades de classe.

On trouve également dans les deux corpora d'autres verbes de perception appartenant au sous-champ de *voir*, mais en nombre beaucoup plus restreint: en effet, le verbe *repérer* apparaît 3 fois dans le CJ et 2 fois dans le CL; nous n'avons, en revanche, qu'une seule occurrence du verbe *surprendre* dans chacun des corpora.

On voit que, dans tout cela, se dessine une certaine opposition entre stativité et agentivité. Si l'on peut parler de stativité dans le cas de *voir* et *entendre*, cela est moins évident avec *regarder* et *écouter* qui sont supérieurs

en agentivité. Leur caractère agentif est dominant et repérable par le fait que le sujet prête attention à un objet visible ou audible (Ozouf, 2004: 4). En effet, comme le remarque, à juste titre, Franckel (1989: 420), ces deux verbes ont pour spécificité d'impliquer une cible: impossible d'écouter ou de regarder sans détermination préalable, à travers une intentionnalité, de quelque chose à écouter/regarder. Dès lors, avec *regarder* par exemple, la vision nécessite un effort alors qu'avec *voir*, la vision apparaît avec évidence (Ozouf, 2004: 4). *Regarder* est donc une vision délibérée en face de laquelle *voir* représente une vision involontaire. Il est intéressant de voir que *se faire* montre une préférence très marquée pour les verbes de perception reconnus comme non-agentifs. La grande fréquence de *se faire entendre* dans les deux corpora ne doit donc pas surprendre. En revanche, les deux corpora font un usage beaucoup plus restreint des verbes de perception à caractère agentif. On ne trouve, en effet, qu'une seule occurrence de *se faire regarder* dans le CL:

- (7) Quelqu'un, jadis, a construit un palais avec des balcons pour sortir, regarder, *se faire regarder*; (FRANTEXT, /Bianciotti, H./*Sans la miséricorde du Christ*/1985, page 222)

et seulement 3 occurrences de *se faire écouter* dans le CJ. En voici un exemple:

- (8) A coup sûr, ce serait une grande et heureuse révolution si la France parvenait à faire essaimer tous les ans deux ou trois mille esprits pourvus de connaissances politiques, ayant un titre pour *se faire écouter*, et des arguments pour faire comprendre que toutes les questions sont difficiles et la plupart des solutions complexes (LMD, mars 1999, page 24)

La combinaison *se faire entendre* apparaît, on le rappelle, 165 fois au sein du CJ, ce qui révèle une grande différence en ce qui concerne la productivité de ces deux verbes relevant du domaine de l'audition. Il est d'ailleurs intéressant de noter que les verbes *écouter* et *regarder* ne font pas partie de la liste de verbes communs aux deux corpora. Ils ne sont donc pas recensés dans le tableau 1. Comme nous le verrons plus en détail ci-dessous, *se faire* peut justement servir à conférer aux verbes intrinsèquement non-agentifs auxquels il s'associe une dimension agentive qui est bien visible dans des exemples du type:

- (9a) Il y avait plusieurs salons aux plafonds assez hauts, avec des buffets où des serveurs en veste blanche versaient à boire aux invités, qui étaient venus en foule *se faire voir* les uns des autres, (...) (FRANTEXT, Ormesson, J. D' /*La douane de mer*/1993, page 180)
- (9b) Et nous allâmes tous les deux, de conserve, côte à côte, sous le même parapluie, jusqu'à un de ces cafés de Saint-Germain-des-Prés où se retrouvaient, depuis la guerre, pour se cacher et *se faire voir*, les philosophes, les actrices, les éditeurs et les amoureux. (FRANTEXT; Ormesson, J. D' /*La douane de mer*/1993, page 54)

Ces deux exemples, issus du CL, sont intéressants: ils montrent que le verbe *voir* tend justement à prendre, grâce à *se faire*, le sens 'fort' de (*se faire*) *regarder*, signe de l'importance que prend l'action d'«être vu» pour le référent du sujet grammatical. Il est bien évident que ce dernier participe activement à la survenue du procès puisqu'il fait tout pour attirer l'attention sur lui. L'intentionnalité se présente alors comme une propriété majeure. Il est important de noter que *se faire* est compatible avec des prédicats non agentifs et peut leur ajouter de l'agentivité. Le seul ajout de *se faire* est capable, en effet, de «réorienter la signification d'un verbe faisant de *voir* un équivalent de *regarder*» (Ozouf, 2004: 24).

En revanche, dans les exemples qui suivent: le verbe *voir* perd son sens perceptif premier et il est bien clair que la dimension agentive présente dans les exemples précédents disparaît:

- (10a) Si j' avais l'âge requis, ce n' est pas Josette que je collerais comme une limace (elle pourrait aller *se faire voir*, cette geluche), c'est Josy. (FRANTEXT, Bayon/*Le lycéen*/1987, page 196)
- (10b) Que Mister Walsh et sa machinette aillent *se faire voir* ailleurs, ça ne les regarde pas. (LMD, mai 1986, page 8)

Se faire voir renvoie ici à l'expression fortement lexicalisée «aller se faire voir ailleurs» qui signifie: *se faire brutalement, désagréablement éconduire*. *Se faire voir* prend, dans ce cas, une valeur fortement péjorative qui est à rapprocher d'une autre expression que l'on retrouve également dans les deux corpora:

- (11a) Qu'il aille *se faire foutre*, rien à battre. (FRANTEXT, Bayon/*Le lycéen*/1987, page 341)

(11b) Dans un petit magasin, on vend une pancarte pour mettre dans son bureau: «En las horas de trabajo, las visitas al carajo» (pendant les heures de travail, que les visiteurs aillent *se faire foutre*). (LMD, mai 1982, page 8)

Il s'agit, on le voit, de l'expression *aller se faire foutre* dans laquelle le verbe transitif *foutre*, pris au sens figuré, vulgaire, s'emploie également pour signifier à quelqu'un qu'on ne veut plus le voir, qu'on ne veut plus entendre parler de lui. On peut donc constater que *se faire* peut s'associer à un même verbe pris soit dans son sens littéral, soit dans son sens figuré. Tel est le cas, on vient de le montrer, du verbe *voir* qui peut parfaitement se joindre à *se faire* sans son caractère sensoriel premier. Comme nous le verrons dans les pages qui suivent, en étudiant de plus près le vocabulaire argotique présent dans le CL, *se faire* semble, en effet, très friand de verbes employés dans le sens métaphorique.

2.1.3. Les verbes non communs aux deux corpora

Dans les pages qui suivent, nous nous proposons de poursuivre l'étude des restrictions sémantiques qui pèsent sur la construction en *se faire*, en prenant en considération les verbes non communs aux deux corpora que nous avons laissés volontairement de côté jusqu'à présent.

2.1.3.1. Verbes du CL et CJ rangés par ordre alphabétique

Dans le tableau qui suit, nous avons rangé par ordre alphabétique les 148 verbes recueillis dans le CJ:

Tableau 4: liste des verbes recueillis dans le corpus de LMD

(r)acheter	chasser	dénoncer	exciser	livrer	publier	respecter	titulariser
accoler	coiffer	dépasser	exempter	loger	rabrouer	ressentir	tourmenter
accuser	coloniser	désavouer	exploser	lyncher	racketter	restituer	trouer
admonester	compri- mer	dispenser	exposer	manger	rafler	retirer	tutoyer
agréer	concéder	doubler	exterminer	matra- quer	rançon- ner	retoquer	verser
agresser	condam- ner	duper	extorquer	menacer	recenser	retourner	vilipender
amputer	conférer	éconduire	faucher	mitrailler	recevoir	revolvé- riser	virer
analyser	confier	écouter	financer	nommer	recon- duire	rouer	voter
anéantir	consacrer	écrire	frapper	noter	réélire	rouler	
apostro- pher	conseiller	édifier	harpon- ner	nourrir	réem- ployer	rudoyer	
apprécier	considé- rer	éditer	héberger	obéir	régulari- ser	séques- trer	
arnaquer	conspuer	éjecter	implanter	octroyer	rejeter	sermon- ner	
arrêter	coopter	élever	imprimer	offrir	rembour- ser	shiner	
assister	coter	éliminer	infiltrer	ouvrir	remettre	soulever	
attraper	critiquer	embriga- der	interrom- pre	parrainer	rémuné- rer	stériliser	
attribuer	déborder	enfermer	jouer	plébis- citer	rendre	surnom- mer	
avalier	débou- lonner	enregis- trer	légaliser	prêter	repré- senter	taper	
avancer	décimer	entretenir	leurrer	procla- mer	répri- mander	taxer de	
blanchir	dédicacer	escroquer	licencier	projeter	réprimer	teindre	
cambrio- ler	dégom- mer	étriller	ligaturer	protéger	reprocher	tester	

Les 109 verbes du CL ont été regroupés dans le tableau suivant:

Tableau 5: liste des verbes recueillis dans le corpus FRANTEXT

blesser	crever	écorcher	greffer	miner	planter	rire	véhiculer
bourrer	croire	égorger	gronder	monter	poisser	rôtir	vider
branler	crucifier	embrasser	habiller	montrer	ramasser	sabrer	vomir
braquer	cueillir	enculer	happer	moucher	ramener	saloper	zigouiller
bronzer	cuire	enfoncer	hisser	mouliner	ramoner	secouer	
celer	décom- mander	engueuler	hospita- liser	mousser	ratatiner	serrer	
cha- touiller	décrire	entailler	ignorer	mutiler	rattraper	siffler	
chier	défoncer	essorer	indiquer	opérer	réciter	sortir	
choper	dégouli- ner	étaler	injecter	palper	recoudre	sucer	
clouer	dénumé- roter	expliquer	inoculer	peindre	réduire	tartir	
consoler	déposer	extraire	jouir	peloter	regarder	têter	
couillon- ner	dérouiller	flinguer	laver	péter	réintégrer	toucher	
couler	désarçon- ner	fourrer	lécher	pincer	remonter	tringler	
couper	désintoxi- quer	friction- ner	manoeu- vrer	piquer	rentrer	trombo- ner	
cracher	éclater	gauler	massacrer	plaindre	répéter	tromper	

Comme on peut le constater, le nombre des verbes recensés est considérable. On ne peut s'empêcher de constater, en survolant ces deux tableaux (les cellules en grisé), que les verbes dénotant une situation «désagréable» pour le référent du sujet occupent une part importante dans la liste des verbes qui sont associés à *se faire*. Les 95 verbes communs aux deux corpora montrent également la dimension détrimentale de cette construction, comme en témoigne le tableau qui suit:

Tableau 6: liste de verbes communs aux corpora FRANTEXT et LMD

(r)accompagner	avorter	descendre	faire	oublier	remplacer	tuer
abattre	balayer	désigner	foutre	pardonne	renverser	valoir
accepter	baptiser	désirer	hacher	passer	renvoyer	violier
acclamer	battre	dévor	huer	payer	repérer	voir
accorder	briser	écraser	humilier	photographe	saigner	voler
admettre	casser	embau-cher	incinérer	piéger	sauter	
aider	chapitrer	emmener	injurier	piétiner	sentir	
aimer	circoncire	engager	inscrire	porter	servir	
appeler	coincer	engloutir	insulter	prendre	soigner	
applaudir	comprendre	entendre	interpeller	présenter	surprendre	
arracher	conduire	enterrer	inviter	prier	tabasser	
assassiner	confisquer	envoyer	lire	reconnaitre	tailler	
attaquer	connaître	épingler	mettre	réformer	tatouer	
attendre	construire	exploiter	mordre	refouler	tirer	
avoir	craindre	expulser	naturaliser	remarquer	traiter	

2.1.3.2. La dimension 'détrimentale' de *se faire*

En ce qui concerne *se faire*, l'étude de notre corpus nous a permis de constater que les procès associés sont, dans une écrasante majorité, valués négativement pour le sujet affecté. En être 'affecté' signifie que ce sujet a peu de chances d'être l'origine volontaire de ce procès. Les cases en grisé des trois tableaux donnés ci-dessus signalent, en effet, les verbes qui décrivent, selon nous, une action néfaste pour le sujet de (*se*) *faire*. On notera qu'il s'agit, le plus souvent, de verbes renvoyant à une **agression de nature physique** (cf. *infra*, ex. (12a)), **verbale** (cf. *infra*, ex. (12b)) ou **psychologique/interpersonnelle** (cf. *infra*, ex. (12c)):

- (12a) Mais le soir, après 20 heures, les femmes n'osent plus se promener seules. Elles prétendent qu'elles *se font agresser* par des bandes de jeunes tziganes. Plaisanteries douteuses, grossièretés, vols à l'arraché. (LMD, octobre 1990, page 6)
- (12b) Il est riche, il a une antichambre chez sa belle-mère où il insulte les solliciteurs et une antichambre chez les ministres où il *va se faire insulter*. (FRANTEXT, Ormesson, J. D'/*La douane de mer*/1993, Pages 138-139)
- (12c) La sécurité, c'est le droit de ne pas *se faire voler* son portefeuille, ou de ne pas trop souffrir des embouteillages. (LMD, avril 1995, page 16)

Il semble, tout d'abord, que le français ait su se procurer une expression lexico-grammaticalisée, grâce au verbe *faire*, pour permettre la thématization de l'objet indirect (cf., *supra*, ex. (12c): [X *voler* Y à Z → Z *se faire voler* Y par X]⁽⁸⁾) qui n'a pas de forme passive comme en anglais⁽⁹⁾.

Il semble, par ailleurs, que *se faire-Inf* constitue un outil linguistique privilégié servant à exprimer la souffrance d'un être vivant, qui subit une action évaluée comme néfaste, inopportune ou «désagréable» et qui possède par conséquent le statut de victime. En effet, *se faire* apparaît fréquemment dans des contextes violents où l'on décrit une attitude destinée à nuire personnellement à **une autre personne** ou à **soi-même**. Ces verbes d'appréciation négative constituent près de 47% des verbes de l'ensemble du corpus LMD et près de 52% de la totalité des 109 types de verbes enchâssés sous *se faire+Vinf* relevés dans FRANTEXT).

(8) Il est bien évident que des verbes bitransitifs tels que *voler* ne peuvent pas entrer dans une construction passive avec *être* Vpp: **j'ai été volé mon portefeuille dans le métro* à comparer avec: *je me suis fait voler mon portefeuille dans le métro*. *Se faire* permet donc de construire un passif à partir de l'objet prépositionnel, ce qui prouve bien qu'«une même langue peut avoir plusieurs passifs» (Lazard, 1994: 254), inégalement grammaticalisés.

(9) Le déséquilibre entre le portugais et l'anglais est ici patent. Il semblerait, en effet, que ce dernier ne réduise pas, contrairement au français et au portugais, la passivabilité du verbe à sa transitivité directe puisqu'il accepte, bien volontiers, outre les passives «canoniques» du type de (1a): (1) a. *Mary gave a book to Peter*; b. *a book was given to Peter by Mary*, des constructions passives associées à des compléments d'objets indirects du type de (2b): (2) a. *Mary gave Peter a book*; b. *Peter was given a book by Mary*. S'il est vrai que «la plupart des langues ont des règles très strictes qui interdisent toutes l'emploi de l'objet au datif comme sujet d'une phrase passive» (Jespersen, 1977: 222), il est vrai aussi qu'en anglais, «on a de plus en plus tendance à réserver le rôle de sujet au terme qui désigne une personne [...] pour des raisons d'ordre émotionnel. C'est pourquoi on emploie facilement le passif pour dire *she was promised an apple*, «on lui a promis une pomme» ou *he was awarded a good medal*, «on lui a décerné une médaille d'or» (idem). En revanche, on sait qu'en français, seul l'objet direct est subjectifiable au passif, d'où, par exemple, l'agrammaticalité des séquences qui suivent: (3) **Marie a été volée ses bijoux*; (4) **une enquête a été procédée*; (5) **cette situation a été profitée* (exs de François, 1998: 17).

Comme le remarque très justement Jean-Marie Muller (2002: 27) «Faire violence, c'est faire mal, c'est faire du mal. Faire violence, c'est **faire souffrir**. Faire violence, c'est aussi se faire du mal et **se faire souffrir** (...)» (souligné par nous). *Se faire* semble rapporter aussi bien à des violences que l'on reçoit des autres (exs: *se faire insulter, se faire battre, se faire expulser, se faire giffler*, etc) que celles que l'on s'inflige à soi-même (exs: *se faire vomir, se faire maigrir, se faire exploser, se faire avorter*, etc.). On appelle violence tout ce que l'on reçoit comme tel. Il peut s'agir d'une parole (*se faire insulter, injurier, menacer, humilier*), d'un mensonge ou d'une trahison (*se faire arnaquer/escroquer/avoir*), d'une séparation, d'une rupture (*se faire plaquer/lâcher*)... Il peut s'agir de coups (*se faire battre/agresser*), de viols (*se faire violer*) ou autres tortures (*se faire maltraiter/torturer/étrangler*).

Faire preuve de violence, c'est tenter d'atteindre un but en utilisant la force physique ou psychique pour forcer quelqu'un, contre sa volonté. Mais pour que l'on parle de violence, il faut évidemment qu'il y ait l'intention de porter atteinte à l'intégrité physique ou morale de l'autre. Blessé quelqu'un par inadvertance n'est pas un acte de violence. Le frapper volontairement, même à cause d'une perte de contrôle, constitue au contraire une manifestation de violence. On acceptera bien plus volontiers l'emploi de *se faire* dans le deuxième cas (celui où la violence/blessure infligée est délibérée) que dans le premier cas où la blessure reçue est accidentelle. C'est ce qui expliquerait, pour Kupferman (1995: 68), le contraste d'acceptabilité entre les deux exemples qui suivent:

(13a) Sam *s'est fait blesser* à la hanche par son adversaire

(13b) *Paul *s'est fait blesser* à la guerre

Il semblerait, pour cet auteur, que la combinaison du verbe *blesser* avec la locution *se faire* ne soit possible que dans un contexte où l'entité qui exerce l'acte violent cause volontairement du tort à celui qui subit cet acte. Il serait donc tout à fait naturel de dire que si l'on pense que la guerre est réellement faite pour tuer:

(13c) Sam *s'est fait tuer* à la guerre

La notion de «**responsabilité**» du sujet de l'énoncé (cf. par exemple l'affirmation de Tasmowski-de Ryck *et alii* 1987: 49)⁽¹⁰⁾ est très souvent

(10) La «responsabilité» dont Tasmowski-De Ryck & van Oevelen (1987) parlent est néanmoins réfutée par certains linguistes (cf. Labelle 2002, par exemple) avec des exemples trouvés sur

proposée dans les études consacrées à *se faire*, pour rendre compte de la lecture causative non intentionnelle de *se faire* qui émerge d'une situation dans laquelle le sujet de l'énoncé cause sans en avoir eu l'intention, mais par simple maladresse, imprudence, négligence ou inattention, une situation qui lui est néfaste:

- (14) Madame Stein était inquiète et fatiguée. Elle n'avait plus de bonne. Celle qui travaillait chez elle depuis longtemps *s'était fait arrêter*, faute d'avoir un permis de travail valable pour Johannesburg. Elle *s'était fait expulser*: son passe n'avait de valeur que pour le secteur de Rustenburg (1) et, comme tant d'autres, on l'avait reconduite manu militari chez elle. (LMD, Septembre 1989, page 23)

Alors que dans un énoncé du type de (15) *il s'est fait applaudir pour ses talents d'orateur*, on peut reconstruire une intention du sujet syntaxique («*il s'est efforcé, en parlant, de convaincre ou séduire*»), même si cette intention reste implicite; en (14) l'adjonction des expressions soulignées pose le sujet comme «étant pour quelque chose»⁽¹¹⁾ dans la situation réalisée. Nous avons donc affaire à un schéma intermédiaire, où le sujet renvoie à un

Internet mettant en jeu des sujets inanimés qui échappent donc à ce critère. En effet, comme le fait remarquer Novakova (2009), les cas de sujet non animé, bien que très peu fréquents (moins de 3% des résultats), existent dans les corpora qu'elle a consultés pour rendre compte du fonctionnement de *se faire*. Elle cite les exemples suivants: (a) *La neige se fait désirer dans certains coins de l'Europe* (*Le Monde*); (b) *Les classiques cassettes vidéo VHS sont ainsi sur le point de se faire dépasser par les DVD* (*Le Figaro*). Pour cette linguiste, ici, ce n'est plus par le rôle sémantique du sujet (instigateur volontaire ou involontaire du procès) qu'on peut rendre compte des différentes nuances de sens entre la construction en *se faire Vinf* et le passif *êtreVé*, substituables dans ces contextes. Ce sont, à son avis, des paramètres aspectuels au service des visées discursives qui entrent en jeu. En choisissant *se faire*, le locuteur présente le procès comme inaccompli (se faire désirer) ou en déroulement (*être sur le point*), ce qui est en harmonie avec le profil aspectuel de *se faire*, le passif, lui, présentant le plus souvent le procès comme accompli.

(11) Ce schéma participatif de la part du sujet, c'est-à-dire son action antérieure à l'événement, Veacock (2008b) le nomme l'«avènement». Comme le signale cette linguiste, l'avènement peut aboutir à un événement ou à une suite d'événements prévus par le sujet (c'est la cas dans *Paul s'est fait livrer une nouvelle machine à laver*). Mais l'avènement peut également aboutir à un événement ou à une suite d'événements non prévus par le sujet (ce sont tous les cas de prédicats jugés «négatifs» tels que *Paul s'est fait renverser*). Pour cette linguiste (2008a: 2210), *se faire + Vinf* permet, en effet, d'identifier l'agentivité antécédente du sujet animé amenant ou justifiant des conséquences bénéfiques ou fâcheuses pour lui. Les notions de «cause», de «responsabilité», de «volonté» et de «faute» sont donc supplémentaires et n'accaparent pas l'agentivité primaire du sujet animé dans *se faire*.

«responsable»⁽¹²⁾. Mais en (15), on a la valeur «sujet bénéficiaire» et en (14), la valeur «sujet détrimentaire».

Ce type de construction construit cette dimension détrimentale en s'aliénant, le plus souvent, à des verbes aspectuels qui semblent comporter dans leur sémantisme le trait 'désagréable'. Tel est le cas du verbe *faillir* que les dictionnaires définissent en utilisant les paraphrases *être sur le point de* ou l'adverbe *presque*:

(16a) [...] elle était la seule qui avait failli *se faire violer* pour de bon. (FRANTEXT, Seguin, F./L'arme à gauche/1990, page 67)

En effet, comme le note Gaatone (1983: 163), «on serait étonné [...] de lire *il a failli survivre* ou *il a failli réussir à l'examen*, sauf, bien entendu, dans un sens ironique». Le verbe *risquer* semble avoir également de fortes affinités avec *se faire*:

(16b) Il risquait à chaque instant de *se faire tuer* pour ce qu'il détestait. (FRANTEXT, Ormesson, J. D'/Le bonheur à San Miniato/1987, page 108)

Les exemples que nous avons recueillis dans nos corpora montrent que l'emploi de *se faire* est possible dans ce contexte et cela perturbe bien souvent la sensibilité des locuteurs non francophones. Pour ces derniers, il est tout à fait intolérable de dire par exemple qu'une femme s'est fait violer car la violation correspond justement à une infraction qui implique l'absence de consentement de la victime. Chacun dispose, en effet, d'un droit absolu de choisir les termes qu'il juge appropriés à l'expression de sa propre pensée mais il faut bien avouer que l'expression française *se faire violer* est plutôt troublante, voire choquante: pour un locuteur natif du portugais ou espagnol, ce genre de construction supposerait inévitablement une intentionnalité difficilement envisageable ici. A moins de considérer, à l'instar de Sinner *et alii* (2005: 161), que «la différence, qui doit bien exister, entre *elle s'est fait violer* et *elle a été violée* semble ténue et tenir plus de la représentation qu'une langue donne de la réalité que de la réalité elle-même», il n'est pas

(12) Le repérage posant un «responsable» est glosable par: «il n'y a pas intentionnalité consciente du S₁ (= sujet syntaxique de *se faire*), mais la situation actualisée l'est en partie à cause du S₂». Pour Cottier (1985), la notion de responsabilité pose le S₁ comme «impliqué» dans le processus. Elle est grossièrement équivalente à celle de «déclencheur non-intentionnel».

simple, en effet, de comprendre ce qui peut bien amener la langue française à utiliser *se faire* pour décrire un acte d'une telle violence. Il semblerait que cette expression heurte tout autant les féministes françaises. Voici ce que nous avons trouvé sur l'une des pages du forum (au titre on ne peut plus révélateur) des chiennes de garde:

«Tu n'ignores pas que tu es sur le forum ouvert par une association qui pense que le langage est tissé d'idéologies, particulièrement l'idéologie machiste. Donc, à la forme employée «**se faire violer**», nous préférons dire et entendre «**être violé-e**». En fait, nous préfererions ne plus avoir à entendre cette expression et nous luttons activement contre.».

L'emploi de *se faire* dans ce cas peut paraître inopportun car il laisse entendre qu'en cas de viol, la femme est fautive, qu'elle était consentante⁽¹³⁾. Aussi n'est-il pas étonnant de constater qu'une telle expression se voit souvent taxée de misogynie, de machisme primaire. Il est intéressant de voir qu'une expression aussi profondément inscrite dans la langue française suscite, en réalité, de fortes réactions de la part de certains locuteurs. Ces derniers ont très certainement compris que la violence des mots entraîne les maux de la violence et qu'il serait fort pernicieux de vouloir «trouver de l'agent dans le chef de cette personne (qui est abusée sexuellement), dans la mesure où il y aurait une portion de factitif» (Sinner *et al.*, 2005: 163). Toute cette problématique ne se pose pas en portugais car cette langue a choisi de ne pas faire voisiner *fazer-se* et des mots décrivant des actes aussi abominables pour la femme.

De même, si on considère un exemple comme celui qui suit, il semble qu'il soit difficile de comprendre la sémantique de cette construction sur la base du sens propre du verbe *faire* car le sujet global de (17) «n'est certes pas responsable au sens normal, tant s'en faut, puisque l'événement lui arrive de manière totalement imprévue» (Kokutani, 2005: 216).

- (17) ... des personnes qui n'avaient rien à voir avec la manifestation, des passants, *se sont fait bousculer, pourchasser* à cheval et même *matraquer* dans certains cas» (Kokutani, 2005: 218)

(13) En effet, comme le note, très justement, Muller (2002 : 230): «Le verbe *se faire* permet aussi une orientation passive mais l'effet 'passif' est perturbé par l'interprétation plus ou moins 'volontaire' suggérée par *faire* ».

Ce sont des exemples de ce type qui amènent certains linguistes à souligner qu'il, existe, en effet, en français, un ensemble d'énoncés en *se faire* «qui s'interprètent dans un sens passif, et où l'idée de responsabilité, même très atténuée, ne paraît pas adéquate» (Tasmowski et Oevelen, 1987: 49). Pour rendre compte de ces énoncés où le référent du sujet est impliqué indépendamment de sa volonté dans un enchaînement causal dont il est uniquement le patient, Kokutani (2005) propose l'étiquette «PASSIF-FATALISTE», que Cottier (1985) gloserait très certainement de la manière suivante:

- (i) 'il est arrivé aux passants, malheureusement, qu'on les a bousculés, pourchassés, matraqués'

Dans ce type d'énoncés, *se faire* perd sa valeur causative et semble fonctionner davantage comme un simple opérateur de localisation de l'événement par rapport au sujet:

- (ii) 'il y a localisation d'une situation par rapport aux passants, ces derniers apparaissant comme détrimentaires du processus.'

Mais pour Kokutani (2005), dans ce type d'exemples, le sujet est toujours en cause parce qu'il est là, parce qu'il est «localisé dans le contexte, dans la progression des événements» (p. 216). Pour ce linguiste, la notion de «responsabilité» du sujet «parce qu'il est là» (2005: 215) lors d'un événement serait mieux expliquée par cette notion de «caractérisation causale»⁽¹⁴⁾. Nous avons bel et bien affaire à un enchaînement de cause à effet: hasard, fatalité de la rencontre, pourquoi cette victime et pas une autre, c'est la personne qui passait par là au mauvais moment. Sa seule responsabilité est d'avoir été dans un lieu à connotation négative propice à l'agression, au mauvais moment.

(14) Comme le précise à juste titre Veacock (2008a: 2210), «utiliser «l'instigateur» ou «cause première» (Tesnière, 1959) ainsi que «responsable» pour qualifier le rôle du sujet animé dans des événements «désagréables» est pervers dans les cas de *se faire violer*, *se faire séquestrer*, etc.». La notion de «caractérisation causale» proposée par Kokutani permet d'expliquer l'emploi de *se faire* dans des exemples comme (17) sans recourir à la notion de «responsabilité du sujet» qui reste difficilement démontrable dans les procès «désagréables».

2.1.3.3. Le registre «familier» du tour *se faire-Inf*

Si l'on considère, à nouveau, les trois derniers tableaux donnés ci-dessus, on s'aperçoit que *se faire* s'allie aussi bien à des verbes qui désignent des actes officiels qu'à des registres que certains spécialistes (cf., par exemple, Cellard *et al.*, 1991) regroupent sous la mention *non conventionnel*, c'est-à-dire ceux qu'on considère, en général, comme marginaux par rapport à la norme linguistique: il s'agit donc essentiellement des mentions *fam.*, *arg. pop.* et *vulg.* des dictionnaires.

On ne peut s'empêcher de noter que les deux corpora se distinguent assez nettement quant au type de registre de langue adopté: alors que les verbes extraits du CJ relèvent, pour la plupart, d'un langage de spécialité lié à la politique (*se faire réélire, voter, coopter, représenter, proclamer, parrainer, plébisciter, nommer, dépasser*), au droit et à l'administration (*se faire chasser, condamner, admonester, enregistrer, légaliser, recenser, régulariser, accuser, exempter, titulariser, retirer, concéder, octroyer, agréer, ...*), ou bien encore à l'économie ou au commerce (*se faire financer, coter, rembourser, rémunérer, etc.*), ceux du CL relèvent, dans la plupart des cas, du registre argotique qui multiplie à volonté les situations obscènes (*se faire tringler, trombonner, planter, défoncer, bourrer, fourrer, sucer, ...*), scatologiques (*se faire chier, tartir, etc.*), voire criminelles (*se faire massacrer, égorger, écorcher, mutiler, flinguer, révolvériser, mitrailler, zigouiller, ...*).

La tournure en *se faire* n'hésite donc pas à sortir des domaines impartis à la fonction officielle (droit, administration, ...)⁽¹⁵⁾ pour s'épanouir dans la rue et ses lieux de loisirs (écoles, stades, bars, restaurants, cités ...). Cette construction est donc un terrain privilégié pour l'observation du passage d'un registre de langue à un autre. Ce passage d'un registre à l'autre à l'intérieur d'une même langue renvoie aux variétés de cette langue, c'est-à-dire à ses différents usages. Il semblerait que l'argot s'infilte partout, même dans des contextes semi-officiels, comme en témoignent les verbes suivants que nous avons trouvés dans le CJ, et où l'appréciation portée est "négative":

- (18) Elle disait qu'elle avait découvert un Franz inconnu, qui ne supportait pas de perdre. Il était pâle et buté. Et soudain il attrapait la calotte de

(15) Après avoir observé un échantillon de données extraites d'articles de *Le Monde* (publiés en 1999 et 2000) pour rendre compte des conditions de réalisation de *se faire* et de *se voir*, Glawogger (2001) arrive, en effet, à la conclusion que les tours en *se faire* se distinguent très nettement de ceux en *se voir* en ce sens qu'ils apparaissent pour la plupart dans des circonstances défavorables pour le sujet grammatical et qu'ils y sont effectivement utilisés pour reproduire un français de registre parfois très familier.

- Shavel, la tendait vers les copains, Une petite aumône m'sieu-dames. Pour un pauv' mec qui *s'est fait arnaquer*. (LMD, août 1990, page 18)
- (19) Kitunda élabore une tactique rationnelle pour s'en emparer, mais ses milliers de soldats, emportés par une frénésie aveugle, *se font faucher* par le feu des mitrailleuses. (LMD, septembre 1982, page 23)
- (20) Comme celui d'envisager la mort d'un autre cadre qui a l'emploi qu'on voudrait (qu'on devrait!) avoir: «Upton Ralph Fallon avait mon boulot. Et s'il *se faisait virer*, s'il tombait trop malade pour continuer à travailler et s'il mourait? Ne pouvais-je pas le tuer? Pour défendre ma famille, ma vie, mon crédit. Par autodéfense en réalité.» (LMD, mars 1999, page 30)
- (21) Tailleurs béninois vendant des pagnes, toutous ghanéennes vendant leurs charmes, colporteurs dioulas vendant de tout, ce monde hétéroclite vit en bonne intelligence avec pour seule crainte celle de *se faire rafler* au cours des nombreuses opérations de ratissages de la gendarmerie ou de l'armée. (LMD, août 1982, page 27)
- (22) Plus césarien que César, plus royaliste que le roi! Ça finissait par le fourrer dans un mauvais cas, le Poncepi, s'il n'était pas à la hauteur. Pouvait *se faire dégommer*. (LMD, avril 1988, page 28; 29)
- (23) Dès lors, il faut trouver des boucs émissaires, à l'intérieur (les opposants sont passésistes, arriérés et froussards) et à l'extérieur. Ici, la France a été plus particulièrement dans la ligne de mire. Mais elle n'a pas été la seule, contrairement à l'idée reçue. Dans le dernier épisode, ce fut aussi à l'Europe de *se faire étriller*. (LMD, janvier 1994, page 14; 15)

Pour le TLFi, les verbes présentés dans les exemples ci-dessus appartiennent, en effet, à des registres de langue non standard: (*se faire*) *arnaquer* est argotique, (*se faire*) *faucher* appartient au registre populaire, (*se faire*) *virer*, (*se faire*) *étriller* et (*se faire*) *dégommer* relèvent du registre familier. Les deux derniers verbes sont à prendre, en outre, au sens figuré. Le niveau lexical est celui qui permet le plus facilement de distinguer les registres de langue par le jeu de la synonymie (ex.: *se faire dérober, voler, cambrioler, escroquer, arnaquer, faucher, rouler, ...*). Il suffit pourtant d'ouvrir des dictionnaires pour se rendre compte assez vite que poser des frontières est problématique: on note, par exemple, qu'un même mot n'est pas forcément étiqueté de la même manière, «classé» dans le même registre: tout dépend de la politique linguistique du dictionnaire, de son discours par rapport à la norme. Si on reprend, par exemple, le verbe *arnaquer* mentionné ci-dessus, on s'aperçoit qu'il est «familier» pour *Le Nouveau Petit Robert* (2007:

140)⁽¹⁶⁾, «populaire» pour le *Dictionnaire de Notre Temps* (1991: 87) alors que LTF le trouve «argotique». Outre les argots parlés au sein de différents groupes socio-professionnels, nous ne devons donc pas ignorer l'importance d'une variété argotique utilisée au niveau de toute la société, dont le vocabulaire finit par passer dans la langue familière, voire dans la langue courante tout en conservant une "nuance" argotique. C'est exactement ce que nous dit Červenková (2001: 78) dans son article portant justement sur l'influence de l'argot sur la langue commune:

Il faut prendre en considération que l'argot a parcouru, pendant les siècles, un long chemin d'évolution et qu'il a considérablement influencé la langue commune, normale, et, dans certains cas, a pénétré dans les autres niveaux de la langue: par le français populaire et familier jusqu'au français littéraire.

L'existence de pratiques argotiques apparaît comme une constante des langues (Gadet, 2002: 5). Alors que certains mots sont restés purement argotiques car ils n'ont pas franchi justement la frontière *argot-langage courant* et ils restent ainsi incompréhensibles pour la majorité des gens, d'autres mots ont, eux aussi, commencé leur carrière dans l'argot «fort» mais se laissent peu à peu apprivoiser en entrant, en dépit des réticences puristes, dans un domaine plus vaste, celui d'une familiarité courante. Les dictionnaires perçoivent, en général, ces nuances et c'est pourquoi ils qualifient les mots de façon soit argotique (ARG), soit familière (FAM), soit populaire (POP). À cet égard, l'entrée de ces mots dans la langue standard marque leur sortie de la sphère étroite de leur communauté d'origine. Selon Červenková (2001: 78), «l'utilisation des éléments argotiques est parfois causée par l'état immédiat de l'usager, qui veut exprimer sa colère, son refus ou son dédain. En effet, le vocabulaire argotique est capable d'exprimer les nuances entre divers sentiments des hommes, qu'ils soient négatifs ou positifs». Il n'est pas étonnant que ce registre de langue ait ses lettres de noblesse chez certains écrivains. Comme le note, à ce propos, Sourdot (2006: 189) dans son article consacré à la problématique de «l'intégration stylistique de l'argot dans le roman contemporain», il est possible d'envisager diverses possibilités d'utilisation de l'argot à des fins stylistiques:

«Entre le texte purement argotique (et le risque d'hermétisme y afférent) et l'œuvre parsemée çà et là de tournures argotiques, entre les ballades en argot

(16) Mais dans l'édition bien plus ancienne de *Le Petit Robert* (1969: 90), *se faire arnaquer* est considéré comme «populaire».

de François Villon ou la pièce de théâtre «Matou de Pantruche» de Gérard Legrand et les romans de Darien ou de Zola, simplement parsemés de tournures argotiques, toutes les options sont ouvertes».

Selon cet auteur, c'est ce critère quantitatif qui peut servir à différencier «l'argot dans la littérature» et «la littérature en argot» (distinction reprise à François, 1975). Balzac en fait un usage assez prudent (dans *le Père Goriot*, *les Illusions Perdues*, *Splendeurs et misères des courtisanes*, ...) ainsi que Victor Hugo qui intitule *l'Argot* le VIIe livre de la IVe partie de son roman *Les misérables*, 1862. Du côté de la prose, au XXe siècle, certains auteurs en font un usage bien plus intensif: Frédéric Dard, l'auteur de *San Antonio* ou Albert Simonin prennent, en effet, leurs lecteurs pour de réels virtuoses de la langue verte. Chez Raymond Queneau (*Zazie dans le métro*, *Les fleurs bleues...*), les mots d'argot n'y manquent pas, et Louis Ferdinand Céline donne, lui aussi, une coloration argotique à ses romans, moins par amour du peuple, dont il est issu, que par dégoût d'une certaine société. Il semblerait que les auteurs représentés dans la base textuelle FRANTEXT n'hésitent pas non plus à combiner *se faire* à une multitude de verbes argotiques dont la mise en situation est, parfois, le seul moyen d'éclairer le lecteur sur leur signification. Tel est le cas du verbe *torcher* dans l'exemple suivant:

- (24) Nous étions pour les républicains: ils *se sont fait torcher*. (FRANTEXT, Ormesson, J. D'/Le bonheur à San Miniato/1987, Page 221)

qui est visiblement utilisé dans le sens de «se faire battre». On notera qu'un tel usage du verbe *torcher* est ressenti comme FAM. et VIEILLI dans *Le Nouveau Petit Robert* (ci-après, LNPR) (2007: 2574), ce qui signifie que ce verbe employé dans cette acception est encore compréhensible de nos jours, mais ne s'emploie plus naturellement dans la langue parlée courante; dans le *dictionnaire de l'argot français* (p. 803), un tel emploi du verbe est précédé de l'abréviation *Vx.* (vieux), ce qui veut bien dire qu'il s'agit d'un emploi désuet, obsolète (qui n'est donc plus en service). Dans un tel contexte, on comprend pourquoi il y a un renouvellement constant des verbes argotiques qui sont rapidement usés. C'est ce qui explique l'importante **polysémie** et l'importante **synonymie**. Un verbe a très souvent plusieurs acceptions (que le contexte différencie). En français familier, (*se faire*) *rouler* par exemple signifie à la fois «se faire avoir» et dans un sens plus étroit «se faire voler» (ex. emprunté à LNPR (2007: 2275) : «*c'est bien trop cher, vous vous êtes fait rouler*»); le verbe (*se faire*) *baiser* sert à désigner l'action de «se faire trom-

per» (ex. de LNPR 2007: 209: «*il s'est fait baisé*») de même que le fait de «se faire posséder (sexuellement)» (ex. «*les femmes, c'est juste bon à se faire baisé*» (Beauvoir)). On constate, par ailleurs, qu'on dispose de plusieurs verbes pour un référent déterminé. L'argot offre, en effet, des gammes de synonymes pour désigner une même réalité. Par exemple au lieu des verbes (*se faire*) *arrêter* (cf. *supra*, ex. 14), (*se faire*) *prendre* ou (*se faire*) *attraper*:

- (25a) Ludo faillit *se faire prendre* en déminant la literie piégée par Tatav. (FRANTEXT, Queffelec, Y./*Les nocés barbares*/1985, page 86)
- (25b) Implicitement, on se repose sur l'idée que ceux qui volent beaucoup finiront bien par *se faire attraper*: il suffirait alors de les mettre hors d'état de nuire par un emprisonnement assez long pour régler peu à peu le problème. (LMD, juin 1988, page 26; 27)

qui relèvent plutôt du registre standard, on peut employer des synonymes argotiques, comme en témoignent les exemples suivants que nous avons recueillis dans le CL:

- (25c) Le jour où leur envie de se laver et de dormir dans un lit devint irrésistible, ils rentrèrent chez eux et *se firent cueillir* aussitôt. (FRANTEXT, Rolin, J./*L'organisation*/1996, Page 70)
- (25d) C'est là que *se fit* par exemple *épingler* à plusieurs reprises l'un des délinquants les plus bêtes de S., Ringo, un type au long et lourd visage criblé de pustules, qui, entre deux casses invariablement loupés et deux séjours en prison, tirait gloire de gagner sa vie en «ramonant des vioques». (FRANTEXT, Rolin, J./*L'organisation*/ 1996, Page 76)
- (25e) Un pillage lamentable... Et en traînant dans le magasin contrairement aux directives... Ces rigolos ont failli *se faire coincer* au moment où ils emmagasinaient le magasin dans leurs valises. (FRANTEXT, Bayon/*Le lycéen*/1987, page 270)
- (25f) On savait s'y prendre de manière à ne jamais *se faire poisser*. (FRANTEXT, Bayon /*Le Lycéen* /1987, Page 44)
- (25g) Pauvre Nabokov, dit Cecilia, il a fini par *se faire piquer*. (FRANTEXT, Sollers, P. /*Le cœur absolu*/1987, page 342)
- (25h) Or c'est ce que précisément, paraît-il, s'aventurait à élucubrer le dénommé Bogdanov avant de *se faire ramasser* par Lénine! (FRANTEXT, Kristeva, J./ *Les samourais* /1990, page 80)

- (25i) Durant mon été de totale délinquance, j'ai vu des chats (et des fûtés pourtant) *se faire choper* et *blesser* grièvement par des pièges tendus que pour nous. (FRANTEXT, Forlani, R. /*Gouttière*/ 1989, Page 272)
- (25j) Je décide de pas rester là, un jour on va *se faire serrer*, les poulets on les voit jamais mais ils sont partout. (FRANTEXT, Belloc, D. /*Kepas*/1989, Page 139)
- (25k) Mets ta ceinture, elle a dit, on va *se faire gauler* par ces pourris de gendarmes... (FRANTEXT, Pouy, J-B/*La clef des mensonges*/1988, Page 77)

La richesse synonymique de l'argot s'explique par le caractère essentiellement émotif, affectif de ce langage; elle est aussi en fonction directe de son renouvellement rapide. Il y a de ce point de vue là une grande liberté et inventivité: le lexique argotique est, en effet, riche en jeux sur les signifiés qui ne sont autres que les tropes de la rhétorique, qu'il s'agisse de métaphores impliquant une comparaison (ex. *se faire faucher* pour se faire voler, *se faire entuber* pour se faire duper) ou de métonymies qui prennent le plus souvent l'effet pour la cause (ex. *se faire descendre* pour se faire tuer, *se faire suer* pour se faire importuner). Il est intéressant de voir qu'un verbe appartenant au registre courant prend un sens figuré en argot⁽¹⁷⁾. Tel est le cas du verbe *moucher* qui désigne dans l'exemple ci-dessous l'action de «se faire remettre vertement à sa place, se faire réprimander» (LNPR 2007: 1643):

- (26a) A chaque coup, il essaie de jouer les terreurs, mais avec les routiers, tu penses, il *se fait moucher*! (FRANTEXT, Thérôme, V./*Bastienne*/1985, Page 85)

On remarquera, une fois encore, que le sens de «réprimande» est rendu, dans nos deux corpora, par une multitude de verbes qui relèvent de registres différents:

- (26b) La solution est simple: désintéressez-vous de la politique ainsi comprise, cessez donc de voter et vous vous épargnerez de tels tracas... Le citoyen qui agira ainsi *se fera* vertement réprimander. (LMD, août 1989, page 1)
- (26c) Ce qui lui a valu de *se faire gronder* dans la langue barbare que parlait Maria. (FRANTEXT, Forlani, R./*Gouttière*/1989, Page 357)

(17) Dans LPR (2007: 1643), *se faire moucher* est considéré comme FIG. FAM. Mais on trouve cette expression dans le dictionnaire de l'argot français et ses origines (2001: 537).

- (26d) On a pu voir à la télévision un évêque du Sud *se faire chapitrer* par sa vieille mère du Nord, qui lui reprochait de croire au paradis, «alors qu'il existe chez nous, grâce à Kim II-sung, grand leader...» (LMD, avril 1986, page 13)
- (26e) Déjà, cependant, les spectateurs, las de *se faire sermonner*, désertaient les salles militantes. (LMD, septembre 1982, page 25)

Alors que les 4 verbes qui précèdent relèvent du registre courant, ceux qui suivent sont classés comme FAM. dans LNPR:

- (26f) ... elle a jamais dit qui c'était mon père et puis c'est pas vrai tout ça... même qu'elle a dû dire à Micho faut pas y aller même qu'elle a dit faut faire vite et qu'ils ont dû *se faire engueuler*... (FRANTEXT, Queffelec, Y./*Les noces barbares*/1985, page 193)
- (26g) La plus grosse, un peu penaude, *se faisait remonter les bretelles*, parce qu'elle n'avait rien vu. (FRANTEXT, Brisac, G. /*Week-end de chasse à la mère*/1996, Page 53)

Se faire remonter les bretelles fait partie des multiples locutions verbales qui se construisent avec *se faire*. Pour véhiculer le sens sous-jacent à une telle locution, le français met à notre disposition d'autres expressions du type: *se faire sonner les cloches, se faire secouer les puces, se faire tirer les oreilles, se faire passer un savon, se faire clouer le bec* que LNPR 2007 classe dans le registre familier, bien qu'elles soient largement utilisées et bien connues par une grande partie des membres de la communauté linguistique.

3. Considérations finales

Ce qui est important est que ces exemples sont (presque) tous détrimentaires, en ce sens que les événements violent un tabou et sont donc considérés négativement.

Il est fort probable que, même avec l'appui du contexte et de la situation, le lecteur non-initié aux tournures argotiques puisse se sentir dérouté devant cette accumulation de verbes familiers, populaires et argotiques qui jalonnent notre CL et dans une bien moindre mesure notre CJ. Avec *se faire*, l'écrit et le parlé, le littéraire et le journalistique, le formel et l'informel, le «grossier» et le «sublime» se côtoient naturellement. C'est ce qui fait la vitalité de cette construction qui, contrairement à ses homologues ibéri-

ques (*fazer-se/hacerse* + Vinf), s'est spécialisée dans la construction d'une diathèse «maléfactive», «détrimentaire» renvoyant à un 'mode participatif atélique' (Veacock, 2008b: 18) du sujet.

Nous avons pu constater, à travers l'exploitation de nos deux corpora, que ce type de construction peut renvoyer à une interprétation réfléchie (de type: «le référent du sujet a agi sur lui-même», cf. *supra*, exs. (2a)-(2c)) ou préférentiellement à une relation (de type: «le sujet a été affecté, en bien ou en mal, par un processus le concernant») renvoyant à un bénéficiaire (lorsque l'action est positive pour le sujet) ou à un détrimentaire (si la situation lui est néfaste) qui peut correspondre aussi bien à un objet direct (cf. *supra*, exs. (25a)-(25k)) qu'à un objet indirect (cf. *supra*, ex. (12c)). Nous avons également montré que ce sujet détrimentaire peut, bien souvent, être tenu pour responsable de ce qui lui arrive (on obtient dans ce cas une signification du type «être par sa propre faute le patient du procès»), même lorsqu'il est impliqué dans un événement *a priori* indésirable (comme dans *en essayant de rattraper son ballon, l'enfant s'est fait renverser par une voiture*). Il semblerait donc que certaines constructions en *se faire* soient plus proches du pôle actif (si le sujet agit sur lui-même pour son bénéfice ou son détriment (ex. *il a trouvé le moyen de se faire renvoyer*)), et d'autres plus proches du pôle passif (si le sujet subit un sort funeste (ex. *il se promenait tranquillement quand il s'est fait tirer dessus*)).

Il est bien clair que ce n'est pas seulement la syntaxe de *se faire* qui provoque la domination, plus ou moins accusée, du pôle actif ou passif mais que d'autres paramètres (notamment d'ordre extra-linguistiques) jouent un rôle prépondérant. Avec *se faire*, «la volonté du sujet peut se manifester nettement [...] ou au contraire s'effacer totalement, au point que ce sujet peut être interprété comme un patient qui subit l'action sans rien y pouvoir» (Blanche-Benveniste, 2007: 164). Le lexique semble être le véritable structeur de l'interprétation, même s'il est vrai que des indices contextuels peuvent annuler la valeur préférentielle normalement induite par le type de procès, au profit de la valeur opposée (cf. *supra*, exs. (2a)-(2c), les cas de faux détrimental). Le travail du linguiste consiste alors «à démêler la proportion variable d'activité et de passivité et de dégager les facteurs lexicaux, grammaticaux ou contextuels qui influent sur l'effet de sens» (Melis, 1990: 30).

Références bibliographiques

- ARAÚJO, S. (2008), *Entre l'actif et le passif: se faire/fazer-se. Syntaxe, sémantique et pragmatique comparées français-portugais*, Thèse de Doctorat, Universidade do Minho/ Université Paris 7 – Denis Diderot.
- ARAÚJO, S., J. J. Almeida, I. Dias, A. Simões (2010), «Apresentação do projecto *Per-Fide*: Paralelizando o Português com seis outras línguas», *Linguamática*, v2 n2, pp. 71-74.
- BLANCHE-BENVENISTE, C. (2007), «Les énoncés à causatifs réfléchis», dans Rousseau André; Bottineau Didier & Roulland Daniel (éds.), *L'énoncé réfléchi*, Rennes, Presses Universitaires de Rennes, pp. 155-173.
- CELLARD, J. ; Rey, A. (1991), *Dictionnaire du français non conventionnel*, Paris, Hachette.
- ČERVENKOVÁ, M. (2001), «L'argot, son influence sur la langue commune et les procédés de sa formation», in *Opera romanica 2*. Jihočeská univerzita, České Budějovice: pp. 49-55.
- COTTIER, E. (1985), *De quelques verbes causatifs anglais et français en tant qu'opérateurs et types de repérages*. Thèse de Doctorat de 3^e cycle sous la direction d'Antoine Culioli. Université Paris VII.
- CREISSELS, D. (2003-2004), *Cours de syntaxe générale*, [Accessible en ligne à l'URL: http://lesla.univ-lyon2.fr/article.php3?id_article=562].
- FRANCKEL, J.J. (1989), *Étude de quelques marqueurs aspectuels du français*, Genève-Paris, Droz.
- FRANÇOIS, D. (1975), «La littérature en argot et l'argot dans la littérature», *Communication et langages*, n° 27, Paris, Retz, pp. 5-27.
- FRANÇOIS, J. (1998), «La passivité des objets: rôles prototypiques et transitivité», *Travaux de linguistique* 35, pp. 21-37.
- FRANÇOIS J. (2000), «Désémantisation verbale et grammaticalisation, (se)voir employé comme outil de redistribution des actants», *Syntaxe & Sémantique* N° 2, pp. 159-175.
- GAATONE, D. (1983), «Le désagréable dans la syntaxe», *Revue romane*, 18 (2), pp. 161-174.
- GADET, F. (1992), *Le français populaire*, Paris, PUF.
- GLAWOGGER, I. (2001), *Verbalperiphrasen mit passivischer Diathese (se faire/se voir + Infinitiv/Partizip II) – Korpusanalyse anhand der Tageszeitung Le Monde*.
- JESPersen, O. (1977), *La philosophie de la grammaire*, Paris, Les Éditions de Minuit, coll. Arguments.

- KOKUTANI, S. (2005), «Sur l'analyse unie de la construction 'se faire + infinitif' en français», dans Bat-Zeev Shyldkrot Hava & Le Querler Nicole (éds.), *Les Périphrases Verbales, Lingvisticae Investigationes, Supplementa 25*, pp. 209-227.
- KUPFERMAN L. (1995), «La construction passive en 'se faire'», *Journal of French Language Studies*, 5, pp. 57-83.
- LABELLE M. (2002), «The French non canonical passive in *se faire*», dans Haraguchi, Shosuke *et al.* (éds.), *Proceedings of Linguistics and Phonetics 2002*, Tokyo, Charles University Press and Meikai University.
- LAZARD, G. (1994), *Lactance*, Linguistique Nouvelle, Paris, PUF.
- MELIS, L. (1990), *La voie pronominale*, Champs linguistiques, Duculot, Paris - Louvain-la-Neuve.
- MULLER, C. (2002), *Les bases de la syntaxe – syntaxe contrastive, Français – Langues voisines*. Pessac, Presses universitaires de Bordeaux.
- MULLER, J.-M. (2002), *De la non-violence en éducation*, Paris, UNESCO, Ouvrage publié avec le concours de l'institut de recherche sur la résolution non-violente des conflits (IRNC), [Disponible sur: <http://unesdoc.unesco.org/images/0012/001272/127218f.pdf>].
- NOVAKOVA, I. (2009), «La construction *se faire*+Vinf: analyse fonctionnelle», dans *La Langue en contexte. Actes du Colloque international «Représentation du sens linguistique (RSL IV)»*, Helsinki, 28-30 mai 2008, pp. 107-120.
- OZOUF, C. (2004), «Caractère différentiel et relation d'équivalence entre *voir* et *regarder*», *Cahiers du CRISCO numéro 16*, [Accessible en ligne: <http://elsap1.unicaen.fr/cahiers/cahier16.pdf>].
- ROGGERO, J. (1984), «Le passif, le causatif et quelques autres formes étranges», Cercle Linguistique d'Aix-en-Provence, *Travaux 2, Le passif*, Publications de l'Université de Provence, pp. 35-37.
- SINNER, C., D. Van Raemdonck (2005), «Faire» et «se faire», c'est toujours faire? Le médio-factitif en espagnol et en français», in Carsten Sinner & Georgia Veldre (éds.) *Diathesen im Französischen / Les diathèses en français*. Frankfurt am Main *et al.*: Lang, pp. 155-175.
- SOURDOT, M. (2006), «L'intégration stylistique de l'argot dans le roman contemporain», *Revue d'Études Françaises* n° 11, pp. 189-196.
- SPANG-HANSEN, E. (1967), «Quelques périphrases passives du français moderne». *Revue Romane*, N° spécial 1, *Actes du 14e Congrès des Romanistes Scandinaves dédiés à Holger Sten*, Copenhague, Akademik Forlang, pp. 139-147.
- TASMOWSKI-DE RYCK L. & VAN OEVELEN HILDEGARD (1987), «Le causatif pronominal», *Revue romane*, 22 (1), pp. 40-58.
- TESNIÈRE, L. (1959), *Éléments de syntaxe structurale*, Paris, Klincksieck.

- VEELOCK, C. (2008a), «*Se faire + infinitif: valeurs pragmatico-énonciatives d'une construction "agentive"*», *Congrès Mondial de Linguistique Française*, Paris, France, pp. 2201-2217.
- VEELOCK, C. (2008b), «*Sur le caractère primordial de l'agentivité dans les "périphrases passives" en se faire, se laisser, se laisser, se voir + infinitif*», *Cahiers de l'Association for French Language Studies*, 14.1., pp. 5-23.
- WILLEMS, D. (2000), «*Les verbes de perception et le passif*», in L. Schøsler (éd): *Le passif*, Actes du colloque international, Institut d'études Romanes, Université de Copenhague du 5 au 7 mars 1998, *Études Romanes de l'Université de Copenhague* n° 45. Copenhague, Danemark, pp. 171-184.

A DESCRIPTIVE STUDY OF BRAZILIAN OFFENSIVE PHRASES

Ana C. Bastos-Gee*
abastos@ilch.uminho.pt

In this paper I describe aspects of the grammar of DP-of-DP offensive phrases in Brazilian Portuguese, classifying the offensive words in three morphosemantic classes. The first part of the paper focuses on characteristics that unify the offensive words into one single class of expressive content constructions. I discuss properties such as: speaker orientation, impossibility of modification by a degree adverb, headedness of the phrases, reversibility, and strength of the determiners. The second part of the paper focuses on characteristics that support a classification of these offensive words in three classes: expressive abstract nouns, epithets, and swear words. I discuss properties such as: morphosemantic characteristics, gender agreement, and number agreement.

Key-words: Brazilian Portuguese, descriptive grammar, offensive phrases, expressive abstract nouns, epithets, swearwords.

Neste artigo descrevo aspectos da gramática de sintagmas ofensivos do tipo DP-de-DP em Português Brasileiro, classificando as palavras ofensivas em três classes morfossemânticas. A primeira parte deste artigo enfoca características que unificam as palavras ofensivas em uma classe única de construções com conteúdo expressivo. Discuto propriedades como: a orientação para o falante, a impossibilidade de modificação por um advérbio de grau, o núcleo dos sintagmas, a reversibilidade e a “força” dos determinantes. A segunda parte deste artigo enfoca características que apoiam uma classificação dessas palavras ofensivas em três classes: substantivos abstratos expressivos, epítetos e palavrões. Discuto propriedades como: características morfossemânticas, a concordância de gênero e a concordância de número.

Palavras-chave: Português Brasileiro, gramática descritiva, sintagmas ofensivos, substantivos abstratos expressivos, epítetos e palavrões.

* Universidade do Minho, Centro de Estudos Humanísticos, Braga, Portugal

1. Introduction

This paper describes constructions with offensive nouns and nominalized adjectives, such as certain expressive abstract words with ironic interpretation, epithets and swearwords in Brazilian Portuguese.

- (1) [A beleza/ galinha/ droga da Maria] fugiu.
 The(FEM) beauty/ hen/ drugof- the(FEM) Mary vanished.
 ‘That lazy/ promiscuous/ piece of shit of Mary vanished’

Each one of the offensive nouns in (1) above is representative of a different class of offensive words in Brazilian Portuguese. The expressive abstract noun *beleza* ‘beauty’ is used ironically to describe *Maria* ‘Mary’ as ‘vain, snobbish or lazy’; the epithet *galinha* ‘hen’ is a pejorative way of characterizing a person who has many lovers; and the swear word *droga* ‘drug’ expresses speaker’s strong disapproval in a more general fashion, similarly to the English expression ‘piece of shit’, which I use to translate all members of this class.

The term *expressive content* is used here to refer to words and phrases that carry emotional content, such as anger, surprise, affection, etc. This is the case of the offensive words studied in this paper, which are used to express speaker’s anger, bother, or irony. In addition to the cases exemplified in (1), words with expressive content are often used as interjections or vocatives. I limit the scope of this paper to the cases in which they are used within a traditional nominal phrase that conforms to the schema DP₁-of-DP₂. In (2) I show the labels I use to refer to the different items in this construction.

- (2) D₁ N₁ of-D₂ N₂
 A droga da Maria
 the drug of-the Mary
 ‘that piece of shit of Mary’

The term *traditional nominal phrase* is a neutral term used here to refer to a phrase that has a noun as its semantic head and to replace the term *determiner phrase* (DP), since recent studies of phrase structure have argued for the availability of additional phrases higher than DP in the nominal phrase (Ormazabal 1991, Ogawa 2001, Bastos-Gee 2011, among others). In (2) there is a traditional nominal phrase formed by two other traditional nominal phrases. The first noun in the linear order is the expressive word

and its label is N_1 . The determiner that immediately precedes N_1 is D_1 . The second noun in the linear order is the semantic head of traditional nominal phrase, as I show in section 2.3, and its label is N_2 . The second determiner in the structure is D_2 and it is often contracted with the preposition *de* 'of'.

This paper has two goals. The first goal is to describe properties that are common to all constructions containing offensive words in Brazilian Portuguese. The second goal is to describe the properties that justify a classification of offensive words, based on morphosemantic properties and agreement patterns. The generative grammar framework, terminology and methodology are used to guide the description of the collected data.

The main language studied in this paper is Brazilian Portuguese; although, whenever available in the literature, I refer to comparable cases in other languages, especially Spanish. Brazilian Portuguese was chosen as object of study, because to the best of my knowledge there is no detailed description and classification of offensive words available in the literature of this language.⁽¹⁾ The data was collected through grammaticality tests applied to speakers of Brazilian Portuguese, from different regions of Brazil, and also includes my own linguistic judgments.

This paper is organized in two parts. In section 2, I describe the properties that unify the offensive words into one large class. These properties are: speaker orientation, impossibility of modification by a degree adverb, headedness of the phrases, reversibility, and strength of the determiners. In section 3, I describe three different classes of constructions with offensive word, focusing on their morphosemantic characteristics, gender agreement, and number agreement.

2. Common properties of constructions with offensive words

As mentioned above, offensive words in Brazilian Portuguese can be classified in expressive abstract nouns, epithets and swear words. In this section, I discuss properties that are common to these three classes of offensive words, when they are within a DP-of-DP structure. These properties are speaker orientation, also referred to as main clause interpretation or widest scope, impossibility of degree modification, inverse headedness, interaction with the movement of other phrases out of the traditional nominal phrase, impossibility of reversibility, and the definiteness issue.

(1) This paper results from the development of many descriptive aspects originally reported in Bastos-Gee 2011, and presents new facts and new generalizations.

2.1. Offensive words are speaker oriented

Expressive abstract nouns, epithets and swear words are interpreted as a semantic contribution of the speaker of the sentence, but syntactically they are clearly within the limits of a traditional nominal phrase. This observation is not new, and it has been made for many languages. Just to cite a few cases, Aoun, Choueiri, and Hornstein (2001: 386) refer to this property of epithets in Lebanese Arabic as a “main clause” interpretation, and Aoun and Choueiri (2000:2–3) attest that some epithets in that language have an extra definite marker, a characteristic that distinguishes them from other nominal-internal items, but which reinforces the claim that they are items within a traditional nominal phrase themselves. Similarly, Potts (2005) observes for English that expressions carrying expressive content, such as expressive adjectives and epithets, have “widest scope”, and Huddleston and Pullum (2002:553) show that expressive adjectives in this language behave like strictly attributive adjectives, such as *former* and *premier*, in the sense that they must be prenominal and can co-occur with other adjectives in the traditional nominal phrase. Potts (2005) also shows that in German, expressive adjectives are case-marked just like all other adjectives.

This property of being at same time interpreted as part of the discourse-layer, on one hand, and internal to the traditional nominal phrase, on the other hand, is also present in Brazilian Portuguese.

- (3) A Maria disse que o chefe mencionou que o beleza do motorista sumiu.
 the Mary said that the boss mentioned that the beauty of-the driver vanished.
 ‘Mary said the boss mentioned that lazy driver vanished’
- (4) A Maria disse que o chefe mencionou que o filho da mãe do motorista sumiu.
 the Mary said that the boss mentioned that the son of-the mother of-the driver vanished.
 vanished.
 ‘Mary said the boss mentioned that son of a bitch of a driver vanished’
- (5) A Maria disse que o chefe mencionou que o merda do motorista sumiu.
 the Mary said that the boss mentioned that the shit of-the driver vanished.
 ‘Mary said the boss mentioned that piece of shit of a driver vanished’

On one hand, the negative opinion towards the driver in the sentences above is perceived as a contribution of the speaker, not as Mary’s or her boss’s. On the other hand, *beleza* ‘beauty’ (expressive abstract noun), *filho da mãe* ‘son of a bitch’ (epithet) and *merda* ‘shit’ (swearword) are clearly within the traditional nominal phrase. It is actually not true for Brazilian Portuguese that items with expressive content behave exactly like other modify-

ing items within the traditional nominal phrase, since such constructions have some special properties, which will be discussed next. However, the fact that a determiner precedes them strongly suggests that the expressive nouns are within the limits of the traditional nominal phrase.

2.2. Degree modification is not permitted

Di Tullio and Saab (2006)⁽²⁾ classify constructions with epithets in Spanish in two classes: attributive and referential. In a nutshell, referential N_1 in Spanish appear within a definite verb-argument containing a full DP_2 or proper name as N_2 while attributive N_1 in Spanish appear within an indefinite predicative expression containing a bare noun as N_2 . According to their description, it is possible to make a clear-cut distinction between these two types of constructions in Spanish, which means that some expressive words are intrinsically referential while others are intrinsically attributive. This is not the case for Brazilian Portuguese, as exemplified below with the swear-word *droga* ‘drug’.

(6) Within verb-argument

Eu comprei a droga da casa/ uma droga de (uma) casa.

I bought the drug of-the house/ a drug of a house

‘I bought that piece of shit of a house.’

(7) Within predicative expression

(Uma/ que) droga de casa, essa do João!

A/how drug of house this of-the John

‘What a piece of shit of a house this one of John’s is!’

The very same offensive word that appears in (6) also appears in (7). The example in (6) complies to the above conditions for referential constructions in Spanish while the example in (7) complies with the above conditions for attributive constructions. Furthermore, the verb-argument containing the offensive word can be definite or indefinite, and N_2 can even be a bare noun in (6).

One of the properties that distinguishes between referential and attributive constructions in Spanish, discussed by Di Tullio and Saab (2006) is modification by degree adverbs. N_1 in referential constructions admits modification by degree adverbs in Spanish while N_1 in attributive constructions does not allow it. Some examples are the following:

(2) I refer the reader to Saab (2006) and Di Tullio and Suñer (2004) for developments in the analysis and description of the data in Spanish.

Modification by degree adverbs is not possible in indefinite predicative expressions modifying bare nouns either.⁽³⁾

(13) Expressive abstract noun

Uma	(* mu ito/ t ão)	<u>beleza</u> de motorista,	esse mais velho.
a	very/ so	beauty of driver,	this more old.
'Such a very good driver, this old one.'			

(14) Epithet

Uma	(* mu ito/ t ão)	<u>mosca morta</u> de motorista,	esse mais velho.
a	very/ so	fly dead of driver,	this more old.
'Such a very stiff driver, this old one.'			

(15) Swearword

Uma	(* mu ito/ t ão)	<u>merda</u> de motorista,	esse mais velho.
a	very/ so	shit of driver,	this more old.
'Such a piece of shit of a driver, this old one'			

Even if the indefinite D_1 were not present, these sentences would still be unacceptable with a degree adverb.

From this point on, I limit the scope of this paper to the argumental cases, both definite and indefinite, as exemplified in (6), since most of the restrictions on offensive content within predicative expressions may be a reflex of constraints on bare nouns in Brazilian Portuguese, and therefore only indirectly related to the expressive content. I leave this suggestion open to further research.

For the sake of completeness of this description, when it comes to intensifying the degree of the expressive content in Brazilian Portuguese, the only process available is the recursion of DPs with expressive content, which I refer to as a “re-cursing” process.

(16) Swearword/ swearword/ swearword

A <u>merda</u> da <u>droga</u> da <u>bosta</u>	do motorista	sumiu de novo.
The shit of-the drug of-the crap	of-the driver	vanished of again
'The piece of shit of the driver vanished again.'		

“Re-cursing” is possible for all three classes of expressive content, and mixing members of different classes is perfectly acceptable.

(3) Abstract words lose their ironic interpretation in predicative expressions, although their content is still expressive. Also, there are very few epithets that can take a bare noun as their N_2 . The epithet *mosca morta* ‘dead fly’ is an exception.

2.3. The second noun is the head

In DP-of-DP constructions, for instance, *o filho do João* ‘the son of John’s’ or *o filho do mecânico* ‘the son of the mechanic’, the traditional head of the structure is the first noun in the linear sequence, *filho* ‘son’. This is, however, not the case for DP-of-DP constructions with offensive words, in which the N_2 is the traditional head.⁽⁴⁾ The classic tests to show the head of these constructions involve anaphor binding and subject-verb agreement. The results of these tests are presented below.

The following examples with anaphor binding show a contrast between sentences without expressive content in (17)-(18) and sentences with expressive content in (19).

- (17) O João_j /o mecânico_i se_{i/j} machucou na porta.
 The John/ the mechanic himself hurt in-the door
 ‘John/ the mechanic hurt himself in the door.’
- (18) O filho_i do João_j/ do mecânico_i se_{i/j} machucou na porta.
 The son of-the John/ of-the mechanic himself hurt in-the door
 ‘The son of John’s/ of the mechanic hurt himself in the door.’
- (19) O **filho-da-mãe** do João_j/ do mecânico_i se_{i/j} machucou na porta.
 The son-of-the-mother of-the John/ of-the mechanic himself hurt in-the door
 ‘The son of a bitch of John/ of the mechanic hurt himself in the door.’

Se ‘himself/herself’ is an anaphor. In (17), the anaphor can bind the c-commanding DP *o João* ‘John’ and *o mecânico* ‘the mechanic’. In (18), the anaphor can bind the c-commanding DP *o filho do João/ do mecânico* ‘the son of John/ the mechanic’ as a whole, but it cannot bind the second DP *o João* ‘John’/ *o mecânico* ‘the mechanic’. The second DP is more embedded within the traditional nominal domain, and it does not c-command the anaphor. The case in (18) is different from (19). In (19), the anaphor can bind the second DP *o João* ‘John’ and *o mecânico* ‘the mechanic’. One may conclude from these facts that the structure of (18) and (19) are different. In fact, these facts are compatible with a number of structural analyses proposed in the literature for other languages. Just to cite a few among them,

(4) Among the first studies on reverse headedness is Selkirk (1977), which studies a type of pseudopartitive construction in sentences such as “She broke/ drank a bottle of wine”. The verbs *to break* and *to drink* show different semantic restrictions when selecting their complements, the verb *to break* selects a solid complement and the verb *to drink* selects a liquid complement. In Selkirk’s (1977) analysis, *bottle* is the head of the complement of *to break* in the “container reading” but *wine* is the head of the complement of *to drink* in the “content reading”. To account to the difference between container and content readings, she proposes that each reading will correspond to a different structure.

there are analyses in which DP_1 and DP_2 are in a small clause configuration (Den Dikken 2006, Suñer 1990, and Español-Echevarría 1997), others in which both DP_1 and DP_2 are modifiers of an abstract head (Di Tullio and Saab 2006) and others in which DP_1 is a pre-nominal modifier of N_2 (Bastos-Gee 2011). It is beyond the scope of this paper to discuss each of them in detail.

The second test used to identify the head of DP-of-DP structures is subject-verb agreement. When the first and second DPs have different number values, the verb agrees in number with the head of the subject.

(20) Subject-verb agreement

O chefe desses taxistas estacionou/ *estacionaram na frente da
 the boss of-these taxi-drivers parked(SING)/ parked(PL) in-the front of the
 minha casa.
 my house
 ‘the boss of taxi drivers parked in front of my house.’

In (20), the verb agrees in the singular with *chefe* ‘boss’.

In the case of DP-of-DP with offensive words, there are only a few cases in which the first and second DPs can have different number values. These restrictions will be described in section 3.3. For now, the cases that do allow the first and the second DPs to have different number values provide a piece of evidence for the second noun as the traditional head of the DP-of-DP subject.

(21) Subject-verb agreement

Essa merda desses taxistas *estacionou/ estacionaram na frente da
 this shit of-these taxi-drivers parked(SING)/ parked(PL) in-the front of-the
 minha casa.
 my house
 ‘those piece of shit of taxi drivers parked in front of my house.’

In the above case, the verb must agree with N_2 *taxistas* “taxi drivers”, not with N_1 *merda* “shit”, which reinforces the conclusion that N_2 is the head of the traditional nominal phrase in (21).

2.4. The order between DP_1 and DP_2 is not reversible

Reversibility of the linear order is a property often found in DP-of-DP constructions. Moro (1997) and (2000) on his study of dynamic antisymme-

try investigates phrases such as *books of this type* and *this type of books*. In Moro's analysis, *books* and *this type* are generated in a small clause configuration, and either *book* or *this type* moves into the specifier of the preposition *of*.

- (22) a. [... of [_{sc} books this type]
 b. [books [of [_{sc} t this type]
 c. [this type [of [_{sc} books t]

Suñer (1990) and Español-Echevarría (1997) argue that the possibility of inversion in nominal domains containing epithets in Spanish can best be analyzed by Moro's style analysis. Some examples of inversion in Spanish phrases are the following:

- (23) a. Un libro de porquería/ una porquería de libro
 A book of garbage/ a garbage of book
 'that piece of shit of a book'
 b. Un libro de mierda/ una mierda de libro
 A book of shit/ a shit of book
 'that piece of shit of a book'

Di Tullio and Saab (2006) criticize the empirical coverage of their analysis. According to them, inversion is possible for attributive epithets only and there are exceptions even among members of the attributive class, as shown below.⁽⁵⁾

- (24) Una birra de cerveza/ *una cerveza de birra
 a rubbish of beer/ a beer of rubbish
 'that piece of shit of a beer'

As for Brazilian Portuguese, inversion of the type described above is nearly inexistent and I could find only one case that allows it, as shown below.

- (25) Um livro de merda/ uma merda de livro
 A book of shit/ a shit of book
 'that piece of shit of a book'

(5) All translations from Spanish are my own.

Except for this one case, the order DP_1 -of- DP_2 is not reversible at all in Brazilian Portuguese. This suggests that *de merda* ‘of shit’ may be a frozen expression in the language instead of the result of a syntactic process.

2.5. Determiners agree in “strength”

In this section, I describe a property of the determiners in DP-of-DP structure with offensive items: “strength agreement”. Before we go into this property, it is important to clarify that only high determiners can appear in constructions with expressive content. The following example shows high and low pre-nominal modifiers in BP.

- (26) High determiners low determiners
 As/ aquelas/ umas poucas/ muitas/ várias modelos viajaram.
 The/ those/ a few/ many/ several models traveled
 ‘Few/ many/ several models traveled.’

From this point on, when I mention to the property of “strength” of determiners, I refer to high pre-nominal determiners only.

The term “strength” refers to the classification of determiners into strong and weak determiners. Strong determiners, such as demonstratives, induce definiteness effects while weak determiners, such as indefinite articles, do not induce definiteness effects. The classic test to show this distinction involves movement of wh-phrases out of the traditional nominal domain.

- (27) Strong determiner: demonstratives
 *De quem você rasgou [essa/aquela foto t_i]?
 Of whom you tore [this/that picture t_i]
 ‘Who is the person such that you tore apart this picture of him?’
- (28) Weak determiners: indefinite, as well as low modifiers
 De quem você rasgou [uma/ muitas/ várias foto(s) t_i]?
 Of whom you tore [a/ many/ several picture t_i]
 ‘Who is the person such that you tore apart his picture?’

Strong determiners do not allow movement of wh-phrases out of the traditional nominal domain, as in (27) above, while weak determiners allow it, as in (28). As for the definite article, Brazilian Portuguese has two homophonous versions of it: a strong definite article and a weak definite article (see Bastos-Gee 2011 for discussion).

(29) Strong definite article

*De quem você rasgou [a foto preferida t_i]?
 Of whom you tore the picture favorite t_i?
 'Who is the person such that you tore apart his favorite picture?'

(30) Weak definite article

De quem você rasgou [a foto t_i]?
 Of whom you tore the picture t_i?
 'Who is the person such that you tore apart his picture?'

Speakers of Brazilian Portuguese see a contrast between (29) and (30). Under a semantic viewpoint, one of the differences between them is that in (29) we have a specific picture while in (30) the picture is less specific. The existence of two versions of the definite article within the same language is a common property among Romance languages (Torrego 1987, Ormazabal 1991, Vernaud and Zubizarreta 1992, Longobardi 1994, Ticio 2003, Tellier and Valois 1995, among others.) and the implications of it for the study of the definiteness effect has been extensively studied in the generative literature from both syntactic and semantic viewpoints.

The generalization for Brazilian Portuguese is that D₁ agrees with D₂ in "strength". This generalization holds for argumental cases only, since all the predicative cases have a bare noun as N₂. To the best of my knowledge, the facts described in this section are new observations. Constructions with expressive nouns within the traditional nominal phrase show a definiteness effect when D₂ is a definite article or a demonstrative. This is true for all semantic types of expressive nouns.

- (31) Eu despedi a/essa/*uma/∅ gracinha **da/ dessa** modelo
 I fired the/ this/ a/ ∅ little-grace of-the/ of-this modelo
 'I fired that snobbish model'
- (32) Eu despedi a/essa/*uma/∅ mosca-morta **da/ dessa** modelo
 I fired the /this/a/ ∅ fly-dead of-the/ of-this modelo
 'I fired that stiff model'
- (33) Eu despedi a/essa/*uma/∅ merda **da/ dessa** modelo
 I fired the /this/a/ ∅ shit of-the/ of-this modelo
 'I fired that piece of shit of a model'

As illustrated in (31)-(33), if D₂ is a definite article or demonstrative, then D₁ must be a definite article or demonstrative. In this case, D₁ cannot be an indefinite article, and it cannot be ∅.

The results for a weak D₂ are shown below.

- (34) Eu despedi a /*essa/ uma/*∅ gracinha de **uma** modelo
 I fired the /this/a/ ∅ little-grace of a modelo
 'I fired that snobbish model'
- (35) Eu despedi a /*essa/ uma/*∅ mosca-morta de **uma** modelo
 I fired the /this/a/ ∅ fly-dead of a modelo
 'I fired that snobbish model'
- (36) Eu despedi a /*essa/ uma/*∅ merda de **uma** modelo
 I fired the /this/a/ ∅ shit of a modelo
 'I fired that snobbish model'

If D_2 is an indefinite article, then D_1 can be a weak definite article or an indefinite article. In this case, D_1 cannot be a demonstrative or \emptyset .

2.6. Summary of the section

In this section I discussed a number of properties that are commonly found in the literature of DP-of-DP constructions in other languages, especially Spanish. I showed that the three classes of offensive words in Brazilian Portuguese have many properties in common.

Table 1: Common properties

	Speaker-orientation	Degree adverbs	N2 as head	Reversibility	Strength "agreement"
Abstract nouns	Yes	No	Yes	No	Yes
Epithets	Yes	No	Yes	No	Yes
Swear words	Yes	No	Yes	No, with one exception	Yes

So far, this paper described properties that all offensive words share. In the next section, I provide a detailed description of each of the morphosemantic classes, pointing out exceptions, if needed, and emphasizing generalizations. I also show patterns of gender and number agreement.

3. Differences between classes of offensive words

There are three morphosemantic classes of offensive nouns in Brazilian Portuguese. The main characteristics that distinguish them are whether they were originally abstract nouns, concrete nouns or interjections. As discussed previously, whether the expressive word modifies an individual or a group has been considered one of the most important properties to the characterization of expressive content in Spanish. Different from Spanish, the ability to combine with an individual or a group does not contribute to establish a clear-cut distinction between two classes in Brazilian Portuguese. There is rather a gradation, which I describe as main tendencies for each class in addition to pointing out exceptions, if needed.

3.1. Morphosemantic properties

Brazilian Portuguese has quite a few nouns that can carry expressive content in the construction studied here, as well as some nominalized adjectives. Before I describe the syntactic and morphosyntactic properties that justify a three-way classification, I discuss the morphosemantic basis to this classification of offensive words into expressive abstract nouns, epithets and swear words. Let us start with expressive abstract nouns.

(37) *A minha irmã/ a Maria/ a minha gata/ a minha chave é uma **beleza/ gracinha**.*

The my sister/ the Mary/ the my cat/ the my key is a beauty/ little-grace
'my sister/ Mary/ my cat/ my key is very good.'

(38) Expressive abstract nouns

A **beleza/ gracinha/ lindeza** da minha irmã/ Maria/ da minha gata/ da minha chave
the beauty/ little-grace/ beauty of-the my sister/ Mary/ of-the my cat/ of-the my key
desapareceu de novo
vanished again.

"that no-good of a sister of mine/Mary/ a cat of mine/ a key of mine vanished again"

Abstract nouns with expressive content, such as in (37) and (38), are usually combined with very specific classes of nouns. The ones presented above can characterize persons, animals and objects. In (37), they are not in a prenominal position within the traditional nominal phrase, and their meaning is positive. When in a definite traditional nominal phrase as in (38), these abstract nouns can only convey irony; they express the property of being 'no good'. When combined with human individuals, the prop-

erty of being ‘no good’ is sometimes interpreted as ‘vain, snobbish or lazy’, depending on the context.

One special note should be made with respect to the presence of diminutive or augmentative morphemes in these constructions. Diminutives and augmentative morphemes are often used in Portuguese to convey expressive content, and they are often added to abstract nouns (and epithets, as well). For instance, the abstract word *beleza* ‘beauty’ is used with feminine diminutive for women, *a belezinha* ‘the little pretty’ and with masculine augmentative for men, *o belezão* ‘the big pretty’.

Additional examples of abstract nouns that can only characterize very specific types of nouns are *asco* ‘disgust’ and *fiasco* ‘failure’, which combine with situations and events, and with food, respectively.

- (39) a. O **fiasco** da festa/ apresentação/ campanha me deu dor de cabeça.
 the failure-of-the party/ presentation/ campaign me gave acke of head.
 “the failure of a party/ presentation/ campaign gave me a headache.”
 b. O **asco** da sopa/ comida me deu dor de cabeça
 the disgust-of-the soup/ food me gave acke of head.
 “the disgust of a soup/ food gave me a headache.”

A few expressive abstract nouns, like *amorzinho* ‘little love’ and *maravilha* ‘wonder’, can only be combined with bare nouns in the position of N_2 . In this case, they have a kind of ironic interpretation when used in a definite/ demonstrative traditional nominal phrase, as shown below. This ironic interpretation is better translated by the phrase “the so called” in English.

- (40) a. O **amorzinho** de menina mentiu para mim.
 the little-love of girl lied to me.
 “that so-called love of a girl lied to me.”
 b. Essa **maravilha** de cidade tem altos índices de criminalidade
 this wonder of city has high rates of criminality
 “this so-called wonder of a city has high rates of criminality.”

As mentioned above, this paper focus on expressive nouns, but it is important to notice that a large number of attributive adjectives can be “nominalized” and occupy the position of N_1 in the structure. The following are just a few examples with attributes that are originally negative and positive.

(41) Nominalized adjectives

A **horrorosa/ antipática** da minha irmã/ Maria sumiu de novo.
 the ugly / unpleasant of-the my sister/ Mary disappeared again.
 “that lazy/ antisocial of a sister of mine/ Mary disappeared again.”

(42) Nominalized adjectives

A **bonita/ inteligente/ simpática** da minha irmã/ Maria sumiu de novo.
 the pretty/ intelligent/ pleasant of-the my sister/ Mary disappeared again.
 “that lazy/ dumb/ antisocial of a sister of mine/ Mary disappeared again.”

If the nominalized adjective denotes a negative attribute, its original meaning remains the same, but if it denotes a positive attribute, then the ironic interpretation is the only one possible.⁽⁶⁾

The most salient property of the members of this class is the ironic interpretation that is present when they are in a traditional nominal phrase that is the argument of a verb, especially definite arguments. Expressive abstract nouns lose their ironic interpretation within a predicative expression, as well as in post-nominal position, as will be shown below.

Opposing to the class of abstract nouns, the class of epithets has nouns that were originally concrete nouns, as exemplified below.

(43) Epithets

A **mosca-morta/ galinha/ banana/ laranja** da minha irmã/ Maria parou de trabalhar.
 The fly-dead/ hen/ banana/ orange of-the my sister/ Mary stopped of work
 “That stiff/ promiscuous/ wimp/ fool of a sister of mine/ Mary stopped working.”

(44) Epithets

*A **mosca-morta/ galinha/ banana/ laranja** minha impressora parou de trabalhar.
 The fly-dead/ hen/ banana/ orange my printer stopped of work
 “That stiff/ promiscuous/ wimp/ fool of a printer of mine stopped working.”

Epithets, in the strict sense I use in this paper, were originally concrete nouns used metaphorically to persons to describe an attributed quality. The contrast between (43) and (44) shows that those epithets and a large number of others collected during my research can only and exclusively be combined with human individuals, with exceptions for pets and pet-objects when they are given anthropomorphic characteristics by the speaker, and only if they can be characterized by the property described by the epithet. One curious aspect of epithets is that each of them has a very specific mean-

(6) A large number of attributive words are listed as both adjective and noun in glossaries and dictionaries. Therefore, even if I intend to have a list of nouns only, it is inevitable to have some nominalized adjectives among them; some are expressive abstract nouns and others, epithets. The most reliable criterion to separate expressive abstract nouns from epithets is whether they are originally concrete or abstract.

ing that sometimes can no longer be associated with the literal meaning of the term that they have originated from. In the examples above, for instance, *mosca-morta* 'dead fly' is someone who is stiff and not lively, *galinha* 'hen' is a promiscuous person, *banana* 'banana' is a wimp, and *laranja* 'orange' is a fool who was blamed for something he or she did not do, especially in a money laundering operation.

Additional members to the list of epithets, in strict sense, are the following ones, accompanied by literal translations only: *animal* 'animal', *anta* 'tapir', *babaca* 'idiot', *babão* 'fool', *bacana* 'rich, cool', *barata tonta* 'flustered cockroach', *bundão* 'big bottom', *burro(a)* 'donkey', *cachorro(a)* 'dog', *canalha* 'scoundrel', *chato(a)* 'pubic louse', *cornio(a)* 'cheated man', *doido(a)* 'crazy', *escroto(a)* 'scrotum', *filho(a)-da-mãe* 'son of the mother', *filho(a)-da-puta* 'son of the whore', *idiota* 'idiot', *mula* 'donkey', *pamonha* 'cornbread', *pentelho(a)* 'pubic hair', *peste* 'pest', *porre* 'drunk', and *puto(a)* 'whore'. There is also the epithet *gato(a)* 'cat', which does not convey an offensive characterization *per se*, but it is slightly vulgar, meaning 'sexually attractive'. The epithets above can be used generally for any person, but there are some epithets that are specially used to express prejudiced views against minorities based on gender, ethnicity and/or sexual orientation. All of them are epithets in the strict sense used here. The ones that are offensive to homosexuals are only combined with male persons, such as *boiolo* 'gay', *cuzão* 'big bottom', *frutinha* 'little fruit', and *veado* 'deer'; the ones that are offensive to women are only combined with female persons and usually originate from animal designations, *perua* 'female turkey', *piranha* 'piranha', *vaca* 'cow', *cadela* 'bitch', *égua* 'mare', etc.; and the ones that are offensive to people who live in the countryside are *matuto(a)* 'countryside person', *caipira* 'countryside person', *zé bedeu* 'nosy Jack', *zé ruela* 'smallroad Jack', *bicho-do-mato* 'beast of the jungle', etc.

It is not my intention to provide a complete glossary of epithets in Brazilian Portuguese. My goal is to show that this class is large and specialized. The most salient property of this class is that these offensive words were originally concrete nouns, which were used in a figurative or metaphoric sense to convey expressive content towards human individuals. The great majority of epithets cannot be combined with bare nouns, which in Brazilian Portuguese are used to refer to kinds of things. Nowadays, they have an expressive meaning, which is not always transparent by looking into their literal definition.

Let us now move on to the swearwords.

(45) Swear words

A **merda/ bosta/ porcaria/ porra/ droga** da minha irmã/ Maria parou de trabalhar.
 the shit/ crap/ garbage/ semen/drug of-the my sister/ Mary stopped of work
 All: “the piece of shit of a sister of mine/ Mary stopped working.”

(46) Swear words

A **merda/ bosta/ porcaria/ porra/ droga** da minha impressora parou de trabalhar.
 the shit/ crap/ garbage/ semen/drug of-the my printer stopped of work
 All: “the piece of shit of a printer of mine stopped working.”

Finally, swear words shown in (45) and (46) were originally interjections of anger. The ones shown above were originally concrete nouns, but being concrete or abstract is not the most important property, since they are now somewhat empty semantically. In a DP-of-DP structure, they can be freely used for any type of noun – person, animal, object, situation, etc. – and they all mean pretty much the same thing, i.e. they are all used to express speakers strong negative feelings toward something or someone, similarly to the English ‘piece of shit’ expression that I used to translate them.

The following table summarizes the semantic relation between type of offensive noun and type of N_2 .

Table 2: N1 versus N2

N1 \ N2	Human individual	Anthro-po-morphized pets/ pet-objects	animals	objects	places	concepts	Kinds (bare nouns)
Abstract nouns	Some	Some	Some	Some	Some	Some	Some
Epithets	Yes	Yes	No	No	No	No	A few
Swear words	Yes	Yes	Yes	Yes	Yes	Yes	Yes

As mentioned previously, abstract nouns can only be combined with N_2 that have specific semantic properties, for instance, either human individuals or situations. Epithets can only be combined with human individuals. Swear words can be combined with any kind of noun, including people, animals, objects, concepts, etc. There are no semantic restrictions on what swear words can combine with.

A syntactic piece of evidence for the distribution presented above can be seen below with respect to post-nominal occurrences of expressive content.⁽⁷⁾

(47) Abstract nouns and epithets

Eu tenho	um advogado	beleza/	gracinha/	mosca-morta/	galinha/
I have	a lawyer	beauty/	little-grace/	dead-fly/	hen/
		banana/	laranja		
		banana/	orange		
“I have a lawyer, who is nice/ nice/ stiff/ promiscuous/ wimp/ fool”					

(48) Swear words

*Eu tenho	um advogado	merda/	bosta/	porcaria/	porra/	droga
I have	a lawyer	shit/	crap/	garbage/	semen/	drug
All: “I have a piece of shit of a lawyer”						

Post-nominal occurrences of expressive words are different from reversibility, because there is no DP-of-DP structure in the examples above. There is only one nominal domain and the expressive nouns are in an adjective-like position. Swearwords, due to their origin as interjections, cannot occupy adjective-like positions within the traditional nominal phrase. The examples in (47)-(48) show a clear separation between expressive abstract nouns/ epithets, on one hand, and swearwords, on the other hand.

To summarize to information in this section, one can say that generally speaking, swearwords can combine with any kind of N_2 , epithets can only combine with human individuals or animals/ pet-objects with anthropomorphic properties, and expressive abstract words are a heterogeneous class, in which some expressive nouns can only combine individuals, others can only combine with objects, and so on.

3.2. Gender agreement mismatch

With respect to agreement between DP_1 and DP_2 , the paradigm depends on the morphosemantic class of the expressive word and on morphosyntactic properties of N_2 . In general lines, agreement between DP_1 and DP_2 is optional for swear words, but obligatory for expressive nouns and epithets. There are a few exceptions among epithets, which I describe case by case.

Before I discuss the actual data, I will review the basic properties of gender agreement in the Brazilian traditional nominal phrase. As in other

(7) Notice that in a post-nominal position, expressive abstract nouns lose the “bothering” interpretation and can be interpreted as a positive aspect.

Romance languages, determiners, nouns and adjectives agree in gender within the traditional nominal phrase in Brazilian Portuguese, as exemplified below.

- (49) O pat-o pret-o versus a pat-a pret-a
 The(MASC) duck-MASC black-MASC the(FEM) duck-FEM black-FEM
 ‘The black duck’

In the above example, the noun *gato(a)* ‘cat’ and the adjective *preto(a)* ‘black’ are both biform, i.e. they have two distinct inflectional forms for gender. In the above example, the distinction is between the inflectional morpheme *-o*, masculine, and the inflectional morpheme *-a*, feminine. Brazilian Portuguese, like other Romance languages, distinguishes between masculine and feminine only, and when the reference to mix-gender groups is necessary, the suffix *-o* is used.⁽⁸⁾ The examples in (50)-(54) show different classes of nouns and ways of identifying gender in Brazilian Portuguese.

- (50) heteronyms: *o homem* ‘the(MASC) man’ and *a mulher* ‘the(FEM) woman’
 (51) inflectional biforms: *o pato* ‘the(MASC) duck’ and *a pata* ‘the(FEM) duck-FEM’
 (52) uniforms: *o/a pianista* ‘the(MASC)/ (FEM) pianist’
 (53) epicene: *a zebra macho/ fêmea* ‘the(FEM) zebra male/ female’
 (54) grammatical gender only: *a mesa* ‘the(FEM) table’

One of the most common ways of distinguishing natural gender in Brazilian Portuguese is through the semantic relation between two nouns with different stems, as in (45). The cases in which natural gender is expressed by inflectional forms are exemplified in (46) and referred to as *biform* nouns. Different from those, Brazilian Portuguese has some nouns that have only one form, but natural gender can be identified by the masculine or feminine determiners that accompany them, as in (52). The traditional grammar also mentions *epicene* nouns like the one in (53), which are usually animal common nouns with only one form for masculine and feminine, but to which the terms *macho/fêmea* ‘male/female’ can be added in order to disambiguate the biological gender of the animal. The last two cases are considered contextual disambiguation of the natural gender, not morpho-

(8) This analysis is actually controversial in traditional and structuralist grammars. There are two lines of analysis: Câmara Jr. (1984), Macambira (1987, 1992), among others, argue that the morpheme *-a* is an inflectional morpheme for feminine and *-o* is a neutral form that indicates the class of the noun (thematic vowel). Kehdi (1990), among others, argues that there is a true opposition between the morpheme *-a*, feminine, and the morpheme *-o*, masculine.

logical properties. Importantly, the gender of the *epicene* noun itself does not change after the terms *macho/fêmea* ‘male/female’ are added to the phrase. Finally, Brazilian Portuguese does not have neuter gender; even nouns that do not express biological gender are classified as masculine or feminine, as in (54) above.

Also relevant to my description is that some adjectives in Brazilian Portuguese are biform for gender and some are uniform, as exemplified below.

- (55) Homem bonit-o/ útil *versus* mulher bonit-a/ útil
 man pretty-MASC/ useful woman pretty-FEM/ useful
 ‘handsome/ useful man’ ‘pretty/ useful woman’

In (55) the adjective *bonito(a)* ‘pretty’ is biform, i.e. it has two forms for masculine and feminine while the adjective *útil* ‘useful’ has only one.

The properties described above for gender in nouns and adjectives are important, because nouns with expressive content are a mixed class of lexical items. They have some properties of nouns and some properties of adjectives. I now describe the main data regarding gender agreement in constructions with expressive content.

One of the most fascinating pieces of data regarding gender agreement in constructions with expressive content is the so-called “gender mismatch”, which consists of apparent gender agreement between D_1 with N_2 , which ignores N_1 although it is syntactically closer to D_1 . This phenomenon has been observed for French by Milner (1978) and for Spanish by Di Tullio and Saab (2006), Di Tullio and Suñer (2004), among others. In Spanish, differently from Brazilian Portuguese, the gender mismatch is only possible for what they call referential epithets. This restriction is not found in Brazilian Portuguese, where the apparent gender mismatch is possible for nouns belonging to any of the three classes of expressive nouns found in this language, as shown below.⁽⁹⁾

- (56) Abstract noun/ epithet/ swear word
 O **beleza/ banana/ merda** do meu irmão fugiu.
 the(MASC) beauty(FEM)/banana(FEM)/shit(FEM) of-the(MASC) my brother vanished.
 ‘That promiscuous/ wimp/ fool of a brother of mine vanished.’

(9) I show later that whether N_2 is human plays a role in this paradigm, because [+human] individuals have natural gender.

In Brazilian Portuguese, this apparent gender mismatch targets a very specific group of nouns: they are all expressive nouns, feminine in their literal meaning, and ending with the suffix *-a* (or *-e*, such as *peste* ‘pest’ and *gilete* ‘gillette’). The morphological endings *-a*, *-e*, *-ista* and *-nte* are common endings for words that belong the group of nouns that are uniform for gender, as exemplified above in (52). These are nouns, in which the natural gender is identified by the determiner that accompanies them. It is important to notice that the masculine gender of D_1 above is still dependent of the gender of N_2 .

There are two exceptions to the morphological group described above: *anta* ‘tapir’ and *mula* ‘donkey’. These two epithets are always feminine and agree in feminine gender with their determiners.

There are no cases of gender mismatch between N_1 and D_1 with masculine N_1 as exemplified below. ⁽¹⁰⁾

(57) Swear word/ epithet/ epithet

O/	*A	caralho/ animal/ asno	d-a	Maria
the(MASC)/the(FEM)	dick(MASC)/animal(MASC)/donkey(MASC)	of-the(FEM)	Mary	sumiu.
vanished.				
‘That dick/ animal/ idiot of Mary vanished.’				

The example in (57) shows that if N_1 is masculine, a feminine D_1 cannot precede it. This means that the so-called gender mismatch only targets feminine nouns. It is, however, a fact that the great majority of expressive nouns in Brazilian Portuguese are feminine.

Among the **epithets**, there are many that are biform nouns and many that were originally biform adjectives used now expressively. The epithet *burro* ‘donkey’ is an example of these cases.

(58) Epithet

a. O	burr-o	d-o	João sumiu.
the(MASC)	donkey-MASC	of-the(MASC)	John vanished.
‘That stupid John vanished.’			
b. *a	burr-a	d-o	João sumiu.
the(FEM)	donkey-FEM	of-the(MASC)	John vanished.
‘That stupid John vanished.’			

(10) Thanks to Alberto Guerreiro (p.c.) who pointed out to me the colloquial neologism *a caralha* ‘the(FEM) dick(FEM)’, which derives a feminine form of the masculine swear word. His suggestion is that nouns that do not conform to the general rule for biform nouns might become regular in the historical evolution of the language.

(59) Epithet

a. *O	burr-o	d-a	Maria sumiu.
the(MASC)	donkey-MASC	of-the(FEM)	Mary vanished.
'That stupid Mary vanished.'			
b. a	burr-a	d-a	Maria sumiu.
the(FEM)	donkey-FEM	of-the(FEM)	Mary vanished.
'That stupid Mary vanished.'			

Biform epithets do not show gender mismatch, but the gender of N_1 , as well as D_1 , is clearly dependent on the gender of N_2 .

One exception to rule that all epithets have obligatory agreement between DP_1 and DP_2 is a small group of noun-adjective compound epithets. These epithets have optional agreement. To the best of my knowledge this is a new observation.

(60) A	mosca-morta/ barata-tonta	do João	sumiu.
the(FEM)	fly-dead/ flustered cockroach	of-the(MASC) John	vanished.
'That stiff/ flustered John vanished.'			
(61) O	mosca-morta/ barata-tonta	do João s	umiu.
the(MASC)	fly-dead/ flustered cockroach	of-the(MASC) John	vanished.
'That stiff/ flustered John vanished.'			

This small class of epithets behaves similarly to swearwords as for the gender agreement, in the sense that they show optional agreement, not mandatory.

The class of **swear words** behaves differently from epithets and abstract nouns. To the best of my knowledge, this observation is also new, recently made on earlier versions of this work (Bastos-Gee 2011). Consider (62)-(64).

(62) Swear word

O	merda	d-o	homem/ pat-o	sumiu.
the(MASC)	shit(FEM)	of-the(MASC)	man/ duck-MASC	vanished
'that piece of shit of a man/ duck vanished'				

(63) Swear word

A	merda	d-o	homem/ pat-o	sumiu.
the(FEM)	shit(FEM)	of-the(MASC)	man/ duck-MASC	vanished
'that piece of shit of a man/ duck vanished'				

(64) Swear word

*O	merda	d-o	carro	sumiu.
the(MASC)	shit(FEM)	of-the(MASC)	car(MASC)	vanished
'that piece of shit of car vanished'				

(65) Swear word

A	merda	d-o	carro	sumiu.
the(FEM)	shit(FEM)	of-the(MASC)	car(MASC)	vanished

“that piece of shit of car vanished”

In the above examples, gender agreement between D_1 and N_2 (i.e. the so-called “gender mismatch” between D_1 and N_1) is only possible if N_2 has natural gender, but not possible if N_2 has only grammatical gender. D_1 can be masculine in (62) agreeing with the nouns *homem* ‘man’ and *pato* ‘duck’, but it cannot be masculine in (64) agreeing with *carro* ‘car’. Notice that this issue does not arise with epithets, because they can only be combined with human individuals, i.e. elements with natural gender, to start with.

It is important to clarify that this is not a case of “semantic” agreement in the traditional sense of “gender sylepsis” between D_1 and N_2 ; this is a real syntactic issue. One way of testing for the issue is to use an epicene, i.e. a noun that is either masculine or feminine, but allows natural gender disambiguation through the adjectives *macho* ‘male’ and *fêmea* ‘female’. One example is provided below.

(66) A cobra macho sumiu.

the(FEM) snake male vanished
 “the male snake vanished.”

The word *cobra* ‘snake’ is grammatically feminine, and the adjective *macho* ‘male’ is used to contextually identify natural gender when the distinction is relevant, for instance, to contrast it with a female snake in the same context. The test has the following format. In traditional nominal phrases with expressive content, if D_1 agrees with the grammatical gender of N_2 , it should be feminine; if D_1 agrees with the contextually salient masculine gender, it should be masculine.

(67) Swear word

A/	*o	merda	d-a	cobra macho	sumiu.
the(FEM)/	the(MASC)	shit	the(FEM)	snake male	vanished

“that piece of shit of a male snake vanished.”

As shown above, D_1 cannot agree with the semantic idea of masculine; the only agreement possible is the syntactic agreement. This shows that the gender mismatch is a syntactic phenomenon, and not a semantic, contextual phenomenon.

Another way in which **swearwords** are different from epithets and abstract nouns is that the so-called “gender mismatch” is optional for swear words, even when N_2 has natural gender, as shown in (62)-(65) above. As can be seen in the examples with swear words, there is always the possibility of D_1 and N_1 agreeing internally within their own traditional nominal phrase, and having a different gender from N_2 , even when N_2 is human. There is no such optionality with epithets and abstract nouns, where either D_1 agrees with N_2 in gender or both D_1 and N_1 agree with N_2 in gender when N_2 has natural gender, exception made to noun-adjective compounds, as mentioned above.

The following table summarizes my proposal for the gender feature of nouns with expressive content in Brazilian Portuguese.

Table 3: gender agreement

Type of expressive noun	D1-N1 gender ‘mismatch’	DP1-N2 gender agreement
Feminine abstract nouns, ending -a and -e beleza ‘beauty’	Yes	Mandatory, if N2 has natural gender
Feminine epithets, ending -a and -e banana ‘banana’	Yes	Mandatory
Exceptions: Anta ‘tapir’, mula ‘donkey’	No	No
Type of expressive noun	D1-N1 gender ‘mismatch’	DP1-N2 gender agreement
Masculine abstract nouns, epithets, and swear words Fiasco ‘failure’, asno ‘donkey’, and caralho ‘dick’	No	No
Biform epithets burro(a) ‘donkey/ idiot’	No	Mandatory
Noun-Adjective compound epithets mosca-morta ‘dead fly’	Optional	Optional
Swear words merda ‘shit’	Optional, if N2 has natural gender No, if N2 has grammatical gender	Optional, if N2 has natural gender No, if N2 has grammatical gender

In the next section, I describe the properties of gender agreement.

3.3. Number agreement mismatch

Before I discuss the actual data, I will review the basic properties of number agreement in the Brazilian traditional nominal phrase. In Brazilian Portuguese, like other Romance languages, determiners, adjectives and nouns agree in number within the traditional nominal phrase.

- (68) O pato preto *versus* o-s pato-s preto-s
 The(SING) duck(SING) black(SING) the-PL duck-PL black-PL
 ‘The black duck(s)’

In Brazilian Portuguese, plural nouns, adjectives and determiners are marked with the inflectional suffix *-s* (and its allomorphs) that is added to the stem of the word, after the feminine suffix if there is one. Singular forms are indicated by the absence of the morpheme *-s*.⁽¹¹⁾

When it comes to number agreement in traditional nominal phrases with expressive content, the paradigm is considerably simpler than the one presented for gender agreement: Number agreement is obligatory when N_1 is an abstract word or an epithet, and optional when N_1 is a swear word. The following examples illustrate definite traditional nominal phrases but the generalization is also true for indefinite traditional nominal phrases in argumental positions. Also, the distinction between uniform/ biform epithets does not play a role here.

- (69) Abstract noun
 A-s beleza-s/ *a beleza d-a-s modelo-s
 The(FEM)-PL beauty-PL/the(FEM) beauty of-the(FEM)- PL model-s
 viajaram.
 traveled.
 ‘those snobbish models traveled.’
- (70) Epithet
 A-s banana-s/ *a banana d-a-s modelo-s
 The(FEM)-PL banana-PL/ the(FEM) banana of-the(FEM)- PL model-s
 viajaram.
 traveled
 ‘those wimpy models traveled.’

(11) In non-standard Brazilian Portuguese, the plural marker is usually retained in the first word in the traditional nominal phrase in the linear order, which is usually the determiner, but not exclusively.

(71) Swear word

A-s merda-s/ a merda d-a-s modelo-s
 The(FEM)-PL shit-PL/ the(FEM) shit of-the(FEM)- PL model-s
 viajaram.
 traveled
 ‘those shitty models traveled.’

In (69)-(70) above, abstract nouns, epithets and their respective determiners must agree with N_2 in number. In (71), the swear words display optional number agreement in the same way they display optional gender agreement, as discussed in the previous section. One empirical difference between the gender and the number of constructions with swear words is that there is no “number mismatch”, in the sense I used the term “mismatch” for gender.

A piece of evidence for it is that there are subtle differences in the interpretation of the traditional nominal phrases containing singular swear words and plural swear words in (71) above. In *a merda das modelos* ‘the shit of the models’, the speaker is upset with the models as a group. In *as merdas das modelos* ‘the shits of the models’, the speaker attributes the negative property to each of the members of the group. This leads me to the conclusion that in both cases, D_1 agrees internally with N_1 and this agreement is semantically relevant for interpretation.

The following table summarizes my proposal for the number feature of nouns with expressive content in Brazilian Portuguese.

Table 4: number agreement

Type of expressive noun	D_1 - N_1 number ‘mismatch’	DP_1 - N_2 number agreement
Expressive abstract nouns <i>beleza</i> ‘beauty’	No	Yes, mandatory
Epithets, including Noun-Adjective compounds <i>banana</i> ‘banana’, <i>mosca-morta</i> ‘dead-fly’	No	Yes, mandatory
Swear words <i>merda</i> ‘shit’	No	Optional

The partial conclusion for this section is that number agreement in traditional nominal phrases with expressive content depends on the seman-

tic type of the expressive N_1 . Expressive abstract nouns and epithets show mandatory agreement with N_2 , while swear words show optional agreement with N_2 , since their number value affects interpretation.

4. Final remarks

In this paper, I described a number of the properties of DP-of-DP constructions containing nouns that carry expressive content, in order to establish differences between this type of DP-of-DP construction and other kinds of DP-of-DP constructions. Additionally, I briefly compared the properties found in Brazilian Portuguese to the properties found in the literature about epithets and other offensive words in other languages, especially Spanish. This paper also includes a detailed description of three classes of offensive words in Brazilian Portuguese, as for their morphosemantic properties and the properties of the noun they modify. This classification of offensive words into three classes was crucial to describe complex patterns of gender and number agreement within the traditional nominal domain.

5. References

- AOUN, J. and L. CHOEIRI. (2000). Epithets. *Natural Language and Linguistic Theory* 18, 1-39.
- AOUN, J., L. CHOEIRI and N. HORNSTEIN. (2001). Resumption, Movement and Derivational Economy. *Linguistic Inquiry* 32, 371-403.
- BASTOS-GEE, Ana C. (2011). *Information structure within the traditional nominal phrase: the case of Brazilian Portuguese*. University of Connecticut: Doctoral Dissertation.
- CÂMARA JR., Joaquim Mattoso. (1984). *Estrutura da língua portuguesa*. 14. ed. Petrópolis, Vozes.
- DEN DIKKEN, Marcel. (2006). *Relators and linkers: The syntax of predication*, Predicate Inversion, and copulas. Cambridge, MA: MIT Press.
- DI TULLIO, A. and SUÑER, A. (2004). *Los 'nombres de cualidad' en la estructura del SD*. Presented at the XXIV Congrès International de Linguistique et de Philologie Romanes (CILPR). Aberystwyph, August 1-6.
- DI TULLIO, Ángela and Andrés SAAB. (2006). *Dos clases de epítetos en el español: sus propiedades referenciales y distribución sintáctica*. Proceedings of the XIV Congreso Internacional de la Asociación de Lingüística y Filología de América Latina (ALFAL), Universidad Autónoma de Nueva León.

- ESPAÑOL-ECHEVARRÍA, M. (1997). Two aspects of the sentential syntax of N/A of a N DP's: predicate raising and subject licensing. In: A. Schwegler, B. Tranel and M Uribe-Exteberria (eds.) *Romance Linguistics: Theoretical perspectives*. Amsterdam: John Benjamins.
- HUDDLESTON, Rodney and Geoffrey PULLUM. (2002). *The Cambridge Grammar of the English Language*. Cambridge University Press.
- KEHDI, Valter. (1990). *Morfemas do português*. São Paulo: Ática, Série Princípios.
- MACAMBIRA, José Rebolças. (1992). *Português Estrutural*. São Paulo: Pioneira.
- MILNER, Jean-Claude. (1978). *De la syntaxe à l'interprétation: Quantités, insultes, exclamations*. Paris: Éditions du Seuil.
- MORO, Andrea. (1997). *The Raising of Predicates: Predicative Noun Phrases and the Theory of Clause Structure*. Cambridge: Cambridge University Press.
- MORO, Andrea. (2000). *Dynamic Antisymmetry*, MIT Press, Linguistic Inquiry Monograph 38.
- OGAWA, Yoshiki. (2001). *A Unified Theory of Verbal and Nominal Projections*. Oxford: Oxford University Press.
- ORMAZABAL, Javier. (1991). *Asymmetries on wh-movement and some theoretical consequences*. Ms., Storrs: University of Connecticut.
- POTTS, Chris. (2005). *Expressive content as conventional implicature*. Proceedings of the Thirty-Third Annual Meeting of the North East Linguistic Society. BookSurge Publishing, December 7.
- SAAB, Andrés. (2006). Concordancia ad sensum y elipsis nominal en español: Un análisis morfosintáctico. *Revista de Lingüística y Literatura*, Universidad Nacional del Comahue, n. 34, p. 45-63, 2006.
- SELKIRK, E. (1977). Some remarks on noun phrase structure. In: P. Culiver, T. Wasow and Akmajian (eds.), *Formal Syntax*. New York: Academic Press, 285-316.
- SUÑER, A. (1990). *La predicación secundaria en español*. Doctoral dissertation. Universidad Autónoma de Barcelona.

MENÇO OU MINTO? REGULARIZAÇÃO DE PARADIGMAS VERBAIS

MENÇO OU MINTO? REGULARIZATION OF VERBAL PARADIGMS

Maria João Colaço*

m.j.colaco@hotmail.com

Esperança Cardeira**

ecardeira@hotmail.com

As mudanças sofridas pelos verbos em português ocorreram em diversas fases da história da língua, dando muitas vezes origem a irregularidades no paradigma verbal.

O objectivo deste trabalho é estudar um grupo de verbos de padrão especial que apresentam um lexema para a primeira pessoa do singular do presente do indicativo e para todas as formas do presente do conjuntivo oposto ao lexema dos outros tempos: SENTIO > *senço*, ARDEO > *arço*, AUDIO > *ouço*, MENTIO > *menço*, PETIO > *peço*, PERDO (*PERDEO) > *perço*. No entanto, cada um destes verbos evoluiu de modo autónomo, não apresentando as mesmas características.

Procura-se determinar os fenómenos responsáveis pela evolução destas formas verbais, a fim de averiguar a época em que houve variação, analisando cada uma das formas pertinentes. Para isso, recorreu-se ao uso de *corpora* constituídos por documentos escritos do português arcaico.

Palavras-chave: variação, mudança, analogia, morfologia verbal.

The changes suffered by the verbs in Portuguese occurred in different periods of the history originating often irregularities in the verbal paradigm.

The main objective of this article is to study a group of special pattern verbs which present a particular lexeme for the first person singular present indicative and all

* Faculté des Langues et des Cultures Etrangères - Universidade de Estrasburgo, França.

** Faculdade de Letras da Universidade de Lisboa – Departamento de Linguística Geral e Românica, Portugal.

the present subjunctive forms that distinguish themselves from the other tenses: SENTIO > *senço*, ARDEO > *arço*, AUDIO > *ouço*, MENTIO > *menço*, PETIO > *peço*, PERDO (*PERDEO) > *perço*. However, each one of these verbs evolved independently, not showing the same characteristics.

This study aims to determine the phenomena involved in the evolution of these verbal forms and to investigate within which epoch was variation registred, analysing each one of the pertinent forms. In order to ascertain these questions, *corpora* constituted by archaic Portuguese written documents were used.

Keywords: variation, change, analogy, verbal morphology.

*

1-Introdução

A variação e a mudança são duas das características que definem a natureza das línguas naturais. Ao longo dos séculos o português sofreu mudanças e conheceu períodos de instabilidade em que diferentes formas estiveram em concorrência. Assim, um estudo variacionista será, seguramente, o melhor caminho a seguir na análise diacrónica de uma mudança linguística.

Este trabalho tem como objetivo a observação de um grupo de verbos que, pela sua evolução fonética, constituiu um tipo de padrão especial e que, influenciados por diversos fatores, nomeadamente fatores de ordem psicológica e cultural, passaram por um período de variação que resultou na regularização de uns, mas não de outros. Este período abrangeu cerca de três séculos (séculos XIII, XIV e XV).

A fim de compreender o modo como se realizou a alteração na morfologia verbal deste grupo de palavras, é necessário estudar o período em que o fenómeno se deu e os fatores que deram origem a mudanças não fonéticas, como, por exemplo, a analogia.

Com base em documentos literários e notariais galegos e portugueses, far-se-á a sistematização e a análise da variação que se regista para este grupo de verbos que apresentaram, em resultado da ação das leis fonéticas, o radical de primeira pessoa do singular do presente do indicativo e de todas as pessoas do presente do conjuntivo fechado por sibilante, diferente do radical das restantes formas verbais.

2-Os Verbos de Padrão Especial em Português

A evolução fonética regular está na base de grande parte das irregularidades no paradigma verbal, que ainda hoje se verificam na língua. Estes verbos irregulares a que Mattos e Silva (2008: 412-425) dá o nome de ‘verbos de padrão especial’ são por esta autora agrupados por tipo de irregularidade em quatro subgrupos diferentes:

Subgrupo 1: verbos que apresentam variação no lexema das formas do não-perfeito e têm lexema específico para as formas do perfeito, com ou sem variante (p. ex. *dizer*);

Subgrupo 2: verbos que apresentam lexema invariável para as formas do não-perfeito e têm lexema específico para as formas do perfeito (p. ex. *saber*);

Subgrupo 3: verbos que apresentam variação nos lexemas do não perfeito, sendo o lexema das formas do perfeito a variante mais generalizada do lexema do não-perfeito (p. ex. *arder, crescer*);

Subgrupo 4: verbos de PP especial, tradicionalmente chamado participípio forte (p. ex. *abrir*).

O primeiro subgrupo é constituído por verbos em que a consoante final do radical se modifica foneticamente em contacto com a semivogal [j], embora não na primeira pessoa do singular do presente do indicativo nem nas formas do presente do conjuntivo. A consoante oclusiva latina [k] sonorizada pelo contexto intervocálico mantém-se, enquanto nas restantes formas fricativiza para [dz] > [z] (cf. DICO > *digo* vs. DICIS > *dizes*). A oclusiva [g], em verbos como *cingir*, também se manteve no conjuntivo e na primeira pessoa do indicativo, mas alterou-se para a africada [dʒ] e mais tarde para a fricativa [ʒ] nas outras formas do presente do indicativo. Esta distinção rapidamente desapareceu na língua, tendo a consoante [g] sido substituída por [ʒ] por analogia com as outras formas. Contudo, em *erguer* a regularização foi diferente da dos restantes verbos, uma vez que a consoante [g] do radical da primeira pessoa predominou sobre as outras (cf. *erges* > *ergues*) (Piel 1989: 24-25).

No subgrupo 2 constam verbos como *saber, prazer, caber* e *dar*, que são por sua vez subcategorizados em dois tipos. No tipo a) agrupam-se os verbos em que se deu uma metátese da semivogal [j] para a sílaba anterior nas formas do presente do conjuntivo. É o caso do verbo *saber*: *sab[j]a* (lat. SAPEAT) > *saiba*. Nas formas dos tempos do perfeito deu-se também uma metátese (SAPUI- > *soub-*), que resultou do deslocamento do *u* para a sílaba anterior. No tipo b) encontramos o verbo *dar*, que apresenta duas vogais

temáticas diferentes: VTa para os tempos do não-perfeito e VTe para os tempos do perfeito (cf. *dás* vs. *deste*) (Mattos e Silva 2008: 412-425).

Ao terceiro subgrupo, constituído pelos verbos que têm um lexema para a primeira pessoa do singular do presente do indicativo e para todas as formas do presente do conjuntivo oposto ao lexema dos outros tempos, pertencem:

(i) verbos como ARDEO > *arço*, AUDIO > *ouço*, MENTIO > *menço*, PETIO > *peço*, em que a sibilante é proveniente da palatalização dos conjuntos latinos [tj] e [dj]. Esta mudança fonológica deu-se pela presença da semivogal [j], que afeta as consoantes correspondentes a <c>, <t> e <d>, transformando-as em <ç> = [ts] > [s], fenómeno que não só se manifestou em formas verbais, como também em palavras de outras classes (Piel, 1989: 25). Contudo, observando o elenco de verbos em que esta palatalização se deu no português arcaico, constatamos que, por analogia com as outras pessoas do presente do indicativo, alguns destes verbos readotaram as consoantes oclusivas [d] e [t] (*ardo* e *mintto*), ao passo que outros mantiveram a forma correspondente à evolução fonética da palavra (*ouvir*, *pedir* e *medir*) (Maia, 1995: 24-25);(1)

(ii) verbos incoativos como *acaecer*, *conhocer*, *nacer*, *crecer* (Mattos e Silva, 2008: 412-425), que apresentaram evoluções diferentes entre si até aos dias de hoje. Este tipo de verbos provém de uma construção latina que consistia na inserção de um infixos <sc> que se associava à vogal temática dos verbos, conferindo-lhes, assim, um valor de início de ação, pelo que a este tipo de verbos se dá o nome de verbos incoativos (do lat. *inchoare*). Esta partícula incoativa, que se podia aplicar a verbos de qualquer uma das quatro conjugações latinas, dava sempre origem a verbos de terceira conjugação (tema em *ĕ*, -SCERE), que se fixaram em português na segunda conjugação, como seria de esperar (Costa, 2007: 13-27). Deste modo, coexistiam dois verbos com a mesma base, mas com um valor aspetual diferente cf. DORMIRE (*dormir*) vs. ADDORMISCERE (*adormecer*, ‘começar a dormir’). Hoje em dia, alguns destes verbos nem sempre contêm este valor incoativo (MERESCERE > *merecer*), embora esta dicotomia se mantenha para muitos dos verbos desta classe.(2)

(1) Curiosamente, esta palatalização da consoante final dos verbos, que se deu em galego-português, mas não em castelhano (*peço* vs. *pidto*), ocorreu nesta língua em palavras de outras classes: PLATEA > *plaza*, CAPITIA > *cabeza* e PUTEU > *pozo* (Piel, 1989).

(2) Na verdade, já no latim vulgar se verificava esta perda semântica do infixos incoativo em certos verbos (Väänänen, 1967: 146)

A evolução fonética que a sequência <sc> sofreu deu origem a diferentes irregularidades. Este morfema, constituído por duas consoantes, era produzido em latim como uma fricativa línguodental surda [s], na posição de coda da primeira sílaba, e como uma oclusiva velar surda [k], na posição de ataque da segunda sílaba. Esta última consoante já terá chegado ao português como uma consoante africada predorsodental [ts], tendo resultado mais tarde na fricativa [s] que hoje existe na língua. Visto que nos primórdios do português a consoante correspondente a <s> ainda não havia palatalizado em [ʃ], sendo provavelmente produzida como [ʂ], pelo menos em grande parte das regiões de Portugal (tal como se verifica hoje em certas regiões setentrionais, onde ainda se conservam as sibilantes apicais), esta terá sido assimilada pela consoante [s], correspondente a <c> /<ç>, simplificando assim num só fone, devido ao elevadíssimo número de traços distintivos que passaram a partilhar (Costa, 2007: 13-27). Assim, esta sequência passou a representar-se maioritariamente como <c> /<ç>, representação substituída, em muitos verbos, pela grafia etimológica, o que influenciou inclusivamente a maneira como estas palavras são pronunciadas. Portanto, esta mudança fonética foi provocada por fatores externos à língua, nomeadamente histórico-culturais, uma vez que a sequência <sc> foi recuperada por volta dos séculos XVI / XVII para várias palavras de diferentes classes, numa tentativa de reaproximação do português às suas raízes latinas. Esta tentativa não abrangeu todos os verbos incoativos existentes no português,⁽³⁾ o que resultou nas diferentes formas que hoje existem na língua (*nasco* > *nasço* vs. *pareasco* > *pareço*).

Relativamente ao subgrupo 4, este é constituído pelos verbos de participio passado irregular, tradicionalmente chamados participios fortes, já que são acentuados no seu radical e não na vogal temática, tal como acontece com os participios regulares. Também estes se subdividem.

(i) No primeiro grupo, encontramos verbos como *abrir*, *escrever* e *cobrir*, cuja construção de participio passado provém do étimo latino ao qual acrescentamos os morfemas nominais de género e de número (cf. *abr- / abert-*).

(ii) No segundo grupo, o participio passado constrói-se a partir do radical único do verbo, acrescentando-lhe os morfemas de género e número. A este subgrupo pertencem verbos como *aceitar*, *juntar* e *salvar* - cf. *aceitar / aceite* (Mattos e Silva, 2008: 412-425). Durante o período arcaico estas formas coexistiram com outras (*acesas / acendudas*). O mesmo se verifica

(3) Aliás, dialectalmente, a <nascer> corresponde ainda hoje *na[s]er*, com fricativa predorsodental ou apicalveolar (não palatal).

nos dias de hoje, em que muitos participios passados irregulares são substituídos pela sua vertente regular, dependendo do contexto em que surjam (*morto / matado*). Outros, ainda, desapareceram em favor de formas regulares que eliminaram a forma etimológica do participio passado (*coit- / cosid-*) (Mattos e Silva, 2008: 412-425).

Como vimos, as irregularidades inerentes aos verbos de padrão especial do português são variadíssimas. Por essa razão, pretende-se estudar e analisar mais detalhadamente os verbos pertencentes ao terceiro subgrupo, que apresentam o lexema da 1ª pessoa do presente do indicativo e de todas as pessoas do presente do conjuntivo terminado em sibilante. Pretende-se, igualmente, estudar os fenómenos que levaram à regularização de muitas destas formas. Entre eles, encontra-se a analogia, que explica algumas mudanças morfológicas destes verbos.

3-A analogia

A questão das mudanças analógicas é muito importante para este estudo, uma vez que justifica a maioria das regularidades em paradigmas onde, devido a mudanças resultantes de leis fonéticas, encontraríamos irregularidades.

O que são, então, as leis fonéticas e o que é a analogia? As mudanças fonéticas são regulares e afectam qualquer signo dentro do mesmo contexto fonológico, independentemente da categoria da palavra ou da função morfológica que o constituinte afetado desempenha na palavra em que se assinala este tipo de mudança. Este ‘carácter cego da evolução’ dos sons justifica-se pelo facto de que as mudanças fonéticas seguem uma lógica diferente da lógica do sistema linguístico, isto é, operam meramente a nível da mudança sonora (Lucchesi, 1998: 77-78). As mudanças analógicas, contrariamente às leis fonéticas, têm uma motivação mais pragmática e são reflexo da representação mental que os falantes têm das palavras que conhecem, estabelecendo relações associativas entre elas. Vejamos o modo como estas mudanças ocorrem e o tipo de relações subjacentes a este fenómeno.

3.1-A ação niveladora da analogia nos paradigmas verbais

Silva (1998: 23) exemplifica os diferentes tipos de regularização analógica dentro do quadro da flexão verbal do português:

- 1- por modelos próprios do verbo – geralmente quando ocorre a regularização dos radicais que, resultantes de leis fonéticas, se tinham tornado estranhos à maioria da conjugação (*arço, ardes* > *ardo, ardes*);
- 2- por esquemas alheios, de outros verbos, mas ainda dentro da mesma classe. Está neste caso a regularização dos morfemas sufixais que exercem uma força coercitiva de conjunto (*dei, deste, dei* (arc.) > *dei, deste, deu*); mas também o podemos ver no aparecimento de algumas formas fracas de perfeito de verbos que primitivamente tinham perfeitos fortes, exercendo-se a força reguladora simultaneamente no radical e nas desinências (*jouve, crive* > *jazi, cri*).

As regularizações que se operam no grupo de verbos que forma o nosso objeto de estudo, isto é, os verbos pertencentes ao subgrupo 3, segundo a classificação de Mattos e Silva (2008), constituem o primeiro tipo de processo de analogia, de acordo com o quadro acima apresentado. Para além de apresentar a classificação dos diferentes tipos de analogia, dentro do quadro da flexão verbal do português, Silva (1998: 21) refere, ainda, que a mudança analógica nos paradigmas verbais (e não só) deriva de associações que se estabelecem na mente do falante e que podem ser de três tipos:

- (i) **Funcionais** – é o caso da desinência *-eu*, que substitui *-ei* na terceira pessoa do singular dos pretéritos perfeitos fracos da segunda conjugação. O falante associa este morfema a uma função específica, destruindo, deste modo, a ambiguidade anteriormente existente.
- (ii) **Semânticas** – aplica-se a verbos com significados semelhantes ou opostos, que adotam as características um do outro. Exemplo deste tipo de fenómeno é o participio passado do verbo *ouvir*, muitas vezes produzido como *ouvisto*, à imagem de *visto*.
- (iii) **Morfológicas** – um verbo afectado por este tipo de associação é *jazer*, já que, durante o período arcaico, apresentava o radical JASC- na primeira pessoa do singular do presente do indicativo e todas as pessoas do presente do conjuntivo à semelhança dos verbos incoativos, que também apresentavam dois radicais diferentes nas formas do presente.

A analogia, que atua no sentido de ‘nivelar’ a língua, corrigindo as irregularidades produzidas pelas leis fonéticas, dá muitas vezes origem a novas irregularidades, criando formas que concorrem com a forma original resultante da sua evolução fonética. Esta situação provoca um desequilíbrio no sistema, uma vez que as novas formas ficam em variação com as antigas, o que pode resultar no desaparecimento de uma das formas ou dar origem a mudança, numa tentativa de voltar a equilibrar o sistema, através de um emprego restrito e particular para cada uma das formas resultantes, na alte-

ração semântica de alguma ou no desaparecimento quer da forma analógica quer da forma etimológica (Alonso, 1989: 75-78).

A analogia é tradicionalmente classificada em três diferentes tipos:

- (i) **nivelação analógica** (que resulta na diminuição da alternância nos paradigmas) – um caso ilustrativo é precisamente o verbo *arder*, que constitui um dos verbos em estudo neste trabalho. Visto que a evolução fonética das suas formas deu origem a uma irregularidade - neste caso, a existência de dois lexemas para as formas do presente - um dos lexemas foi eliminado, passando a existir uma só forma para o radical do presente (*arço*, *ardes* > *ardo*, *ardes*).
- (ii) **extensão analógica** (que resulta na criação de alternância em paradigmas em que ela não existia) – este tipo de analogia revela-se em palavras como *acordos*, cuja vogal do radical se pronuncia frequentemente como [ɔ], à imagem do que acontece com palavras como *porcos* (*p[ɔ]rcos*). A alternância vocálica que existe entre as palavras *p[o]rco*, *p[ɔ]rca* e *p[ɔ]rcos* deve-se a uma harmonização vocálica que se deu na palavra *porcu*, devido à presença da vogal final [u], o que não ocorre com as duas outras palavras, já que as suas vogais finais ([a] e [o]) não são altas. Porém, a palavra *acordo* não sofreu nenhuma harmonização, sendo que a vogal [o] do radical provém da evolução fonética da palavra latina, tanto no singular quanto no plural.
- (iii) **criação analógica** (que consiste na adição ao léxico de uma palavra que se relaciona com outra já existente) – um exemplo deste tipo de criação é a palavra *cheeseburger*, que surgiu pela substituição de um segmento da palavra *hamburger* – *ham* (fiambre) – que, apesar de não ser um constituinte morfológico da palavra, foi interpretado como tal (Marquilhas, 1996: 578-579)

Para os estruturalistas, a analogia constitui um tipo de mudança que não pressupõe uma alteração no sistema (Alonso, 1989: 76). De facto, consiste na imitação regular por parte dos falantes, criando determinadas construções que se aproximam de outras já presentes na língua. Deste modo, dá-se uma substituição de formas anteriormente existentes no sistema linguístico. Como este tipo de regularização não é sistemático, afeta apenas algumas formas e, visto que a analogia é um fenómeno de natureza ‘psicológica’ e não ‘mecânica’ (como são as leis fonéticas), a sua ocorrência é imprevisível. Assim, a sua abrangência não é total, o que se traduz numa ‘atuação esporádica’ (Marquilhas, 1996: 578-579). Trata-se de um fenómeno que não depende simplesmente de uma tentativa de equilibrar formas onde se encontra ‘desproporção’, mas é reflexo de um conjunto de regras já estabelecidas no sistema gramati-

cal dos falantes, que relacionam forma e função. Estas regras acabam por se estender a novos casos, aos quais não eram anteriormente aplicadas.

O tipo de mudança analógica que diz respeito aos verbos aqui estudados é a nivelção analógica, pois estas regularizações (*arço*, *ardes* > *ardo*, *ardes*) acabam com a existência de dois radicais verbais para o presente do indicativo e do conjuntivo em algumas formas. Pretende-se avaliar cada uma das formas verbais deste tipo e analisar a sua evolução, tentando perceber que fenómenos ocorreram na regularização de algumas delas, de que maneira atuou a analogia nestas mudanças e em que casos não é possível explicar linguisticamente as alterações que se deram no paradigma verbal.

4-Dados linguísticos

Neste ponto pretende-se fazer uma análise sistemática da evolução dos verbos que apresentam dois radicais para o presente – um para a 1ª pessoa do singular do presente do indicativo e todas as formas do conjuntivo e um para as restantes formas do presente do indicativo.

Antes de passar ao estudo dos dados recolhidos é importante fazer uma breve descrição dos documentos que constituem o *corpus* deste trabalho. Estes foram retirados de *corpora* em formato digital, que estão disponíveis *online*: *Corpus informatizado do Português Medieval*, *Tesouro Medieval Galego-português* e *Cantigas Medievais Galego-Portuguesas*, sendo que os textos selecionados abrangem o período entre os séculos XIII e XVI, e são provenientes de diversas regiões de Portugal e da Galiza. Estamos conscientes de que a distribuição geográfica destes documentos peca por haver mais registos de textos galegos⁽⁴⁾ e do norte de Portugal do que do sul. Em contrapartida, procurámos utilizar documentos de diversos tipos, de modo a melhor avaliar a evolução das formas verbais em estudo em diversos registos textuais.

Importa acentuar que o estudo que se realizará será mais qualitativo do que quantitativo, devido à referida discrepância na proveniência dos textos a estudar ou a falta de informação referente à sua origem, já que há muitos textos cuja procedência não é certa. Assim, a análise dos dados a nível geográfico contemplará apenas as ocorrências de textos cuja proveniência se conhece. A análise diacrónica destes verbos será, por isso, o principal alvo deste trabalho.

(4) Em alguns documentos galegos ocorrem formas em castelhano; estas formas serão, naturalmente, excluídas da análise.

4.1-Verbo *sentir*

Antes de observar o comportamento deste verbo no período arcaico, apresenta-se abaixo um quadro com todas as ocorrências⁽⁵⁾ pertinentes para o presente estudo, ou seja, todas as formas do presente do conjuntivo e 1ª pessoa do singular do presente do indicativo.

Tabela 1- *Sentir*

	Presente do indicativo		Presente do conjuntivo	
	Forma etimológica	Forma variante	Forma etimológica	Forma variante
P1	<i>senço</i> (2)	<i>sinto</i> (39), <i>seno</i> (5)		
P2				<i>sintas</i> (1)
P3			<i>sença</i> (3)	<i>sinta</i> (8), <i>seno</i> (2)
P4				<i>sintamos</i> (7)
P5			<i>sençades</i> (3)	
P6				<i>sintam</i> (2), <i>sentam</i> (1)

Há, para este verbo, dois radicais que concorrem com o radical etimológico, cuja diferença reside precisamente na vogal do radical. Note-se que a forma do radical que se perdeu, SENT-, é a mesma das restantes formas do presente (cf. *sent-es*) e, no entanto, não foi a que vingou, mas sim SINT-. Isto mostra-nos que a analogia actua de dois modos diferentes. Além da nivelção analógica na regularização deste grupo de verbos, com a generalização de uma só consoante final do radical de todas as formas do presente, deu-se também uma extensão analógica, que criou uma alternância vocálica, anteriormente inexistente. Poderá pensar-se que estes verbos, cuja vogal do radical se alterou, tenham seguido o modelo de verbos como *ferir*, em que encontramos dois radicais diferentes, com a mesma distribuição (*firo* vs. *feres*). Contudo, não há nenhum indício suficientemente forte que nos permita afirmar que esta terá sido a associação que levou à alternância vocálica. Outros fatores poderão ter influenciado a evolução destas formas.

(5) Por não as considerarmos pertinentes para o presente estudo não registamos, para este e para os outros verbos, as alografias do tipo *sinto* ~ *synto*.

Registam-se poucas formas do presente para o verbo *sentir*. Porém, podemos ver que há mais ocorrências de formas analógicas com o radical SINT- (58) do que SENT- (8). Note-se, também, que a presença do radical etimológico SENÇ- é muito pequena no *corpus* observado (8).

Vejam os dados a distribuição destes dados ao longo dos séculos em que houve variação.

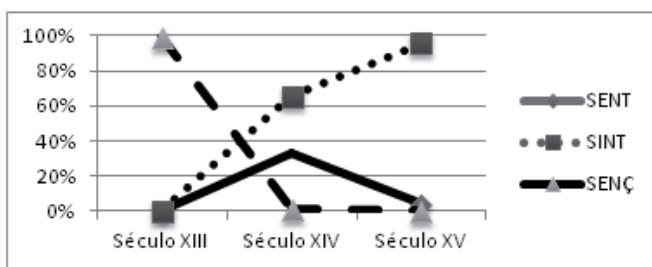


Figura 1 – *Sentir*

No século XIII ainda não há registos de variação, encontrando-se apenas formas verbais terminadas em sibilante. Destas formas etimológicas há ainda uma ocorrência no século XIV, a par das duas formas analógicas acima referidas (SINT- e SENT-), que predominam neste período. De acordo com estes dados, o século XIV parece ser o período em que se regista maior variação para este verbo, havendo um número bastante significativo de formas com o radical SENT- (33,3%), o que nos leva a crer que a extensão analógica que ocorreu nestas formas do verbo *sentir* possa ter sido posterior à nivelção analógica que se manifestou nos restantes verbos deste grupo, eliminando qualquer tipo de variação paradigmática para o presente. Contudo, não se poderá afirmar categoricamente que tenha sido essa a ordem da evolução destas formas verbais. Os dois processos analógicos sofridos por este verbo podem ter aparecido na língua mais ou menos no mesmo momento, de um modo independente. Devido à escassez de dados referentes a este período, não nos é possível tirar conclusões claras neste sentido. No entanto, o facto de o radical SENT- se registar predominantemente no século XIV leva-nos a colocar esta hipótese, especialmente porque os restantes verbos deste grupo, à exceção do verbo *mentir*, apenas sofreram um tipo de analogia – o nivelamento analógico.

No século XV só se registam duas ocorrências do radical SENT-. Como seria de esperar, as formas etimológicas deste verbo já se encontram extin-

tas, o que significa que durante este período a pouca variação existente aponta, já, para a estabilização da mudança.

4.2-Verbo *perder*

Tal como acima vimos para o verbo *sentir*, também *perder* apresentou, durante o período arcaico, três radicais diferentes para as formas do presente do conjuntivo e primeira pessoa do singular do presente do indicativo. Ao contrário de todos os outros verbos aqui em estudo, *perder* foi o único em que se deu uma mudança não fonética no lexema que não resultou na eliminação da alternância da consoante final do radical, antes numa alteração dessa mesma consoante.

Tabela 2-*Perder*

	Presente do indicativo		Presente do conjuntivo	
	Forma etimológica	Forma variante	Forma etimológica	Forma variante
P1	<i>perço</i> (18)	<i>perco</i> (12), <i>perdo</i> (2)		<i>perca</i> (3)
P2			<i>perças</i> (1)	<i>percas</i> (1)
P3			<i>perça</i> (49)	<i>perca</i> (297) <i>perda</i> (21)
P4			<i>perçamos</i> (3)	<i>percamos</i> (9) <i>perdamos</i> (2)
P5			<i>perçades</i> (6)	<i>percades</i> (207), <i>perdades</i> (2) <i>perdaes</i> (1)
P6			<i>perçam</i> (5)	<i>percam</i> (73)

Para este verbo, registaram-se bastantes ocorrências, como se pode ver, o que já nos permitirá fazer uma análise diacrónica e geográfica destas formas verbais.

Olhando para o conjunto de todas as ocorrências vemos que, à semelhança do que se observou para *sentir*, as formas resultantes da ação da extensão analógica sobre este paradigma não se fixaram na língua, tendo o radical PERC- substituído as outras duas formas concorrentes (PERÇ-, PERD-). De facto, esta é a forma que mais vezes se registou, seguida da diretamente resultante da forma etimológica. As formas em PERD- são as menos frequentes na nossa amostra (4%).

Analisemos, então, como se distribuem estas formas diacronicamente:

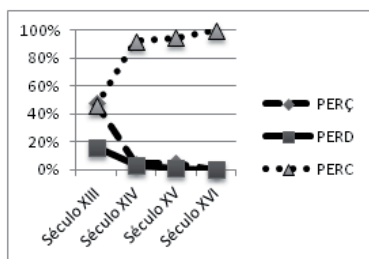


Figura 2-Perder: dados portugueses

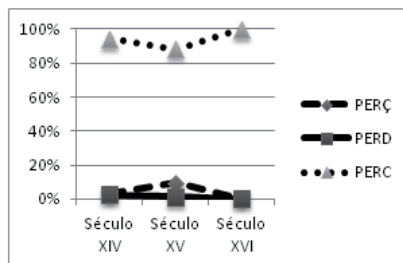


Figura 3-Perder: dados galegos

Relativamente aos documentos portugueses do século XIII, 48% das ocorrências que se registam mantêm a sua forma etimológica, o que mostra que este foi o século que testemunhou o maior período de variação, sendo que o surgimento das formas variantes tem que ter sido anterior a esta época, embora no nosso *corpus* não se registem dados anteriores referentes a este verbo.

Também se observa a existência de dois radicais não etimológicos, embora a forma PERD-, igual ao radical das restantes formas do presente, ocorra com menor frequência (18% de todas as ocorrências) do que a forma PERC-, que prevaleceu.

As ocorrências referentes ao século XIV demonstram que o lexema etimológico corresponde a 4,2% das ocorrências de *perder* neste século, o que já aponta para uma preferência dos falantes pelo uso do radical PERC- neste grupo de formas do presente. O uso da forma PERD- diminuiu bastante relativamente ao século XIII, constituindo agora apenas 3,5% das ocorrências.

No século XV, o radical etimológico PERÇ- representa apenas 4,5% destas formas verbais, não havendo uma grande diferença quanto ao uso desta forma relativamente ao século anterior. Quanto ao radical proveniente da ação da extensão analógica neste paradigma (PERD-), é nítido que o seu uso sofreu um decréscimo acentuado, registando-se apenas duas ocorrências no nosso *corpus* (0,7%), pelo que se pressupõe que nesta época provavelmente já não pertencesse ao léxico ativo dos falantes. Mais uma vez, observa-se que o radical PERD-, igual à forma usada para as restantes pessoas do presente do indicativo, é eliminado na língua, em favor de uma outra forma em que se regista alternância no paradigma do presente.

Relativamente ao século XVI, podemos afirmar que os poucos dados recolhidos (18 ocorrências) corroboram as tendências que os demais dados faziam prever, sendo que todos eles correspondem à forma do radical que hoje se regista. Contudo, não é certo que durante este período não houvesse alguma variação. O reduzido número de ocorrências não nos permite presumir tal com absoluta certeza, mas tudo indica que no século XVI o radical PERC- já se tivesse estabelecido na língua como a única forma gramaticalmente aceite para a primeira pessoa do presente do singular do indicativo e todas as pessoas do presente do conjuntivo.

Em relação aos textos galegos, observa-se que o radical analógico PERD- surge com baixíssima frequência no século XIV, acabando por não ocorrer nos séculos seguintes nos documentos observados, o que é curioso, já que essa foi a forma que se fixou na língua galega. A forma etimológica também regista um decréscimo em favor da outra forma variante, PERC-.

4.3-Verbo *arder*

Arder é um dos verbos que pertencem ao elenco de verbos de padrão especial pertinentes para este estudo. No entanto, a sua ocorrência na língua corrente, contrariamente a verbos como *perder*, *pedir* ou *ouvir*, que pertencem ao vocabulário quotidiano dos falantes, é bastante reduzida, pelo que as formas de *arder* constituintes do nosso *corpus* se registaram também em número muito reduzido. Assim, far-se-á a descrição e a distribuição diacrónica e regional das poucas ocorrências deste verbo, tendo em conta que dificilmente se poderá fazer alguma generalização a respeito das diferentes fases por que passou durante o período de variação.

Tabela 3-*Arder*

	Presente do indicativo		Presente do conjuntivo	
	Forma etimológica	Forma variante	Forma etimológica	Forma variante
P1				
P2				
P3			<i>arça</i> (3)	<i>arda</i> (1), <i>arca</i> (1)
P4				<i>ardamos</i> (1)
P5				<i>ardades</i> (1)
P6				

Vemos nesta tabela que apenas se obtiveram sete formas deste verbo; o radical etimológico e o analógico surgem praticamente com a mesma frequência. Observa-se também uma forma com o radical ARC-, que, por ser uma ocorrência isolada, talvez se possa considerar como uma questão meramente gráfica, uma vez que a alternância <c> ~<ç> era bastante frequente nesta época. Porém, vale sublinhar que a insuficiência de dados não permite averiguar qual terá sido a história desta e das restantes formas de *arder*.

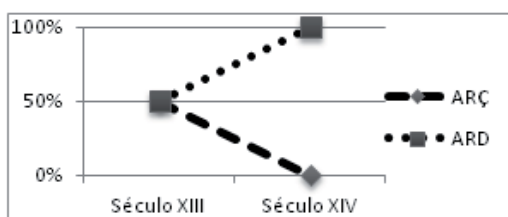


Figura 4-Arder

Todas as ocorrências se distribuem pelos séculos XIII e XIV; apenas uma provém de um documento do século XIV. No século XIII encontramos o mesmo número de formas etimológicas e variantes (3 ocorrências), embora, devido aos poucos dados encontrados, não possamos ver qual seria a verdadeira tendência deste verbo durante este período. A única forma detetada para o século XIV é formada com o radical ARD-, tal como nos dias de hoje. Contudo, somos levados a supor que, à semelhança do que se verifica com os outros verbos, as duas formas tenham estado em concorrência até mais tarde.

Relativamente à distribuição geográfica destas formas verbais, não será possível fazer a sua sistematização, já que se tem conhecimento da origem de apenas três das ocorrências.

4.4-Verbo *ouvir*

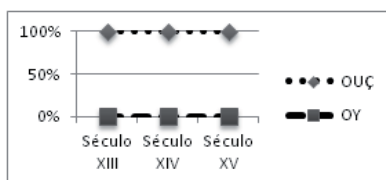
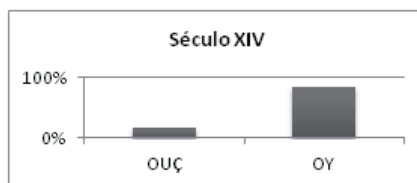
Neste estudo, apenas se assinalam dois radicais alternativos das formas relevantes do presente (OUÇ- e OY-). Note-se que o lexema OIÇ-, que hoje se encontra em variação com a forma OUÇ-, ainda não se regista durante o período arcaico, pelo menos nos documentos por nós analisados. Abaixo apresenta-se o quadro com todos os dados encontrados referentes a *ouvir*.

Tabela 4-*Ouvir*

	Presente do indicativo		Presente do conjuntivo	
	Forma etimológica	Forma variante	Forma etimológica	Forma variante
P1	<i>ouço</i> (59)	<i>oyo</i> (103)	<i>ouça</i> (2)	
P2			<i>ouças</i> (3)	
P3			<i>ouça</i> (14)	<i>oya</i> (6)
P4			<i>ouçamos</i> (3)	
P5			<i>ouçais</i> (3), <i>ouçades</i> (10)	
P6			<i>ouçam</i> (18)	<i>oyam</i> (8)

Olhando para a distribuição destas formas, concluímos que os radicais variantes apenas se registam na primeira pessoa do singular do presente do indicativo e na terceira do singular e do plural do presente do conjuntivo. Contudo, estas formas constituem 53% de todas as ocorrências obtidas.

Vejamos agora como se distribuem estas formas diacronicamente:

Figura 5-*Ouvir*: dados portuguesesFigura 6-*Ouvir*: dados galegos

Para este verbo, encontram-se formas referentes aos séculos XIII, XIV e XV nos documentos portugueses. Quanto aos galegos, os dados obtidos são todos provenientes de textos que datam do século XIV. A distribuição destas formas é nítida – a forma etimológica permaneceu na língua portuguesa, aparentemente sem ter estado em variação com a forma OY-, que encontramos em abundância nos textos galegos e que constitui o radical utilizado hoje nesta língua.

4.5-Verbo *pedir*

A seguir a *perder*, *pedir* é o verbo em relação ao qual se encontram mais ocorrências, o que nos permite uma análise mais detalhada dos fenómenos envolvidos nesta mudança.

Tabela 5-*Pedir*

	Presente do indicativo		Presente do conjuntivo	
	Forma etimológica	Forma variante	Forma etimológica	Forma variante
P1	<i>peço</i> (118)	<i>pido</i> (84)	<i>peça</i>	
P2				
P3			<i>peça</i> (14)	<i>pida</i> (3)
P4			<i>peçamos</i> (3)	<i>pidamos</i> (2)
P5			<i>peçais</i> (1)	<i>pidades</i> (1)
P6			<i>peçam</i> (7)	

Os dois radicais que se apresentam encontram-se em número bastante equilibrado, sendo que a forma analógica corresponde a 41% de todas as formas. Importa referir que, à semelhança do que se apontou em relação ao verbo *ouvir*, em *pedir* o radical PID- foi o que prevaleceu na língua galega, pelo que a proveniência dos documentos em que se encontra cada uma das formas estudadas é de grande relevância na análise deste verbo. É provável que o uso de cada um dos radicais seja característico de determinadas regiões, esperando-se que haja uma preferência pelo radical variante pelo menos na Galiza e possivelmente no norte de Portugal.

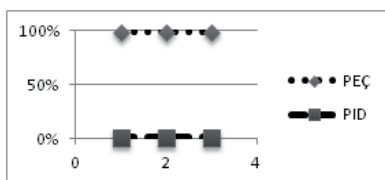


Figura 7- *Pedir*: dados portugueses

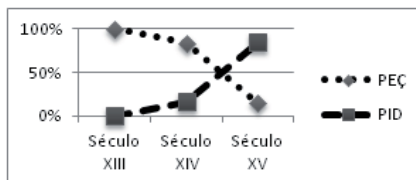


Figura 8-*Pedir*: dados galegos

De acordo com as nossas expectativas, as formas analógicas de *pedir* predominam nos textos galegos, correspondendo à totalidade das formas

encontradas, o que nos leva a concluir que este verbo poderá nunca ter tido um radical concorrente dentro do que hoje constitui o território português. Não será, no entanto, de excluir a hipótese de que possa ter havido alguma variação nas regiões portuguesas circundantes da Galiza, apesar da inexistência de documentos comprovativos.

Relativamente ao percurso que este verbo teve na Galiza, não se regista variação no século XIII: as três formas encontradas são construídas com o radical etimológico, não se distinguindo, portanto, das formas portuguesas. No século XIV surgem as primeiras formas construídas com o radical PID-, embora em número reduzido (16%), o que demonstra que a forma PEÇ- seria ainda a menos marcada, durante este período. É no século XV que o lexema que hoje constitui o radical do presente do verbo galego *pedir* se generaliza. Apesar de ainda se assinalarem algumas formas verbais etimológicas (15%), já se adivinha o domínio da nova forma, que se manterá na língua galega até hoje. Este verbo parece seguir o padrão evolutivo do verbo *ouvir*, que já analisámos acima. A distinção que passa a existir entre os radicais portugueses e os galegos poderá ter origem na influência do castelhano sobre a língua galega, uma vez que as formas inovadoras que surgem no galego são semelhantes às castelhanas já existentes.

4.6-Verbo *mentir*

Apesar de pertencer ao vocabulário corrente dos falantes, este é um verbo para o qual não se registaram muitas ocorrências. Contudo, tentaremos examinar a sua evolução ao longo dos séculos, em Portugal e na Galiza.

Tabela 6-*Mentir*

	Presente do indicativo		Presente do conjuntivo	
	Forma etimológica	Forma variante	Forma etimológica	Forma variante
P1	<i>menço</i> (1)	<i>minto</i> (2)		<i>mintta</i> (1)
P2				
P3			<i>mença</i> (2)	<i>mintta</i> (2)
P4				
P5			<i>mençades</i> (4)	
P6				

A partir da distribuição que acima se apresenta, podemos perceber que a forma etimológica e a variante tiveram uma presença relativamente equilibrada, ao longo deste período, não havendo registo de um uso preferencial notório de uma forma em detrimento da outra. No entanto, sabemos que cada século conta a sua história, bem como cada lugar. Assim, observemos o modo como estas formas se distribuem:

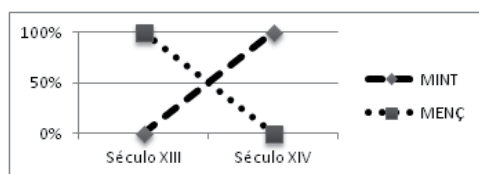


Figura 9-Mentir

Curiosamente, apenas se obtiveram dados relativos ao século XIII e ao século XV, pelo que o período de variação e transição não se regista. Porém, somos levados a calcular que, durante o século XIV, as formas etimológicas tenham coexistido com as formas analógicas.

5-Conclusões

A partir da sistematização e análise que foi feita relativamente a este grupo de verbos que, em resultado da sua evolução fonética, constituem um tipo de verbos de padrão especial, cujo radical das formas do presente do conjuntivo e primeira pessoa do singular do presente do indicativo difere do radical das demais formas verbais, chegámos à conclusão de que não é possível estabelecer um padrão evolutivo comum a todos os verbos.

Ainda assim, podemos subdividir este grupo em quatro padrões distintos:

1º padrão – constituído pelos verbos *ouvir* e *pedir*, para os quais se regista variação a partir do século XIII, sendo que, até esse momento, os lexemas OUÇ- e PEÇ- ainda não haviam sofrido a ação da analogia que, mais tarde, se manifestou. Observamos também que as formas analógicas que se assinalam apenas se encontram em textos galegos. Assim, com margem para dúvidas, podemos supor que as formas etimológicas chegam ao português sem sofrer alterações, já que também não passaram por um perí-

odo de variação no território onde hoje se fala português. Vale referir que as formas alternativas que surgiram nos textos galegos para estes verbos são as mesmas que se registam na língua atual. Em galego, o verbo *oír* não regularizou e o radical do presente do conjuntivo e primeira pessoa do singular do presente do indicativo terminado em semivogal, OI-, difere do radical das restantes formas, O- (*oio, oes, oia*), ao passo que no caso do verbo *pedir* existe um radical verbal específico para todas as pessoas do presente, que difere do radical dos restantes tempos (*pido, pides, pedi*).

2º padrão – constituído pelo verbo *arder*, que durante o seu período de variação apenas apresentou dois radicais variantes para as formas do conjuntivo e primeira pessoa do singular do presente do indicativo – a forma etimológica e a forma resultante da ação de nivelção analógica, que acabou por se generalizar, tanto em português, como em galego.

Poderemos imaginar que a nivelção analógica tenha sido mais eficaz para este verbo do que para os restantes, devido à sua reduzida frequência de uso, comparado com os outros verbos (*perder, pedir, ouvir, mentir e sentir*), que constituem vocábulos do quotidiano dos falantes. Embora o verbo *arder* não seja dos mais raros, emprega-se em situações peculiares. Assim, observamos que o significado, a pragmática, a iteração ou qualquer outra relação associada às palavras interferem em questões morfológicas e fonológicas, o que justifica a existência de quatro padrões distintos neste grupo de verbos.

3º padrão – constituído pelos verbos *sentir* e *mentir*, para os quais se registam três radicais alternativos durante o período de variação no português arcaico (SENÇ-, SENT-, SINT- e MENÇ-, MENT-, MINT-) – um correspondente à forma etimológica, um outro proveniente de uma nivelção analógica (SENT-, MENT-) e um resultante, provavelmente, de uma extensão analógica derivada da influência dos verbos que apresentam alternância vocálica nos radicais do presente, como *ferir, dormir e correr* (SINT-, MINT-). Esta última forma, que eliminou a alternância da consoante final do radical mas introduziu uma alternância vocálica, foi a que se fixou em ambas as línguas.

4º padrão – constituído pelo verbo *perder*. Relativamente a este verbo, como já foi referido, não se encontra uma explicação óbvia para a adoção do radical PERC-, que concorreu com a forma etimológica PERÇ- e com PERD-, resultante da nivelção analógica, e que, apesar de se distanciar da morfologia do radical das restantes formas, foi adotada pelos falantes, mantendo assim a alternância paradigmática já antes existente. Parece-nos, contudo, plausível assumir que a variação gráfica <c> vs. <ç> existente na

época possa estar na origem da criação deste radical. Em galego, foi o radical PERD- que, pela ação da nivelação analógica, prevaleceu, eliminado, assim, qualquer alternância em todas as formas verbais.

Curiosamente, observamos uma ocorrência do uso de um radical terminado em <c> relativa ao verbo *arder* (*arca*). Este caso isolado pode levar-nos a pensar que a consoante alternativa seja apenas resultado da dita variação gráfica, representando, portanto, a consoante sibilante do radical etimológico. O facto de não haver mais registos desta forma corrobora a hipótese de que não existiu realmente um radical terminado em consoante oclusiva para este verbo, o que significa que a variação gráfica não terá afetado o verbo *arder*. Será então legítimo perguntarmo-nos se este tipo de variação gráfica pode constituir a verdadeira causa da variação e posterior mudança das formas do verbo *perder*? De facto, não seria um caso único na história da língua portuguesa. Recordemos como a inovação gráfica dos verbos incoativos influenciou a nossa língua, alterando a pronúncia da consoante final do radical (*nacer* > *nascer* [naser] > [naʃser] / [nɐʃer]). Além disso, a frequência de uso do verbo *perder* é nitidamente superior à do verbo *arder*, como acima já indicámos, o que levou à existência de dois padrões evolutivos distintos para estes dois verbos.

A nível geográfico, observa-se que os documentos galegos apontam para tendências diferentes, que adivinham desde cedo a separação linguística que veio a dar-se posteriormente. Esta distinção começa a partir do século XIII, acentuando-se nitidamente no século XIV, sendo que é no século XV que termina o período de variação, apesar da existência de raras ocorrências de formas etimológicas nos documentos galegos desta época, que seriam, com certeza, consideradas formas marcadas e conservadoras para a maioria dos falantes, visto que a grafia está sempre ‘um passo atrás’ da realidade linguística oral. Vale referir que esta separação entre galego e português se deveu, em grande parte, a fatores externos. No caso do galego, a influência do castelhano não pode ser menosprezada. Os verbos *oír* e *pedir* são exemplos muito concretos desta influência, dado que estas formas variantes, idênticas às castelhanas, e que encontramos em documentos galegos desde o século XIII, não surgiram em documentos portugueses em nenhum momento do período em que os verbos pertencentes a este grupo sofreram variação. Isto significa que enquanto a variação e mudança que os verbos portugueses sofreram se pode justificar pela ação da analogia (nivelação analógica ou extensão analógica), as mudanças verbais do galego mostram como outros fatores, como o contexto histórico, podem ter um papel determinante na evolução de uma língua. Não poderemos, porém,

afirmar que todas as mudanças que ocorreram nestas formas verbais galegas tenham tido a mesma causa. A regularização do verbo *perder* (*perdo*, *perdes*, etc.), que não se verificou na língua portuguesa, parece ser um caso típico de regularização por nivelção analógica. Além do mais, as formas verbais castelhanas do presente não são iguais às formas galegas (*pierdo* vs. *perdemo*), havendo alternância paradigmática relativa ao acento.

A partir da análise que fizemos de cada verbo concluímos que, embora cada um tenha a sua história própria, todos eles passaram por um período de grande variação e instabilidade no século XIV, em que vários radicais verbais coexistiam no mesmo sistema linguístico. Observámos, também, que foi no século XV que se estabeleceram as mudanças que permaneceram, tanto em português como em galego, até aos dias de hoje. Estas observações confirmam, assim, as propostas de periodização que apontam para uma fase de instabilidade que, nos finais do século XIV, separa o português antigo do português médio. Ao longo do período médio (ou seja: ao longo do século XV, quando as ocorrências das formas arcaicas se tornam raras) são selecionadas as variantes que se imporão no português clássico do século XVI. Historicamente, a este período de seleção de variantes corresponde a definitiva separação política e linguística entre o português de raiz setentrional, ligado à Galiza, e o português centro-meridional. Enquanto a Galiza sofria a ocupação castelhana, Portugal definia-se como reino independente. A localização da corte na Estremadura, elegendo as variedades centro-meridionais como modelo linguístico, será determinante para o processo de standardização do português.

Na história de uma língua interferem, indubitavelmente, fatores de natureza linguística e extra-linguística. É, por isso, necessário termos em conta fatores diversificados como a analogia, principal motor de regularização dos paradigmas dos verbos portugueses deste grupo; a influência de outras línguas, nomeadamente a influência do castelhano em determinados verbos galegos; a grafia, que terá sido a causa da criação do atual radical do verbo *perder*, bem como outros fatores de ordem psicológica e social, que levaram à regularização de alguns verbos, mas não de outros do mesmo tipo.

Este grupo de verbos é, assim, uma prova de que a língua é dependente da ação do Homem, das suas necessidades, do seu conhecimento e do contexto cultural em que existe. É impossível dissociar uma língua dos seus falantes, da sua história e da sua cultura: a língua é reflexo da toda a experiência humana.

Referências:

- ALONSO, Emílio R. (1989), *Las estructuras gramaticales desde el punto de vista histórico*, Madrid, Editorial Sintesis.
- COSTA, J. Vasco Oliveira (2007), *A evolução gráfica e fonológica da sequência <-s.c(e/i)-> dos verbos latinos*. Dissertação de Mestrado em Linguística Portuguesa apresentada à Faculdade de Letras da Universidade de Lisboa, inédita.
- LUCCHESI, Dante (1998), *Sistema Mudança e Linguagem*, Lisboa, Edições Colibri.
- MAIA, Clarinda de Azevedo (1995), “Sociolinguística histórica e periodização linguística. Algumas reflexões sobre a distinção entre ‘português arcaico’ e ‘português moderno’”, *Diacrítica* 10, pp. 3-30.
- MARQUILHAS, Rita (1996), “Mudança linguística”, in I. H. Faria, E. R. Pedro, I. Duarte e C. Gouveia (orgs.), *Introdução à Linguística Geral e Portuguesa*, Lisboa, Editorial Caminho, pp. 563-588.
- MATTOS E SILVA, Rosa Virgínia (2008), *O Português Arcaico: Uma aproximação*, Lisboa, Imprensa Nacional-Casa da Moeda.
- PIEL, Joseph-M. (1989), *Estudos de Linguística Histórica Galego-Portuguesa*, Lisboa, Imprensa Nacional-Casa da Moeda [1944: A Flexão Verbal do Português (Estudo de Morfologia Histórica)].
- SILVA, Antonino da Conceição de Almeida (1998), *O sistema verbal do português arcaico. O papel da analogia na regularização dos paradigmas verbais do português arcaico da segunda fase*. Dissertação de Mestrado em Linguística Portuguesa apresentada à Faculdade de Letras da Universidade de Coimbra, inédita.
- VÄÄNÄNEN, Veikko (1967), *Introduction au latin vulgaire*, Paris, Klincksieck.

Referências das fontes documentais:

- BROCARDO, Maria Teresa (ed.) (1994), *Crónica do Conde D. Pedro de Meneses*. Dissertação de Doutoramento. Lisboa: F.C.S.H., pp. 333-693.
- CABANA OUTEIRO, Alexandra (ed.) (2003), *O Tombo H da catedral de Santiago. Documentos anteriores a 1397*, Valga, Concello de Valga.
- CAL PARDO, Enrique (ed.) (1984), *Monasterio de San Salvador de Pedroso en tierras de Trasancos. Colección documental*, A Coruña, Deputación Provincial.
- CAL PARDO, Enrique (ed.) (1999), *Colección diplomática medieval do Arquivo da Catedral de Mondoñedo. Transcripción íntegra dos documentos*, Santiago de Compostela, Consello da Cultura Galega.
- CASTRO, Ivo et alii (eds.) (1985), *Vidas de Santos de um Manuscrito Alcobacense* (Cod. Alc. CCLXVI / ANTT 2274). Lisboa, I.N.I.C., pp. 16-52; 59-83.

- CINTRA, Luís Filipe Lindley (ed.) (1951), *Crónica Geral de Espanha de 1344*, Lisboa, I.N.C.M.
- COMESAÑA MARTÍNEZ, María Ángela (ed.) (1995), *O tombo do Hospital e Ermida de Santa María do Camiño de Pontevedra*, Pontevedra, Museo de Pontevedra.
- DURO PEÑA, Emilio (ed.) (1967), “El monasterio de San Salvador de Sobrado de Trives”, *Archivos Leoneses* 21, 49, pp. 7-86.
- DURO PEÑA, Emilio (ed.) (1972), *El Monasterio de S. Pedro de Rocas y su colección documental*, Ourense, Instituto de Estudios Orensanos “Padre Feijóo”.
- DURO PEÑA, Emilio (ed.) (1977), *El Monasterio de San Esteban de Ribas de Sil*, Ourense, Instituto de Estudios Orensanos “Padre Feijóo”.
- FERNÁNDEZ DE VIANA Y VIEITES, José Ignacio (ed.) (1995), *Colección diplomática del monasterio de Santa María de Ferreira de Pantón*, Lugo, Servicio de Publicaciones de la Diputación Provincial de Lugo.
- FERNÁNDEZ SALGADO, Antonio (ed.) (1999), *A documentación medieval de San Bieito do Campo*. Tese de licenciatura apresentada na Universidade de Santiago de Compostela, inédita.
- FERRO COUSELO, Xesús (ed.) (1967), *A vida e a fala dos devanceiros. Escolma de documentos en galego dos séculos XIII ao XVI*, 2 vols, Vigo, Galaxia.
- GARCÍA ORO, J. (1987), “Viveiro en los siglos XIV y XV. La Colección Diplomática de Santo Domingo de Viveiro”, *Estudios Mindonienses* 3, pp. 11-132. [Publicado como livro pelo Concello de Viveiro, 1988].
- GARVÃO, Maria Helena (1992), *Foros de Garvão. Edição e Estudo Linguístico*. Dissertação de Mestrado apresentada à Faculdade de Letras da Universidade de Lisboa, pp. 65-99.
- GONZÁLEZ GARCÉS, Miguel (ed.) (1988), *Historia de La Coruña. Edad Media*, A Coruña, Caixa Galicia.
- GRAÑA CID, M^a del Mar (ed.) (1990), *Las órdenes mendicantes en el obispado de Mondoñedo. El convento de San Martín de Villaoriente (1374-1500)*, Separata de Estudios Mindonienses.
- LAGARES, Xoán Carlos (ed.) (2000), *E por esto fez este cantar. Sobre as rúbricas explicativas dos cancioneiros profanos galego-portugueses*, Santiago de Compostela, Laivento.
- LOPES, Graça Videira (2002), *Cantigas de Escárnio e Maldizer dos Trovadores e Jograís Galego-Portugueses* (Edição digitalizada).
- LÓPEZ FERREIRO, Antonio (ed.) (1975 [1895]), *Fueros municipales de Santiago y de su tierra*, Madrid, Ediciones Castilla, pp. 695-710.
- LORENZO, Ramón (ed.) (1975, 1977), *La traducción gallega de la Crónica General y de la Crónica de Castilla*, Ourense, Instituto de Estudios Orensanos “Padre Feijóo”.
- LORENZO, Ramón (ed.) (1985), *Crónica troiana. Introducción e texto*, A Coruña, Fundación “Pedro Barrié de la Maza, Conde de Fenosa”.

- LUCAS ÁLVAREZ, Manuel e LUCAS DOMÍNGUEZ, Pedro (eds.) (1996), *El priorato beneditino de San Vicenzo de Pombeiro y su colección diplomática en la Edad Media*, Sada / A Coruña, Ediciós do Castro (Publicações do Seminário de Estudos Galegos).
- LUCAS ÁLVAREZ, Manuel e LUCAS DOMÍNGUEZ, Pedro (eds.) (1988), *San Pedro de Ramirás. Un monasterio femenino en la Edad Media*, Colección diplomática, Santiago, Publicacións de Caixa Galicia.
- LUCAS ÁLVAREZ, Manuel e LUCAS DOMÍNGUEZ, Pedro (eds.) (1996), *El monasterio de San Clodio do Ribeiro en la Edad Media: estudio y documentos*. Sada / A Coruña, Ediciós do Castro (Publicacións do Seminario de Estudos Galegos).
- MAIA, Clarinda de Azevedo (1986), *História do Galego-Português. Estado Lingüístico da Galiza e do Nordeste de Portugal Desde o Século XIII ao Século XVI (Com Referência à Situação do Galego Moderno)*, Lisboa, Fundação Calouste Gulbenkian / Junta Nacional de Investigação Científica e Tecnológica.
- MARTINS, Ana Maria (ed.) (1994.), *Clíticos na História do Português-Apêndice Documental*, vol. 2, Dissertação de Doutoramento apresentada à Faculdade de Letras da Universidade de Lisboa.
- MARTINS, Ana Maria (ed.) (2000), *Documentos Notariais dos Séculos XII a XVI*. Edição digitalizada.
- NETO, João António Santana (ed.) (1997), *Duas Leituras do Tratado Ascético Místico Castelo Perigoso*. Dissertação de Doutoramento, São Paulo, Faculdade de Filosofia, Letras e Ciências Humanas, USP, Edição revista por Irene Nunes.
- NOVO CAZÓN, José-Luis (ed.) (1986), *El priorato santiaguista de Vilar de Donas en la Edad Media (1194-1500)*, A Coruña, Fundación “Pedro Barrié de la Maza, Conde de Fenosa”.
- PARKER, Kelvin M. (ed.) (1975), *Historia Troyana*, Santiago, Instituto “Padre Sarmiento”.
- PARKINSON, Stephen (ed.) (1983), *Arquivo de textos notariais em português antigo*, Oxford, Edição digitalizada.
- PENSADO TOMÉ, José Luís (ed.) (1958), *Os Miragres de Santiago. Versión gallega del Códice latino del siglo XII atribuído al papa Calisto I*, Madrid, C.S.I.C. (Anexo LXVIII da Revista de Filología Española).
- PÉREZ RODRÍGUEZ, Francisco Javier (ed.) (2004), *Os documentos do tomo de Toxos Outos*, Santiago de Compostela, Consello da Cultura Galega.
- PIEL, Joseph (ed. crít.) (1942), *Leal Conselheiro*, Lisboa, Livraria Bertrand. Edição digitalizada, revista por João Dionísio e Sandra Alvarez.
- PIEL, Joseph (ed. crit.) (1944), *Livro da Ensinança de Bem Cavalgar Toda Sela*, Lisboa, Bertrand. Edição digitalizada, revista por João Dionísio.

- PIEL, Joseph-Maria (ed.) (1988), *A demanda do Santo Graal*. Edição concluída por Irene Freire Nunes, com introdução de Ivo Castro, Lisboa, Imprensa Nacional – Casa da Moeda.
- PORTELA SILVA, Ermelindo (ed.) (1976), *La región del obispado de Tuy en los siglos XII a XV. Una sociedad en expansión y en la crisis*, Santiago, Tip. El Eco Franciscano, Separata de *Compostellanum* 20.
- RODRIGUES, M^a Celeste M. (1992), *Dos Costumes de Santarém*. Dissertação de Mestrado apresentada à Faculdade de Letras da Universidade de Lisboa.
- RODRÍGUEZ GONZÁLEZ, Ángel (ed.) (1989), *Livro do Concello de Pontevedra (1431-1463)*, Pontevedra, Museo de Pontevedra.
- RODRÍGUEZ GONZÁLEZ, Ángel (ed.) (1992), *Libro do Concello de Santiago (1416-1422)*, Santiago de Compostela, Consello da Cultura Galega.
- RODRÍGUEZ NÚÑEZ, Clara (ed.) (1989), “Santa María de Belvís, un convento mendicante femenino en la Baja Edad Media (1305-1400)”, *Estudios Mindonienses* 5, pp. 335-485.
- ROMANÍ MARTÍNEZ, Miguel (ed.) (1989-93), *La colección diplomática de Santa María de Oseira (1025-1310)*, 3 vols, Santiago de Compostela, Tórculo Edicións (1989, 1989, 1993).
- ROMANÍ MARTÍNEZ, Miguel e RODRÍGUEZ SUÁREZ, M^a del Pilar (eds.) (2003), *Libro tumbo de pergamino. Un códice medieval del monasterio de Oseira*, Santiago de Compostela, Tórculo Edicións.
- SPONER, Margot (ed.) (1932-4), “Documentos antiguos de Galicia”, *Anuari de l’Oficina Románica de Lingüística i Literatura (Barcelona)* 7, pp. 113-92.
- SPONER, Margot (ed.) (1932-4), “Documentos antiguos de Galicia”, *Anuari de l’Oficina Románica de Lingüística i Literatura (Barcelona)* 7, pp. 113-92.
- TATO PLAZA, Fernando R. (ed.) (1999), *Libro de notas de Álvaro Pérez, notario da Terra de Rianxo e Postmarcos*, Santiago de Compostela, Consello da Cultura Galega.

“A BEACHED WHALE POSING IN LINGERIE” CONFLICT TALK, DISAGREEMENT AND IMPOLITENESS IN ONLINE NEWSPAPER COMMENTARY

Isabel Ermida*
iermida@ilch.uminho.pt

Abstract:

This article investigates the correlation between the expression of disagreement and politeness concerns in the comment boards of the British *Mail Online* newspaper website. It focuses on a specific news report case – that of an overweight female model suing an advertising agency for using her image to promote body shame. The readers’ linguistic response to the legitimacy of her claims is closely entangled with their reaction to how her body looks on the photograph used in the ad campaign and reproduced in the article. Bearing in mind the diversity of posts, which range from encouraging remarks to insulting observations, the paper aims to discuss the expression of disagreement and, ultimately, the emergence of conflict talk in a specific case of Internet discourse (asynchronous online commentary), the central topic of which is women’s body image. More specifically, it wishes to examine, from the perspective of Politeness Theory, how the various expressions of a specific face-threatening act – expressing disagreement – contributes to creating and reifying representations of femaleness and of physical perfection or desirability.

Key words: Politeness, Disagreement, Conflict, Face, Female body, Internet discourse, Asynchronous online newspaper commentary.

*

Introduction

If compared with face-to-face conversation, the verbal interaction in online newspaper comment boards exhibits peculiar features. To begin with, its asynchronous nature challenges the turn-taking system and the adjacency sequences which Conversation Analysis has proven to characterise every-

* Universidade do Minho, Centro de Estudos Humanísticos, Braga, Portugal

day dialogue (cf. Sacks *et al.* 1974), whereas its public and multiparty quality affects the formulation of opinions and the negotiation of agreement and disagreement. Brown and Levinson's Politeness Theory (1987) has successfully approached these issues, as well as those of face redress and conflict avoidance in interpersonal contacts. However, the anonymous character of the conversational exchanges online (*vd.* Donath 1999), together with the impersonal nature of the long-distance electronic contact, encourage a type of linguistic behaviour based on impolite and aggressive elements which tend to be dispreferred in face-to-face dialogues. Besides, when the topics are related not only to subjective and aesthetic preferences but also to ethical and moral issues, the conflictual potential increases and the confrontation assumes more serious and grievous forms.

The present article aims at analysing a specific case of Internet talk, at a time when the cybernetic space seems to win over face-to-face contact. It focuses on the comment pages of the British *Mail Online* newspaper website, where readers freely express their opinions about the topics covered. The article in question, published in Nov. 2011, reports on the case of a morbidly obese female model who is suing an advertising agency for allegedly misusing her image to promote body shame. The problematic ad campaign shows her in a sensuous photograph, reclining suggestively in a sexy outfit which reveals her more than abundant curves, with the following slogan: "Did your wife SCARE you last night?" The focus of this paper is the newspaper readers' linguistic response to the legitimacy of her claims, given that her parallel pornographic activities may suggest a publicity stunt. This response is closely entangled with their reaction to how her body looks on the photograph used in the ad campaign and reproduced in the article. Bearing in mind the complexity of posts, which range from encouraging remarks to insulting observations, the paper aims to discuss the linguistic forms and the discursive strategies of disagreement through which a specific case of asynchronous online commentary constructs female identity in general and women's body image in particular.

The article starts by briefly discussing the state-of-the-art regarding linguistic approaches to online communication. Secondly, it offers a synopsis of the methodological framework of Politeness Theory, especially as far as face management, conflict and disagreement are concerned. Then, the textual analysis section examines the online dialogues according to a threefold division: backgrounded (or implicit) disagreement, hedged disagreement and foregrounded (or explicit) disagreement. The exercise aims at debating the management of the speech act of disagreeing in terms of the face

of speakers and hearers as well as of the construction of the female body in Internet discourse.

1. Overview of online conversation criticism

Computer-mediated communication has been the object of numerous academic studies, especially since its generalisation towards the end of the 1990's. Many such studies are based on the classic premises of Conversation Analysis (CA, Sacks *et al.* 1974, Atkinson & Heritage 1984, van Dijk 1985), which concentrate on face-to-face contexts of linguistic interaction. Unlike these, online conversations are, owing to their asynchronous character, temporally and spatially divided. Besides, the management of the interactive space, which CA analyses in terms of turn taking, overlap, interruption, and so on, cannot apply to the delayed context of cybernetic communication. Similarly, questions such as register informality, with regard to the oral vs. written dichotomy, as well as the problems of temporariness, anonymity and multimodality, have also deserved critical attention, proving that computer-mediated dialogues assume a very specific nature, one that differs from regular conversations.

Before the emergence and generalisation of debate forums, it was mainly the email that absorbed the analysts' attention as far as online communication goes. Studies by Yates (1996), Baron (1998) and Crystal (2001) tried to integrate this interactive genre into existing communication models. Thus, they regarded it in one of two ways: a) as a message written in the traditional format but transmitted through a new electronic medium (including job applications, online hotel bookings, family letters); or b) as a form of oral speech which happens to be written for transmission purposes. Other contributions (Ferrara *et al.* 1991, Maynor 1994, Collot e Belmore 1996) have identified a mixture of different influences and styles in email exchanges – a sort of an "e-style", which is neither oral nor written, but a combination of the properties of both registers.

Baron (2003) also seeks to establish a comparison between email language and face-to-face conversations. The similarities she finds include informality (use of contracted forms, preferred coordinate clause constructions), succinctness (short messages intended for short answers) and the presupposition of temporariness which underlies communication. Among the differences stand out, in the three parameters, more radical informal usages in email than in face-to-face exchanges (such as colloquial forms of treatment, frequent omission of greetings, use of direct speech acts), a

greater variability of the response time (which is due to the asynchronous nature of email and, even if extended, is acceptable, unlike face-to-face exchanges which demand instant response), and the fact that email can be printed, edited and stored, unlike oral exchanges which, unless recorded, are purely ephemeral.

With the appearance of plural and simultaneous participation sites – such as chat-rooms, forums, newsgroups – the discursive nature of cybernetic communication acquires fresh nuances and captures new criticism. Marcoccia (2004), for instance, examines the structure of the so-called “online polylogues”, whose public and asynchronous character makes them differ from interpersonal conversations. Firstly, the message exchanges in virtual forums are often truncated and the conversational sequences tend to be short. Unlike what happens at “live” interactions, the participants’ online speech often appears on an incorrect position of the sequential structure of the conversation. The participation pattern also differs from face-to-face exchanges, insofar as the author (the actual producer of the message) can hide behind the speaker (the persona holding the nickname that appears on the screen).

Other studies aim at understanding how plural online interaction allows participants to use – and to develop – communicative and pragmatic competencies. Lewis (2005), for instance, analyses French and British forums of political discussion (a case of what he calls “many-to-many interaction”) and he reveals a tendency for the fragmentation of communication, which changes from a multi-party interaction to a number of overlapping dialogues (each taking place between two speakers). On a different note, Clarke (2009) focuses on the discursive construction of interpersonal relationships among trainee teachers, by using a Critical Discourse Analysis framework to understand the legitimisation strategies used. Montero-Fleta et al (2009) examine the degrees of formality in two types of chat-rooms (one, Catalan, discussing football; the other, British, discussing the Palestinian crisis), whereas Savas (2011) attempts to understand the variation in the stylistic options of non-presential synchronic forums in terms of individual and contextual differences.

The occurrence of (im)politeness in computer-mediated communication – which the present article aims at discussing – has also been studied. Donath (1999) discusses the extent to which construction of “face” is important for the members of online communities, whose profiles seem to be carefully managed and which correspond, not to what those members are, but to what they wish to convey. Other scholars have ventured that

projecting an image of sophistication, education and courtesy has stopped being of interest to cybernauts. Baron (2003) precisely discusses the decline in “public face” in contemporary American society and the reflection the phenomenon has had in computer-mediated communication. The decrease in social class differences, the increase in inter-class mobility, the disconnection between schooling and economic success and the strong emphasis on youth culture have caused “less impetus to learn the fine points of etiquette or dress up for job interviews” and less public interest in developing “the sophisticated thought and language that higher education traditionally nurtures” (*idem, ibid.*: 90).

Upadhyay (2010) also investigates the language of the Internet, in particular computer-mediated reader responses (the same focus as the present article), with regard to the study of impoliteness, as well as identity. He posits that respondents may use linguistic impoliteness in three ways: a) strategically to communicate disagreements, b) to argue against an out-group’s ideological views, or c) to discredit ideological opponents. Upadhyay also defends that the employment of impoliteness is linked to the respondent’s identification with, or rejection of, a group’s ideological position and goals. Eisenchlas (2011) finds other reasons to explain the occurrence of impoliteness in online talk. He claims that the democratic nature of the Internet, the anonymity of its participants and the discontinuous character of the interaction turn the notions of hierarchy and deference, as well the corresponding questions of face, into more negligible concerns than those existing in face-to-face communication.

In a comprehensive, groundbreaking book, entitled *Cyberpragmatics: Internet-Mediated Communication in Context* (2011), Francisco Yus also relates the language of Internet – which he analyses in its various manifestations and corresponding genres, from email and chatrooms to blogs and personal webpages – with politeness issues. A curious fact which Yus (2011: 257) points out is that the online use of politeness strategies can be “chosen by the user or they can be imposed by a moderator of the system used”. Besides, Yus (2011: 270) claims that the Internet “is particularly appropriate for an analysis of the trans-cultural differences in the use of politeness.” After all, users from all over the world employ this electronic medium to communicate, both asynchronously (as in email) and synchronously (as in chatrooms). The specific nature of online talk influences the use, or misuse, of politeness. As Yus (2011: 263) puts it, “the lack of physical co-presence and the reduced nonverbal contextual support influence the choice of a particular politeness strategy”. But if the Internet lacks resources otherwise

available to speakers in everyday communication, it provides alternative conversational crutches that are non-existent in other written media, such as letters or the press. As Yus notes (2011: 168), an emoticon can soften the propositional content of an utterance or disguise its illocutionary force, much along the lines of what politeness does.

Having set the genre of texts on which we will focus, it is necessary to briefly discuss the theoretical tools to use in the analysis. Next, we shall review the concept of face, a central one to politeness theory, as well as to studies of disagreement and identity.

2. Face management

The vast and ever-growing scholarship on politeness attests the undeniable vitality of the topic. Notwithstanding the diversity of the approaches, most of the existing research reaffirms the impact of Brown and Levinson's pioneering work. Actually, the core concepts of the Theory of Politeness (1987) continue to inspire a steady flow of academic studies.

Brown and Levinson's Theory of Politeness rests upon the notion of 'face', which echoes the common expression "losing face" (and, by extension, getting embarrassed or humiliated – 1987:61). The concept is borrowed from Goffman (1967: 306), who defines it as "an image of self delineated in terms of approved social attributes" and, consequently, as "the positive social value a person effectively claims for himself". Hence the parallel notion of "face work", which covers all the actions an individual undertakes so as to act in accordance with the demands of face. These actions serve to counterbalance any "incidents" which may threaten face and to protect the speaker from his or other speakers' embarrassment. The culturally determined repertoire of face protection strategies is dual: "defensive" (protecting our own face) and "protective" (protecting other people's face) (Goffman 1967:310).

The theoretical implications of Goffman's fertile concept are numerous, complex, and sometimes polemical, as is the case of its purported universality (see, e.g. Mao 1994 and Earley 1997). Goffman (1967:319) does concede that "underneath their differences, people everywhere are the same", and he claims that every society encourages its members "to be perceptive, to have feelings attached to self and a self expressed through face". Much Anglo-American bibliography on politeness, as Bargiela-Chiapinni (2003:1462) points out, tries to redefine the relative weight of the concept of "face" in the West as opposed to its manifestations in other cultures. In Japan, empathy, reciprocity and dependence are overriding principles (cf.

Hill *et al.* 1986, Ide 1989), whereas in China face acquires quite different nuances (Lim 1994).

Brown and Levinson defend the universality of face in the very title of the 1987 book (*Politeness: Some Universals of Language Usage*), as well as in explicit passages in the body of the text: "we are assuming that the mutual knowledge of members' public self-image or face, and the social necessity to orient oneself to it in interaction, are universal" (1987:62). Although they admit that "the core concept is subject to cultural specifications of many sorts" (1987:13), Brown and Levinson establish the concept of Model Person, which is allegedly "a reasonable approximation to universal assumptions" (1987:84). Face, or the "public self-image" which the Model Person wishes to project, is twofold: positive and negative. *Positive face* is the wish that our image be appreciated and accepted by others, whereas negative face is the wish that our space be protected from intrusion and imposition. The former bears on the construction of identity and consensus, and on our attempt to be "ratified, understood, approved of, liked or admired", whereas the latter has to do with territorial integrity and our attempts not to be disturbed.

In social interaction, some situations of verbal and non-verbal communication intrinsically threaten face or, in Brown and Levinson's words (1987:65), "run contrary to the face wants of the addressee and/or of the speaker". These situations consist in the so-called face-threatening acts, which jeopardize, on the one hand, either the speaker's or the hearer's face, and on the other hand, either their positive or negative face. For instance, when a speaker apologises, s/he is admitting to incorrect behaviour, thus threatening his/her positive face, but if the speaker thanks the hearer, the former jeopardizes his/her negative face, by incurring in a debt and allowing the hearer to ask for a make-up favour later on. On the hearers' end, a criticism by the speaker is obviously a threat to positive face, since it indicates that the speaker does not approve of the hearer's behaviour, whereas a request threatens the hearer's negative face, since it invades the hearer's territory by asking him/her to do something. In face of such threats, speakers may simply refrain from doing the face-threatening act altogether.

However, if speakers cannot avoid doing it or if they deem it preferable to run the risk, they may use certain strategies to minimize such a risk and protect the mutual vulnerability of face. In this case, they have two possibilities to choose from (cf. Brown and Levinson 1987:69): 1) to do the face-threatening act "off record", that is, in an indirect way, by providing the hearer with hints such as metaphor, irony, and understatement; and 2) to do it "on record", that is, explicitly and ostensibly, in one of the following ways:

- a) Doing the act “without redressive action, baldly”, that is, in a direct and clear way, without any mitigating forms (for instance, and order would be performed by means of a straight imperative form)
- b) Doing the face-threatening act “with redressive action”, that is, by employing mitigating action, be it *positive politeness* (for instance, by using in-group identity markers, avoiding disagreement, exaggerating interest, telling jokes, etc.) or *negative politeness* (for example, apologising, being indirect, not coercing, impersonalizing, etc.)

This article will discuss how a major strategy of positive politeness – seeking agreement – gets to be manipulated in the comment pages of the *Daily Mail* online. But before doing so, a few preliminary observations regarding the issues of conflict and disagreement are in order.

3. Politeness, conflict and disagreement

At the outset of their book (1987:1), Brown and Levinson’s state: “(...) politeness, like formal diplomatic protocol (...), presupposes [a] potential for aggression as it seeks to disarm it”. And they add that politeness “makes possible communication between potentially aggressive parties” (*ibid.*). The role of polite behaviour in taming down aggressive urges is a key factor to successful social interaction. Identifying possible symptoms of conflict requires a “constant vigilance” and the knowledge of what Brown and Levinson call a “precise semiotics of peaceful vs. aggressive intentions” (*ibid.*) As such, politeness constitutes an important form of “social control” (1987: 2).

This conception of politeness as an antidote to conflict and aggression is also present in Leech (1983), whose Principle of Politeness also aims at preventing or solving any incompatibility between speakers. In Leech’s words (1983:82), the regulatory role of this principle is to keep “the social equilibrium and the friendly relations which enable us to assume that our interlocutors are being cooperative in the first place”. Leech’s Principle of Politeness is divided into six maxims, namely Tact, Generosity, Approbation, Modesty, Agreement and Sympathy. The Agreement maxim, in particular, reads: a) “Minimize disagreement between self and other”, and b) “Maximize agreement between self and other”. Minimizing disagreement is listed first because, as Leech ventures, “avoidance of discord is a more weighty consideration than seeking concord” (1983: 133). Leech also holds that “there is a tendency to exaggerate agreement with other people, and to mitigate disagreement by expressing regret, partial agreement, etc.” (1983:

138). Other ways speakers employ to avoid disagreement is the use of indirect speech acts, by means of such strategies as modalization and passivization. In short, the more indirect the speaker, the more polite and the more capable of avoiding conflict.

In Brown and Levinson's model, the idea of conflict management is closely related to the notion of power, one of the three "sociological factors" which are crucial to "determining the level of politeness which a speaker will use to an addressee" (1987:15), namely: Power, Distance and Ranking of the Imposition. The more powerful the speaker is, the less room there will be for conflict to arise. Indeed, in many situations of social interaction the role of the relative power of participants is crucial, especially if the interlocutor is "eloquent and influential, or is a prince, a witch, a thug, or a priest" (1987:76), in which case he may well impose "his own plans and his own self-evaluation" (*ibid.*). And, of course, by imposing our own opinion and evaluation we threaten the positive, as well as the negative, face of the hearers, in that we show we do not respect or appreciate their opinion, and intrude upon their territory by requiring them to accept ours. One thing is to seek agreement; another is to *impose* it.

The issue of agreement vs. disagreement is actually well established in Brown and Levinson's model. 'Seeking agreement' and 'avoiding disagreement' are complementary strategies that aim at ascertaining common ground between the speaker and the hearer, thus "indicating that S and H both belong to some set of persons who share specific wants, including goals and values" (1987: 103, see also 112-3). So important is the desire to agree, or to appear to agree, with the hearer that speakers may resort to 'token agreement'. In an earlier study by Sacks (1973) a "Rule of Agreement" explains "the remarkable degree to which speakers may go in twisting their utterances so as to appear to agree or to hide disagreement" (Brown and Levinson, 1987: 114). This is the case of responding to a preceding utterance with "Yes, but...", rather than a blunt "No". We will see occurrences of token agreement, or what I call "disagreeing by agreeing" in the present corpus.

In 2004, Miriam A. Locher wrote an important contribution to the study of disagreement from the point of view of Politeness Theory. In *Power and Politeness in Action: Expressing Disagreement in Oral Communication*, she writes: "Conflict can be argued to link the exercise of power, politeness and disagreement on a general level" (2004: 94). Along the lines of Waldron and Applegate (1994:4), who define verbal disagreements as "a form of conflict" insofar as they are "characterized by incompatible goals, negotiation, and

the need to coordinate self and other actions”, Locher claims that disagreeing speakers are in conflict not only in terms of content but also in terms of protecting the hearer’s face, as well as their own. Getting their point across without sounding presumptuous or injuring the hearer’s image causes friction and potentially leads to conflict, which is where politeness comes in. Locher (2004) establishes eight categories for expressing disagreement, with different degrees of politeness: the use of hedges, giving personal or emotional reasons for disagreeing, the use of modal auxiliaries, shifting responsibility, stating objections in the form of question, the use of but, repeating an utterance by a next or the same speaker, and non-mitigated disagreement. We will see how some of these categories occur in the corpus under focus.

Previous studies on disagreement have shed light on other interesting facets of the phenomenon. To begin with, disagreements need not be negatively charged, or psychologically detrimental. Schiffrin (1984: 329) elaborates on “sociable arguments” and claims that expressing disagreement may be part of the expected speech situation and thus be a source of enjoyment. In a later study (1990: 241), she mentions that although arguments may at first sight look like the epitome of conflict talk, they may in fact be regarded as a “cooperative” or “competitive” way of speaking. Other authors also view disagreement – or its strongest version, argument – as a reaction that is not necessarily a dispreferred or negative one. Charles and Marjorie Goodwin (1990: 85), for instance, claim that “despite the way in which argument is frequently treated as disruptive behaviour, it is in fact accomplished through a process of very intricate coordination between the parties who are opposing each other”. And Kotthoff (1993: 193) defends that “once a dissent-turn-sequence has been displayed, opponents are expected to defend their positions”, showing fewer reluctance markers, which makes disagreement become a “preferred” response.

Yet, disagreeing does carry a risk of confrontation and hence of negative psychological tension. Kakavá (1993: 36) explains the vulnerability of the disagreeing speaker, who may face criticism and be the object of counter-attack:

Since disagreement can lead to a form of confrontation that may develop into an argument or dispute, disagreement can be seen as a potential generator of conflict. Not only can disagreement create conflict, but it can also constitute conflict since an argument is composed of a series of disputable opinions or disagreements.

Perhaps because of this conflict potential, disagreement tends to be regarded as a dispreferred answer. Pomerantz (1984) distinguishes between weak and strong disagreements and claims that the former, as dispreferred answers, resort to such delaying strategies as hesitations, requests for clarification, "no talk", turn prefaces, partial repeats and other repair initiators (1984: 70). Strong disagreements, on the other hand, do not make use of any of these devices.

In *Learning Politeness: Disagreement in a Second Language* (2009), Ian Walkinshaw offers a useful taxonomy of disagreement, which includes four categories. By way of example, the fictional situation is that of a speaker responding to a question of whether he likes a second-hand couch:

- a. Explicit / direct disagreement: "I don't like this couch at all." Having only one literal meaning, this FTA will only be performed "if the speaker is not concerned with retaliation from the hearer" (Walkinshaw 2009: 73)
- b. Disagreement hedged with positive politeness: "It's a nice couch, but I don't like it." In this case, disagreement is softened by expressing appreciation of the hearer's likes, wants and preferences.
- c. Disagreement hedged with negative politeness: "You've obviously set your heart on it, but I don't like it." This includes the mitigating strategies oriented towards the hearer's desire to act freely as s/he chooses.
- d. Implied disagreement: "Um, well, it's certainly an interesting colour..." This roughly corresponds to Brown and Levinson's "off-record" strategies, such as hinting and vague, unfinished sentences, which free the speaker of just one communicative intention, and thus of the responsibility for the FTA.

This taxonomy is to some extent reminiscent of Scott's (2002) distinction between two primary types of linguistic disagreement: 'backgrounded' versus 'foregrounded' disagreement, which "exist on a continuum of increasing explicitness and escalating hostility" (2002: 301). (We will see in the textual analysis section how this continuum operates in the present corpus.) Scott also divides the latter into two patterns of disagreement in terms of targets, namely "collegial disagreement" versus "personal disagreement" (which includes "ad hominem attacks").

In *Impoliteness: Using Language to Cause Offence* (2011), Culpeper establishes a further parallel between Politeness Studies and what he correctly labels Conflict Studies. Two subfields of research into conflict are directly related to impoliteness. First, interpersonal conflict bears on the existence of difference or incompatibility between people and on the interaction between parties that have opposing viewpoints, interests or goals.

Secondly, conflict in discourse involves any type of verbal or non-verbal opposition taking place in social interaction, ranging from disagreement to disputes. Both subfields are subsidiary to impoliteness in the following way:

If impoliteness involves using behaviours which attack or are perceived to attack positive identity values that people claim for themselves or norms about how people think people should be treated, then it involves ‘incompatibility’, ‘expressing opposing interests, reviews, or opinions’, ‘verbal or non-verbal opposition’ – it is intimately connected with conflict. (Culpeper, 2011: 5)

A final bibliographical cornerstone that requires mention is a very recent special issue of the *Journal of Pragmatics* (Sep. 2012), devoted to the study of Disagreement. In its Introduction, Angouri and Locher (2012) propose four key premises to theorising disagreement. The first is the everyday nature of the phenomenon. The second is that disagreement may be not only tolerated but expected, instead of being an exceptional speech act. The third is that disagreeing is not necessarily negative. And the fourth is that disagreeing “will have an impact on relational issues (face-aggravating, face-maintaining, face-enhancing)” (2012: 1549). Although some of the articles in the volume revolve around institutional settings and workplace scenarios (cf. Marra 2012, Angouri 2012), others crucially focus on online discourse. Langlotz and Locher (2012), for one, identify instances of the expression of emotional stance in news website postings through conceptual implication, explicit expression, and emotional description. And Bolander (2012) analyses the use of (dis)agreement in personal / diary blogs, arguing that even though the participation framework of blogs encourages explicitness, there is a greater need to signal responsiveness explicitly when readers address other readers than when readers address bloggers.

We will next see the forms which the participants in online conversations – the focus of the present study – use to express agreement and disagreement, and we will try to identify the different linguistic strategies used to express alignment and approval, on the one hand, and confrontation and rebuttal on the other. Perhaps the former bear on the feeling of a shared experience of events and situations (on the concept of networked community, see e.g. Castells 2000) and the latter on the influence of the factors of Distance, Anonymity and Third-Party opinion. The following sections seek to test these hypotheses.

4. Sample description

The corpus of texts under focus in the present article is a set of online written productions of a dialogical nature – or, along Maroccoia's lines (2004), of a "polylogical" character, since they involve more than two participants, and many of the utterances respond to more than one speaker at a time. They were published on the days following the publication of an article on the *Mail Online* newspaper website, entitled "Supersized model describes horror after discovering her 'scary' image was used to promote dating site for men who want to cheat" (Nov 9, 2011). The article crucially includes an illustration, shown next:



Fig. 1 – Illustration of the *Mail Online* article

Together with the image is the following subtitle: "Big is beautiful: Jacqueline, who unwittingly features in the ad, believes the marketing material 'suggests, blatantly, that fat people are patently undeserving of love and loyalty.'" These cues are not only obviously provocative but successfully so. The number of responses to the article is, not surprisingly, high and the intensity of the opinions expressed quite remarkable, in such a way that the discursive situation becomes a very interesting case study of argumentative discourse in general and "multi-party argument" in particular (Maynard

1986, Goodwin and Goodwin 1993: 100). At the same time, it lends itself to an analysis of the management of face and the violations of the politeness strategies usually at work in face-to-face interactions. On the whole, the corpus includes 369 posts, each of which containing one or more sentences, which amount to 18757 words.

5. Content discussion

The core of the article is that the morbidly obese model, Jacqueline, claims that a picture from a photo-shoot she unsuspectingly made was used, without her permission, to advertise an extra-marital affairs website and to promote body shame. The fact that the photo-shoot was patently suggestive in sexual terms, together with the circumstance that Jacqueline runs a pornographic website, are rather incriminating factors that do not help her case. The complex range of reactions to the article spans not only her physical appearance – from which spring judgements as to body perfection, desirability and health – but also the ethical issues involved in her side activities as a porno figure.

Therefore, the expression of disagreement covers a broad scope of topics, which are raised, debated, dropped and recovered along the argumentative sequence. More specifically, the various targets of disagreement in the corpus are the following:

- a. Disagreeing about obesity being beautiful (The newspaper article subtitle uses the expression “Fat is beautiful”)
- b. Disagreeing about obesity being ugly or deserving attack (Jacqueline’s court case rests upon the complaint of body shame caused by obesity)
- c. Disagreeing about objectification of women (Many commentators view Jacqueline as a female scapegoat)
- d. Disagreeing about adultery (The advertisement is for a dating agency for married men)
- e. Disagreeing about hypocrisy (Jacqueline complains of her body being misused even though she exploits it herself for sexual purposes)

The complexity of the analysis also bears on the fact that the disagreement turns occur in a sequence that is not dyadic, but multiparty, admitting different strands of discursive input. Besides, the speech act of disagreeing is closely related to, and actually overlaps with, other speech acts, such as criticising, protesting, reprimanding, etc. Actually, from a speech-act the-

ory perspective, the analysis of the linguistic sequences is particularly challenging, since the illocutionary force continuum ranges from warnings and overt condemnations to encouragements and expressions of support, whose boundaries are often blurred. At an interpersonal level, two sides build up from the beginning of the exchange: a support side and a rejection one, a divide which immediately feeds a potential for disagreement and conflict. From a Conversational Analysis perspective, the asynchronous nature of the exchanges makes the turn-taking system inoperative, as well as the occurrence of interruption or overlap. Yet, it is interesting to note that the turns do occur in response to previous ones. Actually, some posts constitute direct replies to others, whereas other posts exist independently, ignoring previous conversational input. It should also be noted that long turn-takes alternate with very short ones – sometimes, even one-word replies (as is the case of "Oink!").

From the standpoint of the Theory of Politeness – which is the primary theoretical tool in hand – the texts exhibit a great degree of complexity. As an anonymous forum, the comment pages of the newspaper encourage the free expression of thought, which may explain why some posts are such ostensive infringements of the politeness principles, especially positive politeness. Also, the fact that the respondents are speaking about a third party – Jacqueline, who is not taking part in the polylogue – may partly be the reason for the expression of strong, confrontational opinions, which often subvert the interactional protocol presiding over daily face-to-face conversations. As Leech (1983: 133) claims, "politeness towards an addressee is generally more important than politeness towards a third party".

We will divide the analysis into three parts, in accordance with the three major patterns of disagreement found in the texts.

5.1 Backgrounded disagreement: Implicitness and indirectness

Since disagreeing has the potential to damage another person's face, the speaker will normally avoid performing the FTA, or try to minimize its impact. Going "off-record" allows the speaker to remain in the realm of the implicit, where "it is not possible to attribute only one clear communicative intention to the act" (Brown and Levinson 1987: 211). Off-record strategies demand inferential efforts by inviting conversational implicatures, since they typically break Grice's cooperative maxims (*ibid*: 213). Scott (2002: 74) calls off-record disagreeing "backgrounded disagreement" and confirms

- I’m a roly polly pudding and pie and i kisses the boys and squashed them with my enormous mellons and then i flattened them with my enormous backside – Tom, Chicago.

Another metaphor, mixed with metonymy (part for the whole), compares Jacqueline with a country’s province, by means of a reference to the postal code she would deserve if she were one:

- If she lived in Britain, she’d get her own postcode. – Adam Mann, UK

The use of rhetorical questions, another case of infringement of the Quality maxim, helps the speaker to soften the strength of the disagreement. Brown and Levinson (1987: 223) explain that “to ask a question with no intention of obtaining an answer breaks the sincerity rule on questions”. In the next passage, the commentator disagrees that an obese woman may be a legitimate target of adultery by counter-arguing that marriage also affects men’s body negatively. The commentary exhibits a series of rhetorical questions, signaled with several question marks at a time, which “leave their answers hanging in the air” (*ibid*: 223) and save the speaker the responsibility for the propositional content of the utterance. The use of hedges, such as “I presume”, “maybe”, and “just (a thought)”, also indicates that the speaker is going off-record:

- Men saying that they would dump their wives if they got obese when they had been a size 12 at marriage – I presume still have their full heads of hair, lithe muscular bodies, full sets of healthy teeth and the sex drive they had when THEY got married??? Or are they balding, pot bellied, couch potatoes with bad breath and more hair coming out of ears and nose than on the whole of their head, who do more talking about sport and sex than doing it??? Maybe your wife is disappointed with your looks too but loves you anyway?? Just a thought. – Me again, Newbury, Berkshire

A very similar case takes the opposite argumentative stance, defending the male’s point of view. The disagreement, once again, resorts to rhetorical questions, mixed with hedges such as “really”, “some” and “surely”:

- So, according to her, if a couple marries when she is like a size 6 and then balloons to a size 32 and then the husband cheats, because he is no longer attracted to his wife that the entire situation is HIS fault? Surely it is her fault as well?! – Malia, Honolulu, Hawaii, USA

The use of irony, by which the speaker says the opposite of what s/he means, also expresses implicit disagreement. The corpus exhibits several instances of ironical utterances. In the following passage, a commentator disagrees with the idea that being fat deserves derision, by using the adjectives “interesting”, “glad” and “sane” in a patently ironical way:

- Interesting, the fat-bashing. But identity-theft is okay? Glad there are some sane people commenting. – Amber, London

A very good example of irony is the following utterance, where a female commentator shows her dislike for obesity and her implicit disagreement regarding the corresponding values:

- It’s official. I’m never eating, ever again. – Tina, Derby

Likewise, the following dialogue sequence between two commentators – Shaun and Andy – reveals an ironical response on the part of the speaker who initiated the exchange. Shaun initially disagrees with the purported “hate” of which Jacqueline is a victim, and Andy replies with a play on words regarding the expression “having too much on one’s plate”. Shaun’s closing remark shows that, contrary to what he says, he obviously did not “like” Andy’s “enthusiasm” for his joke. The use of the plural pronoun helps to signal his dual illocutionary intention:

- What’s with the hate on her? Doesn’t she have enough on her plate already?
– Shaun, Liverpool, 09/11/2011 15:29 / From looking at the size of her Shaun, I don’t think that there’s ever much left on her plate..... – Andy, Frodsham, 09/11/2011 16:23 / We like your enthusiasm for our joke thank you Andy. – Shaun, Liverpool, 9/11/2011 16:53

Unlike irony, sarcasm does not mean the opposite of what is said (thus violating the quality maxim); instead, it means *more* than what is actually expressed (thus violating the quantity maxim), which is why it also goes by the name “understatement”. So as to express disagreement that Jacqueline is beautiful, the four following commentators resort to sarcastic understatement:

- Note to self: initiate strenuous exercise program immediately. – K.M., Coeur d’Alene, Idaho, USA
- Is this the American version of TWIGGY ??? – Frank, Slough

- Somehow I have a nervous feeling about a possible missing little pet's location during the photoshoot. – Jon, NY
- Supersize model. The oxymoron of the 21st century. – Cat Lover, London
- I look forward to their next ad: "Did your HUSBAND scare you last night?" Bet that could create some publicity, too – with the right illustration of course. – Damascena, Town

The following is a very curious case which uses understatement to criticize understatement. In reply to an earlier comment by a woman ("God forbid if a woman puts on a bit of weight! – Pam, Rutland), a male commentator replies:

- Quite possibly the understatement of the year.....size 32 is now 'a bit of weight'. – Adam, Glasgow

Two other commentators employ the adjective "enormous" sarcastically, and they establish a synecdoche between her 'problem', or her 'issue' (something that is part of her), being "as big as" she is:

- Well she is correct about there being an enormous problem. – Taylor, New York City USA
- Use of the word "enormous" when describing her issue was an unfortunate word choice... – Lauren, London

Hinting is also a typical strategy of indirectness, and as far as disagreement goes, it also plays the function of backgrounding the illocutionary force of the utterance. Hints violate the Relation maxim in that they require the recipient to establish the relevance of the utterance to the issue in hand. The following line hints at the practical difficulties that an overweight female body can pose to a male lover. Though the impoliteness is not overt, it is obvious that no redress is offered, only the off-record statement of a (tabooed) fact:

- You will never get in there. – Steven, Surrey

A particularly interesting comment happens when a reader responds to Jacqueline directly, by using the second-person pronoun. Unlike most other comments, he speaks *to* her, not *about* her. He begins by quoting her words in inverted commas, and goes on to use several off-record devices,

such as rhetorical questions, ellipsis (in the form of suspension marks) and irony (saying she is a “complicated woman” whom no one understands):

- “I find the very idea that there exists a business based solely around the facilitation of infidelity appalling.” Let me guess, you still accept paid memberships from married men? Thought so. Lovely lark this one, doesn’t like her body being exploited.... unless it’s her own porn site. Doesn’t like infidelity.....unless the infidelity is visiting her site. She is a complicated woman and no one understands her but her ice cream sandwiches. – Greg, Canada

Finally, another off-record strategy Brown and Levinson mention and which may be used to express implicit disagreement is the employment of generalizations. In the following comment, a speaker points out that although everybody nowadays criticizes smoking, criticizing obesity is still generally condemned:

- It’s socially acceptable to tell smokers now that smoking will kill them and that it’s awful and yucky. It wasn’t always acceptable back in the day to speak aloud about it. When will it become acceptable to tell people “you are overweight and killing yourself?!” – Odell, Tampa, FL,

Likewise, the following comment is constructed on the basis of generalization. The use of discursive subjects like “people” or “you” (which plays the role of impersonal “one”) helps the speaker express her disagreement in a backgrounded way:

- People know that if they’re overweight it’s very unhealthy. Broadcast and print media carry many stories and statistics about the health risks of obesity. Unless you are talking about heart-to-heart talks with people to whom you are close or doctors talking to patients, I hope it’s never socially acceptable to tell people they’re fat and going to die. It’s rude, cruel, and pointless. Criticizing strangers because you believe they are overweight won’t make them change their behavior. – Carol, U.S.A.

The same can be said of the following general statements, the first about the diversity of the human body in general, the second about the “fat *acceptance* movement”, which is another case of impersonalization (through nominalization):

- Women (just like men) come in all sizes. Fat, thin and in-between. These ads are horribly degrading to women in general. I see her point. – rajapatee, orlando fl
- The fat acceptance movement is literally killing people, and those behind it should be stopped. – Kate, Washington DC, USA

5.2 Disagreement hedged with positive and negative politeness:

The occurrence of comments that involve disagreement but attempt to soften its strength in a polite way cover several linguistic strategies, some of which are aimed at positive face, others at negative face. On the whole, the corpus exhibits several examples of the two cases. Positive politeness strategies are oriented toward the "positive self-image that [the speaker] claims for himself" (Brown and Levinson 1987: 70) and can be used to redress the face-threat inherent in disagreement. Walkinshaw (2009:73ff) also studies the use of positive politeness in disagreements, which he regards as an attempt to attenuate disagreement by expressing appreciation of the hearer's likes, wants and preferences.

The two following examples are interesting cases of a speaker disagreeing that morbid obesity may be regarded as beautiful while protecting the face of moderately obese readers. Indeed, so as to shield the positive self-image which the latter naturally claim for themselves, the speakers preface their disagreement by introducing an apology (see Locher 2004: 134), in the first case, and a concession in the second ("Big may be beautiful..."), both of which are followed by *but*:

- No offense to big women, *but* this is disgusting. – Amanda, Houston, TX
- Big may be beautiful, *but* morbidly obese is disgusting, no apologies.

Claiming common ground by using in-group identity markers and first-person plural pronouns (we, us) is a positive politeness strategy which Brown and Levinson (1987: 107, 127) also stipulate: "by using an inclusive 'we' form, when S really means 'you' or 'me', [the speaker] can call upon the cooperative assumptions and thereby redress FTAs". The following speaker tries to assuage her disagreement by using a solidarity strategy and pretending she is part of a group (see also Chilton 1990: 217). In other words, she disagrees with the views expressed by a group but pretends she is part of that group and that their self-image is hers too. The use of "let's" (contrac-

tion of ‘let us’), in particular, implies a mitigated order, with the ‘exercitive’ illocutionary force (cf. Austin) disguised:

- This is a human being *we* are talking about. Why do people put so much emphasis on what’s on the outside? Shouldn’t *we* be more tolerant and understanding? *We* don’t know a person’s circumstance as to why or how they became that way. *Let’s* stop hurting each other and fix *ourselves* instead. – Victoria, Los Angeles, CA USA

Another curious example of the use of in-group identity markers, in particular address forms, comes in reply to a previous comment by a male participant (“You cannot pile on a trazillion pounds then accuse your partner of being shallow if they no longer desire you.” – Toni, Herts). The respondent resorts to a term of endearment (which may have some ironical undertones) so as to soften the strength of her disagreement, which is otherwise constructed on the basis of negation (note the recurrence of the adverb *not*):

- ‘Desire’ *isn’t* love, *sweetie!* Love must fuel desire, *not* the other way around. (...) The best sex is about love, *not* lust, and *not* just about what a person looks like. They’re still the same person inside, *no* matter what the outside is like. – ShowMeTheChiffon, UK.

The use of hedges is a common politeness strategy that attenuates the threat to face by tentativizing the illocutionary force underlying the utterance. As Brown and Levinson explain (1987:145), a hedge is “a particle, word or phrase that modifies the degree of membership of a predicate or a noun phrase in a set; it says of that membership that it is *partial*, or true only in some respects (...)”. The following speaker tries to redress the positive face of fellow-commentators by pretending not to be certain of his/her disagreement. Hence the use of “I don’t know”, “all that”, “just” and “well”:

- *I don’t know if* I am *all that* against body shaming. This is not *just* very unattractive, it’s a serious health issue. What ever happened to the days of old when women were normal sized with curves? I’m not talking about the curves of Mt Everest but *well, just* nice. Now it’s either stick thin to the point of bones or so fat you want to throw up! – Carer, Austin

Along similar lines, the following comment includes such hedges as "I wish people would stop trying...", instead of using, say, a direct imperative ("Stop trying..."):

- I wish people would stop trying to make out like it's fine to look this way, it's not natural. – Lucy, NY

A common discursive strategy found in the corpus is disagreeing by agreeing. This of course helps save the face of the hearer whose opinion we do not actually share. Brown and Levinson (1987:114-5) refer to "pseudo-agreement" in situations where a speaker begins by stating agreement but "carries on to state his own opinion which may be completely contrary to that of the first speaker". The use of 'but' (or any other adversative conjunction, like 'however') usually ensues. 'But' is actually a key marker of disagreement. In grammar, it is the quintessential oppositional particle, and in pragmatics its impact depends on the position within the utterance. The two following cases both employ 'but' following the explicit use of the verb "agree":

- I *agree* that women are under a lot of pressure to be 'perfect' and I myself constantly feel pressure to lose weight in order to effectively be 'a better person' *but* this girl is disgusting. – mimi, surrey.
- C'mon, as a woman I have to *agree* that if I was married to someone who let themselves go that badly, I would look elsewhere, too. A size 12 is one thing, *but* morbidly obese is another. It's unhealthy and unattractive, and if I were to get that big not due to a medical condition, I wouldn't blame my husband for leaving me. – Lynn, Tampa, Florida, USA

The examples of disagreement through pseudo-agreement abound in the corpus. Consider the following reply to a previous comment, which is itself a disagreement hedged with an apology ("I'm sorry, no one should be proud of being over weight" – laci, Essex):

- *Agreed. However*, no one should be proud of being underweight, either. Both extremes present health issues, but being very thin is sold as beauty in today's society. – Just Sayin', USA.

A similar dyadic exchange takes place when a female commentator criticizes men for being shallow ("Most men I feel are just shallow and do not think with their brain but what is between their legs!" – Emma, Kent):

- Emma, whilst I *agree* that there are men who are shallow, I can also say that there are women who are the same. Women also make catty remarks about men AND other women. – Bloke, Here

Other expressions of pseudo-agreement do not use the verb “agree” explicitly:

- *I can see* all the moral problems *but* that is a seriously funny ad! – JJ, Sydney.
- *I’m all for* the fact that woman come in all shapes and sizes *but* when you know it’s interfering with your health, it’s time to sort it out. – Me, Essex.

Incidentally, the use of ‘no’, the ultimate disagreement marker, may actually mean ‘yes’. Consider a possible reply to the statement: “She is not healthy”. Saying “No, she isn’t” is actually equivalent to “Yes, you’re right”. The following passages use ‘no’ as expression of agreement followed by ‘but’ and the corresponding disagreement (which in the first example ends on a rather ‘impolite’ interjection – “Jeeze grow up”):

- As for the size of the woman, *no* she is not a healthy size at all (*however* some men do like this) *but* do people on here have to remark about it like they’re still in primary school? Jeeze grow up. – FW, UK
- *No*, being morbidly obese is not healthy *but* it’s not your business either. – See it all the time, USA

Occurrences of negative politeness bear on linguistic strategies directed at respecting the recipient’s freedom of action, that is, their negative face, their wish not to be intruded upon or hindered. Cases of disagreement being mitigated with negative politeness are not abundant in the corpus, but they do exist. The use of apologies is a typical case of negative politeness, as Brown and Levinson state, and they pop up now and again:

- *Sorry* but size 32 is not “beautiful” - Deeze, North.
- I’m *sorry* no one should be proud of being over weight – Laci, Essex
- Beauty is health, and whether you be under or overweight you just do not fall in the healthy category, *sorry!* Let’s stop kidding ourselves! – Lee, London

The use of “sorry” is sometimes strengthened by an affectionate term (“love, “Jackie”), again with ironical tinges:

- *Sorry love* but if my nearest and dearest became a size 32 through simple over-eating I would definitely go get a new one.
- No, *love, sorry*, you are NOT beautiful. You are obese and that is UGLY – Louise, Gillingham.
- *Sorry*, but someone this fat is more often seen as horrific by the general populace than as, *gag*, "beautiful". Saying this woman is beautiful and attractive is like saying the same thing of a stumbling drunk or "Oooh look at the handsome meth-head!" *Sorry Jackie* – you're horribly obese and there's nothing attractive about it. – Shwa, Denver,

A different discursive situation is the following. By telling (fat) female readers that nobody is preventing them from being fat, the commentator hedges his disagreement that obesity may be attractive, thus expressing his attempt to protect their negative face. In other words, he expresses his respect for their freedom to be as they prefer:

- A size 2 woman who sees this ad sees the message: "If I don't stay small, he will cheat." Sounds like it's working to me. Nobody's telling you that you can't be fat, ladies. Just don't expect that men will look at you as if you're not. It's that simple. – John, Northampton, PA USA

5.4 Foregrounded disagreement: explicitness and directness

Scott's (2002) category of "foregrounded disagreement", later taken up by Walkinshaw (2011), echoes Brown and Levinson's "bald-on-record" strategies for doing an FTA, which involve "doing it in the most direct, clear, unambiguous and concise way possible" (1987: 69). The reasons Brown and Levinson offer to explain such strategies – urgency, efficiency, negligible threat to face and vastly superior power of the speaker (*ibid.*) – differ from the ones Locher (2004: 143) presents to account for what she calls "unmitigated disagreement":

- a) when it is more important to defend one's point of view than to pay face considerations to the addressee (see also Kotthoff 1993);
- b) in contexts where the relationship of the interactants minimizes the potential risk of damage to the social equilibrium;
- c) when the speakers wish to be rude, disruptive or hurtful (see also Beebe 1995 and Culpeper 1996)

In the *Mail Online* corpus, it seems that overt impoliteness results from the spatial and temporal discontinuity of the verbal exchange, together with its anonymity. As communication is not face-to-face, with no eye contact or personal knowledge of the interactants, face concerns seem to weaken. Anonymity, in particular, seems to invest speakers with a certain sense of power, which is also linked to the notion that no retaliation – apart from a verbal one – is possible on the Internet. Curiously, Brown and Levinson (1987: 97) also mention this possibility when they state: “non-redress occurs [...] where S’s want to satisfy H’s face is small, because S is powerful and does not fear retaliation or non-cooperation from H”.

Some disagreements are so direct that the speaker simply states: “I don’t agree with you”. This is the case of the following comment, which is a reply to an earlier comment, referred to above (“I’m all for the fact that woman come in all shapes and sizes but when you know it’s interfering with your health, it’s time to sort it out.” – Me, Essex)

- *Don’t agree with you.* Size 16-18 is fine... there are a bunch of 16s in our women’s rugby team who are fitter than the average man tbh [to be honest] and they can all tie their boots up. – Man Of The People, UK

Other explicit disagreements are simply an interjection (of repulsion and nausea):

- Yukeeeeeeeeeeeeeeeeeeeee – Bob, England,

Albeit explicit too, the range of other bald-on-record strategies used in the corpus to express disagreement is slightly more elaborate, including lexical choice meant to demean the hearer’s face. Adjectives, in particular, are an easy way to foreground disagreement:

- The male mouthpiece for this dating site sounds like a *dirty, misogynistic* little creep and the words on that advert are utterly *offensive* and *demeaning* to women. – FW, UK,
- Imagine what the *dog-faced* men who use this site and mock this woman look like? No, don’t, it will put you off your lunch....- Violet Brown, Birmingham
- Lol. You’d have to be really really *sad* to engage in anything like this – male or female!!!! – Bexxx, LDN
- Wow and men really visit that site, they must be *short-sighted* or *desperate*!!!!- Dodger, Staffordshire, England

- *Disgusting*. I’m thinking that if anyone finds that sexy they are seriously *insecure* with themselves. – Silvermist, Capo Beach
- Completely *reprehensible*. Perhaps his next campaign could feature an amputee or a burns victim. – Lola, Manama
- *Sickening*, all around. – Pauline, Washington
- Juicy jackie.com??? I feel *sick* at the very thought. – Nobody, Bristol
- That is *frightening*. She should be *ashamed* of herself for looking that way. Even a genetic predisposition to being overweight does not explain a woman weighing 400 lbs. – Charlene, Arlington, tx, usa

Besides adjectives, the choice of nouns can also be a concise way of disagreeing – and insulting, by the way:

- “She says the ad is yet another unwelcome case of body shaming”. She must be an *idiot* then! What did she think was going to happen when she posed for the photographs? – PC99, UK

Calling Jacqueline a whale, which can be regarded as a metaphorical strategy for the speaker to go off-record, as we saw above, also occurs in obviously on-record passages of explicit disagreement. In particular, the use of the noun “whale” together with the adjective “beached” happens three times, and in all of them the disparaging intention is all but obvious:

- Somehow I find it hard to feel sorry for a *beached whale* posing in lingerie. If she didn’t know how dreadful she looked when it was taken, she is more stupid than she looks, and that is quite a bit. You can’t get that fat without eating until you are bloated time and time again. She is a disgusting piece of work. – Brian Williams, Dover
- Who was talking about rugby players? They are all tall and stocky, but they are not *beached whales* like this woman. No athlete or sportsperson is-unless they are sumo wrestlers! So dont twist peoples’ words, face facts. – Hannah, Bham UK
- Being that morbidly obese is not attractive in the slightest and it’s disgusting to see somebody flaunt her misused body. IT IS NOT OKAY to look like a *beached whale*. Size does matter when it comes down to health, and I have no doubt that her health is slowly deteriorating as we speak. – Appalled, London

The following comment is a reply by a woman to an earlier comment by a man. The use of the sardonic noun “billy-no-mates” is aggravated by

the adjectives 'sad' and 'lonely', and the noun 'loser' is made all the more aggressive by the use of block capitals:

- She should be grateful for any attention she gets. – Alex / Likewise Alex. Only attention you'll get are red arrows. But I guess it's better than nothing, you *sad lonely billy-no-mates. LOSER.* – Lou, London

The use of straight imperatives to give advice (which in itself is a threat to the hearer's negative face) aggravates the nature of the FTA, as happens in the following example, which exhibits some mitigating input at the beginning ("I don't think...") only to go bald-on-record in the second sentence:

- I don't think it's right that anyone should let themselves get in this state (medical conditions excluded) but her condition clearly starts at 6am when she opens the fridge door. *Gain* some willpower and self respect and *stop* complaining. – Mimi, Surrey

One of the strategies Brown and Levinson stipulate for protecting the hearer's negative face is "Don't presume /assume", which includes "avoiding presumptions about H, his wants, what is relevant, or interesting or worthy of his attention – that is, keeping ritual distance from H" (1987: 144). In the next passages, the speakers explicitly presume to have knowledge about the recipient's personality, likes, preferences and way of being. At the same time, they ostensibly violate one of Leech's (1983) politeness maxims, the approbation maxim (minimize dispraise of other):

- How can you have so little self respect to let yourself end up like that! – Rod Steele, Shaftesbury
- She looks like that because of gluttony, self-indulgence and a lazy refusal to exercise. If she says differently, she's lying. – Pete, Brighton

An interesting form which explicit disagreement assumes in the corpus is what Goodwin and Goodwin (1990: 97) call "content shift within argument". In the following passage, the speaker openly disagrees with the interpretation the readers have made of the article and performs a 'repair strategy' (Sacks et al. 1974) to restore what she believes is the right argumentative line:

- Missing the point people!!!!!! This is not about the size of the lady involved. It is about the use of her image to promote infidelity; more to the point a

website that helps you actively go out and look for someone to cheat with.
Fat or thin how happy would you be with that? – Polly, Germany

Much along the same lines, the comment shown next starts with a foregrounded expression of disagreement, in which the adverb of negation is signaled in block capitals, followed by a repair utterance in which the speaker gives her own reading of the subject in hand:

- Whether she’s fat or not is NOT in question here. It’s about the fact that sleazy Ashley Madison used her image WITHOUT her permission and did so in an embarrassing and derogatory way. – Jade, Los Angeles, USA

It should be noted that these cases of different topics being raised in the middle of an argument are a symptom of disaffiliation not only towards one party involved in the dispute, but also with its opponent party. Maynard (1986) calls this phenomenon “non-collaborative opposition” and explains it as follows:

In the first place, we have seen that disputes, although initially produced by two parties, do not consist simply of two sides. Rather, given one party’s displayed position, stance, or claim, another party can produce opposition by simply aligning against that position or by aligning with a counterposition. This means that parties can dispute a particular position for different reasons and by different means. It is therefore possible for several parties to serially oppose another’s claim without achieving collaboration. (Maynard 1986: 280)

One final aspect concerning the explicit expressions of disagreement in the corpus deserves mention. It is a fact that, as Locher has remarked (2004: 113), speakers often give personal or emotional reasons for disagreeing. And, by opening up about themselves and actually assuming a self-disclosure mode, they make their face more vulnerable. That is perhaps the reason for the defensive stance which is noticeable in utterances such as the following:

- I am a big and beautiful woman. I am comfortable in my skin and I refuse to be ashamed. If anyone is gross it is those that destroy others’ self-confidence with their hateful cruel words. There is beauty in ALL body types, skin color, and gender. If you can’t see that, something is wrong with you.
– Carol, Chicago

A comment by a male participant who signs “Slim Jim Loves Fat Pat”, and which reads “Most women over size 14 are clinically obese (do the math if you don’t believe me – BMI of 30 or more is obese)”, receives a set of emotional, personal responses which express utter disagreement:

- I think that statement is downright rude! I am a 19-year-old size 14-16 woman. I am by no means obese with a BMI of 21! Size 14 women can just be big boned or curvy women. I have a 26” waist and naturally large hips, does that make me fat, just because I’m labelled with a number 14?! And yes, you said MOST women, but I highly doubt I’m alone. – Elouise, Cambridge

The next female reader also replies to Slim Jim by giving personal, detailed information about her own body size and weight, and by insisting on the repair strategy of claiming that she is healthy rather than obese:

- Slim Jim – size 14 women are not obese, unless they are deluding themselves about their clothes size. I am a size 12-14 with a BMI of 24.9 which makes me just about within healthy range. Before I had children, I was a size 10 and my BMI was 20, which made me very slim and well within the healthy range. – Kath, Cardiff

Personal statements can take the opposite stance: instead of expressing disagreement towards intolerance to obesity, they express disagreement towards its tolerance:

- I was bitterly ashamed of myself when I let myself go to a size 18 and quickly snapped myself out of it back to a 10/12. Being life threateningly overweight is not attractive. This woman is just as bad of a role model as the dangerously underweight models, if not worse. – Rebecca, Gloucestershire

Likewise, the following respondent disagrees with all those defending fat people by setting her own example: that of a person who is “trying her hardest” to lose the excess weight:

- Since giving up smoking I have put on a bit of extra weight but am nowhere near this size.....And I am trying my hardest to get rid of the excess pounds. I know I’m overweight and feel uncomfortable but she must find problems in a lot of the ‘normal’ things she has to do everyday. – Sue, Oxford

Interestingly, the following commentator seems to be aware that personal, emotional comments regarding Jacqueline tend to come from people who have weight issues themselves. So, she anticipates – and corrects – the potential wrong inference as follows:

- Just listen to yourselves on here. Jacqueline makes a lot of very good points about how the female body is perceived and she is right. We have young girls who are absolutely obsessed with their shape and weight. *And before some smart alec [sic] says 'I bet you must be fat', no I'm not fat at all, far from it* – Female, UK

Finally, a very curious explicit disagreement strategy is the blunt expression of approval of Jacqueline's looks. In a forum that challenges obesity in all its aesthetic, psychological and health-related aspects, to say that she is, indeed, attractive is a major rebuttal of the opinions expressed. Next is a sample of utterances that assume an adversarial stance by actually agreeing, rather than disagreeing, with the main topic of the text – that of obesity:

- More to love I guess. – anita daeoph, usa
- Hmmmm, full complete bedfeast !!!!!!!!!!!!!!!!!!!!! yummy !!!!!!!!!!!!! – Tony, Essex
- Jackie certainly is juicy 8] – fezza, England
- Own goal surely, there's nothing scary about the gorgeous Jacqueline. – Mark, Huddersfield
- I like em rotund!! – Bejeeeeeeesus

Conclusion

This article has investigated the expression of disagreement in a corpus of online commentaries to a news report case. Several patterns have emerged in what has proven to be a good example of conflict talk, with alignments and confrontations taking place along the adversarial spectrum. However, the management of the oppositional input has also addressed face concerns, and rather often for that matter, though by no means always.

On an interpersonal level, the corpus has proven to belong to the category of "multi-party argument" (Maynard 1986, Goodwin and Goodwin 1993: 100) even though dyadic sequences were also identified. In other words, speakers interacted either by replying directly to another speaker's utterance or by posting unaddressed comments.

In terms of content, the complexity of the analysis has involved a few concurring factors. First, the construction of the argument revolves around more than one topic. The initial cue, offered by the newspaper itself, is that “big is beautiful”, but the range of disagreement expands to other topics, such as the health-problems regarding obesity, the objectification of women, man-woman relationships, adultery, pornography and hypocrisy. This I have proposed to call “multi-topic argument”. Its complexity derives from the fact that an initial topic prompts the appearance of several topic strands in an equally complex sequence: being raised, debated, dropped and recovered along the argumentative axis. Another analytical challenge has been the fact that the speech act of disagreeing may be mixed with other forms of illocutionary force. Indeed, the disagreements occurring in the corpus often overlap with other speech acts, such as criticising, protesting, reprimanding, etc, making the illocutionary boundaries somewhat vague. Thirdly, the variety of topics has shown to be related to a broad range of “reasons” for agreeing or disagreeing. From ethical and moral reasons to health reasons, social reasons and aesthetic reasons, the opinions expressed reveal a great assortment of motivations for stepping forward and expressing a contrary opinion. Crucially, personal reasons rank high, as is the case of fat people strategically siding with the obese model and confronting her critics (on disagreements as strategic moves, see Upadhyay 2010, above).

Interestingly enough, the representations of femaleness that emerge throughout this complex argumentative process seem to reify, by and large, existing stereotypes of physical perfection and desirability. In fact, although a few voices do get to be raised in favour of alternative standards of female beauty and attractiveness, the overwhelming bulk of comments prove to support the widespread ideals of thinness and slenderness. Besides this reinforcement of prevailing conceptions of women’s body image, the texts also denote the surfacing of another sort of stereotype, that of the equation of the female subject with a sexual object. Indeed, Jacqueline’s side professional activities are more often than not viewed as conflicting elements that undermine her cry for justice and reduce the legitimacy of her claims. Had she been a married mother of four, perhaps the opposition to her complaints would have been less noticeable.

The diversity of ideological, psychological or personal motivations detected has revealed two types of interpersonal stance: a defensive stance (typically exemplified by obese readers) and an aggressive one. Whatever the case, the analysis has confirmed Angouri and Locher’s (2012: 1549) claim that disagreeing does have “an impact on relational issues (face-aggra-

vating, face-maintaining, face-enhancing)". The situations in which replies – including insulting ones – are directed at specific interactants illustrate an aggressive attempt at debasing the opponent's face. Other comments, meanwhile, have revealed face concerns of various strengths.

The analysis has shown that the expression of disagreement in the corpus ranges from "backgrounded" forms, that is, covered, implicit, or mild disagreement, to "foregrounded" forms, i.e., overt, explicit, or unmitigated disagreement (Scott 2002), admitting however in-between manifestations, hedged with positive and negative politeness (Walkinshaw 2011). Occurrences of overt impoliteness, with expression of bald-on-record rebuttals, provocations and even insults may be related to the anonymity of its participants and the discontinuous character of the interaction, both in spatial and temporal terms, which encourage the abandonment of face concerns (Donath 1999, Eisenclas 2011, Yus 2011). Knowing that no retaliation is possible, speakers feel free – and powerful enough (Brown and Levinson 1987: 97) – to attack their opponents' face.

Finally, the analysis has suggested another – crucial – explanation for the impoliteness occurring in the corpus: namely, what I propose to call, along Leech's lines (1983: 133), "third-party factor". Disagreeing with, criticising, debasing and deriding an absent party is much less risky than doing so in face-to-face conversation. The fact that Jacqueline is not taking part in the polylogue makes her face more negligible, and it strengthens the respondents' boldness to show their disaffiliation. This may explain the openly conflictual and confrontational nature of some of the comments, which the protocol ruling over regular daily interaction usually tends to soften.

Bibliography:

- ATCHISON, J. and Lewis, D. M. (Eds.) (2003), *New Media Language*. London: Routledge.
- ANGOURI, Jo (2012), Managing disagreement in problem solving meeting talk. *Journal of Pragmatics*. Volume 44, Issue 12, September 2012, 1565–1579.
- ANGOURI, Jo and Locher, Miriam A. (2012), "Theorising Disagreement". *Journal of Pragmatics*. Volume 44, Issue 12, September 2012, 1549-1720.
- ATKINSON, Maxwell & Heritage, John (Eds.) (1984), *Structures of Social Action: Studies in Conversation Analysis*. Cambridge: C.U.P.
- AUSTIN, J. L. (1962), *How to Do Things with Words*. Oxford: O.U.P.

- BARGIELA-CHIAPPINI, Francesca (2003), "Face and politeness: new (insights) for old (concepts)". *Journal of Pragmatics* 35: 1453-1469.
- BARON, Naomi (1998), "Letters by phone or speech by other means: the linguistics of email". *Language and Communication* 18: 133-170.
- BARON, Naomi (2003), "Why email looks like speech". In: Aitchison & Lewis (Eds.), pp. 85-94.
- BLUM-KULKA, Shoshana (1990), "You don't touch lettuce with your fingers: Parental politeness in family discourse". *Journal of Pragmatics* 14 (2): 259-288.
- BOLANDER, Brook (2012), "Disagreements and agreements in personal/diary blogs: A closer look at responsiveness." *Journal of Pragmatics*. Volume 44, Issue 12, September 2012, 1607-1622.
- BROWN, Penelope & Levinson, Stephen (1987), *Politeness: Some Universals of Language Usage*. Cambridge: C.U.P.
- CASTELLS, M. (2000), *The Rise of the Network Society*. Oxford: Blackwell. 2nd Ed.
- CHERRY, Roger D. (1988), "Politeness in written persuasion". *Journal of Pragmatics* 12: 63-81.
- CHILTON, Paul (1990), "Politeness, politics and diplomacy". *Discourse and Society*, Vol.1(2): 201-224.
- CLARKE, Matthew (2009), "The discursive construction of interpersonal relations in an online community of practice". *Journal of Pragmatics* 41: 2333-2344.
- COLLOT, M. and Belmore, N. (1996), "Electronic language: a new variety of English". In: S. Herring (Ed.): 13-28.
- CRYSTAL, David (2001), *Language and the Internet*. Cambridge: C.U.P.
- CULPEPER, Jonathan (1996), "Towards an anatomy of impoliteness". *Journal of Pragmatics* 25: 349-367.
- CULPEPER, Jonathan (2011), *Impoliteness: Using Language to Cause Offence*. Cambridge: C.U.P.
- DONATH, J. (1999), "Identity and deception in the virtual community". In: Kollock and Smith (Eds.), pp. 31-59.
- EARLEY, Christopher (1997), *Face, Harmony and Social Structure: An Analysis of Organizational Behavior across Cultures*. New York: Oxford University Press.
- EISENCHLAS, Susana A. (2011), "On-line interactions as a resource to raise pragmatic awareness", *Journal of Pragmatics* Volume 43, Issue 1, January 2011, Pages 51-61.
- FERRARA, K., H. Brunner, G. Whitmore (1991), "Interactive written discourse as an emergent register". *Written Communication* 8 (1): 8-34.
- GOFFMAN, Erving (1967), *Interaction ritual: Essays on face-to-face behavior* (Penguin, Harmondsworth). Partially reissued in: A. Jaworski & N. Coupland (Eds.) (1999), *The Discourse Reader*. London: Routledge, pp.306-320.

- GOODWIN, Charles and Goodwin, Marjorie (1990), "Interstitial argument". In: Allen Grimshaw (ed.), *Conflict Talk*. Cambridge: CUP. 85-117.
- HARRIS, Sandra (2003), "Politeness and power: Making and responding to 'requests' in institutional settings". *TEXT* 23:1: 27-52.
- HERRING, S. (Ed.) (1996), *Computer Mediated Communication: Linguistic, Social, and Cross-Cultural Perspectives*. Philadelphia: John Benjamins.
- HILL, Beverly, Sachiko Ide, Shoko Ikuta, Akiko Kawasaki and Tsunao Ogino (1986), "Universals of linguistic politeness: Quantitative evidence from Japanese and American English". *Journal of Pragmatics* 10-3: 347-371.
- IDE, Sachiko (1989), "Formal forms and discernment: Two neglected aspects of linguistic politeness". *Multilingua* 8: 223-248.
- KAKAVÁ, Christina (1993), *Negotiation of Disagreement by Greeks in Conversations and Classroom Discourse*. Washington: Georgetown University.
- KOLLOCK, P. and Smith, M. A. (Eds.), *Communities in Cyberspace*. London: Routledge.
- KOTTHOFF, Helga (1993), "Disagreement and Concession in Disputes: On the Context Sensitivity of Preference Structures". *Language in Society*. Vol. 22, No. 2 (Jun., 1993), pp. 193-216.
- LANGLOTZ, Andreas, Locher, Miriam A. (2012), "Ways of communicating emotional stance in online disagreements". *Journal of Pragmatics*. Volume 44, Issue 12, September 2012, 1591-1606.
- LEECH, Geoffrey (1983), *Principles of Pragmatics*. London: Longman, 1999.
- LIM, Tae-Sep (1994), "Facework and Interpersonal Relationships". In: S. Ting-Toomey (Ed.): 89-108.
- LITTLE, G. D. and Montgomery, M. (Eds.) (1994), *Centennial Usage Studies*. Tuscaloosa: University of Alabama Press.
- MAO, LuMing Robert (1994), "Beyond politeness theory: 'Face' revisited and renewed". *Journal of Pragmatics* 21: 451-486.
- MARCOCCIA, Michel (2004), "On-line polylogues: conversation structure and participation framework in Internet newsgroups". *Journal of Pragmatics* 36: 115-145.
- MARRA, Meredith (2012), "Disagreeing without being disagreeable: Negotiating workplace communities as an outsider". *Journal of Pragmatics*. Volume 44, Issue 12, September 2012, 1580-1590.
- MAYNARD, Douglas W. (1986), "Offering and soliciting collaboration in multi-party disputes among children (and other humans)". *Human Studies*. 1986, Volume 9, Issue 2-3, pp 261-285.
- MAYNOR, N. (1994), "The language of electronic email: written speech". In: G. D. Little and M. Montgomery (Eds.), pp. 48-54.
- MONTERO-FLETA, Begoña, Anna Montesinos-López, Carmen Pérez-Sabater, Ed Turney (2009), "Computer mediated communication and informalization of discourse: The influence of culture and subject matter". *Journal of Pragmatics* 41, 770-779.

- PEREZ DE AYALA, Soledad (2001), "FTAs and Erskine May: Conflicting needs? – Politeness in Question Time". *Journal of Pragmatics* 33: 143-169.
- POMERANTZ, Anita (1984), "Agreeing and disagreeing with assessments: some features of preferred/dispreferred turn shapes". In: J. M. Atkinson & J. Heritage (Eds.), *Structures of Social Action*. Cambridge: Cambridge University Press, 57-101.
- SACKS, Harvey, E. Schegloff, G. Jefferson (1974), "A Simplest Systematic for the Organization of Turn-taking for Conversation". In: Asa Kasher (ed.) (1998), *Pragmatics: Basic Concepts*. London: Routledge. vol. V.
- SAYGIN, Ayse Pinar and Cicekli, Ilyas (2002), "Pragmatics in human-computer conversations". *Journal of Pragmatics* Volume 34, Issue 3, March 2002, 227–258.
- SAVAS, Perihan (2011), "A case study of contextual and individual factors that shape linguistic variation in synchronous text-based computer-mediated communication". *Journal of Pragmatics*. Volume 43, Issue 1, January 2011, 298–313.
- SCHIFFRIN, Deborah (1984). Jewish argument as sociability. *Language in Society* 13: 311–335.
- SCOTT, Suzanne (2002), "Linguistic feature variation within disagreements: An empirical investigation". *Text* 22(2) (2002), pp. 301–328
- SEARLE, John (1969), *Speech Acts. An Essay in the Philosophy of Language*. Cambridge: C.U.P.
- SLUGOSKI, B.R. (1985), *Grice's theory of conversation as a social psychological model*. PhD. Oxford.
- TING-TOOMEY, S. (Ed.) (1994), *The Challenge of Facework: Cross-Cultural and Interpersonal Issues*. Albany: State University of New York Press.
- UPADHYAY, Shiv R. (2010), "Identity and impoliteness in computer-mediated reader responses." *Journal of Politeness Research*. Volume 6, Issue 1, Pages 105–127.
- VAN DIJK, Teun (Ed.) (1985), *Handbook of Discourse Analysis*. Volume 3: *Discourse and Dialogue*. London: Academic Press.
- WALDRON, Vincent R., Applegate, James L. (1994), "Interpersonal Construct Differentiation and Conversational Planning: An Examination of Two Cognitive Accounts for the Production of Competent Verbal Disagreement Tactics." *Human Communication Research*, v21, n1 pp. 3-35.
- WALKINSHAW, Ian (2009) *Learning Politeness: Disagreement in a Second Language*. Bern: Peter Lang.
- YATES, S. (1996), "Oral and written linguistic aspects of computer conferencing." In: S. Herring (Ed.), pp. 29-46.
- YUS, Francisco (2011), *Cyberpragmatics: Internet-Mediated Communication in Context*. Amsterdam: John Benjamins.

NOTA DE MORFOLOGIA HISTÓRICA DO PORTUGUÊS: SUFIXO *-Ó†

NOTES ON PORTUGUESE HISTORICAL MORPHOLOGY: SUFFIX *-Ó

Przemysław Dębowiak[‡]
pdebowskiak@gmail.com

Existe em português um grupo de palavras oxítonas terminadas em -ó ([ɔ] tónico) que não são nem empréstimos, nem interjeições, nem onomatopeias. É uma trintena de substantivos de ambos os géneros gramaticais, alguns empregados na linguagem do dia-a-dia (*avó*), outros mais raros, especializados ou dialetais (*eiró* ‘enguia’ ou ‘pequena eira’), outros ainda já caídos em desuso (*lançó*). Segundo as fontes etimológicas, estes vocábulos foram herdados do latim, provindo de diminutivos formados com os sufixos *-ĕdulus / -a / -um* e *-īdulus / -a / -um*; nalguns casos, a origem deles não foi estudada. O objetivo do trabalho é demonstrar, com base numa análise formal e semântica destes termos, a existência de um sufixo *-ó, ainda produtivo no português medieval, que terá tido, entre outros, valor diminutivo. Posteriormente, o sufixo postulado terá perdido vitalidade, estando hoje em dia completamente lexicalizado num número reduzido de palavras como as do corpus.

Palavras-chave: português, morfologia histórica, derivação sufixal, diminutivo.

In Portuguese, there is a group of oxytone words ending in -ó (stressed [ɔ]) that are neither loans, nor interjections, nor onomatopoeias. These are around thirty nouns

[†] Uma versão menos extensa e ligeiramente modificada deste texto foi apresentada, sob o título *Sobre algumas palavras portuguesas terminadas em “-ó”*, no Colóquio Internacional *Horizontes do Saber Filológico* que teve lugar de 16 a 17 de novembro 2012 na Universidade Sveti Kliment Ohridski em Sófia (Bulgária).

[‡] Universidade Jagellónica de Cracóvia.

of both grammatical genders, some of them used in everyday language (*avó* ‘grandmother’), others rarer, specialised and dialectal (*eiró* ‘eel’ or ‘small treshing-floor’), and others which have already fallen into disuse (*lançó* ‘lancette’). According to etymological sources, these terms were inherited from Latin and come from diminutives formed with suffixes *-ĕolus / -a / -um* and *-ĭolus / -a / -um*; in some cases, their origin was not studied. Based on a formal and semantic analysis of these words, the paper aims to demonstrate the existence of a suffix **-ó*, still productive in Medieval Portuguese, that might have had a diminutive meaning, among others. The postulated suffix might have subsequently lost its vitality and is nowadays completely lexicalised in a reduced number of terms like the ones from the corpus.

Keywords: Portuguese, historical morphology, suffixal derivation, diminutive.

*

Introdução

1. O português é tradicionalmente designado como uma língua de acento livre, o que significa que o acento não é estável a nível de palavras (Cunha & Cintra, 1984: 55–62; Mateus, Frota & Vigário, 2003: 1050–1054; Mateus, Falé & Freitas, 2005: 279–287; Borregana, 2008: 36–38). Quando uma palavra tem acento na última sílaba, é oxítone ou aguda (p.ex. *além*, *partir*); quando na penúltima, é paroxítone ou grave (p.ex. *barco*, *texugo*); se for acentuada na antepenúltima sílaba, então é chamada proparoxítone ou esdrúxula (p.ex. *próximo*, *escândalo*). Se se incluírem também os enclíticos, às vezes é marcada com acento a quarta sílaba, contando do fim do vocábulo, dando-se-lhe então o nome de bisesdrúxulo (p.ex. *lavávamos-nos*, *chamava-se-lhe*). Estatisticamente, as palavras portuguesas são na sua maioria paroxítonas (Espada, 2006: 97), o que também justifica o uso dos diacríticos na ortografia portuguesa: acentuam-se (portanto, distinguem-se) na grafia precisamente aqueles vocábulos que não são paroxítonos.

2. No léxico português existem várias palavras oxítonas terminadas em *-ó* ([ɔ] tónico). Não é uma terminação muito frequente; num dicionário de rimas da língua portuguesa (Castilho, 1894: 239) só encontramos 70 entradas com tal terminação, e num dicionário inverso (Andrade, 1993: 620–622) – 146. Porém, tendo completado o corpus com vocábulos acha-

dos noutros dicionários (DA, DLP, DPLP), constatamos que apresentam uma grande diversidade quanto à origem⁽¹⁾. Nomeadamente:

- a) palavras monossilábicas de proveniência latina que resultam de uma síncope de consoantes intervocálicas e uma crase posterior de vogais tornadas vizinhas:
- *dó* (s.m.) ‘compaixão, pena’ ← *dōlu-* (s.m.);
 - *jó* (s.m.) ‘travessa que limita avante e à ré os bancos dos remadores’ ← *jūgu-* (s.n.);
 - *mó* (s.f.) 1. ‘pedra pesada e redonda para moinho ou lagar’, 2. ‘pedra para amolar instrumentos cortantes ou perfurantes’ ← *mōla-* (s.f.);
 - *nó* (s.m.) ‘laço apertado’ ← *nōdu-* (s.m.);
 - *pó* (s.m.) ‘poeira’ ← **pūlu-* (← **pūlvu-* ← *pūlvīs* (s.m.));
 - *só* (adj.m./f.) ‘sozinho, sozinha’ ← *sōlu-*, *sōla-* (adj.);
- b) empréstimos, p.ex.:
- *faráó* (s.m.) ‘soberano do antigo Egito’ ← lat. *Phārāo* (-ōn);
 - *icipó* (s.m.) ‘arbusto brasileiro’ ← tupi;
 - *portaló* (s.m.) ‘lugar por onde se entra no navio ou por onde se mete a carga’ ← cat. *portaló*;
 - *trenó* (s.m.) ‘carruagem sem rodas, própria para deslizar sobre o gelo’ ← fr. *traîneau*;
- c) interjeições, p.ex.:
- *bó!* ‘ora essa! nessa!’;
 - *fó!* que exprime repugnância (no português informal da Madeira);
 - *tó!* que serve para chamar ou afastar certos animais;
- d) onomatopeias, p.ex.:
- *cocó* (s.m.) ‘galo ou galinha’ (termo informal);
 - *cocorocó* (s.m.) ‘voz imitativa do canto do galo’;
 - *fofó* (s.m.) ‘ventosidade expelida pelo ânus; flato, peido’ (no português informal da Madeira);
- e) termos informais provenientes da linguagem afetiva, p.ex.:
- *cocó* (s.m.) ‘excremento’;
 - *Tó* (s.m.) ‘António’;

(1) As etimologias consultaram-se com dicionários etimológicos (DELP e REW); todas as palavras latinas foram verificadas no dicionário do latim de Gaffiot (G); as formas contemporâneas dos vocábulos portugueses e os seus significados provêm de dicionários da língua portuguesa (DA, DLP e DPLP).

No texto usam-se as seguintes abreviações e símbolos: **adj.** = adjetivo; **adv.** = advérbio; **ant.** = antigo; **cat.** = catalão; **dial.** = dialetal; **esp.** = espanhol; **f.** = feminino; **fr.** = francês; **gal.** = galego; **lat.** = latim; **leon.** = leonês; **m.** = masculino; **n.** = neutro; **PB** = português do Brasil; **PE** = português europeu; **pl.** = plural; **pt.** = português; **s.** = substantivo; **sg.** = singular; < = provém de; ⇒ = resulta em; < = deriva de; > = é a base de derivação para; <> = relacionado com.

- f) palavras resultantes da apócope, p.ex.:
- *cafundó* (s.m.) ← *cafundório* (s.m.) ‘lugar ermo’ (no PB);
 - *freijó* (s.m.) ← *frei-jorge* (s.m.) ‘quiri; árvore leguminosa do Brasil’ (no PB);
- g) vocábulos cuja etimologia não está suficientemente, ou de maneira alguma, esclarecida, p.ex.:
- *badalhó* 1. (s.m./f.) ‘casta de figueira’, 2. (adj.m./f.) ‘diz-se de casta de figueira’;
 - *beilhó* (s.m./f.) ‘bolo frito de abóbora, farinha e açúcar’ (também *belhó* (s.m.), *beilhós* (s.m./f.), *bilhós* (s.f.));
 - dial. *bilhó* (s.f.) 1. (ant.) ‘castanha assada ou cozida e descascada’, 2. ‘criança gorda e muito baixa’ (Beira, Trás-os-Montes)⁽²⁾;
 - *filó* (s.m.) ‘tecido reticular; espécie de cassa’⁽³⁾;
 - *ló* (s.m.) ‘tecido fino como escumilha’ (também em *pão-de-ló* ‘variedade de bolo fofo e doce?’);
 - *queiró* (s.f.) ‘urze do mato’ (também *queirós*, *queiroga*, ambos s.f.)⁽⁴⁾;
 - *quitó* (s.m.) 1. (ant.) ‘espécie de espada’, 2. (dial.) ‘pessoa muito pequena’.

Apresentação e análise do corpus

3. A este grupo de palavras portuguesas oxítonas terminadas em *-ó* pertence igualmente um conjunto de vocábulos que têm uma característica em comum, sobressaindo da sua etimologia: provêm de diminutivos latinos proparoxítonos formados com o sufixo *-ŭlus / -a / -um*, que antecedido das vogais anteriores [e], [i] ou da semivogal [w], adotava a forma *-ĕŏlus / -a / -um* ou *-ĩŏlus / -a / -um*⁽⁵⁾. Segue-se uma análise histórica formal e semântica destes termos, constituindo o núcleo do nosso corpus, junto com

(2) Na literatura etimológica fala-se das palavras *beilhó* e *bilhó* como se proviessem do mesmo étimo (o que não é óbvio), mas, na nossa opinião, não se conseguiu chegar a nenhuma solução satisfatória. Veja-se: Vasconcellos, 1895: 133; Viana, 1906a: 139–140; Vasconcellos, 1927: 126; DELP s.v. *Belhó*, *bilhó*.

(3) DELP (s.v. *Fio*) propõe um étimo “**filolu-*” (por **filĩŏlu-* ‘pequeno fio’), o que se explicaria semanticamente (de ‘matéria’ a ‘tecido feito dela’, veja-se 3.11.), mas é difícil de sustentar do ponto de vista fonético, dada a presença do *-l-* intervocálico.

(4) Veja-se Vasconcellos, 1895: 181–182.

(5) Em função do contexto fonético, este sufixo podia ter também outras formas, entre as quais dominava *-ĕllus / -a / -um*. Outro sufixo diminutivo era bastante frequente em latim: *-cŭlus / -a / -um*, com a sua variedade posterior *-cĕllus / -a / -um*. Vejam-se gramáticas latinas, p.ex. Meillet & Vendryes, 1924: 117; Otrębski & Safarewicz, 1937: 244–247; Palmer, 1988: 236–237.

a dos seus predecessores não-diminutivos (caso se tenham conservado em português).

3.1. pt. *arrió* (s.m.) 1. ‘pedrinha para jogar o alquerque’, 2. (mais usado no pl.) ‘jogo do alquerque’, 3. ‘pelouro’, 4. (no PB) ‘espécie de fava’ ← lat. **rādiōla*- (s.f.) ‘raiozinho’ < **rādīa*- (s.f.) ‘raio, linha’ → pt. *raia* (s.f.) ‘linha, traço, risca’.

As variantes são: *arriol*, *arriós* (ambas s.m.). Todas as formas apresentam aglutinação do artigo feminino (**a rió* → *arrió*). O diminutivo desenvolveu aceções figuradas: ‘risca’ → ‘jogo de pedrinhas sobre uma tábua riscada ou um solo com riscas traçadas’ → ‘pedrinha’ → ‘espécie de fava’.

Forma sinónima da mesma origem: esp. *rayuela* (s.f.) ‘espécie de jogo’; também gal. *riola* (s.f.) ‘fila de pessoas’, *raiola* (s.f.) 1. ‘conjunto de raios de sol que às vezes aparecem entre as nuvens quando o céu está coberto’, 2. ‘aberta de sol num dia chuvoso’, 3. ‘calor morno do sol’.

[Viana, 1906a: 92–94; DELP s.v. *Arrió*, *arriól*, *arriós*, s.v. *Raia*²; DA s.v. *Arriós*, s.v. *Raia*¹; DLP s.v. *arrió*, s.v. *raia*¹; DPLP s.v. *arrió*, s.v. *raia*; DRAG s.v. *raiola*, s.v. *riola*; RAE s.v. *rayuela*].

3.2. pt. *avó* (s.f.) ‘mãe do pai ou da mãe’ ← lat. **āvīōla*- (s.f.) ‘avó’ < *āvīa*- (s.f.) ‘avó’ → pt. \emptyset .

A forma *āvīa* é o feminino de *āvus* (s.m.) ‘antepassado, avó’ que também não se conservou em português, mas cujo diminutivo **āvīōlu*- (s.m.) deu pt. *avó* (s.m.) ‘pai do pai ou da mãe’. Em ambos os casos o diminutivo, decerto de origem afetiva, substituiu completamente o primitivo.

Formas sinónimas da mesma origem: gal. *avoa* (s.f.), *avó* (s.m.); esp. *abuela* (s.f.), *abuelo* (s.m.); também fr. *aïeule* (s.f.), *aïeul* (s.m.) ‘antepassado, ascendente’.

[G s.v. 1 *āvīa*, s.v. *āvus* (*āvos*)⁽⁶⁾; REW 823, 830, 839; DELP s.v. *Avó*; DA s.v. *Avó*, s.v. *Avô*; DLP s.v. *avó*, s.v. *avô*; DPLP s.v. *avó*, s.v. *avô*; DRAG s.v. *avó*; RAE s.v. *abuela*, s.v. *abuelo*; TLF s.v. *aïeul*, *eule*, *euls*, *eules*].

3.3. pt. *corrijó* (s.f.) ‘planta herbácea espontânea em Portugal, também conhecida por carrajó, língua-de-ovelha, tanchagem, etc.’ ← lat. **corrīgīōla*- (s.f.) ‘pequena correia’ < *corrīgīa*- ‘correia; chicote; laço de calçado’ → pt. *correia* ‘tira de couro’.

(6) Ainda que não note **āvīōla*, diminutivo de *āvīa*, encontra-se nele a forma *Aviōla* com o sentido de ‘sobrenome romano’.

Existem também as formas *correjola*, *corrijola*, *corrigiôla*, *corriola* (todas s.f.) que designam outras espécies de plantas, mas provêm do mesmo diminutivo latino. O derivado desenvolveu um sentido figurado.

Formas sinónimas da mesma origem: gal. *correola* (s.f.); esp. *correhuela*, *correyuela* (s.f.); cat. *corretjola* (s.f.).

[G s.v. *corrígia*; REW 2253; DELP s.v. *Correia*; DA s.v. *Correia*, s.v. *Corrijola*, s.v. *Corrigiôla*, s.v. *Corriola*; DLP s.v. *carrijó*, s.v. *correia*, s.v. *correjola*, s.v. *corrigiôla*, s.v. *corrijó*, s.v. *corrijola*, s.v. *corriola*; DPLP s.v. *correia*, s.v. *carrijó*, s.v. *corriola*; DRAG s.v. *correola*; DRAE s.v. *correhuela*, *correyuela*; GDLC s.v. *corretjola*].

3.4. pt. *eiró* (s.m./f.) ‘enguia’ ← lat. **hydrēōla*- (s.f.) ‘pequena hidra’ < *hydra*- (s.f.) ‘cobra de água; hidra de Lerna’ → pt. ∅.

Outras formas do diminutivo conservado são: *iró*, *eiról*, *iról*, *eirós*, *irós* – todas de género feminino; *eró*, *eirogo* – de género masculino⁽⁷⁾. O primitivo não sobreviveu em português, enquanto o derivado se conservou com um sentido figurado, um tanto minorativo.

Formas sinónimas da mesma origem: gal. *airoa* / *eiroa* (s.f.).

[G s.v. *hydra*; Vasconcellos, 1895: 144–145⁽⁸⁾; Viana, 1906a: 377⁽⁹⁾; REW 633⁽¹⁰⁾; DELP s.v. *Eiró*²; DA s.v. *Eiró*¹; DLP s.v. *eiró*¹, s.v. *iró*, s.v. *irós*; DPLP s.v. *eiró*; DRAG s.v. *airoa*, s.v. *eiroa*].

3.5. dial. pt. *eiró* (s.m./f.) ‘eira de piso térreo’ (Trás-os-Montes) ← lat. *ārēōla*- (s.f.) ‘pátio pequeno’ < *ārēa*- (s.f.) 1. ‘superfície, solo unido; pátio’, 2. ‘espaço para bater o trigo’, etc. → pt. *eira* (s.f.) ‘terreno liso ou empedrado onde se põem a secar e se trilhavam ou desgranavam legumes ou cereais’.

Observa-se uma restrição territorial e semântica do diminutivo face ao primitivo.

(7) Os nomes parecidos: *eiroga*, *airoga*, *oirega*, *teiroga*, todos s.f., designam outro peixe, nomeadamente uma espécie de raia (DLP s.v. *eiroga*, s.v. *teiroga*²). A nosso ver, resultam da confusão de *eiró* com *teira* (veja-se 4.15.).

(8) Propõe dois étimos, ambos errados, sem prover explicações suplementares, mas admitindo ter dúvidas a respeito deles: lat. “*areolus* / *areola*” ou “*hariolus* / *hariola*”.

(9) Com a seguinte explicação da primeira das etimologias sugeridas por Vasconcellos (1895: 144–145): “De *areola* < *areia*, por serem as *eirós* transportadas vivas nas selhas, envolvidas em areia molhada”.

(10) Repete a primeira das etimologias propostas por Vasconcellos (1895: 144–145), sustentada por Viana (1906a: 377).

Forma correspondente da mesma origem: cat. **erola** (s.f.) 1. ‘terreno onde se semeia legumes’, 2. (ant.) ‘pequena superfície plana em cima de uma montanha’.

[G s.v. *ârëa*, s.v. *ârëöla*; Vasconcellos, 1895: 145; REW 626, 632⁽¹¹⁾; DELP s.v. *Área*, s.v. *Eiró*¹; DA s.v. *Eira*, s.v. *Eiró*²; DLP s.v. *eira*, s.v. *eiró*²; DPLP s.v. *eira*, s.v. *eiró*; GDLC s.v. *erola*].

3.6. pt. **enxó** (s.f. no PE, s.m. no PB) ‘instrumento de carpinteiro para desbastar tábuas ou peças de madeira’ ← lat. *asciöla-* (s.f.) ‘pequena enxó’ < *ascia-* (s.f.) 1. ‘machado, enxó’, 2. ‘enxada’, 3. ‘colher de pedreiro’, 4. ‘martelo do britador de pedra’ → pt. \emptyset .

O diminutivo retomou o sentido do primitivo que não se conservou em português.

Formas sinónimas da mesma origem: gal. **aixola** (s.f.); esp. **azuela** (s.f.); cat. **aixol** (s.m.).

[G s.v. *ascia* (*ascëa*), s.v. *asciöla*; Vasconcellos, 1887–1889: 304⁽¹²⁾; Viana, 1906a: 396; REW 696, 698⁽¹³⁾; DELP s.v. *Enxada*; DA s.v. *Enxó*¹; DLP s.v. *enxó*; DPLP s.v. *enxó*; DRAG s.v. *aixola*; RAE s.v. *azuela*; GDLC s.v. *aixol*].

3.7. pt. **filhó** (s.m./f.) 1. ‘bolo de farinha e ovos, frito em óleo ou azeite, e normalmente passado por calda de açúcar’, 2. ‘tira fina de massa de farinha e ovos, que, depois de frita, se polvilha com açúcar e / ou canela’ ← lat. **föliöla-* (s.f.) ‘pequena folha; bolo folhado’ < *fölia-* (s.f.) ‘folha, folhagem’ (pl. de *fölium* (s.n.)) ‘folha, cada uma das partes que constituem a verdura dos vegetais’, tornado sg.f.) → pt. *folha* (s.f.) ‘órgão das plantas responsável pela captação de luz e trocas gasosas com a atmosfera’.

Outra forma é **filhós** (s.f.). A diferenciação semântica entre o primitivo e o diminutivo é evidente, este último tendo desenvolvido um sentido figurado pela comparação da estrutura do bolo folhado com uma folha.

Formas sinónimas da mesma origem: gal. **filloa** (s.f.); leon. **fiyuela** (s.f.); esp. **hojuela** (s.f.).

(11) Assinala a existência do lat. *ârëöla*, mas não nota a palavra portuguesa entre os seus continuadores românicos.

(12) Grafia a palavra como *enchó*.

(13) Embora note a existência do lat. *asciöla*, não enumera a palavra portuguesa entre os seus continuadores românicos.

[G s.v. *1 fōlia*, s.v. *fōliolum*, s.v. *fōlium*; Vasconcellos, 1887–1889: 304; Vasconcellos, 1895: 133⁽¹⁴⁾; Viana, 1906a: 461⁽¹⁵⁾; REW 3413, 3415; DELP s.v. *Filhó*, s.v. *Folha*; DA s.v. *Filhó*, s.v. *Folha*; DLP s.v. *filhó*, s.v. *folha*; DPLP s.v. *filhó*, s.v. *filhós*, s.v. *folha*; DRAG s.v. *filloa*; RAE s.v. *hojuela*].

3.8. dial. pt. **grejó** 1. (s.f.) ‘pequena igreja, capela’, 2. (s.m., ant.) ‘lâmpada eterna da igreja’ (Trás-os-Montes) ← lat. *ecclēsīōla*- (s.f.) ‘pequena igreja’ < *ecclēsīa*- (s.f.) ‘assembleia; conjunto (dos fiéis); igreja’ → pt. *igreja* (s.f.) 1. ‘conjunto dos fiéis’, 2. ‘edifício dedicado ao culto da religião cristã’.

Existem variantes **grijó**, **igrejó**, **igrejola** (s.m./f.). O diminutivo guardou o seu valor original, embora só se tenha conservado nos dialetos portugueses setentrionais.

Forma sinónima da mesma origem: cat. **esglesiola** (s.f.).

[G s.v. *ecclēsīa*, s.v. *ecclēsīōla*; Vasconcellos, 1895: 168; Viana, 1906a: 519; Viana, 1906b: 4–5; REW 2823; DELP s.v. *Igreja*; DA s.v. *Grejó*, s.v. *Grijó*, s.v. *Igreja*, s.v. *Igrejó*, s.v. *Igrejola*; DLP s.v. *grejó*, s.v. *igreja*; DPLP s.v. *grejó*, s.v. *grijó*, s.v. *igreja*; GDLC s.v. *esglesiola*].

3.9. pt. **ichó** (s.m./f.) ‘armadilha para coelhos ou perdizes’ ← lat. *ōstīōlu*- (s.n.) ‘pequena porta’ < *ōstīu*- (s.n.) ‘entrada, porta’ → pt. *o*.

Existem outras variantes do diminutivo conservado, todas oscilantes em género: **ichós**, **chó** (devida a uma aférese) e **enchó** (resultado da atração formal a *enxó*, veja-se 3.6.).

Formas sinónimas da mesma origem: gal. **ichó** (s.m.); esp. **orzuelo** (s.m.).

[G s.v. *ostīolum*, s.v. *1 ostīum*; Vasconcellos, 1887–1889: 304; Viana, 1906a: 396⁽¹⁶⁾; Viana, 1906b: 1; REW 6116, 6117; DELP s.v. *Ichó*; DA s.v. *Enxó*², s.v. *Enxós*, s.v. *Ichó*; DLP s.v. *enchó*, s.v. *ichó*; DPLP s.v. *enchó*, s.v. *ichó*; DRAG s.v. *ichó*; RAE s.v. *orzuelo*²].

3.10. pt. ant. **lançó** (s.m./f.) ‘lanceta aguda’ ← lat. *lancēōla*- (*lancōla*-) (s.f.) ‘pequena lança’ < *lancēa*- (s.f.) ‘lança, chuço, hasta’ → pt. *lança* (s.f.) ‘chuço, hasta, pique’.

O diminutivo guarda o seu sentido original.

(14) Inicialmente (1887–1889) propõe um étimo errado sem dar explicações (lat. “*filiolum*”), depois (1895) hesita entre lat. “*filiolum*” e “*foliola*”, mas não decide qual dos dois é mais plausível.

(15) Admite a palavra vir tanto do lat. “*folliola* (< *follis* ‘fole’)” como “*foliola* (< *folium* ‘folha’)”.

(16) Não reconhece na forma que grafa *enxó(s)* uma variante da palavra *ichó*, acreditando que se trata de “uma aceção especial de *enxó* ← latim *asciōla*, diminutivo [sic] de *ascia*”.

Forma sinónima da mesma origem: esp. *lanzuela* (s.f.); também gal. *lanzoa* (s.f.) ‘ranúnculo’ – aceção de origem metafórica.

[G s.v. *lancĕa*, s.v. *lancĕōla*; Vasconcellos, 1910: 334; REW 4878, 4883; DELP s.v. *Lança*; DA s.v. *Lança*, s.v. *Lançó*; DLP s.v. *lança*; DPLP s.v. *lança*; DRAG s.v. *lanzoa*; RAE s.v. *lanzuela*].

3.11. pt. *linhó* (s.m.) 1. ‘fio grosso para coser calçado, lona, etc.’, 2. ‘tecido que imita o de linho’ ← lat. **linĕōlu-* (s.n.) ‘fio de linho; pedaço de tecido de linho’ < *linĕa-* (*linĭa-*) (s.f.) 1. ‘fio de linho, cordão, cordel’, 2. ‘linha de pesca’, etc. → pt. *linha* (s.f.) 1. ‘fio para coser’, 2. ‘linho’, 3. ‘fio de pesca’, etc.

Existe também a forma *linhol* (s.m.). Semanticamente, tanto o diminutivo como o primitivo continuam as aceções latinas.

Formas sinónimas da mesma origem: gal. *liñó* (s.m.) 1. ‘linha de erva, centeio, trigo, etc., que o trabalhador deixa atrás dele ao segar’, 2. ‘costura de fio gordo feita à mão que une os panos da rede’, 3. ‘arte de pesca consistente num fio, geralmente de nylon, cunha chumbada e um ou vários anzóis empatados no extremo’; esp. *liñuelo* (s.m.) ‘cabo ou ramal das cordas e tranças’; cat. *llinyol* (s.m.) ‘fio grosso para coser calçado’; fr. *ligneul* (s.m.) ‘fio empregado pelos sapateiros e seleiros’.

[G s.v. *linĕa*; REW 5061, 5062, 5073; DELP s.v. *Linha*; DA s.v. *Linha*, s.v. *Linhol*; DLP s.v. *linha*, s.v. *linhol*; DPLP s.v. *linha*, s.v. *linhó*, s.v. *linhol*; DRAG s.v. *liñó*; RAE s.v. *liñuelo*; GDLC s.v. *llinyol*; TLF s.v. *ligneul*].

3.12. pt. ant. *mosteiró* (s.m.) ‘pequeno mosteiro’ ← lat. *mōnastĕriōlu-* (s.n.) ‘pequeno mosteiro’ < *mōnastĕriū-* (s.n.) ‘mosteiro, habitação de monges ou monjas’ → pt. *mosteiro* (s.m.) ‘abadia, convento’.

Tanto o primitivo como o diminutivo guardam as suas aceções originais.

[G s.v. *mōnastĕrium*, s.v. *mōnastĕriolum*; REW 5656; DELP s.v. *-mon(o)-*; DA s.v. *Mosteiro*, s.v. *Mosteiró*; DLP s.v. *mosteiro*, s.v. *mosteiró*; DPLP s.v. *mosteiro*, s.v. *mosteiró*].

3.13. pt. *pió* (s.m./f.) 1. ‘correia que as aves de volataria trazem nos pés’, 2. ‘peia’ ← **pĕdiōla-* (s.f.) ‘grilhão para os pés⁽¹⁷⁾ < *pĕde-* (s.m.) ‘pé, pata, garra’ → pt. *pé* (s.m.) ‘parte final dos membros, sobretudo posteriores, dos vertebrados terrestres’.

Outras formas são *peó* e *piós* (ambas s.m.). O derivado tem aceções figuradas.

(17) Semanticamente, este derivado reconstruído corresponde à palavra latina *pēdica* (G s.v.).

Formas sinónimas da mesma origem: gal. *pioga* (s.f.), *piola* (s.f.); esp. *pihuela* (s.f.).

[G s.v. *pēs*; Vasconcellos, 1887–1889: 304; Vasconcellos, 1895: 180; Viana, 1906b: 276; REW 6355, 6439; DELP s.v. *Pé*; DA s.v. *Pé*, s.v. *Peó*, s.v. *Pió*; DLP s.v. *pé*; DPLP s.v. *pió*, s.v. *piós*; DRAG s.v. *pioga*, s.v. *piola*; RAE s.v. *pihuela*].

3.14. pt. *terço* 1. (adj.m./f., ant.) ‘dizia-se do último animal que nasce da mesma ninhada’, 2. (s.m.) ‘falcão macho’ ← lat. **těrtiölu-* (adj.) ‘terceirozinho’ < *těrtiu-* (adj.) ‘terceiro’ → pt. *terço* 1. (adj.) ‘terceiro’, 2. (s.m.) ‘cada parte de um todo dividido em três partes’.

Outras formas do mesmo vocábulo são *terçô* (adj.m./f., s.m.), *treçó*, *terçogo* (s.m.). A diferença semântica entre o primitivo e o diminutivo é evidente, tendo este último desenvolvido sentidos figurados (por as pessoas crerem que o macho é o terceiro a nascer numa ninhada, ou por o macho do falcão ser de um terço mais pequeno do que a fêmea).

Formas sinónimas da mesma origem: esp. *terzuelo* (s.m.) 1. ‘um terço’, 2. ‘falcão macho’; fr. ant. *terçuel* (s.m.) ‘macho de algumas aves de rapina; falcão macho’ (> fr. *tiercelet* (s.m.) ‘falcão macho, gavião macho’).

[G s.v. *1 tertius*; Vasconcellos, 1887–1889: 304; REW 8678, 8679; DELP s.v. *Três*; DA s.v. *Terço*, s.v. *Terçó*, s.v. *Terçô*, s.v. *Treçó*; DLP s.v. *terço*, s.v. *terçogo*¹; DPLP s.v. *terço*, s.v. *terçó*, s.v. *terçô*; RAE s.v. *terzuelo*; TLF s.v. *tiercelet*].

3.15. pt. *terçó* (s.m.) ‘hordéolo, pequeno tumor no bordo das pálpebras’ ← lat. **trītīcēölu-* (s.n.) < *trītīcēu-* (adj.) ‘de trigo’ → pt. *ø*.

Outras formas do diminutivo, todas de género masculino, são: *terçol*, *terçolho*, *treçolho* (estas duas devidas seguramente a uma contaminação com *olho*), *terçogo*; funciona também um empréstimo do latim, *hordéolo* (s.m.). Conservou-se apenas o derivado com um sentido figurado, provindo da semelhança do hordéolo com um grão de trigo.

Formas sinónimas da mesma origem: gal. *tirizó*, *tricó* (s.m.).

[G s.v. *trītīcēus*; Vasconcellos, 1887–1889: 304; REW 4179, 8924; DELP s.v. *Trigo*; DA s.v. *Terçol*; DLP s.v. *terçogo*², s.v. *terçol*, s.v. *terçolho*, s.v. *treçolho*; DPLP s.v. *terçó*, s.v. *terçolho*; DRAG s.v. *orizól*].

3.16. dial. pt. *tinó* (s.m.) ‘doença, moléstia cutânea; comichão da pele’ (Trás-os-Montes) ← lat. *tīnēöla-* (s.f.) ‘piolho’ < *tīnēa-* (s.f.) ‘traça; verme intestinal’ → pt. *tinha* (s.f.) ‘doença cutânea que ataca o couro cabeludo e o pelo’.

Tanto o primitivo como o diminutivo desenvolveram um sentido metafórico.

Forma da mesma origem: esp. *tiñuela* (s.f.) ‘cuscuta, parasita do linho’.

[G s.v. *1 tñëa*, s.v. *tñëöla*; Viana, 1906b: 481; REW 8746, 8747; DELP s.v. *Tinha*; DA s.v. *Tinha*¹, s.v. *Tinhó*; DLP s.v. *tinha*; DPLP s.v. *tinha*; RAE s.v. *tiñuela*].

4. Além das palavras que provêm de diminutivos latinos⁽¹⁸⁾, encontramos outras às quais se podia atribuir uma etimologia analógica, mas com grandes dificuldades, e isso por algumas razões. Em primeiro lugar, todas viriam de formas latinas hipotéticas, reconstituídas. Em segundo lugar, estes vocábulos maioritariamente não têm correspondentes formais noutras línguas românicas que se pudessem ter originado de um mesmo étimo latino, inclusive na onomástica. Portanto, preferimos incluí-los num grupo separado, sem decidir por enquanto qual a sua etimologia.

4.1. *aguilhó* (s.m.) ‘antigo toucado de mulher’ <> *aguilhão* (s.m.) 1. ‘bico ou ponta perfurante’, 2. ‘ferrão dos insetos’.

[DELP s.v. *Aguilhão*⁽¹⁹⁾; DA s.v. *Aguilhão*, s.v. *Aguilhó*; DLP s.v. *aguilhão*; DPLP s.v. *aguilhão*, s.v. *aguilhó*].

4.2. dial. *bolinhó* (s.m.) ‘espécie de pão-de-ló retangular, coberto de açúcar’ (Minho) <> *bolinha* (s.f.) ‘pequena bola’, *bolinho* (s.m.) ‘pequeno bolo’.

Existem variantes *bolinhol*, *bolinholo* (s.m.).

[DELP s.v. *Bola*; DA s.v. *Bolinhol*, s.v. *Bolinholo*; DLP s.v. *bolinhol*; DPLP s.v. *bolinho*, s.v. *bolinhó*, s.v. *bolinhol*].

4.3. dial. *bugalhó* (s.m.) ‘espécie de planta espontânea em Portugal, venenosa para os carneiros’ (Minho) <> *bugalha* (s.f.) ‘noz-de-galha com tubérculos’, *bugalho* (s.m.) ‘noz-de-galha esférica’.

[DELP s.v. *Bugalha*, *bugalho*; DA s.v. *Bugalha*, s.v. *Bugalho*, s.v. *Bugalhó*; DLP s.v. *bugalha*, s.v. *bugalho*, s.v. *bugalhó*].

(18) REW (5267) nota também pt. *malhó* ‘cadarço, atacador’ (< lat. *malleölu-* (s.m.) 1. ‘martelinho’, 2. ‘ramo podado de vinha ou de árvore, em forma de martelo’ < *malleu-* (s.m.) ‘martelo, malho’), palavra que não encontramos em nenhuma das fontes consultadas. Note-se, porém, que existem as formas galegas *amalló* (s.m.) e *amalloa* (s.f.), ambas com o mesmo sentido e com etimologia idêntica.

(19) Machado propõe um étimo latino: “*aculeölu-*, diminutivo de *acüleu-*, influenciado por *aquiliöne-*”, ou seja, **äcülëölu-* (s.m.) < *äcülëu-* (s.m.) ‘ferrão da abelha’ (G s.v. *äcülëus*). No entanto, parece-nos prescindível recuar até ao latim.

4.4. cortiço (s.m.) ‘ganga; ave galinácea de arribação, pouco maior que a rola’ <> *cortiça* (s.f.) ‘casca do sobreiro, do sobro e da azinheira’, *cortiço* (s.m.) ‘cilindro de cortiça dentro do qual as abelhas fabricam cera e mel’.

Outras formas portuguesas que nomeiam a mesma espécie de ave são *cortiçol* (s.m.) e *cortiçola* (s.f.).

[DELP s.v. *Córtex*; DA s.v. *Cortiça*, s.v. *Cortiço*, s.v. *Cortiçó*; DLP s.v. *cortiça*, s.v. *cortiço*, s.v. *cortiçó*, s.v. *cortiçol*; DPLP s.v. *cortiça*, s.v. *cortiço*, s.v. *cortiçó*, s.v. *cortiçol*].

4.5. ant. faceiro (s.m.) ‘travesseiro para repousar a face’ <> *faceira* (s.f., adj.) 1. ‘carne da parte lateral do focinho da rês’, 2. (dial.) ‘veiga, terra plana de lavoura’ (Trás-os-Montes), 3. ‘bonacheirona, simplória’; *faceiro* (s.m., adj.) 1. ‘casquilho; enfeitado com coisas vistosas mas sem valor’, 2. ‘bonacheirão, simplório’.

Outra forma do nome do objeto é *faceiroa* (s.f.). Existe também sinónimo *faceirão* (s.m.).

Formas sinónimas aparentadas: esp. *faceruelo* (s.m., ant.).

[DELP s.v. *Face*; DA s.v. *Faceira*, s.v. *Faceiro*, s.v. *Faceiro*, s.v. *Faceirôa*, DLP s.v. *faceira*, s.v. *faceiro*; DPLP s.v. *faceira*, s.v. *faceiro*; DRAG s.v. *face-ruelo*].

4.6. ant. figueiró (s.f.) ‘pequena figueira’ <> *figueira* (s.f.) ‘árvore da família das moráceas cujo fruto (figo) é comestível’.

Existem formas analógicas, mas designam outras espécies vegetais: pt. *figueiroa* (s.f.) ‘variedade de pereira’; esp. *higueruela* (s.f.) ‘trevo bituminoso’.

[DELP s.v. *Figo*; DA s.v. *Figueira*, s.v. *Figueiró*, s.v. *Figueirôa*; DLP s.v. *figueira*, s.v. *figueiroa*; DPLP s.v. *figueira*; RAE s.v. *higueruela*].

4.7. dial. galinhó (s.m.) ‘gomo de laranja’ (Trás-os-Montes) <> *galinha* (s.f.) ‘fêmea do galo’.

Existe também a variante *ganhó* (s.m., contração de *galinhó*) e o sinónimo *galelo* (s.m.).

Há formas analógicas, mas que designam uma espécie de ave: gal. *galiñola* (s.f.); esp. *gallinueta* (s.f., em Cuba).

[DELP s.v. *Galo*; DA s.v. *Galinha*, s.v. *Galinhó*, s.v. *Ganhó*; DLP s.v. *galelo*, *galinha*, s.v. *galinhó*, s.v. *ganhó*¹; DPLP s.v. *galelo*, s.v. *galinha*, s.v. *galinhó*, s.v. *ganhó*; DRAG s.v. *galiñola*; RAE s.v. *gallinueta*].

4.8. *giestó* (s.m.) ‘arbusto de flores amarelas, espontâneo em Portugal’ <> *giesta* (s.f.) 1. ‘género de plantas leguminosas, a que pertencem vários arbustos de flores amarelas ou brancas’, 2. ‘vassoura de giesta’.

Forma sinónima aparentada: cat. ***ginestola*** (s.f.).

[DELP s.v. *Giستا*; DA s.v. *Giستا*; DLP s.v. *giesta*, s.v. *giestó*; DPLP s.v. *giesta*; GDLC s.v. *ginestola*].

4.9. *ilhó* (s.m./f.) 1. ‘furo para passar um atacador ou outro fio ou cordão’, 2. ‘aro metálico que guarnece esse furo’, 3. (gíria) ‘ânus’ <> *olho* (s.m.) ‘cada um dos dois órgãos da visão’⁽²⁰⁾.

Funciona também a variante ***ilhós*** (s.m./f.). Antigamente existiam provavelmente as formas **olhó*, **olhol*, **ulhó*.

Forma aparentada: gal. ***illó*** (s.m.) 1. ‘lugar onde a água sai do solo’, 2. ‘terreno muito húmido’.

[Viana, 1906b: 7; DELP s.v. *Olho*; DA s.v. *Ilhó*, s.v. *Olho*; DLP s.v. *ilhó*, s.v. *olho*; DPLP s.v. *ilhó*, s.v. *olho*; DRAG s.v. *illó*].

4.10. *meninó* (s.m.) ‘indivíduo espertalhão, finório’ (informal) <> *menino* (s.m.) 1. ‘criança do sexo masculino’, (adj.) 2. ‘novo, moço’, 3. ‘indivíduo finório, espertalhão’.

[DELP s.v. *Menino*; DA s.v. *Menino*, s.v. *Meninó*; DLP s.v. *menino*, s.v. *meninó*; DPLP s.v. *menino*, s.v. *meninó*].

4.11. *milheiró* (s.m.) 1. ‘espécie de uva preta’, 2. ‘milheiro (planta)’, 3. (dial.) ‘pintassilgo’ (Madeira) <> *milheira* (s.f.) 1. ‘pássaro conirrosto, de asas verdes e cabeça amarela’, 2. ‘milhã; *milheiro* (s.m.) 1. ‘planta que dá milho’, 2. ‘pintarroxo’, 3. ‘uva de bago miudinho’, etc.

Existe a variante ***milheirós*** (s.m.), notada apenas na aceção ‘pássaro’.

[Viana, 1906b: 142; DELP s.v. *Milho*; DA s.v. *Milheira*, s.v. *Milheiro*², s.v. *Milheiró*; DLP s.v. *milheira*, s.v. *milheiro*², s.v. *milheiró*, s.v. *milheirós*; DPLP s.v. *milheira*, s.v. *milheiro*, s.v. *milheiró*].

4.12. dial. *raivó* (s.m.) ‘cogumelo comestível que nasce nos lameiros ao vir do outono, com as primeiras chuvas’ (Trás-os-Montes) <> ? *raiva* (s.f.) 1. ‘doença própria dos cães, caracterizada por acessos furiosos’, 2. ‘fúria, ódio, ira’, 3. ‘espécie de bolo seco; raivinha’.

(20) Vasconcellos (1887–1889: 305) propõe um étimo latino da palavra *ilhó*: “*oculiolum*” (isto é, **öcülüölu-* (s.m.)) ‘pequeno olho’ < *öcülu-* (s.m.) ‘olho’. Porém, não cremos ser necessário recorrer ao latim.

Outra forma portuguesa do mesmo vocábulo é **reivó** (s.m.).

[DA s.v. *Raiva*¹, s.v. *Raivós*, s.v. *Reivós*; DLP s.v. *raiva*, s.v. *raivó*, s.v. *reivó*; DPLP s.v. *raiva*, s.v. *raivó*, s.v. *reivó*].

4.13. dial., ant. **ribeiró** (s.m.) ‘ave ribeirinha’ (Beira Alta) <> *ribeira* (s.f.) ‘rio de pouco caudal e de pequeno curso’, *ribeiro* 1. (s.m.) ‘pequeno rio’, 2. (adj.) ‘diz-se de uma espécie de trigo’.

Outra forma é **ribeirós** (s.m.).

[Viana, 1906b: 369; DELP s.v. *Riba*; DA s.v. *Ribeira*, s.v. *Ribeiro*, s.v. *Ribeiró*; DLP s.v. *ribeira*, s.v. *ribeiro*; DPLP s.v. *ribeira*, s.v. *ribeiro*].

4.14. dial. **ruivó** (s.m.) ‘espécie de tortulho, de folíolos vermelhos e cabeça branca na parte convexa’ (Beira) <> *ruivo* (adj.) ‘que é de uma cor entre o vermelho e o amarelo’.

[DELP s.v. *Ruivo*; DA s.v. *Ruivo*, s.v. *Ruivó*; DLP s.v. *ruivo*; DPLP s.v. *ruivo*].

4.15. **teiró** (s.f.) 1. ‘peça da rabiça do arado que tem mão no dente’, 2. ‘parte da fecharia de algumas armas de fogo’, (s.m.) 3. ‘teima’, 4. ‘má vontade’ <> *teira* (s.f.) ‘peixe acantopterígio (que tem raios espiniformes e duros nas barbatanas)’⁽²¹⁾.

Outras formas portuguesas são **teiroga**, **ateiró** (s.f.), **mateiró** (s.m.) – as últimas duas devidas a uma aglutinação do artigo (*a teiró* → *ateiró*, *uma teiró* → *o mateiró*).

Forma sinónima aparentada: gal. **teiroa** (s.f.).

[DELP s.v. *Teira*; DA s.v. *Ateiró*, s.v. *Teira*, s.v. *Teiró*, s.v. *Teiroga*¹; DLP s.v. *teiró*, s.v. *teiroga*¹; DPLP s.v. *teira*, s.v. *teiró*; DRAG s.v. *teiroa*; RAE s.v. *telera*].

Os vocábulos incluídos neste grupo são, a nosso ver, derivados formados já em português, a partir de palavras vernáculas, e não provêm de diminutivos latinos, como o indicam algumas fontes. Supomos que em português medieval, sem determinarmos um período de tempo preciso, a terminação *-ó* (ou, na época, *-oo* / *-oa*) desempenhava a função de um sufixo independente e servia para formar substantivos denominais. O sufixo **-ó* era decerto diminutivo (como em **4.2.** *bolinhó*, **4.6.** *figueiró*, **4.9.**

(21) O nome deste peixe proviria do lat. **tēlāria-* (adj./s.f.?), um derivado suposto do lat. *tēlum* ‘espada, dardo’ (G s.v. *tēlum*; DELP s.v. *Teira*). É interessante mencionar que do mesmo étimo latino vem esp. *telera* (s.f.), semanticamente sinónimo do pt. *teiró*.

ilhó), mas indicava igualmente outros tipos de relações semânticas: 4.4. *cortiçó*, 4.11. *milheiró*, 4.13. *ribeiró* designam várias espécies de aves, provavelmente a partir do seu modo de alimentação ou habitat; 4.3. *bugalhó*, 4.8. *giestó*, 4.11. *milheiró* são nomes de plantas, originários em designações de outras espécies vegetais com as quais aquelas se parecem; 4.12. *raivó* e 4.14. *ruivó* nomeiam cogumelos, por sua natureza pequenos. A autonomia do sufixo *-ó parece-nos particularmente visível em *galinhó*, *ganhó* (4.7.), que têm como sinónimo *galelo*, formalmente também um diminutivo (de *galo*). Outro exemplo é *faceiró* (4.5.): a existência do seu sinónimo *faceirão* prova, na nossa opinião, que *-ó era intercambiável com outros sufixos. A forma *aguilhó* (4.1.) também se pode dever a uma simples substituição do sufixo e provir de *aguilhão*. Portanto, tratar-se-á de formações paralelas no tempo, sem ter de recorrer a um étimo latino.

5. Finalmente, ao grupo de palavras portuguesas terminadas em -ó poder-se-iam acrescentar também alguns nomes próprios, provindo de nomes comuns.

5.1. No que diz respeito aos topónimos, há muitos nomes de localidades entre os quais se encontram alguns dos vocábulos já mencionados de várias origens: *Queiró* (distrito de Vila Real), *Mó* (Braga). Se levarmos em conta apenas os topónimos formados com o sufixo latino *-ĕllus / -a / -um*, *-iĕllus / -a / -um* (ou com o seu descendente *-ó), veremos nomes que se referem ora à flora local, designando diminutivos ou coletivos, tais como *Pereiró*, *Nogueiró* (ambas no distrito de Braga), *Alijó*⁽²²⁾ (Vila Real), *Teixeiró*, *Milheirós* (Porto), *Figueiró* dos Vinhos⁽²³⁾ (Leiria), ora a construções, como *Grijó* (várias localidades nos distritos de Bragança, Porto, Viana do Castelo), *Mosteiró* de Baixo e *Mosteiró* de Cima⁽²⁴⁾ (Vila Real), *Paçó*⁽²⁵⁾

(22) Segundo Viana (1906b: 481), este nome vem do lat. “*lageola*” que nós identificaríamos como derivado suposto (porque não notado) de *lāgēōs*, *lāgēa* ‘espécie de vinha’.

(23) Alguns destes topónimos têm correspondentes no território espanhol: *Pererueta* (província Zamora), *Noguieruelas* (Teruel), *Milleirós* (Lugo), *Higuerola* (Albacete), *Figuerola* del Camp (Tarragona). Para a explicação da origem deste último, Roca (1961: 207) escreve: “La solución *Ficaria + olu* inclina a un colectivo, a un diminutivo o a un simple distintivo”.

(24) Também há *Mosteiró* no distrito do Porto. Em França, no departamento de Haute-Garonne, existe uma localidade cujo nome tem a mesma origem: *Monestrol*.

(25) Do lat. **pālātīōlu-* cujo sentido seria ‘habitação aparatosa’. Também há *Paçó* no distrito de Viana do Castelo e *Palaçoulo* em Miranda do Douro. Em Espanha encontram-se várias localidades chamadas *Palazuelo*.

(Bragança), *Sequeiró*⁽²⁶⁾ (Porto). A propósito, observe-se que não há provas de que os substantivos comuns **pereiró*, **alijó*, **paçó*, etc. funcionassem em qualquer época na língua geral. Além disso, existem curiosidades de tipo *Coimbró*⁽²⁷⁾ (Vila Real). Há também nomes cuja etimologia ignoramos, como *Travassós* (Braga). Estes topónimos devem ser antigos, dado que só se encontram nas regiões setentrionais de Portugal, o que, no entanto, não impede que possam ter sido exportados para outras terras onde houve colonização portuguesa⁽²⁸⁾.

5.2. Quanto aos antropónimos, apontemos nos apelidos *Queirós* e *Queiroz* (Eça de Queirós!), provindo do já mencionado *queiró(s)*, se calhar através de um mesmo topónimo. Importa mencionarmos também o nome de família *Feijó* (p.ex. João de Moraes Madureira Feijó, autor de *Ortografia, ou Arte de escrever e pronunciar com acerto a Língua Portuguesa*, 1734), originário do lat. *fāsēōlus* (s.m.) ‘feijão’, derivado não-diminutivo de *fāsēlus* ‘id.’. Do sobrenome *Taveira* provém o nome de família *Taveirós* (p.ex. Paio Soares Taveirós, ou Taveiroos, trovador do século XII, autor da famosa *Cantiga da garvaia*). Embora no repertório de apelidos portugueses de Neves (1991) não tenhamos encontrado mais nomes terminados em *-ó*, existem, ou existiram, outros antropónimos deste género, p.ex. *Ulho* (veja-se Vasconcellos, 1887–1889: 305).

Conclusões

6. Da análise do corpus (pontos 3., 4., 5.) ressaltam observações e conclusões de natureza formal e semântica.

6.1. O sufixo latino *-ēōlus / -a / -um* e *-iōlus / -a / -um* evoluiu de uma maneira idêntica nos antigos diminutivos latinos e os resultados desse desenvolvimento também são visíveis nos derivados forjados com ele já em português. O acento moveu-se para o *-ō-*; o *-ē-* e o *-i-*, agora átonos, sincopa-

(26) Com a significação ‘espigueiro’ ← lat. **siccāriōlu-*, diminutivo do adj. *siccārius / -a / -um* ‘próprio para manter num ambiente seco’ que regularmente deu em português *sequeiro* (adj.) ‘seco’, (s.m.) ‘lugar seco; terreno que não é regadio’, e no norte de Portugal também ‘espigueiro’.

(27) Cujo sentido primitivo foi ‘pequena Coimbra’ (Serra, 1986: 100).

(28) Para mais topónimos deste género, veja-se igualmente: Vasconcellos, 1895: 145; Nunes, 1920: *passim*; Losa, 1956: 8–9; Laso, 1969: 214.

ram⁽²⁹⁾; o *-l-* intervocálico regularmente desapareceu (mantendo-se, porém, nalgumas formas paralelas dos vocábulos em questão: **3.1.** *arriol*, **3.3.** *correjola*, *corrijola*, *corriola*, **3.4.** *eirol*, *irol*, **3.8.** *igrejola*, **3.11.** *linhol*, **3.15.** *terçol*, **4.2.** *bolinhol*, *bolinholo*, **4.4.** *cortiçol*, *cortiçola*). O *-ó* ([ɔ] tónico) final das palavras portuguesas, critério de seleção para o nosso corpus, resultou de uma crase do hiato constituído pelo *-o-* tónico e a vogal final dos vocábulos latinos femininos: *oa* → *ó*, e como tal era marca do género feminino. No masculino e neutro, o desenvolvimento paralelo regular deu *-ô* ([o] tónico): *oo* → *ô*, como em *avô*, *Paçô*, *Mosteirô* (veja-se p.ex. Viana, 1906a: 94; Skorge, 1956–1957: 57). O processo de contração dessas vogais terá começado ainda no período do português antigo e acabado no português médio, ou seja, à entrada do século XVI já estava concluído (Castro, 1991: 246; Cardeira, 2006: 49); aliás, pode-se deduzi-lo também das anotações referidas no DELP: *auolo*, *cortizoo* (séc. XII); *avoo*, *Bugaloo*, *Bugaloos* (nomes próprios), *faceutó* (m.), *peyoos* (pl.) (séc. XIII); *aguylhoos* (pl.), *ẽixoo*, *ichoo* (séc. XV); *arrioz* (sg.), *avô*, *enxô*, *filhoos* (pl.), *ilhó*, *treçô* (séc. XVI). A terminação *-ó*, e não *-ô*, das palavras provenientes de antigos masculinos e neutros, bem como a instabilidade do género gramatical de muitas das vozes do corpus (p.ex. **3.4.** *eiró*, **3.6.** *enxó*, **3.7.** *filhó*, **3.9.** *ichó*, **3.13.** *pió*, **4.9.** *ilhó*), deve-se à atração formal dos vocábulos terminados em *-ó*, oriundos do feminino latino, maioritários no corpus⁽³⁰⁾.

Algumas das palavras analisadas apresentam a passagem irregular das antigas vogais pretónicas a *-i-*: **3.1.** *arrió*, **3.7.** *filhó*, **3.8.** *grijó*, **3.9.** *ichó*, **3.13.** *pió*, **3.16.** *tinhó*. É o resultado de uma assimilação regressiva, ou seja, palatalização devida à presença de um som palatal a seguir, tal como “De *ucha* procedem *ucharia* (...) e *ichão*, por *uchão*, com mudança de *u*, absolutamente átono, para *i*, por ser pretónico e ficar antes da consoante palatina *ch* (...)” (Viana, 1906b: 1; veja-se também: Vasconcellos, 1887–1889: 304; Viana, 1906a: 461).

(29) Este desaparecimento do *-ê-* / *-ĩ-* é explicado de várias maneiras: Piel (1975: 46) fala da sua “absorção” [“(…) é sabido que os nomes, comuns e próprios, latinos formados com a ajuda do sufixo *-iōlus* deslocaram, no latim vulgar, o acento da antepenúltima para a penúltima: *-ōLO*, com absorção da semivogal *-j-* (...)”]; Mańczak (1969: 57), no caso de **āviōlus*, crê tratar-se de um desenvolvimento irregular devido à frequência de uso. A nosso ver, poder-se-ia dizer que o *-ê-* / *-ĩ-* deixou pegadas pelo menos nalguns dos vocábulos, influenciando ora a vogal da sílaba anterior (ditongação: p.ex. **3.5.** *eiró*), ora a consoante que o precedia (palatalização: p.ex. **3.7.** *filhó*, **3.11.** *linhó*).

(30) Note-se que em português contemporâneo, para além dos nomes próprios, das palavras *avô* (mais os seus derivados *bisavô*, *trisavô*...), *terçô* (**3.14.** *terçó*), *só* e *nhô* (contrações irregulares de *senhor*, devidas à frequência de uso), não existem vozes de origem latina que sejam terminadas em *-ô*.

Além disso, certos vocábulos têm formas paralelas com -g- epentético (3.4. *eirogo*, 3.14. *terçogo*, 3.15. *terçogo*, 4.15. *teiroga*) que apareceu para eliminar o hiato -oo ou -oa. O fenómeno terá ocorrido antes do século XVI, isto é, na época em que as palavras em questão ainda eram paroxítonas.

6.2. Quanto à morfologia, além das oscilações em género gramatical, nalguns casos também se observa uma passagem das formas do plural ao singular. Deve-se isso à analogia aos plurais *noz-es*, *voz-es*, e ao facto de a terminação -ó aparecer raramente no léxico português, ficando assim suscetível de ser eliminada. As novas formações funcionam como sinónimos das palavras de base e têm novas formas regulares do plural. Deste modo, *arriós* (pl. *arrioses*) designa o mesmo que *arrió* (3.1.), *eirós* e *irós* (pl. *eiroses*, *iroses*) correspondem semanticamente a *eiró* e *iró* (3.4.), e *ilhós* (pl. *ilhoses*) tem o mesmo sentido que *ilhó* (4.9.). Há mais pares deste género: 3.7. *filhó* ~ *filhós*, 3.9. *ichó* ~ *ichós*, 3.13. *pió* ~ *piós*, 4.11. *milheiró* ~ *milheirós*, 4.13. *ribeiró* ~ *ribeirós* (veja-se: Vasconcellos, 1887–1889: 304; Vasconcellos, 1895: 144–145; Viana, 1906a: 94; Viana, 1906b: 7).

6.3. Do ponto de vista semântico, os vocábulos recolhidos são nomes de plantas (3.1. *arrió*, 3.3. *corrijó*, 4.3. *bugalhó*, 4.6. *figueiró*, 4.8. *giestó*, 4.11. *milheiró*, 4.12. *raivó*, 4.14. *ruivó*), animais (3.4. *eiró*, 3.14. *terçó*, 4.4. *cortiçó*, 4.11. *milheiró*, 4.13. *ribeiró*), edifícios e espaços (3.5. *eiró*, 3.8. *grejó*, 3.12. *mosteiró*), instrumentos (3.6. *enxó*, 3.10. *lançó*, 4.15. *teiró*), moléstias (3.15. *terçó*, 3.16. *tin hó*). Além disso, há nomes relativos a vários objetos, em geral pequenos (3.1. *arrió*, 3.9. *ichó*, 3.11. *lin hó*, 3.13. *pió*, 4.1. *aguilhó*, 4.5. *faceiró*, 4.7. *galinhó*, 4.9. *ilhó*), dois termos designando pessoas (3.2. *avó*, 4.10. *meninó*) e dois termos culinários (3.7. *filhó*, 4.2. *bolinhó*). Pode-se dizer que em geral os valores primitivos do diminutivo, ou seja, indicação de pequenas dimensões, semelhança e / ou atitude afetiva, continuam mais ou menos transparentemente visíveis nas palavras do corpus.

Em português contemporâneo, os vocábulos do corpus pertencem maioritariamente ao léxico especializado (p.ex. 3.9. *ichó*, 3.13. *pió*, 4.4. *cortiçó*, 4.11. *milheiró*, 4.15. *teiró*), regional (caraterístico dos dialetos portugueses setentrionais, como 3.5. *eiró*, 3.8. *grejó*, 3.16. *tin hó*, 4.2. *bolinhó*, 4.3. *bugalhó*, 4.7. *galinhó*, 4.12. *raivó*, 4.14. *ruivó*), ou já antiquado, caído em desuso (3.10. *lançó*, 3.12. *mosteiró*, 4.5. *faceiró*, 4.6. *figueiró*, 4.13. *ribeiró*). Há igualmente alguns termos de emprego raro, mas conhecidos comumente, tais como 3.4. *eiró*, 3.7. *filhó*, 3.11. *lin hó*, 3.15. *terçó*, 4.9. *ilhó*, 4.10.

meninó; apenas o termo *avó* (3.2.) se usa na linguagem corrente do dia-a-dia⁽³¹⁾.

Quando em português coexistem os continuadores tanto do primitivo como do diminutivo latinos (situação presente na maioria dos casos), ambas as palavras são semanticamente independentes uma da outra. Os diminutivos desenvolvem mais frequentemente sentidos figurados, derivados mais ou menos diretamente das aceções dos primitivos, às vezes já em latim. Isso diz respeito também aos derivados com o sufixo *-ó. Como se vê pelos vocábulos portugueses do corpus, esta evolução semântica pode ser ligeira (3.8. *igreja* – *grejó*, 3.10. *lança* – *lançó*, 3.11. *linho* – *linhó*, 3.12. *mosteiro* – *mosteiró*, 4.6. *figueira* – *figueiró*, 4.8. *giesta* – *giestó*), mas é capaz de ir tão longe que nem se sente qualquer ligação entre eles (3.7. *folha* – *filhó*, 3.13. *pé* – *pió*, 3.14. *terço* – *terçó*, 4.4. *cortiça* / *cortiço* – *cortiçó*, 4.7. *galinha* – *galinhó*, 4.9. *olho* – *ilhó*). Apenas três pares de palavras designam praticamente o mesmo: 3.5. *eira* – *eiró*, 3.16. *tinha* – *tinhó* e 4.11. *milheira* / *milheiro* – *milheiró*; no entanto, tal como ocorre em termos com sentidos muito próximos, um deles tende a limitar o seu uso – neste caso são os derivados que sofrem restrições territoriais.

Há igualmente poucos casos em que do par de palavras latinas sobrevive apenas o diminutivo, ora com significado figurado (3.4. *eiró*, 3.9. *ichó*, 3.15. *terçó*), ora com aceção do seu primitivo, o que implica uma substituição completa deste último (3.2. *avó*, 3.6. *enxó*).

Acrescente-se que numerosos diminutivos latinos, junto com os seus primitivos, seguem os mesmos caminhos da evolução semântica também nas outras línguas românicas (veja-se Hakamies, 1951: *passim*).

7. Em português contemporâneo, a terminação -ó, contrariamente aos seus homólogos românicos da mesma origem com valor diminutivo e afetivo (p.ex. esp. -*uelo*, -*uela*; cat. -*ol*, -*ola*; também pt. -*ol*, -*ola* de importação estrangeira), não goza do estatuto de um sufixo autónomo, estando ausente nas gramáticas, manuais, dicionários e outros trabalhos sobre a derivação sufixal (Vasconcellos, 1928: 437; Piel, 1940a e 1940b; DELP s.v. -*ol*^l; Skorge, 1956–1957 e 1958; Ali, 1971: 232–248; Cunha & Cintra, 1984: 90–103; Vilela, 1994: 66–85; Villalva, 2003; Borregana, 2008: 105–109). A existência de *-ó no repertório dos sufixos nominais do português medieval e as suas funções talvez se comprovassem num estudo aprofundado do léxico dos textos antigos portugueses que esclarecesse a época até à qual apareciam

(31) Estas afirmações comprovam-se pelos dados que dizem respeito à ocorrência dos vocábulos analisados no corpus do português CETEMPúblico.

novos vocábulos terminados em *-ó* e com que sentido. Fazemos questão de repetir que o sufixo postulado tinha com certeza absoluta um valor diminutivo (topónimo *Coimbró!*), mas veiculava igualmente outros sentidos (uma vez que alternava p.ex. com o sufixo *-ão*, como em *feijão* ~ **feijó* (→ apelido *Feijó*), além dos exemplos já mencionados). Com o tempo, o sufixo **-ó* deixou de ser produtivo e não conseguiu ganhar autonomia, estando hoje em dia só lexicalizado nalgumas palavras como as do nosso corpus.

Referências

- ALI, Manuel SAID (1971), *Gramática Histórica da Língua Portuguesa* (7.ª edição melhorada e aumentada; Estabelecimento do texto, revisão, notas e índices pelo Prof. Maximiano de Carvalho e Silva), Rio de Janeiro, Edições Melhoramentos.
- ANDRADE, Ernesto D' (1993), *Dicionário inverso do português*, Lisboa, Edições Cosmos.
- BDELIC = COROMINES, Joan (2011), *Breve diccionario etimológico de la lengua castellana* (tercera edición muy revisada y mejorada), Madrid, Editorial Gredos [1961].
- BORREGANA, António Afonso (2008), *Gramática – Língua Portuguesa*, Lisboa, Texto.
- CARDEIRA, Esperança (2006), *O Essencial sobre a História do Português*, Lisboa, Caminho.
- CASTILHO, Eugenio DE (1894), *Dicionário de rimas luso-brazileiro*, Revisto, e augmentado com um prefacio e um compendio de metrificação por Antonio Feliciano de Castilho, Terceira Edição Refundida e copiosamente accrescentada, Lisboa, Livraria Ferreira.
- CASTRO, Ivo (1991), *Curso de História da Língua Portuguesa*, Lisboa, Universidade Aberta.
- CETEMPÚBLICO = CORPUS DE PORTUGUÊS CETEMPÚBLICO, disponível em <http://www.linguateca.pt/CETEMPUBLICO/>, consultado em 15/09/2012.
- CUNHA, Celso, CINTRA, Luís Filipe LINDLEY (1984), *Nova Gramática do Português Contemporâneo*, Lisboa, Edições João Sá da Costa.
- DA = DICIONÁRIO ABERTO, disponível em <http://www.dicionario-aberto.net>, consultado várias vezes em agosto de 2012.
- DELP = MACHADO, José Pedro (1952–1959), *Dicionário Etimológico da Língua Portuguesa* (2 vol.), Lisboa, Confluência.
- DLP = DICIONÁRIO DA LÍNGUA PORTUGUESA (2008), Porto, Porto Editora.
- DPLP = DICIONÁRIO PRIBERAM DA LÍNGUA PORTUGUESA, disponível em <http://www.priberam.pt/dlpo/>, consultado várias vezes em agosto de 2012.
- DRAG = DICIONÁRIO DA REAL ACADEMIA GALEGA, disponível em <http://www.realacademiagalega.org/dicionario#inicio.do>, consultado várias vezes em agosto de 2012.

- ESPADA, Francisco (2006), *Manual de Fonética. Exercícios e Explicações*, Lisboa–Porto, Lidel.
- G = GAFFIOT, Félix (1934), *Dictionnaire Latin–Français*, Paris, Hachette.
- GDLC = *GRAN DICCIONARI DE LA LENGUA CATALANA*, disponível em <http://www.diccionari.cat/>, consultado várias vezes em agosto de 2012.
- HAKAMIES, Reino (1951), *Étude sur l'origine et l'évolution du diminutif latin et sa survie dans les langues romanes*, Helsinki, Academiae Scientiarum Fennicae.
- LASO, Abelardo MORALEJO (1969), “Toponímia gallega de cereales de cultivo”, *Cuadernos de Estudios Gallegos* XXIV/72–73–74, pp. 206–224.
- LOSA, António (1956), *A Dominação Árabe e a Toponímia a Norte do Douro*, Braga, Edições Bracara Augusta.
- MAŃCZAK, Witold (1969), “Survivance du nominatif singulier dans les langues romanes”, *Revue Romane* 4/1, pp. 51–60.
- MATEUS, Maria Helena MIRA, FROTA, Sónia, VIGÁRIO, Marina (2003), “Prosódia” in M.H. Mira Mateus et al., *Gramática da Língua Portuguesa* (7.^a edição), Lisboa, Caminho, pp. 1035–1076.
- MATEUS, Maria Helena MIRA, FALÉ, Isabel, FREITAS, Maria João (2005), *Fonética e Fonologia do Português*, Lisboa, Universidade Aberta.
- MEILLET, Antoine, VENDRYES, Joseph (1924), *Traité de grammaire comparée des langues classiques*, Paris, Librairie Ancienne Édouard Champion.
- NEVES, Vítor Manuel LEAL PEREIRA (1991), *Apelidologia popular portuguesa*, Lisboa.
- NUNES, José Joaquim (1920), *A vegetação na toponímia portuguesa*, Coimbra, Imprensa da Universidade.
- OTRĘBSKI, Jan, SAFAREWICZ, Jan (1937), *Gramatyka historyczna języka łacińskiego*, cz. 1, Warszawa, Komitet Wydawniczy Podręczników Akademickich przy Ministerstwie Wyznań Religijnych i Oświecenia Publicznego.
- PALMER, Leonard R. (1988), *The Latin Language*, Norman, University of Oklahoma Press.
- PIEL, Joseph Maria (1940a), “A formação dos nomes de lugares e de instrumentos em português”, separata do *Boletim de Filologia* VII, pp. 1–17.
- PIEL, Joseph Maria (1940b), “A formação dos substantivos abstractos em português”, separata de *BIBLOS* XVI/1, pp. 1–29.
- PIEL, Joseph Maria (1975), “Novos fragmentos de toponímia galega oriunda de nomes latinos de senhorios rurais medievos”, *Verba* 2, pp. 45–58.
- RAE = *DICCIONARIO DE LA LENGUA ESPAÑOLA* (VIGÉSIMA SEGUNDA EDICIÓN), Real Academia Española, disponível em <http://www.rae.es/rae.html>, consultado várias vezes em agosto de 2012.
- REW = MEYER-LÜBKE, Wilhelm (1911), *Romanisches Etymologisches Wörterbuch*, Heidelberg, Carl Winter's Universitätsbuchhandlung.

- ROCA, Pedro CATALÁ Y (1961), “Fitotoponimia mayor de la arquidiócesis de Tarragona (Cataluña)”, *Studia Onomastica Monacensia* III, pp. 199–216.
- SERRA, Pedro CUNHA (1986), “A influência árabe na Península Ibérica. Aspectos da sua dimensão e profundidade”, separata de: Adel Y Sidarius (ed.), *Islão e Arabismo na Península Ibérica: Actas do XI Congresso da União Europeia de Arabistas e Islamólogos*, Évora, Universidade de Évora, pp. 97–112.
- SKORGE, Sílvia (1956–1957), “Os sufixos diminutivos em Português”, *Boletim de Filologia* 16, pp. 50–90, 222–305.
- SKORGE, Sílvia (1958), “Os sufixos diminutivos em Português”, *Boletim de Filologia* 17, pp. 20–53.
- TLF = *LE TRÉSOR DE LA LANGUE FRANÇAISE INFORMATISÉ*, disponível em <http://atilf.atilf.fr/>, consultado várias vezes em agosto de 2012.
- VASCONCELLOS, Carolina MICHAËLIS DE (1887–1889), “Etimologias portuguesas”, *Revista Lusitana* I, pp. 298–305.
- VASCONCELLOS, Carolina MICHAËLIS DE (1895), “Fragmentos etimológicos”, *Revista Lusitana* III, pp. 129–190.
- VASCONCELLOS, Carolina MICHAËLIS DE (1910), “Mestre Giraldo e os seus tratados de Alveitaria e Cetraria. Parte II – Estudos etimológicos”, *Revista Lusitana* XIII, pp. 222–432.
- VASCONCELLOS, José LEITE DE (1927), “Observações ao “Elucidario” do P.º Santa Rosa de Viterbo”, *Revista Lusitana* XXVI, pp. 111–146.
- VASCONCELLOS, José LEITE DE (1928), *Opúsculos*, Volume I, *Filologia*, Parte I, Coimbra, Imprensa da Universidade.
- VIANA, Aniceto DOS REIS GONÇÁLVES (1906a), *Apostilas aos Dicionários Portugueses*, Tõmo I, Lisboa, Livraria Clássica Editora.
- VIANA, Aniceto DOS REIS GONÇÁLVES (1906b), *Apostilas aos Dicionários Portugueses*, Tõmo II, Lisboa, Livraria Clássica Editora.
- VILELA, Mario (1994), *Estudos de Lexicologia do Português*, Coimbra, Livraria Almedina.
- VILLALVA, Alina (2003), “Formação da palavras: afixação” in M.H. Mira Mateus et al., *Gramática da Língua Portuguesa* (7.ª edição), Lisboa, Caminho, pp. 937–967.

REVISITANDO A VARIAÇÃO ENTRE DITONGOS NASAIS FINAIS ÁTONOS E VOGAIS ORAIS NA COMUNIDADE DE FALA DO RIO DE JANEIRO

REVISITING THE VARIATION BETWEEN FINAL UNSTRESSED NASAL DIPHTHONG AND ORAL VOWELS IN THE SPEECH COMMUNITY OF RIO DE JANEIRO

Christina Abreu Gomes*
christina-gomes@uol.com.br

Cássia Mesquita**
cassiameq@hotmail.com

Taís da Silva Fagundes***
taisdasilva.f@gmail.com

Esse trabalho revisita a variação entre ditongos nasais finais átonos e vogais orais no português brasileiro com o objetivo de verificar o papel do item lexical e de contextos estruturais na variação a partir de dados de uma amostra da comunidade de fala do Rio de Janeiro conjugando a Sociolinguística Variacionista e a Fonologia de Uso. Os resultados são de um estudo piloto com 12 falantes da Amostra Censo 2000. A análise estatística através do Rbrul revelou efeito do item lexical e de condicionamento prosódico. A distribuição das variantes por faixa etária indica uma situação de variação estável. No entanto, é possível identificar a implementação da variante vogal oral quando se observa o comportamento de itens lexicais específicos.

Palavras-chave: mudança sonora; item lexical; ditongo nasal.

* Christina Abreu Gomes é pesquisadora do CNPq e Professora Associada 3 da Universidade Federal do Rio de Janeiro, vinculada ao Departamento de Linguística e Filologia e ao Programa de Pós-Graduação em Linguística, Rio de Janeiro, Brasil.

** Cássia Mesquita é graduanda em Letras Português-Inglês na Universidade Federal do Rio de Janeiro, Departamento de Anglo-Germânicas, Rio de Janeiro, Brasil.

*** Taís da Silva Fagundes em Letras Português-Espanhol na Universidade Federal do Rio de Janeiro, Departamento de Neolatinas, Rio de Janeiro, Brasil.

This work revisits the variation between final unstressed nasal diphthongs and final oral vowels in Brazilian Portuguese in order to verify the role of the lexical item and that of the structural contexts based on data from a spoken sample of the speech community of Rio de Janeiro according to the Variationist and Usage-based Approaches. The results are from a pilot study with 12 speakers from Censo 2000 sample. The statistical analysis with Rbrul revealed the effect of lexical item and prosodic constraint. The distribution of the variants by age groups indicates a situation of stable variation. However, it is possible to identify the implementation of the oral vowel variant when the behavior of specific lexical items is observed.

Keywords: sound change; lexical item; nasal diphtong.

*

1. Introdução

A alternância entre ditongo nasal final e vogal átona é uma variável estudada em amostras de fala de algumas variedades do Português Brasileiro (PB) com tratamento dos dados dentro da perspectiva da sociolinguística variacionista, como nos trabalhos de Votre (1978) e Guy (1981) para a comunidade de fala do Rio de Janeiro; Battisti (2002), Schwindt e Bopp da Silva (2010) e Schwindt, Bopp da Silva e Quadros (2012) com dados de diversos municípios de estados da região sul do Brasil, e Silva et al. (2012) sobre a variedade falada em Belo Horizonte. Diversos também são os estudos que tratam do status fonológico da nasalidade no português. À interpretação pioneira de Mattoso Câmara (1977/1949) dos ditongos nasais constituídos de vogais orais seguidas de arquifonema nasal, soma-se a de Mateus (1975) para o português europeu e seguem-se reinterpretações do caráter bissegmental da nasalidade tendo como base a teoria autosegmental (Wetzels, 1997), teoria da otimalidade (Battisti, 1997) e geometria de traços (Bisol, 1998). As diversas análises advogam uma representação abstrata de vogal oral seguida de segmento ou traço nasal, que se superficializa em função de um processo de atribuição de nasalidade à vogal e sua posterior supressão.

A questão da nasalidade no PB também foi tratada por Albano (1999:30-32) tendo como base o modelo da Fonologia Articulatória de Brownman & Goldstein (1992). Nessa abordagem, o gesto articulatório é a unidade de análise (e representacional). Com base nos dados acústicos de

Souza (1994), a autora mostra que a vogal não se nasaliza por completo, mas apresenta uma configuração de vogal oral no início, formantes de vogal oral, e de vogal nasal no final. Os dados também mostram que se a vogal é longa a nasalidade intrusiva ou murmúrio nasal pode não aparecer, e, se aparece, tem duração inversa à da vogal. De acordo com a Fonologia Articulatória (ou Gestual), um gesto de abertura vélica se inicia após o gesto vocálico, se sobrepõe a este e termina antes do gesto de abertura. Segundo Albano, a presença ou não da nasalidade vai depender do grau de sobreposição entre o gesto consonantal seguinte e os gestos vocálico e de abertura vélica. O grau de sobreposição pode variar em função do contexto segmental, prosódico e pragmático. Assim, a interpretação de Albano captura a gradualidade da nasalidade e o fato de que no início da implementação fonética a vogal não é nasal desde o seu início. Um tratamento estritamente abstrato (ou fonológico) não permitiria capturar essa dinâmica do processo.

É objetivo deste estudo discutir a natureza da variação ditongo nasal átono e vogal oral no Português Brasileiro a partir de dados de produção espontânea de falantes nativos da comunidade de fala da cidade do Rio de Janeiro. A alternância tem sido tratada nos trabalhos sobre o PB como um caso de processo variável condicionado por fatores de ordem fonológica e morfológica. Pretende-se, com a introdução do item lexical como um grupo de fatores específico situar o papel do item lexical de forma independente assim como situar melhor os efeitos do contexto fonético-fonológico e morfológico. O estabelecimento do item lexical como variável explicativa é possível com a utilização do Programa Rbrul (Johnson, 2009), um instrumental estatístico que permite a análise de variáveis aninhadas, também definidas como *random effects*.

Outro objetivo é observar, em dados levantados de amostra de fala mais recente do PB, a afirmação de Votre sobre o processo irreversível de perda da nasalidade. O trabalho pioneiro de Votre (1978) apontou que a ocorrência de vogais orais alternando com ditongos nasais, observada na comunidade de fala do Rio de Janeiro, como em *órfão ~ órfu*, é um processo de mudança nos sentido da perda da nasalização e redução do ditongo que se implementa na língua desde o português arcaico. Para Votre, a perda da nasal é uma tendência irreversível.

Os dados deste estudo piloto foram levantados de 12 falantes da Amostra Censo 2000 do Programa de Estudos sobre Uso da Língua (PEUL), sediado na Universidade Federal do Rio de Janeiro. Os 12 falantes foram

selecionados em função de sua distribuição nas variáveis de estratificação da amostra. Esses aspectos serão abordados na seção 4(1).

O estudo se desenvolve dentro dos pressupostos da Sociolinguística Variacionista no que diz respeito ao tipo de dado utilizado e à análise da variação. A interpretação dos resultados se ancora na perspectiva teórica da Fonologia de Uso. Nesta perspectiva, a gramática é definida como contendo heterogeneidade sistemática. Em outras palavras, a variabilidade observada no uso é resultante da variabilidade do sistema, conforme definido por Weinreich, Labov e Herzog (1968). Além disso, o modelo confere à variação sociolinguística um caráter representacional (Bybee, 2001, 2010; Pierrehumbert, 2003, 2012; Foulkes e Docherty, 2006; Gomes e Silva, 2004).

Nas seções a seguir são apresentados os aspectos teóricos envolvidos no estudo da variação e mudança fonológica e da natureza da alternância sincrônica, incluindo a perspectiva da Fonologia de Uso, a metodologia de coleta, análise dos dados e discussão dos resultados.

2. Questões teóricas da mudança sonora e modelagem da variação sonora

O debate em torno da unidade da mudança sonora, se o segmento ou o item lexical, remonta ao século XIX e pode ser sumarizada nas posições divergentes de neogramáticos e dialetologistas. Neogramáticos propuseram um conjunto de princípios que regem a mudança fonética, definidos a partir do conceito de lei fonética: 1) a mudança é regular, isto é, ocorre sem exceções; 2) é foneticamente gradual, isto é, se implementa foneticamente de maneira gradual e imperceptível, e lexicalmente abrupta, isto é, afeta todas as palavras do léxico que apresentam o contexto de implementação da mudança; 3) ocorre somente em função de motivações fonética; 4) não admite exceção, e as aparentes exceções podem ser explicadas pela analogia e empréstimos linguísticos. Então, a visão neogramática clássica pode ser sumarizada como sendo uma mudança de uma determinada classe de sons motivada foneticamente que afeta, ao mesmo tempo, todas as palavras que contêm aquela classe de sons (cf. Bloomfield (1984/1933) sem exceção.

(1) Um estudo piloto se refere a um estudo de um subconjunto de falantes de uma amostra para uma primeira testagem de hipóteses e variáveis independentes. A expectativa é a de que se há uma boa distribuição de falantes pelas variáveis de estratificação da amostra e se as hipóteses e variáveis independentes foram bem definidas, os resultados obtidos no estudo piloto tenderão a se manter com o acréscimo de mais dados de outros falantes do restante da amostra.

Já para os dialetologistas, a mudança se propaga tendo como unidade o item lexical (Bloomfield, 1984/1933:328). Os estudos dialetológicos procuravam opor o caráter regular das leis fonéticas à irregularidade das diferenciações observadas nas variedades dialetais. À proposta dos dialetologistas do século XIX, acrescentam-se, nos anos 60, trabalhos com evidências de mudanças que não se enquadram no modelo neogramático. Os novos difusionistas do século XX defendem que o item lexical é a unidade da mudança, e as mudanças sonoras são definidas como foneticamente abruptas e lexicalmente graduais. Wang (1969) e Chen & Wang (1977) apresentaram evidências de que exceções à mudança sonora podem estar relacionadas ao resultado da sobreposição de duas mudanças ao invés da atuação da mudança e da analogia, isto é, o resultado irregular se deve a uma outra mudança que atingiu itens que poderiam ter sido atingidos por uma primeira mudança, mas que não haviam sido atingidos ainda. Também apontam que há mudanças que não se enquadram no tipo lexicalmente abrupto e foneticamente gradual como é o caso da cisão tonal no Chaouzhou, variedade falada na China, em que homônimas se dividiram em duas classes tonais sem motivação fonética ou gramatical. As mudanças resultantes também não puderam ser explicadas como casos de analogia. Para Wang (1969), difusão lexical diz respeito à maneira como a mudança sonora se implementa, ou seja, neste caso, a unidade da mudança é o item lexical. A mudança pode afetar numa primeira fase somente alguns itens enquanto outros permanecem sem ser afetados, havendo, portanto, uma implementação gradual no léxico com resultado regular em que todos os itens relevantes são afetados. Mas, por razões como competição de outra mudança para o mesmo contexto, ou mesmo uma motivação extralinguística, a mudança pode deixar de operar e não atingir a todos os itens lexicais relevantes.

Labov (1981, 1994) procura resolver a controvérsia entre a perspectiva neogramática de mudança sonora e a dos difusionistas defendendo que é a natureza da mudança que determina o tipo, se tendo o segmento ou o item como unidade de mudança. No entanto, para Labov, a maioria das mudanças seria do tipo neogramático. Segundo Labov, difusão lexical não exclui a possibilidade de condicionamento fonético, uma vez que a seleção de uma palavra no processo de mudança pode ter uma motivação fonética. A questão é que se há de fato difusão lexical é porque o condicionamento fonético, ou mesmo de outra natureza estrutural, não é suficiente para dar conta da mudança.

Oliveira (1991) questiona o fato de que os neogramáticos conheciam somente o produto da mudança, isto é, seu resultado final. Portanto, uma

mudança sonora pode se implementar gradualmente no léxico e o resultado afetar todos os itens que satisfazem as condições estruturais da mudança. Oliveira propõe que a mudança é lexicalmente gradual e que pode ser motivada foneticamente. Assim, o autor reinterpreta a posição neogramática como a descrição do resultado de uma mudança, e não como a natureza do processo propriamente dito. Para Oliveira, qualquer mudança sonora se implementa lexicalmente.

Bybee (2002) enquadra a visão dicotômica entre visão neogramática e difusionista em um modelo de gramática que separa léxico de gramática. Em contraposição, na perspectiva dos Modelos baseados no Uso, ou especificamente, na Fonologia de Uso (Bybee, 2001; Pierrehumbert, 2003, 2012; Silva e Gomes, 2007), a questão pode ser colocada de maneira que segmento e item não estão em oposição, uma vez que, neste modelo, se postula que a gramática é emergente das representações sonoras das palavras no léxico. Essas representações incluem o detalhe fonético, adquirido em função da experiência dos falantes de produzir e ouvir em diversos contextos linguísticos e sociais.

Essa concepção de representação da variabilidade e gradualidade fonética observada no sinal da fala é capturada em um modelo de representação conhecido como Modelo de Exemplares. Esse modelo foi trazido da Psicologia para a Linguística por Johnson (1997). Nessa proposta, cada categoria é representada por um conjunto de ocorrências, as instâncias de produção e percepção. As representações são, portanto, extremamente detalhadas e estão organizadas em um mapa cognitivo de maneira que as instâncias semelhantes estão próximas e as diferentes estão mais distantes (Pierrehumbert, 2001), e delas emergem representações mais abstratas como categorias fonéticas, moldes silábicos, lexicais, relações morfofonológicas (Pierrehumbert, 2003).

Conceber a representação da variação de acordo com o Modelo de Exemplares, segundo Bybee (2002), é uma forma de capturar a gradualidade da mudança, uma vez que o falante gradualmente atualiza suas representações, que afetam itens e não somente segmentos. Uma competição de efeitos de diversas ordens, que vão da frequência de ocorrência do item lexical ao contexto estrutural e social, atuam na propagação da mudança no léxico. Essa modelagem teórica captura assim o papel do item e do contexto fonético-fonológico, rompendo com o caráter dicotômico do papel do item lexical e do contexto gramatical no condicionamento da mudança sonora. Ainda, com relação à frequência de ocorrência do item lexical, já apontada por Schuchardt no século XIX (cf. Bybee, 2001), a hipótese é a de que, como

a mudança na articulação se implementa gradualmente na produção, então quanto mais frequentes forem os itens mais sujeitos estarão de serem atingidos pela mudança.

Para a modelagem teórica da variação sonora, a proposição dos Modelos baseados no Uso de tratar a variação de maneira representacional implica na adoção da ideia de que as variantes fazem parte do conjunto de representações de um item. Soma-se a isso a proposição da organização interna do léxico em redes de conexões lexicais por semelhança sonora e semântica entre os itens lexicais. Gomes e Silva (2004) e Foulkes e Docherty (2006) também apontam o Modelo de Exemplares como uma proposta interessante para capturar a variação sonora, tal qual proposta por Pierrenumbert (1994, 2003).

3. Estudos sobre a alternância ente ditongos nasais e vogais orais no PB

O estudo de Votre (1978), sobre a comunidade de fala do Rio de Janeiro, considerou conjuntamente ditongos nasais tônicos e átonos tanto de nomes quanto de formas verbais, e, de acordo com sua análise, a retenção da nasalização é um processo condicionado por fatores fonológicos e morfológicos.

Guy (1981), além de estabelecer o mesmo envelope da variação de Votre, ditongos finais átonos e tônicos, por outro lado, tratou o processo como desnasalização, defendendo que seria necessário identificar dois processos distintos, um sintático e outro fonológico, concorrendo para a ocorrência de formas variantes alternando a realização do ditongo nasal e vogal oral. De um lado, há um processo sintático, uma regra variável, que pode resultar em formas verbais marcadas e não marcadas quanto à informação de número-pessoa, como em *Eles comem* ~ *Eles come*, como também de formas como *Eles fizeram* ~ *Eles fez*. De outro, concorre uma regra variável fonológica que também pode resultar em alternâncias de formas verbais com terminação desnasalizada e que podem ou não indicar ausência de marca morfológica, como, em, respectivamente, *Eles come* e *Eles comeru*, ou também afetar nomes, como em *homem* ~ *homi*. De acordo com Guy, trata-se de um processo de variação estável condicionado somente por fatores fonológicos. Ambos os estudos foram realizados com falantes não escolarizados da comunidade de fala do Rio de Janeiro, de amostra gravada na década de 70, com inclusão de universitários somente no trabalho de Votre. No entanto, estudos subsequentes a esses trabalhos, nas décadas seguin-

tes, Naro (1981), Scherre e Naro (1997, 1998), forneceram uma descrição exaustiva do uso variável da marca de concordância verbo-sujeito, em especial na comunidade de fala do Rio de Janeiro, mostrando a competição de fatores de diversas naturezas, identificando a especificidade da alternância de marcação de número-pessoa no português brasileiro sem o estabelecimento de motivação fonético-fonológica e relação com um processo fonológico de desnasalização.

Battisti (2002) e Schwindt e Bopp da Silva (2010) mostraram que a sílaba tônica bloqueia a desnasalização e trataram da alternância de vogais orais e ditongos nasais átonos considerando tanto as ocorrências em nomes quanto em verbos. Battisti (2002), comparando dados de 2 municípios de cada um dos três estados da região sul do Brasil, conclui que há diferença entre as localidades, havendo um percentual mais alto de desnasalização em Santa Catarina que nos estados do Rio Grande do Sul e Paraná, e que há favorecimento de ocorrência de vogal oral quando o contexto anterior é uma consoante nasal, o contexto seguinte é uma vogal, a vogal núcleo do ditongo é a vogal *e* e em nomes terminados em *-gem*, como *viagem*, *passagem*. A autora mostra que a variável classe de palavra se refere ao comportamento de alguns itens lexicais, em especial nomes terminados em *-gem* e verbos no pretérito perfeito, em que o ditongo nasal não tem status morfológico. Quanto ao contexto seguinte, o efeito da vogal parece configurar um efeito prosódico e não segmental, uma vez que em geral nos dados a vogal seguinte correspondia a um contexto átono que leva também a um efeito de sandhi externo com a eliminação do ditongo ou da vogal oral como em *for. imbora (foram embora)*. E em relação ao contexto anterior, a autora menciona que a grande maioria dos dados com consoante nasal precedente ao ditongo são instâncias da palavra *homem*.

O estudo de Schwindt e Bopp da Silva (2010) compara dados de todas as cidades que compõem o banco de dados VARSUL, sobre variedades faladas na região sul do Brasil, e seus resultados corroboram os achados de Battisti para condicionamento estrutural. O VARSUL compreende amostras de fala espontânea representativas de três estados da região Sul do Brasil (Paraná, Santa Catarina e Rio Grande do Sul). Em cada estado há amostras de quatro cidades, cada uma composta por um conjunto de 24 falantes estratificados em função de sexo, escolaridade e idade. O Projeto Censo/PEUL serviu de modelo para a constituição dessas amostras (Collischonn e Monareto, 2012). Em estudo mais recente, Schwindt, Bopp da Silva e Quadros (2012) procuram aprofundar a questão do condicionamento morfológico através do estabelecimento de uma variável que contempla diferentes itens lexicais

em categorias como nomes terminados com –gem morfológico (ex. reportagem), nomes terminados em –gem na raiz (origem), nomes (homem) e diferentes tipos de verbos definidos em função do tempo e do modo. Nesse estudo, os resultados corroboram os anteriores indicando um grande favorecimento de ocorrência da vogal oral em nomes terminados em –gem,

Os estudos realizados com os dados do VARSUL revelam um forte efeito de item lexical. No entanto, o desenho metodológico da análise estatística oblitera o efeito do item uma vez que este sempre é analisado em função de algum tipo de categoria, como, por exemplo, “nomes terminados em –gem morfológico”, “nomes terminados em –gem parte da raiz” ou “tempo verbal”, e nunca individualmente.

Mais recentemente, Silva et al. (2012) revisitaram a alternância de ditongos nasais finais átonos para as formas verbais a partir da perspectiva dos Modelos baseados no Uso em dados da comunidade de fala de Belo Horizonte. Segundo os autores, as realizações com ditongos nasais e vogais orais das formas verbais (*falam* ~ *fala*, *cantavam* ~ *cantava*, *falaram* ~ *falaru*) são parte de uma mesma categoria, isto é, estão representadas no léxico. A alternância estaria relacionada a uma reorganização de rotinas motoras associadas às variantes. O estudo detectou a importância da frequência de ocorrência do item lexical, no caso, da forma verbal, no gerenciamento da alternância, em que verbos mais frequentes tendem a ocorrer com a vogal oral.

Com relação à alternância observada sincronicamente no PB de formas verbais de 3ª pessoa plural que ocorrem com o ditongo nasal final átono alternando com a vogal oral, Oliveira (1983) procura explicar a variação sincrônica como sendo resultado de um resíduo histórico. Para Oliveira a variação sincrônica é resultante de diferentes processos de mudança que ocorreram no português arcaico e que geraram na língua as formas alternantes sincrônicas. Segundo o autor, textos do português antigo registram as seguintes vogais nasais átonas finais, -õ, -ã e -ẽ resultantes da evolução, respectivamente, de -unt, -(b)ant e -ent. Essas vogais teriam sofrido um processo de desnasalização, gerando a alternância vogal oral ~ vogal nasalizada. Em um momento posterior as formas nasais simples das vogais *e* e *a* teriam se ditongado, respectivamente em -ëy e -ão, resultando na alternância com as vogais correspondentes orais que já alternavam com as vogais nasalizadas *e*, assim, as alternâncias verificadas sincronicamente seriam resíduos históricos, isto é, um caso de variação presente na língua há muito séculos, cujas variantes são o resultado de diferentes processos de mudança muito anteriores à variação observada no Português Brasileiro contemporâneo. Dentro dessa hipótese, ainda, a forma -ão de 3ª pessoa do presente

do indicativo seria resultante de uma analogia com as formas de 3ª pessoa do perfeito e imperfeito do indicativo. Assim, para Oliveira, essa hipótese tem a vantagem de explicar porque, do ponto de vista sincrônico, o ditongo nasal átono final *-ão* ora alterna com *-a*, como em *eles falam ~ eles fala*, ora com *-u* em *eles responderam ~ eles responderu*, sem atribuir a uma regra ou processo fonológico sincrônico que resulta em diferentes outputs, pois ora [-ãw] está em alternância com [-a], como em *Eles falam ~ Eles fala*, ora com [-u] com em *responderam ~ responderu*.

Os estudos sincrônicos do português, de um lado, apontam para a importância do item lexical, mesmo que indiretamente através de categorias morfológicas, ao passo que a interpretação diacrônica de Oliveira aponta para a possibilidade de as variantes sincrônicas serem resultantes de processos de mudanças diferentes atuando em momentos diferentes da língua.

4. Envelope da variação, objetivos e metodologia

Na variedade do Rio de Janeiro, o ditongo nasal final átono alterna com vogal nasalizada e vogal oral, especialmente nos casos em que o ditongo faz parte da marca morfológica de número e pessoa, como em *fizeram ~ fizeram ~ fizeru*. Na análise aqui apresentada, a variação foi tratada como binária opondo as realizações nasais à variante vogal oral. A vogal oral átona na variedade de fala do Rio de Janeiro tem realização reduzida, podendo também ser ensurdecida.

Diferentemente dos trabalhos mencionados na seção anterior, não foram considerados em conjunto os casos em que a variante vogal oral também corresponde à ausência de marca de concordância entre verbo e sujeito. Então, alternâncias do tipo *Eles cantam ~ Eles canta* ou *Eles mudam ~ Eles muda* não foram computadas como dados para este estudo, isto é, tanto a realização com a marca quanto a não-marcada não foram incluídas no *corpus* levantado, uma vez que, conforme já atestado nos trabalhos sobre concordância verbal variável sobre o PB e mencionado na seção anterior, há efeitos de outra natureza atuando, e que não podem ser estendidos aos nomes, o que implicaria na ausência de ortogonalidade dos grupos de fatores e a consequente impossibilidade de avaliar outras motivações que competem na realização variável da concordância verbal de terceira pessoal do plural no PB.

Foram então incluídos todos os casos de ocorrência de qualquer uma das três variantes de nomes, como em *homem*, *jardinagem*, *órfão*, e de formas verbais cuja alternância na realização da sílaba átona final nasal com vogal oral não implica ausência de marca morfológica de concordância, como *estiveram* ~ *estiveru*, *cantarem* ~ *cantari*. No entanto, foram encontrados somente casos de pretérito perfeito na amostra pesquisada, já que não houve ocorrência de formas de subjuntivo que se enquadram no critério de inclusão. A checagem das transcrições foi feita por duas auxiliares de pesquisa, alunas da graduação dos cursos de Letras da UFRJ. Os casos duvidosos foram excluídos do corpus e constituíram menos de 1% do total. Foi obtido um total de 322 dados que correspondem a ocorrências de ditongo nasal, vogal nasal e vogal oral.

Conforme mencionado na Introdução, o objetivo principal foi o de verificar em uma amostra mais recente a afirmação de Votre (1978) relativa à implementação de um processo de desnasalização de ditongos no Português Brasileiro, uma vez que houve diferença de interpretação nos trabalhos de Votre e de Guy, respectivamente, se mudança em progresso ou variação estável. O outro objetivo é o de discutir a natureza dos condicionamentos da variação através da observação tanto de variáveis estruturais, como o contexto fonético-fonológico, quanto o papel do item lexical em relação às tendências de realização das variantes. Conforme já mencionado, essa é uma questão central para o estudo de variação e mudança sonora, cujos resultados podem fornecer subsídios para a discussão não só da natureza da mudança, mas também para a discussão em torno da modelagem da gramática.

Os dados foram coletados da Amostra Censo 2000. Trata-se de uma amostra aleatória da comunidade de fala da cidade do Rio de Janeiro, constituída no ano 2000, estabelecida dentro do modelo laboviano de coleta de fala espontânea e de estratificação dos indivíduos, composta por 32 falantes distribuídos em função de escolaridade, sexo e faixa etária. Essa amostra segue essa estratificação por ter sido concebida com o objetivo de permitir o estudo de processos de mudança linguística em tempo real, comparando duas amostras de fala de uma mesma comunidade em dois momentos diferentes no tempo, e, por isso, mantém a mesma estratificação da Amostra Censo 1980 PEUL/UFRJ, organizada vinte anos antes. A estratificação por faixa etária leva em conta quatro níveis: 7 a 14 anos, 15 a 25 anos, 26 a 49, e 50 anos ou mais. Foram considerados 3 níveis de escolaridade, que correspondem ao ensino fundamental e ensino médio, divididos em 3 grupos: de 1 a 4 anos, de 5 a 8 anos, e de 9 a 11 anos. Os oito primeiros anos correspon-

dem ao ensino fundamental e os três últimos anos ao ensino médio, nível requerido para o ingresso no ensino universitário. Uma subamostra de 12 indivíduos foi definida para este estudo. Os indivíduos foram escolhidos aleatoriamente dentre os 32 indivíduos da amostra de maneira que fossem distribuídos em função das faixas etárias, dos níveis de escolaridade e sexo. Não foi possível uma boa distribuição entre homens e mulheres.

Os 12 indivíduos que compuseram a amostra para esse estudo estão distribuídos da seguinte maneira:

Tabela 1. Distribuição dos indivíduos pelas variáveis sociais estudadas

Escolaridade	1 a 4 anos		5 a 8 anos		9 a 11 anos		Total
	Homem	Mulher	Homem	Mulher	Homem	Mulher	
Idade							
15 – 25	–	1	2	–	1	–	4
26 – 49	1	1	–	–	1	1	4
50 ou mais	1	–	1	1	–	1	4
Total	4		4		4		12

Os dados foram submetidos à análise estatística do programa Rbrul (Johnson, 2009). O Rbrul é um programa elaborado com a finalidade de executar as mesmas tarefas que o Goldvarb, mas permite também considerar o papel do item lexical como uma variável de análise, uma vez que o programa inclui a análise de variáveis aninhadas (*random effect*), isto é, efeito de variáveis cujos elementos têm propriedades avaliadas por outras variáveis independentes (item lexical e tamanho do item, por exemplo). Além disso, a inclusão de elementos ou fatores de variáveis desta natureza não acrescenta informação ao modelo.

5. Resultados

Os dados foram submetidos à análise de regressão logística pelo programa Rbrul com o objetivo de avaliar a significância estatística das distribuições observadas para as variáveis independentes *contexto seguinte*, *vogal núcleo*

do ditongo, distância da sílaba tônica seguinte, tamanho do item lexical, status morfológico e para as variáveis aleatórias *item lexical* e *indivíduo* (*random effect*). O referido programa também seleciona as variáveis relevantes no condicionamento da variação. As variáveis sociais consideradas foram as relativas à estratificação da amostra: sexo, faixa etária e escolaridade. Com relação à variável *status morfológico*, devido ao envelope da variação adotado, nenhuma das variantes tem status de marcador morfológico. Aqui foi considerado que no caso de ocorrências de pretérito perfeito como *falaram* ~ *falaru*, a alternância é parte de um morfema, embora sua ausência não tenha consequências para a informação de número-pessoa. Essa variável tem respaldo no estudo de Guy & Boyd (1990) sobre a ausência de -t e -d em inglês (t/d deletion), em que se estabeleceu a variável status morfológico com uma categoria intermediária entre haver ou não status morfológico para essas consoantes, como em, respectivamente, *plant* (planta), em que a consoante é parte da raiz, e *worked*, em que a consoante indica forma de passado de *work*, trabalhar. A categoria intermediária correspondente à classe dos verbos irregulares, como *slept* (forma de passado de *sleep*, dormir) e *left* (forma de passado de *leave*, deixar, sair), em que -t e -d não são exatamente um morfema mas fazem parte de uma forma com informação de passado.

O valor de aplicação foi a realização da variante vogal oral. Foram registradas 143 ocorrências de vogal oral em um total de 322 dados, correspondendo a 45% dos dados.

Com relação às variáveis testadas, o programa selecionou a variável aleatória *item lexical*, e as variáveis independentes, na seguinte ordem, escolaridade ($p=0.00103$), idade ($p=0.00345$) e distância da sílaba tônica seguinte ($p=0.0354$). Os resultados das variáveis independentes estão apresentados na Tabela 2 a seguir.

Tabela 2. Variáveis selecionadas no condicionamento da vogal oral

Variáveis	Aplicação/N	%	Peso Relativo
Escolaridade			
1 a 4 anos (Ensino Fundamental I)	77/132	58	0,720
5 a 8 anos (Ensino Fundamental II)	43/87	50	0,620
9 a 11 anos (Ensino Médio)	27/103	26	0,163
Idade			
15 a 25 anos	56/129	43	0,336
26 a 49 anos	39/71	54	0,789
50 anos ou mais	52/122	42	0,491
Variáveis	Aplicação/N	%	Peso Relativo
Distância da sílaba tônica seguinte			
sílaba seguinte tônica	17/35	36	0,191
distância de 1 sílaba	57/137	40	0,497
distância de 2 sílabas	40/88	47	0,666
distância de 3 sílabas ou mais	33/62	55	0,683
Total	143/322	45	

Com relação às variáveis sociais, os resultados indicam que há uma tendência de ocorrência da variante oral entre falantes de escolaridade mais baixa. Já os resultados para faixa etária não são indicativos de mudança em progresso, mas sim de variação estável, já que não há diferença entre os mais velhos e o mais novos dessa escala na realização da variante oral.

O efeito da distância da sílaba seguinte indica um condicionamento que envolve uma posição de um constituinte silábico em posição prosódica fraca acentuada pela maior distância da sílaba tônica do vocábulo seguinte. Conforme pode ser observado na Tabela 2 quanto maior a distância maior a probabilidade de ocorrência da variante oral. Essa variável foi avaliada em Votre (1978) e os resultados obtidos neste trabalho replicam os obtidos em seu trabalho.

Com o objetivo de avaliar melhor o papel do item lexical, variável selecionada na primeira rodada no Rbrul, foi realizada outra rodada no Rbrul com as mesmas variáveis da primeira, contendo somente os itens que tiveram acima de três ocorrências no corpus levantado. Esse procedimento foi adotado uma vez que, se há somente duas ocorrências ou uma do item, não

é possível avaliar tendência das duas variantes em questão. Novamente as mesmas variáveis foram selecionadas. A Tabela 3 abaixo apresenta a lista dos itens que tiveram mais de três ocorrências no corpus com os respectivos percentuais e pesos relativos. Para facilitar a visualização os itens foram agrupados nas duas categorias de nome e verbo.

Tabela 3. Distribuição das ocorrências de vogal oral por item lexical

Item Lexical	Frequência	%	Peso Relativo
Passagem	8	87	0,872
Homem	68	73	0,862
Viagem	3	33	0,645
Jovem	25	12	0,178
Ontem	14	7	0,257
Jovens	3	0	–
Jardinagem	6	0	–
Compraram	3	100	0,865
Falaram	5	100	0,845
Fizeram	9	67	0,674
Deram	4	75	0,663
Conheceram	4	50	0,585
Pegaram	5	60	0,543
Tentaram	4	50	0,482
Começaram	19	36	0,472
Tiveram	4	0	0,401
Vieram	4	25	0,391
Deixaram	4	50	0,388
Ficaram	5	20	0,363
levaram	11	36	0,329
Botaram	8	25	0,286
Apertaram	3	0	0,223
Foram	16	6	0,207
Total			

Por se tratar de dados obtidos em amostra de fala espontânea, é esperado que alguns itens ocorram com mais frequência que outros. Com relação à variação, observa-se que alguns itens são mais propensos a ocorrer com a vogal oral, enquanto outros tendem a ocorrer com o ditongo nasal. A frequência dos itens no corpus não pode ser tomada como parâmetro para a avaliação do efeito de frequência, uma vez que se trata de um corpus coletado de 12 falantes. Foi então verificada a frequência de ocorrência de alguns dos itens da Tabela 3 no corpus de fala do corpus LAEL/PUC-SP (<http://corpusbrasileiro.pucsp.br>). Foram obtidos os seguintes valores: *homem*, 17.758 ocorrências; *jovens*, 8.519; *jovem*, 3.040; *jardinagem*, 19 ocorrências. A distribuição de alguns dos verbos quanto à frequência de ocorrência foi a seguinte: *fizeram*, 5.875; *falaram*, 467; *deram*, 2.324; *botaram*, 58 e *foram*, 58. O ideal seria fazer a contagem de frequência na Amostra Censo, mas esta não apresenta as condições técnicas para este tipo de contagem. As frequências obtidas foram tomadas como indicativas para a presente análise.

Observa-se na Tabela 3 que o item *homem*, o item mais frequente no corpus LAEL/PUC e no corpus estudado, foi majoritariamente produzido com a vogal final oral, e os itens *jardinagem*, *jovem* e *jovens* tenderam a ocorrer com o ditongo nasal. Entre as formas verbais, somente *compraram*, *falaram*, *fizeram* e *deram* tenderam à realização com vogal oral. A princípio não parece haver uma relação muito nítida entre frequência do item e tendência à realização da vogal oral, uma vez que a frequência de ocorrência de *homem* e *jovens* é mais próxima do que, por exemplo, *jardinagem*, um item de baixíssima frequência, e o comportamento dos dois primeiros itens é diferente. Essa distribuição, tomada apenas como indicativa, remete ao fato de que diversos aspectos do uso, além da frequência, podem concorrer para que um item lexical esteja mais propício a ser afetado por uma mudança sonora. A hipótese de Bybee, apresentada na seção 2, é a de que itens mais frequentes sejam mais atingidos que os itens menos frequentes.

Comparando-se o comportamento dos itens terminados em *-gem* da Tabela 3, observa-se que estes têm comportamentos distintos, com *passagem* favorecendo a variante oral e se diferenciando de *viagem* e *jardinagem*, portanto, não evidenciando um comportamento uniforme de uma categoria. É possível que o resultado obtido para esses nomes nos trabalhos de Battisti e Schwindt e Bopp da Silva, anteriormente mencionados, seja devido ao comportamento de alguns itens específicos com essa característica.

Outro aspecto importante diz respeito à variação dos itens nos indivíduos em que o item ocorreu. A Tabela 4 a seguir apresenta as ocorrências

da palavra *homem* por indivíduo. Observa-se um comportamento diferenciado por indivíduo. Para alguns a variante oral é a variante predominante (ou a única realizada) e em outros a variante ditongo nasal tem o mesmo tipo de ocorrência. Considerando que a entrevista sociolinguística é um recorte das possibilidades de produção do falante em um espectro estilístico amplo, o comportamento categórico ou semi-categórico observado em relação às duas variantes pode ser tomado como indicativo de uma forte tendência para a realização de uma das duas variantes para o falante.

Tabela 4. Ocorrência da vogal no total de realizações da palavra *homem* por indivíduo

Indivíduo	Apl/Total
M.	2/7
Mu	1/ 4
Fl	2/4
Pat	11/11
And	9/13
Ra	15/16
Adr	9/13
Total	49 /68

Do ponto de vista de uma interpretação baseada na Fonologia de Uso, considerando a representação em exemplares, todas as instâncias de realização das duas variantes, que fazem parte da experiência do falante em ouvir e produzir esses itens, estão disponíveis como parte da representação detalhada desses itens. A tendência majoritária de produção do item com a variante oral pode ser indicativa de uma representação central com a vogal oral.

Conforme mencionado anteriormente, o resultado para a distribuição por faixa etária é indicativo de variação estável. No entanto, o comportamento observado para alguns itens lexicais pode indicar uma mudança na direção da representação central com a vogal oral.

Os resultados obtidos neste estudo indicam o efeito prosódico no enfraquecimento de ditongos nasais que ocorrerem em sílaba átona final propiciando um espalhamento no léxico da variante com vogal oral reduzida,

o tipo mais frequente das sílabas átonas finais do português brasileiro, ao mesmo tempo em que apontam para uma implementação da variante oral por difusão lexical. A ampliação dos dados, com o levantamento dos demais falantes da Amostra Censo 2000, vai permitir a verificação dos resultados encontrados neste estudo piloto bem como irá proporcionar mais ocorrências dos itens já encontrados ou mesmo a ocorrência de novos itens.

6. Considerações Finais

Neste trabalho, foram apresentados os resultados de um estudo piloto sobre a alternância de ditongo nasal final átono com vogal oral, excluindo dados de alternância cujas ocorrências como vogal oral correspondem à ausência de marca de concordância verbal. Essa decisão metodológica se deveu ao fato de que esse tipo de alternância é condicionado por efeitos relacionados a um processo cuja definição envolve competição de efeitos específicos da alternância morfossintática.

Os resultados apontaram condicionamento prosódico bem como o efeito do item lexical. A alternância do ditongo nasal com vogal oral pode estar relacionada à posição fraca do ditongo em sílaba átona final, o que possibilita o espriamento da variante oral reduzida, o tipo mais frequente de sílaba átona no português brasileiro.

O comportamento diferenciado de itens lexicais em relação à frequência de ocorrência de uma das variantes implica problema para a adoção exclusivamente da variação como um processo, uma vez que aparentemente o processo tem frequência de aplicação diferente em função do item lexical. Conforme já observado em Gomes e Melo (2009) e Melo (2012), o efeito do item lexical indica que uma modelagem teórica que considera que as variantes fazem parte da representação das palavras no léxico acomoda melhor os achados. As representações detalhadas das palavras estão inseridas em um léxico dinâmico organizado em redes de relações lexicais estabelecidas em função de semelhanças sonoras, semânticas ou ambas (Bybee, 2001 e Pierrehumbert, 2012). Mais do que uma variável independente no estudo da variação sociolinguística, o papel do item lexical no condicionamento da variação subsidia a discussão sobre a natureza da variação na gramática. A adoção da hipótese representacional da variação sociolinguística tem sido discutida e considerada em diversos trabalhos (Pierrehumbert, 1994, 2002; Foulkes & Docherty, 2006). Esse estudo piloto fornece evidências para que se considere não só a importância do item lexical no estudo

da variação sonora, como também que se considere o comportamento do item lexical para verificar processos de mudança.

Os resultados aqui apresentados precisam, no entanto, ser checados em um conjunto maior de dados. Além disso, o estudo do item lexical pode utilizar outras formas de obtenção de dados para compensar a aleatoriedade das ocorrências em uma amostra de fala espontânea e possibilitar o controle de variáveis como frequência de ocorrência, familiaridade, especialização, entre outras.

Referências Bibliográficas

- ALBANO, Eleonora C. (1999) O português brasileiro e as controvérsias da fonética atual: pelo aperfeiçoamento da Fonologia Articulatória. *D.E.L.T.A.*, v. 15, n. especial, p. 23-50.
- BATTISTI, Elisa. (1997) *A nasalização no português brasileiro e a redução dos ditongos nasais átonos: uma abordagem baseada em restrições*. 187fls. Tese (Doutorado em Letras) – Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 1997.
- BATTISTI, Elisa. (2002) A redução dos ditongos nasais átonos. In: Bisol, Leda & Brescancini, Claudia. *Fonologia e variação: recortes do português brasileiro*. Porto Alegre: EDIPUCRS, p.183-202.
- BISOL, Leda. (1998) A nasalidade, um velho tema. *D.E.L.T.A.*, v. 14,n. especial, p. 27-46.
- BLOOMFIELD, Leonard. (1984) *Language*. Chicago: The University of Chicago Press.1a ed. 1933.
- BROWMAN, C. e L. GOLDSTEIN (1985) Dynamic modelling of phonetic structure. In: V. FROMKIN (org.) *Phonetic Linguistics*. Nova Iorque:Academic: 35-53.
- BYBEE, Joan.(2001) *Phonology and Language Use*. Cambridge: Cambridge University Press.
- BYBEE, Joan. (2002) Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change*, 14: 261-290.
- BYBEE, Joan. (2010) *Language, Usage and Cognition*. Cambridge: Cambridge University Press.
- CHEN, Matthews & WANG, William S.Y. (1977) Sound change: actuation and implementation. *Language* 51, p. 117-152.
- FOULKES, Paul, and DOCHERTY, Gerard (2006). The social life of phonetics and phonology. *Journal of Phonetics* 34, p.409–438.
- GOMES, Christina A. ; MELO, Marcelo A. S. L.(2009) . [pakerladu]: o papel do léxico e do contexto fonético no enfraquecimento da fricativa em coda. In: VI Congresso

- Internacional da ABRALIN, 2009, João Pessoa. VI Congresso Internacional da ABRALIN. João Pessoa: Editora Idéia, 2009. v. 1. p. 722-726.
- GOMES, Christina A. & SILVA, Thais C. (2004). Variação Linguística: antiga questão e novas perspectivas. *Linguagem*, v. 1, n.2, ILAPEC, Macapá, p. 31-41
- GUY, Gregory G. (1981) Linguistic variation in Brazilian Portuguese: aspects of the phonology, syntax, and language history. PhD Dissertation, University of Pennsylvania.
- GUY, Gregoriy G. & BOYD, Sally. (1990) *The development of a morphological class. Language Variation and Change*. n.2, p. 1-18.
- JONHSON, Daniel E. (2009) Getting off the GoldVarb Standard: Introducing Rbrul for Mixed-Effects Variable Rule Analysis. *Language and Linguistics Compass* 3/1, p. 359–383.
- LABOV, William (1981). Resolving the Neogrammarian Controversy. *Language* 57 (2), p. 267-308.
- LABOV, William (1994). *Principles of linguistics change: internal factors*, Cambridge, Blackwell, v. 1.
- MATEUS, Maria H. M.(1975) *Aspectos de fonologia portuguesa*. Lisboa: Instituto de Alta Cultura.
- MATTOSO CÂMARA JR., Joaquim. (1977) *Para o estudo da fonêmica portuguesa*. Rio de Janeiro: Padrão.
- MELO, Marcelo A. S. L. (2012) Desenvolvendo novos padrões na comunidade de fala: um estudo sobre as fricativas em coda no Rio de Janeiro. 2012. Dissertação de Mestrado, Universidade Federal do Rio de Janeiro.
- NARO, Anthony. J. (1981) The social and structural dimensions of syntactic change. *Language*, vol 57, p. 63-98.
- OLIVEIRA, Marco Antonio de. (1983) Resíduos Históricos Como Um Caso de Variação. *Ensaio de Linguística*, v. 9, n.1, p. 230-245, 1983.
- OLIVEIRA, Marco Antonio de (1991) . The Neogrammarian Controversy Revisited. *International Journal of the Sociology of Language*, v. 89, n.1, p. 93-105.
- PIERREHUMBERT, Janet. (1994). Knowledge of Variation, In *Papers from the Parasession on Variation, 30th meeting of the Chicago Linguistic Society*, Chicago Linguistic Society, Chicago, 25 pp.
- PIERREHUMBERT, Janet. (2003) Probabilistic Phonology: discrimination and robustness. In: BOD, Rens, HAY, Jennifer, JANNEDY, Stefanie (eds). *Probabilistic Linguistics*. Cambridge/Massachusetts, MIT Press, p. 177-228.
- PIERREHUMBERT, Janet. (2012) The Dynamic Lexicon. In: Cohn, Abigail, Huffman, Marie, & Fougeron, Céline (eds.) *Handbook of Laboratory Phonology*. Oxford University Press, 173-183.
- SCHERRE, Maria Marta P. & NARO, Anthony J. (1997) A concordância de número no português do Brasil: um caso típico de variação inerente. In: HORA, Dermeval da (org) *Diversidade lingüística no Brasil*, João Pessoa, Idéia, p. 93-114.

- SCHERRE, Maria Marta P. & NARO, Anthony J. (1998) Restrições sintáticas e semânticas na determinação do controle da concordância verbal de número em português. *Fórum Linguístico*. Universidade Federal de Santa Catarina, Centro de Comunicação e Expressão, Pós-graduação em Linguística, Florianópolis, Imprensa Universitária. Vol 1, p. 45-71.
- SCHWINDT, Luis Carlos; Bopp da SILVA, T. B. (2010) Panorama da redução da nasalidade em ditongos átonos finais no português do sul do Brasil. In: Bisol, Leda; Collischonn, Gisela. (Org.). *Português do sul do Brasil: variação fonológica*. Porto Alegre: EDIPUCRS, p. 13-33. (disponível em www.pucrs.br/edipucrs/portugues-dosuldobrasil.pdf).
- SCHWINDT, Luis Carlos; Bopp da SILVA, T. B., Quadros, E. S. (2012) O papel da morfologia na redução da nasalidade na redução de ditongos átonos finais no Português do Sul do Brasil. In: Lee, S. H. *Vogais além de Belo Horizonte*. FALE/UFMG. p.349-359.
- SILVA, Thais C., Fonseca, Marco S., Cantoni, Maria (2012) A redução do ditongo [ãw] pósônico na morfologia verbal do português brasileiro: uma abordagem baseada no uso. *Letras de Hoje* v. 47, n. 3, p. 283-292.
- SILVA, Thais C., Gomes, Christina A. (2007) Representações múltiplas e organização do componente linguístico. *Fórum Linguístico*, v. 4, n. 1, p.147-177.
- SOUSA, E. M. G. (1994). Para a caracterização fonético-acústica da nasalidade no português do Brasil, Dissert de mestrado, LAFAPE-IEL-UNICAMP
- VOTRE, Sebastião. (1978) Aspectos da variação fonológica na fala do Rio de Janeiro. Tese de Doutorado. Universidade Federal do Rio de Janeiro.
- WANG, William S.-Y. (1969). Competing changes as a cause of residue. *Language* 45, p. 9-25.
- WEINREICH, Uriel; LABOV, William; HERZOG, Marvin. (1968) Empirical Foundations for a theory of language change. In: LEHMAN, Winfred P; MALKIEL, Yakov. *Directions for historical linguistics: a symposium*, Austin-London, Univesity of Texas Press.
- WETZELS, W. Leo. (1997) The lexical representation of nasality in Brazilian Portuguese. *Probus*, v. 9, n. 2, p. 203-232.

MECANISMOS DE TEXTUALIZAÇÃO E CONSTRUÇÃO TEXTUAL: PARA UMA ABORDAGEM SÓCIO-DISCURSIVA DO *CARTOON*

TEXTUALIZATION AND TEXTUAL CONSTRUCTION MECHANISMS: TOWARDS A SOCIODISCURSIVE APPROACH OF *CARTOONS*

Audria Leal*

audria_leal@yahoo.com.br

Ana Caldes**

anacaldes@gmail.com

Os textos, como resultados das atividades humanas, possuem na sua organização parâmetros que dependem do contexto situacional, estrutura e regras do sistema linguístico. Estes parâmetros são dependentes das decisões do produtor dos textos, entre as quais figuram a escolha do género e, consequentemente, da sua arquitetura interna. A partir desta ideia, Bronckart (1999) desenvolve um modelo de análise de textos, procurando explicar as operações psicológicas que estão subjacentes à produção textual. Neste artigo, procuraremos mostrar, através da análise do género *cartoon*, o papel que os mecanismos de textualização desempenham na construção textual. Para isso, em primeiro lugar, iremos apresentar os pressupostos teóricos que fundamentam a nossa análise para, em seguida, analisarmos um modelo de *cartoon*, procurando refletir sobre a relação entre o plano de texto e os processos de coesão e conexão nele presentes.

Palavras-chave: género textual, plano de texto, processos de coesão e conexão

As they result from human activity, texts have in their own organisation some parameters which depend both on the situational context and on the structure and rules of the linguistic system. Those parameters are also dependent on the decisions taken by the author of the texts, among which we have to consider the choice of the

* Centro de Linguística da Universidade Nova de Lisboa, Lisboa, Portugal.

** Centro de Linguística da Universidade Nova de Lisboa, Lisboa, Portugal.

text genre and, therefore, of its internal architecture. Bearing this in mind, Bronckart (1999) developed a model of text analysis aiming at explaining the psychological operations which underpin textual production. In this paper, we are trying to show, through the analysis of the *text genre* cartoon, how the mechanisms of textualization work towards textual construction. To do so, and first, we will present the theoretical assumptions which are the core of our analysis. Second, we will examine a cartoon, trying to think about the relationship between the text plan and the cohesion and connection processes.

Keywords: text genre, text plan, cohesion and connection mechanisms

*

1. Introdução

No contexto deste artigo, assumimos que os textos – objetos empíricos que dão conta das práticas linguísticas efetivas – constituem produções sociais em estreita relação com a conjuntura espaciotemporal em que são produzidos. Assumindo ainda que esta conjuntura tem implicações a nível da organização e estrutura textuais, os textos revelam ser, por natureza, realidades complexas, em que a interação entre as dimensões semiótica e contextual aparece, em maior ou menor grau, de forma evidente.

Assim, a construção textual, não sendo gratuita, resulta antes da mobilização de recursos linguísticos pontuais, os quais estabelecem, com os parâmetros dessa conjuntura (de entre eles destacam-se, além do espaço e do tempo de produção, os objetivos pragmáticos que subjazem ao texto) uma relação inevitavelmente dialógica.

É, pois, nesta linha que nos propomos refletir – partindo da abordagem sociodiscursiva proposta por Bronckart – sobre o papel que os mecanismos de textualização exercem na construção textual, sendo que para tal tomaremos como objeto de análise um exemplar do género *cartoon*.

Começaremos por tecer algumas considerações relativas aos pressupostos teóricos de base do Interacionismo Sociodiscursivo (doravante ISD), no sentido de explicitar não só como o conceito de texto é aí tratado, mas também dar conta dos critérios que permitem categorizar os mecanismos de textualização como um dos níveis de organização da estrutura textual. De seguida, passaremos à análise dos textos selecionados, onde se pretende tanto proceder à aplicação das considerações desenvolvidas como, igualmente, verificar e atestar a sua rentabilidade em termos de reflexão linguística, tendo em conta o género textual em causa.

2. Pressupostos teóricos

No quadro do ISD, proposto por Jean-Paul Bronckart, os aspetos que norteiam a abordagem defendida pelo autor têm como base princípios teóricos que, como o próprio afirma, vão ao encontro de uma “psicologia da linguagem”, no contexto da qual se impõe considerar que “(...) qualquer que seja o tamanho das unidades linguísticas (do fonema ao texto), elas devem ser tomadas, em última instância, como condutas humanas (ou como propriedades das condutas humanas) ...” (Bronckart, 1996-1999: 13). Esta disciplina desenvolve, por sua vez, uma orientação epistemológica coerente com a do “interacionismo social”, corrente das ciências humanas e sociais na qual se inscreve e da qual se destacam autores como Vygotsky e Voloshinov.

Tomando como eixo central a interdependência entre a construção dos processos psicológicos e a dimensão coletiva e social da vida humana, o interacionismo social postula alguns princípios de base, com os quais se articula a abordagem proposta pelo ISD, e de que pretendemos aqui dar conta, de forma muito sintética:

a) um deles assenta na ideia que a compreensão do Homem só pode ser levada a cabo a partir de uma compreensão integral das várias dimensões, de vária ordem (biológica, social, cultural, cognitiva, etc.) que o constituem; nesta linha, o ser humano deve se entendido como um todo, uma realidade psicopsicológica;

b) outro princípio sublinha o papel decisivo das atividades sociais e coletivas na definição das condutas humanas enquanto “quadros organizacionais e mediadores” através dos quais os indivíduos interagem com o meio que os rodeia⁽¹⁾;

c) por outro lado, interessa referir a importância de que se reveste a linguagem verbal na gestão das atividades coletivas, na qualidade de instrumento necessário ao desenvolvimento de tais atividades⁽²⁾.

(1) Como refere Bronckart, no seguimento de Leontiev, «(...) ces activités constituent (...) les cadres organisant et médiatisant l'essentiel des rapports entre les individus et leur milieu: elles constituent en conséquence l'élément premier ou majeur de l'environnement humain». (cf. Bronckart, 2004: 114-115).

(2) Este princípio surge bem explícito em Bronckart (1996-1999: 21), quando o autor, ao apontar para uma definição do conceito de “interacionismo social”, refere o seguinte: «A expressão interacionismo social designa uma posição epistemológica geral, na qual podem ser reconhecidas diversas correntes da filosofia e das ciências humanas. Mesmo com a especificidade dos questionamentos disciplinares particulares e com as variantes de ênfase teórica ou de orientação metodológica, essas correntes têm em comum o fato de aderir à tese de que as propriedades específicas das condutas humanas são o resultado de um processo histórico de

Partindo de uma abordagem que Bronckart designa de descendente, isto é, orientada a partir das atividades sociais para as atividades de linguagem e, daí, para os textos e os seus recursos linguísticos⁽³⁾, o autor assume (no seguimento de Roulet *et a.l.*)⁽⁴⁾ a dimensão social como constitutiva dos factos linguísticos sendo que, nesta ótica, os textos são perspetivados na qualidade de produções sociais, onde se revelam, de forma mais ou menos explícita, as relações de interdependência entre as produções de linguagem e o seu contexto acional e social⁽⁵⁾.

No sentido em que se aceita a interação entre as atividades de linguagem e as atividades sociais, que aquelas comentam, a linguagem humana adquire, no âmbito do ISD, o estatuto de “produção interativa”⁽⁶⁾, à qual se pode reconhecer uma dupla valência: se, por um lado, a linguagem aparece como um instrumento semiótico possibilitador das relações entre os sujeitos e destes com o seu ambiente de ação, por outro, assume-se, ao nível mais complexo da questão, que ela possui um *papel fundador* das condutas humanas⁽⁷⁾. Refere o autor, a propósito: “(...) a linguagem não é (apenas) um meio de expressão de processos que seriam, eles, estritamente psicológicos (perceção, cognição, sentimentos, emoções), *mas (...) é, na realidade, um instrumento fundador e organizador desses mesmos processos, em todo o caso nas suas dimensões especificamente humanas*” (Bronckart, 2005: 39, sublinhado nosso).

A conceção de linguagem apresentada insere-se numa perspetiva que o autor designa de *logocêntrica* ainda que, como o próprio refere, se trate de um *logocentrismo relativo*. Do lugar central que se atribui à linguagem no desenvolvimento e funcionamento humanos, importa considerar que, nesse processo, são colocados em interface quatro “sistemas” distintos – os sistemas da língua, da atividade social, das operações psicológicas e o sistema textual/discursivo – os quais operacionalizam interações de ordem diversa, a seguir referidas:

a) Um primeiro tipo de interação, entre a atividade social e o sistema da língua:

socialização, possibilitado especialmente pela emergência e pelo desenvolvimento dos instrumentos semióticos».

(3) Cf. Bronckart (2005a: 61).

(4) Cf. Bronckart, 2005a: 61).

(5) Cf. Bronckart (1996-1999: 14).

(6) Cf. Bronckart (1996-1999: 34).

(7) Cf. Bronckart (2005b: 152).

Assumindo que a atividade linguística tem como função primária assegurar o funcionamento e a gestão das atividades sociais, interessa considerar que essa tarefa só é possível através da construção de instrumentos semióticos – os textos – os quais, atendendo aos recursos disponibilizados por uma dada língua natural, configuram essas atividades⁽⁸⁾.

A multiplicidade de esferas de atividade existentes, a que se associam interesses e objetivos próprios de cada uma, justifica a utilização de modalidades específicas de semiotização, definíveis em termos de *gêneros ou espécies de texto*. Aceitando que todo o texto participa de um gênero específico, ele mesmo regulado pelas propriedades da situação sociocomunicativa em que a atividade de linguagem tem lugar, citamos Bronckart que sublinha, a propósito, o seguinte: “Ce sont globalement les propriétés de l’ activité langagière dans son rapport à l’ activité ordinaire, qui déterminent les formes de réalisation effective de la langue que constituent les différents genres de textes” (Bronckart, 2005b: 154).

b) Um segundo tipo de interação, a operar entre o sistema textual/discursivo e o sistema da língua:

Tratando-se da estruturação, os textos obedecem a formas de composição e de planificação – ao nível da articulação das unidades textuais – responsáveis pela sua coesão e coerência. No entanto, no quadro da língua natural em que cada texto é produzido, essas regras são aplicadas em conformidade com os recursos disponíveis na língua em que o processo de semiotização ocorre. Por outro lado, o processo de produção textual implica a realização de escolhas relativamente à seleção e à organização das unidades composicionais que, integrando o texto, lhe permitem cumprir a sua função no âmbito da situação sociocomunicativa específica em que este se enquadra⁽⁹⁾. De entre as operações psicolinguísticas levadas a cabo neste processo (de que podemos destacar as operações de conexão e coesão de unidades linguísticas), interessa referir também a existência de uma outra modalidade de organização do texto, os “*tipos de discurso*”⁽¹⁰⁾, traduzidos como *configurações particulares de unidades e de estruturas linguísticas*⁽¹¹⁾ – mobilizadas em função dos recursos utilizáveis em uma dada língua – e entendidas como adequadas para o contexto enunciativo em que se inscrevem.

(8) Cf. Bronckart (2005b: 153).

(9) Cf. Bronckart (2005b: 154).

(10) O conceito, que será retomado num ponto mais avançado desta comunicação, será tratado mais adiante.

(11) Cf. Bronckart (2005a: 66).

c) Um terceiro tipo de interação, entre o sistema social e o sistema psicológico:

O cruzamento, por natureza inevitável, entre as representações individuais formuladas pelos sujeitos, a partir da sua experiência sobre o mundo, e as representações coletivas, criadas a partir dos preconstruídos sociais, dá lugar à construção de uma “plataforma” gnoseológica que coloca em *interface* os dois tipos de conhecimento que acabámos de mencionar. É, pois, sobre o diálogo entre as dimensões social e psicológica/individual das atividades humanas que tem lugar o desenvolvimento e a expressão do indivíduo enquanto ser comunicante, no contexto da conjuntura sociohistórica em que a sua atuação se inscreve; é também a partir desse diálogo que assenta, num movimento de (re)avaliação / (re)definição contínuas, o funcionamento das práticas linguísticas no seu estatuto de expressões da dimensão humana.

3. A arquitetura interna dos textos

Segundo Bronckart, a organização do texto assemelha-se a um *folhado* constituído por três níveis sobrepostos, a saber, a infraestrutura geral do texto, os mecanismos de textualização e os mecanismos enunciativos⁽¹²⁾.

Ao nível mais profundo, na infraestrutura, incluem-se as operações de planificação que permitem dar conta não só do plano geral do texto – responsável pela organização de conjunto do conteúdo temático – como também da constituição de sequências textuais, como os tipos de discurso, e as modalidades de articulação possíveis entre eles. Dotados de grande flexibilidade, os planos de texto podem possuir configurações variadas, justificáveis, por sua vez, em função do género em que o texto se inscreve, a natureza do conteúdo temático e os parâmetros da produção. No entanto, na perspetiva bronckartiana, é a combinação entre os diversos tipos de discurso, a par de outras formas de planificação, que desempenha um papel essencial na definição do plano de texto⁽¹³⁾.

Um segundo nível, intermédio, do *folhado textual*, integra o que se designou chamar mecanismos de textualização, sendo que é sobre estes que incidirá, de forma mais desenvolvida, o nosso trabalho. Os mecanismos de textualização têm como função assegurar a organização temática do

(12) Cf. Bronckart (1996-1999: 119-135; 2005a: 66-68).

(13) Cf. Bronckart (1996-1999: 249).

texto, marcando – através dos processos de conexão e de coesão (nominal e verbal)⁽¹⁴⁾ – o seu plano de conjunto.

O terceiro nível, o mais superficial de todos, devido ao facto de se encontrar pouco dependente da linearidade e da progressão temática do texto, é o dos mecanismos enunciativos. Estes mecanismos, encarregues de manter a coerência pragmática do texto, permitem o estabelecimento/distribuição das vozes e dos posicionamentos enunciativos, dando a ver não só as instâncias responsáveis pelo que aí se expressa, mas também a atitude dessas instâncias (traduzida em termos de julgamentos ou avaliações) relativamente ao(s) objeto(s) do conteúdo temático comentado no texto⁽¹⁵⁾. Nesta linha, é pela realização dos mecanismos enunciativos que se manifesta o *tipo de compromisso enunciativo* assumido pelo locutor do texto que, em todo o caso, e porque é sempre um compromisso construído na interação entre os interlocutores da enunciação (instância enunciativa e seu destinatário)⁽¹⁶⁾, dota o texto de coerência pragmática.

4. Os mecanismos de textualização

Como já aqui se referiu, cabe aos mecanismos de textualização assegurar, juntamente com a organização do plano geral, constituído ao nível da infraestrutura, a coerência temática do texto⁽¹⁷⁾. A progressão do conteúdo temático é, pois, conseguida – no que diz respeito a tais mecanismos – através da mobilização de processos isotópicos (assentes na constituição de cadeias de unidades linguísticas) visando a organização semiótica desse conteúdo⁽¹⁸⁾. Tendo em conta o papel que se atribui aos mecanismos de textualização no âmbito do modelo da arquitetura interna dos textos proposto por Bronckart, os mesmos são classificáveis em três tipos, já anteriormente apontados:

– *Mecanismos de conexão*, sendo estes realizados por organizadores textuais que marcam as articulações da progressão temática do texto, seja ao nível do plano geral do texto, da transição entre tipos de discurso ou

(14) Estes processos, apelidados de “isotópicos” possibilitam a construção de cadeias de unidades linguísticas que, por sua vez, contribuem como já se referiu, para a coerência temática do texto, pela retoma/progressão das unidades textuais que integram o espaço do texto.

(15) Cf. Bronckart (2005a: 68).

(16) Cf. Bronckart (1996-1999: 120).

(17) Cf. Bronckart (2005a: 67).

(18) Cf. Bronckart (1996-1999: 259).

sequências ou de articulações da sintaxe local do texto – têm como função, portanto, dar conta das relações que se estabelecem entre os diferentes níveis de organização do texto, segundo processos de segmentação, demarcação/balizamento e empacotamento de unidades⁽¹⁹⁾;

– *Mecanismos de coesão nominal*, a que se atribuem as funções quer de introdução, no cotexto, de novas unidades de significação que o autor designou de “unidades-fonte”, quer a retoma (ou não) dessas unidades no decurso do texto. Neste último caso, o processo de retoma é construído a partir de cadeias anafóricas capazes de assegurar a progressão do conteúdo temático apresentado⁽²⁰⁾;

– *Mecanismos de coesão verbal*, que garantem a organização temporal dos processos expressos no espaço textual e são essencialmente realizados pelos tempos verbais⁽²¹⁾, os quais possibilitam a marcação das relações de continuidade, descontinuidade ou oposição entre os objetos temáticos a que os mesmos se reportam⁽²²⁾.

5. Mecanismos de textualização e tipos de discurso

No trabalho de reflexão teórica sobre a natureza e o funcionamento das práticas linguísticas, desenvolvido por Bronckart no âmbito do ISD, os *tipos de discurso* revestem-se de uma importância fulcral na organização do texto ao assumirem-se como uma das suas dimensões constitutivas.

No sentido de explicitarmos o conceito de *tipos de discurso* diremos, por conseguinte, que ele remete para unidades linguísticas infraordenadas que, a título de “segmentos”, entram na composição textual e dão conta de modos distintos de enunciação. Estes modos, por sua vez, “(...) descrevem atitudes de locução gerais que se traduzem, no quadro de uma língua natural, por configurações de unidades e processos linguísticos relativamente estáveis⁽²³⁾ e passíveis de classificação. Na sequência desta possibilidade postula-se, assim, classificá-los em quatro tipos segundo, por um lado, um critério de conjunção ou de disjunção relativamente à articulação (ou não) entre as coordenadas de organização do conteúdo temático e as da situação de produção e, por outro, segundo um critério de implicação ou autonomia

(19) Cf. Bronckart (1996-1999: 264).

(20) Cf. Bronckart (1996-1999: 268; 2005a: 67).

(21) Cf. Bronckart (2005a: 67).

(22) Cf. Bronckart (1996-1999: 273).

(23) Cf. Bronckart (2005a: 69).

no que respeita à presença (ou não), no espaço textual, da relação expressa entre a instância responsável pela produção do texto e o próprio ato de produção deste⁽²⁴⁾. O quadro seguinte (figura 1) esquematiza o que se acaba de apontar:

Relação ao ato de produção	Coordenadas gerais		dos mundos
	Conjunção		Disjunção
	EXPOR		NARRAR
	Implicação	<i>Discurso interativo</i>	<i>Relato interativo</i>
Autonomia	<i>Discurso teórico</i>	<i>Narração</i>	

Figura 1. Os tipos de discurso (reproduzido de Bronckart, 1996-1999: 157)

Embora os mecanismos de textualização tenham como função marcar a organização de conjunto do texto, do qual fazem parte, a par de outros tipos possíveis de unidades textuais, os tipos de discurso, isto não significa necessariamente uma correspondência exata entre eles, sendo que o espaço de ocorrência das marcas linguísticas que dão conta de ambas pode não coincidir. A este respeito sublinha Bronckart que se, por um lado, os mecanismos de textualização conseguem atravessar ou transcender as fronteiras delimitadas pelos tipos de discurso, marcando precisamente a transição entre eles, por outro, as marcas de textualização⁽²⁵⁾ que realizam esses mecanismos (ou que realizam a mesma função) podem ser de ordem diversa, sendo suscetíveis de variar consoante o tipo de discurso em que se incluem⁽²⁶⁾.

A ocorrência das marcas de textualização que verbalizam cada um dos processos isotópicos de conexão e de coesão nominal/verbal pode depender, segundo o autor, dos tipos de discurso em que essa ocorrência se manifesta. Assim, é a especificidade enunciativa que caracteriza cada tipo em particular que parece ditar a ocorrência desta ou daquela marca no contexto da unidade linguística de que faz parte⁽²⁷⁾; no entanto, é de notar que se este é um facto que se verifica no funcionamento dos três tipos de mecanismos, como acabámos de referir, ele surge, no entanto, de forma mais explícita, ao

(24) Cf. Bronckart (2005a: 70; 2005b: 154).

(25) O conceito designa, segundo Bronckart, as diversas unidades linguísticas que concretizam, do ponto de vista linguístico, as funções realizadas pelos mecanismos de textualização (cf. Bronckart, 1996-1999: 260).

(26) Cf. Bronckart (1996-1999: 260).

(27) Cf. Bronckart (1996-1999: 267-268; 271).

nível do processo de coesão verbal, onde a influência dos tipos de discurso sobre as modalidades de realização desses processos se manifesta ainda em maior grau, relativamente aos outros dois mecanismos de textualização mencionados.

6. Análise textual

Antes de iniciarmos a análise que se segue, é importante salientar que consideramos o texto como produto da interação humana, em situação de comunicação, em que estarão em interação, para além dos aspetos linguísticos, fatores sociais, culturais e históricos. Desse modo, procuraremos fazer uma análise linguística do texto tomando-o como parte do complexo universo de ações humanas. Também é importante esclarecer que os géneros plurisemióticos fazem intervir aspetos não linguísticos que interagem com os níveis linguísticos. É o caso, por exemplo, de textos como os *cartoons*, em que os fatores não verbais são partes constitutivas desse género, sendo que isso nos leva a considerar, na sua análise, a relação entre o sistema linguístico e a imagem.

O primeiro texto para análise (figura 2) encontra-se impresso no jornal *Diário de Notícias*, do dia 8 de fevereiro de 2006. Trata-se de um exemplar do género *cartoon* publicado sob a rubrica “Cravo e Ferradura”, onde de forma recorrente se apresenta um leque de personagens, nos seus quotidianos, a comentarem determinados assuntos da realidade historicosocial coincidentes com a localização espaciotemporal de produção do texto. Debrucemo-nos, agora, sobre o primeiro *cartoon* seleccionado:

Texto 1:



Figura 2. *Cartoon* “Cravo e Ferradura”

a) A conexão e o plano geral do texto

Relativamente ao plano geral do texto, podemos apresentar várias unidades textuais que surgem bem delimitadas neste *cartoon* – o título da rubrica (“Cravo e Ferradura”) e a identificação do autor do texto (“Bandeira”). Além disso, e ocupando um papel central na organização da mancha gráfica do texto, podemos apontar uma distribuição de elementos verbais e icónicos semelhante àquela que é apresentada pela estrutura das histórias em “quadrinhos” ou vinhetas; este aspecto releva do facto de o género textual em causa ter como característica marcante e reveladora da sua identidade a presença da imagem e ser, por conseguinte, reconhecido como um género de natureza icónica ou icónico-verbal. O que aparece representado neste texto é, assim, uma ação comunicativa – em que a mudança de turno/de fala aparece marcada pela estrutura e pela ordem dos balões – fortemente vinculada ao contexto sociocultural em que a própria se inscreve.

Por outro lado, poderá dizer-se, ainda neste ponto, que a organização do plano de texto acompanha a organização do conteúdo temático, delimitando as diferentes fases da interação verbal e da situação que essa mesma

interação descreve e dá a conhecer: começa-se, pois, por indicar os responsáveis pelas acções descritas – “os caricaturistas” – sendo que este elemento, que funciona como ponto de partida para a leitura do texto, será posteriormente colocado em destaque através do percurso traçado no/pelo plano de texto. O esquema que se segue (figura 3) dá conta do processo que acabamos de referir:

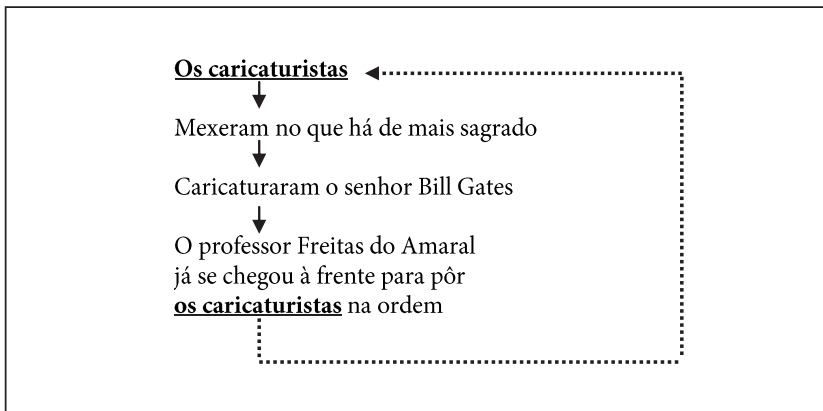


Figura 3. Esquema de análise textual

b) A coesão nominal

Como já foi dito, a coesão nominal é marcada por dois grupos de unidades que se relacionam ou, melhor dizendo, que “tecem” o texto: são os elementos com a função de introdução dos elementos novos, conhecidos por “unidades-fonte” e os elementos com função de retomada dessas “unidades-fonte”.

A primeira unidade-fonte observada no texto é a expressão “professor Freitas do Amaral”, sendo que esse elemento não voltará a ser retomado ao longo do texto.

Um segundo elemento que podemos identificar é a unidade-fonte “os caricaturistas”. Essa unidade será retomada no texto, mais adiante, pelo pronome “eles” (“Que foi que eles fizeram agora?”) e retomada elipticamente em “Ø mexeram” e “Ø parece que caricaturaram...”, dando a ver a ideia de que o elemento “os caricaturistas” será a principal unidade-fonte do texto, como veremos mais adiante. Este é ainda um caso da chamada anáfora pronominal.

A terceira unidade-fonte é também um dos pontos importantes do texto e uma das responsáveis pela criação do humor. O autor do *cartoon* estabelece, como uma pista para a construção do humor e da crítica presente no texto, a correlação entre os segmentos “... o que há de mais sagrado” e “... o senhor Bill Gates”. Estes dois elementos não ocorrem como um caso de anáfora, em que uma unidade é retomada pela outra, mas sim como um mecanismo de correferência que trabalha, neste texto, para o estabelecimento da coesão. Isso porque, segundo Campos e Xavier (1991), a anáfora existe quando pelo menos um dos termos tem autonomia referencial. Esse caso não ocorre na relação de Bill Gates com o sagrado. Ou seja, não existe previamente uma correferência entre esses dois termos. Assim, na relação entre estas duas unidades, só entendemos “o sagrado” relacionado com “Bill Gates” a partir da correferência construída pelo texto. Este facto, juntamente com outros elementos, nomeadamente, o deíctico “agora”, contribuem para a construção do conteúdo temático e do contexto sociohistórico semiotizados no texto.

c) A coesão verbal

Segundo o que temos vindo a explicitar, os mecanismos responsáveis pela coesão verbal estabelecem as organizações temporal e hierárquica dos processos que são essencialmente realizados pelos tempos verbais. Assim, em relação à temporalidade dos verbos, assumimos uma organização em que os verbos do passado possuem valor de anterioridade, os do presente de simultaneidade e os do futuro de posteridade. Essa temporalidade é marcada pela relação entre o momento da produção e o momento do processo expresso pelo verbo e, acrescenta Bronckart, também o que Reichenbach (1947) chama de “momento psicológico de referência”⁽²⁸⁾, este último sendo aplicado por Bronckart, com adaptações, à análise de texto. Não nos deteremos, no entanto sobre a aplicação deste ponto focando-nos no facto de Bronckart considerar, na análise da coesão verbal, quatro pontos essenciais na construção da coerência verbal: temporalidade primária e secundária e contraste global e local.

Para a análise da coesão verbal é essencial perceber o mundo discursivo que é semiotizado no texto. Ao analisarmos o tipo de discurso desse *cartoon* podemos observar, na primeira unidade textual, a presença de marcas linguísticas que apontam para o tipo de discurso “narração” (com verbos que estabelecem um valor de anterioridade) – são essas marcas o uso do pretérito perfeito (“se chegou”), tempo verbal característico da narração e a ausência de pronomes na primeira e na segunda pessoas do singular e do plural.

(28) Cf. Bronckart (1996-1999: 276).

Contudo, nas unidades seguintes, os parâmetros ligados ao conteúdo temático são interpretados à luz dos critérios de validade do mundo ordinário, o que aponta para o discurso interativo: a ausência de uma origem espaciotemporal explícita, a presença da frase não-declarativa “Que foi que eles fizeram agora?” e a própria imagem que apresenta os balões marcando uma mudança de turno de fala dentro da interação.

Sabemos que o discurso interativo apresenta como tempo-base o presente e o pretérito perfeito, com valor de simultaneidade para o presente e com valor de passado próximo para o pretérito perfeito simples. Desse modo, é essencial conhecer a duração do ato de produção quando falamos de mecanismos de coesão verbal. No entanto, isso não impede a construção de eixos de referência temporal distintos dessa duração de produção. É o caso desse exemplo: ao combinarmos os dois elementos – tipo de discurso e coesão verbal – observamos que a localização do processo visível no tempo verbal aponta para a construção de um valor anterior muito próximo do tempo em que ocorre a interação. Podemos mesmo falar numa localização de anterioridade (pretérito), mas incluída entre dois limites de duração, marcando uma localização de simultaneidade. Esta questão é reforçada quando analisamos o deítico “agora”.

Este deítico possui, neste texto, um papel central. Semanticamente, o “agora” expressa um corte que divide o tempo em dois instantes que se opõem⁽²⁹⁾: um “antes” e um “depois”. Neste caso, o “agora” faz referência a um acontecimento ocorrido no passado (fazer caricaturas de uma personalidade sagrada, neste caso, Maomé), mas que também ocorre no presente (fazer caricatura de personalidade “sagrada” no presente, neste caso, Bill Gates) e que co-ocorre no presente. Essa ocorrência, que será localizada à esquerda em relação ao tempo T_0 , refere-se ao facto de os caricaturistas terem elaborado *cartoons* sobre Maomé ocasionando, assim, uma crise religiosa que se traduziu numa crise política. Desse modo, num primeiro momento, o “agora” orienta para a inferência sobre a crise que envolveu os caricaturistas e os muçulmanos (ocorrência do passado).

No entanto, como referimos anteriormente, esse “agora” não só apresenta o que se passou, como também faz referência a acontecimentos que co-ocorrem no presente. Assim, e num segundo momento, o “agora” apresenta um significado do tipo “de novo”, “novamente”, “outra vez”, sendo que esta ideia é corroborada na unidade posterior com a expressão “Mexeram no que há de mais sagrado”. Podemos então concluir que o “agora” não só possibilita que haja uma inferência sobre a crise anterior causada pelos cari-

(29) Cf. Sousa (2000).

caturistas, como apresenta a ideia de que essa ação volta a ocorrer no presente, caracterizando-se, assim, por ser um intervalo fechado à direita sem espessura no tempo futuro; ao fazer a oscilação entre os tempos passado e presente, o deíctico permite, em consequência, fazer o corte e a retoma dos acontecimentos descritos em relação ao momento de produção do texto. Esse ponto é essencial para percebermos o processo de coesão verbal desse texto e a sua relação com a construção do mundo discursivo “expor”, marcando, no texto, uma relação entre os processos de coesão verbal e a criação de mundos discursivos.

O segundo *cartoon* analisado (figura 4) foi impresso no jornal Público, do dia 04 de fevereiro de 2006. O autor do *cartoon* é o Luís Afonso. Este mesmo autor é responsável por duas séries: a “Sociedade Recreativa”, publicada na revista *Pública* que acompanha o jornal *Público* aos fins de semana, e a série “Bartoon” que é publicada diariamente. Este *cartoon* em particular foi publicado sem ser associado a série em questão e evoca o mesmo contexto social do *cartoon* anterior: a visita de Bill Gates a Portugal, acrescida do contexto das eleições presidenciais. Vejamos de seguida a análise deste segundo *cartoon*.

Texto 2:

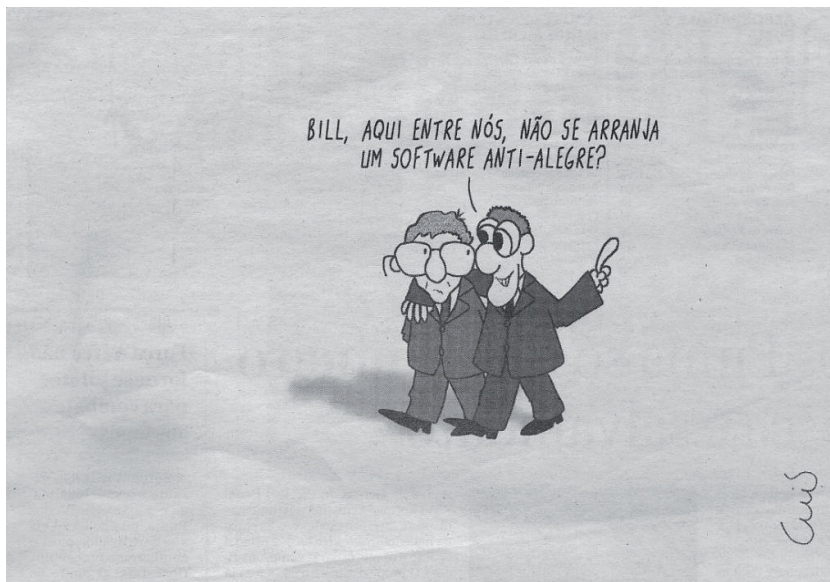


Figura 4: *Cartoon* de Luís Afonso

a) A conexão e o plano geral do texto

Em relação ao plano geral do texto, diferente do exemplo anterior, este *cartoon* apresenta-se centralizado, não havendo, de fato, cenário envolvente. As unidades textuais, apresentam-se, assim, de forma conjunta, com exceção da assinatura do autor, que se encontra no canto direito do texto. Como já referido ainda no exemplo anterior, o destaque dado à mancha gráfica e sua participação na organização temática do género reforça o seu carácter plurissemiótico. Neste exemplo, apesar de não haver mudança nos turnos da fala marcados por balões, podemos depreender, a partir da imagem, que se tratar de uma encenação de diálogo entre as personagens Bill Gates e Sócrates, este último então primeiro-ministro de Portugal. Portanto, a interação entre os personagens do texto é fortemente centralizada numa única cena, não havendo propriamente um percurso de leitura quadro-a-quadro.

b) A coesão nominal

Como já especificado, a coesão nominal, segundo Bronckart (1999), é marcada linguisticamente por um elemento que introduz uma unidade-fonte e os elementos que retomam esta unidade-fonte. Contudo, em géneros plurissemióticos, é essencial considerar que não só os elementos verbais são considerados referentes, como também os não verbais podem assumir este papel. De fato, neste texto, só é possível uma compreensão do funcionamento textual se entendermos que a própria imagem fornece informações sobre a identificação de unidades-fontes relevantes. Na parte linguística, a primeira unidade-fonte identificada é “Bill”, referindo-se a Bill Gates, e que será, de certa forma, retomado quando o personagem utiliza o “nós”. Este caso é interessante, pois, este “nós” inserido na expressão “aqui entre nós” retoma tanto o elemento verbal “Bill” e a imagem do Bill Gates como a imagem que representa o primeiro-ministro “Sócrates”, além de que mostra uma intimidade entre os dois personagens corroborada pela imagem na qual estão abraçados. Portanto, a unidade de referência, “nós”, remete, sem dúvida, para as imagens do texto, ocorrendo uma interação direta entre o verbal e o não verbal. Além disso, o próprio termo *software* pode ser correlacionado com “Bill”, referindo-se ao fato de este ser o criador da *Microsoft*. Podemos então considerar o termo “software” procede a uma retomada da unidade-fonte “Bill”, uma vez que esta relação é essencial para a compreensão do humor do *cartoon* e, conseqüentemente, para a funcionalidade do género textual. Um segundo elemento-fonte que podemos identificar é o

termo “Alegre” em “Anti-Alegre”. Esta unidade-fonte não será retomada no texto, mas, sem dúvida, introduz um elemento novo, “Manuel Alegre”, então candidato à presidência da República pelo mesmo partido do Sócrates. O “anti” que acompanha o nome “Alegre” sugere uma interpretação em que, apesar de Alegre ser candidato pelo mesmo partido do Sócrates (o Partido Socialista – PS), o primeiro-ministro não apoiou a candidatura de Manuel Alegre, e por isso, no texto, solicita a Bill um *software* “Anti-Alegre”. Portanto, a identificação de Anti-Alegre como uma unidade-fonte é imprescindível para a compreensão do texto, pois retomará o contexto político que põe em foco a desavença, dentro do PS, entre Sócrates e Manuel Alegre. Portanto, as correferências construídas neste texto são responsáveis pela construção do humor e da crítica.

c) A coesão verbal

Como já salientámos, a análise da coesão verbal passa necessariamente pela construção do mundo discursivo que é semiotizado no texto. De fato, temos vindo a afirmar que também a imagem participa na construção dos parâmetros ligados ao conteúdo temático, dando informações sobre o tempo da produção, o papel social dos que participam na interação, bem como estabelece referências temáticas. Assim, a própria imagem participa na construção do mundo discursivo.

Neste segundo exemplo, as unidades linguísticas remetem para parâmetros ligados ao conteúdo temático que são interpretados à luz dos critérios de validade do mundo ordinário, o que aponta para o discurso interativo. Assim, a presença da frase interrogativa, “Bill, aqui entre nós, não se arranja um software Anti-Alegre?”, indica a ausência de uma origem espaciotemporal explícita. Além disso, a própria imagem, que apresenta uma interação entre os personagens, mostra a encenação como sendo construída no mesmo instante temporal da enunciação. Estes dois fatores são corroborados ainda pelo uso do verbo no presente do indicativo “não se arranja”, o que indica uma localização de simultaneidade em relação ao T_0 . Estes fatores indiciam para a conjunção entre a temporalidade explícita verbalmente no texto e o próprio tempo da situação de enunciação. Além disso, a própria imagem reforça esta conjunção, *inclusive* com a implicação dos agentes-produtores (Bill Gates e Sócrates) na cena construída. Isto indicia a presença do mundo discursivo “expor implicado”.

Outro ponto interessante é a presença do deítico espacial “aqui” que, não sendo propriamente um marcador temporal, reforça a simultanei-

dade da ação de linguagem construída no texto. No entanto, é de destacar que o “aqui” também é considerado como parte da expressão “aqui entre nós”, o que reforça o grau de intimidade, que é estabelecida pelo autor Luís Afonso, entre Bill Gates e Sócrates, o que, a nosso ver, reforça a implicação dos agentes-produtores na cena enunciativa. Todos estes fatores convergem para a criação do “expor implicado” com a presença do discurso interativo, pois, de fato, o discurso interativo apresenta como tempo-base o presente além da implicação dos agentes-produtores do discurso. Assim, é essencial analisar, não só o verbal, como também a interação entre este verbal com o não verbal para percebermos a relação entre os processos de coesão verbal e a criação de mundos discursivos.

7. Conclusões

Levando em consideração o papel decisivo das práticas linguísticas na concretização das relações com o meio, são os textos que, na qualidade de objetos concretos e atestáveis, nos dão conta da natureza dessas relações. Referindo-se a eles como os *correspondentes empíricos das atividades linguísticas*, realizados segundo os recursos disponíveis numa dada língua natural, Bronckart classifica os textos de *unidades comunicativas globais* cuja especificidade composicional depende, em parte, das características da situação de interação e da conjuntura sociohistórica em que se inscrevem, em parte, das atividades sociais que comentam e no contexto das quais são gerados⁽³⁰⁾.

Face ao exposto, e a partir da análise dos textos, concluímos que os mecanismos de textualização não só participam na organização temática do texto, como também têm um importante papel na construção das funções comunicativas do género. Nesse caso, verificamos que a coesão textual é peça-chave para construir o humor. A análise desses textos mostraram que a coesão, geralmente pensada em termos de relações anafóricas, é aqui estabelecida pela correferência, que por sua vez encaminha para o estabelecimento do humor. Além disso, atestamos a existência entre a relação “tipo de discurso” e “coesão verbal”, sendo este facto também essencial para o entendimento do texto e, conseqüentemente, para a construção do humor. Outro ponto interessante é que a própria imagem participa na construção das correferências do texto. Neste género, a interação do verbal com o não verbal é a responsável pelos processos interpretativos do leitor e, com isso,

(30) Cf. Bronckart (2004: 115).

fundamental para a apreensão do humor e da crítica. Esta última aceção, apesar de não enfatizada pelo interacionismo sociodiscursivo, configura-se como um ponto relevante na contribuição dos estudos do funcionamento da linguagem.

Deixamos aqui, através desta reflexão, portas para serem abertas e caminhos a serem seguidos. De facto, o estudo da coesão textual está longe de se esgotar e, por isso, deve ter um lugar importante na análise linguística, podendo ser esclarecedor para o entendimento de questões que envolvem os mecanismos de funcionamento das línguas.

8. Referências

- BRONCKART, Jean-Paul (2006) *Atividade de linguagem, discurso e desenvolvimento humano*, Anna Rachel Machado e Maria de Lourdes Meirelles (orgs.), Campinas: Mercado de Letras.
- , (2005a) Os géneros de texto e os tipos de discurso como formatos das interações de desenvolvimento. In Fernanda Menéndez (org.) *Análise do Discurso*, Lisboa: CLUNL/Hugin Editores, pp. 38-79.
- , (2005b) Les différentes facettes de l'interactionnisme sócio-discursif. *Calidoscópico*, vol.3, nº 3, Set. /Dez. 2005, pp. 149-159.
- , (2004) Commentaires conclusifs. Pour un développement collectif de l'interactionnisme socio-discursif. *Calidoscópico*, vol. 2, nº 2., Jul. / Dez., pp. 113-123.
- , (1996-1999) *Atividade de Linguagem, textos e discursos. Por um interacionismo sócio-discursivo* São Paulo: Editora da PUC-SP / EDUC (trad. port. de *Activité langagière, textes et discours. Pour une interactionnisme sócio-discursif*, Lausanne: Delachaux et Niestlé). São Paulo: Editora da PUC-SP / EDUC.
- CAMPOS, Maria H. & Xavier, Maria F. (1991) *Sintaxe e semântica do português*. Lisboa: Universidade Aberta.
- COUTINHO, Maria Antónia (2007), "Organizadores textuais", disponível em <http://www.fcsh.unl.pt/cadeiras/texto/Organizadores%20textuais.pdf>
- , (2005) Para uma linguística dos géneros de texto. *Diacrítica* 19/1, Braga: Universidade do Minho, pp. 73-88.
- , (2003) *Texto(s) e competência textual*, Lisboa: FCG-FCT.
- COUTINHO, M. A.; Leal, A.; Pinto, R.; Teixeira, C.; & Caldes, A. (2009) La dynamicité de la langue dans des textes de différents genres. In *Texto!*. Disponível em: <http://www.revue-texto.net/index.php?id=2103>.
- DOLZ, Joaquín e Schneuwly, Bernard (1996) Genres et progression en expression orale et écrite. Éléments de réflexion à propos d'une expérience romande. *Enjeux* (revue de Didactique du Français), pp. 37-38, 48-74.

- DUARTE, Inês (2003) Aspectos linguísticos da organização textual. In Maria H. Mateus *et alii* (orgs.) *Gramática da Língua Portuguesa*, 5ª ed., Lisboa: Editorial Caminho, pp 85-123.
- KOCH, Ingedore (2001) *Desvendando os segredos do texto*, São Paulo: Cortez.
- , (1998a) *A coerência textual*, 10ª ed., São Paulo: Contexto.
- , (1998b) *A coesão textual*, 10ª ed., São Paulo: Contexto.
- SCHNEUWLY, Bernard (1994) Genres et types de discours: considérations psychologiques et ontogénétiques. In Y. Reuter (ed.) *Les interactions lecture – écriture*, Berne: Peter Lang, pp. 155-174.
- SOUSA, Otilia da C. e (2000) Linguística, Filosofia da Linguagem e Operações Cognitivas: a propósito da noção de presente. In *Cadernos de Filosofia*, nº 7, Lisboa: Edições Colibri, pp. 113-122.

GRAMATICALIZAÇÃO E ESPECIALIZAÇÃO FUNCIONAL: O CASO DO CONECTOR *POIS*

GRAMMATICALIZATION AND FUNCTIONAL SPECIALIZATION: THE CASE OF THE CONNECTOR *POIS*

Maria da Conceição de Paiva
paiva@club-internet.fr*

Maria Luiza Braga
malubraga@terra.com.br**

Nosso objetivo neste artigo é discutir a hipótese de que a gramaticalização de orações obedece a um *cline* de [-integradas] > [+integradas], ou seja parataxe > hipotaxe, com base em um estudo diacrônico do conector explicativo *pois*. Mostramos que, embora no português contemporâneo, as orações introduzidas por *pois* só admitam a posposição, até o século XVIII, elas podiam ser antepostas, interpostas ou pospostas à oração núcleo. A perda desta flexibilidade, que, aparentemente, viola a trajetória prevista por princípios mais gerais, é discutida à luz de características discursivas das orações introduzidas por *pois*. A análise destas propriedades conduz à conclusão de que este conector sofre uma especialização funcional. Em estágios anteriores da língua, ele podia introduzir tanto orações explicativas que codificavam informação nova como orações que introduziam informação já compartilhada pelos interlocutores. Ao longo do tempo, *pois* se especializa na introdução de informação nova.

Palavras chave: Conector *pois*, ordenação, gramaticalização

Our aim in this article is to discuss the hypothesis that the grammaticalization of clauses follows the cline [-integrated] > [+integrated], i.e., parataxis > hypotaxis, based on a diachronic study of Portuguese explicative connector *pois*. We show

* Professora do Programa de Pós-graduação em Linguística, Faculdade de Letras, Universidade Federal do Rio de Janeiro e pesquisadora do CNPq, Rio de Janeiro, Brasil.

** Professora titular do Programa de pós-graduação em Linguística da Faculdade de Letras, Universidade Federal do Rio de Janeiro e pesquisadora do CNPq, Rio de Janeiro, Brasil.

that, although the clauses introduced by *pois* can only be postposed in contemporary Portuguese, up to the 18th century, they were allowed to occur before or after their nuclear clauses. The loss of that flexibility, which apparently violates the path assumed by more general principles, is discussed considering the discursive characteristics of the clauses introduced by the connector *pois*. The analysis of those properties leads us to conclude that the observed changes are the result of a functional specialization of *pois*. In former stages of the language, this connector could introduce clauses that codify both new and given information. Through the time, however, *pois* specializes in introducing new information.

Key words: connector *pois*, position, grammaticalization

*

1. Introdução

Os estudos sobre a origem e evolução de conectores e de períodos complexos vêm acumulando evidências que permitem aferir a validade de princípios relacionados aos processos de gramaticalização. Dentre esses pressupostos, destaca-se o da unidirecionalidade que, aplicado à articulação de orações, prevê um *cline* de evolução na forma parataxe > hipotaxe. Em outros termos, estruturas mais dependentes e mais integradas derivam de empregos anteriores com menor grau de dependência e de integração. Evidências de estudos mais recentes permitem problematizar esta hipótese, atestando múltiplos casos de evolução de estruturas mais integradas para estruturas menos integradas.

Neste artigo, buscamos fomentar esta discussão, a partir de uma análise diacrônica das orações introduzidas pelo conector *pois* no português, mais particularmente no que se refere às suas restrições de posição. A partir de dados atestados nos diferentes períodos desta língua, sustentamos que a trajetória de *pois* pode ser descrita como um caso de especialização funcional que envolve, principalmente, a natureza da informação codificada pelas orações que ele inicia. Mostramos que a perda de flexibilidade das orações encabeçadas por *pois*, ao longo da história do português, reflete a perda de uma possibilidade discursiva, qual seja, a de codificar orações com informação velha, o que conduz à especialização de *pois* como introdutor de orações com informação nova e, necessariamente, pospostas à oração núcleo.

Os dados analisados foram coligidos em diferentes textos do século XIII ao século XX. Até o século XVIII, o *corpus* é constituído por textos de

gêneros diferenciados.⁽¹⁾ Para os séculos XVIII e XIX, a amostra é constituída do gênero cartas e inclui subgêneros diferenciados como cartas pessoais, administrativas e de negócios, cartas de leitores e redatores de jornais de diferentes cidades brasileiras. Finalmente, a amostra para o século XX, compreende cartas de leitores publicadas em quatro jornais de grande circulação na cidade do Rio de Janeiro.⁽²⁾

O artigo está organizado da seguinte forma: na seção 2, caracterizamos o conector *pois* no português contemporâneo e destacamos algumas particularidades sintáticas e semânticas das orações que eles introduzem. Na seção 3, retomamos pontos centrais acerca dos processos de combinação de orações e da gramaticalização de cláusulas, enfatizando a necessidade de concebê-los numa perspectiva não dicotômica que considere um *continuum* de integração. Na seção 4, analisamos a evolução das orações encabeçadas por *pois* no que tange à sua posição no período e examinamos a correlação entre ordenação e domínio da relação de causalidade. Na seção 5, apresentamos argumentos para uma interpretação das mudanças depreendidas na posição das orações com *pois* em termos de especialização funcional deste conector. Seguem-se as considerações finais.

2. O conector *pois* no português contemporâneo

Nas descrições gramaticais do português contemporâneo, *pois* é incluído, mais frequentemente, no conjunto das conjunções coordenativas, tanto na sua função de introdutor de orações explicativas como de orações conclusivas. (Cegala, 1970; Cunha, 1970). Os valores semânticos instanciados nas orações com *pois* remontam às origens do português, como já desta-

(1) Para os períodos arcaico e clássico, foram utilizados os seguintes textos: século XIII: Tempos dos preitos (TP). Textos notariais (TN), Afonso X, Foro Real (FR); século XIV: Textos notariais em Clíticos da história do português (CHP), Crônica geral de Espanha (CGE); século XV: Crônicas do Conde D. Pedro de Menezes (DPM) e Orto do esposo (OE); para o século XVII: Crônicas do rei de Bisnaga (CRB), Cartas de Dom João III (CD); século XVII: Epanáforas de variada língua portuguesa (ELP) e Cartas familiares de F. M. Melo (FMM) As versões utilizadas correspondem às que foram organizadas para o *Corpus* Informatizado do Português Medieval (CIPM) (disponível no endereço : <http://cipm.fcsh.unl.pt/>) e as que constituem o Corpo Tycho Brahe, da USP (disponível em . <http://www.tycho.iel.unicamp.br/~tycho/>).

(2) Estas cartas integram um corpus mais amplo (Amostra do Discurso Midiático), composta de textos representativos de diferentes gêneros jornalísticos publicados nos jornais cariocas *O Globo*, *Jornal do Brasil*, *Extra* e *Povo*. Esta amostra foi organizada pelos membros do Programa de Estudos do Uso da Língua (PEUL), sediado na UFRJ e estão disponíveis em www.letras.ufrj.br/~peul

caram, por exemplo, Barreto (1999), Lima, (2002), Paiva e Braga (2013) e derivaram de empregos anteriores deste elemento como advérbio temporal. Segundo interpretação mais corrente, *pois* deriva da partícula latina *post* (*post* > *pos* > *pois*), que podia funcionar como advérbio ou preposição, com valores locativos e temporais (Said Ali, 2001 [1921]; Mattos e Silva, 1989, 2006; Barreto, 1999; Lima, 2002). A forma fonte *post* teria dado origem, não apenas ao conector *pois*, como também à locução conjuntiva causal/explicativa *pois que* e as temporais *depois de*, *depois que* (cf. Nunes, 1975; 1962). Para Nascentes (1955), todavia, a evolução de *pois* como conector envolveria uma mistura dos contextos de *post e* da forma *postea*, já no baixo Latim. Segundo Barreto (1999), adotando a posição de Corominas e Pascual (1991), a forma *post* assimilou o valor semântico da locução conjuntiva latina *postquam*, sinalizadora de posterioridade de um estado de coisas em relação a outro, mas capaz de introduzir também uma explicação. Assim, a evolução de *pois* exemplificaria uma trajetória de gramaticalização advérbio > preposição > conjunção, através de sucessivas etapas de recategorização que, envolvem, dentre outras mudanças, a fixação do elemento no início da oração.

No português arcaico, coexistia com *pois* explicativo e conclusivo, o emprego deste elemento como conector temporal (Olinda, 1991; Barreto, 1999; Lima, 2002; Braga & Paiva, 2013), o que indica sua polissemia desde períodos mais remotos. A passagem de conector temporal para conector causal/explicativo é interpretada por Lima (2002) como uma consequência natural da interdependência entre as noções de tempo e causa, ou seja: se A precede B, A pode ser entendido como a causa de B (cf. Paiva 1991, 1996), um processo de mudança semântica sustentado em evidências translinguísticas (Traugott & König 1991; Traugott 1995, 2010).

Como já destacado por diversos autores (Olinda, op. cit.; Barreto, op. cit.; Paiva & Braga, op. cit.), nos séculos XIII e XIV, *pois* podia alternar com *pois que/poys que* na introdução tanto de orações causais/explicativas e conclusivas como de orações temporais. Segundo Barreto, já no século XVI, o uso dos dois conectores com valor temporal é suplantado pelo das locuções conjuntivas *depois de/depois que*.

Embora não constitua foco deste artigo, vale mencionar que a multifuncionalidade de *pois* é mais ampla, incluindo usos deste elemento como continuativo, marcador discursivo (reforço de afirmação), ou partícula fática para Lima (2002), que, ao que tudo indica, lançam raízes no português antigo (Barreto, op. cit.; Cândido 2009). É plausível que estes diferentes usos de *pois* sejam o ponto extremo de um *continuum* espaço > tempo >

texto, produtivo no desenvolvimento de elementos de conexão interoracional (cf. Barreto, *op. cit.*; Cândido, *op.cit.*).

No português contemporâneo, o conector *pois* se particulariza tanto nas suas propriedades sintáticas como semântico-discursivas. Diferentemente de *porque*, elemento mais produtivo e versátil, capaz de instanciar relações causais nos domínios referencial, epistêmico e conversacional (cf. Sweetser, 1990; Paiva, 1996; Dancyeger & Sweetser, 2000, Braga & Paiva 2010), o uso de *pois* é mais restrito, encabeçando, mais frequentemente, orações que apresentam uma evidência para uma conclusão (domínio epistêmico) ou uma justificativa para um ato de fala (nível conversacional). Em outros termos, orações com *pois* realizam uma causa da enunciação (Lobo 2003, 1999) ou nos termos de Lopes (2004), uma causa explicativa. Esta particularidade demonstra bem que *pois* ocupa apenas parte do espaço conceitual da causalidade, o que se evidencia na impossibilidade de parafrasear qualquer ocorrência de *porque* por *pois*.⁽³⁾

À maioria das análises das orações introduzidas por *pois* subjaz um pressuposto de simetria entre posição e significado. Assim, a interpretação de *pois* como conector explicativo decorre, em grande parte, da sua posição no início da oração, enquanto a interpretação conclusiva emerge mais naturalmente com a sua posposição ao verbo. Considerando que essas duas relações se imbricam, por constituírem duas formas distintas de perspectivização do mesmo raciocínio inferencial, (*Pedro não está em casa pois as luzes estão apagadas/As luzes estão apagadas, Pedro não está, pois, em casa*), pode-se presumir que elas são originadas por uma mesma cadeia de mudanças semânticas.

Do ponto de vista sintático, a inclusão das orações explicativas encabeçadas por *pois* no rol das coordenadas não é tão consensual, visto que, como destaca Lobo (2003:54), as estruturas com *pois* “possuem um estatuto pouco claro a meio caminho entre a coordenação e a subordinação”. Por um lado, as orações com *pois* compartilham diversas propriedades com outros tipos de orações coordenadas: não admitem mudança de posição, não se incluem no escopo de uma negação, de uma partícula focalizadora ou de um elemento modalizador, não podem ser objeto de interrogação, não admitem operação de clivagem e não podem ser encadeadas por uma conjunção coordenativa ou constituir argumentos de um predicado (cf.

(3) A possibilidade de substituir *pois* por *porque* decorre naturalmente da ambiguidade deste conector que pode instanciar relações causais em diferentes domínios (cf. Paiva 1991, Lopes 2004, Paiva & Braga, 2010).

Peres, 1997; Lobo, 2003; Matos, 2005; Lopes, 2004; Peres, 1997; Braga & Paiva, 2011).

Como o conector *pois* não partilha todas as propriedades das conjunções coordenativas, alguns autores o excluem deste conjunto (Quirk et al 1985, Matos 2005). Matos (2003), por exemplo, convoca particularidades como o fato de *pois* não poder coordenar sintagmas abaixo do nível da oração ou ligar mais de dois constituintes oracionais para classificá-lo como subordinativo. Lobo (2003) e Lopes (2004), ao contrário, o catalogam como elemento de coordenação. Para Lopes (2004), as particularidades no comportamento de *pois* não chegam a comprometer sua análise como elo de ligação paratática, uma posição partilhada neste estudo.

Outros autores incluem *pois* em um grupo distinto, como é o caso de Bechara (1999) e de Peres e Mascarenhas (2008). Bechara (op. cit), embora reconheça que *pois* explicativo partilhe algumas propriedades dos coordenativos, o insere no conjunto das unidades adverbiais, que podem estabelecer relações interoracionais ou intertextuais. Peres e Mascarenhas (2006), conjugando argumentos sintáticos e semânticos, consideram que as orações com *pois* são instâncias de estrutura de suplementação, ou seja, em que uma oração estabelece uma relação semântica de dependência semântica com outra sem que haja integração sintática entre elas.

Enunciados construídos com *pois* se particularizam igualmente quanto às suas propriedades discursivas: apresentam a relação causal entre A e B como pressuposta, isto é, “conforme às expectativas” (Lopes 2004:24), assegurada como possível pelo conhecimento compartilhado pelos interlocutores. Constituem, na maioria das vezes, instanciações de estados de coisas mais gerais, normalmente relacionados por causa-efeito. Aproximam-se, portanto, de um raciocínio baseado na “normalidade das relações que se instauram entre os estados de coisas” (Lopes (op. cit: 36), garantida pelo nosso conhecimento de mundo. A natureza pressuposicional da relação entre A e B é independente do tipo de informação codificada pela oração encabeçada por *pois*, que pode introduzir informação dada ou nova. Há argumentos, portanto, para considerar que, em períodos complexos com *pois*, no português contemporâneo, instanciam-se dois atos de fala distintos e independentes, frequentemente acompanhados, inclusive, de pausa entre as duas orações.

3. Gramaticalização de orações complexas

Segundo a hipótese de unidirecionalidade, central nos estudos de gramaticalização, pode-se postular que os processos de combinação de orações seguem uma trajetória de [- integração] > [+ integração], ou seja, estruturas hipotáticas emergem de estruturas paratáticas. Nesta perspectiva, a evolução de processos de combinações de orações obedece à unidirecionalidade, segundo o qual estruturas mais gramaticais se originam de estruturas menos gramaticais ou lexicais.

De fato, a proposta acima é simplificadora, se consideramos que uma concepção dicotômica das formas de articulação de orações, traduzida na oposição coordenação/subordinação, se mostra limitada para explicar a evolução da articulação de orações e dos elementos conectores. Diversas objeções podem ser levantadas quanto à dicotomia coordenação/subordinação: a concepção destes conceitos em termos de primitivos (Haiman & Thompson, 1984); sua caracterização com base em critérios semânticos e sintáticos que podem conflitar entre si; o fato de que ela não consegue distinguir classes coerentes de orações e, menos ainda, os casos fronteiros⁽⁴⁾; a dificuldade de aplicar os critérios a línguas com outras formas de codificação de uma mesma relação semântica (Cristofaro 2003). Uma das críticas mais severas diz respeito à amplitude do termo subordinação para tratar, indistintamente, orações que funcionam como argumentos da oração núcleo e as denominadas adverbiais, que não possuem função argumental.

Diversas alternativas propõem uma divisão tripartida, (parataxe, hipotaxe, subordinação), para dar conta da natureza particular das orações adverbiais e das relativas explicativas em relação às completivas e adjetivas restritivas (Halliday 2004; Mathiessen & Thompson 1988; Hopper & Traugott 1993, 2003). Nesta perspectiva, a parataxe relaciona dois núcleos independentes, ambos assertivos, e constituiria uma forma mais simples de combinação de cláusulas; na hipotaxe, a margem ou satélite preserva certa independência sintática, pois não integra a estrutura argumental da cláusula núcleo; na subordinação, uma cláusula constitui um termo sintático daquela com que se liga. Na formulação de Hopper e Traugott (2003), as diferenças entre estes processos podem ser esquematizadas como:

(4) Ver, por exemplo, o caso das comparativas e consecutivas que, segundo várias análises, devem ser excluídas do grupo das adverbiais, por compartilharem propriedades tanto das orações relativas e completivas como das coordenadas. (cf. Brito & Matos 2003)

Parataxe >	Hipotaxe >	Subordinação
- dependência	+ dependência	+ dependência
- encaixamento	- encaixamento	+ encaixamento

(Reproduzido de Hopper & Traugott 2003 : 78)

De forma ainda mais radical, diversos autores propõem escalas mais detalhadas, ou seja, um *continuum* de dependência/independência entre as orações. Nesta perspectiva, maior ou menor dependência entre duas orações resulta da confluência de diversas propriedades, cada uma delas independentemente motivada. É o caso, por exemplo, de Lehmann (1988) para quem diferentes graus de vinculação entre duas orações se definem a partir da combinação de parâmetros como “dessentencialização” da oração; gramaticalização do verbo principal; entrelaçamento das orações; grau de explicitude do elo interoracional. A vantagem desta concepção é a de fornecer descrições mais adequadas de diferenças tipológicas.

Uma concepção não binária dos processos de articulação de orações permite compreender de forma mais clara as mudanças no uso de conectores. Assumindo que princípios atinentes aos processos de gramaticalização podem ser estendidos para o estudo da evolução de conectores e elos interoracionais, pode-se pressupor a generalidade da trajetória unidirecional [- integrado] > [+ integrado] (cf. Hopper e Traugott, 1993) na combinação de orações.

Uma primeira objeção a esta hipótese é colocada em Heine e Kuteva (2007), para quem, na verdade, o esquema acima constitui apenas um dos mecanismos de evolução de períodos complexos, aquele em que duas orações independentes são integradas em uma única sentença. Como consequência desta integração, as duas orações podem apresentar propriedades comuns, como: referentes compartilhados (mais frequentemente o referente do sujeito); simultaneidade ou adjacência temporal; compartilhamento de localização espacial e contorno entonacional único. Este processo é particularmente produtivo no desenvolvimento de relativizadores e de complementizadores.

O desenvolvimento de conectores que introduzem orações adverbiais envolve, predominantemente, um mecanismo de expansão, ou reinterpretação de sintagmas adverbiais, nominais ou verbais, como elos de conexão interoracional. Este é o caso de *pois*, que, como vimos na seção anterior,

deriva de uma fonte adverbial com valores locativos e temporais. Na grande maioria dos casos, essa mudança é acompanhada de mudanças semânticas, igualmente unidirecionais, no sentido de [- subjetivo] > [+ subjetivo] (Traugott & König, 1991), ou, na versão proposta por Traugott (2003, 2010) [- subjetivo] > [+ subjetivo] > [intersubjetivo]. Nos seus estágios iniciais como conector, um elemento seria convocado para a expressão de relações entre estados de coisas (enunciado) e se estenderia, gradualmente, para sinalizar relações no mundo das crenças e atitudes (enunciação) e, numa última etapa, se estenderia para usos intersubjetivos, ou seja, como forma de regulação das relações entre os interlocutores.

Esta hipótese de mudança semântica tem sido questionada em diversos trabalhos. Contrariando o *cline* previsto, são identificados casos em que significados subjetivos emergem antes de significados objetivos, referenciais, que surgem apenas em estágios mais avançados da língua. Assim Paiva e Braga (2013) mostram que, além de marginal, o uso de *pois* para relações causais no domínio referencial, [- subjetivo], é mais tardio. Nos estágios iniciais do português, o uso de *pois*, ao que tudo indica, fica mais restrito às relações nos domínios epistêmico e conversacional. Concluem que, se o uso de *pois* como conector temporal, ainda frequente no período arcaico, pode ser interpretado em termos de subjetivização, há fortes evidências de que as mudanças semânticas subsequentes operam no sentido de [+ subjetivo] > [- subjetivo].

Um exemplo ilustrativo é o estudo diacrônico de Evers-Vermeul sobre os pares de conectores causais/explicativos *omda/want*, *dus/ daarom* do holandês. A autora destaca dois pontos: a possibilidade de intercambialidade entre eles, pelo menos até o holandês médio, e a acentuada estabilidade das propriedades destes elementos, ao longo de 800 anos. As mudanças no uso destes conectores envolvem uma redistribuição de acordo com os domínios de causalidade. A autora atesta que, embora o par *omdat* e *want*, por exemplo, pudesse alternar no holandês antigo, ao longo do tempo, o primeiro se especializa na expressão de relação causal no domínio do conteúdo e o segundo no domínio epistêmico. Conclui que, mais do que subjetivização, a evolução de um conector envolve uma especialização em relação a outros membros do conjunto a que ele pertence. A esta especialização pode estar subjacente tanto subjetivização como objetivização.

Uma outra questão, até certo ponto superposta à natureza mais ou menos subjetiva da relação causal, envolve o estatuto sintático das orações introduzidas por estes elementos: na sua trajetória de gramaticalização, eles obedecem a um *cline* de paratático > hipotático? Análises de diversas

línguas permitem depreender padrões regulares de mudança de uso dos conectores no sentido de ligações paratáticas para ligações hipotáticas. No entanto, outros estudos têm apontado evidências contrárias a essa trajetória, a partir de análises que atestam não apenas o desenvolvimento de estruturas paratáticas a partir de estruturas hipotáticas, ou mesmo subordinadas (cf. Hopper & Traugott 1993; Harris & Campbell 1995; Frajzyngier 1996; Ziegeler 2004; Gunther 2010; König & Van der Avera 1988, Pereira et ali., 2010; Evers-Vermeul, 2005). Pereira, Paiva e Braga, por exemplo, defendem que a formação das locuções conjuntivas temporais *na hora que, no dia que* resultam de uma complexa reanálise de construções relativas. Concluem, então que, neste caso, ter-se-ia um desenvolvimento de estruturas menos integradas, hipotáticas, a partir de estruturas subordinadas. Uma hipótese semelhante é aventada por Fiéis e Lobo (2008) para algumas conjunções causais/explicativas do português, particularmente para *pois*. Comparando propriedades sintáticas das orações introduzidas por este conector em diferentes estágios do português e com base na possibilidade de que elas sejam antepostas ou precedidas de uma conjunção coordenativa, as autoras concluem que, em estágios anteriores do português, tais orações se comportavam como subordinadas periféricas. Propõem, então, que *pois* teria passado por uma mudança de conector subordinativo para conector coordenativo.

4. Diacronia da posição das orações explicativas encabeçadas por *pois*

Como foi visto na seção 1, uma particularidade das orações encabeçadas por *pois*, no português contemporâneo, é a sua posição fixa, posposta à oração núcleo. No entanto, em estágios anteriores da língua, elas admitiam flexibilidade (cf. Fiéis & Lobo, 2008), podendo ser antepostas, intercaladas ou pospostas à oração núcleo, como mostram, respectivamente, (1), (2) e (3):

(1) *Pois estamos em tempo de restituir*, restitua-me Vossa Mercê a sua graça. Se espera que eu o mereça, para mi será desesperação essa esperança. Mais há de quinze dias que ainda a pouca saúde que tinha se foi por aí. E segundo eu estou longe dela, tarde tornará (Séc. XVII. Carta pessoal)

(2) As cartas pera as India vos iram dentro nesta semana, prazendo a Noso Senhor. Emcome(dovos muyto que, *pois estam tam* cedo, prestes trabalheis por nõ perderem tempo. E dõ Gonçalves Coutinho he jaa llaa, ha quatro ou çimquo dias, e nõ amda quaa pesoa allgu(a por que ellas ajam de sperar. (Sex. XVI- CRB)

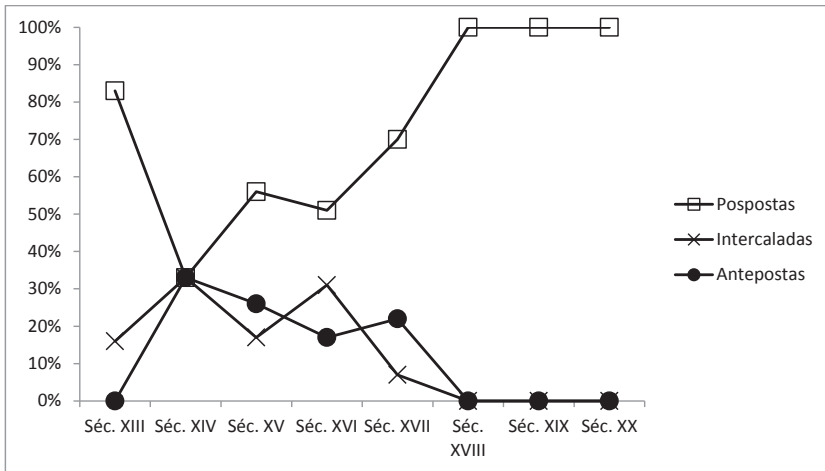
(3) o corpo do home~ he adoptado e co~uinhaul aa me~te e~ deuuda yqualdança, da qual desuayraria, se lhe fosse e~adudo algu~a cousa de afeytame~to, mayorme~te das cousas baixas. E merece pore~ de se vingar Deus do home~, *pois lhe faz e~juria co~ os afeytamentos*. (Séc. XV, OE)

Um aspecto merece ser ressaltado em relação à possibilidade de intercalação. Mesmo se a maioria dos casos são similares a (2), ou seja, interposição da oração com *pois* no interior daquela à qual se liga. são também recorrentes nos estágios iniciais, enunciados como (4), em que a oração interposta é antecedida de um conector coordenativo.

(4) E isto convem que seja em tamanho numero que, posto que lhe tan asynha na~o venha socorro, que se possa mamter, ca, *pois ha serventia de vosso rregno na~o pode ser sena~o per agua*, he de emtemder que na~o aveis de ter o vemo a vosso mamdado, mas cuydaí que se pode seguyr tall azo que estara~ os navios em vossos rregnos tres & quatro meses & no~ averem tempo de viage~ (séc. XV, DPM)

Casos como (4) poderiam ser analisados como anteposição da oração encabeçada por *pois*. (Fiéis & Lobo, 2008). Acreditamos, no entanto, que, do ponto de vista discursivo, eles constituem casos de inserção de uma oração explicativa no interior de outra oração causal, no caso a introduzida por *ca*. Um argumento favorável a esta interpretação é a tendência à sequenciação de diversos conectores no português arcaico. É necessário considerar, ainda que uma parte significativa de dados desta natureza envolve a coocorrência com *ca*, elemento que, dentre outras funções, podia ser usado para sinalizar continuidade textual (cf. Barreto, 1999).

A análise de 294 orações, distribuídas de forma bastante desigual ao longo de oito séculos (século XIII ao século XX), permite constatar o enriquecimento da ordem nos enunciados construídos com o conector *pois*, na forma de oração núcleo + oração causal/explicativa, a partir do século XVIII, como mostra o gráfico 1:

Gráfico 1- Distribuição das orações com *pois* de acordo com a posição

No que se refere à anteposição das orações introduzidas por *pois*, observa-se sua predominância nos séculos XIV (33%) e XV (26%), períodos em que se verifica também maior ambiguidade categorial e semântica do item *pois*: seu emprego original como advérbio coexiste com seu uso como conector (cf. Lima, 2002; Barreto 1999; Paiva & Braga, 2013) e, nesta função, *pois* podia instanciar relações semânticas diferenciadas como tempo, causa e conclusão (Paiva & Braga, 2013). Nos séculos XVI e XVII, observa-se relativa estabilidade de orações antepostas com *pois*, com índices aproximados de 22%. A partir do século XVIII, não se atestam mais ocorrências de enunciados em que a oração explicativa preceda a oração núcleo.

A tendência para a interposição de orações introduzidas por *pois* ao longo do tempo é mais irregular, com picos mais acentuados nos séculos XIV (33%) e XVI (31%). Já no século XVII, decresce de forma significativa a frequência de orações com *pois* intercaladas (7%), possibilidade que desaparece completamente no século XVIII.

A posposição de orações com *pois* à oração núcleo se destaca como ordem não marcada, pelo menos em termos de frequência, já no século XIII. De fato, dos 6 casos de *pois* explicativo coligidos neste momento, 5 são de posposição. No entanto, a posposição de *pois* decresce nitidamente no século XIV, mantém-se relativamente estável entre os séculos XV e XVI,

com percentuais próximos de 50%. A partir do século XVII, momento em que o índice de posposição alcança 70%, a direção da mudança parece ser irreversível e a posposição torna-se categórica, a partir do século XVIII. Neste sentido, este estudo confirma a afirmação de Fiéis e Lobo, acerca do enrijecimento da ordem das explicativas com *pois*.

A questão central, a nosso ver, é se a flexibilidade observada em períodos anteriores do português poderia ser interpretada como evidência de mudança no estatuto sintático do conector *pois* no sentido de introdutor de orações hipotáticas para orações paratáticas. Um primeiro aspecto a considerar é a inter-relação entre a ordenação das orações e o domínio em que se instaura a relação de causalidade. Como já discutido em Paiva (1991), uma disposição icônica, consoante ao pressuposto de antecedência da causa ao efeito, seria mais previsível para as orações que codificam causa estrita, ou seja, em que A e B constituem estados de coisas. Embora ao longo de todo o período em análise o emprego de *pois* prevaleça para relações nos domínios epistêmico e conversacional, ou seja, introduzindo explicações e justificativas (cf. Paiva & Braga, 2013), podem ser atestadas ocorrências, ainda que marginais, de enunciados que admitem uma interpretação de causa no domínio referencial, como no exemplo (5):

(5) Sahimos no dia 25, ás 8 $\frac{3}{4}$ e a 27 ás 8 $\frac{1}{2}$ da manhã, encontramos uma pequena canoa de casaca, chegando ás 8 e 50' em um porto onde havia três canoas. Desembarcamos depois de ter eu dado algumas providências *pois*, desconfiava não serem estas canoas de bakairis. (Séc. XIX, Cartas Administrativas)

Em (5), a oração encabeçada por *pois* parece se situar numa fronteira bastante tênue entre causa estrita e explicação, admitindo duas interpretações possíveis. Na primeira delas, introduz uma explicação (a desconfiança de que as canoas encontradas representassem algum perigo) que justifica a necessidade de providências prévias ao desembarque. Considerando, no entanto, a organização mais narrativa do trecho e o fato de que tomar providências constitui um estado de coisas agentivo e intencional, não fica excluída uma leitura da oração com *pois* como a causa das precauções adotadas para o desembarque. Esta interpretação é favorecida, ainda, pela presença da oração temporal que precede a oração com *pois*.

Ao que tudo indica, a posposição das orações encabeçadas por *pois* é inteiramente independente do domínio em que se instaura a relação de causalidade. Um ponto merece, porém, ser destacado. No período em que

se constata maior flexibilidade na posição da oração com *pois*, qual seja, do século XIV ao século XVII, há diferenças mais sutis na distribuição para as diferentes posições das orações com *pois*.

Tabela 1- Ordenação de orações com *pois* de acordo com o domínio da causalidade

Século	Epistêmica			Referencial			Conversacional		
	Antep.	Interc.	Posp.	Antep.	Interc.	Posp.	Antep.	Interc.	Posp.
XIII	0	0	3 100%	0	0	1 100%		1 50%	1 50%
XIV	1 50%	1 50%	0			1 100%	0	0	0
XV	4 31%	2 15%	7 54%	-	-	-	2 20%	2 20%	6 60%
XVI	-	4 50%	4 50%	-	-	-	6 22%	7 26%	14 52%
XVII	8 17%	3 6%	37 77%	1 14%	1 14%	5 72%	11 31%	3 9%	21 60%

Devido à severa limitação no número de dados para os séculos XIII e XIV, nos restringimos aos períodos subsequentes. Ainda assim, é necessária cautela na interpretação dos resultados resumidos na tabela 1, em razão, sobretudo, do número muito escasso de *pois* no domínio referencial nos séculos XV e XVI. O ponto digno de nota diz respeito, portanto, ao uso de *pois* nos domínios epistêmico e conversacional. Para o domínio epistêmico, observa-se um nítido corte entre os séculos XIV e XV, por um lado, e o século XVII, por outro, principalmente no que se refere aos valores para a posposição: índices muito próximos nos séculos XV e XVI e aumento significativo no século XVI (77%). Este aumento é acompanhado do recuo de orações com *pois* intercaladas e antepostas.

Para as orações com *pois* no domínio conversacional, configura-se uma situação mais compatível com estabilidade da posposição, associada a altos

índices desde o século XV. Quanto à anteposição das orações com *pois*, surpreende seu aumento no século XVII (31%), após um período de estabilidade, com índices similares para os séculos XV e XVI. As orações com *pois* intercaladas, por sua vez, apresentam uma trajetória mais regular, com frequências próximas para os séculos XVI e XVII e nítida redução no século XVIII. Os fatos mais relevantes parecem ser, portanto, o enrijecimento na posição de orações introduzidas por *pois* e o decréscimo da possibilidade de intercalação.

5. Da hipotaxe à parataxe ou especialização semântico-discursiva?

Os resultados discutidos até este ponto constituiriam, em princípio, argumentos favoráveis à postulação de uma trajetória hipotaxe > parataxe para o conector *pois*). Uma objeção a esta conclusão é que, apesar da sua flexibilidade até o século XVII, as orações encabeçadas por *pois* resistem a operações como inserção no escopo de uma interrogação ou de uma negação ou, ainda, de um advérbio enunciativo. Além disso, não podem constituir complemento de uma outra oração nem serem clivadas.

Outras propriedades apontam pistas para uma reflexão. Dentre eles, destaca-se a independência prosódica das orações com *pois*. Ainda que a imprecisão da pontuação em textos de estágios anteriores imponha cautela, a alta frequência de exemplos como (6) é mais compatível com uma interpretação de estatuto paratático das orações encabeçadas por *pois*.

(6) Duarte Coelho me dise a muyto bõa vomtade e obras que achara no duque de Nemurs, irmão do duque de Saboya, meu muyto amado e preçado irmão, pera as cousas de meu serviço, e muyto bõo fora tardes levado carta para elle. E porem, *pois* a nam levastees, aguora volla mando na forma que a mamdaeis pedyr (séc.XVI, CDJ)

No trecho (6), a oração com *pois*, separada por vírgulas daquela com que se liga, adquire características de um comentário parentético, adicional, acrescentado de forma a garantir a naturalidade do desejo do infante.

Um caso mais extremo de independência prosódica das orações explicativas com *pois* se verifica nos contextos em que ela se segue a ponto final, como no exemplo (7):

(7) Ainda mal, porque para acreditar, o que disser nesta Relação, tenho já tão curto numero de testemunhas, que eu serei só o autor della. *Pois* dos poucos que deste naufragio escaparão vivos, são hoje mortos, quasi todos. (Séc.XVII, ELP)

Embora entre os séculos XIV e XV ainda possam ser atestados alguns casos de ausência de vírgula antes da oração com *pois*, a independência prosódica destas orações é sistemática ao longo dos oito séculos. Como já postulado por Guimarães (1987), independência prosódica é o reflexo de uma estrutura discursiva na forma de tópico-comentário. Nos termos do autor (Guimarães, op. cit:79), “se o tema é uma oração e o comentário outra, então cada oração deve corresponder a um grupo entonacional”. Alguns conectores, *pois* dentre eles, se tornam marcadores de comentário, explicando-se, assim, muitas das suas particularidades sintáticas. De forma semelhante, Lopes (2004:5) postula que *pois*, assim como *porque* e *que*, só podem introduzir informação não pressuposta, ou seja, constituírem uma asserção. Tal restrição parece corresponder bastante bem aos usos de *pois* no português brasileiro contemporâneo. Entretanto, em estágios anteriores do português, essa restrição não era categórica, podendo-se atestar casos de orações explicativas introduzidas por *pois* que codificam informação inferível ou já introduzida no discurso anterior, como nos trechos exemplificados a seguir:

(8) se o marido a no~ quis(er) accusar nen er q(ui)s(er) q(ue) seya doutri~ accusada, nenguu no~ seya recebido por accusador en tal feyto, **ca pois** el quer p(er)duar a ssa molh(er) este peccado, no~ e' deryto q(ue) outri~ a demande ne~ sub(re) el acuse nenhu~a cousa. (sec. XIII, FR)

(9) E, segundo conta Lucha[m], o que escreveo esta estoria, depois que foro~ co~pridos os cinque a~nos, ma~daronlhe dizer os Roma~a~os que se tornasse, se no~ que o nom receberia~ mais por senador.E elle, com despeyto, nom o quis fazer mais disse que, **pois elle era sanador**, que tomava elle outros cinque a~nos; e e~ este cinquo a~nos conquistou tod[a] Espanha. (Sec.XIV, CGE)

No exemplo (8), a informação veiculada pela oração com *pois* é recuperável do discurso anterior, mais especificamente, do conteúdo da oração condicional onde se especificam as condições para a aplicação da regra estabelecida. Através de um processo inferencial, pode-se concluir que, em caso de o marido querer perdoar sua mulher pelo pecado cometido, nenhuma outra pessoa terá o direito de acusá-la. Neste caso, a justificativa

que assegura a validade do ato de fala diretivo realizado a seguir, toma por base uma informação já compartilhada pelos interlocutores.

O exemplo (9) é ainda mais interessante, já que a oração com *pois*, intercalada entre dois complementizadores ligados ao verbo *dizer* e o seu complemento, retoma uma informação já apresentada no discurso anterior, portanto, compartilhada.

Os exemplos acima apresentam configuração sintagmática e discursiva bastante similar. Todos eles admitem uma paráfrase, por exemplo, por *já que*, conector mais especializado na introdução de informação compartilhada pelos interlocutores ou pressuposta. Estes fatos permitem levantar a hipótese de que as restrições à posição das orações com *pois* no português contemporâneo, mais do que uma trajetória de mudança no grau de integração da oração, reflete a perda de uma possibilidade discursiva, qual seja, a de introduzir informação velha ou inferível e conseqüente especialização semântico-discursiva como elemento de introdução de informação nova, configurando, assim, enunciados do tipo tema-rema. Este movimento poderia estar associado ao surgimento e generalização de *já que*, o que tudo indica, a partir do século XVII (cf. Fieis & Lobo 2008; Paiva & Braga 2011).

Uma evidência adicional desta configuração é a frequente presença de elementos anafóricos nas orações encabeçadas por *pois* antepostas ou interpostas, reforçando seu *status* de informação compartilhada, como no exemplo (10), em que, na oração explicativa com *pois* grifada, o SPrep “*de aquele naufrágio*” remete ao acidente descrito na oração anterior.

(10) Otros muytos soldados de importancia ficârão sepultados entre aquellas aguas; dos quaes eu desejei trasladar os nomes, pois não podia os ossos, a estas letras, para immortal memoria delles: (...) *Pois Deos me livrou do risco de aquelle naufragio*, os livrasse eu se pudesse aelles, tambem do naufragio do esquecimento (Séc. XVII, ELP)

Ainda que não esteja excluída a possibilidade de orações com *pois* antepostas ou intercaladas codificarem informação nova, a grande maioria delas (47/59 = 79%) se caracteriza pelo traço [+ informação dada].

Um outro aspecto merece reflexão: em muitos desses casos de anteposição, as construções com *pois* antepostas e intercaladas se aproximam dos contextos mais característicos da locução conjuntiva *pois que*, como exemplifica (11):

(11) *Poys que offyzio dos escriuaes e' publico e (co)munal p(er)a todos, mandamos que a todos aquellos q(ue) dema~dare~ carta p(er)a s(eus) p(re)ytos, assy p(er) mandado dos alcaydes coma por os iuyzes como p(er) si dalgu~as (con)pras ou de uendas dos omees que a[i]a~ d(e) faz(er), faças sen outro e~longame~to nenhuu e no~ as leyxe d(e) faz(er) por amor ne~ por desamor nenhuu ne~ p(er) medo ne~ p(er) uergonha d(e) nenguu. (Sec. XIII, FR)*

Um dado relevante é que nos 47 períodos complexos com *pois que* coligidos nos mesmos textos, a oração encabeçada por esta locução codifica informação compartilhada ou apresentada como tal. Não se pode excluir, portanto, a possibilidade de que, durante a coexistência entre os dois conectores, *pois* tenha incorporado propriedades de *pois que*, o que explicaria, inclusive, a intercambialidade entre eles em alguns contextos, durante um certo período.⁽⁵⁾ Dada a recorrência muito mais significativa de *pois* como introdutor de orações que expressam informação [- dada], este uso teria se generalizado, ou nos termos de Traugott (2003), semanticizado ao longo do tempo.

6. Considerações finais

Ao longo deste artigo apresentamos algumas evidências para uma questão controversa no que se refere à evolução de períodos complexos, qual seja, a hipótese de que estruturas [+ integradas], hipotáticas se originam de estruturas [- integradas], paratáticas. Considerando alguns aspectos da evolução do conector causal/explicativo *pois*, apresentamos alguns argumentos que permitem discutir esta direcionalidade.

O surgimento de *pois* conectivo a partir de uma base adverbial exemplifica bastante bem um processo de gramaticalização no seu sentido mais clássico. O posterior desenvolvimento dos diversos valores/funções de conector *pois*, ao contrário, suscita questões atinentes ao seu estatuto sintático. Mostramos, com base numa análise quantitativa, que suas restrições sintáticas no português contemporâneo são melhor explicadas se consideramos as propriedades discursivas das orações introduzidas por *pois*, notadamente no que se refere ao tipo de informação que elas codificam.

A análise permitiu mostrar que a perda de flexibilidade da oração encabeçada por *pois*, nítida a partir do século XVIII, mais do que o indício de

(5) Para Fagard (2009), *pois que* constituiria um estágio anterior que teria dado origem *pois* através da perda de *que*. Esta interpretação pode ser discutida, se considerarmos que a intercambialidade entre os dois conectores é limitada a certos contextos (cf. Paiva & Braga, 2013).

uma mudança de estatuto sintático das orações que ele introduz, pode ser interpretada como uma mudança discursivamente motivada. No período, bastante longo em que as orações introduzidas por *pois* admitiam posição variável, pôde ser atestada igualmente maior flexibilidade no seu estatuto informacional: orações explicativas encabeçadas por *pois* serviam tanto à apresentação de informação nova como de informação dada. Como se pode esperar, esta maior flexibilidade discursiva encontra paralelo na forma de linearização do período complexo: orações explicativas com informação dada ou inferível são mais frequentemente antepostas ou intercaladas e orações com informação nova, mais frequentemente pospostas à núcleo (Paiva, 1991). Pode-se, então, concluir que o enrijecimento na posição de orações explicativas com *pois* implica uma cristalização do seu papel discursivo, ou seja, o de introduzir informação nova.

Evidentemente, uma verificação mais rigorosa desta hipótese requer situar o conector *pois*, em cada um dos seus estágios de evolução, no paradigma em que ele se insere, a fim de identificar em que medida a perda de uma possibilidade discursiva resulta de uma possível divisão de funções entre os diferentes elementos coexistentes. Podemos presumir que as mudanças operadas nos enunciados explicativos com *pois* não sejam indiferentes à longa coexistência com a locução *pois que* e à possível intercambialidade entre eles. Evidências mais seguras para esta conclusão teriam que levar em conta também outros movimentos no conjunto de conectores causais, dentre eles, o desaparecimento de *ca*, a alta frequência do uso de *porque* e a emergência de *já que*.

Referências

- BARRETO, T. (1999), *Gramaticalização das conjunções na história do português*, Instituto de Letras, Universidade Federal da Bahia, Salvador, Tese de doutorado em Letras.
- BECHARA, E (1999), *Moderna Gramática Portuguesa*, 37ª ed. Rio de Janeiro, Lucerna.
- BRAGA, M. L & PAIVA, M. C. (2011), “Gramaticalização e gramática de construções: estabilidade e instabilidade no uso de orações complexas de causa em tempo real”, *Revista Letras & Letras*, 27, pp.51-70.
- CANDIDO, F. M. (2009), *Os diferentes padrões das construções com “pois”*, Araraquara, UNEP, Dissertação de Mestrado.
- CEGALLA, D. P. (1970), *Novíssima gramática da língua portuguesa*, São Paulo, Editora Nacional, 1970.
- COROMINAS, J & PASCUAL, J. A. (1991), *Diccionario crítico etimológico castellano e hispánico*, Madrid, Gredos.

- COUTINHO, I. L. (1962), *Gramática Histórica*, Rio de Janeiro, Acadêmica.
- CRISTOFARO, S. (2003), *Subordination*, Oxford, Oxford University Press, 2003.
- CUNHA, C. F. (1970), *Gramática do português contemporâneo*, Belo Horizonte, Bernardo Alvares.
- DANCYGIER, B. & SWEETSWER, E. (2000), “Constructions with if, since, and because: Causality, epistemic stance, and clause order”, in E. Couper-Kuhlen & B. Kortmann (eds), *Cause, condition, concession and contrast: Cognitive and discourse perspective*. Berlin: Mouton de Gruyter, pp. 111-142.
- EVERS-VERMEUL, J. (2005), *Connections between form and function of Dutch Connectives: Change and acquisition as windows on form-function relation*, Dissertation UiL OTS, Universiteit Utrecht.
- FAGARD, B. (2009), “Grammaticalisation et renouvellement : conjonctions de cause dans les langues romanes”, *Revue romaine de linguistique* 54, pp. 21-43.
- FIÉIS, A & LOBO, M. (2008), *Para uma diacronia das orações causais e explicativas do português*. Comunicação. XXIV Encontro Nacional da APL, Braga, Portugal.
- FRAJZYNGIER, Z. (1996), *Grammaticalization of the complex sentence: a case study in Chadic*, Philadelphia, John Benjamins.
- GUIMARÃES, E (1987), *Texto e argumentação*, Capinas, Pontes.
- GUNTHNER, S. (2010), “From subordination to coordination? Verb second position in German causal and concessive constructions”, *Pragmatics* 32, pp. 33- 56.
- HAIMAN, John & THOMPSON (1988), “Introduction” in J. Haimann & S. Thompson (ed), *Clause Combining in Grammar and Discourse*, Philadelphia, John Benjamins, pp. X-XIII.
- HEINE, B.; CLAUDI, U. & HÜNNEMEYER, F (1991), *Grammaticalization: a conceptual framework*, Chicago, Chicago University Press.
- HALLIDAY, M. A. K. (2004), *An introduction to functional grammar*, London, Arnold, 1994.
- HARRIS, A. & CAMPBELL, L. (1995), *Historical syntax in cross-linguistic perspective*, Cambridge, Cambridge University Press.
- HEINE, B. & KUTEVA, T. (2007), *The genesis of grammar: a reconstruction*, Oxford, Oxford University Press.
- HOPPER, P; TRAUGOTT, E. C. (2003), *Grammaticalization*, 2. ed. Cambridge, Cambridge University Press.
- KÖNIG, E. & VAN der AUWERA, J. (1988), “Clause integration in German and Dutch: conditional and concessive”, in J. Haiman & S. Thompson, (eds), *Clause combining in Grammar and Discourse*, Amsterdam/Philadelphia, John Benjamins, pp.101-133.
- LEHMANN, Christian (1988), “Towards a Typology of Clause Linkage” in S. Thompson & J. Haiman (orgs), *Clause Combining in Grammar and Discourse*, Philadelphia, John Benjamins, 1988. p. 181-225.

- LIMA, J. P de (2002), "Grammaticalisation, subjectification and the origin of phatic markers", In: W. Wischer & G. Diewald (eds), *New reflections on grammaticalisation*, Amsterdam/Philadelphia, John Benjamins, pp. 363-378.
- LOBO, M. (2003), *Aspectos da sintaxe das orações adverbiais do português*, CLUL, Universidade Nova de Lisboa, Lisboa, Dissertação de doutoramento.
- LOPES, Helena C. (2004), *Aspectos sintáticos, semânticos e pragmáticos das construções causais: contributo para uma reflexão sobre o ensino da gramática*, Universidade do Porto, Porto, Dissertação de doutoramento.
- MATTHIESSEN, C.& THOMPSON, S (1988), "The structure of discourse and 'subordination'", in J. Haiman & S. Thompson (eds.) *Clause combining in grammar and discourse*, Amsterdam/Philadelphia, John Benjamins, pp. 243-281.
- MATOS, G. (2005), "Coordination vs. subordination adverbiale - propositions causales en portugais", Disponível em <http://www.cavi.univ-paris3.fr//ilpga/colloque-coord-subord-2005/pretextes/index.html>.
- MATTOS E SILVA, R, V (1989), *Estruturas Trecentistas. Para uma Gramática do Português Arcaico*, Lisboa, Imprensa Nacional/Casa da Moeda. 1989.
- MATTOS E SILVA, R. V. (2006), *O português arcaico: morfologia e sintaxe*, 2. ed, São Paulo, Contexto.
- NASCENTES, A (1995), *Dicionário etimológico da língua portuguesa*, v. 1, Rio de Janeiro, Livraria Acadêmica, 1955.
- NEVES, M. H. de M. (1999), *Gramática de usos do português*, São Paulo, Editora da Unesp.
- NUNES, J. J. (1975), *Compêndio de Gramática Histórica Portuguesa, Fonética e Morfologia*, Rio de Janeiro, Livraria Clássica Editora.
- PAIVA, Maria C. de (1991), *Ordenação das cláusulas causais: forma e função*, Rio de Janeiro: FL/UFRJ, tese de doutoramento.
- PAIVA, M. C. de (1995), "Empregos de *Porque* no Discurso oral", *D.E.L.T.A.*11, São Paulo, PUC- SP, pp. 278-30.
- PAIVA, M. C DE (1996), "Pressupostos Semânticos e discursivos da relação de causalidade" in: Alzira Macedo et alii. (org.), *Variação e discurso*, Rio de Janeiro, Tempo Brasileiro, pp. 51-62.
- PAIVA, M. C. A. & BRAGA, M. L. (2010), "Cláusulas *porque*: da sintaxe ao discurso in M. C. Mollica, (Org.), *Usos da linguagem e sua relação com a mente humana*, Rio de Janeiro, Editora Tempo Brasileiro, 2010, pp. 55-71.
- PAIVA, M. C.& BRAGA, M. L. (2013), "Evolução de *pois* e *pois que* no português: uma trajetória de subjetivização?" in Maria Maura Cezário & M. Angélica F. da Cunha, *Linguística centrada no uso: uma homenagem a Mário Martelotta*, Rio de Janeiro, Mauad/Faperj, pp.97-112.

- PEREIRA, M. H., BRAGA, M. L. & PAIVA, M. C. (2010), "Gramaticalização das construções (prep)1 + (det) + N + (prep)2 + que" in Lourenzo Vital & Sueli Coelho, *Estudos de processos de gramaticalização em português: metodologias e aplicações*, Campinas, Mercado das Letras, pp. 173-200.
- PERES, João Andrade (1997), "Sobre Conexões Proposicionais em Português, in A. M. Brito, F. Oliveira, I. P. de Lima e R. M. Martelo (org.), *Sentido que a Vida Faz. Estudos para Óscar Lopes*, Campo de Letras, Porto, pp. 775-787.
- PERES, J.; MASCARENHAS, S. (2006), "Notes on sentential connections (predominantly), in Portuguese", *Journal of Portuguese Linguistics* 5, pp. 113-169.
- QUIRK, R. Sidney GREENBAUM, Geoffrey LEECH & Jan SVARTVIK (1985), *A Comprehensive Grammar of the English Language*, London/New York, Longman.
- SAID ALI, M. (2001 [1921]), *Gramática histórica da língua portuguesa*, 8. ed, São Paulo, Melhoramentos.
- SWEETSER, E. (1990), *From etymology to pragmatics: metaphorical and cultural aspects of semantic structure*, Cambridge, Cambridge University Press, 1990.
- TRAUGOTT, E. C. (1989), "On the rise of epistemic meanings in English: an example of subjectification in semantic change", *Language* 65, pp. 31-55.
- TRAUGOTT, E. (1995), "Subjectification in grammaticalization", in D. Stein & D. Wright, (ed.), *Subjectivity and Subjectivisation*, Cambridge: Cambridge University Press, 1995, pp. 37-54.
- TRAUGOTT, E. (2003), "From subjectification to intersubjectification" in Raymond Hickey (ed), *Motives for language change*, New York, Cambridge University Press, pp. 124-142.
- TRAUGOTT, E. (2010), "(Inter)subjectivity and (inter) subjectification: a reassessment", in K. Davidse et alii (org), *Subjectification, intersubjetification and grammaticalisation*, Berlin/New York, Mouton de Gruyter, pp. 29-74.
- TRAUGOTT, E. & KÖNIG, E. (1991), "The semantics-pragmatics of grammaticalization revisited" in E. Traugott & B. Heine, (eds). *Approaches to grammaticalization*, Amsterdam/Philadelphia, John Benjamins, pp.189-218.
- ZIEGLER, D. (2004), "Redefining unidirectionality: is there life after modality?" in O. Fischer; M. Norde & H. Perridon (ed), *Up and down the cline: the nature of grammaticalisation*, Amsterdam, John Benjamins, pp. 115-126.

SNS COMO RÓTULOS EM LIVROS DIDÁTICOS DE HISTÓRIA DO BRASIL: SIMPLES OU COMPLEXOS?

Vera Lúcia Paredes Pereira da Silva
veraparedes@terra.com.br

Gabrieli Pereira Bezerra
gamari@ig.com.br

Este artigo pretende identificar e analisar sintagmas nominais inespecíficos que atuam como rótulos em um *corpus* de livros didáticos de História do Brasil. Além disso, propõe discutir, a partir da comparação com gêneros dos domínios acadêmico e jornalístico, o grau e o tipo de complexidade dos sintagmas nominais sob análise, a partir da hipótese de que a complexidade do sintagma nominal pode contribuir para a identificação do gênero em que se insere.

Palavras-chave: Sintagmas nominais; rótulos; referência; gêneros discursivos.

This article identifies and analyzes unespecific Noun Phrases used as labels in a *corpus* constituted by textbooks of Brazil's History. Furthermore, it proposes to discuss the degree and type of complexity of those NPs, taking into account evidence from other genres, from academic and journalistic domain. The hypothesis that the complexity of the Noun Phrase may contribute to the identification of discourse genres is our point of depart.

Key words: Noun phrases; labels; referentiation; discourse genres.

*

1. Introdução

Este artigo visa à análise da estrutura e da função de sintagmas nominais (doravante SNs) inespecíficos que atuam como rótulos numa coleção de livros didáticos contemporâneos de História do Brasil. Discute-se, também, o grau de complexidade desses SNs, a partir da proposta de que a complexidade do SN pode ser um parâmetro adicional na identificação dos gêneros de discurso (cf. Paredes Silva, 2009, 2011). Nesse sentido, é estabelecida

uma comparação com a complexidade de SNs em outros gêneros dos domínios acadêmico e jornalístico já analisados.

O uso de sintagmas nominais inespecíficos é uma das estratégias de referenciação empregada pelo produtor de um texto. Essa estratégia, nomeada por Francis ([1994] 2003) de “rotulação”, pode assumir a função de introduzir objetos de discurso, conectar, sumarizar e organizar as partes do texto de diferentes extensões, ligando o que foi dito ao que será dito, contribuindo, assim, para a progressão textual. Além disso, o rótulo pode atribuir avaliação aos segmentos textuais em que está inserido. Vejamos um exemplo:

- (1) A escravidão estimulou **pensamentos racistas**. Os brancos olhavam para a situação dos escravos e chegavam a uma conclusão absurda: “Eles foram escravizados porque são seres humanos inferiores aos europeus.”⁽¹⁾
Schmidt (2003:214)

No exemplo (1), temos um rótulo prospectivo (ou catafórico) constituído de um nome-núcleo- *pensamentos* e um modificador- *racistas*. Esse rótulo remete às informações subsequentes para que tenha seu significado devidamente compreendido, sendo também responsável por introduzir uma avaliação do produtor do texto sobre a informação-suporte⁽²⁾.

Ao mesmo tempo, buscamos correlacionar a noção de referenciação à análise de gêneros e ao estudo dos SNs. Para Koch (2003), a competência sociocomunicativa dos falantes torna-os capazes de discernir o que é adequado ou não em determinada situação social, bem como de diferenciar os gêneros: anedota, poema, atas, cartas pessoais, notícias jornalísticas e outros. Torna-os, também, aptos a identificar as sequências que predominam em um texto: narrativa, expositiva, argumentativa ou descritiva, por exemplo.

Dentro dessa perspectiva, o livro didático pode ser classificado como *gênero secundário*, já que é uma unidade funcional utilizada em situações de “convívio cultural mais complexo” (cf. Bakhtin, 2003:263). Além disso, nesses livros, podem ser incorporados vários gêneros- notícia, banda desenhada, charge, etc. e convivem várias sequências textuais - argumentativas, expositivas, narrativas, etc.

(1) Somente os exemplos retirados do *corpus* em análise estão numerados e com o rótulo em negrito.

(2) Apothéloz & Chanet (2003) introduzem o termo *informação-suporte*. Para os autores, o termo se refere à proposição recuperada pela operação discursiva que designam de nomeação.

Este artigo está assim constituído: numa primeira parte, mencionamos os aspectos teóricos que fundamentam a análise. Em seguida, descrevemos o *corpus* e a metodologia utilizada. Seguem-se a análise propriamente dita e as considerações finais. As referências bibliográficas fecham o texto.

2. Fundamentação teórica

2.1. Relevância do SN no texto escrito formal

A análise da distribuição e função de sintagmas nominais em gêneros da fala e da escrita (cf. Paredes Silva 2005, 2007a, 2007b, 2008) tem revelado a relevância dessas construções em diversos gêneros investigados, seja por sua alta incidência, seja por seu papel de principais condutores do tema/tópico discursivo dos textos. Segundo Castilho(2010), já os gramáticos gregos apontavam que “os substantivos são o fundamento do texto, pois não se pode construir um texto sem utilizar essa classe”. (Castilho, 2010:455). Tem-se observado, em especial, que gêneros jornalísticos opinativos (cf. Beltrão, 1980) como o *editorial*, o *artigo de opinião*-e os gêneros acadêmicos escritos (tais como *artigos publicados em periódicos especializados*, *resumos de teses*, e *resumos para congressos*) são terreno particularmente fértil para o estudo dos SNs complexos. Afinal, é através dos SNs que se constroem os referentes no discurso, no processo de referenciação, e tais gêneros tendem a apresentar alto grau de complexidade.

O uso de SNs - particularmente aqueles com formas nominalizadas- é associado ao texto escrito formal e ao discurso científico (cf. Basílio, 1996). Chafe(1982), um dos pioneiros no estudo das relações fala/escrita, dá especial destaque às nominalizações, ao fazer uma relação de traços opositivos das duas modalidades. O fato de permitirem a compactação de informação numa “unidade ideacional”⁽³⁾ a partir de uma base nominal é característico do texto escrito mais elaborado (cf. Chafe, 1982) ou de gêneros de fala que lançam mão de estratégias mais próprias da escrita (cf. Tannen,1982), já que esse tipo de construção exige um nível de planejamento que a fala natural (*on line*) não permite.

Assim, os SNs tem papel de destaque na continuidade tópica e na progressão temática do texto, assim como no seu propósito informativo. Os SNs que atuam como rótulos, também chamados encapsuladores, sendo compactadores, são responsáveis por mudar ou ligar os tópicos e contribuir,

(3) “O mecanismo de integração (de informação) mais característico em nossas amostras de língua escrita foi a nominalização” (Chafe: 1982, p.39)

também, na preservação da continuidade textual ao introduzir as informações novas dentro das velhas. Além disso, a escolha do SN a ser empregado revela o propósito comunicativo do produtor do texto.

Biber (1988, 1995) numa análise multidimensional dos gêneros, baseia-se em padrões de co-ocorrência de traços linguísticos computados em textos de diferentes gêneros. Entre os traços contabilizados, as *nominalizações* aparecem associadas à dimensão *informativa* do discurso.

No seu critério de dimensões ao longo das quais situa os gêneros, Biber (*op. cit.*) salienta que a escolha precisa do item lexical e a cuidadosa integração de informação são aspectos ausentes de um discurso mais interativo, com maior grau de envolvimento entre os participantes, que seria, para o autor, a contrapartida da dimensão informacional do discurso (ilustrada, pelo autor, na escrita, por uma carta pessoal, por exemplo).

Estudos anteriores sobre SNs complexos em vários gêneros do domínio jornalístico e acadêmico (Paredes Silva, 2011,2012, Bastos, 2013, por exemplo) tem revelado uma escala de complexidade, a partir de critérios como: número de itens lexicais componentes do SN, presença ou não de modificadores, número e tipo de elemento encaixado à direita, presença de nominalizações e eventuais argumentos. Foi no domínio acadêmico que se encontraram índices mais altos de complexidade de SNs.

Observe-se o exemplo abaixo, extraído de uma resenha:

- (i) O estudo apresenta *o embate ideológico das diferentes concepções sobre Reforma Agrária entre os entrevistadores e o entrevistado*.
(Melo, Rev. ANPOLL, n.9)

O SN destacado apresenta sete itens lexicais, havendo três nominalizações capazes de projetar argumentos à direita, complexificando o SN.

Assim, essa complexidade varia consoante o gênero em causa. Por exemplo, entre os gêneros jornalísticos investigados em Paredes Silva 2011, os artigos de opinião foram aqueles que apresentaram mais SNs complexos (cerca de 70%), enquanto que as crônicas, de estilo mais leve e coloquial, apresentam uma incidência bem mais baixa (30%).-

2.2. O SN na constituição do texto

Assim, se inicialmente tomamos como SNs complexos aqueles constituídos de mais elementos do que o determinante (artigo ou demonstrativo) e o nome núcleo, ou seja, aqueles que podem apresentar, além de determi-

nantes e quantificadores, modificadores à esquerda ou à direita do núcleo, constatamos que os casos de fato mais complexos são aqueles em que o SN se estende à direita, pelo acréscimo de complementos exigidos pelos itens que os constituem, itens esses identificados como nomes valenciais⁽⁴⁾. Incluem-se, ainda, os SNs de mais de um núcleo coordenado. O exemplo em itálico abaixo ilustra esses usos:

- (i) *O aumento crescente das demandas sociais e a complexidade da sociedade moderna* reduziram drasticamente *a capacidade de investimento dos estados*
(Artigo de Opinião_ O Globo)

No exemplo acima, de alguma forma o núcleo é enriquecido, seja pela presença de modificadores (crescente, sociais, moderna), seja pela presença de vários Sintagmas Preposicionados (doravante SPreps). Entretanto, não é esse o padrão predominante dos SNs que funcionam como rótulos aqui analisados, como pretendemos demonstrar a seguir.

Assim, neste trabalho estamos considerando SNs complexos aqueles que envolvem mais de dois itens lexicais ou apresentam algum tipo de encaixe, como um SPrep ou uma oração adjetiva, conforme observamos abaixo:

- (2) As greves geralmente eram organizadas pelos sindicatos. Na República Velha, foram fundados inúmeros *sindicatos*⁽⁵⁾. Mas não era nada fácil organizar um sindicato, promover uma greve. Washington Luís, que foi presidente do Brasil de 1926 a 1930, cunhou **uma frase que ficou famosa**: “*A questão social é um caso de polícia*”⁽⁶⁾. Percebeu o que ele quis dizer? Que se os trabalhadores estivessem passando fome e fizessem greve (“a questão social”), isso deveria ser resolvido pela polícia, que baixaria a paulada nos grevistas...
Schmidt (2003: 79)

No exemplo(2) acima, o SN *uma frase que ficou famosa* exerce o papel de rótulo ao necessariamente remeter ao segmento subsequente para ter seu significado compreendido.

(4) Adotamos, como Moura Neves (1996, 1999), a expressão *nomes valenciais* para nomes que apresentam uma estrutura argumental, podendo selecionar argumentos tal como um verbo.

(5) Grifo do autor.

(6) Grifo do autor.

Os rótulos apresentam outras funções importantes na organização textual, além da capacidade intrínseca, de referir ao que foi dito e/ou ao que será dito, citada no exemplo acima. Uma delas é a de mudar ou ligar os tópicos e contribuir, também, na preservação da continuidade textual ao introduzir as informações novas dentro das velhas. Koch (2003) também acrescenta que os rótulos desempenham ainda uma função cognitivo-discursiva relevante, porque ao remeter à informação-suporte, resumizam-na. Nesse sentido, muitos rótulos são expressos por nominalizações (cf. Francis ([1994]2003) e Zamponi(2003). Dessa forma, o uso de rótulos é um recurso de coesão lexical muito comum em textos escritos, principalmente de natureza argumentativa.

2.3. O livro didático e a questão dos gêneros

Como se sabe, Bakhtin (2003:262) define os gêneros do discurso como “tipos relativamente estáveis de enunciados”. Acrescenta, ainda, que há uma heterogeneidade de gêneros, mas os falantes podem discernir os gêneros mais estabilizados, pois estes estariam diretamente ligados às situações sociais.

Através da distinção que estabeleceu entre os gêneros primários (simples) e secundários (complexos), ampliou-se a concepção de gênero, que até então compreendia exclusivamente a produção literária, para integrar desde os gêneros do cotidiano à tese científica. Estariam incluídos entre os gêneros primários do discurso: o diálogo, a carta, a interação face-a-face, o bilhete, o relato cotidiano dentre outros. Já nos gêneros secundários inserem-se aqueles que estão relacionados às esferas mais complexas: romance, editorial, tese, anúncio, livro didático, palestra etc. Koch (2003:54) assinala que os enunciados por nós produzidos são situados sócio-historicamente, ou seja, estão imbricados nas situações sociais, sendo estas responsáveis pelos gêneros, os quais apresentarão características definidoras próprias e estão sujeitos a mudanças devido às transformações sociais e também a novos métodos de organização, o que se coaduna com a teoria dos gêneros de Bakhtin.

Marcuschi (2008:154) adota e defende a ideia de Bakhtin de que nos comunicamos por intermédio de textos que são realizações de um gênero. Acrescenta ainda que “em consequência, estamos submetidos a tal variedade de gêneros textuais, a ponto de sua identificação parecer difusa e aberta, sendo eles inúmeros, tal como lembra muito bem Bakhtin, mas não infinitos.” Marcuschi (2008) cita como exemplos de gêneros: sermão, carta

comercial, carta pessoal, bilhete, reportagem, aula expositiva, romance, horóscopo etc.

Ao conceito de gênero de discurso, acrescenta-se ainda a noção de *domínio discursivo*. Marcuschi (2008:23) emprega essa expressão para “designar uma esfera ou instância de produção discursiva.” Segundo o autor, citando Bakhtin, “o domínio discursivo constitui muito mais uma “esfera da atividade humana” do que um princípio de classificação de textos e indica *instâncias discursivas* (grifo do autor) (Marcuschi, 2008:155). Dessa forma, entende-se que um domínio discursivo pode dar origem a diversos gêneros. No domínio religioso, por exemplo, podemos ter: novenas, ladainhas, jaculatórias, sermões etc. Já no pedagógico, temos o livro didático, que na visão bakhtiniana poderia ser categorizado como pertencente a um gênero secundário, pois além da convivência de tipos textuais, há também a incorporação de outros gêneros sem que percam sua identidade: a charge, a transcrição de uma notícia, a banda desenhada entre outros, todos a serviço do propósito de ensinar.

A classificação do livro didático como gênero é problemática, como reconhece Marcuschi (2008), e não há unanimidade nesse ponto. Diferentemente da nossa abordagem, o autor considera o livro didático um suporte, e não um gênero. Argumenta ele que o livro, de um modo geral, é sempre suporte⁽⁷⁾. Ao mesmo tempo, porém, reconhece haver elementos muito específicos no livro didático e uma “funcionalidade típica”. Seria suporte para os vários gêneros que lá se podem abrigar. Mas esse, afinal, não é um traço dos chamados “gêneros secundários” como o romance, de acordo com Bakhtin? Além disso, há a questão do propósito comunicativo do gênero, da função primeira que desempenha, que não pode ser esquecida. Fica lançada, pois, a questão.

A multiplicidade de trabalhos desenvolvidos, nos dias atuais, em torno da questão dos gêneros, ocorre devido, principalmente, aos referenciais nacionais de ensino (Parâmetros Curriculares Nacionais) que enfatizam a importância do estudo dos gêneros textuais, os quais são considerados imprescindíveis na formação do aluno como leitor/ produtor de textos. Para Beth Marcuschi & Costa Val (2008:9) “trabalhar adequadamente um gênero seria levar os alunos a considerar seu suporte, sua esfera de circulação e os leitores a que se dirige.”

(7) “Entendemos aqui como suporte de um gênero um *lôcus* físico ou virtual com formato específico que serve de base ou ambiente de fixação de um gênero materializado como texto” (Marcuschi, 2008:174)

3. *Corpus e metodologia*

A partir da constatação de que os livros didáticos de Língua Portuguesa têm sofrido mudanças a fim de se adequar aos debates em torno do ensino de língua materna, às contribuições trazidas pelos Parâmetros Curriculares Nacionais (PCNs) bem como aos subsídios fornecidos pelas avaliações sistemáticas do MEC⁽⁸⁾, consideramos pertinente analisar um *corpus* constituído por uma coleção de livros didáticos de História do Brasil, uma vez que os PCNs de História ressaltam que um dos objetivos do ensino dessa disciplina é o contato com o texto, o desenvolvimento da capacidade interpretativa. Portanto, o *corpus* deste artigo é constituído da uma coleção de livros didáticos de História do Brasil – *Nova História Crítica*, de Mário Furlley Schmidt, organizada em quatro volumes, destinados ao 6º, 7º, 8º e 9º anos do Ensino Fundamental.

A escolha dessa coleção foi motivada pelo fato de ser amplamente adotada na rede de ensino da cidade do Rio de Janeiro e, por essa razão, ser um dos livros didáticos mais vendidos na atualidade⁽⁹⁾. No entanto, uma polêmica formou-se em torno da coleção ao ser reprovada, na avaliação do MEC no ano de 2007, para alunos de Ensino Fundamental (5ª a 8ª séries) e aprovada para os alunos de Ensino Médio, já que as coleções foram avaliadas por equipes distintas, provenientes de diferentes universidades⁽¹⁰⁾.

Para realização deste artigo, utilizamos parte dos dados coletados para a tese de Bezerra 2010, que consistiu em uma análise comparativa, abrangendo três períodos da história do Brasil - pré-regime militar, período do regime militar e período contemporâneo. Optamos aqui por utilizar somente um dos autores devido à exiguidade de espaço.

A partir da leitura dos capítulos correspondentes aos principais temas recorrentes em livros de História do Brasil, identificamos, conforme nosso objetivo inicial, duzentos e nove SNs que funcionavam como rótulos nos textos de Schmidt, já que atendiam ao critério proposto por Francis

(8) Ministério da Educação e Cultura

(9) Desde 1998, quando esse livro passou a integrar a lista de recomendação do MEC, cerca de 28 milhões de livros foram vendidos para crianças e adolescentes do Ensino Fundamental e Médio de todo o país, segundo parecer da editora Nova Geração.

(10) A primeira coleção foi avaliada pela Universidade Federal do Rio Grande do Norte e a segunda pela Universidade Federal Fluminense. Houve, então, o questionamento de que a reprovação envolveria aspectos políticos, e não apenas pedagógicos. Essa discussão acabou ganhando as páginas dos jornais em setembro de 2007, despertando o interesse do público em geral, e chegou a ser debatida durante semanas em jornais e telejornais.

(1994:192) “... um elemento nominal inerentemente não-específico cujo significado no discurso necessita ser precisamente decifrado.”

A fim de analisarmos esses duzentos e nove rótulos identificados no *corpus*, usamos, ao mesmo tempo, uma metodologia quantitativa e qualitativa, examinando-os sob vários aspectos: seu caráter retrospectivo ou prospectivo, a presença ou não de modificador, o determinante empregado, a semântica do nome-núcleo e a localização desse SN no texto.

Para aferir a importância e simultaneamente a correlação entre esses aspectos, utilizamos uma análise quantitativa que, através dos cálculos percentuais, atribuiu maior confiabilidade e permitisse lidar melhor com os dados e suas características, conduzindo a afirmações empiricamente comprovadas.⁽¹¹⁾

4. Análise dos rótulos

Nos livros didáticos de História do Brasil analisados, identificamos que o rótulo pode ser constituído somente do nome-núcleo ou ser precedido de determinante e/ou modificador. Vejamos o esquema abaixo:

- Nome
- Determinante + Nome
- Determinante + Modificador(es) + Nome+ Modificador(es)

Conforme se sabe, o modificador pode ocorrer tanto à esquerda quanto à direita do nome-núcleo: à esquerda temos o adjetivo e à direita podemos ter um adjetivo, SPrep ou uma oração relativa. Estamos designando como *determinante*, de modo bastante amplo, o elemento à esquerda que precede o nome-núcleo, mas não o qualifica tal como faz um adjetivo⁽¹²⁾. Abaixo, observamos o rótulo constituído somente do nome-núcleo.

- (3) Tudo isso nos revela que **idéias** como “os negros se submetiam com mais facilidade que os índios” e “os negros sempre foram mais passivos, aceitando humildemente sua situação” não passam de um preconceito idiota (e será que existe algum preconceito que não seja idiota?).
Schmidt (2003:208)

(11) Para esses cálculos, lançamos mão de parte do pacote estatístico GOLDVARB.

(12) Estamos utilizando o termo determinante de forma bastante abrangente, para incluir artigos, possessivos, demonstrativos, à semelhança dos especificadores de Castilho (2010:454.)

- (4) Você notou como iam se acumulando **motivos** para o exército rejeitar o império?⁽¹³⁾
Schmidt (2003:310)

A análise do *corpus* evidenciou, contudo, que o rótulo na sua configuração mínima, isto é, constituído apenas do nome nuclear, é pouco frequente nos dados, havendo apenas três ocorrências. No primeiro exemplo, temos um rótulo prospectivo na forma plural - *idéias*-, que remete a uma série de informações colocadas entre aspas pelo autor do texto, o que, de certa forma, contribui para que o leitor localize as informações e, ao mesmo tempo, o exime da responsabilidade de incorporá-las. Já no segundo exemplo, o rótulo *motivos* insere-se em um contexto que trata da relação existente entre os militares e a queda do império, retomando todas as informações anteriores à ocorrência do nome-núcleo. Como observamos nos dois exemplos acima, o rótulo, por ser inerentemente inespecífico, remete a outras porções textuais para a sua realização lexical, funcionando anafórica e/ou cataforicamente.

- (5) Os positivistas sonhavam com a “ditadura dos cientistas”, e alguns militares sonhavam com a “ditadura militar”. Os fazendeiros paulistas adotavam o liberalismo político inspirado no regime dos EUA e no darwinismo social, uma espécie de “cada um por si no vale-tudo do mercado”. Os republicanos radicais imaginavam que o novo regime faria como o Brasil o que a revolução de 1789 tinha feito com a França: iria estabelecer a liberdade e a igualdade, garantir os direitos dos cidadãos.
Todas essas questões eram importantes.
Schmidt (2003:37)

- (6) O papa Pio IX era contrário à maçonaria por **dois motivos**: ele não aceitava as idéias místicas dos maçons nem suas propostas políticas liberais.
Schmidt (2003:311)

Quanto ao exemplo (5) notamos que, na primeira parte do texto, é apontada uma série de questões sobre o sistema de governo. A presença do quantificador *todas* mostra que a retomada diz respeito à totalidade do que foi exposto. Além disso, o demonstrativo *essas*, dentre outros aspectos aponta a localização na porção textual das questões abordadas.

(13) O rótulo *motivos* remete a toda seção em que aparece inserido, por isso optamos por transcrever somente o trecho no qual ele ocorre.

Já os numerais fazem parte da configuração de muitos rótulos prospectivos, como no exemplo (6) acima, pois, ao funcionarem como organizadores textuais, fornecem pistas ao leitor de como o texto será estruturado, servindo como diz Francis ([1994] 2003:223) para “... sequenciar estágios de um argumento...”. Por isso, o uso do numeral *dois* como parte integrante deste rótulo, antecipa ao leitor a quantidade de informação que ele deve procurar na sequência, para que possa entender a rejeição do Papa Pio IX à maçonaria. Percebe-se, então, nesses casos, que os determinantes que acompanham os nomes-núcleo adicionam informações a ele, assim como os modificadores, conforme os exemplos a seguir:

- (7) A escravidão estimulou **pensamentos racistas**. Os brancos olhavam para a situação dos escravos e chegavam a **uma conclusão absurda**: “Eles foram escravizados porque são seres humanos inferiores aos europeus”. Daí foi um passo para acreditarem que negros e índios faziam parte de uma raça inferior que merecia ser dominada pelos brancos. **Essa idéia falsa** se espalhou no século XIX, e até hoje, na entrada do século XXI, ainda tem gente que acredita nela.
Schmidt (2003:214)
- (8) As greves geralmente eram organizadas pelos sindicatos. Na República Velha, foram fundados inúmeros *sindicatos*⁽¹⁴⁾. Mas não era nada fácil organizar um sindicato, promover uma greve. Washington Luís, que foi presidente do Brasil de 1926 a 1930, cunhou **uma frase que ficou famosa**: “*A questão social é um caso de polícia*”⁽¹⁵⁾. Percebeu o que ele quis dizer? Que se os trabalhadores estivessem passando fome e fizessem greve (“a questão social”), isso deveria ser resolvido pela polícia, que baixaria a paulada nos grevistas...
Schmidt (2003: 79)
- (9) E o que a Inconfidência Mineira e a Conjuração Baiana tinham de diferente? Para começar, **a situação das pessoas** nas duas regiões. Em Minas Gerais, o tormento era a decadência do ouro e os impostos absurdos cobrados por Portugal. Você se lembra de que os inconfidentes queriam que a revolta acontecesse no dia da derrama? Na Bahia, o problema era outro. Você se recorda que no final do século XVIII, os preços internacionais do açúcar subiram. Então, muitos senhores de engenho voltaram a plantar bastante cana.

(14) Grifo do autor.

(15) Grifo do autor.

Schmidt (2003: 101)

No exemplo (7), temos a ocorrência de três rótulos, os quais têm seus nomes nucleares modificados por adjetivos, que imprimem nesse contexto um valor axiológico negativo. Já em (8), o nome-núcleo *frase* é modificado pela oração relativa *que ficou famosa*, o que confere ao SN alguma complexidade. No exemplo (9), o termo *situação* é modificado pelo sintagma preposicionado *das pessoas*.

O emprego de modificadores, conforme a Tabela 2 abaixo, ocorre em 44% dos dados analisados; nos demais casos, que representam 56% dos dados, o nome-núcleo aparece sozinho ou acompanhado de determinante, o qual também tem um papel de destaque no emprego dos rótulos devido principalmente ao caráter organizador.

Tabela 2: Frequência do uso dos modificadores em Schmidt

Sem modificador	118=56%
Com modificador	91=44%
Total	209

Já a Tabela 3 nos mostra a posição dos modificadores e ressalta que a posição à direita do nome-núcleo é ocupada em 78% dos casos. No entanto, é preciso destacar que o uso do adjetivo é favorecido, pois há apenas seis ocorrências de SPrep e três de orações relativas no *corpus* analisado. O fato de termos modificadores expressos num único item lexical e preferencialmente à direita, sua posição canônica, confere uma dimensão de “baixa complexidade” a esses SNs. Veja-se, inclusive, que em apenas 2% dos casos temos modificadores nas duas posições.

Tabela 3: Posição do modificador em relação ao nome-núcleo

Direita	71=78%
Esquerda	18=20%
Direita e esquerda	2=2%
Total	91

A partir da análise do exemplo (7), percebemos que, ao avaliar ou mesmo reavaliar a informação-suporte, o modificador pode atribuir um valor axiológico, quer negativo quer positivo ao nome-núcleo, tal como

sinaliza Conte (2003:181) ao discutir o fenômeno do encapsulamento anafórico. Segundo a autora, a avaliação dos conteúdos do contexto pode tanto ser operada pelo nome-núcleo quanto pelo adjetivo avaliativo. Vejamos alguns exemplos adicionais:

- (10) Você sacou a relação existente entre Estado absolutista, mercantilismo, apoio do Estado à burguesia e expansão marítima? Agora, **uma informação preciosa**: o primeiro Estado absolutista a se formar na Europa foi Portugal! A conclusão que você pode tirar é que Portugal foi pioneiro na expansão marítima porque foi o primeiro país europeu a construir um Estado absolutista capaz de apoiar as navegações.
Schmidt (2003:97)

Destacamos, no exemplo (10), o alto valor atribuído pelo modificador *preciosa*, dado ao conteúdo subsequente, a fim de que o leitor possa concluir o motivo que levou Portugal ao pioneirismo na expansão marítima. Em oposição, temos o exemplo seguinte (11), no qual o rótulo *amarga surpresa*, responsável por introduzir uma avaliação negativa, antecipa ao leitor que a informação-suporte não corresponde à expectativa de Cabral.

- (11) Do Brasil, Cabral seguiu para a Índia, seu principal objetivo. Mas os árabes já estavam informados do sucesso de Vasco da Gama e, então, pressionaram as autoridades indianas a não comerciarem com os portugueses. Quando Cabral chegou à Índia, teve **a amarga surpresa**. Os indianos não queriam comerciar. Preferiam os mercadores árabes. E agora? O que fazer? Viajar tanto tempo para nada?
Schmidt (2003:107)

Portanto, como podemos observar nos exemplos acima, o adjetivo adiciona significados ao nome-núcleo, nesses casos atribuindo um caráter avaliativo – positivo/negativo. O uso de modificadores tem um papel relevante, já que o rótulo não pode ser considerado apenas a partir do seu nome-núcleo, mas sim por todos os elementos que o constituem, ou seja, o SN como um todo.

Entretanto, Lapa (1970:107) nos assinala a necessidade de termos cuidado com o uso dos adjetivos, já que alguns adjetivos não acrescentam conteúdos por serem um “caracterizador banal que serve para tudo”.

O exemplo abaixo ratifica o comentário de Lapa (1970) ao ter como modificador o adjetivo *interessante*. Segundo Lapa (*op. cit.*), é um adjetivo a evitar. Por que teria Schmidt (2003) usado, então, esse recurso? Precisa-

mos ter em mente o propósito do *corpus* sob análise – ensinar – e o público alvo – alunos do ensino fundamental - que são leitores em formação, para compreender as razões do autor, ou seja, este se preocupou em empregar uma linguagem que chamasse a atenção do seu público leitor a fim de tentar atraí-lo.

- (12) **Uma pergunta interessante:** por que os brasileiros nordestinos não foram utilizados nas fazendas de café do Sudeste?
Schmidt (2003:287)

Também Francis ([1994] 2003:217) discute essa questão ao mostrar que alguns adjetivos acrescentam pouco conteúdo ao núcleo que acompanham. Por exemplo, para a autora, o adjetivo *básica* não seria a melhor opção para acompanhar o nome-núcleo *verdade*, pelo fato de não existirem “...verdades que não são básicas...”. Assim parece que os dois autores criticam o uso do lugar comum. Vejamos o exemplo que a autora apresenta:

A pressão determina o ritmo e, embora possa ser previsível por uma boa pesquisa de opinião, ela não nos diz nada sobre qualquer eleição que não seja a de hoje em Monmouth. Contudo, ninguém tem interesse *nessa verdade básica*. A comunidade política está viciada em toda corrida de cavalo que puder encontrar.
(Francis, *op. cit.*)

A partir da análise dos constituintes dos rótulos apresentada acima, podemos considerar que tais SNs podem apresentar certa complexidade (exemplos (7) e (8)), já que “entende-se como SNs complexos aqueles que, além de determinante e núcleo, apresentam modificadores e/ou complementadores. (Paredes Silva, 2011). Entretanto, comparativamente com outros gêneros, o seu grau de complexidade é relativamente baixo. A natureza coloquial do texto de Schmidt(2003) é um dos fatores que poderia ser apontado como responsável pela incidência de SNs de baixa complexidade entre os dados.

No tripé (conteúdo temático, estrutura organizacional, estilo), base do gênero, segundo Bakhtin, esse coloquialismo caracterizaria o estilo, as escolhas lexicais, como se percebe nos exemplos a seguir. Trata-se de estratégias utilizadas pelo autor nesta busca pela aproximação e atração do leitor para o conteúdo veiculado. Por exemplo, temos o uso de gírias (*mixuruca*, *sacou*), o uso de frases feitas (*tudo está bem quando acaba bem*) e o uso do marcador discursivo **bem** que, segundo Risso (1999:260), é comum em textos da fala, exercendo a função de iniciar um turno de resposta em estru-

ras conversacionais, introduzindo o que será dito a seguir. Isso confere um tom de conversa ao texto escrito.

Puxa, se era tão vantajoso o emprego do trabalho livre, por que então o Brasil teve escravos?

Schmidt (2003:289)

Vão trabalhar a vida toda, sem parar. No final da vida, velhinhos, com um salário de aposentadoria bem **mixuruca**, ainda terão de ouvir: “Esse velho é pobre porque nunca foi chegado ao trabalho”..

Schmidt (2003:208)

Perceba um aspecto muito importante: os italianos e os alemães que foram para o Sul tiveram o privilégio de ir para a única região onde ocorreu uma razoável distribuição de terras... **sacou?**

Schmidt (2003:286)

Os próprios fazendeiros participavam das comemorações. Era como dissessem: **tudo está bem quando acaba bem.**

Schmidt (2003:296)

Bem, você sabe que, se o escravo fizesse corpo mole, o capataz dava logo uns tapas no coitado.

Schmidt (2003:212)

Schmidt (2003)⁽¹⁶⁾ busca contar os vários ângulos do fato que é exposto, a fim de construir com o aluno a sua “versão” da História, como nos diz o autor no manual do professor nos trechos abaixo (p.14-15):

Como podemos estimular os alunos a estudar História, se os livros didáticos permanecem com uma linguagem seca, complicada, monótona? Estilo pesado, talvez encobrindo alguma frustração acadêmica e incapacidade de revelar ao estudante as conexões do estudo com suas aspirações cotidianas. Como é que alguém pode ter coragem de escrever um livro para 5ª ou 6ª série – para meninas e meninos de 10 a 12 anos de idade, ainda no começo da sua vida intelectual – com uma linguagem apropriada às teses de doutorado? [...]

Por isso, não hesitamos em utilizar uma linguagem em estilo próximo do coloquial. Frases curtas, vocabulário simples (mas não ingênuo ou empobrecido), estilo dinâmico, tudo aquilo que dê a impressão de estarmos “batendo um papo” com o leitor. O livro deve ser amigo, aquele a quem nos dirigimos, com

(16) A partir de conversas informais com professores de História do Brasil e com os próprios alunos, observa-se que seus livros têm uma boa receptividade entre público jovem.

quem concordamos e discordamos, e que enxergamos de maneira diferente a cada leitura.

Diante da necessidade apontada acima pelo próprio Schmidt(2003) de que seu livro didático⁽¹⁷⁾ tenha uma linguagem próxima à de um adolescente, identificamos dentre os dados analisados apenas quatorze SNs de maior complexidade, os quais, apesar serem categorizados dessa forma, não apresentam uma carga informativa tão densa quanto um texto escrito formal, devido principalmente à ausência de nomes valenciais que causariam o prolongamento do SN à direita. Vejam-se os exemplos: o primeiro do *corpus* em análise e o segundo de Paredes Silva(2011), já mencionado em 2.2.

(13) Há **uma coisa que chama bastante a atenção**: o mapa do século XVI mostra muitas coisas que o mapa do século XV simplesmente ignorava.
Schmidt (2003:93)

(i) *O aumento crescente das demandas sociais e a complexidade da sociedade moderna* reduziram drasticamente a capacidade de investimento dos estados.
(Opinião_ O Globo)

Além disso, como o rótulo, geralmente, está localizado em pontos que Conte(2003) chama de “nodais”, como o início de um parágrafo, a plena compreensão do rótulo torna-se fundamental para que o leitor, principalmente esse leitor em formação, identifique informações essenciais no desenvolvimento do texto. Esse fato também justificaria o favorecimento do emprego de SNs de menor complexidade no papel de rótulos no *corpus* analisado, que apresenta um caráter mais informal, visando à aproximação de seu leitor. No exemplo(6), retomado como (14) abaixo, vimos que o rótulo *dois motivos* ocorre em início de parágrafo, preparando o leitor para o que será desenvolvido a seguir, como uma estratégia de organização do texto.

(14) O papa Pio IX era contrário à maçonaria por **dois motivos**: ele não aceitava as idéias místicas dos maçons nem suas propostas políticas liberais.
Schmidt (2003:311)

(17) Isso nos mostra a capacidade do gênero livro didático de se modificar e reorganizar conforme as mudanças sociais, pois, em outras coleções analisadas em Bezerra(2010) de décadas anteriores, não se evidenciava a preocupação com o público leitor.

Podemos observar que nos trechos analisados, há variados SNs complexos, mas não possuem o traço característico do rótulo: ser inerentemente inespecífico. Vejam-se os exemplos abaixo:

A cidade havia crescido, as ruas agora eram iluminadas por lampião a gás, em vez do **velho lampião fedorento a óleo de baleia**, havia bondes puxados por burros, lojas e mais lojas. A limpeza urbana tinha começado a ser feita pela firma Aleixo Gary, que possuía empregados (lixeiros) logo apelidados de garis. Apareciam fábricas, bancos, companhias de seguro, estradas de ferro, empresas capitalistas. Desde 1850 havia **uma linha regular de navios a vapor ligando o Rio a Londres**. Era a modernidade chegando de mansinho na capital do Império.

Schmidt (2003: 292)

Observamos que os SNs complexos *o velho lampião fedorento a óleo de baleia e uma linha regular de navio a vapor ligando o Rio a Londres*, em comparação com o exemplo apresentado em Paredes Silva(2011) acima apresentam quanto à semântica e à sintaxe do nome-núcleo um menor grau de complexidade, pois privilegiam o emprego de nomes designadores de pessoas, lugares e objetos; no lugar de nominalizações, as quais possuem densa carga informativa e, ao mesmo tempo, requerem uma maior capacidade interpretativa.

O baixo grau de complexidade dos SNs que funcionam como rótulos em livros didáticos encontra paralelo nos artigos de divulgação científica. Em dissertação recente, Bastos(2013) estudou SNs complexos em exemplares de duas revistas de divulgação científica⁽¹⁸⁾ vendidas em bancas de jornais do Rio de Janeiro: Superinteressante e Galileu⁽¹⁹⁾. Apesar da alta incidência de nominalizações (mais de 50% dos casos) raramente explicitavam os argumentos projetados. Tal como os livros didáticos, a proposta dessas revistas é atingir um público jovem, com assuntos (temas) da atualidade. Seguem-se dois exemplos. No primeiro deles, de tipo menos frequente, SNs complexos cujo núcleo são formas nominalizadas projetam seus argumentos, estendendo o SN para a direita:

(18) Seu estudo, baseado em artigos de capa de 20 revistas, não se restringiu a SNs /rótulos, mas compreendeu SNs em função de sujeito e de objeto, alguns dos quais funcionavam como rótulos. Seus objetivos eram, além de verificar o grau de complexidade dos SNs de tais artigos, correlacionar esse aspecto com a caracterização do gênero em questão. (cf. Bastos, 2013)

(19) As revistas escolhidas são aquelas de maior tiragem e aceitação pelo público jovem.

- (i) Apesar de existir *uma variação diária nas concentrações de hormônios masculinos e femininos (...)* não existe *nenhuma relação comprovada entre esse sobe e desce e picos diários de fertilidade ou libido*
Bastos(2013:55)

Já no segundo exemplo, do tipo mais frequente, apesar de haver uma nominalização funcionando como elemento encapsulador -*controle*, nominalização essa que poderia projetar um ou mais argumentos, estes se depreendem do co-texto e não se explicitam, tornando o SN menos complexo:

- (i) Contra esse tipo de atraso, psicólogos (...) têm se rendido ao que eles chamam de ‘vontade estendida’ – ferramentas externas que ajudam a nossa parte que quer trabalhar. Um estudo do M.I.T. mostra que esse *controle* não precisa vir necessariamente de cima para baixo. Confrontados com as opções de entregar todos os trabalhos no fim do semestre ou acertar datas de entrega diferentes, a maioria preferiu a segunda alternativa. Moral da história: em vez de se expor à vagabundagem involuntária, eles preferiram [um *controle* externo] para fazer o que racionalmente queriam – estudar.”
Bastos(2013:55)

Ao mesmo tempo, artigos acadêmicos publicados em periódicos especializados, pertencentes ao domínio acadêmico; ou artigos de opinião, do domínio jornalístico, conforme os dois exemplos que se seguem, respectivamente, são aqueles em que mais se destacam os SNs complexos, fazendo valer a máxima de Chafe(1982) sobre a importância das nominalizações no texto escrito formal e na função de compactação de informação.

- (i) *As duas seções que se seguem* dedicam-se, respectivamente, a expor e discutir *esses dois movimentos da filosofia da linguagem de Wittgenstein: a elicitación e a crítica de uma determinada compreensão hegemônica da linguagem e do significado..(Sobre a estabilidade do significado em Wittgenstein)*
(Artigo publicado na Revista Veredas- UFJF)
- (ii) Lula comemorou com uma festa *a doação de R\$13 milhões dos “meus amigos da Febraban” para as cisternas do Fome Zero*
(Opinião – Chega de sábios, precisa-se de bravatas – Globo)

Observe-se, no último exemplo acima, um caso raro: uma nominalização que projeta todos os argumentos do verbo relacionado, externo e

internos: X (“meus amigos da Febraban”) doa Y (R\$13 milhões) a Z (para as cisternas do Fome Zero).

Por outro lado, gêneros jornalísticos de estilo mais informal, como a crônica, também se destacam por explorar preferencialmente os modificadores na constituição do SN. No entanto, rótulos também ocorrem. Vejam-se os exemplos, respectivamente:

- (i) (...) pediu um prato de queijo minas sem tempero.
(Crônica_ Arthur Xexeo O Globo)
- (ii) Há *uma civilização carioca de delícias que aos poucos se esvai*.
(Crônica_ Joaquim Ferreira dos Santos- O Globo)

É certo que em tais estudos não nos restringimos ao exame dos rótulos, mas dos SNs complexos, de um modo geral. Entretanto, é possível fazer algumas comparações, que auxiliam na identidade dos gêneros de discurso e domínios discursivos examinados.

5. Considerações Finais

Observamos no *corpus* de livros didáticos analisado que há SNs os quais atuam dentro do texto retomando ou apontando informações, sendo, então, responsáveis por organizar as partes em que se inserem, bem como introduzir simultaneamente objetos de discurso. Tais SNs assumem o papel de rótulos.

Ao verificar a constituição dos rótulos, notamos que estes podem ter apenas o nome-núcleo, assim como vir acompanhado de determinante e/ou modificador. Diante da primeira possibilidade, os que se apresentam com mais de dois itens lexicais já indicariam certo grau de complexidade.

A análise dos dados nos mostrou que, dentre os rótulos identificados, a maior parte não poderia ser categorizada como SNs complexos, já que 56% dos dados eram constituídos apenas de nome-núcleo ou nome-núcleo acompanhado de determinante. Desta forma, o *corpus* analisado apesar de ser constituído por texto escrito, revelou uma baixa complexidade quanto ao emprego dos rótulos, o que pode ser relacionado à própria função organizadora do rótulo no texto. A partir dessa constatação, vimos também que os rótulos eram constituídos em sua maioria de nomes-núcleos genéricos e os modificadores empregados visavam antes a estabelecer valores axiológicos, relevantes na condução de um discurso argumentativo.

Na comparação com o gênero mais próximo⁽²⁰⁾ (em termos de propósito comunicativo e público alvo) o artigo de divulgação científica, também encontramos baixo grau de complexidade, em termos de tendência ao prolongamento à direita do SN, com Spreps ou orações encaixadas. Diferenciam-se, entretanto, pela alta incidência de nominalizações, neste último, muitas delas funcionando como rótulos.

O SN complexo é sem dúvida um recurso da escrita formal, comprovando Chafe (1982), mais presente em textos dirigidos a leitores mais proficientes. Por outro lado, aqueles que se aproximam de artigos em periódicos especializados, resenhas de livros, *abstracts* de teses supõe-se que sejam leitores familiarizados com esses recursos e preparados para enfrentá-los, diferentemente dos jovens que pretendemos conquistar para a leitura, nesses tempos de reinado da informática.

6. Referências

- APOTHELOZ, D. & CHANET, C. (2003), Definido e demonstrativo nas nomeações. In: CAVALCANTE, M., BIASI-RODRIGUES, B; CIULLA, A. (orgs.) *Referenciação*. São Paulo: Contexto, p.131-76.
- BASILIO, M. (2004). *Formação e classe de palavras no português do Brasil*. São Paulo, Contexto.
- BASTOS, M.Ximenes (2013), *O uso de Sintagmas Nominais complexos em artigos de divulgação científica* Dissertação de Mestrado em Linguística . Rio de Janeiro,UFRJ.
- BAKHTIN, Mikail (2003), *Estética da criação verbal*. 4ed. São Paulo: Martins Fontes.
- BELTRÃO, L.(1980), *Jornalismo opinativo*. Porto Alegre, Sulina.
- BEZERRA, Gabrieli Pereira (2010), *Sintagmas nominais como rótulos em livros didáticos de História do Brasil*. Tese de Doutorado em Linguística. Rio de Janeiro: UFRJ.
- BIBER, D. (1998), *Variation across speech and writing*. Cambridge:Cambridge University Press.
- CASTILHO,A.(2010), *Nova Gramática do Português Brasileiro*. São Paulo, Contexto.
- CHAFE, W. (1982), Integration and involvement in speaking, writing and oral literature. In: TANNEN,D. (ed) *Spoken and written language: exploring orality and literacy*. Norwood, N.J. Ablex,p.35-53.
- CONTE, M. (2003), Encapsulamento anafórico. In: CAVALCANTE, M. , RODRIGUES, B & CIULLA, A (orgs.).*Referenciação*. São Paulo: Contexto, p.177-190.

(20) Considerando-se aqueles gêneros que já foram investigados sob a mesma perspectiva, da complexidade dos SNs.

- FRANCIS, Gill (2003), Labelling discourse: an aspect of nominal-group lexical cohesion. In: COULTHARD, Malcolm. *Advances in Written text analysis*. London: Routledge, p.83-101, 1994./ Trad. Monica Cavalcante *et al.*;revisão Alena Ciulia. In: CAVALCANTE, M., RODRIGUES, B & CIULLA, A (orgs.).*Referenciação*. São Paulo: Contexto, p.191- 228.
- KOCH, Ingedore (2003), *Desvendando os segredos do Texto*. São Paulo: Cortez.
- LAPA, M.R. (1970) *Estilística da Língua Portuguesa*. Porto Alegre: Acadêmica.
- MARCUSCHI, B. & COSTA VAL, C. (2008), Gêneros textuais no espaço extra-escolar e na sala de aula. In: *Na ponta do lápis*. São Paulo, n° 9, jun, p. 8-9.
- MARCUSCHI, Luiz Antonio (2008),*Produção Textual, análise de gêneros e compreensão*. São Paulo: Parábola Editorial.
- Parâmetros curriculares nacionais: história, geografia/ Secretaria de Educação Fundamental* (1997),– Brasília: MEC/SEF.
- PAREDES SILVA, V.L. (2005), *Referencia(ção), repetição e gêneros jornalísticos*. Participação em mesa-redonda do XIII Encontro da Assel-Rio. UFF, outubro/2005.(mimeo).
- .(2007a), A continuidade de referência em gêneros da escrita e da fala no português brasileiro. In: LOBO, M & COUTINHO, M. A (orgs) *XXII Encontro Nacional da Associação Portuguesa de Lingüística. Textos Selecionados*. Lisboa, APL.
- .(2007b), Continuidade de referência: nomes, pronomes e anáfora zero em gêneros da escrita e da fala. *Revista Lingüística n. 3* Programa de Pós graduação em Lingüística da UFRJ.
- .(2008), Desfazendo um mito: a repetição na escrita e suas funções. In:RONCARATI, C.& ABRAÇADO, J.O *Português Brasileiro 2* .Niterói, EDUFF, p.334-344.
- .(2011), *O uso de sintagmas nominais complexos em gêneros jornalísticos e acadêmicos*.
- Palestra ministrada no Curso de Pós-graduação em Educação da Universidade do Minho. Braga,17 de junho de 2011. (17 p. mimeo)
- Risso, Mercedes Sanfelice (1999), Aspectos textuais-interativos dos marcadores discursivos de abertura *bom, bem, olha, ah*, no português culto falado. In: NEVES, Maria Helena de Moura(org.). *Gramática do Português Falado*. São Paulo: Humanitas/FFLCH/USP; Campinas: Editora da Unicamp, p. 259-296.
- ZAMPONI, Graziela (2003), *Processos de referenciação: anáforas associativas e nominalizações*. Tese de Doutorado. Campinas: UNICAMP.

Corpus

- SCHMIDT, Mario Furley (2003), *Nova História Crítica .v.1-4*. São Paulo: Nova Geração.

“COMPONENTE” COMO SUBSTANTIVO UNIFORME DE DOIS GÊNEROS

Iva Svobodová*
iva.kilianova@tiscali.cz

O presente estudo tem como principal objetivo analisar o género gramatical da palavra *componente*. A necessidade de identificar o género deste substantivo prende-se com o facto de ter aumentado, nas últimas décadas, a frequência de uso nas áreas científica, política, cultural, social, industrial, informática e ainda outras. As nossas dúvidas têm-se projetado cada vez mais no plano de praticamente todas as disciplinas linguísticas (morfologia, sintaxe, semântica e lexicologia). A conclusão da nossa pesquisa indica que os fatores mais decisivos na interpretação do género do nome em questão são os diafásico e semântico. Verificámos, ao mesmo tempo, uma parcial influência da frequência de uso. O nosso estudo baseia-se sobretudo em dois pilares: a linguística *de corpora*, que contribuiu para a criação do material estudado, e a extensa investigação de Maria Carmen Frias e Gouveia, que forneceu os pontos-chave do quadro metodológico da nossa investigação.

Palavras chave: *componente*; género gramatical; Maria Carmen Frias e Gouveia.
Key-words: *componente*; gramatical gender; Maria Carmen Frias e Gouveia.

*

1. Notas introdutórias

A presente análise tem como principal objetivo fazer uma profunda reflexão sobre uma série de questões que se relacionam com o género gramatical da palavra *componente*. O nosso estudo baseia-se em pesquisas de Maria Carmen de Frias e Gouveia que recordou ser „o género gramatical um dos

* Departamento das Línguas e Literaturas Românicas da Faculdade de Letras da Universidade de Masaryk, Brno, República Checa.

aspectos da gramática mais merecedor de atenção, apresentando-se com grande vitalidade e de capital importância comunicativa na língua portuguesa“ (Corbett: 1991 *apud*: Frias e Gouveia: 2007, 2011). Frias e Gouveia, que trabalhou de uma forma minuciosa todos os possíveis aspectos relacionados com o emprego do género gramatical, focalizou, entre outros, as dificuldades que este pode acarretar no caso de palavras com uma certa oscilação do género gramatical e que originam um grande espectro de questões, que se mostram relevantes para o uso do género gramatical: as diferenças entre as variedades europeia e americana do Português (Frias e Gouveia: 2007), o fator histórico (Frias e Gouveia: 1998, 1999, 2004, 2005, 2006), o fator diafásico e estilístico-pragmático (Frias e Gouveia: 1998) e o fator lexicológico relacionado com o género dos estrangeirismos (Frias e Gouveia: 2004), entre outros. O que nos levou a conhecer a obra de Maria Carmen de Frias e Gouveia foi a necessidade de identificar o género da palavra *componente* porque, por mais que custe acreditar, mesmo depois de consultar e estudar atentamente os dicionários de língua portuguesa mais prestigiados, não conseguimos identificá-lo univocamente. A necessidade de encontrar o género desta palavra prende-se com o facto de ter aumentado a frequência do seu uso nas áreas científica, política, cultural, social, industrial, e ainda outras. Para além disso, as nossas dúvidas têm-se projetado cada vez mais no plano de todas as disciplinas linguísticas (morfologia, sintaxe, semântica e lexicologia).

Os métodos que M. C. Frias e Gouveia aplicou ao estudo do género gramatical serviram-nos de modelo de análise e contribuíram para a fiabilidade e complexidade da nossa pesquisa. O problema da palavra *componente*, como veremos mais adiante no enquadramento teórico, encaixa-se no tema dos nomes uniformes de dois géneros, e traz mais um exemplo de dificuldades que podem originar incompatibilidades gramatical e lexical no eixo sintagmático, geradas pelas oposições a nível tanto semântico como diatópico nas variedades brasileira e europeia. (Frias e Gouveia: 2011).

Por maior que seja o valor e o prestígio dos dicionários consultados, pensamos que a presente análise poderá contribuir para a área da lexicografia portuguesa. Neste sentido, estamos completamente de acordo com Frias e Gouveia que aponta, ao longo dos seus trabalhos, para uma heterogeneidade e oscilação que existe na indicação do género gramatical pelos dicionários de língua portuguesa (*idem*: 2011).

2. Enquadramento teórico

O tema do nosso trabalho, como já foi referido, diz respeito ao problema do género gramatical dos substantivos uniformes. Além dos *a*) substantivos uniformes de um género, que são formados, sobretudo, por nomes inanimados e cujo género gramatical evoluiu do latim, existem *b*) substantivos uniformes de dois géneros, chamados ‘sobrecomuns’ ou ‘comuns de dois’ (que possuem o género natural), o que se abona no caso dos seres sexuados (*o/a cônjuge, o/a pianista*) e *c*) os substantivos ‘épícenos’, os quais possuem apenas um género para ambos os sexos. Nos três grupos existe um conjunto de palavras de género hesitante, originado por fatores históricos (*a eczema x o eczema*), geográficos (*PB o grama x PE a grama*) e diafásicos (*as jeans x os jeans*). Além disso, verifica-se, na linguagem popular, a tendência para biformizar alguns nomes como, por exemplo, *a melra, a gaia, a chefe* (Frias e Gouveia:1998). Fora dos três grupos referidos, geralmente, menciona-se mais um: o das *d*) palavras uniformes, cujo sentido está estreitamente vinculado ao género gramatical. As gramáticas citam os seguintes exemplos: *o/a cabeça o/a caixa, o/a capital, o/a cisma, o/a guarda, o/a guia*, etc. (o artigo, no nosso caso, como veremos, será muitas vezes o único indicador gramatical e, implicitamente, também semântico). Frias e Gouveia afirma que dentro deste grupo se encontram palavras que ainda não têm género fixo e que 1. tendem a fixá-lo (Frias e Gouveia: 2005) ou 2. a alterá-lo. Estão, no primeiro caso, por exemplo, *amálgama* e *hélice* que tendem a fixar-se como femininas; no segundo, entre outras, *guia* e *grama*. (*ibidem*). Quanto à palavra *componente* não conseguimos encaixá-la nem num nem noutra caso. Não sabemos se o género desta palavra tende a fixar-se ou a alterar-se ou se se trata, antes, de um lema não claramente dicionarizado por carecer, na sua explicação dicionarizada, de uma clara definição pragmático-semântica e de qualquer restrição diatópica.

Segundo as nossas premissas poderá, conforme o género do quarto tipo dos substantivos uniformes de dois géneros (d), mudar também o grau de genericidade, o que normalmente ocorre sobretudo nos nomes sexuados (b), onde o género masculino é evidentemente genérico, apesar de haver tendência para a biformização (*árbitro x árbitra*). Nos épícenos (c), tanto o feminino como o masculino podem ser genéricos (*o tigre x a águia*). Nos nomes inanimados uniformes que mudam de sentido conforme o género, provavelmente não existirá nenhuma regra geral, sendo o carácter genérico, de acordo com as nossas premissas, idiossincrático. Mas para tal afirmação teríamos que proceder a um outro tipo de pesquisa e analisar todos os

nomes uniformes e biformes de dois géneros etimologicamente idênticos e estudar a compatibilidade de abstração com outras unidades gramaticais (ou classes lexicais). De acordo com as nossas hipóteses e no que à palavra *componente* diz respeito, no Português Europeu, o feminino será mais genérico do que o masculino, no Português do Brasil, o contrário.

3. Descrição metodológica

Vejam, em síntese, como se processou a nossa investigação. No caso de *componente* tivemos que decidir, primeiro, se é palavra uniforme de um ou de dois géneros. Por mais paradoxal que possa parecer a um falante nativo, não foi tão fácil encontrar uma resposta unívoca. Com efeito, os resultados da pesquisa que realizámos nos dicionários checo-portugueses, não coincidiam com as experiências que adquirimos a título de falante de PLE (português língua estrangeira). Os dicionários bilingues existentes no contexto da lexicografia L1-L2 e L2-L1⁽¹⁾ definem o nome *componente* como substantivo de género único, i. e. masculino, o que talvez se deva à influência da variedade brasileira, como veremos mais adiante.

Continuámos a consultar outros dicionários portugueses (*Dicionário da Língua Portuguesa Contemporânea da Academia das Ciências de Lisboa*, *Dicionário Houaiss*, *Dicionário Aurélio*, *Dicionário Priberam*, *Dicionário Aulete*) que admitem a possibilidade de serem usados ambos os géneros, contudo, não os delimitam pragmaticamente.

Para estabelecermos a delimitação semântica e diatópica precisa dos dois géneros gramaticais atribuídos à palavra *componente*, realizámos mais dois passos importantes: o primeiro não se mostrou eficiente; o segundo, ao invés, ajudou-nos a orientar a nossa investigação. Assim, primeiro, consultámos os dicionários das outras línguas românicas, supondo que poderíamos encontrar oposições genéricas analógicas do nome *o/a componente* em francês, espanhol e italiano.

Ao consultar os dicionários naquelas línguas, enfrentámos, parcialmente, os mesmos problemas da inexatidão da delimitação pragmática (semântica, diafásica e diatópica), tal como nos dicionários portugueses.

Consequentemente, tivemos que excluir o aspeto contrastivo ao nível das línguas românicas e procurar outras fontes fiáveis, as quais, como veremos, acabaram por constituir o pilar da nossa pesquisa, adequado às nossas

(1) L1 (língua materna) – L2 (língua estrangeira), de acordo com Felix Córdoba, *apud* Buzek (2010) e (2011).

finalidades, i. e., recorremos aos *corpora* Linguateca e *Corpus* do Português e analisámos todas as ocorrências encontradas, tendo sempre em atenção o contexto utilizado. Além destes dois *corpora* linguísticos, consultámos também o *corpus* paralelo Interkorp, o qual, não obstante, nos ofereceu apenas 600 ocorrências de *componente* quer na função nominal quer na adjetival, sem a possibilidade de delimitação geográfica geral dos dados, facto que nos levou a eger o *Corpus* do Português como instrumento de análise contrastiva diatópica do género gramatical de *componente* (i.e., nas duas variedades geográficas: a do Português do Brasil e a do Português Europeu -doravante PB e PE) e o *corpus* Linguateca como instrumento de análise do género desta palavra usada exclusivamente no PE. O *Corpus* do Português ofereceu-nos 400 ocorrências nominais para o PB e 200 para o PE. O *corpus* Linguateca⁽²⁾ ofereceu-nos 4000 ocorrências nominais e adjetivais, das quais retirámos uma amostra bastante representativa de mais de 1000 construções nominais que foram divididas, depois de uma classificação minuciosa (que teve que ser feita em muitos casos manualmente), em dois grupos, de acordo com o género masculino ou feminino.

4. Lematização de *o/a componente* nos dicionários de língua portuguesa

Proceda-se à descrição pormenorizada do processo de investigação e da metodologia aplicada. A primeira fase da nossa pesquisa foi de natureza lexicográfica. Como referimos no quadro metodológico, começámos por consultar os dicionários bilingues L1 (língua materna) – L2 (língua segunda) (Buzek:2011). Para nossa surpresa, não encontrámos nos dicionários bilingues o género feminino – caso um tanto peculiar, conforme veremos mais à frente. Para os fins da nossa análise, consultámos o *Dicionário da Língua Portuguesa Contemporânea* (2001), da Academia das Ciências de Lisboa, o qual, curiosamente, apesar de admitir a possibilidade do uso de ambos os géneros, não os distinguiu pragmático-semanticamente, incluindo os dois sob um único lema. Este dicionário documenta o uso do género feminino num único caso (1) e as demais explicações semânticas falam a favor do género masculino. Recorremos, por isso, aos dicionários de língua portuguesa *Priberam* e *Aulete*, sendo esse último o único que especifica pelo menos uma área, em que o nome aparece no género feminino, i.e., em astronomia, dando o exemplo de *a componente chaundleriana*. Curiosa-

(2) Costa, Santos & Cardoso (2008).

mente, encontramos num outro dicionário de língua portuguesa, no *Aurélio*, o género inverso nesta combinação, i.e., o *componente chaundleriano*. O dicionário que com maior detalhe especifica o uso do género da palavra *componente*, é o *Houaiss*. Este trabalha o lema em questão mais detalhadamente, indicando a possibilidade dos dois géneros e incluindo algumas colocações e combinações. Aponta para o uso em uma ou outra área científicas, indicando até as possibilidades de sinonímia em alguns casos (*componente fonológico x fonológica*, *componente semântico x semântica*), mas não nos dá resposta definitiva quanto ao uso que corresponde à realidade, porque não efetua a suficiente divisão dos géneros gramaticais. As únicas referências que falam a favor de género feminino são as que se encontram sublinhadas:

Veja-se, em concreto, todas as dicionarizações encontradas de *componente*:

- Dicionário da Língua Portuguesa Contemporânea⁽³⁾
componente s. m. ou f. (do latim *componens*, *componentis*), part. pres. de componere – ‘colocar juntamente’). **1.** Elemento que entra na composição de alguma coisa. *O verbo é um dos componentes da frase. O curso tem uma componente eminentemente prática.* **2. quím.** Elemento de um corpo composto. *O hidrogénio é um dos componentes da água.* **3. tecnol.** Elemento que entra na composição de um circuito electrónico. **4. fís.** Cada uma das forças que actuam simultaneamente sobre um corpo, cujo efeito é equivalente ao da resultante. **5.** Membro de uma classe, corpo, instituição... *Os componentes do grupo de escuteiros actuaram com destreza e sentido prático.* (DLPC: 890).
- Dicionário Priberam da Língua Portuguesa:⁽⁴⁾
componente s. 2 g. Elemento (que concorre com outros numa composição). [Mecânica] Cada uma das forças que actua simultaneamente para produzir entre todas uma resultante.
- Dicionário Aulete⁽⁵⁾
componente s. 2 g. Alguém ou algo que faz parte de um conjunto organizado; parte ou elemento de um sistema: os componentes de um computador. Membro de um grupo: os componentes do coral. *Quím.* Substância que compõe um sistema químico (e necessária para caracterizá-lo) [Tb.

(3) As descrições metalinguísticas citadas, sendo diretas, não foram adaptadas ao Novo Acordo Ortográfico.

(4) <http://www.priberam.pt/dlpo/default.aspx?pal=componente>.

(5) <http://aulete.uol.com.br/componente>.

us.como adj.] *sf. Astron.* Qualquer das estrelas que compõem um sistema estelar. [F.: Do lat. *componens, entis.*]

- Dicionário da Língua Portuguesa Aurélio ***componente*** (do lat. *componente*).
adj. 1. Que entra na composição de alguma coisa.
s. m. 2. Aquilo que entra na composição de alguma coisa. 3. Parte elementar de um sistema 4. *Fis.Quím.* Num sistema, qualquer das espécies químicas, entre as que o constituem, incluída no menor grupo de substâncias necessárias para caracterizar quimicamente o sistema (Cf.nesta acepção constituinte.)
Componente chandleriano, *Astr.* V. período de chandler.

- Dicionário Houaiss da Língua Portuguesa⁽⁶⁾
componente Adj. 2g s. 2g 1. Que ou o que compõe ou ajuda na composição de algo <os c. de um motor> <os c. de um sistema filosófico>; 2. que ou quem é membro de uma classe, de uma instituição, de um corpo, etc. <os elementos c. do corpo policial> <uma organização esquerdista com c. que se pretendiam autônomas>; 3. em ciência e tecnologia, diz-se de ou parte constituinte de um sistema; 4. *ELECTR.* diz-se de ou qualquer dispositivo com características definidas que disponha de terminais através dos quais possa ser conectado a outros componentes para formar um sistema; 5. *GRAM.GENER.* diz-se de ou cada uma das partes constituintes da gramática de uma língua: componente sintático, componente fonológico, componente semântico; 6. *GRAM.GENER.* diz-se de ou cada uma das partes do componente sintático: componente da base e componente transformacional; 7. *QUÍM.* Diz-se de ou substância que compõe um sistema químico *componente categorial; *GRAM.GENER.* 1. mesmo que base categorial; 2. a parte da base que contém as categorias sintáticas existentes na língua e o sistema de regras para a construção dos sintagmas e da estrutura básica da oração, determinando as relações gramaticais entre os seus constituintes. *Cf. Categoria sintática. c.chandleriano ou c.chandleriana *ASTR.* m.q período de Chandler, c. fonológico ou c.fonológica; *GRAM.GENER.* inventário de fonemas e traços distintivos de uma língua e sistema de regras que interpretam foneticamente as cadeias geradas pelos componentes sintático e transformacional, em termos de traços distin-

(6) Incluímos o Dicionário da Língua Portuguesa do Instituto de Antônio Houaiss nos dicionários consultados para obtermos o maior número possível de lematizações. Não obstante, estamos conscientes de que a publicação contém expressões “pejorativas e preconceituosas”, pratica racismo contra ciganos e tem outras explicações bastante imprecisas. Curiosamente, como se vê, para a nossa investigação foi exatamente este dicionário que nos proporcionou a lematização mais trabalhada da palavra *componente*, o que, por outro lado, comprova o facto de ser o dicionário mais trabalhado no contexto da lexicografia portuguesa.

tivos; componente lexical. GRAM.GENER. m.q.léxico; c. semântico ou c.semântica GRAM.GENER.1. sistema de regras que realizam a interpretação semântica das frases geradas pelo componente sintáctico 2.LING. em análise semântica opoencial, unidade mínima de significação que entra na formação do sentido de um morfema ou de uma palavra (p.ex.: [+humano], [+adulto], [+fêmea] são componentes semânticos do significado de mulher); traço semântico, marcador semântico. c. sintagmático ou c.sintagmática GRAM.GENER. sistema de regras que determinam as estruturas frásicas possíveis na língua sintaxe; componente sintáctico ou c.sintáctica GRAM.GENER. parte da gramática que gera, por meio de regras formalizadas, todas e apenas as sentenças possíveis de uma língua. Uso como termo da GRAM.GENER. no Brasil é ger.empr. no masc. Etim. Lat. *Componens, entis*, part. pres. de *componere*.

5. Lematização de o/a componente nos dicionários de língua espanhola, francesa e italiana

Como já se referiu na parte introdutória e no quadro metodológico, um dos possíveis indicadores do género de *componente* poderia ser a analogia com as interpretações semânticas de *componente* em outras línguas românicas. Procedemos à investigação das elaborações dicionarísticas de *componente/composante* em dicionários monolíngues de língua espanhola, italiana e francesa. Apesar de terem sido encontradas especificações muito relevantes, infelizmente, tivemos que enfrentar o mesmo problema de insuficiente identificação genérica.⁽⁷⁾ Vejamos apenas algumas das definições para não nos desviarmos do tema fulcral do nosso trabalho. Sublinhamos, novamente, todas as referências relativas ao género feminino.

Em espanhol, verificam-se ambos os géneros, sendo, em todo o caso, mais frequente o uso do masculino. No *Diccionario de la Real Academia Española*⁽⁸⁾, encontramos as seguintes explicações: „En matemáticas, se usa como sustantivo femenino para nombrar cada una de las partes en que se descompone un objeto matemático, como, por ejemplo, un vector, y se emplea frecuentemente en meteorología para referirse a la dirección de los vientos: «*Vientos flojos con predominio de la componente oeste*» (NCastilla [Esp.] 13.5.99). No debe extenderse este uso en femenino a otros ámbitos, como ocurre en los ejemplos siguientes: «*Esta unidad coordina todas las*

(7) Constate-se que se trata de um problema geral que aparece também na área de lexicografia de outras línguas: em espanhol (veja-se Ivo Buzek: 2010; 2012) ou em francês (veja-se A. Polická: 2010, 2011), por exemplo.

(8) <http://lema.rae.es/dpd/>.

componentes del computador» (Pérez/Pino *Computación* [Chile 1982]); «*Se intenta justificar la revitalización de la UEO como medio de fortalecimiento de la componente europea de la OTAN»* (País [Esp.] 2.8.88).⁽⁹⁾

Em italiano, existem também os dois géneros, mas o problema é que os dicionários de qual se dispõe, não são homogêneos quanto ao seu tratamento léxico-gramatical. Alguns mencionam em primeiro lugar o género masculino como o mais frequente, tendo um espectro mais largo de variação semântica e pertencendo ao campo geral (usado no sentido de elemento, participante, membro em estruturas como *il componente della commissione, il componente essenziale della pasta*), mas também ao campo técnico e linguístico (*il componente sintattico di una lingua*). Mas é possível usar o género feminino também num sentido mais geral de componente não químico (*le componente del pensiero di un autore*).⁽¹⁰⁾

Em francês, o uso do género feminino da palavra *composant* sinaliza, de acordo com *Le Trésor de la Langue Française Informatisé*⁽¹¹⁾, um maior grau de abstração, podendo ser usado em sentido figurativo:

“2. *Au fig., emploi subst. fém. Une composante; les composantes d’un problème. Faire à travers les expériences vécues par la totalité des catégories sociales de la population l’inventaire exhaustif des composantes du loisir*”;
 “3. *Si le puritanisme n’exprime pas l’essence du protestantisme, il n’y a pas moins dans l’âme protestante une composante puritaine; la production ne saurait avoir pour but la seule jouissance.*”⁽¹²⁾

O género feminino, de acordo com a mesma fonte lexicográfica, também é usado na matemática (*composante(s) vectorielle(s)*); na área da linguística e da gramática generativa (em que também mais raramente pode ser usado o masculino) e na área da mecânica em que *composante* tem também um sentido mais abstrato: (*une composante. “Les trois composantes de l’accélération de translation du centre de gravité de la tête.”*). Por outro lado, no que à área da química e aos produtos eletrónicos diz respeito, são mencionados os seguintes exemplos no masculino: *un composant. le composant*

(9) Para fazer uma análise mais profunda sobre o uso do género em espanhol, recomendamos que se use o *corpus* <http://corpus.rae.es/creanet.html>.

(10) <http://garzantilinguistica.sapere.it/>, <http://www.wordreference.com/iten/componente> e http://dslo.unibo.it/coris_eng.html.

(11) <http://atilf.atilf.fr>.

(12) *Ibidem*.

doué de la plus grande activité biologique, les composants d'un circuit électronique (tels que résistances, condensateurs, transistors.⁽¹³⁾)

De acordo com os exemplos citados, pode-se deduzir que, ao passo que o género masculino é usado mais frequentemente no caso dos nomes concretos e técnicos (por exemplo, nas áreas da eletrónica e química), o uso do feminino é mais frequente nas áreas da literatura, filosofia, linguística, artes, mas também na área da mecânica, em que representa um nome abstrato.⁽¹⁴⁾ Pode antecipar-se que os nossos resultados, relacionados com o uso real dos géneros em português, irão mais ao encontro da palavra *composante* do francês.

6. Diferenças históricas

Frias e Gouveia dedicou uma considerável parte dos seus artigos ao tema da evolução do género gramatical do Português arcaico para o moderno (Gouveia: 1993, 2005) de palavras que mudaram de género no curso da evolução da língua portuguesa „ou porque fixaram um determinado género, quando havia hesitação, ou porque a terminação da palavra determinou essa fixação ou ainda porque a língua culta exerceu alguma influência“ (Frias e Gouveia: 2005; 550). “É um facto que encontramos atualmente palavras cujo género difere do antigo ou que o alteraram até aos nossos dias, em muitos casos tendo ainda género duvidoso” (*ibidem*). Pressupomos que, no caso de *componente*, não deve ter havido um género etimológico, sendo “*componens*“ o particípio de ‘praesentis activ’ e, como tal, deve ter sido considerado como adjetivo ‘amorfo’⁽¹⁵⁾, mas trigenérico com o mesmo sufixo para os três géneros (neutro, feminino e masculino), adquirindo assim o género do substantivo com o qual se une. Tendo desaparecido o neutro, restaram só dois géneros, adquirindo *componente* o género do núcleo substantivo. Assim pode deduzir-se que, através da derivação imprópria, *componente* português se tornou um substantivo bigenérico com um uso variável. De acordo com os exemplos encontrados, no PE, prevalece o género feminino enquanto que, no PB, o masculino. Pressupomos que o PB é a pré-fase do português europeu moderno e que pode ter havido a mudança de género do masculino para o feminino no PE, enquanto que, no PB, ficou tal como

(13) *Ibidem*.

(14) *Ibidem*.

(15) *Adjetivo amorfo* é um termo usado para os adjetivos uniformes, de acordo com a terminologia de romanistas praguenses: (Zavadil, Čermák:2010:176).

se tinha fixado depois de lhe ter sido atribuído o género masculino. Neste caso, podemos admitir que o género de *componente* do PE seguiu o género gramatical de *composant* francês/a, e parcialmente de *componente* italiano, como se pode ver no parágrafo anterior. Em PE, encontramos uma vasta escala de exemplos a serem analisados na próxima secção do presente trabalho. Contudo, para podermos confirmar e comprovar as nossas hipóteses, seria necessário fazer uma análise diacrónica do género gramatical de *componente* e dispor de documentos históricos. Como os textos escritos que consultámos e que tivemos à disposição não contêm a palavra em questão, recorreremos aos *corpora* históricos possíveis (a Linguateca – Vercial e ao *Corpus* do Português), que, infelizmente, não nos proporcionaram uma amostra significativa de ocorrências. O *Corpus* do Português ofereceu-nos 10 exemplos do século XIX, todos no feminino e todos usados por um único autor. É escusado lembrar que tal análise não seria fiável mas, sim, passível de suscitar grandes dúvidas. O *corpus* Linguateca Vercial não encontrou nenhuma ocorrência da palavra em questão. Maria Carmen de Frias e Gouveia apoiou muitas vezes as suas análises em relatórios médicos dos séculos passados podendo assim observar, com base em materiais fiáveis e autênticos, a evolução do género gramatical no que à terminologia médica diz respeito (por exemplo, *eczema*, *laringe*, *entorse*, etc) (*ibidem*). É nossa firme convicção que o único procedimento válido, no caso da palavra *componente*, seria estudar os materiais arquivados de produtos alquímicos, agrícolas de alimentação e áreas de longa tradição. Faltando-nos este tipo de documentos, omitiremos este aspeto diacrónico, tendo todavia consciência de que este terá desempenhado, no caso aqui estudado, um papel primordial. Devido a estes obstáculos relativos à insuficiência de materiais autênticos, no nosso trabalho, apenas se pode desenvolver um estudo sincrónico. Este, como veremos a seguir, constará de duas partes principais: uma que abordará o problema das diferenças geográficas (ou diatópicas) e outra que nos oferecerá um estudo pormenorizado do uso da palavra *componente* em português europeu contemporâneo.

7. Diferenças diatópicas

O objetivo desta secção será considerar as principais diferenças entre *componente* usado nas variedades europeia e americana do Português. Sobre este aspeto, Frias e Gouveia escreveu uma série de artigos (Frias e Gouveia: 2006, 2008, 2009, 2011). Omitiremos da nossa análise, por não termos acesso à linguagem popular e regional, as diferenças verticais que existem

dentro de cada uma das variedades comparadas. Para já, relembremos os exemplos que Frias e Gouveia inclui nos seus estudos relacionados com o PE e PB: *o grama (PB) x a grama (PE)*, *os media (PE) x a mídia (PB)*, *o duche (PE) x a ducha (PB)*, *o gangue (PE) x a gang (PB)*, *o fondu (PE) x a fondu (PB)*, *o champanhe (PE) x a champanhe (PB)*, *o sanduíche (PB) x a sanduíche (PE)*, *o/a omelete (PB) x a omelete (PE)* (Gouveia: 2007). Como vemos, trata-se de substantivos de um género, com uso divergente em Portugal e no Brasil.

O nosso caso é um pouco diferente. Segundo os resultados que obtivemos, no PB, a palavra *componente* é palavra de um género (hesitante), já que não encontramos nenhuma diferença semânticas acarretadas pela mudança do género gramatical (substantivos uniformes de/do? tipo *a*) no “Enquadramento teórico”). Ao contrário, na variedade europeia, os resultados falam atualmente mais a favor do carácter uniforme mas *bigenérico*, especificando-se cada género gramatical a ser usado em diferente secção semântica. gramatical (substantivos uniformes de/do? tipo *d*) no “Enquadramento teórico”).

O *Corpus do Português* oferece-nos a possibilidade de analisar semanticamente a palavra *componente* de modo separado nas duas variedades comparadas. Procedemos à análise de 237 ocorrências em PE e de 411 em PB. Não conseguimos identificar o género gramatical em 179 casos do PB e em 71 do PE, por faltar na construção um determinante, atributo ou modificador biformes, marcadores do género no caso dos nomes uniformes. Comparámos os contextos em que a palavra ocorre no género masculino e feminino em ambas as variedades da língua portuguesa. Chegámos a revelar que, no PE, o género masculino foi identificado em 44 ocorrências, enquanto que o feminino o foi em 122. No PB, surpreendentemente, ocorreu o contrário: o masculino foi identificado como o género gramatical mais frequente (em 232 casos) e o feminino apenas em 30. Procedendo à análise dos contextos em cada frase, chegámos à conclusão de que, de acordo com os exemplos que tivemos à disposição, no português europeu, o género feminino da palavra *componente* parece genérico, sendo usado num vasto leque de campos semânticos (em áreas como meio-ambiente, ecologia, cultura, economia, linguística, política) e o masculino revelou-se semanticamente restringido ao uso em áreas como eletricidade, produtos eletrónicos, biologia, medicina, farmacologia, química, etc. No português do Brasil, verificou-se quase o contrário: o género feminino revelou-se muito menos frequente do que o masculino. Ao contrário da variedade europeia, contudo, não foram encontradas nenhuma restrições semânti-

cas. O uso do género feminino equivale, a título do espectro semântico, ao uso do masculino, sendo esse muito mais frequente.

Vejam-se alguns exemplos de sintagmas Nome+Adjetivo, tirados do *Corpus* do Português:

Português Europeu 22(m.) < 122(f.)⁽¹⁶⁾

No género masculino foram encontradas **44** ocorrências:

O componente- ativo (do tempero), tóxico, biologicamente ativo, primário (o tolueno), principal (montromonlonite).

No género feminino foram encontradas **122** ocorrências:

A componente – agrícola, ambiental, artificial, áudio, brasileira, cénica, científica, cultural, da velocidade radial, de abastecimento de água, de entretenimento, de formação, de internacionalização, de pequenos comerciantes, de reciclagem, essencial de funcionamento, estratégica, exportadora, expressiva, financeira, genética, histórica, horizontal, humana, importante, inata, individual, noturna, política, sensorial, simbolista, social, subjetiva, técnica, tecnológica, necessária, onírica, política.

Português do Brasil 232 (m) > 30 (f)

Foram encontradas **232** ocorrências no género masculino:

O componente – arbóreo, co-seno, da atividade, das taxas, de contextualização, de ajuste estrutural, de circuito, de geração, de processamento, de traição, de um sistema, do texto, dominante, em questão, genético, gramatical, importante de filmes, informativo, isolado, menos controlado, metodologia, resistente, plástico, político, propósito, racional, sentencial, socioeconómico, ético.

No género feminino foram encontradas **30** ocorrências:

A componente – da força, escalar do gradiente, fonética, fonético-fonológica, hidrostática, horizontal da velocidade, vertical, radiométrica, subjetiva, seno, co-seno.

8. Estudo sincrónico do português europeu

Como pudemos ver nas secção anterior, o problema no caso do *Corpus* do Português foi o facto de este nos proporcionar um número insuficiente de ocorrências usadas em PE, sobretudo no masculino, o que nos impossibilitou obter uma delimitação semântica mais exata de ambos os géneros.

(16) Os exemplos incluídos no presente texto são citações diretas de www.corpusdoportugues.org, mas apesar disso, foram adaptados ao Novo Acordo Ortográfico.

Como pudemos observar, a análise dos resultados das ocorrências encontradas sinalizou-nos pelo menos o grau de extensão semântica dos dois géneros no PE (vejam-se os exemplos mencionados no fim da secção anterior.) O *componente* aparece acompanhado por atributos e modificadores que pertencem às áreas da eletricidade e dos produtos eletrónicos, biologia, química, farmacologia e informática. O género feminino foi encontrado em contextos semanticamente mais vastos (cultural, económico, político, científico, etc.).

Dos valores que o *corpus* Linguateca (Cetem Público) nos oferece, depreende-se que a palavra *componente* apresenta um total de 4629 ocorrências, das quais 1172 documentadas na secção sociedade. Em segundo lugar está a secção cultura, seguida por áreas não determinadas: política, economia, cultura e sociedade, desporto, opinião e informática. O uso de cada género gramatical difere conforme a área, como vemos.

Eis os primeiros resultados de Linguateca, Cetem Público:

Distribuição de '*componente*' por secção

Sociedade	1172	Economia	606
Cultura	970	Cultura	264
Área não determinada	686	Desporto	164
Política	632	Opinião	99
		Computadores	36

Procedemos da seguinte maneira: para ilustrar previamente a diferença do género gramatical de acordo com a frequência de uso e com o campo semântico, pesquisámos o uso da palavra *componente* apenas com os determinantes *o*, *a*, *um*, *uma*, o processo mais fácil de identificar o género gramatical. Chegou-se a uma diferença bem marcada entre o género feminino e o género masculino.⁽¹⁷⁾

(17) As abreviações usadas do CETEMPúblico correspondem aos valores possíveis: pol (política portuguesa e internacional), des (desporto), eco (economia), clt (cultura), opi (opinião), com (informática) e nd (não determinado). Alguns artigos pertencem a mais de uma categoria (clt-soc).

	uma	a	um	o
Sociedade	244	276	11	6
Cultura	234	243	14	7
Cultura-sociedade	54	41	18	7
Não determinado	144	160	15	10
Economia	126	149	9	
Política	109	131	4	3
Desporto	32	40	4	
Opinião	21	12	2	1
Computadores	3	5	2	3

Mas faltava, em todo o caso, procurar outras ocorrências da palavra *componente* acompanhada por determinantes, modificadores e atributos, marcadores do género em casos duvidosos. Quando não foi possível identificar o género gramatical pelo atributo ou modificador mencionado no quadro anterior (por exemplo, no caso de „forte“ „componente“, „como“ „componente“), tivemos que procurar o marcador do género gramatical mais próximo: „o“ „forte“ „componente“/“a“ „forte“ „componente“, e outros. Com este procedimento pretendemos registar um número mais elevado de ocorrências e obter assim a maior objetividade possível:

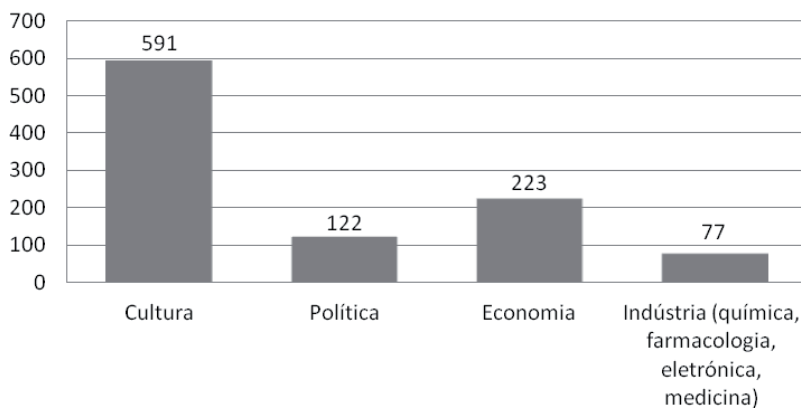
Determinante + <i>componente</i> Preposição + determinante + <i>componente</i> Modificador + <i>componente</i> Preposição + <i>componente</i>	f.	m.	Não identificado
esta/este componente	47	4	
desta/deste componente	17	4	
nesta/neste	3	-	
essa/esse componente	48	5	
nessa/nesse componente	3	-	
dessa/desse componente	10	2	
aquela/aquele componente	4	3	
daquela/daquele componente	-	1	
naquela/naquele componente	-	-	
da/do componente	492	4	
na/no componente	113	-	

Determinante + <i>componente</i> Preposição + determinante + <i>componente</i> Modificador + <i>componente</i> Preposição + <i>componente</i>	f.	m.	Não identificado
a/o componente	1057	34	
uma/um componente	967	79	
segunda/segundo componente	19	2	
sua/seu componente	198	4	
pouca/pouco componente	1	-	
outra/outro componente	48	8	
ligeira/ligeiro componente	1	-	
, (vírgula) componente	6	2	18
como componente	17	-	33
sem componente	2	-	2
forte componente	320	.	40
principal componente	27	15	24
TOTAL:	3063	167	117

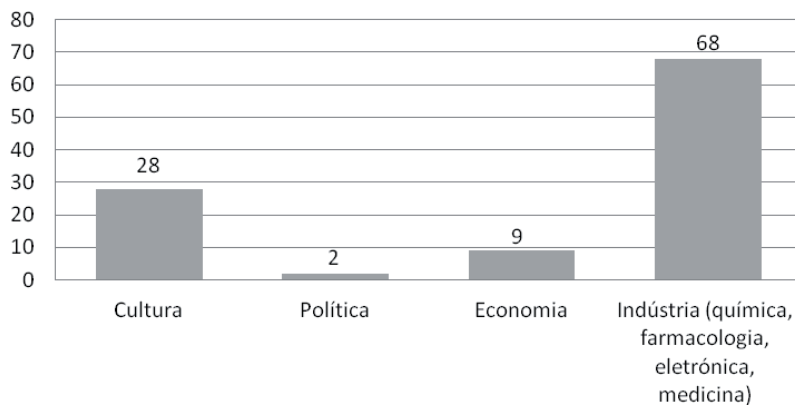
Chegámos a contar 3063 ocorrências do género feminino, 167 do género masculino, sendo as demais ou adjetivos ou substantivos de género não identificável. Da lista das construções que foram analisadas, fizemos a seleção manual de todas as que se repetiram (duas, três ou até mais vezes), como ocorreu com *componente política, cultural, social, a(c)tiva, passiva, máxima, vital, viral, de educação, regionalista, rebelde, moralista, ética, etc.*, chegando o nosso *corpus* a perfazer 1013 construções da palavra „componente“ no género feminino e 90 no masculino. Dividimos as construções encontradas em quatro campos semânticos diferentes – ou secções, como se pode ver no quadro seguinte:

Comparação da atribuição do gênero gramatical à palavra *componente* de acordo com o campo semântico:

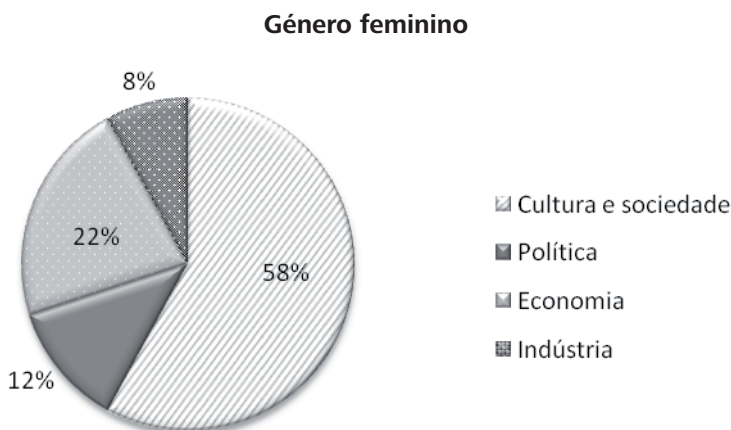
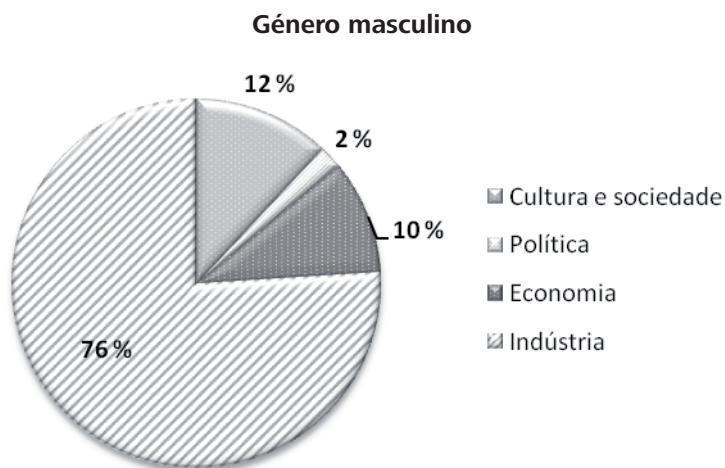
Ocorrências no gênero feminino



Ocorrências no gênero masculino



Divisão porcentual por cada área (percentagem arredondada):



9 . Conclusão:

Os resultados da presente pesquisa contrariaram parcialmente as nossas premissas: pressupunha-se que a oscilação do género gramatical fosse originada, com maior probabilidade, pelo fator da frequência de uso, o que acabou por verificar-se em apenas alguns casos do PB e noutros tantos

casos do PE, mas não se esperava que esta análise acabasse por ser um estudo do carácter „genérico semântico“ (Frias e Gouveia: 1999) da palavra em questão.

Em algumas construções foi difícil encontrar a fronteira exata entre a secção política e económica como, por exemplo, em *componente de cooperação* ou *componente profissional*. Uma vez, *componente* parecia ser a palavra comum e mais abstrata, como no caso de *componente humana* ou *componente portuguesa*. Por esta razão, decidimos analisar os constituintes fráscicos mais próximos como fizemos no caso de *componente portuguesa da produção*, em que só o atributo *da produção* especificou a área de indústria. Esta necessidade levou-nos a observar detalhadamente as construções inteiras e não apenas as que estavam imediatamente pospostas (e em raros casos antepostas) à/ao *componente*. Tendo feito esta análise minuciosa, conclui-se que, no PE, o género está mais vinculado à compatibilidade semântica no eixo sintagmático, o que não raramente nos faz aproximar das chamadas combinatórias ou colocações. Muitas vezes, a palavra *componente* é seguida de aposto de especificação, como se pode ver na lista adicionada em anexo.

Todos os dados parecem indicar que o género masculino da palavra *componente* aponta para as realidades mais concretas existentes nas áreas da química, farmacologia, medicina, alimentação, eletricidade e de produtos químicos, enquanto que o género feminino abrange um grau maior de abstração, o que, também, predeterminará certamente o grau de abstração dos verbos. Muitas vezes o *componente* articula-se com verbos como *colar*, *pegar*, *transportar*, *reduzir*, *ser*, *detetar*, *lançar*, *fornecer*, *encontrar*, *danificar*, etc. Em *a componente*, por outro lado, verificou-se no eixo sintagmático o uso de verbos mais abstratos (por exemplo, *incentivar*, *desenvolver*, *juntar*, *incluir*, etc.).

Ao mesmo tempo, registou-se, na secção de cultura, em que se verificou a prevalência do feminino, em poucos casos, o uso do género masculino, o que talvez se deva à oscilação do género que não acarreta nenhuma diferença semântica. A secção de cultura, não obstante, parece ser a única em que ocorre a oscilação do género no PE, sendo que na secção de política e de economia foram documentadas apenas 2 e 9 ocorrências no masculino, respetivamente, ou seja, um número muito reduzido para podermos constatar o mesmo fenómeno.

Como vimos, a percentagem calculada é absolutamente unívoca no que diz respeito à predominância do género feminino, demonstrando que *componente*, no género feminino, designa uma realidade extralinguística com

maior grau de abstração. Este facto é comprovado pelos números, contradizendo o que estipulam os dicionários, que falam evidentemente a favor do uso do masculino. Este, não obstante, caso apareça, é usado em contextos muito mais específicos, correspondendo 76% das suas ocorrências encontradas nas áreas da informática, química e farmacologia, ou seja, áreas em que foi documentado muito raramente o género feminino.

Mesmo que à primeira vista os números falem a favor do feminino, teremos que proceder, na escolha do género, muito mais cuidadosamente, considerando as áreas em que se usará o nome *componente*. Nas áreas da cultura, política e economia, podemos, sim, partir de números quase unânimes: 591:28 (cultura), 122:2 (política), 223:9 (economia). A área da indústria, não obstante, constitui um ponto problemático, por predominar, de novo, o uso do género feminino: 77:69. Apesar disso, antes de concluir a pesquisa, é preciso acrescentar que a área da indústria teve que ser subdividida em subsecções, conforme o contexto em que *componente* aparece. Assim, obtivemos mais três subcampos 'industriais', sendo o primeiro 'a indústria em geral', onde aparece muitas vezes o género feminino. No segundo subcampo de 'as indústrias farmacológica, química e alimentar' predomina o uso do género masculino, aparecendo muito raramente casos de feminino. Na terceira subsecção da 'informática', contudo, verificou-se apenas o género masculino.

A favor do uso do masculino na área industrial (sobretudo nas áreas da farmacologia, química e informática), fala também a alta percentagem deste campo no espectro semântico geral do género masculino. Dos 100% das ocorrências, foi aqui que se registou a maior parte do uso do género masculino. Aliás, a mais alta de todos os itens comparados, o que corrobora a unanimidade do uso do masculino nesta área.

Anexo: ⁽¹⁸⁾

Política:

f.

a componente

c. anticomunista, c. comunitária, c. das forças nucleares, c. federal, c. governamental, c. interestadual, c. internacional, c. jurídica, c. marcelista, c. marítima, c. mediática, c. militar, c. nacional, c. nominalista, c. ovimbundu, c. parlamentar, c. regional, c. de reparação geográfica naval, c. social, c. socialista, c. supranacional, c. transatlântica, c. ultra-nacional, c. urbana, c. do Sistema de Forças Nacionais, c. de requalificação urbana, c. da segurança política, c. de esquerda, c. de centro dentro da coligação

m.

o componente

c. da cidadania e da democracia, c. importante do bem-estar social do Estado

Cultura:

f.

a componente: c. da atividade teatral, c. afetiva, c. de aproximação ao bairro, à família e à comunidade, c. árabe, c. árabe-islâmica, c. arbitrária, c. arqueológica, c. artística, c. arquetónica, c. ascética, c. autonómica, c. autobiográfica, c. auto-sustentada, c. benéfica, c. bélica, c. Benetton-retardada, c. cabarética, c. céltica, c. civil, c. checa, c. científica, c. cultural, c. «consulta aos telespetadores», c. criativa, c. dançante, c. dramática, c. dramatúrgica, c. design, c. dissuasória, c. eclesiástica, c. editorial, c. educativa, c. emocional, c. ensaística, c. equestre, c. erótica, c. estética, c. europeia, c. experimental, c. extracurricular, c. «fake american band», c. de ficção portuguesa, c. formal, c. formativa, c. «gay», c. geral, c. humana, c. inovadora, c. invisível, c. jamaicana, c. jazzística, c. labiríntica, c. letiva, c. não letiva, c. da literatura melodramática, c. minimalista, c. mística, c. mitológica, c. musical, c. não cíclica, c. operacional, c. de oralidade, c. paisagística, c. pedagógica, c. pedagógico-didática, c. orquestral, c. «pastiche», c. perversa, c. pessoal, c. plástica, c. popular, c. prática, c. profissionalizante, c. psicológica, c. de reconhecimento, c. reivindicativa, c. religiosa, c. romanesca, c. sádica, c.

(18) Exemplos do uso de ‘o/a’ componente no Português Europeu de acordo com as áreas/secções semânticas. Esta lista não é completa e contém as combinatórias mais frequentes, sendo adicionada para possíveis fins lexicográficos.

sexual, **c.** social, **c.** soul e blues, **c.** subtil, **c.** teatral, , teatral, **c.** do temperamento austríaco, **c.** utópica, **c.** visual

m.

o componente: **c.** cómico, **c.** «rock `n` roll», **c.** de aposta; **c.** essencial da finta; **c.** enriquecedor da sexualidade dos envolvidos, **c.** valioso do património cultural, **c.** do processo social, **c.** fulcral do desenvolvimento sustentado

Economia

f.

a componente

c. valor acrescentado bruto, **c.** da capacidade de manutenção de empregos, **c.** comercial, **c.** de crédito a empresas, **c.** do desemprego de importação e distribuição, **c.** do desenvolvimento, **c.** Ecologia, **c.** economia, **c.** económica, **c.** emprego, **c.** empresarial, **c.** externa, **c.** financeira, **c.** fiscalizadora, **c.** indexada, **c.** imobiliária, **c.** importada, **c.** da inflação, **c.** mão-de-obra, **c.** marca própria, **c.** não financeira, **c.** «número de unidades industriais», **c.** «preços», **c.** do Português do Atlântico, **c.** privada, **c.** da procura interna, **c.** de produção, produtora, **c.** das receitas fiscais, **c.** remuneratória, **c.** da rentabilização, **c.** do trabalho barato, **c.** do trabalho clandestino, **c.** da venda direta, **c.** da oferta

m.

o componente.

c. comercial, **c.** excecional (o défice orçamental), **c.**vital (dividendo), **c.** essencial ao sucesso da nova empresa, **c.** no preço do produto, **c.** que oscilará entre os três e os oito por cento

Indústria

Em geral:

f.

a componente

c. aeronáutica, **c.** agrícola, **c.** ambiental, **c.** física, **c.** hidroelétrica, **c.** industrial, **c.** meteorológica, **c.** rodoviária, **c.** rodoferroviária, **c.** tecnológica, **c.** viária

Indústria alimentar:

m.

o componente

c. do milho utilizado no fabrico da cerveja, o «gripz», c. arroz, c. das rações para animais, c. da comida

Saúde, farmacologia, química

f.

a componente

c. alcoólica, c. biológica, c. laboratorial, c. clínica, c. da terapêutica hormonal, c. do sistema imunitário

m.

o componente

c. da alimentação humana, c. do ácido esteárico, c. analógico, c. da «cannabis» célula, c. central, c. chave, c. da dieta alimentar, c. do aminoácido dióxido de titânio, c. complexo, c. da doença de Alzheimer, c. específico, c. essencial, c. estrutural, c. dos glóbulos vermelhos, c. do GN (o etano, o propano, o butano), c. da hemoglobina, c. de uma imunidade protetora, c. da membrana das células sanguíneo, c. das membranas do cérebro, c. do núcleo dos átomos (neutrão), c. presente no vírus, c. principal, c. de uma proteína, c. psico-ativo, c. do sistema imunitário dos doentes, c. tóxico, c. do veneno, c. vital, c. do vírus

Informática, produtos electrónicos:

m.

o componente

c. áudio, c. da bomba de gasolina, c. do computador, c. do computador central, c. da direcção ou da suspensão, c. eléctrico, c. electrónico, c. de «hardware» c. informático, c. do motor, c. numa placa electrónica, c. com o preço mais..., c. de «software», c. do veículo, c. indonésio«standard»,

Referências bibliográficas:

BECHARA, Evanildo: (2001). "Moderna Gramática Portuguesa". Rio de Janeiro: Editora Lucerna.

BUZEK, Ivo: (2010). "La imagen del gitano en la lexicografía española". Brno: Masarykova univerzita.

- BUZEK, Ivo: (2011). "Historia crítica de la lexicografía gitano-española". Brno: Masarykova univerzita.
- CORBETT, Greville: (2011). "O género de alguns vocábulos "problemáticos" do português contemporâneo". In: *Revista Mundo & Letras*, nº 2. (pp. 85-97). São Paulo: F. José Bonifácio. [Disponível online em: http://www.revistamundoeletras.com.br/artigos2011/2011_Artigo07.pdf].
- FIÉVET, Anne-Caroline. PODHORNÁ-POLICKÁ, Alena: (2010). "Argot des jeunes français contemporain des cités en didactique du FLE/S: motivations des jeunes et limites des dictionnaires pour une étude des divergences socioculturelles" In: ABECASSIS Michaël, LEDEGEN Gudrun (éds.): *Les Voix des Français*. Volume I (pp. 159-174). Bern: Peter Lang.
- GOUVEIA, Maria Carmen de Frias: (1998). "Algumas mudanças de género em curso no Português". In: *Actas do XIII Encontro Nacional da Associação Portuguesa de Linguística*. Vol. I (pp.339-352). Lisboa: Associação Portuguesa de Linguística.
- GOUVEIA, Maria Carmen de Frias: (1998). "Algumas observações sobre a linguagem popular e regional no que se refere à categoria de género. Reflexos do género gramatical do Português antigo na linguagem popular". In: *Atti del XXI Congresso Internazionale di Linguística e Filologia Romanza. Morfologia e sintassi delle lingue romanze*. Volume II (pp. 339-349). Tübingen: Max Niemeyer Verlag.
- GOUVEIA, Maria Carmen de Frias: (1999). "A propósito do masculino genérico em Português". In: *Actas do XIV Encontro Nacional da Associação Portuguesa de Linguística*. Vol. II (pp. 21-28). Braga: Associação Portuguesa de Linguística.
- GOUVEIA, Maria Carmen de Frias: (2003). "O género dos estrangeirismos usados na língua portuguesa". In: *Actas do VIII Encontro Nacional da Associação Portuguesa de Linguística* (pp. 411-419). Lisboa: Associação Portuguesa de Linguística.
- GOUVEIA, Maria Carmen de Frias: (2004). "Considerações sobre a categoria gramatical de género. Sua evolução do latim ao português arcaico." In: *Biblos. Revista da Faculdade de Letras. Ocidente. Oriente*. Vol. II (2ª série), (pp. 443-475). Coimbra: FLUC .
- GOUVEIA, Maria Carmen de Frias: (2005). "Para uma descrição do género gramatical em Português". In: *Biblos. Revista da Faculdade de Letras. Cultura e Desenvolvimento*. Vol. III (2ª série), (pp. 201-246). Coimbra: FLUC.
- GOUVEIA, Maria Carmen de Frias: (2005). "A categoria gramatical de género do Português antigo ao Português actual". In: *Estudos em homenagem ao Professor Doutor Mário Vilela*. Vol. II (pp. 527-544). Porto: FLUP. [<http://ler.letras.up.pt/uploads/ficheiros/4584.pdf>].
- GOUVEIA, Maria Carmen de Frias: (2006). "O género dos nomes de animais em Português: descrição e história". In: *Biblos. Revista da Faculdade de Letras. Cidade(s) e Cidadania*. Vol. IV (2ª série), (pp. 381-396). Coimbra: FLUC.

- GOUVEIA, Maria Carmen de Frias: (2007). “Ainda o género gramatical dos substantivos e adjectivos em Portugal e no Brasil”. In: *Biblos. Revista da Faculdade de Letras. Universidade – um passado com futuro*. Vol. V (II série), (pp. 263-276). Coimbra: FLUC.
- GOUVEIA, Maria Carmen de Frias: (2008). “O género gramatical do português: da teoria à prática. Análise da atribuição de género por alunos do 1º Ciclo Universitário”. In: *Biblos, Revista da Faculdade de Letras. Ciências e/ nas Artes*. Vol. VI (II série), (pp. 221-250). Coimbra: FLUC.
- GOUVEIA, Maria Carmen de Frias: (2009). “Reflexos do português antigo na linguagem popular e regional: o género gramatical” In: *Biblos, Revista da Faculdade de Letras. Sociedade em Tumulto*. Vol. VII (II série), (pp. 429-449). Coimbra: FLUC.
- MATEUS, Maria Helena et.al.: (2003). “Gramática da Língua portuguesa” 6.ª edição. Lisboa: Editorial Caminho, S.A.
- NÁDVORNÍKOVÁ OLGA, PODHORNÁ-POLICKÁ Alena, ŠOTOLOVÁ Jovanka, VURM Petr: (2010). “Využití InterCorpu ve vysokoškolských kurzech francouzské filologie (motivace a cíle meziuniverzitní spolupráce romanistů při využití InterCorpu ve výuce)”. In: ČERMÁK František, KOCEK Jan (éds.): *Mnohojazyčný korpus InterCorp: Možnosti studia* (pp. 232-240). Praha: Nakladatelství, Lidové noviny a Ústav ČNK.
- PODHORNÁ-POLICKÁ, Alena: (2011). “L’expressivité et la marque lexicographique: étude comparative franco-tcheque d’un corpus du lexique non-standard. Les marques fam., pop., arg. vs expressivité en lexicographies française et tcheque.” In *Baider Fabienne, Lamprou Efi, Monville-Burston Monique. La marque en lexicographie. États présents, voies d’avenir* (pp. 209-225). Limoges: Lambert-Lucas.
- VILELA, Mário: (1999), “Gramática da Língua portuguesa”, 2.ª ed. Coimbra: Livraria Almedi.
- SANTOS, Diana: (2008). “Corporizando algumas questões”. In Stella E. O. Tagnin & Oto Araújo Vale (eds.), *Avanços da Lingüística de Corpus no Brasil* (pp. 41-66). USP, São Paulo: Editora Humanitas.
- ZAVADIL, Bohumil, ČERMÁK, Petr: (2010). „Mluvnice současné španělštiny. Lingvisticky interpretační přístup.“ Praha: Carolinum.

Dicionários consultados:

- Portugalsko-český slovník, Jindrová, Pasienska. Praha: Leda (2007).
- Česko-portugalský slovník, Jindrová, Hamplová. Praha: Leda (1997).
- Portugalsko-český slovník, Zdeněk Hampl: (1975).
- Dicionário Houaiss da Língua Portuguesa, Houaiss, Antônio, Mauro Villar. Lisboa: Círculo de Leitores: (2002).

Novo Dicionário Aurélio *versão 5.0 – Dicionário Eletrônico* [CD-ROM]. Positivo Informática: (2004).

Novo Aurélio Século XXI: *O Dicionário da Língua Portuguesa*. Ferreira, A. Buarque de Holanda. Rio de Janeiro: Nova Fronteira:(1999).

Dicionários consultados online:

<http://aulete.uol.com.br/>

www.priberam.pt

www.aurelio.pt

Os corpora utilizados:

www.linguateca.pt

www.corpusdoportugues.pt

www.korpus.cz

questões de ética
e política linguísticas

A SITUAÇÃO ATUAL DO ENSINO DA LÍNGUA PORTUGUESA NA GALIZA

Xurxo Fernández Carballido*

xurxo.fernandez@usc.es

A realidade do ensino da língua portuguesa na Comunidade Autónoma Galega é difícil de analisar pela própria realidade linguística, política e cultural do território situado ao norte do rio Minho. Nenhuma outra região do Estado espanhol nem, se calhar, do mundo apresenta uma situação comparável à galega.

A Galiza é um espaço de dobradiça entre o hispano e o lusófono. O território onde nasceu o galego-português tem hoje, por questões históricas, como uma das línguas oficiais o espanhol, para além do galego, a língua própria desse território.

No artigo é feita uma apresentação da situação atual do ensino da língua portuguesa nos diferentes âmbitos educativos, as dificuldades e as potencialidades.

Palavras-chave: Português Língua Estrangeira (PLE); presença da língua portuguesa na Galiza; glotopolítica da língua portuguesa.

The reality of Portuguese teaching in the Galician Regional Community is difficult to be analysed due to the linguistic, political and cultural reality of this land located north of the Minho River. No other Spanish region, maybe no other region in the world, challenges a situation like Galician's.

Galicia is a place between two worlds, the Hispanic and Lusophone. Besides Galician, its native language, the land where Galician-Portuguese was born has currently another official language, Spanish, due to historical reasons.

This paper analyses the current situation of Portuguese teaching in different educational environments, its difficulties and potentials.

Keywords: Portuguese as a foreign language (PFL); Portuguese language presence in Galicia; glottopolitics of Portuguese language.

*

Universidade de Santiago de Compostela. Centro de Línguas Modernas. Galiza.

1. Introdução⁽¹⁾

Os cidadãos galegos possuem a potencialidade natural de dominarem as duas línguas românicas mais importantes do mundo, português e espanhol. Contudo, a realidade é que a presença académica da língua portuguesa nos diferentes níveis educativos na Galiza não é muito relevante.

Com o objetivo de desvendar a situação real da presença da língua portuguesa na Galiza, primeiro é preciso comparar a presença da língua espanhola no sistema educativo português com a implantação da língua portuguesa no Estado espanhol.

Depois é analisada de maneira analítica e estatística a presença da língua portuguesa na Galiza segundo o nível de ensino: universitário, escolas oficiais de idiomas, secundário, básico e ensino não-formal.

A ordem estabelecida para esta descrição vem determinada pela história da penetração da língua portuguesa nos âmbitos académicos, dos departamentos das faculdades de letras, sobretudo desde a área de língua e literatura galego-portuguesa, até às atuais possibilidades de cursos on-line.

2. O espanhol no sistema educativo português. O português no sistema educativo espanhol.

Portugal e Espanha incorporam-se ao processo de construção europeu na década de 80. Esta integração era também linguística. De facto, a construção europeia é definida em grande medida pelos conhecimentos linguísticos dos cidadãos europeus, a denominada Europa das línguas. É por isso que a União Europeia, através das suas diretivas, recomenda o conhecimento de pelo menos duas línguas estrangeiras para além da materna. Iniciava-se um movimento de penetração da língua portuguesa em Espanha e da língua espanhola em Portugal.

Ao longo das últimas décadas a língua inglesa consolidou-se como a língua franca europeia e internacional. Já a eleição da segunda língua estrangeira nos diferentes sistemas educativos europeus tem que ver com a tradição política e económica de cada estado e, geralmente, também a questões de vizinhança ou fronteiriças (*Comissão Europeia*: 2012).

(1) Esta publicação baseia-se na dissertação de mestrado apresentada e defendida no Instituto de Letras e Ciências Humanas da Universidade do Minho no dia 14 de dezembro de 2012, sob o título de *O ensino da língua portuguesa na Galiza*, orientada pelo Professor Doutor da UM Álvaro Iriarte Sanromán.

No caso de Portugal e Espanha, os dois estados têm maioritariamente o francês como segunda língua estrangeira. Por uma questão de tradição, mais do que por razões económicas ou sociais, com excepção, se calhar, da forte emigração portuguesa às regiões francófonas da Europa – França, Suíça, Luxemburgo e Bélgica – ou das comunidades bascas e catalãs, divididas entre Espanha e França.

No entanto, em poucos anos a implementação da língua espanhola em Portugal está a ser notável quando comparada com a presença da língua portuguesa em Espanha. Segundo dados da *Consejería de Educación* da Embaixada de Espanha em Portugal (2012), a evolução do número de estudantes de espanhol no ensino básico e secundário entre o ano académico 1991-1992 e 2011-2012 é vertiginosa: de 35 alunos para 94.924.

Segundo o citado relatório oficial, o mesmo se passa na evolução do número de professores no período 1997-98 / 2011-2012, em que, de 28 professores, passou-se para 91. E se no ano académico 1991-1992 havia apenas 3 escolas a lecionarem espanhol, já no período 2011-2012 eram 617. Esta presença do espanhol no sistema educativo português caracteriza-se pela sua homogeneidade, sem se limitar ao espaço concreto das zonas fronteiriças.

Esta procura do espanhol nos últimos anos em Portugal está ligada à abertura das fronteiras, à desapareição das alfândegas e à movimentação de pessoas e mercadorias dentro do espaço europeu. Hoje, mesmo em período de crise, como o que estamos a viver, a ligação económica entre os dois estados ibéricos é muito forte. Espanha é o principal destino das exportações portuguesas, e Portugal tem um peso em destaque na balança comercial espanhola. Já no caso da Galiza, Portugal não é só o principal parceiro económico galego, como o recetor da maioria dos emigrantes que vivem e trabalham na Galiza, sem contar os trabalhadores portugueses e galegos que passam a fronteira a diário ou por empreitadas sazonais.

A respeito da presença da língua portuguesa no sistema educativo espanhol, a obtenção de dados é muito mais complexa, porque cada comunidade autónoma tem transferidas as competências em matéria educativa e porque a Embaixada de Portugal ou do Brasil não publicam relatórios anuais sobre o ensino do português em Espanha que facilitem o trabalho dos especialistas, planificadores e investigadores.

Em qualquer das hipóteses, e sublinhando a eventualidade dos números, pode ser representativo, segundo dados da *Consejería de Educación*⁽²⁾ da

(2) No Estado espanhol, equivalente autonómico ao Ministério da Educação português.

Junta de Extremadura, o número de 18.000 estudantes de português para o ano académico 2008-2009.

Números surpreendentes numa população envelhecida de 1.000.000 de habitantes, quando contrastados com os perto de 620 alunos que estudam português em ESO (Ensino Secundário Obrigatório) e Bacharelato na Galiza⁽³⁾, com uma população de 2.300.000 habitantes, segundo a *Consellería de Educación da Xunta de Galicia*. Estes dados foram divulgados pelo Governo galego em 2012 no Parlamento da Galiza em resposta parlamentar a uma pergunta sobre a presença do ensino do português no sistema de ensino obrigatório galego do partido político Bloque Nacionalista Galego.

Contudo, que cada comunidade autónoma do Estado espanhol tenha autonomia no desenho do seu sistema educativo faz com que se possam cobrir as necessidades específicas de cada espaço geográfico. Desde esta perspectiva, de todas as comunidades autónomas espanholas fronteiriças com Portugal –Galiza, Castela e Leão, Andaluzia e Extremadura– apenas esta última implementou todas as medidas ao seu dispor para fazer da sociedade extremeña uma região com alto nível de conhecimentos em língua portuguesa, com todas as potencialidades económicas, sociais, culturais, etc., que daí se derivarem.

De facto, com o *Plano Linguaex*, a *Junta de Extremadura* tentou potenciar novos projetos linguísticos para além dos já consolidados: “El creciente desarrollo de relaciones sociales, económicas y culturales con nuestro vecino Portugal, aconsejan otorgar un lugar privilegiado a la enseñanza-aprendizaje de esta lengua.” (Junta de Extremadura. Plano Linguaex 2009-2015, 2008: 15).

A *Junta de Extremadura* promoveu e apoiou numerosas iniciativas, com grandes esforços em recursos humanos, técnicos e económicos. Desde a convocatória de vagas de professores à publicidade institucional. A campanha “*Aprende português. Te abrirá muchas puertas*”⁽⁴⁾, foi o ícone desta política linguística.

A ambição destas linhas programáticas de política linguística estão a dar os seus frutos na Extremadura. Neste sentido, as autoridades portuguesas deveriam fazer maior pressão perante os governos espanhol e as respetivas autonomias, pelo menos as fronteiriças, para que o português ganhe a dimensão que deveria ter.

(3) O segundo ciclo do ensino básico português corresponde a 5º e 6º da Primaria em Espanha. O terceiro ciclo do ensino básico em Portugal corresponde-se com o 1º, 2º e 3º de ESO. O ensino secundário português equivale a 4º de ESO e o Bacharelato.

(4) <http://www.youtube.com/watch?v=NQ2fUZT0vOo>

A este respeito, e no contexto de crise económica generalizada no que é redigido este trabalho, a política linguística portuguesa no exterior não é a melhor, bem pelo contrário, e decisões como a de centralizar as atividades do Instituto Camões na Embaixada de Espanha em Madrid são uma catástrofe.

3. Presença do português na Galiza segundo o nível de ensino

Ao contrário do que foi analisado no caso da Extremadura, na Galiza não houve qualquer tentativa de promoção da língua portuguesa, quer no sistema educativo, quer na sociedade em geral. Várias podem ser as razões que expliquem este estado de inanição, alguns linguísticos, outros políticos ou ideológicos.

A proximidade linguística entre galego e português ou espanhol e português, faz com que a sociedade galega acredite geralmente que a aprendizagem da língua portuguesa não é necessária, porque partem do princípio de que com o galego chega para se “desenrascarem”.

Também parte de uma experiência generalizada a consideração de que em Portugal não é preciso mudar de código linguístico, porque os portugueses adaptam-se com facilidade a se comunicarem em espanhol.

Na Galiza, o denominado conflito normativo, pelo qual se confrontavam duas visões da língua galega diferentes, dentro do sistema linguístico galego-português ou fora dele, fez com que o ensino da língua portuguesa fosse visto como uma ameaça de origem reintegracionista, a corrente que pretende reintegrar o galego no tronco comum da língua portuguesa.

Por outro lado, alguns dos denominados setores lusistas da sociedade galega ainda hoje não acabam de assumir que a língua portuguesa seja apresentada como língua estrangeira na Galiza, que assim seja focada, desde critérios de profissionalidade, rigorosidade e indo ao encontro dos requerimentos e necessidades dos alunos, fora da conflituosidade linguística que caracteriza a sociedade galega.

Das motivações ideológicas para a não promoção da língua portuguesa na Galiza, a mais importante é o facto de a Comunidade Autónoma Galega ser governada, em quase todo o período democrático, por formações políticas de ideologia espanholista, que interpretam a promoção da língua galega e, por extensão, da língua portuguesa, uma marca política dos grupos nacionalistas da esquerda galega. Aliás, a promoção da língua portuguesa

levanta em alguns âmbitos sociais, políticos, educativos e ideológicos ainda muitas suspiçcias, interpretando estes grupos como um cavalo de Troia na espanholidade da Galiza a promoçao da língua portuguesa.

Sem este quadro prévio não podemos compreender muitos dos debates e resistências que existem à volta da implementação e expansão da língua portuguesa na Galiza. É significativo que o argumentário pareça estar sempre a favor da presença do português na sociedade galega, pois é um discurso aparentemente assumido pelas elites políticas e sociais da Galiza. De facto, o atual presidente do Governo galego, Alberto Núñez Feijóo, chegou a comprometer-se a estudar a introdução do português no ensino secundário, demandada pelo na altura embaixador português no Reino de Espanha, Álvaro de Mendoza e Moura, no ano 2009, segundo noticiava o jornal galego *Galicia Hoxe*⁽⁵⁾.

Ou seja, na Galiza há um discurso institucionalizado mas vácuo sobre a geminação e as relações seculares entre os dois povos. Por exemplo, perante a proposta na Comissão de Educação do Parlamento da Galiza por parte do Grupo Parlamentar do Bloque Nacionalista Galego⁽⁶⁾ para ofertar português como segunda língua estrangeira em todas as escolas públicas do secundário, a iniciativa foi chumbada por demonstrar, segundo o Grupo Parlamentar do Partido Popular –grupo maioritário na Câmara galega– “complexo ideológico”⁽⁷⁾, e pondo ao mesmo nível a aprendizagem de hndi ou chinês com o português. Isto é, a língua portuguesa é interpretada em clave ideológica no jogo político galego, representa no quadro nacional galego algo mais do que uma língua.

Esta dupla linguagem à volta das vantagens que supõe o conhecimento do português e o debate identitário sobre o papel da língua portuguesa no desenvolvimento da consciência coletiva galega e no processo de construção nacional galego, volta uma e outra vez.

Assim, mesmo o atual presidente do governo galego, Alberto Núñez Feijóo, conhecido pelas suas resistências e restrições à língua galega através de numerosas iniciativas legislativas e políticas, a mais conhecida e polémica o *Decreto 79/2010, de 20 de maio, para o plurilinguismo no ensino não universitário da Galiza*, quando se sente acurrulado pelas posições mais

(5) <http://www.galiciahoxe.com/ultima-hora/gh/fejoo-comprometese-estudar-introduccion-do-portugues-no-ensino-demanda-lle-expuxo-embaixador-luso-espana/idNoticia-433210/>

(6) <http://www.parlamentodegalicia.es/buscaVideos/asx.aspx?wmv=0000007334-002-30kb.wmv&timecode=01:17:17&duration=00:25:47>

(7) http://arquivo.bng-galiza.org/opencms/opencms/BNG/parlamentogalego/contidos/novas/novas/nova_2750.html

radicais e ultramontanas contrárias à língua galega recorre ao discurso litúrgico sobre as potencialidades da língua galega no espaço lusófono⁽⁸⁾.

3.1. Ensino da língua portuguesa no ensino universitário

Tradicionalmente os estudos de língua e literatura inseriam-se no histórico curso de Filosofia e Letras, onde os estudos de língua, literatura e cultura portuguesa têm uma larga tradição, com a criação do Instituto da Língua Portuguesa nos inícios do século XX na Universidade de Santiago de Compostela.

Com a chegada da democracia ao Estado espanhol e a democratização do ensino público universitário criaram-se novos cursos, como Filologia Hispânica Secção Galego-Português, mas o número de cadeiras de língua portuguesa era muito reduzido pelo que os formandos ficavam com um conhecimento muito superficial e rudimentar da língua portuguesa.

Só no ano 1994 se criou o curso de Filologia Portuguesa na Universidade de Santiago de Compostela, embora após o Plano Bolonha agora esteja integrado no *Grau*⁽⁹⁾ em *Línguas e Literaturas Modernas* com o itinerário em Língua Portuguesa e Literaturas Lusófonas. O curso centra-se em aspetos da língua como gramática histórica galego-portuguesa, dialetologia, gramática, debate normativo, lírica medieval galego-portuguesa, literatura portuguesa clássica, moderna e contemporânea, literaturas lusófonas, etc. Não há, no entanto, cadeiras específicas de didática do PLE ou de elaboração de materiais didáticos, por exemplo.

Na Universidade de Vigo existe o Grau de Estudos em Galego e Espanhol, onde quer os estudantes de galego, quer os de espanhol, devem aprovar obrigatoriamente duas cadeiras semestrais de língua portuguesa no primeiro ano do curso. Também na Universidade de Vigo existe o Grau em Tradução e Interpretação onde há várias cadeiras de tradução e interpretação português / galego e português / espanhol. Embora não seja possível ter a língua portuguesa como língua principal do curso, mas inserida nas especialidades de inglês ou francês.

Na Universidade da Corunha o Grau em Galego e Português obriga a cursar cadeiras de língua e literatura portuguesa no seu plano de estudos.

(8) <http://www.youtube.com/watch?v=MqQVCQehCqs>

(9) Curso, no sistema universitário português.

Por seu lado, as três universidades galegas têm um serviço de línguas denominados Centro de Línguas Modernas (CLM) e que oferecem cursos regulares de línguas para a comunidade universitária.

No caso da USC, o CLM também é responsável pelas provas necessárias que acreditam o nível exigido para as bolsas Erasmus e outros programas de intercâmbio interuniversitário, como o Europracticum, e de colaboração com o Brasil e outros países da Lusofonia.

O CLM da USC também é responsável pela elaboração da Prova de Competência Linguística (PCL) que acredita o nível B1 de uma língua estrangeira, obrigatório para ser licenciado em qualquer curso da universidade compostelana.

Como amostra significativa da evolução da língua portuguesa no espaço universitário, a continuação vão ser analisados um pouco mais detalhadamente os dados da evolução da secção de português do CLM da USC, desde a sua criação no ano académico 2006-2007 até o ano 2011-2012.

Tabela 1: Evolução da língua portuguesa no CLM da USC

Ano	Nível 1 (A1-A2)	Nível 2 (B1)	Curso de verão (A1-A2)	Prova Erasmus	Prova Europracticum ¹	PCL ²
06-07	15	8	10	180	24	–
07-08	9	4	9	196	–	–
08-09	18	4	13	142	43	–
09-10	32	7	20	489	28	36
10-11	26	11	5	492	71	39
11-12	38	47	19	428	70	31
Total	138	81	76	1927	236	106

Repare-se que o CLM da USC foi criado em 1975 com a denominação de Instituto de Línguas, mas o português só se ofertou oficialmente na data referida, 2006-2007, sendo uma das últimas secções incorporadas, junto com chinês.

Ao lado da análise numérica, também se deve fazer uma pequena referência às circunstâncias laborais da mesma, pois é, evidentemente, uma questão que afeta de maneira direta o desenvolvimento da secção. Ao longo destes anos a instabilidade laboral e a mesma continuação da secção foi

posta em causa por mudanças nas políticas de direção e pela própria instabilidade financeira da instituição universitária.

Apesar destas circunstâncias tão adversas, a procura da língua portuguesa no CLM parece consolidar-se ao longo dos anos.

3.2. Ensino da língua portuguesa nas escolas oficiais de idiomas

É nas escolas oficiais de idiomas (EOI) onde se centra a maioria das pessoas que estudam língua portuguesa na Galiza. Este sistema estatal de ensino de línguas para adultos tem na Galiza 10 centros no total, dos quais 8 ofertam português –Lugo, Ourense, Ferrol, Santiago de Compostela, Vigo, Vilagarcia e Pontevedra– e em três –Monforte de Lemos, Ribadeo e Viveiro– ainda não existe a opção.

Segundo os dados elaborados pelo Portal Galego da Língua (PGL)⁽¹⁰⁾ nas EOI da Galiza no ano académico 2011/2012 havia na Galiza 1597 alunos a estudar português⁽¹¹⁾.

Tabela 2: Evolução dos alunos de português nas EOI no período 2010-2012.

Escola Oficial de Idiomas	Ano académico 2010-2011	Ano académico 2011-12	Varição percentual
EOI Corunha	199	220	10%
EOI Ferrol	119	145	22%
EOI Lugo	142	175	23%
EOI Ourense	121	131	8%
EOI Pontevedra	111	144	23%
EOI Santiago de Compostela	283	340	20%
EOI Vilagarcia	78	90	15%
EOI Vigo (com delegação em Tui)	295 (72)	352 (85)	19%
Total alunos	1348	1597	18%

(10) <http://www.pglingua.org/>

(11) Neste vídeo estudantes de EOI explicam os motivos para estudarem português: <http://www.youtube.com/watch?v=A2jnWHZjHRs>

Apesar dos bons dados da língua portuguesa em números absolutos e em percentagem, quer na sua evolução anual quer quando comparados com dados de outras línguas, o governo autonómico não alarga o ensino da língua portuguesa nas EOI.

Quer dizer, línguas em franco retrocesso como o francês, sem apoios político-sociais como o italiano, muito afastadas linguística e culturalmente como o alemão ou o chinês, estão a ter maior apoio institucional face a uma língua com alta procura, com fortes apoios e muito próxima geográfica, linguística e culturalmente. Estes acontecimentos refletem, mais uma vez, a ideia de que para o governo galego o português não é uma prioridade educativa porque “não faz falta”.

3.3. Ensino da língua portuguesa no ensino secundário

Para além do inglês, que funciona de facto como a primeira língua estrangeira, administrativamente os centros podem oferecer alemão, francês, italiano e português como segunda língua estrangeira. No entanto, criar uma matéria de segunda língua estrangeira costuma ser um processo longo e com uns procedimentos burocráticos labirínticos.

Segundo a experiência dos docentes de língua portuguesa, dentro das razões para a criação da matéria de língua portuguesa como segunda língua estrangeira num centro de secundária destacam:

- Vontade e empenhamento pessoal do docente. Normalmente tem que lutar para que essas horas de língua portuguesa sejam compatíveis com os seus horários, com os horários dos colegas e contra outras possibilidades linguísticas, pois os centros raras vezes têm capacidade numérica de ofertar várias línguas.

- Contato com outros colegas ou escolas. Conhecer outras pessoas, a Associação de Docentes de Português na Galiza (DPG)⁽¹²⁾ ou centros que lecionam português serve muitas vezes como rastilho para que o docente se empenhe para a implementação desta cadeira no seu centro de ensino.

- Novas perspetivas profissionais e pessoais para o docente. Os docentes das cadeiras de língua e literatura galegas são os que costumam criar e lecionar português nas escolas do secundário. Lecionar galego na Galiza não é, às vezes, muito gratificante pela situação de postergação em que vive a língua e a cultura galega, pelo que a nível profissional muitos docentes sentem que a cadeira de português pode ser um revulsivo. Por vários moti-

(12) <http://www.dpgaliza.org/>

vos, porque alarga os seus conhecimentos e possibilidades e porque, ao mesmo tempo, os estudantes enquadram a língua galega desde uma outra perspetiva, em termos de dimensão e de utilidade.

Face a esta vontade de criação detetam-se vários impedimentos:

- Falta de apoio da autoridade educativa da Administração pública galega.

- Ausência de especialistas da cadeira e falta de convocatórias públicas de vagas da especialidade de língua portuguesa.

- Oposição de colegas e/ou direção da escola. As razões para esta oposição podem ser diferentes: colegas de outras segundas línguas que veem ameaçado o seu horário ou estabilidade no emprego; consideração de que o português não é uma língua importante para a formação dos alunos.

- Insuficiência no número de horas disponíveis do docente voluntário para lecionar a disciplina.

- Excesso de disciplinas opcionais já oferecidas pela escola.

- Desinformação, desinteresse ou desvalorização em relação à importância da disciplina de língua portuguesa dos estudantes, mães e pais.

Mas uma vez criada a disciplina e as turmas de língua portuguesa na escola, verifica-se que a sua continuidade ao longo dos anos académicos é muito complexa, desaparecendo muitas vezes essa disciplina após um ou dois anos. As principais razões para a sua desaparecimento são, segundo os depoimentos do docentes de português através da DPG:

- Número insuficiente de alunos. No ensino público galego, os rácios para escolha de cadeiras são cada vez maiores, mas por causa da queda demográfica das últimas décadas, cada ano há menos estudantes nas escolas. Portanto, é difícil preencher o número mínimo para formar uma turma.

- Deslocação do docente que iniciou o processo para a implementação da cadeira de língua portuguesa. A deslocação dos docentes é muito elevada no sistema educativo galego, as razões são múltiplas, necessidades da administração ou interesse pessoal do docente. Como consequência, o professor de português numa escola, depois de um complexo processo, pode ser que não continue no ano seguinte. O que já desencoraja muitos docentes a iniciar a burocracia para abrir uma turma de português.

- Ausência de apoio da direção da escola. Dentro das lutas de interesses dos claustros de professores e na rotação de direções é frequente a queda de disciplinas e maiores apoios a umas secções em detrimento de outras.

Mesmo assim, em termos gerais há alguns fatores que favorecem –ou favoreceriam, se existisse vontade política– a criação de turmas de língua portuguesa nas escolas secundárias da Galiza:

– Docentes competentes. O capital humano é dos principais valores que tem a implantação da disciplina de língua portuguesa na Galiza, sobretudo, os que já lecionam galego, pois a sua reciclagem é rápida e efetiva, embora não se possam desvalorizar as dificuldades.

– Uso imediato e efetivo da língua. Esta é a principal vantagem dos estudantes galegos face a alunos de outras regiões ou em comparação com outras línguas.

– Sensação de maior facilidade. Embora o aluno galego, regra geral, tenha um horizonte de espetativas muito superior ao conhecimento real da língua portuguesa, é certo que ultrapassadas as primeiras impressões, o estudante galego tira maior proveito do seu esforço do que no caso de outras línguas. Também em relação a esta questão vai ser determinante a focagem do docente, que deve priorizar a focagem comunicativa à história da língua ou à dialetologia.

– Proximidade. Em todos os sentidos e para tudo. Proximidade na realidade social, cultural e linguística, sobredimensionada pela Lusofonia, mas também na possibilidade de realizar projetos educativos de intercâmbio, viagens, estadias, etc.

Toda esta análise pertence ao âmbito do ensino público galego, mas é preciso ter em consideração o grande número de estudantes galegos que assistem a colégios concertados e privados. Estes centros educativos costumam ter visões da educação mais tradicionais, mas se a Administração pública impelisse a presença da língua portuguesa nos centros públicos, também os concertados e privados se somariam à iniciativa.

Resulta surpreendente que nem a própria Administração responsável pelo ensino saiba o número de estudantes. A *Consellería de Educación* é incapaz de oferecer um dado exato sobre e os centros que lecionam português e o número de alunos que frequentam a cadeira de língua portuguesa. Assim, é a DPG, uma simples associação de docentes, a encarregar de elaborar de maneira amadora e com grandes esforços inquéritos que possam servir de referência para o planeamento das próprias atividades da DPG, para conhecer as necessidades dos docentes, dos centros e dos alunos, e para, em definitiva, conhecer a situação real da língua portuguesa no sistema público galego.

Ao realizar as informações de forma não institucionalizada, há muitos centros que não respondem aos inquéritos, por esquecimento ou falta de interesse. O outro ponto significativo dos dados, é o alto número de centros que não mantêm continuidade de um ano académico para o outro, o que está diretamente ligado com os mecanismos de como são oferecidas estas cadeiras nas escolas.

São aproximadamente 40 os centros a lecionarem português, mas com problemas de continuidade, ao depender exclusivamente da vontade e disponibilidade do ou da docente. A mudança de destino do professor para outra escola implica que a turma de português, independentemente do número de alunos, desapareça. Isto faz com que a disciplina de português, acabe por ser uma cadeira “voluntarista”.

Outro desafio é a passagem do básico para o secundário, pois há muitas escolas do secundário onde não há possibilidade de continuar língua portuguesa aos estudantes que já o fizeram no básico. Também existe o problema de continuidade dentro do próprio percurso do estudante dentro da mesma escola do secundário, com o qual o aluno de português de um ano pode deparar-se com a falta de português no ano a seguir, e isto é sempre uma grande barreira para os professores angariarem alunos e para os alunos cursarem esta disciplina.

Face a esta situação, no ano 2010 foi apresentada uma proposta por todos os representantes sindicais do ensino secundário e apoiada pela DPG e que, como já foi explicado, foi rejeitada em Comissão Parlamentar pelo Grupo Parlamentar do Partido Popular, grupo que detém a maioria do Governo, com dois pontos básicos:

- Implementar a língua portuguesa como segunda língua estrangeira no sistema educativo obrigatório na Galiza.
- A convocatória de vagas específicas de docentes de língua portuguesa.

Estas tomadas de posições político-linguísticas põem em causa o tão cacarejado plurilinguismo linguístico que as atuais autoridades governamentais galegas dizem desejar e reduzem a nada os discursos litúrgicos e vazios da irmandade e da geminação galego-portuguesa que as autoridades galegas, e também portuguesas, sempre utilizam em momentos protocolares.

Enquanto não houver uma política linguística decidida em prol da língua portuguesa, em todos os níveis de ensino, os discursos sentimentais, saudosistas e passadistas são retórica. As oportunidades passam à frente dos galegos, outros não duvidam em apanhá-las. Em relação com esta situação, em março de 2013 foram apresentadas mais de 17.000 assinaturas em abaixo-assinado no Parlamento da Galiza, na Iniciativa Legislativa Popular Paz Andrade para a promoção da língua portuguesa no ensino galego⁽¹³⁾.

(13) <http://www.pglingua.org/noticias/informante/5454-ilp-paz-andrade-atingiu-17000-assinaturas-e-esta-tarde-serao-entregues-no-parlamento-da-galiza>

3.4. Ensino da língua portuguesa na Galiza. O ensino básico

Segundo os dados elaborados pelo Instituto Camões de Vigo em colaboração com a DPG para o curso 2009-2010 havia 433 alunos de português na província de Pontevedra, em duas escolas, e 972 na província de Ourense, distribuídos por nove escolas. Evidencia-se a ausência da disciplina de língua portuguesa nas províncias da Corunha e Lugo e a concentração de alunos nas zonas de forte presença migratória portuguesa na província de Ourense e o caso isolado, de entre todos os outros concelhos arraianos, do concelho minhoto de Tominho.

Estes dados são ainda mais interessantes quando comparados com a presença do português no ensino secundário, que ilustra o que já foi analisado na secção sobre este nível formativo, a falta de continuidade entre o básico e o secundário da disciplina de língua portuguesa.

A maioria dos centros do básico que lecionam português estão inseridos no Programa de Língua e Cultura Portuguesa. Um programa dirigido aos estudantes lusodescendentes e a estudantes galegos dos centros onde o Programa é desenvolvido. Os centros aderidos ao Programa adscvem pessoal docente português enviado pela Embaixada Portuguesa, através do Instituto Camões, com o objetivo de manter a língua e a cultura portuguesa entre as comunidades lusodescendentes, mas também entre os alunos autóctones.

Segundo informações telefónicas que pude apurar desde a *Xunta de Galicia*, para o atual ano 2012-2013, 11 escolas participaram no programa, 10 do básico e uma do secundário, e o Ministério da Educação português enviou 4 professores.

O Programa parece interessante e muito positivo, sobretudo se se puderem alcançar dois objetivos: que os lusodescendentes possam conhecer a língua portuguesa e que os alunos galegos também a adquiram. Mas os dados revelam que a implantação do Programa é insuficiente. A comunidade portuguesa é a de maior importância numérica na sociedade galega, e devem-se acrescentar a importante presença de brasileiros e cabo-verdianos. Quer dizer, 4 docentes e 11 escolas, não refletem a verdadeira dimensão da presença lusófona na Galiza, pelo que o objetivo deveria ser muito mais ambicioso: a implantação da língua portuguesa como segunda língua estrangeira em todos os centros de ensino da Galiza.

3.5. Ensino da língua portuguesa no ensino não-formal

Neste artigo é utilizado o conceito de ensino não-formal para toda a formação oferecida fora do percurso do ensino obrigatório, do ensino superior e das EOI.

Segundo a análise realizada por Figueiredo Capuz (2012:115) para este tipo de ensino não-formal na Comunidade Autónoma da Extremadura, mas adaptando-a à realidade social e administrativa galega, podem-se diferenciar âmbitos de formação que vão ao encontro das necessidades específicas de cada espaço social:

- Formação ocupacional. Destinada a setores profissionais específicos que possam ter contato profissional direto com pessoas de língua portuguesa, médicos, enfermeiros, empregados de balcão, etc.

- Formação contínua. Cursos focados a pessoas em contato com a utilização da língua portuguesa ou que desejam melhorar o seu currículo.

- Escola Galega de Administración Pública (EGAP). Esta instituição da Administração autonómica galega organiza cursos de formação para o funcionariado público, que faz todo o sentido no contexto da Euro-região Galiza-Norte de Portugal.

- Organizações transfronteiriças. Organismos como o Eixo Atlântico ou os gabinetes de iniciativas transfronteiriças deveriam ter uma maior preocupação pelo conhecimento linguístico dos cidadãos da euro-região.

- Sindicatos e associações empresariais. Estas organizações têm um contato direto com a realidade laboral e económica, pelo que para procurar novas oportunidades a competência linguística é importante.

- Associações populares. Estas entidades poderiam ter um papel muito importante na realização de cursos de língua portuguesa nos seus âmbitos de atuação. De facto, a relação entre associações galegas e portuguesas é muito frequente, o mesmo se passa com as representações sindicais e empresariais, mas há a impressão de que essas relações, pela parte galega, são estabelecidas maioritariamente em espanhol e assim secundadas pelas congéneres portuguesas.

Na Galiza, ao contrário do verificado na Extremadura, apenas o setor do associativismo está a desenvolver projetos de ensino e aprendizagem de língua portuguesa. De facto, estão a desenvolver e implementar projetos que tentam paliar as graves deficiências e inibições das diferentes administrações públicas.

De todas elas, é a Associação Galega da Língua (AGAL) a que maior trabalho, iniciativas e sucesso está a ter no desenho e programação de cur-

sos de português de todo o tipo, independentemente ou em parceria com outras organizações. Este trabalho colaborativo entre entidades conhecedoras de realidades complementares deveria ser incrementado, em atividades presenciais ou em formato virtual

3.5.1. Ensino e-learning

A respeito do formato dos cursos, estão-se a consolidar duas possibilidades de ministrá-los, presenciais ou e-learning, dois sistemas compatíveis e que podem ser frequentados por públicos similares ou de perfis muito diferentes, embora nos últimos tempos as barreiras entre usuários e-learning e presenciais sejam cada vez mais etéreas.

Assim, foi criado *Português para nós*⁽¹⁴⁾ um curso de português on-line adaptado às necessidades específicas dos galegos, livre e gratuito. Elaborado pelo grupo Galabra –Grupo de Estudos nos Sistemas Culturais Galego, Luso, Brasileiro e Africanos de Língua Portuguesa– da Universidade de Santiago de Compostela, empresa Imaxin Software e no que também colaborou o professor da UM, Álvaro Iriarte Sanromán.

Após este curso, também desde a plataforma da AGAL e o Grupo Galabra elaborou-se o curso *Português no prato*⁽¹⁵⁾, que segue, a grandes traços o sistema de trabalho descrito com antecedência, mas orientado ao setor dos empregados e funcionários da hotelaria e restauração. Nesse curso participaram 330 alunos em 2011, o que faz ideia da importância dos cursos e-learning.

Também a AGAL, em colaboração com o sindicato Confederación Intersindical Galega (CIG), organiza cursos de português de nível básico para docentes on-line com a participação de 320 e 98 pessoas, nas duas primeiras edições de 2011.

As novas dimensões dos cursos e-learning garantem inúmeras possibilidades aos utentes mas fazem mais complexas as investigações académicas. É muito difícil ou mesmo impossível conhecer todos os cursos que possam existir em linha ou que nível de presença galega há neles. Se calhar os mais interessantes sejam os feitos através da plataforma de ensino à distância do Instituto Camões que oferecem até três cursos de língua portuguesa por semestre e em cujos dados não se assinala a origem dos alunos mas a sua

(14) <http://www.portuguesparanos.org>

(15) <http://portuguesnoprato.com/>

nacionalidade institucional, quer dizer, um galego é indexado como espanhol, com independência do grupo linguístico a que pertencer.

3.5.2. Cursos presenciais

Nas sociedades contemporâneas a administração pública traça as linhas prioritárias em política linguística, estabelece os objetivos a atingir e desenha as medidas para que o conjunto da sociedade chegue a esses objetivos. Se a língua portuguesa fosse uma prioridade para a Administração galega faria o preciso para que a sua presença estivesse assegurada no ensino formal e não-formal.

Na Galiza este empenhamento político nunca teve a intensidade necessária para que tivesse uma presença continuada. De facto, as únicas tentativas foram elaboradas a partir do Projeto Mobilitas, através do programa europeu Interreg III, com cursos de língua portuguesa em diferentes cidades e vilas da Galiza para um público geral.

Com a inibição das diferentes administrações –câmaras municipais, deputações e governo autonómico–, é o associativismo, especialmente a AGAL, a que organiza grande número de cursos, focados de preferência ao público juvenil. Os cursos OPS (O Português Simplês), é um curso de apresentação da língua portuguesa para estudantes galegos do secundário. Durante o ano 2011 foram realizados 73 cursos OPS com a participação de 2122 estudantes.

Já no ano 2012 a oferta também foi alargada às escolas do básico com os cursos Cacimbo, oficinas para que os meninos e meninas tenham o primeiro contato direto com a língua portuguesa e dos que já se realizaram 37 ateliês.

Em geral, estes cursos são focados para que os estudantes tenham uma experiência em português no âmbito das suas escolas, pois para muitos pode ser o único contato com esta língua ao longo de toda a sua vida escolar e, por outro lado, a contratação das oficinas parte normalmente dos professores de galego, que apreciam nestes ateliês o reforço linguístico positivo e, sobretudo, sociolinguístico ao seu trabalho com a língua galega.

A AGAL também organiza cursos de português na cidade do Porto em colaboração com a Universidade do Porto sob a denominação *aPorto*⁽¹⁶⁾ e que, embora sejam cursos presenciais em Portugal, são ideados para galegos. O curso é uma estadia mínima de uma semana com especial atenção à expressão oral. Durante o verão de 2011 foram 99 os inscritos nestes cursos.

(16) www.aporto.org

4. Conclusões

Ao longo deste artigo foi analisada a situação atual da língua portuguesa no sistema educativo galego, do âmbito universitário ao sistema obrigatório de ensino, formação para adultos e não-formal.

A história da presença da língua portuguesa no ensino galego teve início no âmbito universitário, ligado aos temas filológicos galego-portugueses, e sofreu uma expansão a outros níveis, EOI, ensino secundário, básico, e nos últimos tempos ao ensino não-formal, sobretudo de identidades ligadas a organizações sociais, e ao ensino e-learning.

Também é decrescente a presença da língua portuguesa segundo a franja etária e sócio-cultural, regra geral, quanto maior nível de formação e maior idade, maiores conhecimentos em português.

Nas três universidades galegas há uma preferência de temas linguísticos e literários nos cursos de línguas modernas. O mesmo se passa nos cursos de tradução e interpretação. Por outro lado, os CLM das três universidades oferecem uma formação em português com uma perspectiva comunicativa.

No caso galego, é na formação para adultos que há maior número de estudantes de português na Galiza. Nas EOI os estudantes podem atingir o nível B2 de português ao longo de 6 cursos, sempre desde uma ótica comunicativa da língua.

O ensino obrigatório, quer público quer concertado ou privado, representa o calcanhar de Aquiles da implantação da língua portuguesa no sistema de ensino galego. No ensino secundário, a cadeira de língua portuguesa está normalmente ligada ao professorado e à cadeira de língua galega, embora haja turmas independentes. Por outro lado, há aspetos, que sem terem a denominação específica de matéria de língua portuguesa, também estão presentes nos conteúdos linguísticos e literários transversais à cadeira de língua e literatura galega.

No ensino básico há poucas turmas e sem possibilidade de continuidade no secundário. Isto representa um gravíssimo problema, porque mesmo as turmas ligadas a zonas de presença emigratória portuguesa e incluídas no Programa de Língua e Cultura Portuguesa, não podem continuar e acrescentar os seus estudos nos centros do secundário que lhes são atribuídos.

Nos últimos tempos tem ganhado muito peso o denominado ensino não-formal, especialmente o ensino à distância ou e-learning. Pela própria natureza destes cursos é difícil saber a importância do impacto desta oferta educativa na sociedade galega por parte de instituições públicas ou privadas lusófonas, especialmente o Instituto Camões e os seus cursos à distância.

Na Galiza é preciso sublinhar a importância dos cursos on-line desenhados e organizados por organizações sociais, especialmente a AGAL. Em qualquer dos casos é uma linha de trabalho que se perspetiva com muito futuro, pelo que os cursos deveriam ir ao encontro das necessidades dos utentes, ainda que o processo se veja dificultado sem o apoio das administrações públicas.

Em geral, a presença do português na Galiza continua a ser muito fraca em qualquer tipo de ensino, levada a cabo em muitos casos de maneira voluntarista e intuitiva, perante a inibição e falta de apoio das instituições públicas.

5. Referências bibliográficas

- BARRASA, Daniel (2009), “El portugués será el segundo idioma extranjero en los colegios extremeños”, *El Periódico de Extremadura*, disponível em http://www.elperiodicoextremadura.com/noticias/extremadura/el-portugues-sera-segundo-idioma-extranjero-en-los-colegios-extremenos_434378.html consultado em 24/03/2009.
- CAPUZ FIGUEIREDO, Javier (2012), “La enseñanza no reglada de la lengua portuguesa en Extremadura” págs. 111-121, em *Tejuelo*, nº 14, Consejería de Educación, Ciencia y Tecnología de la Junta de Extremadura.
- CHANCELETE DE MACHETE, Rui e António Luís VICENTE (2010), *Língua e cultura na política externa portuguesa. O caso dos Estados Unidos da América*, Lisboa, Fundação Luso-Americana para o Desenvolvimento.
- COMISSÃO EUROPEIA (2012), *First European Survey on Language Competences*, disponível em <http://ec.europa.ec/languages/eslc/docs/en> consultado em 28/06/2012.
- CONSEJERÍA DE EDUCACIÓN DE LA EMBAJADA DE ESPAÑA EN PORTUGAL (2012), *La enseñanza de la lengua española en Portugal. Curso 2011 / 12. Escuelas Públicas de Enseñanza Básica y Enseñanza Secundaria*, disponível em <http://www.educacion.gob.es/portugal/dms/consejerias-exteiores/portugal/publicaciones/informe-el-esp-a-ol-en-portugal-2011-2012/informe%20el%20espa%C3%B1ol%20en%20portugal%202011%202012.pdf> consultado em 16/07/2012.
- CONSEJERÍA DE EDUCACIÓN DE LA JUNTA DE EXTREMADURA (2008), *Plan Linguaex 2009-2015*, disponível em http://recursos.educarex.es/pdf/linguaex/informacion_linguaex.pdf consultado em 23/03/2013.
- FAGIM RODRIGUES, Valentim e José Ramom PICHEL CAMPOS (2012), *O galego é uma oportunidade / El gallego es una oportunidad*, Corunha, Agal.
- GABINETE DE ESTATÍSTICA E PLANEAMENTO DA EDUCAÇÃO (2010), *A internacionalização da língua portuguesa. Para uma política articulada de promoção e difusão*,

- Lisboa, Ministério da Educação, disponível em http://www.gepe.min-edu.pt/np4/?newsId=364&fileName=Internacionalizacao_LP.pdf consultado em 26/03/2013.
- GALICIA HOXE (2009), “Feijóo comprométese estudar introdución do portugués no ensino”, disponível em <http://www.galiciahoxe.com/ultima-hora/gh/fejoo-com-prometese-estudar-introducion-do-portugues-no-ensino-demanda-lle-expuxo-embaxador-luso-espana/idNoticia-433210/> consultado em 02/06/2009.
- IRIARTE SANROMÁN, Álvaro (2010), “Uma língua não se faz só substituindo palavras”, conferência apresentada nas *III Jornadas de Língua*; 4 de Fevereiro de 2010, Ourense, Universidade de Vigo – Campus de Ourense, disponível em <http://tv.uvigo.es/video/23573> consultado em 26/03/2013.
- JORNAL DAS LETRAS (2005), “Aprender Português em terras de Espanha”, Suplemento do JL número 85, 13-26 de abril de 2005.
- RODRÍGUEZ, José Luis (2008), “A Galiza, umha lusofonia de fronteira”, em *Actas do VIII Congresso Internacional da Associação Internacional de Lusitanistas*, Santiago de Compostela, Universidade de Santiago de Compostela.

HOW CAN SOCIOLINGUISTIC DATA BE USED?*

Celeste Rodrigues

celesterodrigues@campus.ul.pt**

Deolinda Simões

Deolinda.Reis.Simoes@at.gov.pt***

Our paper addresses the ethical and legal issues related to the display of the speech data included in sociolinguistic data bases, considering in detail the case of CPE-Var, a corpus of European Portuguese collected by the first author. Privacy, consent, data integrity, anonymisation, authorship, copyright and related topics are addressed. We provide general background of sociolinguistic data bases, discussion of the possible uses of the data, discussion of the consequences of misuse of the data, besides the relevant support to appropriate scientific use of CPE-Var data. Most ethical and legal questions arise when data are to be used outside the main scope of the initial research project. Even though the legal framework covers the first objective of the research, it is questionable if some other research proposals are also covered by the consent obtained. Our conclusion is that some further uses of CPE-Var data are legitimate, but others must be cautiously avoided or discarded.

Key words: ethic and legal rights, sociolinguistic databases, privacy, informed consent, authorship, copyright

O presente artigo trata os problemas éticos e legais relacionados com a divulgação de dados de fala incluídos em bases de dados sociolinguísticos, tomando para o efeito o caso do CPE-Var, um corpus recolhido pelo primeiro autor. Os tópicos discutidos incluem: privacidade, consentimento, integridade dos dados, anonimização, autoria e copyright, entre outros. É fornecido um background das recolhas de dados sociolinguísticos, uma discussão dos usos possíveis desses dados e das

* We thank M. E. Cabrita for a previous linguistic review of the paper.

** University of Lisbon, Arts Faculty (DLGR), Centro de Linguística da Universidade de Lisboa (CLUL), Lisbon, Portugal, under «PEst-OE/LIN/UI0214/2011» Project (Funded by National Fundings delivered by Fundação para a Ciência e a Tecnologia (FCT)).

*** Law and Forensic Sciences expert, Senior Officer of Customs and Tax Authority of the Finance Ministry, Lisbon, Portugal.

consequências dos possíveis usos indevidos, para além de argumentação de suporte dos usos científicos dos dados deste corpus. A maior parte das questões éticas surgem quando se pretende usar os dados fora do âmbito inicial da pesquisa. Embora os fins iniciais da pesquisa estejam legalmente cobertos, é discutível se outros usos serão legítimos. A nossa conclusão é a de que alguns novos usos podem legitimamente ser dados aos materiais, mas outros devem ser avaliados com muito cuidado ou rejeitados.

Palavras-chave: direitos éticos e legais, bases de dados sociolinguísticos, privacidade, consentimento informado, autoria, copyright.

*

1. Introduction

Among ethical problems involved in sociolinguistic research, there are: privacy, anonymisation, copyright and informed consent.

Privacy is one of the major ethical issues discussed here, since private information is mentioned in several interviews. The high spontaneity level of some interviews creates the illusion of confidentiality, allowing private matters to be addressed. Sociolinguistic interviews often share the following characteristics: they are intended to provide useful manageable linguistic data in a format that allows linguistic description, statistical treatment of the data and both social and linguistic profile creation. The researcher tries to get all the information he can to better understand the data afterwards, although he may not show an explicit interest in those particular details. Data, after transcription, are usually included in data bases. Data consist of sound recordings of single individual interviews in pre-established appointments in an environment familiar to the interviewee. CPE-Var interviews have formal discourse, reading tasks and semi-informal discourse. Since the interviews had to be orthographically transcribed by students, privacy of the information concerning the speakers was challenged. In addition, some data is now available in a database, which was created to treat quantitatively the phonetic variants of interviews. This database has personal data of the speakers, although anonymisation requirements are generally obeyed.

Apart from that, nowadays, voices are recognized with high probability in very short samples, if they are submitted to careful acoustic scrutiny. These interviews vary from 60 to 75m long. As a consequence, is it ethically fair to assemble even small parts of the interviews and to diffuse them in public speech databases? Can we use data to study the acoustic charac-

teristics of voice? Is it possible only to reproduce some samples in scientific environments? Which are the limits? Which are the laws that we must observe in this environment?

Speakers own their voices. The researcher who collected those voices is the owner of the interviews. Does the researcher have the right to freely disclose all the content of the interviews? Copyright matters of the researcher seem to interfere with speaker's right to their own voices.

Another ethical question arises when we think about consent. Nowadays, consent is always asked, but not in the 90's. CPE-Var doesn't have a written consent, even though the situation in the recordings was only possible with the voluntary consent of the speaker. What can a lawyer present for/against the use of this kind of data in public databases and in other domains?

Furthermore, how can we assure that the integrity of the data is kept if we allow that data are used by other people? Can we change the data in any way? Can we, for instance, copy silence intervals to the parts of the interaction where speaker proper name appears? Or is it considered manipulation? Which are the rights of the author, as the author of the recordings and of the database where transcriptions appear? Several of these issues are developed below.

Section 2 provides some general sociolinguistic background. Section 3 summarises methodological issues concerning data collection and interview format of CPE-Var. Section 4 mentions major limitations (both ethical and legal) to the use of data that emerged during the research. Section 5 further expands on the solutions implemented to avoid undesired ethical results. Section 6 focuses on the legal background of ethical issues. Finally, in Section 7, we present some concluding remarks and refer to issues that must be taken into consideration in sociolinguistic research.

2. Sociolinguistic background

Among Social Sciences, Sociolinguistics is a sub-field of Linguistics concerned with the correlation of society and language use. Sociolinguistics then combines technical knowledge from Linguistics and several other fields: Sociology, Psychology and behavioural sciences in general, etc. Following this reasoning, sociolinguistic work supposes a wide knowledge of the social tissue and of the behaviour of speakers according to different settings of language use, from the start.

Familiarity with Portugal's social reality and its culture as regards language use, on the one hand, and the linguistic knowledge which rendered such work possible, on the other hand, have led Rodrigues to build CPE-Var corpus (Corpus de Português Europeu – Variação), since such a collection of adequately stratified linguistic data was not available for European Portuguese (EP) in the 1990s. This collection of linguistic data includes native speakers from both major EP dialects and, consequently, may exemplify the behaviour of speakers of two dialects, one closer than the other to standard EP.

EP in mainland Portugal though is acknowledged to have small linguistic variation since speakers from all parts of the country understand one another. Several variation features, namely in the phonetic shape of words and in the prosodic curves that characterise single dialectal varieties can, however, be found. Variation in syntax and in the lexicon is also found, although it is generally considered to be smaller. Several variation phenomena are associated to sociocultural differentiation.

There are two main dialectal and geographical areas in the Portugal's mainland: center and south area is mainly flat and the north is more mountainous. Rodrigues chose two targets: Lisbon area, within the geographical area where linguists have already acknowledged to be found a language variety that generally avoids several linguistic features, known to identify non-standard varieties of EP (in the center and south area), and Braga, which belongs to the northern area where the language has several different (and sometimes stigmatized) linguistic features. The former is often considered to be more innovative and the latter more conservative, as a consequence of language formation and history (Castro 2006, Cintra 1971).

Lisbon, as the capital city, has a considerable part of the country's population (considering the territorial area covered) with sociocultural elite, nearly all governmental bodies, the most important media and road infrastructures. It is also the political and financial center of the whole country. As a consequence, educated speakers from Lisbon or from Coimbra (in center area) tend to use unmarked linguistic features, not only in writing, but also in spoken language (a certain kind of linguistic standard). Like any other dialect, Lisbon dialect has sociolinguistic variation. Only the linguistic features present in the spoken language of the educated speakers tend to spread all over the territory. Some of its features are now found in geographical areas where they were not expected (for instance, the centralization of stressed pre-palatal /e/ is nowadays found in the Alentejo). For all these reasons, the importance of the linguistic features of the dialect spoken in Lisbon is naturally high as such features may spread, even if the speakers

are unaware of this diffusion process. Adoption of standard language features by speakers of other dialects (standardisation) is particularly visible when it is quantitatively relevant. This often happens when these speakers interact with someone from outside their region (be it from the standard dialect or not), as it is the case of the CPE-Var.

Lisbon dialect is different from non-urban areas within the dialect of the center and south. Unlike the rest of the geographical area, it shows a tendency to close initial unstressed front vowels and a tendency towards centralisation in general (either in stressed or in unstressed position). These processes lead to neutralisation of several phonological vowels in certain highly productive contexts. Moreover, the Lisbon dialect has a high level of vowel weakening, culminating very often in vowel deletion. As far as the consonantal system is concerned, the Lisbon area shows an innovative variant related to sonorant /r/. The phonological system also has two contextually motivated variants, a tap in Coda position (*carta*, *letter*) and in onset intervocalic position (*cara*, *expensive*) and a trill in onset (*rapaz*, *boy*, *carro*, *car*) after a coda segment (*melro*, *blackbird*) (see Mateus and Andrade 2000). The trill is realised most commonly in Lisbon as [R], unlike the alveolar pronunciation it shows in some speakers of other dialects (the segments are in free variation). [R] is considered to be an innovative feature of Lisbon speech that is now spreading to the other varieties of the language (it may be noted incidentally that this feature keeps its anterior nature in the nearby regions of the south).

The main goal of Rodrigues's PhD doctoral dissertation was to determine how the features of the standard dialect spread in the dialect used in Braga, as an exponent example of this northerner speech, which means that she had to collect a corpus of Braga speech comparable to the one collected in Lisbon.

Braga is an ancient town, highly rejuvenated by industry and the establishment of university facilities in the past three decades, among other factors. It represents here the north region where Portuguese developed and where some ancient linguistic features are still present today. Salient features of Braga speech include the absence of nasal vowel closing and the pronunciation of a bilabial stop or a fricative for the phonological fricative /v/ (*luvas*, *gloves*) and a small degree of vowel deletion phenomena. Nasal codas are still present in Braga dialect, unlike in the other dialects. Braga also exhibits a typical prosody, quite distinct from the Lisbon speech (Vigário e Frota (2003), among others). Variation can be found among speakers, according to the degree of attention to speech, the kind of interaction established and the gender and the sociocultural profile of the speakers.

3. Description of methodological issues concerning CPE-Var

CPE-Var interviews follow the general lines of sociolinguistic interviews, described by Labov (1981), as will be shown below. The investigation was initially meant to identify the phonological and phonetic features that were subject to variation in the cities of Lisbon and Braga. It involved, then, a sample of interviews of different speakers characterised by specific socio-cultural profiles. Speakers were classified in four education levels and in five age groups. Even with all these differentiation factors, since no income information was collected, the sociocultural profiles were only approached. Several linguistic indices/markers and stereotypes of speaker voice were identified (see Rodrigues, 2003).

Recording quality had to be high to pursue such objectives: very small properties of the acoustic signal. Recordings had therefore to be done in a controlled environment with the speakers' previous agreement. These characteristics rend it possible for us to use the interviews for several other purposes nowadays. The interviews are linguistically rich enough to allow the study of several linguistic topics from distinct perspectives: phonetic, phonological, morphological, syntactic, reading and spontaneous discourse analysis, etc. That is why we have to deal with ethical issues in this new phase of data use.⁽¹⁾

3.1. Data collection

CPE-Var includes 180 single sociolinguistic interviews among speakers of Lisbon and Braga collected from 1996 to the end of 1998. These interviews often include personal information raising a number of ethical and legal problems to the researcher. Sensitive topics present in some interviews and the use of voice materials for identification of acoustic parameters relevant to the identification of speaker's voices raise many ethical concerns: namely information privacy and confidentiality, transcription and recording anonymisation, author's rights, among the most important. Most of the problems may be expected from the beginning, others come around only in the course of the investigation or afterwards.

(1) For example, some orthographic transcriptions of CPE-Var data were useful for papers on EP syntax properties (DUARTE, Inês, Maria João FREITAS, Anabela GONÇALVES, Matilde MIGUEL and Celeste RODRIGUES (2002)) and a part of CPE-Var is currently being used to identify robust acoustic parameters of speaker's voices.

Sociolinguistic interviews often share the following characteristics: they are intended to provide useful manageable linguistic data in a format that allows linguistic description, statistical treatment of the data and both social and linguistic profile creation. Researchers try to get all the information they can to better understand the data afterwards, although they may not show an explicit interest in those particular details. Data, after transcription, are usually included in databases. Data consist of sound recordings of single individual interviews conducted in pre-established appointments in an environment familiar to the interviewee. CPE-Var interviews have formal discourse, reading tasks and semi-informal discourse.

Informal speech samples sometimes include personal information. It would be un-ethical to divulge this information (for instance, original explanations/ideas, appreciation on the moral conduct of public or otherwise recognisable figures, on the speaker's health, contents of an unexpected phone call, etc.) for it was not meant to be revealed though it was recorded. Some speakers avoid sensitive subjects, others do not. Moreover, these statements/opinions combined with personal information given at the beginning of the interaction and the quality of the voice recorded clearly identify the speaker, even though the researcher omits the speaker's name in the transcription or the database. Nowadays, accurate voice recognition occurs in very short samples if they are submitted to careful acoustic scrutiny. These interviews vary from 60 to 75m in length. Therefore, some questions may be raised: is it ethically fair to assemble even small parts of the interviews and to divulge them in public speech databases? Is it possible to reproduce only some samples in scientific environments? Where do we set the limits? Which laws must we comply with in this environment? Must all data collections have informed consent files signed beforehand? Most of them do not have such files, especially the first ones. CPE-Var does not either, even though the situation in the recordings was only possible with the voluntary consent of the speaker. On what grounds can a lawyer argue for or against the use of this kind of data in public databases and in other domains?

3.2. Speakers

Target speakers were born either in Lisbon or Braga and have lived in their cities most of their lives. They have been selected according to their social profile: they are either male or female, from five age groups and four education levels.

Some interviewees had some previous knowledge of the interviewer's work, others did not. Speakers were asked to collaborate in a research that would lead in the first place to the interviewer's PhD dissertation. Speakers were informed of the general research goal, that is, dialectal comparison. Speaker anonymity was assured. Speakers agreed to participate in all the tasks the interview comprised, including reading of a long word-list, reading of a list of sentences and a text. They also agreed to entertain a non-oriented dialogue on topics of their choice. No consent document was signed at the time (1996-1998) since speaker's agreement was regarded as tacit and at that time such document was not required. Their consent is obvious considering their willingness to be interviewed and the fact that several refer to the recording procedure in the course of the interview.

Most speakers relaxed after the reading tasks, allowing sometimes for the linguistic interaction to go on and on, and achieving a degree of attention to discourse next to the vernacular. Other speakers, though, always kept in mind that they were speaking to someone that did not belong to their linguistic variety (as the results of Rodrigues (2003) show). Even though speakers pay different degrees of attention to the way they talk in this last part of the interview, the spontaneous speech collected is of high standard quality. It is informal enough to describe the linguistic differences of the two dialects under analysis, namely regarding both phonological and phonetic properties of the language in its current style (in the above described situation). Linguistic variation phenomena were the main focus of research.

3.3. Interview

Interviews were all made by the same interviewer, keeping as far as possible the same linguistic situation. Interviews had no observers and were made in a quiet room; most recordings were made in an environment familiar to the interviewee.

A PMD Marantz portable recorder equipped with a unidirectional microphone placed in front of the interviewee (at 20 cm distance) was used to capture the recordings. Analogical recordings are now adequately preserved in digital format.

Interviews have the following structure: at the beginning, a part of formal discourse (where speakers identified themselves and their social profile – this part of the interview is the most formal of all parts involving sponta-

neous speech), three reading tasks (550 isolated words presented in separate cards, a list of sentences also presented in separate cards, and a one-page text from the weekly press), followed by an semi-informal dialogue where speakers were asked to relax since their test phase had been accomplished. This structure was used to test if a different order of the standard Labovian procedure would capture more casual spontaneous speech at the end. Labovian sociolinguistic interview tries to capture informal speech style at the beginning and proceeds to more formal styles. We anticipated people's relationship would develop all along the interview, leading to a lower degree of formality in the final part, if the interview is long enough. This assumption was verified, since results of quantitative analysis of variation phonological phenomena show an increase in the use of informal variants in this last part of the interview (Rodrigues, 2003).

3.4. Interviewer

The interviewer belongs to a third dialect variety of EP, and thus all interviewees should feel free to use their native dialect alike. She is a native speaker of the southern dialect of continental EP, but not of the Lisbon area. She acquired EP on the west coast of the Alentejo. As she is a non-standard speaker of EP, all the interviewees should feel equally at ease to make use of their native dialect.

4. Ethical problems

Some interactions conveyed personal details of private life that from the beginning create in the researcher a sense of responsibility as regards disclosure of these aspects of the interview. We believe that not all parts of the dialogue present in the interview should be disclosed, even in an academic environment. Due to the singularity of the interview, the intimacy degree of the participants grew along the interview. Most of the interviews show speaker behaviour close to the vernacular, as expected. Note that most speakers had no previous knowledge of each other and met only once before the recording.⁽²⁾

(2) CPE-Var allows the identification of different attitudes from speakers towards the interviewer and the interview itself. Most speakers accommodate linguistically to the interview situation, some showing more careful speech than others.

There were speakers who mentioned having witnessed sexual assault committed by children, other speakers had been involved in conflicts with fire guns, others revealed aspects of their health and private life, events which the researcher deemed inappropriate to freely disclose. For this reason, one of the most important ethical problems identified is related to secrecy, in other words, the need for the researcher to keep private life information only to herself, if it was disclosed in the interviews. How can private information be kept when it is present in the recording and the recording is going to be transcribed by undergraduate students? How can the researcher assure a speaker's anonymity if speaker's name is spontaneously given during the interview? How can anonymity be preserved if acoustic study of the voice recorded can reveal a speaker's identity?

Another ethical problem relates to the use of the data for objectives different from the ones initially devised. Speakers were informed of the general aims of the research. They cannot, however, be informed of new research goals since 15 years have elapsed, their phone numbers have changed and they cannot be reached anymore. This problem is sensitive since tacit consent was obtained and the law now states the need for signed consent. If written consent had been obtained, the rights of the researcher would have been assured. Does tacit consent have the same value as written consent? We will discuss how Portuguese law rules over this subject matter.

A third ethical problem is related to the right of speakers to their voice in legal recordings, such as the ones we have in CPE-Var. What can one do with the informant's voice? Can we manipulate the interview, for instance, to erase parts of the speech deemed to be intimate or parts where the name of the speaker is given?

A fourth problem concerns the rights of the researcher to the original research. In this particular case, the researcher created her own model of interview (although based in Labovian interviews), she collected the whole set of recordings, she transcribed most of the interviews and she is also the author of the database where samples of the data are included. Are there any limits to her authorship and copyright due to the kind of data involved (namely, voice recordings)? All the tasks named above are time consuming and a lot of effort was made to bring them about and to keep them completely private.

Does the researcher have the right to combine linguistic information and speakers' identity or can she only deal overtly with linguistic information and the speaker's profiles? Can the researcher use the data for academic

purposes or only for the PhD for which it was initially designed? How long are speakers entitled to their privacy? When can the data be freely used?

These are the major ethical topics which will be discussed presently.

The use of CPE-Var in the course of the PhD dissertation raised no ethical problems to the researcher because all parts of the transcriptions presented avoided private or sensitive topics. Moreover, the academic environment is the one initially anticipated. In spite of that, a first problem concerning the use of the CPE-Var interviews came up when transcriptions of the whole interviews were to be made. Complete orthographic transcription was supposed to be made by undergraduates. This is when the researcher started to worry about privacy issues. How could she continue to assure privacy/confidentiality of the interactions and speakers' anonymity? Furthermore, were there in the interviews any hints about the speaker's consent, so that it would be undisputable that they could legally be used?

4.1. Ethical problems in detail

Let us start with the issue of consent since most ethical questions raised above relate to it.

Written consent is always required nowadays although it was regarded as unnecessary in the 1990s. CPE-Var does not have signed informed consent. Sometimes oral consent was given during the interview, sometimes it could be inferred from the interview itself. Tacit consent was always obtained, however, since all the speakers agreed to willingly participate. Furthermore, they read aloud a large amount of reading materials to be clearly captured by a visible tape recorder placed in front of them. Their performance would be impossible had they been coerced. The interviews involved only previously contacted speakers who showed no a priori problems about speaking with the interviewer. No payment or any kind of feedback was promised since the research itself was not funded.

CPE-Var recorded interviews are then admissible data for language study and, furthermore, the voices captured are useful for several purposes. We believe that we have obtained tacit consent since the interviews were previously scheduled and recorded with a table microphone and a portable PMD tape recorder in front of the speaker. The law excludes evidence in court only if information is obtained through private life intrusion, through home, mail or communication violation without the consent of the holder. This kind of evidence would then be considered null and void. None of the

above-mentioned situations occurred during CPE-Var recordings. These recordings would, therefore, be legitimate even in a court of law.

The interviews frequently include the speakers' name, as we have mentioned. Furthermore, speakers may often be identified by their knowledge of the topics mentioned, opinions given, names dropped during the talk, etc. As a consequence, privacy / anonymity / confidentiality issues were at stake, if undergraduate students were to have access to the interviews. Students asked to perform the orthographic transcription task had been previously coached and asked not to disclose any information present in the recordings without the researcher's consent. No problem arose in this transcription phase or ever since, even though we have no guarantee that undergraduate students kept their word. We are fully aware that that kind of speech obtained during the interviews was only possible in those particular circumstances: two people alone, face-to-face, in a dialogue setting that creates the illusion of intimacy. This illusionary sense of intimacy was considered fundamental in order to get the sort of data the research set out to analyse. Our target was semi-informal speech data which are only available when the informants trust the interviewer and feel at ease.

Phonetic transcriptions were made by the researcher herself; therefore, no problems occurred during that phase either. For that matter until conclusion of the researcher's PhD no problems came up in the use of the data of CPE-Var.

A database was built to keep all the relevant data used in Rodrigues's PhD. Data include samples of phonetically transcribed occurrences of words in context, classified linguistically. The classification includes the speakers' sociolinguistic profile. The database is available exclusively to the holder; it is considered private and Rodrigues holds all legal rights to it, namely, reproduction, display, manipulation, etc. See below what the Portuguese law states about copyright.

Some ethical issues came up later, when voice of some informants was to be used to analyse their acoustic properties in order to establish criteria / parameters for voice identification. Would it be morally reasonable to use the data for this specific purpose? The main area of research had been mentioned, but the specific topic addressed was not. We could not use a methodology that would constrain the informant's speech in order to collect good quality informal speech data. In this sense, our means were justified by the objectives of data collection. Concessions were made due to the need to collect spontaneous speech data of appropriate quality. Even if it is reasonable, does the law in any way prevent this kind of research with

this kind of data? How can speakers' anonymity be assured if their voice is under acoustic scrutiny? Forensic Phonetics has developed hugely in recent years and, if careful treatment of appropriate voice data is made, it can provide reasonable evidence of the speaker's identity beyond the shadow of a doubt. We participate in a research project that aims at evaluating the relative importance of acoustic parameters towards speaker identification focusing in EP dialectal data. Can we use CPE-Var data for this research? It is worth noting that the recordings would not exist had Rodrigues not made them - they are therefore a product of her technical work and she may argue that she is the owner of this 'artifact'. More than 15 years have elapsed since data were collected: is this enough to allow for disclosure of the data for different research uses?

Since consent for disclosure of the interviews contents was inferred, all parts of the CPE-Var data that are not deemed private nor can they do the speaker any harm can be used. These data are disclosed by means of the orthographic/phonetic transcriptions but not through the all sound recording of the voices, even though some parts of the recording can be disclosed to exemplify several research issues. This means that crimes, intimate affairs, health condition, etc. mentioned by some speakers are not disclosed. What about the confidentiality of new ideas, theories, phone numbers which are revealed, etc? They were not disclosed in the first attempts to show the data. Meanwhile, after 15 years we do not deem that secrecy of these matters is needed. Time has rendered such secrecy useless (for instance, telephone numbers have changed and are therefore no longer accessible).

Data were only disclosed for scientific purposes in small samples (proportionality) enough to verify the hypotheses. Transcribers were asked not to disclose any of the information present in the interviews lest they face prosecution.

What about anonymisation?

Anonymisation of transcriptions was mainly assured by means of coding procedure. For instance, speaker Maria Melo (fictitious name) was codified as informant number 32 with the profile LF32, meaning that it is a Lisbon female graduate speaker in the 2nd age group. Speakers were thus retrieved straightforwardly. What about the recordings? Although the speaker's name does not appear clearly in all recordings, in some it does either at the beginning or somewhere else.

When the interviewee's first name appears, it can be replaced by another name in the transcriptions (normally, a similar one, for prosodic reasons). What should be done within the recording itself? Should we replace the

name by a silence interval? This changes the recording integrity and may even be termed manipulation. What else can we do? The solution we have adopted is simply, for the purposes of the research project concerning the acoustic analysis of voices from the CPE-Var, not to use those parts of the interview which include the speakers' identification or other identifiable people.

Do we have the legal right to use CPE-Var data for other research uses that could not have been anticipated from the beginning? Are there any limits to using it?

To assure anonymisation only the number of the informant, if possible, is given. If that is not possible, then the substitute (the name by which the proper name was replaced) is used. Otherwise the code attributed to the cell of the speaker is mentioned.

Using this procedure guarantees that our speakers cannot suffer any damage for having placed their trust in the interviewer during the interview. A speaker who might potentially be charged with committing a crime cannot be charged since his anonymity was preserved. This procedure also renders it impossible for any absent person referred to in the interview to bring charges against us since we do not disclose his/her name.

Even though we did not promise speakers any feedback of our collection of data, they may receive feedback, if they so wish. Informants may have access to the publications authored by Rodrigues, including books, scientific papers, presentations to symposia, etc. They can also reach Rodrigues at the University.

5. Ethical and legal rights

Legal limitations to the use of acoustic signal data recorded and their transcription outside the context where they were collected, even if exclusively in academic environment and for research purposes, constitutes a very important and open issue from the Ethics and Law viewpoints.

We will now address the above-mentioned open questions, particularly those that concern the speaker's personal data and such data legal protection – i.e. the human being and his own data –, namely, the right to the reservation of the intimacy of private and family life under Article 26 Nr. 1 of the *Constituição da República Portuguesa*, the Portuguese Republic Constitution.

5.1. Privacy

It is worth noting that the above-mentioned rights have been deemed so important in modern societies that they were afforded constitutional dignity. In addition, according to publications of Professors Gomes Canotilho and Vital Moreira (2007), the normative scope of the fundamental right to the reservation of the intimacy of the private and familiar life must keep in mind the following aspects: respect of each other's behaviour, respect of anonymity and the respect of rules of life in society. These fall under personality rights, the violation of which is punished by law.

Canotilho and Moreira (*idem* p. 182) argue that "the right to the secrecy of human being" (image rights, the right to speak, right to private life) should be intrinsically linked to personality rights: "the constitutional criterion should perhaps start from the concepts of privacy - (Article 26 Nr. 1) and human dignity - (Article 26 Nr. 2 of the Portuguese Constitution) so as to define a concept of privacy of each person, culturally appropriate to contemporary life."

In line with this, Andrade (2004, p. 498) also contends that the latest revision of constitutional law: "raises to the constellation of Rights, Freedoms and Guarantees in criminal matters (Article 26, Nr. 1) the right of every man – and him only, to decide who can record his voice".

He further emphasised: "the full availability of the human person on the spoken word as a direct expression of his/her own personality and dignity", which was enshrined in the Constitution.

Andrade also considers the right to speak from a double dimension perspective:

- a) A positive dimension – meaning the legitimacy to authorise the recording and the audition, freely, with no restrictions;
- b) A negative or exclusive dimension – meaning the freedom to refuse the recording and the audition, with no restrictions.

Andrade argues that the law protects the right to speak as personal goods, as a direct expression of the speaker's personality in communication with other members of society.

In this respect it is worth noting that the fundamental provision of said Article 26 Nr. 1 also grants protection under legal and constitutional law to the image and to the voice, more precisely to the word, by listing these rights, together with the above, such as the right to good name, to personal identity, to the personality development and to citizenship. This means that

the word is protected by the supreme law of the nation by making sure that particular caution must be exercised when capturing voice and image to avoid conflicting with other rights, including the right to privacy.

5.2. Database and Copyright

The Portuguese Code of Copyright and Related Rights of 1985 (Decree-Law Nr. 63/85, March, 14th, amended by Law Nr. 65/2012, December, 20th) aims at protecting authors of literary, scientific and artistic works. In its Article 1, Nr. 1, it defines such works as: “intellectual creations of the literary domain, scientific and artistic in any way externalised.”

This provision protects works from any misuse, as well as from any economic benefits arising out of such misuse or exploitation. The Copyright Code protects patrimonial rights, but also the rights of a personal nature, called moral rights. Influenced by European Union law, the copyright term was extended and works are now protected for a period of 70 years (when before the term was 50 years) to strengthen the preservation of historical and cultural heritage.

It should be noted that along with the industrial property law, protection of the literary, scientific and artistic works as well as the rights of creators, both from the economic and moral standpoints falls under copyright law, one of the areas of intellectual property.

Regarding the contents of the data on CPE-Var, described above, the legal framework should comprehend all the data recorded as forming a body, which must be understood to mean a *database*, and to that extent its legal protection is conferred. Therefore, in our research there are two databases: one is the collection of recorded interviews and the other the database created to process quantitatively the phonetic transcription of word occurrences in the CPE-Var recordings.

Rodrigues, in her capacity of CPE-Var author, is entitled to act in the management, protection and defense of her work, and may authorise its use by third parties.

Nowadays databases enjoy broad legal protection, both under the Portuguese and EU law. Thus, Decree-Law Nr. 122/2000, July, 4th, which transposes into national law EU Directive Nr. 96/9/EC of the European Parliament and the Council, March, 11th, sets out the legal protection regime of databases. The solution now adopted, by overriding EU law, provides double protection: for one, the databases that constitute intellectual creations

are protected by copyright with some special features; and then it provides protection of the investment of the manufacturer of certain databases.

The Portuguese Law then sets out what shall be and is meant by database by defining it in its Article 1, Nr. 2 (subject) as: “a collection of works, data or other materials, arranged in a systematic or methodical way and individually accessible by electronic means or other.”

It also establishes that databases are protected by copyright and that such protection is subject to country of origin, considering the author who is qualified as such by the law of the database country of origin.

The European Community, given the importance of fundamental freedoms referred to in the Charter of Fundamental Rights of the European Union, strengthens the protection of fundamental rights, setting out (under Article 7) respect for private and family life: “Everyone has the right to be respected for his private and family life, his home and communications.” and (under its Article 8, Nr. 1) entitled protection of personal data it expressly lays down: “Everyone has the right to the protection of personal data”, reinforcing and conferring legal dignity to this subject matter.

Portuguese Law assures data integrity and preservation of personal dignity, which is one of the most important legal principles of the Portuguese Constitution. It also assures the right use of the data and, for that reason, good-faith is imperative in all manipulation of the data and in the general use of the data exclusively in academic environment to protect its quality, the author and the speakers’ anonymity.

5.3. Personal data and consent

In addition, in this context the “Law on Personal Data Protection,” Law Nr. 67/98, October, 26th, transposes into the Portuguese legal system Directive 95/46/CE of European Parliament and the Council, October, 24th 1985 on the protection of individuals with regard to the processing of personal data and such data free movement.

Law Nr. 67/98 first sets out the general principle that the processing of personal data must be made in a transparent manner and with respect for private life as well as rights, freedoms and guarantees. In our view, this principle applies to the contents of the interviews and to the personal data reported by respondents during the interview, which were duly protected and complied with such legal requirements.

The Personal Data Protection Law defines, among others, the meaning of the following expressions: personal data, personal data processing and data consent, due to the particular importance of this latter term. Accordingly, the law sets out clearly (Law Nr. 67/98, Article 3 h)) what is meant by consent of the owner of the data: “any expression of intent, free, specific and informed, under which the holder accepts that personal data be processed.”

The interviewee who provides access to the interview and simultaneously authorises its recording is thereby granting his free and spontaneous consent. As argued by Andrade (1992, p. 251) consented recording or its use is excluded from the typicality of illegal events for the following reason: if consent given by the author is valid, it can never be alleged to be null or void on grounds of intrusion of privacy, since consent was given by the holder (a right freely available) and as such the interview has the holder’s effective intervention.

The National Commission for Data Protection (NCDP) was created in Portugal to examine compliance with the legislation. The NCDP is an independent administrative body with power to control and monitor personal data processing, with strict respect for human rights and fundamental freedoms and guarantees set out in the Portuguese Constitution and Law. The NCDP is the National Authority for Control of Personal Data that liaises with the data protection supervisory authorities in other countries.

6. Concluding Remarks

Our paper described in detail the CPE-Var collection of sociolinguistic data. During the processing of the data, several problems arose which posed some ethical dilemmas. In the above sections we have discussed some of those problems from a legal point of view, namely privacy, confidentiality, anonymity, use of data outside the scope of the former investigation for which they had been collected, copyright issues, personal right to voice and image recording, consent, etc.

As regards privacy/confidentiality, it should be noted that CPE-Var speakers have never expressed their wish to privacy. All the recordings were obtained after a brief appointment where speakers were informed of the general research goal and where they gave their consent verbally to the recording. In the 1990s this was deemed adequate procedure to collect legal speech data in Portugal. No image was captured. Since the recorder and a clearly visible microphone were placed before the speakers, all the recordings had the interviewees’ spontaneous collaboration. After the reading

tasks, interviewees spoke freely and for as long as they wanted. Their consent to the recording is obvious.

Orthographic transcriptions of the spoken speech were carried out by undergraduate students under the supervision of the corpus author. Undergraduates were asked not to disclose the names and information present in the interviews. Phonetic transcriptions were carried out by Rodrigues only and therefore no problem arose during that task.

Rodrigues is deemed to be the author of the collection of recordings and she has therefore exclusive power to explore and give access to the data, once data integrity and appropriate use are assured. Recorded voice is considered property of the author and is used exclusively for research purposes. Exemplification is restricted to small excerpts proportional to the needs.

A code was attributed to each speaker in order to keep anonymity in the database created to process the phonetic transcriptions. Speakers have been ascribed an alphanumerical code that combines information about their provenience, age, gender and instruction level.

Rodrigues uses CPE-Var to explore voice quality. For that purpose, she only uses the parts of the talk that do not contain personal data or any other private information.

All the results of scientific exploitation of CPE-Var data are currently available in several publications, providing the social feedback desired in all scientific research.

We believe that speakers collaborated with the researcher in good faith and as a consequence, Rodrigues feels she is under obligation to respect them by making sure that data will only be used in their best interest (that is, by producing no damage to their lives).

Portuguese Law follows EU law in the protection of all fundamental rights of speakers and of the author of the various products of this sociolinguistic research. In Social and Human Sciences most of the above mentioned ethical issues have proved to be particularly sensitive within Medical Sciences. The importance of good care of the human body and the right to receive best treatment are major concerns in this area.

Above all, it is fundamental for each researcher to establish a clear line between what is ethically appropriate/irreproachable behaviour and behaviour which is ethically reproachable. If this line is never crossed and if the researcher acts according to the universal principle of good faith s/he will be able to process ethically all sensitive items presented by the data. The researcher has to reach a compromise between the best interest of the research and the best interest of the speakers in a sociolinguistic research.

Does our methodology pass the ethical tests normally used, say, in Medical Sciences?

- 3) The impartiality test – Would we want someone to apply to us the methodology we have adopted?
- 4) The universalisability test - Would we want our methodology to be applied to any other similar cases?
- 5) The interpersonal justifiability test – Do we have good reasons to justify our methodology options?

Our methodology passes the impartiality test. We would not mind if equal methodology was applied to us, since the interview was previously scheduled and lasted only one hour. Our recordings assume, from the beginning, that a semi-informal dialogue involving an interviewer from a dialect different from the informant's would occur. As it occurs only once, in principle, no damage results to the informant, apart from possible waste of time. Informants could refuse to participate in the study, either at the beginning or during the interview (free will principle). Furthermore, the research's main goal is socially recognisable as good, that is, the acquisition of more scientific knowledge of EP language daily use.

Our methodology also passes the universalisability test. It can be used by other researchers, since it is described in depth in published papers and no damage results from its application to the target speakers (on the contrary, its application enriches society, as more scientific knowledge can be obtained).

It also passes the interpersonal justifiability test since our reasons for choosing this methodology are explainable and easily understandable.

Portuguese literature on ethical issues in sociolinguistic research is scarce. We found no works published on the subject, besides the ones on bioethics, economic and human resources, law or computational engineering, media and philosophical theories were identified. Problems may arise if there is a conflict between two methodologies achieving the same result, one more beneficial than the other. Our methodology is not intrusive (it produces no harm to the speakers) therefore we believe that it is minimally invasive of personal life.

References

- AMARAKONE, Keith and Sukhmeet PANESAR (2006): *Ethics and the Human Sciences*, Elsevier Limited, printed in Italy, p. 211.
- ANDRADE, Manuel da Costa (1992): *Sobre as Proibições de Prova em Processo Penal*, Coimbra, Coimbra Editora, p. 121 and ff.
- ANDRADE, Manuel da Costa (2004): *Consentimento e Acordo em Direito Penal*, Coimbra, Coimbra Editora, ISBN 972-32-0438-X, p. 498 and ff.
- BRITZ, J. J. (un-dated), "Technology as a Threat to Privacy: Ethical Challenges to the Information Profession", <http://web.simmons.edu~chennitNIT-279696-025-Britz.html>.
- CANOTILHO, Gomes and Vital MOREIRA, (2007): *Constituição da República Portuguesa Anotada*, ISBN 9789723214628, p. 182 and ff.
- CASTRO, Ivo (2006) *Introdução à História do Português*, Edições Colibri, Lisboa.
- CINTRA, Luís Filipe Lindley (1971) "Nova proposta de classificação dos dialectos galego-portugueses," *Boletim de Filologia*, XXII, 81-116 (repr. in CINTRA, Luís Filipe Lindley (1983), *Estudos de Filologia Portuguesa*, Lisboa).
- Código do Direito de Autor e dos Direitos Conexos*, 1985 (Decree-Law Nr. 63/85, March, 14th amended by Law Nr. 65/2012, December, 20th).
- Constituição da República Portuguesa*, April 2nd, 1976 – 7th, Constitutional Revision Law Nr. 1/2005, August, 12th.
- CARTA DOS DIREITOS FUNDAMENTAIS DA UNIÃO EUROPEIA – JORNAL OFICIAL (2010/C 83/02), March, 30th, 2010.
- DATTALO, Patrick (2010): "Ethical Dilemmas in Sampling," *Journal of Social Work Values and Ethics*, Volume 7, Number 1, Copyright 2010, White Hat Communications.
- Decree-Law Nr 334/97, November, 27th* [transposing into Portuguese Law *Directive Nr. 93/98/CEE, of the EU Council, October, 29th* - changing the term of copyright and related rights protection, *Código do Direito de Autor e dos Direitos Conexos*].
- Decree-Law Nr 122/2000, July, 4th* [transposing into Portuguese Law *Directive Nr. 96/9/CE, of the European Parliament and of the European Council, March, 11th 1996*, relative to electronic/digital databases protection].
- DUARTE, Inês, Maria João FREITAS, Anabela GONÇALVES, Matilde MIGUEL and Celeste RODRIGUES (2002): "Geometria de traços e distribuição de pronomes sujeito em PE e em PB", III *Workshop do projecto Português Europeu e Português Brasileiro: Unidade e Diversidade na Passagem do Milénio*, Faculdade de Letras de Lisboa, September, 23-25th.
- ISERSON, Kenneth, Arthur B. SANDERS and Deborah MATHIEU (eds.) (2001): *Ethics in Emergency Medicine*, Galen Press, Ltd., 2nd ed.

- LABOV, William (1981) "Field Methods of the Project on Linguistic Change and Variation", in *Sociolinguistic Working Papers*, Nr. 81, p. 28-53.
- MATEUS, Maria Helena and Ernesto d'ANDRADE (2000): *The Phonology of Portuguese*, Linguistics, OUP, Oxford.
- Protecção de Dados Pessoais* – Law Nr. 67/98, October, 26th [transposing *Directive of EU Nr. 95/46/CE, of the European Parliament and European Council, 1995, October, 24th*, into Portuguese Law relative to the protection of individuals concerning personal data processing and their free circulation].
- Protecção Jurídica das Bases de Dados* – Decree-Law Nr. 122/2000, July, 4th [transposing into Portuguese Law *Directive Nr. 96/9/CE, of European Parliament and of European Council, March, 11th*].
- RODRIGUES, Celeste (2003): *Lisboa e Braga: Fonologia e Variação*, FCT-FCG, Lisboa.
- SCHRAMM, Fermin Roland (2004): "A moralidade da prática de pesquisa nas Ciências Sociais: aspectos epistemológicos e bioéticos", *Ciência e Saúde Coletiva*, 9 (3), p. 773-784.
- SIMÕES, Deolinda (2010(11)): *Medidas Legislativas para Protecção da Cadeia Alimentar no Âmbito da Importação e da Admissão*, Master Thesis FMUL (Mestrado em Medicina Legal e Ciências Forenses), Lisboa, pp. 198 (<http://hdl.handle.net/10451/2701>).
- VIGÁRIO, Marina & Sónia FROTA (2003) The intonation of Standard and Northern European Portuguese. *Journal of Portuguese Linguistics* 2-2 (Special Issue on Portuguese Phonology edited by W. L. Wetzels), 115-137.

recensões

PINTO, PAULO FEYTOR (2010), O ESSENCIAL SOBRE POLÍTICA DE LÍNGUA.

Lisboa: Imprensa Nacional – Casa da Moeda. [98 páginas] ISBN 978-972-27-1874-5.

1. Política de Língua?! Exatamente isso. Ou seja, um conjunto de tentativas (medidas e/ou atividades), umas explícitas e outras implícitas, no sentido de regular as práticas linguísticas de e numa comunidade como (respeitantes exclusivamente à comunidade linguística em que nos inserimos) as relacionadas com (i) o Acordo Ortográfico, (ii) o ensino do português no estrangeiro, (iii) o estatuto do português na União Europeia, (iv) programas, manuais e exames de português em Portugal, (v) a terminologia linguística.

2. Trata-se, obviamente, de um título bem-vindo à bibliografia portuguesa sobre a matéria, que infelizmente não é assim tanta. Para além disso, num formato (de bolso) e numa coleção (O Essencial Sobre) que são deveras convidativos para, praticamente num relance, se ficar a saber por que linhas é que se tem andado a coser a relação entre política e língua portuguesa, começando pela “Cultura linguística” (pp. 17-31), continuando pelas “Práticas linguísticas” (pp. 33-47), passando, depois, à “Política de língua” propriamente dita (pp. 49-77) e terminando nos “Eixos da política de língua do Portugal democrático” (pp. 79-92). Porém, para se entender o que aqui se diz, o A. fala, no primeiro capítulo (pp. 11-15), da constituição de um modelo de análise, ou seja: é em 1959 que se define pela primeira vez a planificação linguística como a atividade de elaboração de uma norma ortográfica, descrições gramaticais e dicionários de uma língua, para orientação de falantes e escreventes em comunidades linguisticamente diversificadas; em 1969, estabelece-se a diferença entre ‘planificação do *corpus*’ (regulação da forma das línguas) e ‘planificação do estatuto’ (regulação das suas funções na sociedade); em 1986, introduz-se o ensino de línguas, que viria a constituir a ‘planificação da aprendizagem’; e, por fim, em 2003, é a vez de se incluir a ‘planificação do prestígio’. Porque a planificação do *corpus*, da aprendizagem e do prestígio são uma questão de estatuto, é pois natural que a planificação deste último seja primordial. Registe-se, ainda, que uma política linguística pode não ser explicitada em documentos legais. Por

exemplo (não deixa de ser curioso), o português só em 2001 é que passou a ser oficialmente a língua oficial de Portugal.

3. No segundo capítulo, define-se o conceito de “cultura linguística” e exemplifica-se com o caso da cultura linguística portuguesa do final do século passado (séc. XX). Trata-se, pois, do conjunto de representações e atitudes dos portugueses perante a própria língua (língua portuguesa) e as dos outros (línguas europeias e asiáticas e as *não-línguas*). Por conseguinte, aqui pode ler-se que o português se tornou a língua materna de praticamente toda a população portuguesa, que é uma língua com tradição escrita e literária seculares (8 e 7 séculos, respetivamente), com uma gramática publicada há 500 anos e dicionário monolíngue há 200, que é falada por muitos milhões em todo o mundo e que os portugueses têm dificuldade em aceitar a intervenção de outros nas decisões acerca da língua que consideram primordialmente sua. Que as línguas equiparáveis à língua materna, por serem também línguas oficiais, com tradição literária escrita, eram, entre as línguas europeias, o francês, o inglês e o castelhano e, entre as asiáticas, o chinês mandarim, o japonês e o hindi. Que as não-línguas eram todas as variedades a que se não reconhecia estatuto de língua por não terem tradição escrita consolidada e não serem línguas nacionais oficiais. Entre estas, estavam os dialetos das línguas europeias e asiáticas e os crioulos de base lexical europeia e todos os dialetos isolados (as línguas da África subsariana).

4. No terceiro capítulo, identificam-se e caracterizam-se os constituintes das práticas linguísticas, seguido de diagnóstico das variedades linguísticas, dos falantes e das funções sociais de ambos, em Portugal, entre 1974 e 2004. Concretamente: língua materna e línguas estrangeiras, variedades internas dessas línguas e eventuais variedades de transição, ou seja, todos os dialetos, socioletos, interletos. A caracterização dos falantes tem que ver com a caracterização sociolinguística de todas as variedades. Funções sociais das variedades: administração pública e organismos oficiais, ensino, onomástica, meios de comunicação e novas tecnologias, cultura e agentes económicos. Práticas linguísticas em Portugal, de 1974 a 2004: fronteiras políticas no território/ reorganização das práticas linguísticas; e também, claro, a utilização dos sistemas de escrita. O reportório de cada falante pode incluir a língua materna, língua(s) segunda(s) e língua(s) estrangeira(s); situações de diglossia (característica mais da comunidade do que do indivíduo). Em 25 de Abril de 1974, 99 % da população portuguesa tinha o português como língua materna. Havia uma minoria que falava mirandês e outra, a L(íngua)G(estual)P(ortuguesa). Só 0,4 % eram oriundos de

outros países europeus. A partir de 1979, os africanos, sobretudo falantes de cabo-verdiano, passaram a ser a maior presença de estrangeiros entre nós. E, entre os europeus, predominavam os falantes de francês, castelhano e inglês. Para além da língua materna, em 1981, só 74 % sabia ler e escrever e falavam francês e inglês. No início da década de 1990, havia um número relevante de falantes de outras 11 línguas. No início deste século (séc. XXI), as línguas estrangeiras mais faladas pelos portugueses eram o inglês (36 %), o francês (30 %), o espanhol (10 %), o alemão (5 %) e o italiano (2 %). Apesar de ser uma língua importante no mundo e na Europa, a percentagem de falantes portugueses de inglês era a mais baixa da UE. Depois dos países nórdicos, Portugal era o país em que mais pessoas tinham aprendido línguas estrangeiras enquanto trabalhavam lá, emigradas. O inglês era a língua mais usada para ver filmes e televisão, ouvir rádio e navegar na internet. Era o país onde havia mais trilingues e tetralingues, nos países da UE. Importância relativa de outras línguas maternas faladas em Portugal: cabo-verdiano, ucraniano, russo e, depois, francês, romeno, espanhol, inglês. Também o mirandês e a LGP estavam acima das línguas asiáticas (wu e guzerate). O ensino básico público em Portugal Continental, na viragem do século, confirma um grande aumento de falantes de romeno, ucraniano, russo e búlgaro. Em 2001, a língua materna maioritária, o português, era lido e escrito por 91 % da população, e eram falados três grupos dominantes de variedades : as variedades europeias maioritárias, as variedades brasileiras faladas por quase 1 % da população residente e as variedades faladas pelos portugueses ciganos.

5. No quarto capítulo, delinea-se o quadro teórico que permite a sistematização das medidas implícitas ou explícitas com que as autoridades políticas podem tentar regular o uso de línguas. Análise de cerca de 4000 diplomas legais publicados na 1.^a série do *Diário da República* durante os 30 anos após o 25 de Abril de 1974, incluindo todos os que tinham a palavra *língua(s)*. As tentativas explícitas e implícitas de regulação das práticas linguísticas de uma comunidade (nisto consiste a “política de língua”) podem ser de nível macro (iniciativas do Estado), meso (de grupos ou organizações) e micro (individuais). A explicitação da política linguística, a planificação linguística, materializa-se em diplomas legais de carácter incitativo ou imperativo que podem ter diferentes níveis de intervenção geográfica (internacional, nacional, regional) e jurídica (lei constitucional, lei, decreto-lei, decreto, decreto regulamentar, portaria, despacho normativo). A política linguística explícita processa-se em quatro etapas: (i) preparação da planificação, (ii) formalização das decisões tomadas, (iii) planificação

do desenvolvimento da política linguística e (iv) controlo da atividade de planificação. A planificação linguística compreende a interligação de quatro vertentes, a saber: (i) “planificação do estatuto” cujas decisões podem resultar na oficialização, na nacionalização ou na proibição de uma ou mais línguas e, ainda, conduzir à revitalização de línguas mortas ou em declínio, à manutenção de línguas, à promoção da intercompreensão entre falantes de diferentes línguas ou à difusão da língua junto de quem não a tem como língua materna, especialmente no estrangeiro; (ii) “planificação do *corpus*”, que consiste na padronização da estrutura e do funcionamento de cada língua (normalização ortográfica, descrição gramatical da variedade linguística adotada como padrão, regulação da onomástica, entre outros); (iii) “planificação da aprendizagem”, que diz respeito ao leque de atividades organizadas para a aprendizagem de línguas cujo objetivo é aumentar a qualidade e a quantidade dos seus falantes; e (iv) “planificação do prestígio”, ou seja, todas as atividades conducentes à promoção das medidas contempladas nas vertentes acabadas de referir, especialmente a utilização da(s) língua(s) em contextos formais de grande prestígio e visibilidade nacional e internacional.

6. Por fim, no último capítulo, tendo presente a cultura linguística dos portugueses, as suas práticas linguísticas e a legislação aprovada, apresentam-se seis eixos fundamentais de política de língua do Portugal democrático, que são: (i) “consolidação do português, língua nacional e oficial” (desde a fundação de Portugal, no séc. XII, e com exceção do período entre 1450 e 1650, em que o português cedeu a dianteira ao castelhano, a língua de Camões foi aos poucos assumindo as funções que o latim tivera até aí nos domínios da cultura, religião, administração, comércio, ensino e justiça; instituição, em 2001, do português como língua não materna no ensino básico e reconhecimento constitucional do português como língua oficial); (ii) “restrições onomásticas ancestrais” (manutenção formal de uma política onomástica multissecular: o modelo de regulação dos nomes próprios dos portugueses ainda em vigor remonta a 1496); (iii) “gestão desigual da diversidade linguística” (apenas no âmbito da administração da justiça estava consagrado o direito à língua materna; nos demais, a política linguística portuguesa havia marginalizado de todo todas as línguas maternas minoritárias faladas em Portugal); (iv) “sucessos e insucessos do ensino de línguas estrangeiras europeias” (com a finalidade de se promover a intercompreensão, o maior sucesso da política de aprendizagem de línguas observou-se na aprendizagem de línguas estrangeiras quase exclusivamente europeias); (v) “estrutura institucional confusa, instável e fragmentada” (foram muitos os centros de decisão da política linguística, nas

primeiras décadas de democracia: 9 instâncias legislativas e 14 diferentes organismos); (vi) “influências externas: União Europeia e Lusofonia” (dois fatores externos influenciaram enormemente a política de língua nas primeiras três décadas do regime democrático em Portugal: a adesão à UE em 1986 e o facto de a língua maioritária e oficial ser também língua oficial noutros países do mundo: Brasil, Cabo Verde, Guiné-Bissau, São Tomé e Príncipe, Angola, Moçambique e, mais tarde, Timor-Leste; dimensão internacional da língua, portanto).

7. Termina o volume com um “Glossário” (pp. 93-96), de termos linguísticos ou sociolinguísticos usados ao longo da exposição, apresentando uma definição só para aqueles cujos conceitos não foram explicitados no corpo do texto – o que não deixa de ser uma boa síntese: por um lado, define-se o que não foi explicado e, por outro, remete-se para a(s) página(s) onde certos outros foram devidamente tratados.

8. Seguem-se (terminando o volume) duas páginas de referências bibliográficas (97-98), umas citadas e outras não, a que o leitor poderá recorrer, caso esteja com dúvidas e/ou queira saber mais. Destaco, do próprio autor, *Como pensamos a nossa língua e as línguas dos outros* (2001) e *Política de língua na democracia portuguesa* (2008) e, ainda, por tratarem exclusivamente do assunto em questão, Mateus (coord.) (2002), *Uma política de língua para o português*, Felipe (2005), *Promoção da língua portuguesa no mundo: hipótese de modelo estratégico* e Salomão (2007), *Línguas e culturas nas comunicações de exportação. Para uma política de línguas estrangeiras ao serviço da internacionalização da economia portuguesa*. Estas leituras podem ser complementadas com estas outras: Vários (1983), *Estão a assassinar o português!* (17 depoimentos). Lisboa: Imprensa Nacional - Casa da Moeda; Vários (2005), *A língua portuguesa: presente e futuro*. Lisboa: Fundação Calouste Gulbenkian; Gama, M. (Org.) (2007), *A política da língua portuguesa*. Braga: Centro de Estudos Lusíadas / Universidade do Minho; e Mateus, M.^a H. Mira & Pereira, D. & Fischer, G. (Coord.) (2008), *Diversidade linguística na escola portuguesa*. Lisboa: Fundação Calouste Gulbenkian.

9. Está bem escrito. Bem estruturado. Bem fundamentado. Por isso, adquiri-lo e proceder à sua leitura reflexiva é uma decisão mais do que oportuna.

Henrique Barroso
Instituto de Letras e Ciências Humanas
Universidade do Minho

**RUTH AMOSSY, (2010). LA PRÉSENTATION DE SOI –
ETHOS ET IDENTITÉ VERBALE,**

Paris, Presses Universitaires de France, 235 páginas.

A investigação em torno do conceito de *ethos* não constitui uma questão nova. Da Retórica à Sociologia, a noção de *ethos* tem sido continuamente objeto de inúmeras discussões intradisciplinares e interdisciplinares.

O interesse linguístico no conceito de *ethos* só recentemente se manifestou. No quadro da Análise Linguística do Discurso, Dominique Maingueneau (1984) é o primeiro a desenvolver uma teoria em volta desta noção. Desde a década de 80, os estudos sobre o *ethos* têm proliferado, com propostas de novas perspectivas e abordagens teóricas; designadamente, D. Maingueneau (1991, 1993, 1999), R. Amossy (1994, 1999), Rabatel (1997, 1998), Kerbrat-Orecchioni (1980), entre outros, têm dado contributos indispensáveis para a consolidação do conceito de *ethos* enquanto objeto de direito no âmbito da investigação linguística dos discursos.

Com a publicação de *La présentation de soi – Ethos et identité verbale*, em 2010, Ruth Amossy, Professora Emérita do Departamento de Estudos Franceses da Universidade de Tel-Aviv, articula o conceito de *ethos* com a noção de *identidade verbal*. O objetivo do estudo de R. Amossy centra-se na análise dos mecanismos discursivos de que o locutor faz uso na construção de uma identidade, no seu posicionamento social e no seu desejo de influenciar o outro.

Esta obra distingue-se não só pelo seu contributo no quadro da investigação sobre o *ethos*, mas, igualmente, pela apresentação de um panorama geral dos estudos realizados em torno deste tema nas últimas décadas.

A obra estrutura-se em duas partes distintas: a primeira compreende os três primeiros capítulos e apresenta os principais fundamentos teóricos em que irá assentar a reflexão em torno do *ethos*; a segunda abarca os quatro últimos capítulos e centra-se na análise dos mecanismos que contribuem para a construção da imagem de si nas trocas verbais. Esta segunda parte encontra-se organizada em torno do valor dos pronomes pessoais «je», «tu», «il» e «nous» na construção da imagem de si.

No primeiro capítulo, são articulados o conceito de *ethos* retórico, na herança de Aristóteles, a noção de «présentation de soi» no âmbito da microssociologia de Goffman e de «image de soi» perspetivada no quadro da Análise do Discurso, nomeadamente, pela teoria desenvolvida por Dominique Maingueneau. Considerando a divergência dos conceitos, mas, sobretudo, a sua confluência, é proposta a assimilação da noção de *ethos* à de «présentation de soi», na sua conceção alargada, ou seja, estendida a todas as trocas verbais.

O segundo capítulo retoma os conceitos de *estereótipo* e *estereotipização* desenvolvidos pela investigadora em trabalho anteriores, designadamente, *Les idées reçues. Sémiologie du stéréotype* (1991), *Stéréotypes et clichés. Langue, discours, société* (1997), em colaboração com Anne Herschberg Pierrot, e *L'argumentation dans le discours* (2000). O *estereótipo* é definido como «une représentation collective figée, un modèle culturel qui circule dans le discours et dans les textes» (2010 :46). Estas representações coletivas constituem uma parte integrante de um dado «imaginaire sociodiscursif» e, nesta medida, encontram-se inseridas numa *doxa*. Estes modelos culturais que preexistem no imaginário coletivo são analisados como uma parte integrante da construção do *ethos*. Nesta conceção, o *estereótipo* é entendido como uma construção de “leitura”, na medida em que o locutário reconstrói, tendo por base elementos díspares contidos no discurso, a imagem do locutor em função de um modelo cultural preexistente e o seu próprio conhecimento do mundo.

O terceiro capítulo aborda a questão do *ethos* pré-discursivo, introduzindo este conceito na teoria desenvolvida em torno do *estereótipo* e relacionando-o com o *ethos* discursivo. A imagem pré-discursiva é entendida como «l'ensemble des donnés dont on dispose sur le locuteur au moment de sa présentation de soi» (2010: 73). Neste sentido, o *ethos* discursivo é perspetivado como uma reação ao *ethos* pré-discursivo: o *ethos* que o locutor constrói de si no discurso inclui sempre a imagem que o interlocutor poderá fazer dele, tendo como base a sua inserção numa representação coletiva, num estereótipo. Nesta linha, a re-elaboração do *ethos* pré-discursivo ocupa um lugar de interesse, especialmente no caso de figuras públicas. Esta problemática encontra-se documentada com variados exemplos, desde o discurso político ao discurso literário.

O quarto capítulo centra-se no pronome pessoal «je» e, como tal, a questão da subjetividade na língua, introduzida no quadro linguístico por E. Benveniste (1970) e, posteriormente, desenvolvida por Kerbrat-Orecchioni (1980), é retomada sob a perspetiva da construção do *ethos*. Segundo

a autora, a utilização da primeira pessoa do singular permite não só a manifestação da subjetividade no discurso, mas também a de uma imagem de si, na medida em que esta imagem emerge das marcas linguísticas da inscrição do sujeito no seu enunciado.

No quinto capítulo, a construção do *ethos* é inserida na dinâmica interacional, sendo, assim, perspectivado enquanto um fenómeno de gestão coletiva.

O sexto capítulo trata da problemática das identidades de grupo, da legitimidade do locutor enquanto intermediário da manifestação da identidade de um grupo de indivíduos e da construção de um *ethos* coletivo.

O último capítulo foca a questão do apagamento enunciativo e da responsabilidade nos discursos de terceira pessoa, centrando-se, concretamente, no discurso científico, filosófico e jornalístico.

La presentation de soi – ethos et identité verbale é, sem dúvida, uma obra de referência no âmbito dos estudos em torno do *ethos*. A apresentação sintética dos conteúdos, a variada documentação de casos concretos e a riqueza de temas abordados, tornam este estudo acessível, também, a não especialistas. Esta obra constitui, assim, uma ferramenta indispensável a futuros investigadores na área da Análise Linguística do Discurso.

Micaela Aguiar

Referências

- MOSSY, Ruth (1991), *Les Idées reçues. Sémiologie du stéréotype*. Paris, Nathan.
- AMOSSY, Ruth (1994), «Stéréotype et Argumentation», In *Le Stéréotype. Crise et transformation*. Caen, Presses Universitaires de Caen.
- AMOSSY, Ruth. & HERSCHBERG PIERROT, Anne (1997), *Stéréotypes et clichés. Langue, discours, société*. Paris, Colin.
- AMOSSY, Ruth. (1999), *Images de soi dans le discours – La construction de l'ethos*. Paris, Delachaud et Niestlé.
- AMOSSY, Ruth. (2000), *L'argumentation dans le discours. Discours politique, littérature d'idées, fiction*. Paris, Nathan.
- BENVENISTE, Emile (1970), «L'appareil formel de l'énonciation», In *Langages* 217, pp. 12-18.
- KERBRAT-ORECCHIONI, Catherine (1980), *L'énonciation. De la subjectivité dans le langage*. Paris, Armand Colin.

- MAINGUENEAU, Dominique (1984), *Genèses du discours*. Liège, Mardaga.
- MAINGUENEAU, Dominique (1991), *L'Analyse du Discours*. Paris, Hachette.
- MAINGUENEAU, Dominique (1993), *Le contexte de l'œuvre littéraire. Énonciation, écrivain, société*. Paris, Dunod.
- MAINGUENEAU, Dominique (1999), «Ethos, scénographie, incorporation», In AMOSSY R. (1999), (éd.) *Images de soi dans le discours. La construction de l'ethos*. Lausanne, Delachaux et Niestlé.
- RABATEL, Alain (1997), *Une histoire de point de vue*. Paris, Klincksieck.
- RABATEL, Alain (1998), *La construction textuelle du point de vue*. Lausanne, Delachaux & Niestlé.

SÁNCHEZ REI, XOSÉ MANUEL (2011), *LINGUA GALEGA E VARIACIÓN DIALECTAL*,

Edicións Laiovento, 662 páginas

Xosé Manuel Sánchez Rey é professor titular da Universidade da Corunha e uma referência na área da Filoloxía Galega e Portuguesa, atestada pela variedade e interesse científico dos textos que já publicou. *Lingua galega e variación dialectal* ocupa na sua vasta bibliografía um lugar de destaque.

Para Xosé Manuel Sánchez Rei, “O desenvolvemento da dialectoloxía galega nestes últimos anos foi, na verdade, sorprendente” (p. 19). Creio que é de toda a justiça integrar aí e em lugar de destaque a obra agora apresentada. De facto, constitui uma reflexão alargada, moderna, no domínio da dialetologia, abordando de forma profunda questões teóricas e metodológicas fundamentais para esta área de investigação, acompanhadas de abundante exemplificação. Vale acentuar que a exaustividade que encontramos no tratamento dos diversos temas é reforçada pela bibliografía que se estende por 57 páginas.

Em termos do quadro teórico adotado, é particularmente relevante a opção por uma perspectiva interdisciplinar, eclética mas crítica. O autor afirma explicitamente a importância de disciplinas como a sociologia, a antropologia, a história da cultura, a psicologia social, entre outras, para o estudo da variação linguística.

É ainda um caso exemplar de interação entre a investigação e o ensino, combinando o rigor científico com uma exposição clara e atrativa.

A obra está estruturada em cinco capítulos, antecidos por uma *Introdução*, onde o autor faz a apresentação de alguns vetores teóricos e metodológicos que estruturam o seu trabalho. Integrada, explicitamente, na área científica da dialetologia, e da dialetologia galega em particular, *Lingua galega e variación dialectal* define como objeto de análise “a situación da Galiza e as súas problemáticas específicas” (p. 21), num espaço linguístico que tem no galego-português a sua origem individualizadora e modernamente se assume como um “... punto de vista, que non é outro do que considerarmos o caso do galego e do português como dúas normas, dúas variedades (...) do mesmo sistema linguístico” (p. 24).

Os cinco capítulos estão organizados de forma a cobrir toda a complexa problemática dos estudos dialetológicos, nas vertentes teórica e metodológica, sempre concretizada pela aplicação pormenorizada a diferentes fenómenos do galego.

O *Capítulo 1. A variação linguística* estabelece como princípios teórico-metodológicos o primado da oralidade como objeto de estudo e uma perspetiva de análise não normativa; e detém-se longamente, a propósito da importância da análise das variedades das línguas, nas variáveis e fatores a considerar no estudo dos usos linguísticos que caracterizam os diferentes tipos de variacionismo.

Os mecanismos de mudança são abordados a partir de diversas teorias, com destaque para a teoria das ondas (p. 106) a teoria da difusão lexical (p. 108) ou ainda a explanação dos fatores internos e externos já apontados, segundo o autor, por Labov. A escolarização maciça das populações leva à reflexão sobre a complexidade do papel da escola na questão da variação linguística.

O capítulo encerra com uma perspetiva histórica sobre mudanças registadas ao longo do tempo em diferentes espaços, que explicam a estigmatização do *dialectal*.

O *Capítulo 2. Língua e Dialecto*, que acentua a centralidade da questão em debate, “o variacionismo e a mudança caracterizam calquera sistema linguístico...” (p. 211), continua a reflexão anterior sobre questões terminológicas. *Língua e dialecto* são objeto de uma cuidada atenção, em particular na destrição de fatores e critérios de definição. O autor opta pelo termo *variedade*, sem conotações negativas, “Fuxindo da carga pexorativa que en ocasións implica a denominación de *dialecto* e a asumirmos a existencia de *variedades* dunha lingua, que resulta un termo desprovido deses matices negativos...” (209-210). Este percurso, em si fundamental para a área da dialetologia, leva o autor a uma reflexão sobre o galego ao longo dos séculos e em particular ao “...feito de o galego ser lingua ou en confronto responder mormente a unha variedade dialectal...” (p. 218).

O *Capítulo 3. A Dialectoloxía* mostra com muita clareza as possibilidades que a dialetologia moderna oferece à investigação, decorrentes de teorias e metodologias mais adequadas. Longe, por exemplo, da tradicional limitação a ambientes rurais, e objetos de análise restringidos à análise fonológica, morfológica e lexical, sobressai a dialetologia urbana ou dialetologia social (p. 246). Complementarmente, a segunda parte deste capítulo desenvolve questões centrais de metodologia de investigação. Tal como no capítulo anterior, o autor finaliza com um extenso subcapítulo dedicado à “Breve historia da dialectoloxía galega...”

O *Capítulo 4. Tipoloxía dos dialectalismos e a súa transcendência* propõe uma perspectiva alargada das “diversas clasificacións que se puideren facer sobre os fenómenos dialectais” (p. 389). Assim, são elencados e discutidos os fatores mais representativos para a construção de uma tipologia de dialetismos. Cada um destes fatores é ilustrado por exemplos variados do galego. É particularmente relevante o facto de o autor não se limitar aos tradicionais domínios da prosódia, fonética, morfologia, etc., mas integrar, nomeadamente, “dialectismos pragmático-textuais” (p. 451). Assim, sustenta que “... do confronto de varios etnotextos que posuímos, non parece resultar conflituoso acreditar, aínda que sexa só provisoriamente, en que se detecta algunha sorte de variacionismo xeográfico tamén no ámbito da pragmática.” (p. 452).

O *Capítulo 5. Dialectoloxía, filoloxía e lingüística* retoma e aplica a perspectiva interdisciplinar referida no capítulo inicial e explora as relações entre dialetologia, linguística e filologia. Em particular, aborda as relações da dialetologia com a gramática histórica, o generativismo, o estruturalismo, a pedagogia, a informática, a tradução, a didática da língua, etc. A propósito de dialetologia e onomástica, Sánchez Rey realiza uma análise fina dos “nomes de lugar e de persoa” galegos. A propósito, mostra a repercussão neste domínio de fenómenos como a gheada ou o rotacismo, entre outros, e as soluções adotadas relativamente a diferentes localidades galegas.

O *Capítulo 6. Conclusións* sintetiza a riqueza das informações, propostas e análises realizadas ao longo dos cinco capítulos anteriores. Subscrevo as palavras (modestamente matizadas) do autor, que a meu ver são reveladoras dessa orientação teórica, metodológica e prática que guia todo o texto: “*Lingua galega e variación dialectal* tencionou mostrar a validade da reflexión lingüística nessa liña, fuxindo do arqueoloxismo fácil e a situar a pesquisa dialectolóxica nun degraio diferente ao dunha visión reducionista do variacionismo.” (p. 602).

A exaustividade da investigação, que apenas aflorámos, fazem deste texto um instrumento de trabalho único para quem pretenda estudar a dialetologia galega. Mas não se fica por aqui. É uma referência obrigatória para todos os que se interessam pela variação linguística.

E ficam assim plenamente justificadas as 662 páginas que acima assinaei.

Maria Aldina Marques

diacrítica

revista do centro de estudos humanísticos
série ciências da linguagem

1. Apresentação

Diacrítica – Série Ciências da Linguagem é uma revista universitária, de periodicidade anual, editada pelo Centro de Estudos Humanísticos da Universidade do Minho (CEHUM) e subsidiada pela Fundação para a Ciência e a Tecnologia. A revista está aberta a propostas de publicação de investigadores internos e externos ao CEHUM que se enquadrem no domínio dos estudos linguísticos. Para além de artigos, sujeitos a arbitragem científica, a revista pode publicar igualmente entrevistas e recensões críticas desde que se enquadrem nos parâmetros temáticos e de qualidade estabelecidos pela Comissão Editorial. Os números editados da *Diacrítica – Série Ciências da Linguagem* a partir de 2003 estão disponíveis em linha na página do CEHUM (<http://ceh.ilch.uminho.pt/diacritica.htm>). As normas de publicação na revista encontram-se igualmente acessíveis em <http://ceh.ilch.uminho.pt>

2. Direção e Comissões Editorial, Científica e Redatorial

Diretora:

Ana Gabriela Macedo (ILCH – U. do Minho)
gabrielam@ilch.uminho.pt

Diretores-Adjuntos:

Carlos Mendes de Sousa (ILCH – U. do Minho)
mdesousa@ilch.uminho.pt

Vítor Moura (ILCH – U. do Minho)
vmoura@ilch.uminho.pt

Comissão Editorial:

Pilar Barbosa (ILCH-U. do Minho)
pbarbosa@ilch.uminho.pt

Cristina Flores (ILCH-U. do Minho)
cflores@ilch.uminho.pt

José Teixeira (ILCH-U. do Minho)
jsteixeira@ilch.uminho.pt

Comissão Científica:

Jorge Morais Barbosa (U. Coimbra); António Branco (U. Lisboa); Ana Brito (U. Porto); Ivo Castro (U. Lisboa); Antonia Coutinho (U. Nova de Lisboa); Maria João Freitas (U. Lisboa); Jürgen M. Meisel (U. Hamburgo / U. Calgary); José Luís Cifuentes Honrubia (U. Alicante); Mary Kato (U. Campinas); Rui Marques (U. Lisboa); Fátima Oliveira (U. Porto); Amadeu Torres (U. Católica Portuguesa); Graça Rio-Torto (U. Coimbra); José Luís Rodrigues (U. Santiago de Compostela); Eduardo Paiva Raposo (U. da Califórnia, Sta. Bárbara); Conceição Paiva (Universidade Federal do Rio de Janeiro); Augusto Soares da Silva (U. Católica Portuguesa).

Comissão Redatorial:

A Comissão Redatorial da *Diacrítica – Série Ciências da Linguagem* integra, para cada número da revista, o conjunto de professores, investigadores e especialistas responsáveis pela revisão científica dos artigos propostos para publicação.

3. Arbitragem Científica

Os artigos propostos à *Diacrítica – Série Ciências da Linguagem* para publicação são submetidos à emissão de pareceres por dois avaliadores (ou três, quando necessário) na respetiva área científica em que o texto se enquadra.

Os artigos são enviados sob anonimato aos *blind referees*, internos e externos ao CEHUM, a quem é solicitado que o parecer emitido tenha em conta, de acordo com a ficha de avaliação adotada pela Revista:

- adequação às normas de publicação da Revista;
- adequação do tema do artigo ao âmbito da Revista;
- pertinência;
- originalidade;

- enquadramento teórico;
- metodologia da recolha e tratamento de dados;
- clareza da apresentação;
- argumentação e relação entre hipóteses de partida e resultados.

Os pareceres deverão incluir uma recomendação em relação a possível publicação, entre as seguintes: publicar sem quaisquer modificações; publicar com pequenas modificações; publicar com modificações significativas; o artigo não se revela adequado para publicação.

Será ainda solicitada a indicação de sugestões e sua justificação, com vista a uma otimização da qualidade científica do artigo submetido a parecer, a ser enviadas, sob anonimato, aos Autores.

A aprovação dos artigos terá lugar até 15 de junho de cada ano civil.

4. Instruções para os Autores

- 1) Todos os artigos, entrevistas e resenhas propostos para publicação na *Diacrítica – Série Ciências da Linguagem* devem ser enviados para o endereço diacritica.linguagem@ilch.uminho.pt até **31 de março** de cada ano civil e elaborados de acordo com as Normas de Publicação descritas neste documento.
- 2) Dos artigos a submeter a publicação na revista devem ser remetidas duas versões eletrónicas, apresentadas em tamanho A4:
 - uma versão anónima, em ficheiro Word;
 - uma versão identificada, em ficheiro pdf, com a afiliação de autor e o respetivo endereço eletrónico (conforme consta nas normas de publicação).
- 3) Os textos das entrevistas e resenhas serão enviados em duas versões eletrónicas, ambas identificadas, uma em ficheiro Word e outra em ficheiro pdf.
- 4) Impõe-se que todos os artigos propostos para publicação sejam originais inéditos, não tendo sido anteriormente publicados, completos ou em parte, quer no formato impresso quer no eletrónico.
- 5) Os textos publicados e as imagens (se as houver) são da responsabilidade dos respetivos Autores.

5. Normas de Publicação

Informações Gerais

1. São aceites originais inéditos escritos em língua portuguesa, inglesa, francesa e espanhola.
2. O título, o resumo e as palavras-chave devem ser apresentados no idioma do texto do artigo e nos idiomas português e inglês.
3. Os artigos e as entrevistas não devem exceder 20 páginas (incluindo as notas e as referências).
4. As resenhas críticas não devem ir além dos 10 000 caracteres com espaços.
5. A afiliação de autores deve ser feita a dois níveis, após o título do artigo e em nota de rodapé, devendo conter informação completa sobre os autores.

Obs.: A partir do próximo n.º 26/3, correspondente à edição de 2012 da *Dia-crítica – Série Ciências da Linguagem*, todos os textos propostos para publicação na revista e redigidos no idioma português deverão seguir o disposto no Acordo Ortográfico da Língua Portuguesa de 1990, em vigor desde 2009.

Afiliação de Autores

1. Após o título do artigo, alinhado à esquerda, deve constar o nome do Autor seguido, na linha imediata, do endereço de correio eletrónico.

Fonte: Times New Roman, Tamanho 12, Espaçamento entre linhas: Simples.

2. Em nota de rodapé, com remissão para o nome do Autor, deve ser feita menção à instituição a que pertence (a dois níveis: Universidade e Departamento ou Centro), bem como à cidade e ao país.

Instruções de Formatação

Títulos

1. O título do original deverá estar em Negrito, Times New Roman 14, alinhado à esquerda.
2. Os títulos das secções e subsecções deverão estar em negrito e o tamanho da fonte deve ser Times New Roman 12.

3. Sugere-se a utilização de, no máximo, dois níveis de titulação, sem numeração ou com numeração árabe (*e.g.* 1, 2.2., mas não 3.2.1.).

Resumo/Abstract e Palavras-chave/Keywords

1. O resumo deverá ser escrito depois do título do artigo, sem parágrafo e com 1 cm de recuo a partir das margens esquerda e direita. Não deve constar a designação Resumo.
2. O texto do resumo não deverá exceder 150 palavras.
3. O número máximo de palavras-chave é 6.

Texto

As páginas deverão ser numeradas.

Corpo de Texto:

- Tipo de letra: Times New Roman
- Tamanho: 12
- Espaçamento: 12 pt
- Alinhamento: Justificado
- Indentação de parágrafos: 1 cm
- Margem superior e esquerda: 3 cm
- Margem inferior e direita: 2 cm

Tabelas

Não use formatações, bordas nem sombreados complicados. As tabelas devem ser identificadas com numeração consecutiva e título, aparecendo em cima da tabela (*i.e.* Tabela 1. Título). Sempre que possível as tabelas devem ser orientadas como “Retrato” e não como “Paisagem”.

Figuras

Todas as figuras, incluindo tabelas e equações que sejam imagens, devem ser incluídas no corpo do texto com referência. As figuras devem ser identificadas com numeração consecutiva e título, aparecendo abaixo da figura (*i.e.* Figura 1. Título). Figuras em branco e preto produzem os melhores resultados, de modo que as coloridas devem ser evitadas.

Numeração de Exemplos, Regras e Fórmulas

Cada item a apresentar (i.e. exemplo, regra ou fórmula) deve ser escrito numa linha de indentação separada, com o número entre parênteses. Devem ser usadas minúsculas pequenas para agrupar conjuntos de itens relacionados.

Exemplo:

- (2) a. Breogán. – Deica pasen as eleicións, non sí?
 b. D. Pepito. – Iso mesmo: deica que as eleicións pasen.

No texto, devem-se referir os itens numerados como (2), (2a,b), (2a- c).

Citação de Formas/Exemplos

Podem ser sublinhadas as palavras citadas no texto. Os exemplos numerados não.

As transcrições devem ser feitas dentro de parênteses rectos ou barras fonéticas. Exemplo: o sufixo [k], a palavra *fé* /fɛ/

As referências específicas a grafemas devem ser feitas da seguinte forma: a letra < q >

Devem-se transliterar ou transcrever todas as formas escritas numa língua que não use o alfabeto latino, a não ser que haja um motivo obrigatório para usar a ortografia original.

As formas escritas numa língua distinta da do artigo, devem ser explicadas/ interpretadas depois da 1ª ocorrência, dentro de aspas. Exemplo: As palavras latinas *canis* ‘cão’ e *equus* ‘cavalos’ são nomes

Notas/Epígrafes

As notas deverão ser em Times New Roman 8, com espaçamento de 10 pt e surgirão em pé de página, com a numeração seguida.

O algarismo que remete para a nota deverá ser colocado depois do sinal de pontuação. Exemplo: “como frequentemente pode ser demonstrado.”⁵”

Nas remissões de umas para outras páginas do artigo, usar-se-ão as expressões latinas consagradas (cf. *supra*, cf. *infra*), que virão sempre em itálico e por extenso.

As epígrafes, que deverão ser em itálico, e em Times New Roman 10, só necessitam da indicação do nome do autor, sendo opcional a indicação do título da obra.

Referências

As citações pouco extensas (até três linhas, inclusive) podem ser incorporadas no texto, entre aspas. (Utilizar a seguinte sinalização para aspas: "..."; e no caso de uma citação com aspas dentro de aspas: "... '...' ...")

As citações mais longas serão recolhidas, ficando impressas em Times New Roman 10, sem aspas, alinhadas, à esquerda, pela indentação de parágrafo do texto.

As interpolações serão identificadas por meio de parênteses rectos [].

As omissões serão assinaladas por reticências dentro de parênteses curvos (...).

No texto, o título das publicações será em itálico e o dos artigos, colocado entre aspas.

Nas Referências devem apenas ser mencionados os autores e obras citadas no artigo.

As referências serão sempre feitas no corpo do texto, na forma abreviada da indicação, entre parênteses curvos, do nome do autor, data de publicação e, se for o caso, número de página. Se se tratar de uma citação indirecta, essas indicações serão precedidas da palavra *apud*.

Exemplos:

Um só autor: (Simenon, 1985: 7). Dois autores: (Sjöwall & Wahlöö, 1985). Nota: o símbolo "&" deverá ser utilizado apenas para referências entre parênteses. No corpo do texto deve-se utilizar "e" no idioma no qual o artigo for escrito.

Exemplo: "Martins e Oliveira (2008) afirmam que..." ou "Martins and Oliveira (2008) state that..."

Três ou mais autores: (Doyle et al., 1973). Nota: a primeira referência a textos de mais de dois autores deverá conter o nome de todos (até um limite de cinco autores), e as demais aparições deverão utilizar "et al."

Exemplo:

(a) primeira referência: "Conforme Silva, Martins e Lira (2009),...";

(b) demais referências: "No entanto, Silva et al. (2009) sugerem..."

Citação indirecta: (*apud* Chandler, 1974: 755).

Autor repetidamente citado: (*Idem*, 10) ou (*Ibidem*) no caso de ser citada a mesma obra na mesma página.

Será incluída no final, em Times New Roman 10, com o título “Referências”, se o texto for em português, e com o título “References”, se o texto for em inglês, a lista completa, por ordem alfabética de apelidos de autores, das obras que tenham sido referidas ao longo do texto.

Se houver duas ou mais referências do mesmo autor e do mesmo ano, acrescentar-se-ão à data as letras a, b, etc. : e.g. Van Dine (1946a), (1946b).

Qualquer informação adicional, tal como a indicação da data original de publicação de um artigo, deve ser apresentada entre parênteses rectos [].

O apelido do autor consultado deverá estar formatado com a fonte Maiúsculas pequenas (no Word, menu Formatar, Fonte, opção Maiúsculas pequenas).

Deverá sempre indicar-se a editora e a edição consultada. Poderá também indicar-se, se for considerada relevante, a data da primeira edição. Estas indicações deverão vir no fim da referência, entre parênteses rectos. Quando se trate de traduções, deverá vir sempre indicado o nome do tradutor.

Exemplos:

LIVROS:

Dürrenmatt, Friedrich (1992), *Der Richter und sein Henker*, Zürich, Diogenes [1950]. Dürrenmatt, Friedrich (1993) *O juiz e o seu carrasco*, trad. Fátima Freire de Andrade, Porto, Asa. Boileau, Pierre & Thomas Narcejac (1982), *Le Roman policier*, Paris, Nathan.

COLETÂNEAS:

Alewyn, Richard (1968), “Anatomie des Detektivromans”, in Jochen Vogt (ed.) (1998), *Der Kriminalroman. Poetik.Theorie. Geschichte*, München, pp. 52-72.

REVISTAS:

Jameson, Fredric R. (1970), “On Raymond Chandler”, *The Southern Review* 6, pp. 624-650. Pizer, John (1987), “History, Genre and ‘Ursprung’ in Benjamin’s Early Aesthetics”, *The German Quaterly*, vol. 60, nº 1, pp. 68-87.

DOCUMENTO NA INTERNET:

Apelido, nome próprio (ano), *Título do Documento*, [em linha] disponível no endereço [consultado em data]. CRUP (1996/1997), *Repensar o Ensino Superior*, Texto do Conselho de Reitores das Universidades Portuguesas, disponível em <http://www.crup.pt/>, consultado em 10/07/2005

LOCAIS NA INTERNET E PÁGINAS PESSOAIS OU DE INSTITUIÇÕES:

Nome, [em linha] disponível em endereço [consultado em data]. Centro de Estudos Humanísticos, disponível em <http://ceh.ilch.uminho.pt/>, consultado em 25/01/10

PUBLICAÇÕES EM REVISTAS NA INTERNET:

Apelido, nome próprio (ano), «Título do Artigo», *Título da Revista*, volume, número, número das páginas, [em linha] disponível em endereço [consultado em data].

CASSEY, Andrew J. (2007), “State Trade Missions”, *Paper Seminar*, School of Economic Sciences, Washington State University, Spring 2008: 1-29, disponível em http://www.ses.wsu.edu/seminar/papers_Spring08/Cassey_State_trade_missions_4.pdf, consultado em 04/05/08 or deverá seguir as mesmas normas já referidas para os artigos.

Se faire + Vinf: un outil au service de la construction d'une diathèse 'maléfactive' de l'objet (in)direct

Silvia Lima Gonçalves Araújo

A descriptive study of Brazilian offensive phrases

Ana C. Bastos-Gee

Menço ou minto? Regularização de paradigmas verbais

Maria João Colaço e Esperança Cardeira

"A beached whale posing in lingerie" Confl ict talk, disagreement and impoliteness in online newspaper commentary

Isabel Ermida

Nota de morfologia histórica do português: sufixo -ó

Przemysław Dębowiak

Revisitando a variação entre ditongos nasais finais átonos e vogais orais na comunidade de fala do Rio de Janeiro

Christina Abreu Gomes, Cássia Mesquita e Tais da Silva Fagundes

Mecanismos de textualização e construção textual: para uma abordagem sócio-discursiva do cartoon

Audria Leal e Ana Caldes

Gramaticalização e especialização funcional: o caso do conector *pois*
Maria da Conceição de Paiva e Maria Luiza Braga

SNS como rótulos em livros didáticos de história do Brasil: simples ou complexos?

Vera Lúcia Paredes Pereira da Silva e Gabrieli Pereira Bezerra

"Componente" como substantivo uniforme de dois géneros

Iva Svobodová

Questões de Ética e Política Linguísticas

A situação atual do ensino da língua portuguesa na Galiza

Xurxo Fernández Carballido

How can sociolinguistic data be used?

Celeste Rodrigues e Deolinda Simões

Recensões

Pinto, Paulo Feytor (2010). *O Essencial sobre Política de Língua.*

Henrique Barroso

Ruth Amossy, (2010). *La présentation de soi – Ethos et identité verbale,*

Micaela Aguiã

Sánchez Rei, Xosé Manuel (2011), *Língua galega e variación dialectal,*
Maria Aldina Marques