

Opacité et transparence prosodique

Jean-Pierre Angoujard

jean-pierre.angoujard@univ-nantes.fr

*Université de Nantes – Laboratoire de Linguistique de Nantes -
LLING (France)*

RÉSUMÉ. Cet article est consacré à l'analyse d'un petit ensemble de données accentuelles en arabe (Le Caire, Damas, dialectes bédouins). Ces données ont été récemment utilisées dans McCarthy (2007) pour défendre la notion d'*opacité* et pour argumenter en faveur des *dérivations* (avec structures sous-jacentes et règles ordonnées), y compris au sein des théories de l'optimalité. En opposition à cette approche, nous montrons que l'appel aux dérivations n'est en aucune manière indispensable et qu'une description *transparente* monostratale (avec un niveau unique pour les représentations) est accessible. Notre analyse, qui s'inscrit dans le cadre de la *phonologie déclarative*, confirme que les processus accentuels sont crucialement dépendants de la *substance*, *i.e.*, de la nature et sonorité des segments.

MOTS-CLÉ. Opacité, phonologie déclarative, accentuation, dialectes arabes, substance.

ABSTRACT. In this article, we focus on a small set of facts relating to stress in Arabic (Cairo, Damascus, bedouin dialect). McCarthy (2007) makes use of these facts (clear cases of *opacity*, according to him) to support view involving derivations (underlying structures and ordered rules or constraints) even inside Optimality Theory. In contrast with this approach, we aim at proving that derivations are truly dispensable and that a *monostratal* description is attainable. Our analysis comes within the framework of *Declarative Phonology* and it shows that stress placement is transparent and depends crucially on the nature and sonority of segments.

KEY-WORDS. Opacity, Declarative Phonology, Stress, Arabic dialects, sonority.

1 – Résumé du premier épisode

Le titre de cet article se réfère à une publication ancienne (Angoujard 1983) qui, déjà, remettait en cause une analyse dérivationnelle

de l'accentuation de mot dans le parler arabe du Caire (Kenstowicz 1979).

J.J. McCarthy (1979) avait initialement développé une analyse métrique (sans référence aucune à la qualité des sommets vocaliques) de l'accentuation dans ce parler:

- a) un pied binaire [s w] (fort-faible) est assigné, de gauche à droite, aux paires de syllabes légères ;
- b) l'ensemble (pieds et syllabes non-intégrées) est rassemblé dans un arbre de mot branchant à *droite* et étiqueté comme suit: dans une paire [n₁ n₂], n₂ est s (= fort) si et seulement s'il branche.

On obtenait alors:

- [_s[_s[bu]_w[xa]] [_w[la]]] = [búxala] («avares»), avec un pied binaire regroupant les deux premières syllabes légères;
- [_w[ʕa]_s[_s[mal]_w[ti]]] = [ʕamálti] («tu as fait - fém.»), sans pied binaire regroupant deux syllabes légères (il n'en existe pas ici, de contiguës), mais avec une branche [s w] associant les deux dernières syllabes (la dernière syllabe est non branchante);

et, surtout (caractéristique de ce parler):

- [_w[ʕaa]_s[_s[la]_w[mu]]] = [ʕalámu] («son monde - à lui») (avec un accent pénultième comme pour [ʕamálti], mais, cette fois, *via* la construction d'un pied binaire final; on notera également que seules les voyelles accentuées sont réalisées longues au Caire).

On opposera cette dernière réalisation à celle, largement majoritaire en domaine arabe et présentant une accentuation antépénultième, soit [ʕá:lamu].

Dans son article de 1979, Kenstowicz faisait état d'un ensemble de données qui n'avaient pas été retenues par McCarthy, spécifiquement une série de formes comparables à [ʕalámu] mais accentuées différemment, ainsi: {saafir}{u}/↔/[sáfru] («ils ont voyagé») (et non *[safiru]). Comme l'objectif de Kenstowicz était clairement de conserver et «sauver» l'analyse accentuelle de McCarthy, il a choisi une solution dérivationnelle (et l'on sait qu'une solution dérivationnelle

est *toujours* possible): il suffisait, bien sûr, d’ordonner l’accentuation *après* la syncope vocalique pour que le problème soit résolu (la voyelle susceptible de recevoir l’accent ayant opportunément disparue).

/saafir-u/ $\xrightarrow{\text{syncope}}$ /saafu/ $\xrightarrow{\text{accent}}$ /sáafu/ $\xrightarrow{\text{abrègement}}$ [sáfru]

Il a été montré, d’une part, que la proposition de Kenstowicz ne suffisait pas à résoudre toutes les difficultés rencontrées par l’analyse de McCarthy (cf. Bohas et Kouloughli 1981) et, d’autre part, qu’une analyse non dérivationnelle était accessible, pour peu que l’on reconnaisse le rôle de la substance (de la nature des sommets vocaliques) dans l’établissement des contraintes accentuelles (il n’est pas besoin de s’interroger longtemps pour admettre que {saalam}{u}/ et {saafir}{u}/ se distinguent par la qualité de leur voyelle pénultième).

Dans le cadre développé extensivement dans Angoujard (1990), on retiendra:

- la dernière syllabe est toujours interprétée comme **H** (lourde)¹;
- dans le parler du Caire les pieds binaires sont tous de type [+ –] (avec tête à gauche);
- l’accent, pour l’ensemble des parlers arabes, porte sur l’avant-dernier pied (sur sa partie forte, *i.e.* la tête, s’il s’agit d’un pied binaire);
- on reconnaît trois type de hiérarchies (plutôt naturelles):
 - **H** > **L** (les syllabes lourdes sont plus «lourdes» que les syllabes légères...);
 - [CVC] > [CV:] (les syllabes lourdes “fermées”, qui contiennent 3 segments, sont plus lourdes que les syllabes incluant une voyelle longue avec seulement deux segments)²;
 - [Ca] > [Ci/u] (il est fait référence ici à la hiérarchie de *sonorité*);

¹ Cette interprétation, commune dans les langues qui opposent voyelles longues et voyelles brèves (c’est notamment le cas en latin), peut être rapportée à une *neutralisation* de la distinction de longueur vocalique en position finale de groupe accentuel.

² Cette hiérarchie est plus explicite dans le cadre du modèle rythmique (Angoujard 1997): les syllabes fermées y sont interprétées comme une succession de deux modèles rythmiques, le second modèle incluant un noyau vide, soit [CV C*] (cf. section 3 ci-dessous).

– les *interactions* entre ces hiérarchies construisent une hiérarchie entre les pieds binaires:

$$\begin{array}{ccccccc} [[CVC] [CV]] & > & [[CV:] [Ca]] & > & [[CV:] [Ci/u]] & > & [[CV] [CV]] \\ 4 & & 3 & & 2 & & 1 \end{array}$$

Cette simple hiérarchie admise, il suffit de se référer, pour le parler arabe du Caire, à une contrainte limitant la construction des pieds binaires aux pieds de niveau 2:

Contrainte: (pour Σ_{bin} = pied binaire; n = valeur de Σ_{bin}),

$$\Sigma_{bin} \Rightarrow (n \leq 2)$$

D'où, avec $n = 4$, [madrása]³ et [yiktíbu]; avec $n = 3$, [ʕalámu]; avec $n = 2$, [sáf(i v ø)ru]⁴; avec $n = 1$, [búxala].

2 – Opacité et dérivations (McCarthy 2007)

A propos de son ouvrage publié en 2007, McCarthy écrit: «*This book is about generalizations that are not quite true [because] the truths they state are hidden by other aspects of the system. When a generalization is partially obscured in this way, it is said to be opaque*» (McCarthy 2007: 1).

L'objectif n'est pas tant, pour l'auteur, de mettre en évidence la présence de généralisations «cachées» (non immédiatement accessibles) que de militer, une nouvelle fois, en faveur de solutions dérivationnelles. Soit un exemple simple (retenu dans McCarthy 2007: 146, qui parle de *Levantine Arabic*):

Dans l'arabe de Damas (comme en latin) l'accent de mot porte sur la syllabe pénultième si elle est lourde; si la pénultième est légère,

³ On rappellera que les séquences [obstruante + liquide] ne sont pas interprétées comme des attaques doubles en arabe et que la suite [madra] doit donc être analysée comme [(mad)(ra)].

⁴ Les syllabes initiales de [ʕalámu] et de [sáf(i v ø)ru] sont de type [CV:]. Plus exactement, dans un cadre déclaratif, ces syllabes sont soumises à une alternance [CV:]~[CV].

il porte alors sur l'antépénultième; il porte sur la dernière syllabe uniquement si elle est «surlourde» soit de type [CVCC]. La 3ème personne du masculin pluriel de l'accompli (suffixe {u}) du verbe *katab* «écrire» sera donc : [kátabu] «ils ont écrit». Or, pour la première personne du singulier de l'accompli (qui se construit par suffixation de {t}), on entend à Damas [katábət] «j'ai écrit», avec présence d'un schwa devant le [t] et, surtout, une accentuation pénultième (sur une syllabe légère). La dernière syllabe n'étant pas une «surlourde», on attendrait, sur le modèle de [kátabu], une réalisation *[kátabət].

Comment rendre compte de cette accentuation *opaque*? Pour McCarthy, l'explication, ne peut être que dérivationnelle: l'accentuation prend place avant l'épenthèse de schwa («*The earliest and most successful theory of opacity is the derivation*» - McCarthy 2007: 2).

/katab-t/ $\xrightarrow{\text{accent}}$ /katábt / $\xrightarrow{\text{épenthèse}}$ [katábət]

Pour appuyer son argumentation en faveur des analyses dérivationnelles, McCarthy (2007: 146) retient une série d'exemples qui, s'ils mettent également à contribution le schwa, sont phonologiquement et morphologiquement plus complexes. Les formes retenues ci-dessous appartiennent à l'arabe de Damas où l'accent porte normalement sur une pénultième lourde, comme en (a) et (b) ci-dessous (la forme entre barres obliques représente le découpage morphologique):

- (a) /katab-na/ katábna «nous avons écrit»
- (b) /katab-l-ak/ katáblak «il t'a écrit»

Par contre, une voyelle «épenthétique» (= non lexicale) n'est pas accentuée dans cette position:

- (c) /katab-l-ha/ katábəlha «il lui a écrit (à elle)»

Sauf si la voyelle épenthétique «rompt» un agrégat consonantique de type CCCC:

- (d) /katab-t-l-ha/ katabtəlha «je lui ai écrit (à elle)»

Ainsi donc, selon l'analyse de McCarthy, une forme comme (c) met en échec la «règle» commune d'accentuation (une pénultième lourde est accentuée), et on ne peut pas seulement justifier l'accentuation antépénultième en supposant que les voyelles «épenthétiques» ne sont pas accentuables, puisqu'une telle voyelle reçoit l'accent en (d). Conclusion: il n'y a plus de généralisation transparente et il faut (le phonologue comme le locuteur?) supposer active une dérivation.

Nous voudrions montrer, dans les trois sections qui suivent, qu'une analyse déclarative (Bird 1991; 1995; Angoujard 2006) de l'accentuation dans l'arabe de Damas est parfaitement possible. Nous commencerons (section 3) par exposer l'analyse syllabique basée sur une description par contraintes, *i.e.* le modèle rythmique développé dans Angoujard (1997). La section 4 présentera une analyse déclarative de l'accentuation en arabe et cette dernière sera appliquée (section 5) à l'arabe de Damas (et aux formes retenues ci-dessus).

3 – Le modèle rythmique

Dans le cadre de Angoujard (1997), la syllabe est rapportée à l'interaction entre trois objets: les segments, la grille rythmique et la courbe prosodique⁵. A toute séquence de sons est associée (au sens de la phonologie plurilinéaire) une *grille rythmique*. La grille rythmique est conçue comme la répétition un nombre *n* de fois du *modèle rythmique*, soit (pour \emptyset représentant un syntagme phonologique et *mod* un modèle rythmique)⁶:

$$\emptyset \rightarrow \text{mod}^n$$

On reconnaît (contrainte C_mod) deux types de modèles rythmiques, un modèle rythmique «léger» (L) et un modèle rythmique «lourd» (H) :

$$C_mod: \quad \text{mod} = \text{mod_L Y mod_H}$$

⁵ Nous ne ferons pas usage dans les analyses qui suivront de la courbe prosodique (mais voir Angoujard 2006: 43-45); les segments, représentés dans cet article par les symboles de l'API, correspondent en droit à des *expressions* ou *compositions d'éléments* (Kaye, Lowenstamm & Vergnaud 1985; aussi Angoujard 2006: 36-40).

⁶ On remarquera que l'organisation rythmique (et donc, entre autres, la syllabation) est définie pour la séquence phonologique et non, comme il est souvent retenu, pour les représentations lexicales (les «mots»). Cette approche est cruciale, qui suppose une construction de la structure syllabique (plus généralement prosodique) *en actes* et ne recourt donc à aucune resyllabification.

Un modèle rythmique est composé d'une position (ou *creux* rythmique) initiale type *init* suivie d'une position noyau (ou *sommet rythmique*, type *som*) et, facultativement (il s'agit alors d'un modèle rythmique de type H) d'un autre creux rythmique (position finale, type *fin*). Soit, sous forme de grille rythmique, l'exemple du mot français [paʁti] «parti», composé de deux modèles rythmiques:

x	\bar{x}	x	x	\bar{x}
p	a	ʁ	t	i

Nous aurons donc, universellement, les deux contraintes C_mod_L et C_mod_H décrivant respectivement les modèles rythmiques L et H :

$$C_mod_L: mod_L = \{init \wedge som\}$$

$$C_mod_H: mod_H = \{init \wedge som \wedge fin\}$$

Il est important de rappeler que, dans le cadre de Angoujard (1997), seules les consonnes les plus sonores (glides, liquides, nasales) peuvent être associées à une position finale de modèle H (comme c'est le cas dans le premier modèle de [paʁti] ci-dessus). De ce fait, toutes les séquences $[C_iVC_j]$ où C_j est une obstruante sont interprétées comme la succession d'un modèle L (soit, $[C_iV]$) et d'un second modèle L avec un noyau *vide* (soit $[C_j^*]$).

Les deux contraintes C_mod_L et C_mod_H peuvent naturellement être complétées pour intégrer des valeurs paramétriques propres à telles et telles classes de langue. Ainsi, en arabe (mais c'est également vrai en latin), tout modèle rythmique en position finale de mot⁷ est *interprété* comme H. Cette caractéristique peut être intégrée dans une contrainte $C_mod_L_ar$, en y ajoutant l'information selon laquelle ce modèle de type L ne peut se trouver en fin de mot (représentée ici par une frontière] ω):

$$C_mod_L_ar: \quad mod_L = \left\{ init \wedge som \right\}] \square \wedge \neg (] \square \omega)$$

⁷ Il faut entendre «mot» comme une simplification pour «mot accentuel».

Par ailleurs, en arabe toute voyelle longue apparaît comme brève en position finale de mot, ce qui implique qu'on ne peut trouver dans cette position de modèle H correspondant à [CV:]⁸.

D'autre part, toute séquence [CVC_j] est interprétée *en fin de mot* comme [CVC_j*], même si C_j est une sonante: ainsi, les formes verbales [katab] «il a écrit» et [faʕal] «il a fait» sont-elles métriquement (syllabiquement) identiques.

Nous poserons alors une contrainte C _ mod _ H _ ar qui retient:

a) qu'un modèle H peut être naturellement de type [CVC] mais que dans ce cas il ne peut pas se trouver en fin de mot (où on aurait [CVC*])

b) qu'un modèle H peut être de type [CVC*]⁹

c) qu'il peut être de type [CV] mais seulement s'il se trouve en fin de mot (miroir de la description fournie pour C_mod_/ ar)

$$C_mod_ar: \quad mod_H = \left\{ \begin{array}{l} \text{a): } \{init \wedge som \wedge fin\} \wedge \neg(\square] \omega) \\ \vee \\ \text{b): } \{mod_L \wedge (mod_L \mid som \mid Seg_{\square} \wedge var(\square))\} \\ \vee \\ \text{c): } \{init \wedge som\}_{\square} \wedge (\square] \omega) \end{array} \right\}$$

Les contraintes, dans le cadre de la phonologie déclarative, sont des représentations partielles. Pour expliciter ce point, nous donnons ci-dessous la représentation du mot français [paʁti], qui, comme celle

⁸ On peut certainement trouver étrange qu'une séquence CV soit interprétée comme lourde en fin de mot, tandis que les séquences [CV:] sont «réduites» dans ce même contexte. Ceci devient beaucoup moins étrange lorsqu'on reconnaît que ces deux caractéristiques *conspirent* en vue d'une neutralisation de la longueur vocalique.

⁹ Le noyau vide est décrit, dans le cadre des structures de traits, comme un type *som* ayant une variable (*var*) comme valeur de son attribut SEG.

de toute séquence, est *équivalente* à une conjonction de contraintes non ordonnées, soit¹⁰:

$$\begin{array}{l}
 \langle p_{[1]}, a_{[2]}, \#_{[3]}, t_{[4]}, i_{[5]} \rangle \quad (\text{représentation lexicale}) \\
 \wedge \\
 [[1] \emptyset [2] \emptyset [3] \emptyset [4] \emptyset [5]] \quad (\text{contrainte d'ordre linéaire}) \\
 \wedge \quad \phi \rightarrow \text{mod}^+ \quad (\text{grille rythmique}) \\
 \wedge \quad (C_{\text{mod}} \wedge C_{\text{mod}_L} \wedge C_{\text{mod}_H}) \Rightarrow \\
 \quad \langle p, a, \# \rangle = \{ \text{init} \wedge \text{som} \wedge \text{fin} \} \\
 \quad \langle t, i \rangle = \{ \text{init} \wedge \text{som} \}
 \end{array}$$

4 – L’accentuation en arabe

Au-delà de la syllabe (ou du modèle rythmique), toutes les langues organisent leur prosodie en constituants de niveau supérieur, les *pieds*. Dans le cadre de Angoujard (1990; 1997) les pieds sont maximalelement binaires. Les suites de deux modèles rythmiques de type L sont régulièrement associées pour constituer un pied binaire, soit Σ [L L]. Inversement, un pied binaire ne peut inclure deux modèles H (Σ^* [H H]). La présence ou non de pieds incluant un modèle L et un modèle H est dépendante de la valeur paramétrique retenue par telle ou telle langue : Σ [L H]/[H L]? (*oui/non*).

Tout regroupement en pied binaire est inégalitaire: l’un des deux modèles rythmiques est interprété comme «fort» (+) en face d’un modèle «faible» (-). Les langues se distinguent par le choix d’une hiérarchie [+–] ou [–+]; au sein d’une langue donnée tous les pieds binaires respectent nécessairement la même hiérarchie. Si une langue accepte de construire des pieds de type [L H]/[H L], ceux-ci ne peuvent contredire la hiérarchie [+–]/[–+] définie pour cette langue: autrement dit, une langue de type [+–] ne pourra construire que des pieds [H L].

Les pieds unaires sont interprétés comme «forts» (+).

Pour toute langue nous aurons donc les contraintes ci-dessous¹¹:

¹⁰ Le symbole \emptyset est un opérateur de précedence immédiate.

¹¹ Σ représente le pied ; Σ_{un} et Σ_{bin} respectivement, un pied unaire et un pied binaire ; le pied binaire peut être dominant à gauche (Σ_{bin_C}) ou dominant à droite (Σ_{bin_D}) ; les valeurs + ou – sont associées à un attribut VAL- Σ .

Pieds: $\Sigma = \Sigma_{un} \vee \Sigma_{bin}$
 Pieds_Un: $\Sigma_{un} = \Sigma[[mod]_{[1]} \wedge ([1]|VAL-\Sigma : +)$
 Pieds_Bin: $\Sigma_{bin} = \Sigma_{bin_G} \vee \Sigma_{bin_D}$
max_ Σ_{bin} (le plus grand nombre possible de pieds binaires sont construits.)

Pour prendre l'exemple du parler de Damas, ce parler est de type [+–] et construit tous les pieds incluant un modèle L et un modèle H. Nous y trouverons donc les contraintes :

Pieds_Damas : $\Sigma = \Sigma_{un} \vee \Sigma_{bin_G}$
 (résolution pour Damas de l'alternance ($\Sigma_{bin_G} \vee \Sigma_{bin_D}$)

$\Sigma_{bin_G} \equiv (\Sigma[[mod]_{[1]} \wp [mod]_{[2]}]_{\Sigma})$
 $\wedge ([1] = (L \vee H))$
 $\wedge ([1] |VAL-\Sigma : +) \wedge ([2] |VAL-\Sigma : -)$

Enfin, nous avons vu dans la section 1 que l'accent de mot, pour tous les dialectes de l'arabe, se situe sur l'avant dernier pied :

$$\left[\dots \Sigma |mod \left[\begin{array}{l} \text{POIDS}_{\Sigma} : + \\ \text{POIDS}_{\omega} : 4 \end{array} \right] \wp \Sigma |mod \left[\begin{array}{l} \text{POIDS}_{\Sigma} : + \\ \text{POIDS}_{\omega} : 3 \end{array} \right] \right]_{\omega}$$

5 – La transparence damascène

Soit les deux formes (retenues dans McCarthy (2007) et déjà présentées ci-dessus, section 2):

- (a) /katab-na/ katábna « nous avons écrit »
- (b) /katab-l-ak/ katáblak « il t'a écrit »

Dans ces deux formes l'accent porte « régulièrement » sur une antépénultième lourde. Cette accentuation est directement prédite (aussi par McCarthy) et *transparente*.

Pour (a), où aucun pied binaire ne peut être construit :

$\emptyset \rightarrow mod^+$

$(C_mod \wedge C_mod_L \wedge C_mod_H) \Rightarrow$

$-\langle k, a \rangle = \{init \wedge som\}$ (C_mod_L)

$-\langle t, a, b, * \rangle = \{init \wedge som \wedge init \wedge som \mid var\}$ (C_mod_H b)

$-\langle n, a \rangle = \{init \wedge som\}$ (C_mod_H c)

$\Sigma \wedge \Sigma_{un} \wedge \Sigma_{bin_G} \Rightarrow$

$-\langle k, a \rangle = [VAL_ \Sigma : +] \wedge [POIDS \omega : 3]$

$-\langle t, a, b, * \rangle = [VAL_ \Sigma : +] \wedge [POIDS \omega : 4]$

$-\langle n, a \rangle = [VAL_ \Sigma : +] \wedge [POIDS \omega : 3]$

La forme (b) est largement comparable, avec ${}_{\omega}[_{\Sigma}[ka]_{\Sigma}[tab*]_{\Sigma}[lak*]]$, le dernier pied étant cette fois décrit, de même que l'avant-dernier, *via* la contrainte (C_mod_H b)

Soit maintenant les formes (c) et (d), qui ont la particularité de contenir une voyelle «épenthétique» (non lexicale) et, pour la forme (c), de ne pas être accentuée sur l'avant-dernière syllabe :

(c) /katab-l-ha/ katábølhə «il lui a écrit (à elle)»

(d) /katab-t-l-ha/ katabtølhə «je lui ai écrit (à elle)»

La forme (d) est tout à fait régulière :

(d) = ${}_{\omega}[_{\Sigma}[ka]_{\Sigma}[tab*]_{\Sigma}[_{\sigma}[t*]_{\sigma}[l*]]_{\Sigma}[ha]]$

Un pied binaire est construit sur les deux syllabes à noyau vide adjacentes. Une contrainte très répandue dans les langues (elle est, en particulier, active en français, *cf.* Angoujard 2006: 91-93), impose (d'une manière passablement naturelle) que tout sommet en position forte (+) soit instancié (*i.e.* qu'il soit doté d'une réalisation sonore). Dans Angoujard (2006), cette contrainte est intégrée à la description du type SOM : un sommet qui a une valeur pour le trait POIDS supérieure à 2 ne peut correspondre à une variable (il doit être instancié):

$$som = \left\{ \left[\begin{array}{ll} \text{SEG :} & Seg \\ \text{POS :} & 2 \\ \text{PC :} & Pc \\ \text{POIDS :} & Poids \end{array} \right] \wedge \left((Poids > 2) \Rightarrow (\neg var(Seg)) \right) \right\}$$

Cette instanciation, en l'absence d'information lexicale, fait appel à la voyelle *par défaut*. Dans l'arabe de Damas, cette voyelle est un schwa et il n'y a donc rien à ajouter pour rendre compte de la forme [katabálha].

La forme (c) et la plus problématique (elle fait état d'une accentuation antépénultième alors que sa syllabe pénultième est lourde); c'est aussi la plus intéressante et celle qui, lue comme «opaque», conduit McCarthy à proposer une dérivation.

Il faut tout d'abord rappeler que la voyelle schwa est régulièrement présente (et accentuable) dans l'arabe de Damas, non seulement dans le cas où il s'agit de la réalisation par défaut d'un sommet en position forte (comme nous venons de le voir avec [katabálha]), mais également comme version «réduite»¹² des voyelles moyennes qui n'apparaissent comme telles que dans la dernière syllabe du mot :

[ʕámel]	«il a fait»	[ʕmált]	«j'ai fait»
[sámeʕ]	«il a entendu»	[səmáʕkon]	«il vous a entendu» (plur.) _
[btáktob]	«tu écris (masc.)»	[btəktábəlna]	«tu nous écris (masc.)»

Il existe donc trois origines de réalisations de type schwa (dans les variétés libanaises, on rencontre des réalisations de type [i], cf. «*i-like sounds*» – Cowell 1964, p. 13) :

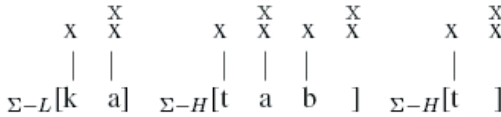
- des alternances lexicales (ə v (e/o))
- des réalisations par défaut de sommets en position forte (+)
- des voyelles dites «de transition» (jamais accentuées et sans effet rythmique)

C'est cette «voyelle de transition» qui est également présente dans la forme [katábəʔ] «j'ai écrit» avec accent pénultième non attendu (cf. ci-dessus, section 2).

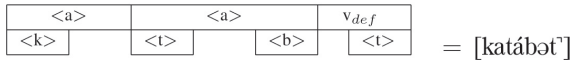
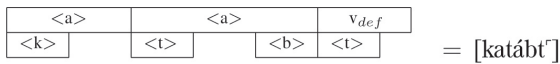
Dans le parler de Damas (et, plus généralement, en Syrie), cette forme peut être réalisée avec une voyelle de transition (soit, [katábəʔ])

¹² Dans le cadre de la phonologie déclarative ces «réductions» vocaliques sont naturellement interprétées comme le résultat de la résolution d'alternances [e/o]–[ə].

ou sans voyelle de transition (soit, [katábɔt] et, cette fois, l’accentuation est, pour tous, en tout point régulière). On remarquera en tout premier lieu (et la solution, à vrai dire, est contenue dans cette remarque) que les deux réalisations peuvent correspondre à une seule et même représentation rythmique (rappelons que les voyelles non lexicales sont interprétées comme la réalisation par défaut d’un sommet vide, cf. Charrette 1991; Angoujard 1997; 2006) :



On peut naturellement se demander pour quelle raison la présence d’une voyelle (même par défaut) ne modifie pas l’organisation rythmique. Une première réponse (on suivrait alors Coleman 1996; 1998) serait de considérer que la présence de cette voyelle par défaut (alors dite, précisément, «de transition») est un phénomène *d’interprétation phonétique* en tout point distinct de la description phonologique, spécifiquement un phénomène *d’interprétation temporelle*. A cette opposition temporelle serait alors associée une opposition de type «fermante» vs «ouvrante» pour l’occlusive finale (dans le schéma ci-dessous, la voyelle de transition est représentée par v_{def}):



Cette approche, systématique dans le cadre retenu par Coleman, a l’inconvénient de dresser une opposition de nature entre le «phonologique» et le «phonétique». Il n’est pas aisé de choisir, pour un processus donné, entre une analyse phonologique et un abandon aux variations (gradient, intervalle temporel). La phonologie déclarative, monostratale, peut être vue comme un encouragement à restreindre au maximum le recours au non phonologique. Après tout, l’ensemble

des contraintes (ou représentations partielles) décrit la représentation entière (l'objet phonique) dont, par exemple, la présence ou non d'une «transition» vocalique fait partie.

En tout état de cause, pour le processus dont la forme [katábət] est exemplaire, une analyse déclarative est certainement possible. De nombreuses langues se réfèrent à une hiérarchie vocalique qui fait usage d'une distinction phonologique entre voyelles «fortes» (généralement les plus sonores, représentées comme V) et voyelles «faibles» (généralement les moins sonores, représentées comme v)¹³. Si on admet l'existence de noyau vide (non instanciés), on retient une hiérarchie

$$V > v > \emptyset.$$

Il est possible d'intégrer à cette hiérarchie l'objet *sommet par défaut* pour obtenir une hiérarchie étendue : $V > v > v_{def} > \emptyset$. Il suffit alors de modifier, en référence à cette hiérarchie, la contrainte décrivant le modèle rythmique H incluant un noyau vide (C_mod_H_ar b, ci-dessus section 3) pour qu'il puisse décrire, non seulement une suite [CVC*], mais également la suite avec voyelle par défaut [CVCV_{def}]. Cette contrainte (C_mod_H_ar b' ci-dessous) ne propose rien d'autre que de mettre en évidence le fait que les locuteurs/auditeurs de Damas ne font pas de distinction prosodique entre un noyau vide et un noyau rempli par une voyelle par défaut. Mais elle met aussi en évidence le fait que ces données peuvent parfaitement être intégrées dans une grammaire monostratale et ne nécessite aucunement le recours à des dérivations.

$$C_mod_H_ar : = b') : \left\{ mod_L \wedge \left(mod_L \mid som \mid Seg_{\square} \wedge \left(\square = \left(var \vee v_{def} \right) \right) \right) \right\}$$

La réalisation par défaut d'un sommet vide appartenant à un pied H de type [CVC*] est facultative pour une forme comme [katábt]~[katábət]; elle est obligatoire pour une forme comme

¹³ Ainsi avons-nous vu (section 1) que le parler du Caire fait usage de la hiérarchie [a] > [i, u]; le français oppose l'ensemble de ses voyelles «pleines» au schwa : $V > [ə]$. Quelle que soit l'opposition retenue au sein de l'échelle de sonorité, aucune langue n'intégrera à l'ensemble v une voyelle plus sonore que celles contenues dans l'ensemble V (la hiérarchie de sonorité ne saurait être contredite).

[katábəlha] (*[katáblha]) par soumission à une contrainte très fréquente dans les langues et qui ne permet pas la succession de deux sommets non instanciés.

6 – Une seconde étude de cas (McCarthy 2007: 179).

Dans son argumentation en faveur d'une phonologie dérivationnelle, McCarthy prend également l'exemple d'un dialecte bédouin déjà bien étudié (Cantineau 1936; Al-Mozayni 1981; Al Mozainy *et al.* 1985 ; Angoujard 1993a).

La caractéristique principale de ce parler tient au fait que les pieds binaires construits sont dominants à droite [-+], ce qui le distingue singulièrement des autres dialectes orientaux et le rapproche, sur ce point, de ceux du Maghreb (Angoujard 1990; 1993a). On trouvera donc la contrainte:

$$\Sigma = \Sigma_{un} \vee \Sigma_{bin_D}$$

Ce parler est «différentiel», c'est-à-dire qu'il fait usage d'une distinction (adossée à la sonorité des segments) entre deux types de sommets vocaliques, que nous représenterons comme V vs v.

Nous avons vu dans la section 1 que le parler arabe du Caire limite la construction des pieds binaires aux pieds incluant un modèle H à voyelle longue (y sont donc exclus tous pieds [HL] de type *[[CVC][CV]]). Ce parler bédouin fait intervenir, pour les pieds de type [L H] (rappelons qu'il s'agit d'une langue [-+]), la voyelle du modèle rythmique L: ce pied ne peut être construit que si son modèle L contient une voyelle de type v (ou, si l'on veut, on ne trouvera pas de pieds *[[CV] H]¹⁴.

Ce parler différentiel retient deux types (deux classes) de verbes trilitères :

- a) type {C₁VC₂VC₃} (correspondant au type {faʕala} de l'arabe classique)

¹⁴ On remarquera que, quelques soient leurs variétés et spécificités, toutes les restrictions (contraintes) susceptibles de peser sur la construction des pieds binaires ont une nature commune: elle se réfère à une hiérarchie des pieds (et toutes peuvent être décrites comme une limitation aux pieds < n) et cette hiérarchie est *substantielle*, dans le sens où elle s'appuie soit sur le *poids* (H > L), soit sur la *sonorité*, soit sur une combinaison des deux.

$$\begin{array}{ccccccc}
 & & \begin{array}{c} \text{x} \\ \text{x} \\ \text{x} \end{array} & & \begin{array}{c} \text{x} \\ \text{x} \end{array} & & \begin{array}{c} \text{x} \\ \text{x} \\ \text{x} \end{array} \\
 & \text{x} & | & \text{x} & | & \text{x} & | & \text{x} \\
 (3) & \Sigma[\sigma[\text{s} & \text{V}] & \sigma[\text{m} & \text{v}]] & \Sigma[\sigma[\text{ʔ} & \text{a} & \text{t} & \text{ }]] = [\text{sámʔa}] \\
 & & [+ & & [- & & + & &] \\
 & & \Rightarrow & \text{pied binaire de type} & [[\text{Cv}] & \text{H}].
 \end{array}$$

$$\begin{array}{ccccccc}
 & & \begin{array}{c} \text{x} \\ \text{x} \\ \text{x} \end{array} & & \begin{array}{c} \text{x} \\ \text{x} \\ \text{x} \end{array} & & \begin{array}{c} \text{x} \\ \text{x} \end{array} \\
 & \text{x} & | & \text{x} & | & \text{x} & | & \text{x} \\
 (4) & \Sigma[\sigma[\text{k} & \text{V}] & \sigma[\text{t} & \text{V}]] & \Sigma[\sigma[\text{b} & \text{a} & \text{t} & \text{ }]] = [\text{ktíbat}] \\
 & & [- & & + & & + & &] \\
 & & \Rightarrow & \text{pied binaire de type} & [\text{L} & \text{L}].
 \end{array}$$

Ces paradigmes mettent aussi en évidence une alternance [a]~[i] très caractéristique de ce parler. L'examen des formes des verbes *sími* ʔ et *kítáb* ont conduit plusieurs auteurs (en particulier Al-Mozainy (1981) et McCarthy 2007) à supposer l'existence d'une règle de réduction «en syllabe ouverte» du type /a/→[i] («raising» dans McCarthy 2007: 177 sq.). Bien entendu cette approche suppose que les formes en question sont *dérivées*:

/sami ʔ/ → [símiʔ]

/katab/ → [kítáb]

Rien ne contraint à admettre l'existence d'une représentation sous-jacente (et donc à imaginer que le locuteur/auditeur se référerait, par exemple, à une imaginaire forme radicale /katab/ pour produire/interpréter les formes conjuguées de l'accompli, alors même qu'il n'apparaît jamais – que l'on n'entend jamais – de [a] entre les deux premières consonnes de ce verbe).

Puisque nous avons affaire à des alternances, autant les traiter pour ce qu'elles sont, des alternances entre représentations¹⁶. La

¹⁶ On pourrait insister en parlant d'alternances entre représentations *de surface*. Ce ne serait pas faux, mais quelque peu superflu dans la mesure où la phonologie déclarative est monostratale et se réfère à un niveau de représentation unique.

particularité de ce parler bédouin tient au fait qu'il ne fait pas seulement usage de l'alternance classique entre les voyelles appartenant à la classe v et zéro :

$$v \equiv ([i]/[u] \vee \emptyset),$$

mais également d'une alternance globale au sein de la classe V (alors que la plupart des dialectes de l'arabe préservent la voyelle [a] de toute alternance prosodique, que ce soit avec une voyelle de sonorité inférieure ou, plus encore, avec \emptyset) :

$$V \equiv ([a] \vee [i]/[u] \vee \emptyset).$$

Ces alternances sont gouvernées par deux contraintes prosodiques. La première contrainte est tout à fait habituelle, qui retient la valeur \emptyset pour tout sommet en position faible dans un pied binaire¹⁷ :

$$(a) L(-) \Rightarrow \emptyset$$

Le seconde contrainte impose, pour les sommets de type V , le choix de la réalisation [i] ([u]) lorsque que le sommet vocalique appartenant à un modèle L est non final et suivi d'un sommet instancié :

$$(b) \left(\left[\begin{array}{l|l} \text{mod}_L & \text{som} \\ \hline \text{SEG} : \langle a \vee i \vee \emptyset \rangle > \mathbb{I} \\ \text{POIDS} : \mathbb{I} \end{array} \right] \wp \left[\text{mod} \mid \text{som} \mid \text{SEG} : \text{Seg}(\mathbb{I}) \right] \wedge (\mathbb{I} > 2) \wedge \text{nonvar}(\mathbb{I}) \right) \\ \Rightarrow \left(\mathbb{I} = [i] \right)$$

Cette contrainte peut naturellement être lue comme une «version» de la règle de réduction *raising* retenue par McCarthy. La différence fondamentale tient au fait que la contrainte est intégrée dans une ana-

¹⁷ On rappellera, par exemple, que c'est une contrainte semblable qui est à l'origine de la «disparition» de la plupart des voyelles latines non accentuées lors de la constitution du français.

lyse non dérivationnelle, qu'elle ne suppose en aucune façon que la forme *kítab*, avant toute réalisation, doit être identifiée comme /kátab/. Elle reconnaît seulement que le locuteur sait que le verbe a une première voyelle de type V et que, dans ce contexte, il s'agit d'un [i].

Les données sont encore un peu plus complexes, dans la mesure où une gutturale adjacente imposera la réalisation d'un [a] et non d'un [i] dans ce contexte. Ainsi réalise-t-on [šáħab] «accompagner», à comparer avec [kítab]. Cette présence de la voyelle basse est justifiée dans Angoujard (1993b) par le fait que les gutturales contiennent l'élément **A** et qu'une contrainte impose à cet élément (de sonorité maximale) d'être associé à un sommet vocalique.

Il n'est évidemment pas indispensable, cette fois encore, de supposer un ordre (et donc une dérivation) qui supposerait une réalisation, probablement intermédiaire, /šíħab/, qu'une règle d'harmonie transformerait en [šáħab]. Il est parfaitement possible de seulement compléter la contrainte (b) ci-dessus en imposant le choix de la voyelle <a> dans un contexte guttural.

McCarthy (2007) utilise un argument supplémentaire en faveur de l'usage des dérivations: la syncope des voyelles hautes. Nous avons déjà rencontré ce processus (pour nous une alternance entre [i]/[u] et Ø) dans une forme comme [sámʕat] (3ème personne du féminin singulier de l'accompli du verbe [símiʕ]). L'absence de la voyelle haute dans cette forme (le choix de Ø au sein de l'alternance lexicale) est directement gouvernée par la contrainte sur les positions faibles à l'intérieur d'un pied binaire (cf. ci-dessus).

Trois formes verbales supplémentaires sont retenues par McCarthy, pour lesquelles il affirme que les voyelles hautes affectées par la syncope se trouvaient dans une position *accentuable* (le tableau ci-dessous contient les formes sous-jacentes supposées) :

/kitib/	ktíb	«il a été écrit»
/kitib-at/	kítbat	«elle a été écrite»
/ʃarib-at/	ʃárbat	«elle a bu»

«Syncope of high vowels from positions where they could otherwise be stressed is evidence of an opaque, counterbleeding interac-

tion between stress and syncope.» (McCarthy 2007: 187). Clairement, [ktíb] est comparée (opposée) à [símiʔ], [kítbat] et [ʃárbat] à [ktíb].

La forme [ʃárbat] est parfaitement régulière, en tout point identique à la forme [sámʃat] (et ne nécessite donc ni forme sous-jacente, ni dérivation). Pour les formes [ktíb] et [kítbat], il convient juste (et c'est la moindre des choses) de se rappeler qu'il s'agit de formes *passives* (on sait que l'accompli passif est construit, en arabe classique, comme {fuʃila}). Autrement dit, et parce qu'il s'agit du passif, ces formes verbales ont pour première voyelle radicale une voyelle de type *v*, soit {C₁vC₂vC₃} et, pour l'exemple choisi, {k,v,t,v,b}. Ce point acquis, les formes attestées sont régulières et les contraintes que nous avons retenues les décrivent très directement, sans nécessiter une syncope préalable à l'accentuation :

$$(5) \quad \begin{array}{ccccccc} & & x & \overset{x}{\underset{x}{x}} & & x & \overset{x}{\underset{x}{x}} \\ & & | & | & & | & | \\ \Sigma[\sigma[k & v] & [\sigma[t & v & b] &] & = [ktíb] \\ & & [- & + & & & \end{array}$$

$$(6) \quad \begin{array}{ccccccc} & & & \overset{x}{\underset{x}{x}} & & & \overset{x}{\underset{x}{x}} \\ & & x & | & x & \overset{x}{\underset{x}{x}} & x & \overset{x}{\underset{x}{x}} \\ & & | & | & | & | & | & | \\ \Sigma[\sigma[k & v]] & \Sigma[\sigma[t & v] & [\sigma[b & a & t] &]] & = [kítbat] \\ & & [+ & - & + & & & \end{array}$$

Nous pensons que les alternances présentes dans les formes verbales de ce dialecte bédouin, comme son accentuation, n'imposent en aucune manière le recours à des dérivations (avec formes sous-jacentes et règles ou contraintes ordonnées). Rien, ici, n'est opaque. Tout est transparent – et cette transparence demande seulement que le locuteur/auditeur de ce dialecte ait été à même de repérer (à partir des formes prononcées et entendues) que la première voyelle peut être, selon le contexte, pour certains verbes, [a], [i/u] ou zéro, et pour d'autres verbes, simplement [i/u] ou zéro. Une tâche qui ne nous paraît pas excessive.

RÉFÉRENCES

- Al-Mozainy, H. Q. 1981. *Vowel Alternations in a Bedouin Hijazi Arabic Dialect: Abstractness and Stress*. PhD thesis. Austin: University of Texas.
- Al-Mozainy, H. Q.; Bley-Vroman, R.; McCarthy, J. J. 1983. Stress shift and metrical structure. *Linguistic Inquiry*. **16**: 135-144.
- Angoujard, J.-P. 1983. Opacité et transparence prosodique. *Recherches Linguistiques*. **11**: 85-113.
- Angoujard, J.-P. 1990. *Metrical Structure of Arabic*. Dordrecht: Foris.
- Angoujard, J.-P. 1993a. Syncope et voyelles accentuées (parler du Hedjaz, Bani Hassan et Hébreu biblique). *Langues Orientales Anciennes. Philologie et Linguistique*. **4**: 103-123.
- Angoujard, J.-P. 1993b. Une analyse de la forme [a ámar] (parler du Hedjaz) dans le cadre de la phonologie déclarative. *Linguistica Communicatio*. **V(1&2)**: 105-119.
- Angoujard, J.-P. 1997. *Théorie de la syllabe. Rythme et Qualité*. Paris: CNRS Editions.
- Angoujard, J.-P. 2006. *Phonologie déclarative*. Paris: CNRS Editions.
- Bird, S. (Ed.). 1991. Declarative Perspectives on Phonology. *Edinburgh Working Papers in Cognitive Science*. **7**. [Centre for Cognitive Science, The University of Edinburgh].
- Bird, S. 1995. *Computational Phonology. A constraint-based approach*. Cambridge: Cambridge University Press.
- Bohas, G.; Kouloughli, D. 1981. Processus accentuels en arabe. *Analyses, Théorie*. **1**: 1-59.
- Cantineau, J. 1936. Études sur quelques parlers nomades d'Orient. *Annales de l'Institut d'Etudes Orientales d'Alger*. **II et III**.
- Charette, M. 1991. *Conditions on Phonological Government*. Cambridge: Cambridge University Press.
- Coleman, J. 1996. Declarative syllabification in Tashlhiyt Berber. In: J. Durand; B. Laks (Eds.). *Currents Trends in Phonology: Models and Methods*. Salford: ESRI/ University of Salford, 175- 216.
- Coleman, J. 1998. *Phonological representations. Their names, forms and powers*. Cambridge: Cambridge University Press.
- Cowell, M. W. 1964. *A Reference Grammar of Syrian Arabic*. Washington DC: Georgetown University Press.
- Kenstowicz, M. 1979. Notes on Cairene Arabic syncope. In: *Proceedings of the 10th Annual Conference on African Linguistics*.
- McCarthy, J. J. 1979. On stress and syllabification. *Linguistic Inquiry*. **10(3)**: 443-465.
- McCarthy, J. J. 2007. *Hidden Generalizations: Phonological Opacity in Optimality Theory*. London/Oakville: Equinox.

Turn on your TAP: memory in language processing

Danuta Gabryś-Barker

danutagabrys@hotmail.com

University of Silesia, Institute of English (Poland)

ABSTRACT. Memory has always constituted an important aspect of research on language acquisition and learning. Traditionally, this research looks at the processes involved in memorising language elements, that is, memory learning strategies in FL vocabulary acquisition or the processes involved in the recall of language, for example retrieval strategies. In descriptions of language processing, models created have highlighted ways of memory functioning at different stages of incoming data-processing, thus focusing on short term memory (STM) and long term memory (LTM) and the relations between the two. Individual learner differences at the level of cognitive styles have also been investigated in relation to memory as a significant variable in learning and language production. The above mentioned areas of research are most frequently investigated by means of different retrospective techniques.

This paper is also concerned with language processing and the ways STM is activated but on this occasion simultaneously with exposure to a language task to be performed. The main objective of the paper is to demonstrate how simultaneous introspection as a research method can inform us about these processes and the functioning of memory. It looks at selected research projects using simultaneous introspection and think-aloud protocols (TAP-s) as ways of collecting and analysing data in a multilingual context of language production. The aim of the paper is not however to discuss the findings and to answer the research questions posed in my own research over the period 1993-2005 but to show the general validity of TAP data in studying memory and language activation processes.

KEY-WORDS: short term memory, on-line processing, thinking aloud protocols, second language acquisition, language activation.

1 – Introspection as verbalised thoughts

1. 1 – *Introspective methods and their characteristics*

1.1.1 – *Description*

The use of introspective methods came into being with the challenge to the hegemony of behaviouristic theory, with its *STIMULUS – RESPONSE* formula seeking to explain the mechanisms and structure of cognitive processes of human mind. The inadequacy of S – R explanations led to the feeling of dissatisfaction among scientists and made them look for methods that would enable them to probe the subjects' internal states. Data obtained in those experiments have become fundamental in psychology, although there are still many doubts raised about “the new methods” of introspection, about their being highly speculative and subjective (Gabryś1995).

To be able to see the validity of these method it is necessary to describe the encoding process which transforms behaviour into data and then back from data into behaviour, indeed to describe cognitive processes of human mind. Ericsson & Simon (1984) claim that:

Verbal behaviour is to be accounted for in the same way as any other behaviour, that is, by developing and testing an information-process model of how information is accessed and verbalized in response to stimuli.

Each verbalization comes from the cognitive process that underlies it. Consequently, verbalization must comply with all the constrains that have been identified for cognitive processes. In turn, cognitive processes consist of a set of sub-processes, which follow one another and are being transformed under the influence of a series of information processes. Information is stored in short term memory (STM - with a limiting capacity of about 15-20 seconds) and in long term memory (LTM - with permanent storage but slow access time). Information just received is stored in STM and is easily retrieved and articulated at the time of performing a task.

Verbalization processes of which all introspective methods make use, reflect cognitive processes in two ways, either:

a. directly, i.e. the time of the task performance is concurrent with the verbalization

or

b. indirectly, i.e. the information is retrieved from STM or LTM after the completion of the task.

In both cases verbalizations become functions of time.

Verbalization of cognitive processes in thinking aloud takes place on three distinct levels of thought processing. Ericsson & Simon (1984) describe them as:

1. the level of the vocalization, i.e. articulation of oral encodings, where no thinking processes take place. In self-directed verbalizations (e.g. in the case of thinking aloud protocols -TAPs) they are individual and depend on the subject's interpretation of the instruction s/he is given or on the semantic content of the task;
2. the level of description or explication of the content, which means "labelling" information and recoding it in an idiosyncratic way, characteristic of a subject/informant;
3. the level of explanation of thoughts and ideas that rush through the subject's mind or any other, even emotional, reactions to the information (task) s/he is to solve. It involves a process of interpretation. (Gabryś1995)

Level 1 verbalization is a *direct* process in which information encoded is vocalized (articulated), i. e., in a language task a phrase or a sentence is pronounced or read aloud by the subject with no cognitive processes taking place. *Level 2* and *level 3* are encoded verbalizations where the information attended to by the subject (level 1) is modified by recording processes, i.e., the subject generates a verbal representation of the information s/he has got stored in his/her mind by means of filtering it for the purposes of the task.

On *level 3* the subject reports his/her ongoing thinking, embracing not only its verbal aspects, i.e. word associations and interferences

from L1 and L3 (for example) but also personal, emotional responses to the task.

The way one (a subject) verbalizes, i.e., the extent to which the above mentioned levels occur in the process of verbalization may be influenced by *the type of instruction* given to the subject. Here are some examples of instructions in various experiments, starting with the first TAP (thinking aloud protocol; all quotations come from Ericsson & Simon 1984):

“Try to think aloud. I guess you often do so when you are alone and working on a problem”.

“Think, reason in a loud voice, tell me everything that passes through your head during your work searching for the solution to the problem”.

To be able to receive a complete report of a thinking process, other types of information are essential, such as are requested in the following instructions:

“The chief thing is to talk aloud constantly from the minute I present a picture, for I want to get everything you happen to think of no matter how irrelevant it may seem”.

“I am not primarily interested in your final solution, still less in your reaction time but in your thinking behaviour, in all your attempts, in whatever comes to your mind, no matter whether it is a good idea or less good idea or a question. Be bold! I do not count your wrong attempts, therefore speak them all out.”

“Do not plan what to say or speak after the thought, but rather let your thoughts speak, as though you were really thinking out loud”.

“In order to follow your thoughts we ask you to think aloud, explaining each step as thoroughly as you can”.

“Do not try to explain everything to anyone else. Pretend there is no one here but yourself. Do not tell about solutions but solve it.”

Each of the instructions, apart from their common intention to make the subject verbalize his or her thoughts, emphasizes a different aspect of the verbalization process. These instructions fulfil the role of a monitor in performing the tasks (Gabryś 1995).

1.1. 2 – *Introspective methods in psychology*

Man has always been interested in and speculated about the nature of the mind and thinking processes. As early as ancient times, philosophers and priests questioned the nature of the cognitive processes going on in the human mind. Pure observation used to be the only source of knowledge – and mostly, self-observation. However, observation has not been and often is still not considered to be a pure scientific method, even when it describes simple natural phenomena (to say nothing of the human mind). Though with time, controlled observation carried out in laboratory conditions or in various research contexts, and described in terms of strictly controlled variables, has become a scientific method capable of offering “indisputable evidence” (Ericsson & Simon 1984). The pioneers of introspective methods emerged at the turn of the 20th century. Most of them used the method to investigate the contents of consciousness (James, Wundt, Doners), reducing it to sensory and imaginal components reported by the subjects. The number of studies undertaken was relatively small. The first *thinking aloud protocols* (TAPs) produced in psychological experiments were simple descriptions of what had been said by the subjects and not their actual verbalized thoughts. As a consequence, they were very selective, biased and interpretative. The first protocols were used by Watson (1920) to illustrate thinking processes involved in a problem-solving situation. They were very unsystematic and full of interpretative remarks (Gabryś 1995).

2 – Introspective methods in second language acquisition research

As noted in Gabryś (1995), research in SLA dealing with error analysis over the past decade reflects a shift of emphasis from the product (i.e., an incorrect language form) to the process which underlies generation of the latter, as well as of correct forms. **Introspective reports**, described as an empirical tool of measurement of human behaviour long employed in psychology, have gained recognition as a mode of research in SLA only recently. Cognitivism in L2 /FL teaching and learning allowed the introduction of introspection as a valid and reliable method of research. It offers a variety of procedures for data

collection that can be classified according to certain pre-established criteria (Faerch & Kasper 1987).

2.1 - The criteria for classification of data collection

Faerch & Kasper (1987), the most ardent advocates of introspective methods in SLA research, propose the following criteria:

1. Object of introspection (according to the type of information gathered):
 - a. cognitive, affective or social aspects
 - b. declarative versus procedural knowledge
 - c. modality of language use (spoken vs. written, receptive vs. productive, combination of the above, e.g. in translation)
 - d. continuity of the process vs. a concrete aspect.
2. Related or unrelated to a concrete action.
3. Temporal relation to action (simultaneous, immediately consecutive, delayed consecutive).
4. Informant training.
5. Elicitation procedure selected:
 - a. degree of structure
 - b. media support
 - c. self-initiated v. other-initiated
 - d. degree (presence or absence) of interaction between the informant and experimenter or between informants themselves
 - e. integration with action (e.g. recorded discussions).
6. Combination of methods.

2.2 – Characteristics of data and focus of research

The object of introspection may refer to various aspects of language production the experimenter wishes to investigate. His/her interest may lie in a cognitive structure of the IL (interlanguage) of the informants. In other cases, when for example L2 achievement is to

be measured, an affective aspect of a language performance may be in focus. In the case of measuring affective aspects involved in L2 learning such as motivation and attitudes, introspection becomes a complementary tool to all kinds of interviews and questionnaires administered.

When performing certain language tasks *two types of knowledge* are being activated by the subjects, that is, *declarative knowledge*, referring to the knowledge of L2 rules (competence), and *procedural knowledge*, referring to the actual production (performance). Introspective methods and especially TAPs make it possible to follow the processes which are being activated, be they either in language learning and the development of receptive or productive skills. However, it is important to note that only *conscious thinking processes* can be stored and verbalized in STM (LTM in the long run). It is obvious that some of the cognitive processes (automatic ones) are lost in TAPs. But in most language tasks, subjects are faced with certain linguistic problems to be solved, so the thinking processes operate on a conscious level, i.e., the solutions do not come automatically but require all sorts of strategies (for example, recall strategies) to be used by the subjects.

The first studies in SLA carried by means of introspection in the early 70's referred to investigating the *competence of native speakers* (i.e. their declarative knowledge) concerning their intuitions about L1 (their mother tongue). They consisted of judgement tests and interviews recorded and transcribed. At the beginning of the 80's, TAPs found their way into L2 research of reception and production (Raabe 1982; 1986; Faerch & Kasper 1986; Schneider & Zimmermann 1986; 1987).

Introspective thinking aloud may focus on either the *overall approach* of the subjects towards the task, in which the whole process of task fulfillment is verbalized - or on a *specific aspect* of the task under consideration, which involves question-asking and interference on the part of the experimenter. The latter is often used in the cases of lexical search performed by the subjects. A part of SLA research, namely that dealing with motivation and establishing profiles of learners, focuses on the description of a certain action or behaviour the subject is involved in when performing the task.

Temporal relationship to action in thinking aloud can either be:

- a. simultaneous - subjects verbalize at the moment of performance;
- b. consecutive - verbalization comes immediately after the fulfilment of the task;
- c. delayed consecutive - there is a time span between the performance and verbalization.

The first type – *simultaneous verbalization* – is used in such tasks as listening and reading comprehension, writing or translation. The phenomenon of speech (communication strategies) can best be investigated by means of immediately *consecutive reporting* (STM is activated here), while *delayed introspection (retrospection)* makes use of diary studies or direct retrospective interviews and discussions.

Another important aspect of the methodology of introspective methods is the subjects' *training in verbalizing* cognitive processes. Although it was claimed by Ericsson & Simon (1984) that lack of training does not influence the reports, it nevertheless influences their completeness. Subjects that have been exposed to the recordings of thinking aloud become more conscious of what verbalization means and what aspects of cognitive processes and which emotional stages are taken into consideration. In addition, such exposure definitely reduces the stress and inhibition caused by the novelty of the method of data elicitation, which for some subjects may seem daunting.

The elicitation procedures utilized differ according to *the degree of structure* imposed on the informants (subjects) by the experimenter. They vary from multiple choice questions (e.g. judgment tests), open – ended questions (eg. in interviews and questionnaires) to “free” verbalization (e.g. in translation or cloze tasks), in the case of which subjects themselves make choices or often are straightforwardly being told to verbalize continuously everything that comes into their mind.

In the case of introspection, informants are often exposed again to the stimulus with which they had performed their task for the first time, that is, an original text or a picture story they had been previously asked to reflect on – with the purpose of activating their LTM. However, some retrospective studies do without recall support.

The data elicitation sometimes involves *interaction between the experimenter and the subjects or between the subjects themselves*. In such cases, interpretation of the data collected becomes much more complex as extra variables such as group structure and its dynamics, personal characteristics and other social and psychological factors influence it (Gabryś 1995).

2.3 – Introspection in translation tasks

One of the SL teaching and learning areas most open to introspective methods of investigation is *translation*, since it is a process which for the most part takes place on the level of consciousness which means it is accessible to verbalization.

Krings (1986) argues that the objections raised against concurrent introspective verbal reports are valuable sources of data, such as:

- the inaccessibility of certain cognitive processes for verbalization,
- the inconsistency of the actual behaviour and verbalization,
- the interfering character of verbalization, i.e., influence on the performance of an informant,
- the incompleteness of reports become irrelevant and inapplicable in the case of translation.

Compared with retrospective methods, *introspection* utilizes information stored by informants in their STM, i.e., information which is still accessible at the time of verbalization. Besides, processes verbalized do not require on the part of the informant any kind of “selection’ or “interference’ or speculation on the language sample, but straightforward information on what he or she is actually doing at the very moment and whatever thoughts pass through his or her mind.

The process of translation embraces *two aspects of language performance*: automated processes on the unconscious level and non-automated processes on the conscious level of processing. Verbal reports make it possible to distinguish between the two aspects. When verbalizing on the conscious level no distortion or change of the structure of language form received takes place, while translation, unlike non-verbal tasks, involves the use of a verbal code, so no extra

encoding is necessary. It is a natural process, and so Ericsson & Simon (1984) call it “level 1 verbalization” (see above).

It is often argued by those who criticize verbal reporting that verbal reports present *incomplete data*. It is obvious that a full report of all the mental (cognitive) processes is indeed not possible, taking into consideration the fact that a part of them is automated, i.e. unconscious. But accompanied by other experimental data, verbal reports come closer to a direct presentation (interpretation) of cognitive models than any other method.

Krings (1986) and other researchers who used translation tasks in collecting verbal report data all stress a low degree of structure of the reports, and Krings (1986) argues that “it is necessary to develop analytical categories step by step and to refine them gradually”. In the case of a translation task, *the basic categories* are those of *translation problems* (the first category) encountered by the informants and *strategies* (the second category) that are used to solve them. Translation problems can be caused either by comprehension problems at one end and production ones at the other.

For *the first category* (translation problems), further differentiation at the level of features can be made for both translation problems and strategies employed by the informants. It may include:

1. a number of problems encountered (for the purposes of comparison of the informants)
2. identification of difficulty areas (often very individual, showing idiosyncratic use of both L1 and L2)
3. successful (correct) solutions of translation difficulties versus unsuccessful (incorrect) ones.

The second category focuses on retrieval strategies, i.e. ways in which informants perform lexical search. Retrieval diagrams have been drawn up to demonstrate step by step strategies verbalized in TAPs. It is especially in this category that thinking aloud gives a lot of information, compared with pure performance data (product) analysis of a translation handed in.

These categories refer to structuring data received in TAPs that can be identified in verbal protocols. Apart from the two mentioned

above, other types of information can be gathered, for example, the informant's attitude towards the solutions of the problems he or she has arrived at - ranging from full acceptability to full dissatisfaction, or justification for certain choices made. Krings emphasizes the importance of *minimal interference* by the experimenter, i.e. no pressure to verbalize should be put on the informant since it could cause certain distortions of the naturalness of the translation process. Data received in TAPs can easily be *complemented* and *validated* by means of other methods, such as word-association tests and judgement tasks.

The results received in verbal reports that are based on translation can obviously be used for *a description of translation processes*, but at the same time they give a valuable insight into retrieval strategies, production problems and make it possible to identify areas of difficulty encountered by the subjects at various level of language advancement (Gabryś 1995).

3 - Introspective methods in second language learning research (examples)

Table 1 presents examples of studies of various aspects of foreign/second language performance when the whole variety of introspective methods and their combinations was deployed.

TABLE 1. Introspective methods in SLA studies (based on Gabryś-Barker 2005)

Research focus:	Studies:
Learning strategies	Naiman 1975, Stevick 1981, Wenden 1982, Cohen 1984
Reading in a foreign language	Cavalcanti 1982, Serrano 1984
Translation	Gerloff 1986, Krings 1986, Schneider & Zimmermann 1986, Schneider & Zimmermann 1987
Lexical search	Schneider & Zimmermann 1986, Haastrup 1987, Williams & Hammarberg 1998, Herwig 2001
Spoken language	Schwartz & Flower 1980, Dechert, Mohle & Raupach 1980/1984, Cohen & Apeh 1981, Faerch & Kasper 1987
Writing in a foreign language	Hayes 1983, Raimes 1985
Language transfer	Dewaele 2001, Jessner 2003

The whole variety of learners' behaviours can be observed on the basis of introspective verbal reports:

- a. the way learners attend to language input
- b. the way they arrive at spoken utterances (speech production)
- c. the way the text is being processed (reading comprehension, the reading process in L2 itself)
- d. the way the text is being generated (the writing process)
- e. the way lexical items are learnt at the stage of input
- f. the way lexical items are retrieved from memory.

The presentation in Table 1 is by no means exhaustive. The studies quoted above are the earliest, the most recent and the best-known ones in SLA research. It is worth mentioning that research in SLA by means of introspective methods has become more and more influential, especially as it gained the status of a scientific research tool not so long ago.

To exemplify the use of the theoretical framework of introspective methods in research, both in psychology and SLA, I would like to describe the way in which I incorporated **concurrent introspection** to advance my studies on multilinguality, or more precisely on the mental lexicon of multilingual language users (Gabryś-Barker 2005).

4 – Multilinguality researched by introspection

I would like to present here several projects I conducted with the aid of simultaneous introspection, in which the complexity of multilingual language interaction in the mental lexicon was observed from various perspectives (Table 2)

TABLE 2. Simultaneous introspection in SLA studies (Gabryś/Gabryś-Barker 1993-2005)

Study:	Type of language tasks used:	Research focus:	Year:
1.	Cloze task (in L2 English) Translation from L1 (Polish) into L2 (English)	* retrieval strategies of restricted collocations * metaphoric interpretations * psychotypology * learner profiles	1993
2.	Translation from L2 (English) into L3 (Italian) (English/Italian)	* language transfer at the level of lexis and syntax * transfer of training * language competence in language transfer	1999a
3.	Translation from L1 (Polish/Portuguese) into L2 (English) Translation from L2 (English) into L3 (German)	* lexical search processes * acquisition versus learning * transfer of training	1999b
4.	Translation (as above)	* the influence of the language of input (L1 versus L2) on the TL output * implicit versus explicit processing * lexical transfer	2005
5.	Translation (as above)	* inner/private speech * activation of individual languages * levels of metalinguistic awareness	2005

The transcripts of the verbalizations in the form of TAPs demonstrate the ways multilingual language users access their mental lexicon(s) and what processing mechanisms can be observed. The data stressing the importance of individual factors in multilingual processing, such as for example learning history or transfer of training, are re-examined in studies 1, 2 and 3 (Table 2).

In study 4, different approaches to the language task are presented, depending on the language of input (L1 or L2). The data exemplify different types of retrieval strategy used by the subjects in their verbalizations. The incorrect lexical solutions are observed, classified and discussed. Study 5, on the other hand, focuses mainly on the

language of thought and language activation/inhibition in different types of comments produced by the subjects when performing thinking aloud during a multi-language translation tasks. (Table 2)

Having analysed the results of the above studies, it became apparent that the phenomenon of multilinguality and in this case multilingual lexical competence, is very much determined by the language awareness of the subjects - in general, and by lexical awareness in particular. Additional survey studies were carried out to determine the subjects' perceptions of the phenomenon and their ability to use their linguistic knowledge in the practice of multilingual lexical development. It was assumed by me that this form of data collection (questionnaires) would exhaustively and explicitly show the subjects' language awareness.

As I say in the concluding chapter of my book (Gabryś-Barker 2005), these studies investigated the enormous complexity of multilinguality, and specifically lexical multilinguality – not only how words and phrases are stored but also how they are accessed and retrieved from memory. The observations on processes involved in the tasks of lexical retrieval (search) were recorded and discussed on the basis of thinking aloud protocols (simultaneous introspection) and retrospective comments by the subjects. They demonstrated that lexical search processes are not only lexical in nature; they also involve syntactic and phonological processing as important elements in lexical competence. TAP data also allowed me to formulate possible explanations for types of relations between languages that were observed in language processing recorded and retrieved from ST memory. They are:

- an L3 acquisition mechanism is different from that of L1 – learning L3 involves a similar mechanism to be activated as is the case with L2 (learning and not acquisition),
- activation of the mechanisms leads to activation of the language itself,
- suppression of L1 as “*non-foreign*” and recourse to another *foreign language* i.e. L2 as a learning/processing strategy.

5 – Sample data: A learner profile

The following examples come from the study conducted in 1995 (Gabryś 1995), which aimed at demonstrating the type of data accessible by means of simultaneous introspection.

TAP 1:

- *ah - I'll read it first - przeczytam (czyta)* When Albert I've won the big prize - *kurcze nie wiem - hm hm - czasy - (czyta)* When Albert entered the office- (*pisze*)Quando Alberto - *enter -entrare- entered - entre - o - r - a - entrato in - past definite - czyli simple past - czyli - passato prossimo - entered - (pisze)entrare - entrato in - l'oficio - l'uficio - nel'oficio - (pisze) nel'oficio - double f - (powtarza) officio-officio-officio.....*

Observations on language processing:

- attendance to input: holistic versus fragmentary
- the use of metalanguage and consciously acquired rules
- automatic recall of nominal phrases
- importance of accuracy in grammar and spelling

TAP 2:

(*reads*) Nobody was working - *to będzie czas ciągly -(powtarza) nobody was working- czas ciągly czyli imperfetto czyli a -work - lavoro- (pisze) nessun - tylko czy to będzie tak jak po angielsku - was working - working - liczba pojedyncza - nikt nie pracował - (pisze) nessun - - nie wiem czy to będzie podwojna negation - bo jak w polskim - (pisze) nessun lavorava - napiszę- a tutti impiegati.*

Observations on language processing:

- examples of transfer between L1, L2 and L3
- focus on grammatical accuracy
- verbalisation exclusively in L1

TAP 3:

(*czyta*) ... were still talking about the lottery - *and - znowu będzie imperfetto- to zreszta podają poniżej - parlare di - parlavano - (pisze)*

impiegati parlavano - vamo - di lottery - lotterija - to bedzie po hiszpańsku - no to ja muszę sprawdzić w słowniku - loteria - parlavano di lotteria - czyli będzie (pisze) della lotteria - della lotteria.

Observations on language processing:

- grammar focus (again)
- use of grammatical rules
- references to Spanish (L4): perceived language distance

TAP 4:

(czyta) Jules - jak to się będzie wymawiało - po włosku - Jules - nie - (czyta) his best friend - I - il - migliore - miglior - (pisze) miglior amico - il suo miglior amico - a jeszcze zapomniałam - (czyta) impiegati parlavano - still still - to jest - nie - ancora - nie pamiętam - muszę sprawdzić - still ancora - nie wiem czemu mi się skojarzyło z już - yet - still - impiegati ancora parlavano - nie wiem czy będzie przed parlavano ...

Observations on language processing:

- focus on word order (L2/L3 language transfer)
- need of assistance (dictionary) in lexical search

To sum up, this data can lead to establishing a learner's profile on the level of:

1. language competence in:

- * attendance to input
- * crosslinguistic consultations (L1, L2, L3)
- * memory and cognitive strategies of recall
- * perceptual styles
- * sources of language interference and typology of errors
- * strong and weak points of a learner
- * degree of internalization of rules and automatization of recall
- * transfer of learning and transfer of training

2. on the affective level:

- * attitude to the language task
- * motivation to perform
- * autonomy in language production

6 – The final comment

The success of verbalization processes, and as a consequence the type of data retrieved, is very much determined by the initial training of the subjects. It entails exposure to the introspective methods and thinking aloud materials. In many cases warm-up exercises are used to attune the subjects to the experiment or to give them some initial practice. Such exercises allow the researcher to intervene to help the subjects in their verbalizations by stopping them from lapsing into silence or by monitoring them initially. However, the extent to which pre-training is given should depend on the informants and their personal characteristics and on the nature of the task to be performed.

REFERENCES

- Bursch, L. 1982. *Fremdsprachenstudien – Frauenstudien*. Tübingen: Stauffenberg.
- Cavalcanti, M. 1982. Using the unorthodox, unreasonable verbal protocol technique. In: S. Dingwale; S. Mann; F. Katamba (Eds.). *Methods and Problems in Doing Applied Linguistic Research*. Lancaster: University of Lancaster, 72-85.
- Cohen, A.; Robbins, M. 1976. Towards assessing interlanguage performance. *Language Learning*. **26**: 45-66.
- Cohen, A.; Aphex, E. 1981. Easifying second language learning. *Studies in Second Language Learning*. **3**: 221-236.
- Cohen, A. 1984. Studying second language strategies. *Applied Linguistics*. **5**: 101-112.
- Dechert, H.; Raupach, M.; Mohle, D. (Eds., 1984). *Second Language Production*. Tübingen: Narr.
- Dewaele, J. M. 2001. Activation or inhibition? The interaction of L1, L2 and L3 on the Language Mode Continuum. In: J. Cenoz; B. Hufeisen; U. Jessner (Eds.). *Cross-linguistic influence in third language acquisition: psycholinguistic perspectives*. Clevedon: Multilingual Matters, 69-89
- Ericsson, K.; Simon, M. 1984. *Protocol Analysis: Verbal Reports as Data*. Cambridge MA: The MIT Press.
- Faerch, K.; Kasper, G. (Eds., 1987). *Introspection in Second language Research*. Philadelphia: Multilingual Matters.
- Gabryś, D. 1993. *Retrieval of Conventional Syntagms in English: Mechanisms and Effects*. PhD Dissertation. Katowice: University of Silesia (unpublished).
- Gabryś, D. 1995. Introspection in second language learning research. *Kwartalnik Neofilologiczny*. **3/95**: 271-291.
- Gabryś, D. 1999a. The phenomenon of transfer in L3 learning from a psycholinguistic

- perspective: a case study. In: M. G. Pinto; J. Veloso; B. Maia (Eds.). *Psycholinguistics on the threshold of the year 2000. Proceedings of the 5.th International Congress of the International Society of Applied Psycholinguistics*. Porto: Faculdade de Letras da Universidade do Porto, 409-413.
- Gabryś, D. 1999b. Cross-linguistic influences in L3-learning. In: M. Wysocka (Ed.). *On language theory and practice*. Katowice: University of Silesia Press, 169-182.
- Gabryś-Barker, D. 2005. *Aspects of multilingual storage, processing and retrieval*. Katowice: University of Silesia Press.
- Gerloff, P. 1986. Second language learners' reports on the interpretive process: think aloud protocols in translation. In: J. House; S. Blum-Kulka (Eds.). *Interlingual and intercultural communication*. Tübingen: Narr, 77-82.
- Haastруп, K. 1985. Lexical inferencing. A study of procedures in reception. *Scandinavian Working Papers in Bilingualism*. 5: 63-86.
- Hayes, J. R.; Flower, L. 1983. Uncovering cognitive processes in writing. In: P. Rosenthal; L. Tamor; S. A. Walmsley (Eds.). *Research on Writing: Principles and Methods*. New York: Longman, 207-220.
- Herwig, A. 2001. Plurilingual lexical organization : evidence from lexical processing in L1-L2-L3_L4 translation. In: J. Cenoz; B. Hufeisen; U. Jessner (Eds.). *Cross-linguistic influence in third language acquisition: psycholinguistic perspectives*. Clevedon: Multilingual Matters, 115-137.
- Jessner, U. 2001. The nature of cross-linguistic interaction in the multilingual system. In: J. Cenoz; B. Hufeisen; U. Jessner (Eds.). *Cross-linguistic influence in third language acquisition: psycholinguistic perspectives*. Clevedon: Multilingual Matters, 45-55.
- Kasper, G. (Ed.). 1986. *Learning, Teaching and Communication in the Foreign Language Classroom*. Aarhus: Aarhus University Press.
- Krings, H. 1986. The Use of Introspective Data in Translation. In: K. Fearch; G. Kasper (Eds., 1987). *Introspection in Second language Research*. Philadelphia: Multilingual Matters, 159-176.
- Naiman, N.; Fröhlich, M.; Stern, H. H.; Todesco, A. 1975. *The Good Language Learner*. Toronto: Modern Languages Centre/Ontario Institute for Studies of Education.
- Nisbett, R.; Wilson, T. D. 1977. Telling more than we can know: verbal reports on mental processes. *Psychological Review*. 86: 231-259.
- Raabe, M. 1982. Ist ne ... pas denn keine doppelte verneimung?. In: C. Gnetzmann; D. Stark (Eds.). *Grammatikunterricht*. Tübingen: Narr, 61-100.
- Raabe, M. 1986. The influence of L1 and L3 in the foreign language classroom. In: G. Kasper (Ed.). *Learning, Teaching and Communication in the Foreign Language Classroom*. Aarhus: Aarhus University Press, 59-69.
- Raimes, A. 1985. *An investigation how ESL students write*. New York: CUNY (Dept. of English, Hunter College).
- Schwartz, J., 1980. The negotiation for meaning; repair in conversation between SL learners of English. In: D. Larsen-Freeman (Ed.). *Discourse Analysis in Second Language Acquisition*. Rowley MA: Newbury House.

- Schneider, K.; Zimmermann, R. 1986. The Collective Learner Tested: Retrospective Evidence for a Model of Lexical Search. In: K. Faerch; G. Kasper (Eds.). *Introspection in Second language Research*. Philadelphia: Multilingual Matters, 177-196.
- Schneider, K.; Zimmermann, R. 1987. Dialogical aspects of individual lexical search. *Multilingua*: **6(2)**: 113-130.
- Serrano, N. S. 1984. Patterns of reading in L1 and L2. In: D. M. Singleton; D. G. Little (Eds.). *Language Learning in Formal and Informal Contexts*. Dublin: Irish Association of Applied Linguistics, 165-176.
- Stevick, E. 1981. Learning a Foreign Language; the Natural Ways. In: M. Hines; W. Rutherford (Eds.). *On TESOL '81*. Washington DC: TESOL, 1-10.
- Watson, J. B. 1920. Is thinking merely the action of language mechanisms? *British Journal of Psychology*. **11**: 87-104.
- Wenden, A. 1982. *Learner training for L2 Learners: A Selected Review of Content and Method*. New York: CUNY.
- Williams, S.; Hammarberg, B. 1998. Language switches in L3 production: implications for a polyglot speaking model. *Applied Psycholinguistics*. **19**: 295-333.

É morfologia?*

Maria Carlota Rosa

carlota@ufrj.br

Universidade Federal do Rio de Janeiro (Brasil)

RESUMO. O artigo focaliza as duas características que levam a questionar se processos que empregam a supressão de material das palavras-base estão ou não no âmbito da morfologia: a não-universalidade e a intencionalidade. Os dois processos com tais características arrolados na literatura lingüística são o *cruzamento vocabular* ou *amálgama* e a *acronímia* ou *siglação*.

A não-universalidade desses processos deriva da necessidade de a língua contar com um sistema de escrita fonológico para que possam surgir. A intencionalidade está ligada à busca de eufonia. Ambas as características levam diversos autores a retirar tais processos do campo da morfologia, ou a deixar em aberto a questão, ou ainda sugerir que são morfologia, mas improdutiva.

Por mais freqüentes que se venham tornando as siglas e os cruzamentos em diferentes línguas, entre elas o português, são processos que não produzem formações cujo significado seja composicional. Sua estrutura é opaca.

Levada em conta, porém, uma definição de morfologia como representação do sistema mental envolvido na formação de palavras (Aronoff & Fudeman, 2005: 239), não há como negar que a probabilidade de uma sigla, especialmente um alfabetismo, ser aceita em português é alta, o que coloca praticamente qualquer seqüência assim formada no campo das palavras potenciais. Nessa perspectiva, é estranho não aceitar tais processos na morfologia.

PALAVRAS-CHAVE. *formação de palavras e criação vocabular – processo não-universal – processo intencional - morfologia- acronímia ou siglação – cruzamento vocabular ou amálgama*

ABSTRACT. This article focuses on two aspects pointed out as attributes of blends and acronyms: they are not universal and they are intentional. Their non-universality derives from the necessity of a writing system. The intentionality is

Linguística - Revista de Estudos Linguísticos da Universidade do Porto - Vol. 4 - 2009, pp. 45-58.

* Versão preliminar deste trabalho apresentada na mesa-redonda "Padrões lexicais: as várias dimensões", no I Simpósio de Estudos Lexicais do Rio de Janeiro/ SEL-Rio, PUC-Rio, em 24 de janeiro de 2007. Agradeço ao Prof. Edson Watanabe (UFRJ) as informações sobre o japonês.

related to the search of euphony. Both aspects are in the origin of a discussion if blending and acroyimization are morphological processes.

If morphology is understood as the mental system involved in word formation (Aronoff & Fudeman, 2005: 239), we have to consider the high probability of a new acronym, specially if an initialism. The high probability makes almost any sequence of letters a potential word in Portuguese. On that perspective is rather strange to say that those processes fit badly as morphological processes.

KEYWORDS. *word formation and word creation – non-universal process – intentional process – morphology – acronyms – blends.*

1 – Introdução

Nos estudos tradicionais sobre o português, já se sugeria a possibilidade de algumas formações não resultarem de um processo morfológico. São observações rápidas, em geral de cunho histórico, relativas à formação do léxico português. Estão nesse caso, por exemplo, Cunha & Cintra (1985: 95), que justificaram não arrolarem *-iça* como sufixo formador de substantivos com base em adjetivos, uma vez que “*justiça* não apresenta propriamente o sufixo *-iça*, porque a palavra é continuação do latim *justitia*” e, por conseguinte, não se tratava de uma palavra formada pelos processos à disposição de um falante do português.

Mais recentemente a questão começou a ganhar espaço na literatura linguística, possivelmente em razão de siglas e acrônimos terem-se tornado comuns em muitas sociedades nas últimas décadas do século XX e, em menor escala, o cruzamento vocabular ou amálgama. Sem procurar fazer uma revisão exaustiva da literatura linguística, a lista de processos de criação vocabular situados fora do âmbito da morfologia varia de autor para autor, mas o cruzamento vocabular ou amálgama e a acronímia ou siglação são recorrentes. Muitos autores não reconhecem como morfológicos tais processos, que consideram marginais, marcados pela supressão de partes de palavras que serviram de base, termo que aqui designa as palavras a que os processos de acronímia ou cruzamento foram aplicados para gerar uma palavra complexa. Mas a questão parece incômoda e esses mecanismos podem ser classificados como não morfológicos ou extragramaticais (Beard 1998: 56; Haspelmath 2002: 25; Booij 2007: 20-23; Villalva 2008: 52-63), como morfologia improdutiva (Aronoff

& Anshen 1998: 246)¹, ou ainda como uma fronteira que não está claramente demarcada (Bauer 2003: 46)². Uma vez fora da morfologia, nem a acronímia, nem o cruzamento vocabular (nem a redução) teriam importância para a teoria morfológica³⁻⁴ (Spencer 1991: 461-462; Corbin 1987: 579).

Uma vez considerados fora da morfologia, tais processos não deveriam ser referidos como *formação*, mas como *criação vocabular* (Spencer 1991: 461-462; Booij 2007: 20) ou ainda como *invenção de palavras* (ingl. *word manufacturing*).

O objetivo deste artigo é discutir os argumentos que sustentam a retirada desse tipo de processos do âmbito da morfologia. Aqui, *morfologia*, como em Aronoff (1976), Aronoff & Anshen (1998) e Aronoff & Fudeman (2005), é o estudo e o sistema mental envolvido na formação de palavras e trata da estrutura interna das palavras complexas potenciais de uma língua. Uma palavra é potencial na medida em que poderia ser gerada por um dos processos de formação de palavras produtivos na língua. As palavras existentes seriam aquelas no léxico, em princípio porque: (a) não têm estrutura interna, como PEIXE ou PEDRA; ou (b) sua estrutura é apenas parcialmente reconhecida, caso de formações em *-ico* como HERMÉTICO, HÉCTICO; ou (c) seu significado não pode ser deduzido das partes componentes, como BOIOLA ('homossexual masculino').

As seções que se seguem situam, primeiramente, o problema. Em seguida, são focalizadas a não-universalidade e a intencionalidade, duas características atribuídas ao cruzamento vocabular e à acronímia, mas com o foco na formação de siglas. Segue-se a discussão. Por fim, algumas considerações encerram o texto.

¹ Aronoff & Anshen (1998: 246): "Some scholars have insisted that the study of morphological productivity should confine itself to the study of words that are produced unintentionally [...]. This rules out entirely the study of unproductive morphology, which resembles more marginal forms of word creation like the formation of blends [...] or acronyms [...], in being more likely to be intentional or noticed".

² Bauer (2003: 46): "They are thus not universal since not all languages are written. Neither do they clearly belong under the heading of morphology although they are included here for the sake of completeness".

³ Spencer (1991: 461-462): "None of these is of any great importance to morphological theory".

⁴ Corbin (1987: 579): "... faut-il l'inclure [*la siglaison*] dans le composant lexical, puisqu'il est possible de construire des dérivés de sigles (*cégétiste*), ou dans le composant phonologique, et permettre aux RCM [*règle de construction de mots – MCR*] d'avoir accès à la sortie de celui-ci?".

2 – Situando o problema

Em 1976, Mark Aronoff trazia a noção de produtividade para o primeiro plano no estudo da morfologia, num quadro teórico em que a morfologia constitui-se na parte da gramática referente à estrutura interna das palavras complexas potenciais e, por essa razão, parte da competência lingüística do falante.

O termo *produtividade* já fora empregado antes, por exemplo, em obras tão antigas quanto a *Gramática Histórica da Língua Portuguesa*, de M. Said Ali, revista e ampliada pelo autor em 1931 e, mesmo sem o emprego do termo, a produtividade surgia em obras como Góes (1938) ou Allen Jr. (1941), por exemplo, que podiam referir um sufixo como *-íssimo*, por exemplo, como o “principal sufixo formador de grau superlativo absoluto synthético” (Góes 1938: 108). Mas era uma noção pré-teórica: a produtividade não era tratada com mais que umas poucas observações assistemáticas.

Diferentemente desses autores mais antigos, Aronoff procurava discutir as “*propriedades muito específicas*” (Aronoff 1976: 35) que traçariam uma linha divisória entre processos produtivos e processos improdutos e, desse modo, retirar do termo *produtividade* o caráter intuitivo com que era empregado. A produtividade representa os processos gramaticais – e portanto regulares – ativos na formação de palavras novas numa língua. Tornava-se central porque “...*the simplest task of a morphology, the least we demand of it, is the enumeration of the class of possible words of a language*” (Aronoff 1976: 17-18).

Nessa obra já clássica, ao perguntar de que as novas palavras eram cunhadas⁵, Aronoff (1976: 20) comentava dois fenômenos que caracterizou como *oddities*, termo que poderia ser traduzido como

⁵ “*What Are New Words Coined From?*”, título da seção 2.2.3. E prosseguia: “*Oddities*. The italicized words [...] above are all rather unusual coinages. Those whose basis is not completely opaque are *blendings* [...] or [...] *portmanteau words*, formed by merging parts of words into a word which meets the phonotactic restrictions of the language. More transparent examples are *smog*, from *smoke* and *fog*, and *chunnel*, from *channel* and *tunnel*. A related type of coining is that of *letter words* and *syllable words*, collectively known as *acronyms*. [...] This type is almost unknown in the languages of the world and was uncommon in our own before this century. [...] It does, in any case, presuppose an alphabet. [...] The device is, in short, very unusual and certainly not a universal fact of language. These two devices form words which have no recognizable internal structure or constituents. This makes them opaque, and hence uncommon.”

‘esquisitices’. Reconhecia como “esquitos” dois tipos de formações opacas – uma vez que têm estrutura interna não recuperável a partir exclusivamente do conhecimento lingüístico do falante –, e que, por essa razão, segundo ele, eram incomuns. Um deles era o *cruzamento* ou *fusão vocabular* (ing. *portmanteau word* ou *blending*), que pode ser exemplificado em português em formações como SACOLÉ (de SACO e PICOLÉ, ‘sorvete em saquinho’) ou BRASIGUAIO (de BRASILEIRO e PARAGUAIO, definido no *Dicionário Houaiss* como ‘que tem características, elementos ou aspectos tanto brasileiros como paraguaios’). O outro eram os *acrônimos* (ing. *acronyms*), exemplificados em português por formações como INCA (para *Instituto Nacional do Câncer*) e formas por vezes denominadas *alfabetismos* ou *inicialismos*, como IPTU (para *Imposto Predial e Territorial Urbano*). *Acrônimo* é aqui sinônimo de *sigla*, uma vez que as pronúncias soletrada ou não que distinguiriam os dois tipos, a despeito de serem, em geral, previsíveis, como demonstra o algoritmo proposto por Barbosa, Rosa, Gonçalves & Resende Jr (2003), podem aplicar-se a uma mesma formação, como ilustram, no português brasileiro, ONG⁶ ([œni'ʒe] ou ['õŋgi]) e USB⁷ ([wesi'be] ou ['uʒbi]).

Bauer (2003:46), como Aronoff (1976:20), classificava ambos os tipos como formações baseadas em alfabeto e, assim sendo, não-universais. Na discussão acerca da produtividade e do papel de processos como acronímia e cruzamento, uma questão seria apontada como uma linha divisória entre esses e os processos gramaticais⁸⁻⁹: a intencionalidade, cuja expressão máxima é a busca pela eufonia.

⁶ Sigla para *organização não governamental*.

⁷ Sigla para *universal serial bus*.

⁸ Beard (1998:56): “Blending, acronymization, and analogical formation also tend to be conscious operations, unlike grammatical derivation. Words like *smog*, *motel*, and *tangelo* are created intentionally by a logical rather than grammatical process: if the reference is part A and part B, then the word referring to it should comprise parts of the words for A and B. Acronyms like *laser*, *scuba*, *aids*, have been converted from phrases to the initial letters of the words in those phrases, which are not part of grammar, then the initials have been phonologically interpreted. *Acquired immune deficiency syndrome*, for example, provides *aids*, which is rendered pronounceable by applying English spelling rules in reverse. The process hence requires considerable conscious activity outside the bounds of grammar. As in the case of clipping, the phrase and the acronym are synonymous, and both remain in the language”.

⁹ No tocante a cruzamentos, Aitchison (1994) defende que não são intencionais e que podem, mesmo, ser encontrados na fala de alguns afásicos.

A seção seguinte revê ambas as características apresentadas na literatura lingüística, tomando a formação de siglas como base para a discussão.

3 – Um processo incomum?

3.1 – A não-universalidade

A não-universalidade deriva da base gráfica das formações, mais evidente nos alfabetismos. Bauer (1983: 238) aponta como evidência para a base gráfica da acronímia o fato de que o valor fonético de cada formativo, em última análise uma letra, é fixado a partir do acrônimo, não da locução com base na qual foi formado. Um exemplo em português: CIMI (*Conselho Indigenista Missionário*), foneticamente [´siml], ganha para o < c > não a pronúncia como oclusiva velar surda [k] que o < c > representa em *Conselho*, mas a pronúncia que <c> tem diante de <i> em português, [s]; por sua vez, [t] passa a [i] e a receber apenas a nasalidade do contexto. O carácter ortográfico de siglas soletradas ou alfabetismos (SBT¹⁰, CNPq¹¹, APPACDM¹²) é evidente. É o *nomen* da letra que é pronunciado, o que torna possível dotar de pronúncia seqüências gráficas que, de outro modo, não representariam seqüências legais na língua, nem fonológica nem ortograficamente.

O carácter gráfico determina a pronúncia, isto é, a preferência por silabar ou soletrar a sigla deriva do número de formativos e da seqüência de consoantes e/ou vogais considerada. Uma sigla em português com dois formativos será soletrada: HU¹³, BO¹⁴; por outro lado, uma sigla de três formativos com a estrutura VCC será soletrada ou não na dependência de a primeira consoante na seqüência ser ou não <s>, <r> ou <l> e ser diferente da segunda: ACM¹⁵, USP¹⁶.

¹⁰ Sigla para *Sistema Brasileiro de Televisão*.

¹¹ Sigla para *Conselho Nacional de Desenvolvimento Científico e Tecnológico*.

¹² Sigla para *Associação Portuguesa de Pais e Amigos do Cidadão Deficiente Mental*.

¹³ Sigla para *hospital universitário*.

¹⁴ Sigla para *boletim de ocorrência*.

¹⁵ Sigla para *Associação Cristã de Moços* ou ainda para referência ao falecido político Antônio Carlos Magalhães (1927-2007).

¹⁶ Sigla para *Universidade de São Paulo*.

A não-universalidade apresenta-se sob três ângulos estreitamente relacionados. Primeiramente, quanto ao surgimento recente do processo. Era desconhecido antes do século XX, uma vez que não se confunde com abreviaturas (a esse respeito, ver Abreu 2004; 2009), e raro ainda em meados desse século. Aronoff (1976: 20) reporta o fato em relação ao inglês¹⁷. Sandmann (1988: 146)¹⁸ refere observação de Piel sobre o francês, quando este publica, em 1966, a edição revista e ampliada da *Gramática histórica da língua francesa: 2ª Parte – Formação de palavras*, de Meyer-Lübke. Cunha (1972: 130-131)¹⁹, por exemplo, demonstrava que as siglas também eram incomuns em português, mas começavam a ganhar espaço. Câmara Jr (1973: 50) limita o processo à formação de nomes próprios, e as locuções-base por ele arroladas eram “*títulos de obras, revistas e jornais em citação [...], partidos [...], departamentos, organizações e até nomes de países [...]*,” mas já apontava uma tendência de ampliação: “*mais recentemente nomes de autores, figuras políticas [...]; etc.*”

Seja como for, a raridade sentida nas décadas de 1960 e 1970 não pode ser invocada quase meio século depois, como demonstrou Abreu (2004; 2009) para o português, nem a relação de causalidade defendida por Aronoff (‘*opaque, and hence uncommon*’) pode ser sustentada. A opacidade é uma característica do processo, mas a ela não pode ser atribuída a causa de essas formações serem mais ou menos comuns — afirmação em que (*in*)*comum* se refere à frequência de formações desse tipo.

A não-universalidade pode ser interpretada não no âmbito da frequência de formações numa linha do tempo, mas no âmbito da ocorrência do processo nas línguas do mundo. As siglas pressupõem

¹⁷ Aronoff (1976: 20): “[t]his type is almost unknown in the languages of the world and was uncommon in our own before this century [...]. It does, in any case, presuppose an alphabet. At present it is most common in the official languages of the major imperialist powers. The device is, in short, very unusual and certainly not a universal fact of language”.

¹⁸ Sandmann (1988: 146): “PIEL (p. 175) acentua, além disso, como aliás já o tinha feito MEYER-LÜBKE (ib.), o caráter atual dessas formações: Às palavras formadas artificialmente com letras iniciais, que nos últimos decênios proliferaram sempre mais também na França (...)”.

¹⁹ Cunha (1972: 130-131 – *ênfase no original*): “Também moderno — e cada vez mais generalizado — é o processo de criação vocabular que consiste em reduzir longos títulos a meras **siglas** constituídas das letras iniciais das palavras que os compõem. [...] Uma vez criada e vulgarizada, a sigla passa a ser sentida como uma palavra primitiva, capaz, portanto, de formar derivados: arenista, emedebista, etc”.

um alfabeto, o que retiraria do processo qualquer possibilidade de universalidade.

A restrição do processo a línguas com sistema de escrita alfabético não parece ser tão extrema, ou melhor, parece poder ser estendida para línguas com sistemas de escrita fonológicos, o que inclui os sistemas de escrita silábicos. A formação de siglas não é desconhecida do japonês, por exemplo, língua cujo sistema de escrita não tem por base um alfabeto, mas um conjunto de caracteres e dois silábrios, e em que o processo parece ser particularmente produtivo com empréstimos: *modern girl* como *mo-ga*, ‘*dance parties*’ como *dan-pa*, ‘*potato chips*’ como *po-te-chi*, ‘*engine stop*’ como *en-su-to-pu*²⁰.

A famosa *Tokyo Kogyo Daigaku*, literalmente,

To -	kyo ²¹	Ko	gyo	Dai	gaku
leste	capital	engenharia	tecnologia	grande	escola

tem como sigla *ToKoDai*, normalmente representado pelo carácter de ‘leste’ (To), de ‘engenharia’ (Ko) e de ‘grande’ (Dai). São só três caracteres²². E o “interessante é que o japonês é sempre silábico e não existe a situação de termos consoantes seguidas”²³.

A impossibilidade de alfabetismos no chinês²⁴ não se mantém no tocante ao sistema romanizado de escrita do chinês conhecido como *pinyin*:

Since the promulgation of the (latin) alphabetic pinyin supplementary writing system, one sees alphabetic acronyms, some of them rather alarming and most of them quite dysphonic (if that's the opposite of euphonic). For instance, Guangzhou Foreign Language Institute could end up as GZWGYXXY for Guangzhou WaiGuo YuYan XueYuan [= WideState OutCountry SpeechTalk LearnYard]; or as GWYX for Guangzhou Waiguo Yuyan Xueyuan [= Guangzhou Foreign Language Institute] if they take the first letter of each disyllabic word rather than the first letter of each monosyllabic morpheme.

²⁰ LINGUIST List: Vol-4-502. 24 JUN 93 From: guy modica - Subject: Re: 4.494 Acronyms

²¹ Originalmente a capital do Japão era Kyoto, mais a oeste.

²² Edson Watanabe (c. p.).

²³ Edson Watanabe (c. p.).

²⁴ LINGUIST List: Vol-4-494. Wed 23 Jun 1993. ISSN: 1068-4875. Lines: 236 Subject: 4.494 Acronyms From: Joel Bradshaw - Subject: Re: 4.482 Qs: Acronyms

Anão-universalidade reflete-se ainda num terceiro aspecto: mesmo numa língua dotada de sistema de escrita fonológico, pressupõe um falante alfabetizado no mínimo medianamente, que deverá não só conhecer o *nomen* das letras, mas que já tenha desenvolvido, no tocante à leitura e à escrita, a competência para segmentar, excluir, substituir e adicionar letras.

Uma língua pode existir sem manifestação escrita e mesmo que conte com um sistema de escrita, ele não será necessariamente fonológico nem estará necessariamente disponível para todos.

3. 2 – A intencionalidade

Câmara Jr. (1973: 50) apontava como uma das tendências da acrografia o fato de essas formações “*associarem-se a raízes da língua parônimas ou homônimas, cuja significação a isso se presta*”. E exemplificava: “SAPS, para *Serviço de Alimentação da Previdência Social* associado à raiz *sap-* de *sapere* ‘saborear’. Mais recentemente, MERCOSUL (*Mercado Comum do Sul*) pode ser associado à raiz *merc-*, presente em MERCAR, MERCADO, MERCANTE, MERCANCIA. Outros exemplos: SciELO (*Scientific Electronic Library Online*), SIGA (*Sistema Integrado de Gestão Acadêmica*), SIGMA, PROHPOR (*Programa para a História da Língua Portuguesa*), CASA (*Comunidade Sul-Americana de Nações*). É um desiderato (nem sempre alcançado) fazer coincidir uma seqüência com uma raiz ou com uma palavra que de alguma maneira remeta o falante para uma unidade de som e significado, o que parece procurar compensar a opacidade de uma formação que não se constitui de raiz e afixos.

A tendência apontada por Câmara Jr. expressa o que intuitivamente um falante considera uma *boa sigla*. GRUDA, para *Grupo de Estudos de Doenças Afetivas*, por exemplo, é percebida como uma “boa sigla”, uma vez que evoca o ‘apego excessivo’, coloquialmente referido como *grude*. A sensação de *boa sigla* deriva de se ter conseguido criar algum tipo de redundância, de modo a associar a forma a um lexema. A associação não pode, porém, evocar uma relação indesejável. Assim, embora os alfabetismos sejam os tipos mais comuns de siglação (Abreu 2004), não se aceitaria reduzir *Conselho Universitário*, por exemplo, a um alfabetismo, pela carga pejorativa que carregaria.

Para criar a associação com outras palavras da língua, a acronímia permite, com razoável grau de liberdade, a manipulação das letras que serão levadas em conta para a formação, sendo possível combinar letras iniciais, sílabas, letras não iniciais e até mesmo seqüências cuja relação com a locução original não é perceptível. A liberdade de extrair da locução algumas letras em detrimento de outras ou mesmo de incluir letras supérfluas (*vide* Abreu 2004; 2009) exprime o quanto há de intencional nesse processo.

A relação entre as letras escolhidas como formativos da sigla e o significado é estabelecida a cada sigla e poderá nem mesmo ser recuperada quando do emprego da sigla. É muito rara a associação a um significado em razão da freqüência, mas, de qualquer modo, não se aplica a alfabetismos. A freqüência com que foram empregados no Brasil nas décadas de 1970 e 1980 *-bras* (para '*Brasil*' ou '*brasileiro*', e.g., PORTOBRÁS, TELEBRÁS, BANCOBRÁS, ELETROBRÁS, SIDERBRÁS, NUCLEBRÁS, RADIOBRÁS) e *embra-* (para '*empresa brasileira*', e.g., EMBRAER, EMBRAFILME, EMBRAMEC, EMBRAPA, EMBRATTEL, EMBRATER, EMBRATUR) fizeram com que esses elementos passassem a ter um significado associado, como um morfema, o que permitiu que informalmente, surgissem termos como ROUBOBRÁS e EMBRA-QUALQUER-COISA.

Por outro lado, a liberdade na manipulação do que será retido e do que será suprimido obscurece por vezes a fronteira entre a acronímia e outros processos, como a redução: PETROBRAS (*Petróleo Brasileiro S.A.*, atualmente *Petrobras Brasil*), por exemplo, seria efetivamente uma formação acronímica? Quanto da base teria de ser mantido, como questiona Cannon (1989), para que se possa ainda falar de sigla?

Em razão da opacidade, o uso freqüente pode levar uma sigla a ser reanalisada como uma palavra primitiva. Um exemplo recente no Brasil é *ibope*. De sigla para o *Instituto Brasileiro de Opinião Pública e Estatística*, passou a designar, como nome, 'bom índice de audiência entre telespectadores' ou ainda 'prestígio'. Com a reanálise, a sigla pode (a) servir de base para a formação de novas palavras (PEEMEDEBISTA, de PMDB, *Partido do Movimento Democrático Brasileiro*; PETISTA, de PT, *Partido dos Trabalhadores*; EMEESSETISTA, de MST, *Movimento dos Trabalhadores Rurais Sem Terra*), até mesmo de um cruzamento vocabular: PETELHO (de PT e PENTELHO, 'petista chato'); e

(b) perder a relação entre as letras e as palavras da base que representam, podendo ser empregada com alguma dessas palavras: *Sistema Siga* (de *Sistema Integrado de Gestão Acadêmica*), *protocolo IP* (*Internet Protocol*), *Banco BMG* (*Banco de Minas Gerais*), *Fundação Funasa* (*Fundação Nacional de Saúde*).

4 – Fora da morfologia?

Haspelmath (2002: 25) argumenta em favor de retirar da morfologia os processos que tomam a escrita por base porque, não obstante criem novas palavras, como as operações morfológicas fazem,

they do not fall under morphology, because the resulting new words do not show systematic meaning-sound resemblances as the sort that speakers would recognize. If we know that *CD* is an abbreviation of *compact disc*, we do not know it as a result of unconscious language acquisition, but because we were explicitly told so. Thus, not all processes of word-creation fall under word-formation, and abbreviations and clippings will play no role in this book. (Haspelmath 2002:25)

Booij (2007: 21), na mesma linha de argumentação, aponta que o uso da supressão leva a que seja impossível usar do conhecimento lingüístico para descobrir o significado da formação resultante. No entanto o mesmo argumento pode ser empregado para palavras que não se questiona se resultam de processos morfológicos. Não é possível, por exemplo, interpretar ADOÇANTE como um ‘substituto químico do açúcar de baixo valor calórico’ a partir de sua estrutura. Nem DECLARAÇÃO como um ‘documento entregue anualmente à Receita’.

É possível contra-argumentar que, mesmo sendo esses significados não previsíveis a partir da estrutura, ainda assim seria possível pressupor ‘algo que adoce’ e ‘algo relacionado a declarar’, o que seria impossível a partir de um *C* ou de um *A*. Fora de qualquer contexto isso é exato; no entanto, em contextos especializados, podemos esperar que *U* esteja por *universidade*, *F* por *Faculdade*, *A* por *associação*, *C* por *conselho* se ocorrerem numa sigla em posição inicial; ou *V* por *vírus* se em posição final (HIV, HPV, EBV)...

A noção de produtividade procura demonstrar os processos que estão à disposição de um falante. É possível estabelecer, por exem-

plo, que é praticamente nula a probabilidade de novo nome abstrato em *-or* (como FULGOR, ALBOR), não obstante já se ter notado que “quanto menos uma regra é produtiva, tanto mais forte é o efeito de seu desempenho lingüístico criativo” (Sandmann, 1992: 74). A probabilidade de novas siglas, especialmente de alfabetismos, é alta, o que fica demonstrado em criações e reanálises coloquiais com fins humorísticos, como ASPONE (*assessor de porra nenhuma*), PDI (*porra da idade*), TPM (*tendência para matar*), QI (*quem indica*) ou em linguagens de especialidades, como PB (*português do Brasil*), HO (*higiene ocupacional*)... Mas uma seqüência de nomes de letras não se torna uma probabilidade de formação para um analfabeto, não obstante a freqüência no quotidiano possa criar as condições para seu reconhecimento e emprego como uma palavra primitiva, caso de CD, DVD, por exemplo. Novamente se pode contra-argumentar que também há processos derivacionais que são produtivos entre falantes escolarizados: por exemplo, a formação de nomes abstratos em *(i/e)-dade* a partir de adjetivos.

Ambos os processos têm características particulares, mas a linha divisória não parece tão clara.

5 – Considerações finais

O distanciamento que formações baseadas em elementos da escrita apresentam em relação aos processos de formação de palavras que tomam por base raízes e /ou afixos permitem entender a discussão acerca de se estariam ou não na morfologia. Diferentes trabalhos vêm demonstrando a natureza desse distanciamento. Gonçalves (2006), focalizando o cruzamento vocabular, demonstra a importância da interface do processo com a fonologia, do mesmo modo que fizera Corbin (1987), em outro quadro teórico, em relação às siglas. Abreu (2009), com base em testes psicolingüísticos, demonstrou que, no tocante às siglas, os falantes distinguem as siglas das palavras e das não-palavras.

Mesmo que se leve em conta sua não-universalidade, ainda assim, para muitas línguas, a probabilidade de novas formações vem crescendo nas últimas décadas, o que implica que são mecanismos que fazem parte da competência lingüística. Sendo assim, é impor-

tante compreendê-los e, se não têm importância para uma teoria da morfologia, isso não pode significar que não têm interesse para estudo.

REFERÊNCIAS

- Abreu, K. N. M. de. 2004. *Um caso de morfologia improdutiva no português do Brasil: a formação de siglas e de acrônimos*. Dissertação de Mestrado em Lingüística. Rio de Janeiro: UFRJ/ Programa de Pós-Graduação em Lingüística.
- Abreu, K. N. M. de. 2009. *Um estudo sobre as siglas do português do Brasil*. Tese de Doutorado em Lingüística. Rio de Janeiro: UFRJ/ Programa de Pós-Graduação em Lingüística.
- Aitchison, J. 1994. *Words in the mind: an introduction to the mental lexicon*. 2nd. ed. Oxford: Blackwell.
- Ali, M. S. [1931] 1971. *Gramática histórica da língua portuguesa*. Rio de Janeiro: Acadêmica.
- Allen Jr., J. H. D. 1941. Portuguese word-formation with suffixes. *Language Dissertation n° 33*, 17(2). April-June, 1941. Baltimore: Linguistic Society of America.
- Aronoff, M. 1976. *Word formation in generative grammar*. Cambridge MA: The MIT Press.
- Aronoff, M.; Anshen, F. 1998. Morphology and the lexicon. In: A. Spencer; A. M. Zwicky (Eds.). *The handbook of morphology*. Oxford: Backwell, 235-247.
- Aronoff, M.; Fudeman, K. 2005. *What is morphology?* Oxford: Blackwell.
- Barbosa, F. L. de F.; Rosa, M. C.; Gonçalves, C. A. & Resende Jr., F. G. V. 2003. Algoritmo para leitura de siglas em um sintetizador de voz. *Anais do XX Simpósio Brasileiro de Telecomunicações*. Rio de Janeiro: IME/PUC-RJ, 672-675.
- Bauer, L. 1983. *English word-formation*. Cambridge: Cambridge University Press.
- Bauer, L. 2003. *Introducing linguistic morphology*. Washington DC: Georgetown University Press.
- Beard, R. 1998. Derivation. In: A. Spencer; A. Zwicky (Eds.). *The handbook of morphology*. Oxford: Blackwell, 44-65.
- Booij, G. 2007. *The grammar of words*. Oxford: Oxford University Press.
- Câmara Jr., J. M. 1973. *Dicionário de Filologia e Gramática referente à língua portuguesa*. 5^a ed. Rio de Janeiro: J. Ozon.
- Cannon, G. 1989. Abbreviations and Acronyms in English Word-Formation. *American Speech*. 64(2): 99-127.
- Corbin, D. 1987. *Morphologie dérivationnelle et structuration du lexique*. Tübingen: Niemeyer.
- Cunha, C. F. da. 1972. *Gramática da língua portuguesa*. 1^a ed. [Rio de Janeiro]: MEC/FENAME.

- Cunha, C.; Cintra, L. F. L. 1985. *Nova gramática do português contemporâneo*. Rio de Janeiro: Nova Fronteira.
- Góes, C. 1938. *Diccionario de affixos e desinencias*. Rio de Janeiro: Francisco Alves.
- Gonçalves, C. A. 2006. A ambimorfemia de cruzamentos vocabulares: uma abordagem por ranking de restrições. *Revista da ABRALIN*. **5(1-2)**: 169-183.
- Haspelmath, M. 2002. *Understanding Morphology*. London: Arnold.
- Houaiss, A.; Villar, M. S. 2001. *Dicionário Houaiss da língua portuguesa*. Rio de Janeiro: Objetiva.
- Sandman, A. J. 1988. *Formação de palavras no português brasileiro contemporâneo*. Curitiba: Scientia et Labor/Ícone.
- Sandman, A. J. 1992. *Morfologia lexical*. São Paulo: Contexto.
- Spencer, A. 1991. *Morphological Theory*. Oxford: Blackwell.
- Villalva, A. 2008. *Morfologia do português*. Lisboa: Universidade Aberta.

Observações sobre o uso dos afixos derivacionais em alguns crioulos de base lexical portuguesa e espanhola

Angela Bartens
abartens@mappi.helsinki.fi
Universidade de Helsínquia (Finlândia)

RESUMO. Na presente contribuição, examinamos a hipótese de McWhorter (1998; 2005) segundo a qual a morfologia derivacional das línguas crioulas seria transparente ou “composicional” com base em dados de seis crioulos de base ibero-românica. Mais precisamente, tratamos apenas da derivação afixal. Constatamos que, dependendo da coexistência com as línguas lexificadoras e/ou de adstrato, os seis crioulos estudados utilizam os afixos derivacionais de maneiras muito distintas. Contudo, na maioria dos casos, a derivação é semanticamente transparente. Existem, porém, em todos os crioulos estudados contra-exemplos ao postulado da transparência ou da composicionalidade semântica.

PALAVRAS-CHAVE:

Abstract. In the present study, we examine McWhorter’s (1998; 2005) hypothesis according to which the derivational morphology of creole languages is semantically transparent or “compositional”. We have drawn our data from six Iberoromance-based creoles. To be precise, in this paper we only deal with derivation by means of affixation. Depending on the coexistence of the creoles with the lexifier and/or adstrate languages, they turn out to employ their derivational affixes in quite distinct ways. In the majority of cases, the derivational processes are semantically transparent. However, all the creoles studied here also feature counter-evidence against the postulate of semantic transparency or compositionality.

KEY-WORDS:

1– Introdução

Sobretudo após a publicação dos trabalhos de McWhorter (1998; 2005)¹, assistiu-se ao reinício do debate sobre as questões seguintes:

Linguística - Revista de Estudos Linguísticos da Universidade do Porto - Vol. 4 - 2008, pp. 59-83.

¹ McWhorter (2005) é uma antologia de artigos anteriores actualizados. Por essa razão, não citamos os artigos individuais.

as línguas *pidgins* e crioulas têm morfologia flexional e derivacional e esta última é realmente “transparente” ou “composicional” como pretende McWhorter, ou não?

Integrando-se nos estudos que se inspiram na hipótese de McWhorter – que estipula que as línguas crioulas se podem distinguir por três critérios intralinguísticos: a morfologia flexional é virtualmente inexistente; os tons não são usados, ou são usados de forma muito restrita, para marcar distinções lexicais e morfossintáticas; e a derivação é sempre composicional (p. ex. McWhorter 2005: 10) –, a presente contribuição trata de apenas uma parte da derivação, a derivação afixal.² Num trabalho anterior (Bartens 2004), investigámos o redobro, cujas funções nas línguas crioulas podem ser flexionais ou derivacionais, segundo Kouwenberg & LaCharité (2003); deixámos de lado a conversão, processo altramente produtivo nos crioulos.

Por exemplo, Plag (2001; 2006), Siegel (2004) e Arends *et al.* (2006) questionam os postulados de McWhorter partindo de uma perspectiva mais lata. Todos concordam com a conclusão de que os postulados de McWhorter não se podem manter.³ No caso específico da morfologia derivacional, afirmam o seguinte: ainda que existam muitos casos de derivação transparente, há contra-exemplos em quantidade suficiente para refutar a hipótese. Enquanto McWhorter (1998: 798) sublinha que a opacidade semântica da derivação é sempre resultado da derivação semântica a longo prazo, Plag (2001: 155-156; 2006: 305) e Braun & Plag (2003) observam que, numa situação de contacto linguístico, é previsível que surja opacidade semântica por outros mecanismos, sobretudo o empréstimo. A conversão tam-

² Entendemos a derivação no sentido de Fries (2005) e Günther (2005): junção de afixos a uma base. No caso dos sufixos, é inequívoco que o sufixo constitui o núcleo da palavra derivada e por isso determina a respectiva classe gramatical. Muitas vezes, também determina a classe das bases às quais se pode juntar. A posição dos prefixos, pelo contrário, não é inequívoca. Fries (2005: 5) observa que alguns autores incluem a conversão na derivação. De qualquer modo, como notamos no texto, a conversão é um mecanismo de formação de palavras típico das línguas crioulas que não consideramos aqui precisamente por esse motivo. Em função da definição da derivação dada acima, também não consideramos as mudanças de classe gramatical que ocorrem quando a acentuação ou a estrutura tonal da palavra mudam, processos esses que têm um papel importante, por exemplo, no papiamentu.

³ Com efeito, Faraclas (2003: 287) chega à conclusão de que o tok pisin manifesta mais complexidade, pelo menos nos usos do sufixo *-pela*, do que as línguas substratais e superstratais, precisamente porque é uma língua *pidgin e/ou* crioula.

bém pode ser vista como um argumento contrário à transparência semântica (Plag 2006: 306).⁴

Mais adiante veremos que em alguns dos crioulos aqui estudados é muito comum que se tenham tomado de empréstimo às línguas lexificadoras palavras morfológicamente complexas e que esses empréstimos inclusive sejam responsáveis pela extensão das listas de afixos derivacionais nas descrições desses crioulos.

Nas secções seguintes, vamos estudar a afixação derivacional de alguns crioulos de base lexical portuguesa e espanhola tomando em consideração a questão da transparência semântica e da produtividade, esta última estando ligada à primeira a tal ponto que Lefebvre (2003: 72) afirma que a morfologia produtiva é por definição semanticamente transparente.⁵ Outra definição da transparência semântica é oferecida por Plag (2006: 305): a transparência implica a possibilidade de inferir o significado de palavras complexas a partir das suas partes. Note-se que McWhorter tem substituído o termo “transparente”, nos seus últimos escritos, por “composicional”. Justificamos este pequeno estudo pelo pouco que se tem trabalhado sobre o tema nos crioulos de base ibero-românica.⁶ Compararemos os crioulos estudados e procuraremos constatar se houve mudanças recentes nos inventários dos afixos derivacionais ou se tais inventários reflectem ainda o estado mais ou menos original da língua.

⁴ Além disso, Plag (2001: 154) declara que a opacidade semântica é uma “non-issue” quando se tem em conta que as pesquisas psicolinguísticas recentes mostrariam que as formas semanticamente opacas estão representadas como formas morfológicamente simples no léxico mental (cf. também McQueen & Cutler 1998).

Arends *et al.* (2006) sugerem outra divisão, indo além da mera distinção entre morfologia flexional e derivacional: uma divisão em morfologia inerente e morfologia contextual. Os crioulos, efectivamente, parecem não ter morfologia contextual (para uma definição dessa oposição, veja-se Arends *et al.* (2006: 235) ou Booij 1993: 30). Muysken (2004: 1655) cita os seguintes factores favoráveis à retenção da morfologia durante o processo de pidginização-crioulização: trata-se de morfemas derivacionais, não há alomorfa, o processo é aglutinativo e sintagmaticamente interpretável. A complexidade da palavra que lista como quinto factor não nos parece importante, pois o próprio autor afirma que (*ibid.*) que a complexidade em si não determina se uma palavra é retida ou não.

⁵ Ainda que Bauer (1983: 19) observe que nem todos os morfemas derivacionais produtivos são transparentes – dando o exemplo da substantivação no inglês *coverage* ‘cobertura’ < *cover* ‘cobrir’ + *-age*, por um lado, e o de *carriage* ‘carruagem, coche’ < *carry* ‘levar’ + *-age*, que o autor considera opaco –, também é preciso considerar a idade do produto da derivação.

⁶ Arends *et al.* (2006) constituem uma excepção, já que incluem o angolano, o papiamentu e o palenquero. Tivemos conhecimento deste artigo só muito recentemente.

2 – Os inventários dos afixos derivacionais dos crioulos estudados

2.1 – O cabo-verdiano

Veiga (1995: 122-129) apresenta uma lista muito extensa de afixos derivacionais. Divide o campo da prefixação em prefixos de origem grega e latina e admite que a produtividade dos primeiros é muito baixa (Veiga 1995: 123). O inventário dos sufixos é muito maior ainda que Veiga não o reconheça, é muito provável que a maior parte das palavras derivadas tenham sido retidas do português na sua forma integral, ou seja, que não se trate de derivações ocorridas no crioulo. Por conseguinte, os afixos correspondentes não são produtivos no cabo-verdiano. Vejam-se, por exemplo, as “derivações” *antibrasu* ‘antebraço’, *bisnetu* ‘bisneto’, *papelaria* ‘papelaria’, *parokianu* ‘paroquiano’, *sosialismu* ‘socialismo’, *studanti* ‘estudante’, *traison* ‘traição’, *pintura* ‘pintura’, *razistenti* ‘resistente’, etc. (Veiga 1995: 122, 125-127).

Quint (2000) e Brüser & Santos (2002) são muito mais cautelosos no que diz respeito à produtividade dos afixos derivacionais. Apresentamos na Tabela 1 os afixos listados por Brüser & Santos (2000: xli-xliv), com a adição dos sufixos *-ista*, *-udu* e *-Vdu* de Quint (2000: 144, 152):⁷

Tabela 1: Afixos derivacionais do cabo-verdiano

afixo	função	exemplo	etimologia/ afixo cor- resp. portu- guês ⁸
<i>dis-, diz-</i>	modificação de verbos e adjectivos, frequentemente anulação do significado da base	<i>diskarapati</i> ‘soltar-se, perder o apoio, cair’ < <i>karapati</i> ‘segurar(-se), agarrar(-se), abraçar’	<i>des-</i>
<i>ra-</i>	repetição de verbos	<i>radjobe</i> ‘remirar, rever’ < <i>djobe</i> ‘olhar, procurar, examinar’	<i>re-</i>
<i>-ada</i>	formação de substantivos a partir de predicados nominais e verbais	<i>totisáda</i> ‘um só golpe no pescoço’ < <i>totis</i> ‘nuca, toutiço’	<i>-ada</i>

⁷ O estado do sufixo *-Vdu*, quer dizer, do participípio passado do português, é geralmente contestado: no caso do angolano, Arends *et al.* (2006: 226) consideram que se trata de um sufixo de flexão enquanto Maurer (1995: 91) o lista entre os sufixos derivacionais (cf. 2.3.).

⁸ Tendo em conta que as funções dos afixos derivacionais são muito semelhantes em português, não as listamos na tabela, mas só o afixo correspondente do português. Fazemos uma excepção no caso do crioulo do Ceilão do século XIX, onde inclusive a forma dos afixos é a mesma que no português (por isso, não é preciso citá-los), bem como no caso do chabacano, onde há diferenças importantes no tocante às funções que lhes correspondem nas línguas de origem.

afixo	função	exemplo	etimologia/ afixo cor- resp. portu- guês ⁸
- <i>ânsa</i>	formação de substantivos a partir de predicados nominais e verbais	<i>konsolânsa</i> 'consolação' < <i>konsola</i> 'consolar'	- <i>ança</i>
- <i>ânti</i>	formação, a partir de predicados verbais, de substantivos e adjectivos para a designação dos sujeitos inerentes a estes verbos	<i>rabidânti</i> 'revendedora' < <i>rabida</i> 'comprar para revender'	- <i>ante</i>
- <i>ariâ</i>	formação de substantivos a partir de predicados adjectivais	<i>salbaxariâ</i> 'maldade, má-criação, travessura' < <i>salbaxi</i> 'grosseiro, malcriado, selvagem'	- <i>aria</i>
-(<i>i</i>) <i>dádi</i> , - (<i>i</i>) <i>ndádi</i>	formação de substantivos a partir de predicados adjectivais	<i>balindádi</i> 'qualidade de servir' < <i>bale</i> 'prestar, servir, ser útil, valer'	- <i>dade</i>
- <i>dor</i> , - <i>dera</i>	formação, a partir de predicados verbais, de substantivos e adjectivos para a designação dos sujeitos inerentes a estes predicados	<i>ledor</i> 'leitor' < <i>le</i> 'ler'; <i>nhemedera</i> 'dentadura, dentes' < <i>nheme</i> 'mastigar'	- <i>dor</i> , - <i>deira</i>
- <i>dura</i>	formação de substantivos a partir de predicados verbais; o resultado designa uma única ocorrência	<i>kansadura</i> 'canseira, cansaço' < <i>kânsa</i> 'cansar(-se)'	- <i>dura</i>
- <i>éns(i)a</i>	formação de substantivos a partir de predicados verbais e adjectivais	<i>intirfirénsa</i> 'interferença' < <i>intirfiri</i> 'interferir'	- <i>ência</i>
- <i>éntu</i>	formação de adjectivos a partir de predicados substantivos	<i>medéntu</i> 'que tem medo' < <i>médu</i> 'medo'	- <i>ento</i>
- <i>eru</i> , - <i>era</i>	formação, a partir de predicados verbais e nominais, de substantivos e adjectivos para a designação dos sujeitos inerentes a estes predicados	<i>faronperu</i> 'fanfarrão, bazofiator' < <i>farónpa</i> 'pompa, fanfarronice, bazófia'	- <i>eiro</i> , - <i>eira</i>
- <i>ésa</i> , - <i>éza</i>	formação de substantivos a partir de predicados adjectivais	<i>morabéza</i> 'amabilidade, afabilidade' < <i>morábi</i> 'amável, afável'	- <i>eça</i> , - <i>eza</i>
- <i>ida</i>	formação de substantivos a partir de predicados verbais	<i>djagasida</i> 'especialidade culinária da ilha de Fogo' < <i>djagasi</i> 'atrapalhar-se, perder a cabeça'	- <i>ida</i>
- <i>inhu</i> , - <i>inha</i>	formação de diminutivos a partir de substantivos, adjectivos e advérbios	<i>bunitinha</i> < <i>bunita</i> 'bonita'	- <i>inho</i> , - <i>inha</i>
- <i>ista</i>	formação, a partir de predicados nominais, de substantivos e adjectivos para a designação dos sujeitos inerentes a estes predicados	<i>kadjetista</i> 'pessoa que pensa em e vai a Calheta regularmente' < <i>Kadjeta</i> 'Calheta (cidade em Santiago)'	- <i>ista</i>
-(<i>a</i>) <i>menti</i>	formação de advérbios a partir de predicados adjectivos	<i>prontamenti</i> 'pronto!' < <i>pronto</i> 'pronto'	-(<i>a</i>) <i>mente</i>
- <i>mentu</i>	formação de substantivos a partir de predicados verbais	<i>kortamentu</i> (<i>di barriga</i>) 'dores de barriga acompanhadas de diarreia' < <i>korta</i> 'cortar', ser cortado	- <i>mento</i>

afixo	função	exemplo	etimologia/ afixo cor- resp. portu- guês ⁹
-on, -óna	formação de aumentativos a partir de substantivos e adjetivos	<i>kason</i> < <i>kása</i> 'casa'	-ão, -ona
-ós, -osu, -ozu	formação de adjetivos a partir de predicados substantivos	<i>skerdós</i> 'canhoto' < <i>skérdu</i> 'esquerdo'	-oso
-son	formação de substantivos a partir de predicados verbais	<i>diskulpason</i> ⁹ 'desculpa' < <i>diskulpa</i> 'desculpar(-se), perdoar'	-são
-udu	formação de adjetivos a partir de predicados substantivos	<i>bokudu</i> 'gárrulo' < <i>bóka</i> 'boca'	-udo
-ura	formação de substantivos a partir de predicados adjetivais	<i>bedjura</i> 'idade avançada' < <i>bédju</i> 'velho'	-ura
-Vdu	formação de adjetivos a partir de predicados verbais	<i>raparádu</i> 'malcriado, que não tem respeito para os outros' < <i>rapára</i> 'reparar, observar, consertar'	-Vdo

Quint (2000: 146) observa que as variantes com *-(i)dádi* e *-éza* são mais recentes que *-indádi* e *-ésa* e considera que as primeiras entraram na língua como empréstimos, enquanto que os últimos seriam produtivos.¹⁰ Pelo menos na formação de nomes de plantas, *-eru/-era* também não seria produtivo, e os nomes de árvores formar-se-iam com *pé di*, p.ex.: *pé di banána* 'bananeiro' (Quinto 2000: 145). A formação de advérbios com *-menti* é, segundo o mesmo autor, um fenómeno de aportuguesamento, com a excepção notável de *antigamenti* 'antigamente' (Quint 2000: 221).

Lamentavelmente, as fontes mais antigas sobre o cabo-verdiano, como Brito (1887), não contêm informação sobre a derivação. Ainda que tenhamos escolhido exemplos das representações da morfologia derivacional de Brüser & Santos (2002) e de Quint (2000) que não

⁹ Tem o sinónimo *diskulpa*.

¹⁰ Quanto aos outros casos de alomorfia, a distribuição de *-inhu/-inha* e *-eru/-era* é condicionada pelo sexo do referente, no caso dos referentes animados somente, e pelo fonema final da base (caso a base dos diminutivos termine em vogal ou ditongo tônicos, intercala-se um /s/). A distribuição de *-on/-óna* é determinada pela fixação no léxico do derivado e talvez pela ênfase (caso a base termine em vogal ou ditongo tônicos, intercala-se um /s/ no crioulo fundo e um /z/ no crioulo leve); a distribuição de *dis-/diz-* é determinada pelo facto de a base começar com consoante ou vogal (Brüser & Santos 2002: xli-xlii, xliiv).

existem no português (pelo menos, não no português contemporâneo), para demonstrar que se trata de afixos produtivos, e se bem que consideremos possível encontrar muitos mais (p.ex.: *berador* ‘pessoa que fica perto da mesa ou vem de visita nas horas de comida para obter a sua parte’ < *bera* ‘beira’ e *minizada* ‘meninos (colectivo?)’ < *minis* ‘meninos’ (Pires & Hutchinson 1983 s.v.), suspeitamos que uma proporção importante dos “derivados” do cabo-verdiano tenham sido retidos do português *in toto*, tanto no caso das retenções mais antigas, como no dos empréstimos recentes. Contudo, sem um estudo quantitativo, é impossível pronunciarmo-nos sobre a magnitude dessa proporção.

No que tange à transparência ou à composicionalidade semântica dos derivados, figuram na Tabela 1 vários exemplos que definitivamente não correspondem a esta categoria (acresce que seria bastante possível alargar essa lista de exemplos): estão neste caso *djagasida* ‘especialidade culinária da ilha de Fogo’, *prontamenti!* ‘pronto!’ (só se utiliza como exclamação), *kortamentu (di bariga)* ‘dores de barriga acompanhadas de diarreia’, *bokudu* ‘gárrulo’, *raparádu* ‘malcriado, que não tem respeito para com os outros’¹¹ e não é evidente porque *nhemedera* significa ‘dentadura, dentes’ e não, p.ex., ‘mulher ou fêmea que mastiga’.

2.2 – O kriyôl da Guiné-Bissau

No quadro que se segue (Tabela 2), apresentamos uma adaptação da lista que Couto (1994: 85-87) faz dos afixos derivacionais mais ou menos produtivos no kriyôl da Guiné-Bissau:¹²

Tabela 2: Afixos derivacionais do kriyôl

afixo	função	exemplo	etimologia ¹³
<i>dis-</i>	modificação de verbos, frequentemente anulação do significado da base	<i>disjunda</i> ‘soltar, abrandar’ < <i>des-</i> <i>junda</i> ‘retesar, esticar’	
<i>-ada</i>	formação de substantivos a partir de predicados verbais	<i>kunsada</i> ‘início’ < <i>kunsa</i> ‘começar’	<i>-ada</i>

¹¹ Encontrámos este exemplo na parte de dicionário de Brüser & Santos (2002).

¹² Por exemplo, Couto (1994: 87) afirma que *-ansa* e *-on* são tão pouco produtivos que as palavras que as contêm são sentidas como unidades não analisáveis pelos falantes. Por conseguinte, não listamos aqueles sufixos.

¹³ Ou seja, correspondência no português, salvo no caso de *-ndV/-ntV*, que provém do mandinga.

afixo	função	exemplo	etimologia
-asku	formação de substantivos abstractos a partir de adjectivos	<i>bunitasku</i> 'beleza' < <i>bunitu</i> 'bonita'	-asco
-dia	formação de substantivos a partir de predicados nominais	<i>kamaradia</i> 'amizade' < <i>kamarada</i> 'amigo, kamarada'	-dia
-dur	formação, a partir de predicados verbais, de substantivos para a designação dos sujeitos inerentes a estes predicados, geralmente nomes de profissões	<i>bajadur</i> 'bailarino' < <i>baja</i> 'bailar'	-dor
-eru	formação, a partir de predicados verbais e nominais, de substantivos para a designação dos sujeitos inerentes a estes predicados, geralmente profissões	<i>kalfanтеру</i> 'carpinteiro' < <i>kalfata</i> 'calafetar'	-eiro
-esa	formação, a partir de predicados nominais e adjectivais, de substantivos, o derivado indicando uma qualidade ou condição abstracta	<i>faimadesa</i> 'fome, avidez' < <i>faima</i> 'esfomear'	-eça, -eza
-ia	formação de verbos a partir de predicados nominais	<i>mandukia</i> 'espancar' < <i>manduku</i> 'pau'	-ear, -iar
-menti	formação de substantivos a partir de predicados verbais	<i>kaímenti</i> 'caída' < <i>kaí</i> 'cair'	-mento
-ndadi	formação de substantivos a partir de predicados nominais e adjectivais	<i>amigundadi</i> 'amizade' < <i>amigu</i> 'amigo'	-dade
-ndV, -ntV	formação de causativos a partir de verbos	<i>ciganta</i> 'fazer chegar' < <i>ciga</i> 'chegar'	mandinga -ndi
-sí	formação de verbos a partir de predicados nominais e adjectivais	<i>bejusi</i> 'envelhecer' < <i>beju</i> 'velho'	-cer
-siñu	formação de diminutivos a partir de substantivos	<i>mamesiñu</i> 'madrastra' < <i>mame</i> 'mãe', <i>papesiñu</i> 'padrastra' < <i>pape</i> 'pai', <i>ermonsiñu</i> 'irmão menor' < <i>ermon</i> 'irmão'	-(z)inho
-uda, -ura	formação de substantivos a partir de predicados adjectivais	<i>larguda</i> 'largura' < <i>largu</i> 'largo'	-ura

Segundo Couto (1994: 85), os derivados com *-asku* e *-eru* não são muito produtivos. No caso das designações de profissões em *-eru*, opina que frequentemente são emprestados tais quais do português. Quando não se tratar de nomes de profissões, pode ocorrer a variante *-era*, p.ex. *kabasera* 'cabaceira' < *kabas* 'cabaça' (ibid.).¹⁴ Entre as

¹⁴ Então parece questionável a presença de *-eru* na Tabela 2. Lembremo-nos também do que diz Quint com respeito à formação dos nomes de árvores no cabo-verdiano. É uma lacuna que Couto traga só este único exemplo do emprego da variante *-era*.

variantes *-uda* e *-ura*, a primeira seria a original mas já não muito produtiva hoje em dia e estaria substituída pela segunda no kriyôl moderno (Couto 1994: 87). Como no caso do cabo-verdiano, o sufixo *-ndadi* parece ser crioulo, a variante *-dadi* mais bem portuguesa. Com efeito, Rougé (1988: 16) considera que por exemplo *koldadi* ‘qualidade’ e *bardadi* ‘verdade’ entraram na forma integral no kriyôl. Por fim, é de notar que o sufixo causativo *-ndV*, *-ntV* tem como fonte primária o sufixo mandinga *-ndi*, também causativo, ainda que a influência convergente do sufixo português *-antar* não é de excluir (Couto 1994: 86; Rougé 1988: 17).

Em quanto à opacidade semântica dos derivados, parece que os afixos produtivos produzem derivados maioritariamente transparentes como há de esperar-se. No caso dos exemplos não transparentes de derivações com o sufixo diminutivo citados na Tabela 2 pode tratar-se de calcos duma língua de adstrato. Por fim, ainda que Couto não diga nada sobre a produtividade do sufixo *-si*, duvido que seja produtivo. A parte do exemplo citado na Tabela 2, os demais derivados listados por Couto (1994: 87; sem bases!) não parecem transparentes, cf. *bensi* ‘benzer, rezar’, *mansi* ‘pernoitar’, *notsi* ‘passar o dia’ [sic], *obidisi* ‘obedecer’, *lebsi* ‘difamar, desrespeitar’.

Como podemos observar, o kriyôl compartilha a maioria dos afixos derivacionais com o cabo-verdiano, com certas divergências: o inventário consta de catorze afixos, não vinte e quatro como no cabo-verdiano;¹⁵ dos afixos de origem portuguesa, *-asku*, *-dia*, *-ia*, *-si* não se citam nos dois inventários cabo-verdianos principalmente consultados;¹⁶ a inexistência do sufixo causativo *-ndV*, *-ntV* no cabo-verdiano explica-se pela menor duração do contacto com as línguas de sub/adstrato; e, comparando as funções citadas nas Tabelas 1 e 2, nota-se que com vários afixos, o campo de aplicação é diferente (em geral, mais restringido) daquele dos afixos homólogos do cabo-verdiano.

¹⁵ Os números 14 e 24 não deveriam considerar-se de forma absoluta, porque é muito provável que nas tabelas figurem afixos cuja produtividade é muito baixa, como, por exemplo, o caso, já mencionado, do kriyôl *-eru/-era*–*eru/-era*.

¹⁶ Veiga (1995: 128), pelo contrário, lista *-ise*, o mesmo sufixo que kriyôl *-si* derivado do sufixo português *-ecer*.

2.3 – O angolar

Dependendo da definição, o angolar tem dois ou três afixos derivacionais, todos sufixos. Maurer (1995: 50-51, 90-91) indica os três sufixos que damos na Tabela 3:

Tabela 3: Afixos derivacionais do angolar

afixo	função	exemplo	etimologia/ correspondência no português
<i>-me(n)tu</i>	formação de substantivos a partir de predicados verbais	<i>kadhamentu</i> 'casamento' < <i>kadha</i> 'casar-se'	<i>-mento</i>
<i>-rô, -dô</i>	formação, a partir de predicados verbais, de substantivos para a designação dos sujeitos inerentes a estes predicados	<i>zirô taba</i> 'fazedor de tábuas' < <i>zi taba</i> 'fazer tábua'	<i>-dor</i>
<i>-ru, -du</i>	formação de predicados adjetivais a partir de predicados verbais; modificação de adjetivos e de <i>môtxi</i> 'muito' no sentido da ênfase	<i>n'djaru</i> 'de pé' < <i>n'dja</i> 'parar-se', <i>fiôru</i> '(muito) frío' < <i>fiô</i> 'frío', <i>môtxiru</i> 'muito(s), muitíssimo(s)' < <i>môtxi</i> 'muito'	<i>-Vdo</i>

Arends *et al.* (2006: 226, 229) consideram que o angolar tem só dois sufixos derivacionais. *-ru* e *-du*, segundo estes autores, são sufixos flexionais, já que não são considerados os usos enfáticos com bases adjetivais, nem a construção com *môtxi*. Ainda que bastante produtivos (contudo, nem todos os adjetivos podem levar o sufixo – cf. Maurer 1995: 51), esses usos enfáticos podem ser considerados semanticamente opacos.¹⁷ No que diz respeito à formação de predicados adjetivais a partir de predicados verbais, parece que muitos participios portugueses terão sido emprestados de maneira holística. O mesmo deve ter ocorrido com vários derivados em *-me(n)tu*. A julgar pelos poucos exemplos dados por Maurer, ambos os tipos de derivações seriam em si transparentes. Por outro lado, a formação de *nomina agentis* com *-rô, -dô* é genuinamente produtiva e transparente ao mesmo tempo.

¹⁷ Note-se que Maurer (1995: 50-51) não traduz a presença do sufixo enfático com 'muito'; fizemo-lo na Tabela 3 para clarificar melhor a função do sufixo. Lorenzino (1998: 137-139) concorda com a análise que Maurer faz de *-ru, -du* e fala das "propriedades idiossincráticas" do sufixo.

Ainda que o angolares esteja em contacto com o crioulo são-tomense e cada vez mais também com o português (Maurer 1995: 2), isso não parece ter influenciado no sistema de afixos derivacionais que continua sendo muito reduzido frente ao português e aos crioulos de Cabo-Verde e da Guiné-Bissau.

2.4 – O papiamentu

Com as descrições do papiamentu, estamos novamente¹⁸ perante uma situação em que certos autores listam uma multiplicidade de afixos derivacionais (por exemplo, Munteanu 1991: 181, 187-189), por oposição a outros que consideram que o papiamentu tem só dois ou três afixos derivacionais produtivos. Esta é a posição de Lenz (1928: 148-150), de Maurer (1998: 182) e de Dijkhoff (1993: 74-80). Kouwenberg & Murray (1994: 22-31) tomam uma posição intermédia. Vejamos a Tabela 4:¹⁹

Tabela 4: Afixos derivacionais do papiamentu

afixo	função	exemplo	etimologia ²⁰
<i>in-</i>	formação de antónimos de predicados adjectivais	<i>inabitá</i> 'inabitado' < <i>abitá</i> 'habitado'	esp., port. <i>in-</i>
<i>re-</i>	repetição de verbos, modificação de verbos no sentido dum caso específico	<i>reboltiá</i> 'revoltar' < <i>boltiá</i> 'voltar'	esp., port. <i>re-</i>
<i>-bel</i>	formação de adjectivos a partir de predicados verbais; o resultado exprime a capacidade de subir a acção do verbo	<i>kurabel</i> 'curável' < <i>kura</i> 'curar'	esp. <i>-ble</i> , port. <i>-vel</i>
<i>-dó</i>	formação, a partir de predicados verbais, de substantivos para a designação dos sujeitos inerentes a estes predicados	<i>fêrdó</i> 'pintor (de paredes)' < <i>fêrf</i> 'pintar'	esp., port. <i>-dor</i>
<i>-ero</i> , <i>-era</i>	formação, a partir de predicados nominais, de substantivos para a designação dos sujeitos inerentes a estes predicados	<i>kunukero</i> 'camponês' < <i>kunuko</i> 'campo'	esp. <i>-ero</i> , <i>-era</i>
<i>-eria</i>	formação de nomes de comércio a partir de substantivos	<i>bukeria</i> 'livraria' < <i>buki</i> 'livro'	esp. <i>-eria</i> , port. <i>-eria</i>

¹⁸ Quer dizer, como no caso do cabo-verdiano anteriormente discutido (cfr. 2.2.).

¹⁹ Nas Tabelas 4, 5 e 6 utilizamos as seguintes abreviaturas: esp. = espanhol, hil. = hiligainão, hol. = holandês, ingl. = inglês, port. = português.

²⁰ Enquanto, por exemplo, Munteanu (1991) considera que o papiamentu é um crioulo de base lexical exclusivamente espanhola, a nossa opinião é a de que estamos perante um crioulo de base lexical espanhola e portuguesa. Por isso, citamos as duas línguas quando não for absolutamente claro que o afixo possa ser originário de apenas uma delas.

afixo	função	exemplo	etimologia
<i>-ista</i>	formação, a partir de predicados nominais, de substantivos para a designação dos sujeitos inerentes a estes predicados, inclusive a designação de membros de organizações	<i>wèlguista</i> 'grevista' < <i>wèlga</i> 'greve', <i>manista</i> 'membro da M.A.N.' < M.A.N.	esp., port. <i>-ista</i>
<i>-isá</i>	formação de verbos a partir de predicados nominais	<i>ferbalisá</i> 'emitir uma garantia' < <i>ferbal</i> 'garantia'	esp., port. <i>-izar</i>
<i>-mente</i>	formação de advérbios a partir de predicados adjectivos	<i>únikamente</i> 'unicamente' < <i>úniko</i> 'único'	esp., port. <i>-mente</i>
<i>-mentu</i>	formação de substantivos a partir de predicados verbais	<i>dalmentu di outo</i> 'choque de carros' < <i>dal</i> 'golpear'	esp., port. <i>-mento</i>
<i>-nsa</i>	formação de substantivos a partir de predicados verbais	<i>yudansa</i> 'ajuda' < <i>yuda</i> 'ajudar'	esp. <i>-nza</i> , port. <i>-nça</i>
<i>-shi</i>	formação de substantivos a partir de predicados verbais	<i>redashi</i> 'bote' < <i>reda</i> 'espalhar um bote'	hol. <i>-tie</i>
<i>-shon</i>	formação de substantivos a partir de predicados verbais; o substantivo designa o resultado ou um caso da acção do predicado verbal	<i>strobashon</i> 'estorvo' < <i>stroba</i> 'estorvar'	esp. <i>-ción</i>

A produtividade da maior parte dos afixos listados é baixa ou até muito baixa. Maurer (1998: 182) considera que somente *-dó* e *-mentu* são plenamente produtivos e diz que os advérbios em *-mente* “quase sempre” se formam como em espanhol (ou português), ou seja, a partir duma base que termina em *-a* enquanto os adjectivos de origem ibero-românica são invariáveis e terminam em *-o* no papiamentu. Contudo, não chega à mesma conclusão que Kouwenberg & Murray (1994: 23, 29), que dizem que *-mente* simplesmente não é um afixo derivacional produtivo. Lenz (1928: 148-150), à época em que publicou o seu estudo, considerava que só *-mentu*, *-dó* e *-ero* eram produtivos. Isso significa que o inventário dos afixos derivacionais quase não mudou durante o século XX, um período porém em que outros aspectos da língua mudaram (incursão do léxico inglês, aumento do léxico de origem espanhola e holandesa, etc.). A grande maioria dos derivados são transparentes. Casos de derivações não transparentes citados na Tabela 4 seriam *reboltíá* 'revoltar' e *ferbalisá* 'emitir uma garantia'. É de notar que, tanto no papiamentu como no angolar, *-mentu* e *-dó* são

os únicos dois sufixos plenamente produtivos, sendo ainda de realçar a sua produtividade também no cabo-verdiano e no kriyôl.²¹ Por fim, notemos o facto de que na lista figura um sufixo de origem holandesa, resultado da convivência de séculos com essa língua de prestígio.

2.5 – O crioulo de Ceilão do fim do século XIX

A decisão de incluir uma variante histórica nesta comparação pode parecer estranha. Justificamo-la com os argumentos seguintes: a inclusão do chabacano, crioulo de base lexical espanhola, exige, de certa maneira, a apresentação da morfologia derivacional de um crioulo asiático de origem portuguesa. Além disso, a morfologia derivacional do crioulo de Ceilão do fim do século XIX apresenta fenómenos interessantes e as pouquíssimas descrições da variedade do século XX não contêm informação sobre o nosso tema. Contudo, é preciso lembrar que se tem assumido que precisamente no decorrer do século XX tiveram lugar umas mudanças tipológicas significativas que aproximaram o crioulo da sua língua de adstrato, o tamil (e, indirectamente, também ao cingalês, língua oficial do Sri Lanka, ainda que menos falado pelos crioulófonos). Essa posição é defendida em Smith (1978; 1979), embora este autor seja da opinião de que (comunicação pessoal, Junho de 2006) que o crioulo do fim do século XIX descrito sobretudo por Dalgado era um crioulo muito aportuguesado. Muito provavelmente, a guerra civil no Sri Lanka tem tido como efeito que a comunidade de falantes do crioulo, tradicionalmente radicada em Batticaloa, no leste da ilha, se tenha dispersado e que já não se possam fazer novas descrições da língua.

Dalgado (1900a: 28-29, 31, 52-53) cita derivações com oito afixos diferentes.²² Um estudo do seu vocabulário revela, porém, que há pelo menos uns oito outros afixos mais ou menos produtivos. Tendo em conta que Dalgado conhecia bem os dialectos portugueses no tempo e no espaço, podemos em princípio excluir que se trate, por exemplo, de regionalismos, no caso dos exemplos citados abaixo, já

²¹ Enquanto os derivados em *-dor* parecem (quase) sempre ser transparentes, não é sempre o caso dos derivados em *-mentu*. Cf. cabo-verdiano *kortamentu di bariga*.

²² *Des-/dis-*, *-ança*, *-ção*, *-dade*, *-dor*, *-dura*, *-mente*, *-oso*.

que Dalgado o teria comentado nas entradas respectivas. Listamos os afixos derivacionais que encontramos no crioulo de Ceilão do fim do século XIX na Tabela 5:²³

Tabela 5: Afixos derivacionais do crioulo de Ceilão do século XIX

afixo	função	exemplo
<i>des-</i> , <i>dis-</i>	formação de antónimos de predicados nominais e adjectivais	<i>desfrutuoso</i> 'infrutoso' < <i>frutuoso</i> 'frutoso'
<i>-ado</i>	formação de adjectivos a partir de predicados verbais ²⁴	<i>devotado</i> 'devoto' < <i>devotá</i> 'devotar(-se)', ingl. <i>devoted</i>
<i>-ança</i> , <i>-ance</i>	formação de substantivos a partir de predicados verbais	<i>accordança</i> , <i>accordance</i> 'accordo' < <i>accordá</i> 'acordar', ingl. (in) <i>accordance</i>
<i>-ando</i>	formação de adjectivos a partir de predicados verbais	<i>purificando</i> 'purificante' < <i>purificá</i> 'purificar'
<i>-ante</i>	formação de adjectivos a partir de predicados verbais	<i>passante</i> 'passado' < <i>passá</i> 'passar'
<i>-avel</i>	formação de adjectivos a partir de predicados verbais; o resultado exprime a capacidade de subir a acção do verbo	<i>honoravel</i> 'honrado, nobre' < * <i>honrá</i> 'honrar', ingl. <i>honourable</i>
<i>-ça a n</i> , <i>-ção</i>	formação de substantivos a partir de predicados verbais	<i>comidação</i> 'convite' < <i>comidá</i> 'convitar', ingl. <i>Invitation</i>
<i>-dade</i>	formação de adjectivos a partir de predicados verbais	<i>feudade</i> 'fealdade' < <i>feu</i> 'feito'
<i>-dor</i>	formação, a partir de predicados verbais, de substantivos para a designação dos sujeitos inerentes a estes predicados	<i>furtador</i> 'ladrão' < <i>furta</i> 'furtar'
<i>-dura</i>	formação de substantivos a partir de predicados verbais	<i>gemeduras</i> 'gemidos' < <i>gemé</i> 'gemer'
<i>-eiro</i>	formação, a partir de predicados verbais, de adjectivos e substantivos para a designação dos sujeitos inerentes a estes predicados	<i>servideira</i> 'servidora' < <i>servi</i> 'servir'
<i>-eza</i>	formação de substantivos a partir de predicados adjectivais	<i>branqueza</i> 'brancura' < <i>branco</i> 'branco'
<i>-mente</i>	formação de advérbios a partir de predicados adjectivos	<i>verdonhadamente</i> 'vergonhosamente' < <i>verdonhado</i> 'envergonhado'
<i>-mento</i>	formação de substantivos a partir de predicados verbais	<i>governamento</i> 'governo, principado' < <i>governá</i> 'governar', ingl. <i>government</i>
<i>-oso</i>	formação de adjectivos a partir de predicados nominais	<i>adulteiroso</i> , <i>adulterioso</i> 'adultero' < <i>adultería</i> 'adulterio', ingl. <i>adulterious</i>
<i>-ria</i>	formação de substantivos a partir de predicados verbais	<i>caçaria</i> 'caçada' < <i>caçá</i> 'caçar'

²³ No caso do crioulo de Ceilão do século XIX, as correspondências portuguesas são tão óbvias – na maioria dos casos, os afixos são inclusive formalmente iguais – que não os listamos à parte. Notemos novamente, porém, a influência do inglês na existência de *dobretes*.

Podemos observar que o uso que o crioulo de Ceilão do século XIX fazia dos afixos derivacionais era produtivo (provavelmente, em graus diferentes, dependendo do afixo) e transparente, eliminando em parte irregularidades do português ou criando dobretes morfologicamente possíveis onde o português possui uma só variante (por exemplo *livrança*, *livrance*, *livração* ‘livramento, libertação’).²⁵ Também vemos a partir dos exemplos (que se poderiam multiplicar) que a influência do inglês também se manifesta nos derivados.²⁶ Por fim, diferentemente da situação nos outros crioulos estudados até agora, a formação dos advérbios é produtiva, já que se faz com base no adjetivo na forma em que foi retida no crioulo – o masculino do português.²⁷

2.6 – O chabacano

Para a descrição dos afixos derivacionais do chabacano, baseámo-nos essencialmente no artigo de Steinkrüger (2003). Tal escolha justifica que tenhamos tratado das variedades do sul das Filipinas faladas em Zamboanga e Cotabato, na ilha de Mindanao, e não das variantes da Baía de Manila, entre as quais hoje em dia estão vivas as variedades de Cavite e Ternate. Em geral, as variedades do sul apresentam mais influência das línguas de adstrato que aquelas da Baía de Manila. Além disso, as principais línguas de adstrato austronésicas das primeiras são as línguas das Visayas (o cebuano e o hiligainão), enquanto que, no caso das segundas, o adstrato é o tagalog.²⁸

Segundo Steinkrüger (2003: 257-259), o chabacano possui vinte afixos derivacionais. Entre tais afixos, conta-se uma proporção muito elevada de prefixos, comparativamente com o que se verifica em

²⁴ Ou emprego de gerúndios e participípios como adjectivos (cf. Dalgado 1900a: 40). Esta observação aplica-se também aos sufixos *-ando* e *-ante*.

²⁵ A formação análoga de participípios que no português são irregulares actua no mesmo sentido. Cumprindo com uma sugestão dum revisor externo, sublinhamos que a forma *livrance* só existe no crioulo de Ceilão do século XIX, não no português-padrão.

²⁶ Obviamente, há muitos empréstimos directos do inglês. Além disso, há empréstimos do holandês e de outras línguas.

²⁷ Uma excepção curiosa é constituída pelas formas *grandomente*, *grandamente* ‘grandemente’ (Dalgado 1900a: 157).

²⁸ Isso é devido às circunstâncias históricas e geográficas da formação dos crioulos de base espanhola nas Filipinas.

relação aos crioulos anteriormente considerados (oito afixos ou 40% do total). O que é ainda mais interessante é que os prefixos provêm das línguas de adstrato, principalmente do hiligainão, enquanto a maior parte dos sufixos (dez entre doze ou 50% do total dos afixos derivacionais) provêm do espanhol (Steinkrüger 2003: 256).²⁹ Com a excepção de *-ero/-era*, os sufixos de origem espanhola só se podem afixar a raízes espanholas enquanto os afixos austronésicos se combinam com qualquer base.³⁰ Por essa mesma razão, têm uma produtividade mais alta. Apesar de Steinkrüger dizer que lista unicamente os afixos derivacionais produtivos (2003: 257), não deixa de reconhecer, na sequência imediata de tal afirmação, que duvida da produtividade de vários sufixos espanhóis, pelo menos dos casos de *esa*, *-ísta*, *-syon*, *-tibo*, *-a/-o* (2003: 259). Por outro lado, considera que por exemplo os sufixos espanhóis *-do/-w* e *-da* são muito produtivos (2003: 259-260).

Tabela 6: Afixos derivacionais do chabacano

afixo	função	etimologia	função na língua de origem	exemplo
<i>ika-</i>	formação de números ordinais a partir de números cardinais	hil. <i>ika-</i>	a mesma ³¹	<i>ika-dos</i> 'segundo' < <i>dos</i> 'dois'
<i>ma-</i>	formação de adjectivos, em geral a partir de substantivos	hil. <i>ma-</i>	formação de adjectivos a partir de vários tipos de raízes	<i>ma-pyédra</i> 'pedregoso' < <i>pyédra</i> 'pedra'
<i>mag-</i>	formação de substantivos e adjectivos colectivos a partir de substantivos	?		<i>mag-ermáno</i> 'entre irmãos, que se refere a irmãos' < <i>ermáno</i> 'irmão'

²⁹ A predominância dos afixos austronésicos entre os prefixos explica-se pelo facto de essas línguas (por exemplo, o tagalog, que está detalhadamente descrito) terem mais prefixos que sufixos ou infixos, o que é ainda mais importante numa situação de contacto linguístico e empréstimo; com poucas excepções, os prefixos têm características de palavras lexicais: não se fundem com a base e conservam acento próprio (Himmelman 2004: 1476).

³⁰ Existem, porém, restrições ainda não muito bem entendidas no que diz respeito à integração das raízes de origem inglesa (Steinkrüger 2003: 259).

³¹ Quer dizer que a função do afixo é a mesma que no chabacano. Preferimos este breve comentário à possibilidade de repetir o texto que figura na colona "função" (no chabacano) por visualizar melhor que efectivamente se trata da mesma função sem alterações nas duas línguas.

afixo	função	etimologia	função na língua de origem	exemplo
<i>maka-</i>	formação dos causativos de substantivos, verbos e adjectivos	hil. <i>maka-</i>	formação de predicados a partir de vários tipos de raízes para a designação da capacidade ou qualidade inerente a estes predicados	<i>maka-bungúl</i> 'que torna surdo' < <i>bungúl</i> 'surdo'
<i>man-</i>	formação de verbos a partir de substantivos e adjectivos	hil. <i>man-</i>	formação de verbos a partir de substantivos	<i>man-amigo</i> 'ser/tornar-se amigos' < <i>amigo</i> 'amigo'
<i>pa-</i>	formação de verbos recíprocos a partir de adjectivos; formação de verbos causativos a partir de verbos; formação de predicados nominais direccionais a partir de substantivos	hil. <i>pa-</i> , hil. <i>pag-</i>	formação de gerúndios a partir de raízes diferentes	<i>pa-alisto</i> 'testar-se mutuamente as capacidades' < <i>alisto</i> 'listo, activo', <i>man-pa-balábak</i> ³² 'ser posto transversalmente' < <i>balábak</i> 'posto transversalmente; pontalete transversal'
<i>paka-</i>	formação de substantivos a partir de predicados adjectivais	hil. <i>paka-</i>	formação de substantivos abstractos a partir de qualquer expressão	<i>paka-alisto</i> 'talento' < <i>alisto</i> 'listo, activo'
<i>ta-</i>	designação do aspecto imperfeito nos verbos; função derivacional em certas combinações	esp. <i>está</i>	3ª pessoa singular do auxiliar <i>estar</i> , aspecto imperfeito com o gerúndio	<i>tampahóven</i> 'comportar-se como uma pessoa jovem' < <i>ta-man-pa-hóven</i> < <i>hóven</i> 'jovem'
<i>-(y)á, -é, -í</i>	formação de verbos a partir de predicados adjectivais	esp. <i>-ar, -er, -ir</i>	desinência do infinitivo	<i>bibí</i> 'viver' < <i>bibo</i> 'vivo'
<i>-a, -o</i>	distinção do sexo dos substantivos	esp. <i>-a, -o</i>	distinção do género dos substantivos e adjectivos	<i>nyéta</i> 'neta', <i>nyéto</i> 'neto'.
<i>-da</i>	formação de substantivos a partir de predicados verbais para a designação da maneira, do acto, do lugar ou do produto da acção	esp. <i>-da</i>	formação de <i>nomina actionis</i> a partir de predicados verbais	<i>abláda</i> 'maneira de falar' < <i>ablá</i> 'falar', <i>entendida</i> 'compreensão, maneira de entender' < <i>entendé</i> 'entender'
<i>-do, -w</i>	formação de adjectivos a partir de predicados verbais e nominais	esp. <i>-do</i>	desinência do participio passado masculino	<i>kurtáw</i> 'cortado' < <i>kurtá</i> 'cortar', <i>mantekádo</i> 'com muita manteiga' < <i>mantéka</i> 'manteiga'

³² Steinkrüger traduz *man-pa-balábak* por 'get set askew'. Cf. tagalog *balagbág* 'crossbeam, crosspiece; placed or lying crosswise, not straight, not parallel', *ibalagbág* 'to put something across the way', *pabalagbág* 'crosswise, awry' (English 2005: s.v.).

afixo	função	etimologia	função na língua de origem	exemplo
<i>-dor, -dora</i>	formação, a partir de predicados verbais, de substantivos para a designação dos sujeitos inerentes a estes predicados	esp. <i>-dor, -dora</i>	a mesma	<i>nadador</i> 'marinheiro' < <i>nadá</i> 'nadar'
<i>-era, -ero</i>	formação, a partir de predicados verbais, de substantivos para a designação dos sujeitos inerentes a estes predicados	esp. <i>-era, -ero</i>	a mesma	<i>kaskaséro</i> 'alguém quem conduz como um louco' < <i>kaskás</i> 'conduzir como um louco' ³³
<i>-esa</i>	formação, a partir de predicados adjetivais, de substantivos para a designação da qualidade destes predicados	esp. <i>-eza</i>	a mesma	<i>rikésa</i> 'riqueza' < <i>ríko</i> 'rico', <i>pobrésa</i> 'pobreza' < <i>pobre</i> 'pobre'
<i>-(h)an</i>	formação de verbos recíprocos a partir de verbos formados com o prefixo <i>man-</i> ; formação de nomes de lugares onde certa acção decorre	hil. <i>-(h)an</i>	sufixo locativo	<i>man-kwénto-han</i> 'contar-se' < <i>kwénto</i> 'conto'; <i>tubuhán</i> 'canal de açúcar' < <i>tubú</i> 'cana-de-açúcar'
<i>-(h)in</i>	formação de adjetivos a partir de predicados nominais	hil. <i>-gin</i> (?)	marcador dum construção passada passiva	<i>myedú-hin</i> 'tímido, temeroso' < <i>myédu</i> 'medo'
<i>-ísta</i>	formação, a partir de predicados verbais e nominais, de substantivos para a designação dos sujeitos inerentes a estes predicados	esp. <i>-ista</i>	a mesma	<i>chabacanísta</i> 'experto em chabacano' < <i>chabacano</i>
<i>-syón</i>	formação de substantivos a partir de predicados verbais	esp. <i>-ción</i>	a mesma	<i>dibisyón</i> 'divisão' < <i>dibidí</i> 'dividir'
<i>-tibo</i>	formação de adjetivos a partir de predicados verbais	esp. <i>-tibo</i>	a mesma	<i>pensatibo</i> 'pensativo' < <i>pensá</i> 'pensar'

A maior parte dos derivados são semanticamente transparentes. Há outros, porém, que não o são, como se verifica, por exemplo,

³³ Steinkrüger (2003: 259) nota que tanto o verbo como o derivado existem no tagalog. Por conseguinte, pelo menos essa derivação não é produtiva porque é pouco provável que o tagalog tenha tomado empréstimos às variedades do chabacano que Steinkrüger estuda.

através da multiplicidade das funções de *pa-*, *ta-* e *-(h)an* ou casos como *nadador* ‘marinheiro’ < *nadá* ‘nadar’ . Enquanto a função do afixo geralmente corresponde, nos outros crioulos estudados até agora, à sua função na língua de origem, este muitas vezes não é o caso do chabacano. Por isso, listámos, na Tabela 6, as funções dos afixos considerados tanto em chabacano como na língua de origem. Por último, ainda que o inventário de afixos derivacionais do chabacano seja considerável, é muito menor que os inventários encontrados quer em espanhol, quer em hiligainão (Steinkrüger 2003: 256).

3 – Discussão

Os seis crioulos estudados utilizam os afixos derivacionais de maneiras muito distintas. O angolár, geralmente classificado como crioulo radical, tem um inventário muito reduzido, de que faz um uso produtivo e transparente, enquanto que os inventários dos outros crioulos são muito mais extensos. A coexistência com a língua lexicadora contribui para um estado de coisas em que palavras derivadas transitam holisticamente de uma língua para a outra. Porventura sempre foi essa a situação deste crioulo, pelo menos em parte, ainda que tenha havido épocas na história do cabo-verdiano em que a presença do português foi menos significativa (cf., p. ex., Bartens 2000). Esta observação parece valer também para o papiamentu, que hoje em dia se encontra muito influenciado pelo espanhol, ainda que se trate de um crioulo de base lexical portuguesa e espanhola. Exemplifica esta situação a existência de autores que afirmam que o sufixo *-mente* se tornou produtivo no decorrer dos últimos cem anos, a par de outros que colocam em dúvida a sua produtividade e defendem que se trata de empréstimos holísticos (ver 2.4.).

Por outro lado, pelo carácter restrito dos contactos do kriyól com o português, existe uma necessidade real de se fazer uso produtivo, e na maior parte dos casos transparente, da sua morfologia derivacional, ainda que obviamente existam diferenças na produtividade de afixos individuais. O contacto prolongado com o adstrato levou à incorporação do afixo causativo tomado de empréstimo ao mandinga (com influência apenas convergente do afixo português). Note-se,

porém, que o inventário do kriyôl é aproximadamente da mesma extensão que o inventário do cabo-verdiano e do inventário que alguns têm postulado para o papiamento – que, aliás, incorporou um sufixo holandês de baixa produtividade durante o seu prolongado contacto com aquela língua da administração.

O crioulo de Ceilão do século XIX encontrou-se numa situação semelhante à do kriyôl actual: o contacto com o português-padrão não era muito forte e decididamente o crioulo não foi submetido a nenhum processo de normalização linguística, o que permitia a coexistência de variantes morfofonológicas, algumas delas emprestadas do crioulo português de Macau, conforme nota Dalgado (1900a) na parte de vocabulário, outras surgidas do contacto com o inglês. Em geral, a morfologia derivacional utiliza-se de maneira produtiva e transparente.

No chabacano, a tendência do distanciamento da língua lexificadora e da aproximação às línguas de adstrato é levada muito mais longe do que no kriyôl ou no crioulo de Ceilão do século XIX: como notámos na secção 2.6., os afixos derivacionais de origem austronésica são muito mais produtivos que os afixos espanhóis, apesar de os derivados não serem transparentes em todos os casos.

Resumindo, podemos dizer que, na maioria dos casos, a derivação é semanticamente transparente. Existem, porém, em todos os crioulos estudados contra-exemplos ao postulado da transparência ou da composicionalidade semântica de McWhorter.³⁴ Neste sentido, os nossos resultados concordam com os obtidos por Plag (2001; 2006), Braun & Plag (2003), Siegel (2004) e Arends *et al.* (2006). Ainda que a extensão do inventário de afixos tenha pelo menos em parte que ver com a relativa basi- ou acrolectalidade do crioulo (muito reduzido no angolar, geralmente considerado basilectal, e extenso no caso

³⁴ Em várias ocasiões, temos mostrado que os crioulos de base lexical ibero-românica não se conformam com os padrões estabelecidos sobretudo a partir dos crioulos de base inglesa que também constituem o campo original de estudo de McWhorter. Por exemplo, os crioulos de base ibero-românica têm participios e outras formas conjugadas dos verbos, inclusive formas de conjuntivo, e muitos interrogativos opacos (Bartens 1995; 1996). Achamos que as configurações tipológicas do próprio inglês, por um lado, e das línguas ibero-românicas, por outro, desempenham um importante papel nesse fenómeno.

do cabo-verdiano e do papiamento, considerados acrolectais),³⁵ a produtividade dos afixos é condicionada pela inexistência da possibilidade de emprestar derivados inteiros da língua lexificadora e pela intensidade do contacto com o adstrato. Neste sentido, o caso do chabacano é sumamente ilustrativo e põe em dúvida uma conclusão à qual seria fácil de chegar a partir dos inventários derivacionais dos demais crioulos estudados, ou seja, que pelo menos em matéria de derivação, estes crioulos seriam línguas neo-ibero-românicas.

Precisamente, excluindo o chabacano do estudo, esta conclusão seria imaginável tendo em conta que a formação de palavras constitui um campo muito conservador numa língua (contanto que alguma morfologia correspondente se conserve durante o processo de criouliização; cf. Lüdtke 1996: 236-237) e que são mais ou menos os mesmos afixos que se conservam em variedades extra-europeias das línguas ibero-românicas, não ou em menor grau reguladas por instâncias normalizadoras, como exemplificado pelo português de Angola, de Moçambique ou de Goa.³⁶ No espanhol americano que obviamente tem as suas instâncias normalizadoras, tem sido notável o aumento da produtividade dos sufixos *-ada*, *-ero/-era* e *-ción*. Por outro lado, só um afixo (possivelmente) emprestado a uma língua indígena, *-eco/-eca* (formação de nomes de lugares e seus habitantes, e adjectivos pejorativos) do nahuatl *-ecatl* é utilizado de maneira corrente no espanhol centro-americano e mexicano (Kany 1962: 85ss.; Buesa Oliver & Enguita Utrilla 1992: 238-246; Lipski 1994).

Por isso, repitamo-lo, é bastante elucidativo observar casos como o do chabacano, que demonstra claramente até que ponto um crioulo pode mudar sob a influência do contacto linguístico,

³⁵ Note-se, porém, que os inventários mais extensos postulados neste estudo não têm a dimensão dos inventários das línguas de super-, sub- ou adstrato, observação que já fizemos a propósito do chabacano.

³⁶ Endruschat (1994: 85-87) realça a produtividade dos afixos seguintes no português angolano: *anti-*, *auto-*, *des-*, *in-*, *inter-*, *não-*, *pós-*, *pre-*, *sobre-*, *sub-*, *-a*, *-ado*, *-dade*, *-dor*, *-eiro/-eira*, *-il/-al*, *-ismo*, *-ista*, *-(iza)ção*, *-izar*, *-sita*, *-ivo*. Döll (1994: 129-131) nota as formações análogas “errôneas” (ou seja, produtivas segundo a nossa interpretação) com os afixos *anti-*, *auto-*, *des-*, *inter-*, *mini-*, *não-*, *pós-*, *pre-*, *pró-*, *sobre-*, *sub-*, *-ante*, *-(iza)ção*, *-eiro*, *-ismo*, *-ista*, *-mento*, *-(íst)ico*, *-ível*, *-izar* no português moçambicano. Dalgado (1900b: 68-69) comenta o aumento dos usos de *-eiro* em detrimento de *-ista* no português regional de Goa.

evidenciando que seria muito precipitado concluir que os crioulos são apenas variantes das suas línguas-mãe (ou lexificadoras), como tem sido corrente assumir em certos círculos crioulisticos, de resto diametralmente opostos aos postulados de McWhorter.³⁷ A única alternativa para salvar aquela teoria seria manter que o chabacano não é um crioulo – e por agora não existem argumentos sustentáveis para afirmá-lo.³⁸

REFERÊNCIAS

- Arends, J.; Verhagen, J.; Van Lier, E.; Dikker, S.; Cardoso, H. 2006. On the presence versus absence of morphological marking in four Romance-based creoles. In: P. Bhatt; I. Plag (Eds.). *The Structure of Creole Words. Segmental, Syllabic and Morphological Aspects*. Tübingen: Niemeyer, 223-241.
- Bartens, A. 1995. *Die iberoromanisch-basierten Kreolsprachen: Ansätze der linguistischen Beschreibung*. Frankfurt/ Main: Peter Lang.
- Bartens, A. 1996. Interrogativa und verwandte Wortarten in den iberoromanisch-basierten Kreolsprachen. In: A. Endruschat & E. Gärtner (Eds.). *Untersuchungen zur portugiesischen Sprache (Beihefte zu Lusorama. 1. Reihe. 7.Band)*. Frankfurt am Main: TFM/Domus Editoria Europaea, 243-262.
- Bartens, A. 2000. Notes on componential diffusion in the genesis of the Kabuverdianu cluster. In: J. McWhorter (Ed.). *Language Change and Language Contact in Pidgins and Creoles. (Creole Language Library 21.)* Amsterdam/Philadelphia: John Benjamins, 35-61.
- Bartens, A. 2004. A comparative study of reduplication in Portuguese- and Spanish-based creoles. In: M. Fernández; M. Fernández-Ferreira; N. Vázquez Veiga (Eds.). *Actas del III Encuentro de ACBLPE*. Madrid/Frankfurt: Iberoamericana/Vervuert, 239-253.
- Bauer, L. 1983. *English Word-formation*. Cambridge: Cambridge University Press.
- Bickerton, D. 1981. *Roots of Language*. Ann Arbor MI: Karoma.
- Bickerton, D. 1984. The language bioprogram hypothesis. *The Behavioral and Brain Sciences*. 7: 173-188.

³⁷ Robert Chaudenson foi o defensor mais árduo dessa tese nos seus escritos mais antigos. Nos anos 1990, suavizou a sua posição relativamente a esta questão.

³⁸ Quer isto dizer que a convergência linguística teria que assumir dimensões ainda muito mais importantes para que o carácter crioulo do chabacano se oblitere. Como acima mencionámos (v. 2.5.), o crioulo de Ceilão do século XX é outro caso de um crioulo que tem experimentado um movimento de convergência importante em direção ao adstrato.

- Booij, G. 1993. Against split morphology. In: G. Booij; J. van Marle (Eds.). *Yearbook of Morphology 1993*. Dordrecht: Kluwer, 27-49.
- Booij, G.; van Marle, J. (Eds., 2003). *Yearbook of Morphology 2002*. Dordrecht: Kluwer.
- Booij, G.; Lehmann, C.; Mugdan, J. (Eds., 2004). *Morphologie: ein internationales Handbuch zur Flexion und Wortbildung/Morphology: an international handbook of inflection and word-formation. 2. Halbband/Volume 2*. Berlin: Mouton de Gruyter.
- Braun, M.; Plag, I. 2003. How transparent is creole morphology? A study of early Sranan word formation. In: G. Booij; J. van Marle (Eds., 2003). *Yearbook of Morphology 2002*. Dordrecht: Kluwer, 81-104.
- Brito, A. P. de (1967[1887]): Dialectos Crioulos-Portugueses. Apontamentos para a Gramática do Crioulo que se Fala na Ilha de Santiago de Cabo Verde. Revistas por F. Adolfo Coelho. In: J. M. Barbosa (Ed.). *Estudos Linguísticos Crioulos. Edição de artigos publicados no Boletim da Sociedade de Geografia de Lisboa*. Lisboa: Academia Internacional da Cultura Portuguesa, 329-404.
- Buesa Oliver, T.; Enguita Utrilla, J. M. 1992. *Léxico del español de América: su elemento patrimonial e indígena*. Madrid: Mapfre.
- Couto, H. H. do. 1994. *O crioulo português da Guiné-Bissau*. Hamburg: Buske.
- Dalgado, S. R. 1900a. *Dialecto indo-português de Ceylão*. Lisboa: Imprensa Nacional.
- Dalgado, S. R. 1900b. Dialecto indo-português de Goa. *Revista Lusitana*. 6: 63-84.
- Dijkhoff, M. 1993. Papiamentu word formation. Amsterdam: University of Amsterdam. Unpublished dissertation.
- Döll, C. 1994. 6. Portugiesisch in Moçambique. In: M. Perl et al. (Eds.). *Portugiesisch und Crioulo in Afrika. Geschichte – Grammatik – Lexik – Sprachentwicklung*. Bochum: Brockmeyer, 110-139.
- Endruschat, A. 1994. "5. Portugiesisch in Angola". In: M. Perl et al. (Eds.). *Portugiesisch und Crioulo in Afrika. Geschichte – Grammatik – Lexik – Sprachentwicklung*. Bochum: Brockmeyer, 66-69.
- English, L. J. 2005. *Tagalog-English Dictionary*. Mandaluyong City: National Book Store/ Congregation of the Most Holy Redeemer.
- Faraclas, N. 2003. The *-pela* suffix in Tok Pisin and the notion of 'simplicity' in pidgin and creole languages: What happens to morphology under contact?. In: I. Plag (Ed.). *Phonology and Morphology of Creole Languages*. Tübingen: Niemeyer, 269-290.
- Fries, N. 2005. Ableitung. In: H. Glück (Hrsg.). *Metzler Lexikon Sprache*. Stuttgart: J.B. Metzler, 5-6.
- Günther, H. 2005. Ableitungsaffix & Ableitungsbasis. In: H. Glück (Hrsg.). *Metzler Lexikon Sprache*. Stuttgart: J.B. Metzler, 6.
- Himmelmann, N. P. 2004. Tagalog (Austronesian). In: G. Booij; C. Lehmann; J. Mugdan

- (Eds). *Morphologie: ein internationales Handbuch zur Flexion und Wortbildung/ Morphology: an international handbook of inflection and word-formation. 2. Halbband/Volume 2*. Berlin: Mouton de Gruyter, 1473-1490.
- Kany, C. E. 1962. *Semántica hispanoamericana*. Madrid: Aguilar.
- Kouwenberg, S.; LaCharité, D. 2003. The meanings of “more of the same”. Iconicity in reduplication and the evidence for substrate transfer in the genesis of Caribbean Creole languages. In: S. Kouwenberg (Ed.). *Twice as meaningful. Reduplication in Pidgins, Creoles and other contact language*. London: Battlebridge, 7-18.
- Kouwenberg, S.; Murray, E. 1994. *Papiamentu*. München/Newcastle: Lincom Europa.
- Lefebvre, C. 2003. The emergence of productive morphology in creole languages: The case of Haitian Creole”. In: G. Booij; J. van Marle (Eds.). *Yearbook of Morphology 2002*. Dordrecht: Kluwer, 35-80.
- Lenz, R. 1928. *El papiamentu, la lengua criolla de Curazao. La gramática más sencilla*. Santiago de Chile: Balcells y Cia.
- Lipski, J. M. 1994. *Latin American Spanish*. London/New York: Longman.
- Lorenzino, G. 1998. *The Angolar Creole Portuguese of São Tomé: Its Grammar and Sociolinguistic History*. München: Lincom Europa.
- Ludtke, J. 1996. 106. Gemeinromanische Tendenzen ‘V. Wortbildungslehre. In: G. Holtus; M. Metzeltin; C. Schmitt (Eds.). *Lexikon der Romanistischen Linguistik. Band 1. Latein und Romanisch. Historisch-vergleichende Grammatik der romanischen Sprachen*. Tübingen: Niemeyer, 235-272.
- Maurer, P. 1995. *L’angolar. Un créole afro-portugais parlé à São Tomé*. Hamburg: Buske.
- Maurer, P. 1998. El papiamentu de Curazao. In: M. Perl; A. Schwegler (Eds.). *América negra: panorámica actual de los estudios lingüísticos sobre variedades hispanas, portuguesas y criollas*. Frankfurt am Main/Madrid: Vervuert/Iberoamericana, 139-217.
- McQueen, J.M.; Cutler, A. 1998. Morphology in word recognition. In: A. Spencer; A. M. Zwicky (Eds.). *The handbook of morphology*. Oxford: Blackwell, 406-427.
- McWhorter, J. 1998. Identifying the Creole prototype: Indicating a typological class. *Language*. **74(4)**: 788-818.
- McWhorter, J. 2005. *Defining Creole*. Oxford: Oxford University Press.
- Munteanu, D. 1991. *El Papiamentu, origen, evolución y estructura*. Bochum: Brockmeyer.
- Muysken, P. 2004. 153. Pidginization, creolization, and language death. In: G. Booij; C. Lehmann; J. Mugdan (Eds.). *Morphologie: ein internationales Handbuch zur Flexion und Wortbildung/Morphology: an international handbook of inflection and word-formation. 2. Halbband/Volume 2*. Berlin: Mouton de Gruyter, 1653-1661.
- Perl, M. et al. 1994. *Portugiesisch und Crioulo in Afrika. Geschichte – Grammatik – Lexik – Sprachentwicklung*. Bochum: Brockmeyer.

- Pires, J.; Hutchinson, J. P. 1983. *Disionariu preliminaru kriolu. Preliminary creole dictionary. Cape Verdean/English*. Boston: Funkul.
- Plag, I. 2001. The nature of derivational morphology in creoles and non-creoles. *Journal of Pidgin and Creole Languages*. **16(1)**: 153-160.
- Plag, I. (Ed., 2003). *Phonology and Morphology of Creole Languages*. Tübingen: Niemeyer.
- Plag, I. 2006. Morphology in Pidgins and Creoles. In: K. Brown (Ed.). *Encyclopedia of Language and Linguistics. Second Edition. Vol. 8*. Amsterdam : Elsevier, 304-308.
- Quint, N. 2000. *Grammaire de la langue cap-verdienne. Étude descriptive et comprehensive du créole afro-portugais des Îles du Cap-Vert*. Paris: L'Harmattan.
- Rougé, J.-L. 1988. *Petit dictionnaire étymologique du kriol de Guinée-Bissau et Casamance*. Bissau: Instituto Nacional de Estudos e Pesquisa.
- Siegel, J. 2004. Morphological simplicity in Pidgins and Creoles. *Journal of Pidgin and Creole Languages*. **19(1)**: 139-162.
- Smith, I. R. 1978. *Sri Lanka Creole Portuguese Phonology*. Trivandrum: Dravidian Linguistics Association.
- Smith, I. R. 1979. Convergence on South Asia: a creole example. *Lingua*. **48**: 193-222.
- Steinkrüger, P. 2003. Morphological processes of word formation in Chabacano (Philippine Spanish Creole). In: I. Plag (Ed.). *Phonology and Morphology of Creole Languages*. Tübingen: Niemeyer, 253-268.
- Veiga, M. 1995. *O Crioulo de Cabo Verde. Introdução à Gramática*. Praia: Instituto Caboverdiano do Livro e do Disco/Instituto Caboverdiano de Cultura.

Variação e mudança no vocalismo átono quinhentista: práticas escriturais e juízos normativos.

Maria Helena Paiva
mhelenapaiva@netcabo.pt

Faculdade de Letras da Universidade do Porto (Portugal)

RESUMO. No prolongamento de estudo recente (Paiva 2008) sobre o vocalismo átono quinhentista, a partir da descrição de gramáticos e ortógrafos coevos, que confirmou a elevação de <a> átono, prossegue-se a pesquisa sobre a elevação do restante vocalismo átono, com base no tratamento informático dos textos metalinguísticos da época, que permitiu o exame exaustivo das grafias, em conexão com os juízos de valor expressos. Para isso, descrevem-se o *statu quo* da variação anterior à relatinização, as consequências desse movimento quer na eliminação de esquemas variacionais anteriores, quer na produção de novos, delimita-se a mudança de <e>><i> e interpretam-se grafias inabituais com <e>, concluindo-se que esse tipo de variação gráfica é uma consequência de <e> representar já [i]. Quanto a [o], a deriva unidireccional que desde a Idade Média conduz a [u], e que a relatinização reforça, está claramente documentada e suscita juízos de valor negativos que não deixam dúvidas sobre a dinâmica em curso.

PALAVRAS- CHAVE. Variação e mudança. Vocalismo átono quinhentista. História do Português. Tratamento informático.

ABSTRACT. Following a recent study (Paiva 2008) of the 16th century unstressed vowels based on the descriptions of contemporary grammarians that confirmed the raising of unstressed <a>, I continue the research about the raising of the remaining unstressed vowels based on the computer processing of the metalinguistic texts of this period, which allowed the exhaustive examinations of the writings, in connection with the expressed value judgements. In order to achieve this, I describe the *statu quo* of the variation before the relatinisation, the consequences of this movement whether on the elimination of the previous schemata of variation or on the production of new ones; I delimit the change of <e>><i>, interpret non usual writings with <e>, and infer that this kind of writing variation derived from the fact that <e> represents already [i]. Regarding [o], the unidirectional drift that since the Middle Ages leads to [u] (and with the relatinisation strengthens) is clearly documented and raises negative evaluation judgements that reveal undoubtedly the ongoing dynamics.

KEY-WORDS. Variation and change. 16th century unstressed vowels. History of Portuguese. Computer processing.

1 – Introdução

A reconstituição do vocalismo átono quinhentista exige, como primeira fase, a análise crítica das informações dos gramáticos e ortógrafos coevos, cujas descrições constituíram o tema de uma indagação recente (Paiva 2008), baseada no tratamento informático do *Corpus Metalinguístico Quinhentista* (a partir de agora, *CMQ*, descrito no anexo): o confronto entre o explícito e o implícito no plano descritivo configura um corpo de certezas, mas também de hipóteses com graus de comprovação variáveis, que por isso devem ser testadas por recurso a outros meios de informação. É essa pesquisa que se continua agora.

Como a elevação do fone átono representado por <a> é um facto incontroverso que foi amplamente confirmado, fixar-nos-emos na problemática inerente aos restantes fones átonos: quanto à vogal átona representada por <e>, há uma concentração de indícios de que esta «não é nem [e] nem [i], mas [i]» (Paiva 2008: 218); quanto ao fone átono representado por <o>, a identificação, ainda que não constante, entre «o pequeno» e [u] ou [w] por Fernão d'Oliveira, comprova, entre outros factos, que «está em curso a elevação de [o] átono», o que suscita a avaliação da amplitude desse movimento (Paiva 2008: 219).

1 – *Statu quo* e relatinização

Dado que a variação <e> ~ <i> e <o> ~ <u> átonos está documentada ao longo de toda a história do Português e se projecta no estado de língua que os textos metalinguísticos de 1536-1540 reflectem, importa caracterizá-los relativamente a essas coordenadas, antes que o movimento de relatinização altere esse quadro.

A primeira condição para que esta indagação seja possível consiste em isolar, na linha contínua da relação latim - português, o que, no plano lexical, se situa no plano da importação linguística e não no da evolução, selecção esta que é efectuada em função dos traços fonológicos que distinguem as formas semicultas e cultas das formas vernáculas; simultaneamente, se o que importa é definir o estado de língua que se designa por Português quinhentista, é indispensável isolar o que, nesse caudal, em que formas cultas e semicultas penetraram ao

longo do tempo com ritmos diversificados, pode ser considerado como importação recente, susceptível portanto de caracterizar uma época. Avaliámos por isso o grau de modernidade do léxico relatinizante incluído no *QMQ*, retendo, como prováveis importações recentes, os vocábulos não documentados com forma culta ou semiculta senão a partir do século XV nas obras de referência; assume-se assim, no plano pontual, a margem de falibilidade do critério histórico, que se dilui, contudo, no plano de conjunto (Paiva 2002, IV: 22-26).

2.1 – <e> e <i>

A variação <e> ~ <i> pretónicos e postónicos, como refere C. Maia, está atestada em período antigo, mas ter-se-á intensificado «nos textos do século XV e XVI», em consequência quer da assimilação da átona à tónica, como em *pidirõ, seguinte*, (1986: 364; 367), quer de dissimilações como *deligencia, deziã, vezjão* (1986: 421). A este núcleo de variação agregam-se outros, como o fechamento de [e] antes de palatal, como *mjlhor* e *sinurio* ‘senhorio’ (Maia 1986: 367). Noutros casos, a forma exclusiva ou predominante contrasta com o português padrão actual por apresentar <e>, em consequência de vários factos, entre os quais a proveniência de Ī, como em *fegura*, ou a variação já existente em latim, como *dereito* < DĒRĒCTUM , que coexistiu com DĪRĒCTUM (Maia 1986:422).

Este tipo de variação persiste no *CMQ*: assim, quando na tónica existe [i] ou [u], <i> poderá resultar de inflexão vocálica, como em *conhicimento* (B:1 oc.) ou *testimunho* < TĚSTĪMŌNĪU- (B, constante: 2 ocs.) ou de dissimilação, como em *vezinho* < VĪCĪNU- (O, constante: 9 ocs.), *dezia*, *deziã* (O, constante no imperfeito: 3 ocs.), *derigida* de *derigir* < DĪRĪĜĒRE (B, 1 oc.), a par de 3 ocs. com <i>); antes de palatal, *milhor* é constante em Oliveira num total de 26 ocs.

O sistema de equivalências próprio do novo estrato lexical caracteriza-se pela regularidade, de tal maneira que pode estabelecer-se uma regra: quanto mais latinizante o Autor, menos são as formas afectadas por estes fenómenos; e mesmo uma sub-regra: quando tal não acontece, há aí uma pista a explorar. Assim, de entre os exemplos apresentados, são predominantes no *CMQ*: *conhecimento* (5 ocs., num total de 7), o imperfeito do indicativo de *dizer* com <i>, como

dizia (14 ocs., num total de 17), *melhor*, cuja variação, representada no *Quadro 1*, ilustra cabalmente os factos referidos: *milhor* era com certeza uma forma muito corrente, que reflecte a realidade oral ainda hoje audível; é a única usada por Oliveira e é largamente dominante em Barros, não obstante o seu pendor latinizante, patenteado por acolher já a forma mais próxima do latim, que é quase a única documentada na segunda sincronia, o que reflecte a consolidação da relatinização no tempo, tanto mais que a única ocorrência é uma forma considerada por Leão como «errada», e «emendada» para *melhor*.

QUADRO 1: MELHOR						
	T	O	B	G	LRT	LRI
	57	26	13	6	10	2
milh-	37	26	10	-	1	-
melh-	20	-	3	6	9	2

No âmbito da relatinização, situam-se na área da mudança por importação enquanto factos de adstrato, e não na área da mudança fonética por evolução, os novos padrões de equivalência entre os grafemas latinos e os grafemas do português quinhentista impresso: enquanto na evolução histórica tradicional, o resultado de Ĩ tónico ou átomo é [e], representado na escrita por <e>, a equivalência que se estabelece agora ignora a diferença entre Ī e Ĭ, e faz corresponder a ambos <i>.

Em importações semicultas ou cultas ao latim, encontra-se raramente, pelo menos no séc. XV, a equivalência Ĭ - <e>, sem dúvida porque os introdutores dos neologismos relatinizantes intuíram por comparação a relação contrastiva e a generalizaram. Assim, *agricultura* < AGRĬCULTURA está documentada como *agrecultura* (*Livro dos Ofícios*), mas só a forma com <i> ocorre no CMQ (B: 1 oc., LRT: 1); em *artifício* < ARTĪFĪCĪUM, a pretónica está representada por <e> na *Crónica de 1344* e no *Leal Conselheiro*, enquanto o vocábulo usado por Barros, num total de 4 ocs., apresenta as duas variantes: *arteficio*: 2 ocs., *artificio*: 2 ocs.; *província* < PRŌVĪNCIA, está representado por *provencia* em Fernão Lopes, *Crónica de D. Pedro*, mas nas 6 ocs. do vocábulo no CMQ, todas apresentam <i> (B: 2 ocs., LRT: 2; LRI :2).

Fenómenos assimilatórios ou dissimilatórios e a diferença entre a equivalência Ĩ - <e> ou <i> imbricam-se, mas têm como resultado final a uniformização por aproximação da forma latina, o que pode ser ilustrado pela variação dos resultados de DĪMĪNŪĚRE, representados no *Quadro 2*, que se distribuem por três esquemas, segundo evoluções prováveis : 1) -<e> - <e>- (*deme-*), por equivalência entre Ĩ e <e>, na segunda sílaba e assimilação na primeira; 2) -<e> - <i> (*demi-*), por equivalência entre Ĩ e <i> na segunda sílaba e dissimilação na primeira; 3) -<i> - <i> (*dimi-*) por identificação com a forma latina.

O quadro evidencia que a latitude máxima da variação é a que se encontra em Oliveira, único em que está presente o esquema -<e> - <e>- , o mais distante da forma latina; a forma intermédia -<e> - <i>- está também representada em Barros em conformidade com o ecletismo que o caracteriza, mas o predomínio da forma mais próxima do latim nos dois primeiros gramáticos anuncia o futuro, expresso pela exclusividade do esquema em LRT.

QUADRO 2: DIMINUIR				
	T	O	B	LRT
	14	8	4	2
-<e> - <e>-	2			
demenuyr		1		
demenuem		1		
-<e> - <i>-	3			
deminuyr		1		
deminuidas		1		
deminuye			1	
-<i> - <i>	9			
diminuir		1		1
diminuido		2		
diminuy		1		
diminuye			2	
diminuimos			1	
diminuem				1

Diminuir ilustra uma tendência de tanto maior amplitude quanto mais se avança no tempo, e que engloba , de entre os casos anteriormente referidos, a substituição de *fegura* e *dereito* por *figura* e *direito*

; nesse caudal, em que à variação <e> ~ <i>, sucede a exclusividade das formas com <i> <Ī, integram-se as numerosas importações recentes ao latim, como, entre muitas outras, *ânimo* < ANĪMUS, *atribuir* < ATTRĪBŪĪRE, *delicado* < DĒLĪCATUS, *dirigir* < DĪRĪGĒRE, *distinto* < DĪSTĪNCTUS, *estímulo* < STĪMŪLUS, *fácil* < FACĪLĪS, *imitar* < ĪMĪTĀRI, *ministério* < MĪNĪSTĒRĪUM, *permitir* < PERMĪTĒRE.

Na escrita, a variação <e> ~ <i> encontra-se em vias de extinção; a relatinização favorece as formas com <i>, que prevalecem, o que, numa perspectiva global, aumenta, na língua dos cultos, a percentagem dessas formas¹.

2.2 – <o> e <u>

A variação <o> ~ <u> anterior ou coeva do CMQ contrasta com a variação <e> ~ <i> por ter carácter unidireccional, no sentido de que o fonema a que as evoluções conduzem é sempre representado por <u>, tanto em posição pretónica como postónica.

O panorama geral, definido a partir do exame sistemático da questão em quatro obras de referência (J.J. Nunes 1956, E. Williams 1975, C. Maia 1986, R. Silva 1989), caracteriza-se pelos traços seguintes:

1º - Inexistência de fenómenos de dissimilação, no sentido <o> < <u> e no sentido <u> < <o>.

2º - Frequência elevada de fenómenos de assimilação de diversos tipos, tendo todos como resultado [u]: a) quando existe [ú] na sílaba seguinte: *custume* < *costume* < CONSUETŪDĪNE- ; b) antes de consoante palatal: *cunhado* < CŌGNATUS, *mulher* < MŪLIERE- ; c) antes de [i] ou [j] contíguo ou da sílaba seguinte: *muito* < MŪLTU-, *durmo* < DORMĪO, *cupro* < COOPERĪO.

¹ A análise da variação *primeiro* ~ *premeiro*, no conjunto da obra de Fernão d'Oliveira, confirma este juízo e matiza-o: o contraste entre a *Gramática* (em que predomina *primeiro* - 61,25 %, num total de 80 ocs.) e o resto da obra, não estudada exaustivamente mas em que se constatou o amplo domínio de *premeiro*, é esclarecido, por um lado, pela existência, na *Gramática*, de duas zonas compactas (*primeiro*, do início ao fl. 30, *premeiro*, do fl. 56 até ao fim (fl. 76), com uma única e elucidativa excepção); por outro lado, na zona intermédia em que as duas variantes convivem, definem-se redes de associações distintas e conotações específicas determinantes do rumo da mudança e das opções posteriores de Oliveira. cf. Paiva, no prelo: § 4. PRIMEIRO, uma encruzilhada de tendências.

3º - Incidência paradigmática, no âmbito do verbo, deste tipo de fenómenos de inflexão vocálica: *cuberto, descuberto*, a partir de *cubro*.

Uma parte das formas resultantes de assimilação está fortemente consolidada e é exterior às áreas semânticas e às esferas culturais onde o fenómeno de relatinização se desenvolve (como *muito* e *cunhado*). Mas são numerosos os casos em que é reposto <o> onde existia em latim Ō ou Ŏ: das 50 ocs. de *costume*, 47 apresentam <o>, sendo as 3 restantes, *custume*, usadas uma por Oliveira (1/37), outra por Barros (1/8) e a terceira por Leão (LRT: 1/1), o que deve ser interpretado como resultante de um hábito não só consolidado como difundido, visto que n' *Os Lusíadas*, das 19 ocs. do vocábulo, 4 são grafadas *custume(s)* (Cunha 1980); das 5 ocs. de *acostumar*, só uma é grafada com <u>, por Oliveira (*acostumamos*), mas *costumar*, que só ocorre na segunda sincronia, (5 ocs, em Gândavo, LRT e LRI) apresenta sempre <o> em posição pretónica. Ora se, anteriormente ao século XVI, estão documentadas formas com <u> nestes vocábulos, e se hoje é [u] átono que é articulado, não é admissível que, apesar do peso da tradição gráfica, tenham sido predominantemente os autores da primeira sincronia que usaram formas com <u>, se essa grafia não reproduzisse a forma oral.

Relativamente às formas decorrentes de fenómenos de analogia, as grafias acusam a expansão de [u]: em *cobrir*, num total de 7 ocs, 4 das quais arrizotónicas, 3 são grafadas com <u>: *cubrir, cuberto(s)*, em Barros; quanto aos compostos, *descobrir* e *encobrir*, respectivamente 5 e 3 ocs., só em Barros estão documentadas as formas do particípio passado *descuberto* (1 oc.) e *emcuberto* (1 oc.); exceptuando uma ocorrência de *cuberto* (55r 20), todas as restantes são metalinguísticas, e inserem-se em contextos que favorecem a depreensão de nexos formais. A relatinização manterá <o> em todas as formas arrizotónicas mas, também nestes casos, não podem subsistir dúvidas quanto ao facto de ser o fonema [u] que é representado em posição pretónica em *cobrir, coberto, descoberto* e *encoberto*.

Quanto a *cumprir* ~ *cumprir* < CŌPLĒRE (C. Maia 1986: 402-403), (5 ocs.), deverá ter-se em conta que a latência da forma latina COMPLĒRE pode entrar, quer a representação de [u] nas formas em que a vogal resulta de inflexão vocálica, quer a extensão analógica

às outras formas, com vogal átona ou tónica: em Oliveira encontra-se a forma que prevalecerá, *cumpre*, mas Barros usa, em ocorrências metalinguísticas, *compriste* (1 oc.) e *compre* (3 ocs.). *Comprimento* (3 ocs.: O.), *comprido* (4 ocs.: O., B., G., LRI.), *compridamente* (3 ocs.: O.) e *comprimeto* (3 ocs.), apresentam formas constantes com <o>.

A segunda grande área afectada pela relatinização é aquela em que se contrapõem <o>, resultado de Ū átono na evolução tradicional e nas formas semicultas, principalmente quando introduzidas durante o século XV, e <u>, que lhe sucede nas importações quinhentistas que se tornam dominantes e que são marcadas pela fidelidade às formas originárias. A observação do conjunto do panorama que o CMQ oferece permite constatar a existência de três conjuntos que, vistos pontualmente, não são necessariamente sucessivos, mas que perspectivados na globalidade, constituem três estádios na expansão da mudança: os extremos caracterizam-se pela inexistência de variação, o primeiro, em que Ū está representado por <o>, e o último, em que a correspondência Ū - <u> está consolidada, enquanto ponto de chegada do movimento que se revela no estádio intermédio em que convivem, como resultados de Ū, <o> e <u>.

O primeiro conjunto constitui uma sobrevivência de resultados regulares na relação latim vulgar – português ou de equivalências que caracterizaram os semicultismos do passado: assim *molher* < MŪLIERE-, de que foi acima referido o resultado com <u> por acção assimilatória da palatal, que contrasta com o quadro que o CMQ patenteia: das 26 ocs do vocábulo (O: 10, B: 11, LRT: 3, LRI: 1), todas são grafadas com <o>, à excepção de uma, *mulher*, que Leão inclui na lista de «palavras que a gente vulgar usa & escreve mal» e que corrige para *molher* (LRT 71r 14). Ou ainda *arroído* < a + RŪGĪTUS (1 oc. : LRT) ou *triângolo* < TRIANGŪLUM (1 oc.: O.).

Os segundo e terceiro conjuntos podem ser ilustrados pelos vocábulos em que figuram os resultados da terminação nominal latina –ŪLUS, -A, -UM, cujo inventário e distribuição foram representados nos Quadros 3 e 4.

No Quadro 3, *capitolo*, *titolo* e *vocabolo* prolongam a tradição anteriormente atestada. Nestes vocábulos, de elevado número de ocorrências, só Oliveira usa <o> com acentuado predomínio, de acordo com a sua concepção de que as importações ao latim, mesmo

cultas, devem harmonizar-se com o que é habitual, ainda que no plano simplesmente gráfico; embora integrado na primeira sincronia, Barros, já propulsor da aproximação formal ao latim, no total das 36 ocs. dos três vocábulos, usa <o> em 7 e <u> nas 29 restantes. Nos autores da segunda sincronia são usadas exclusivamente grafias com <u>.

QUADRO 3: -ŪLUS : Ū> <o> ~ <u>							
		T	O	B	G	LRT	LRI
CAPÍTULO		64	56	1			7
	capitolo	44	43	1			
	capitulo	17	10				7
	Cap.	3	3				
TÍTULO		10	1	6		3	
	titolo	7	1	6			
	titulo	3				3	
VOCÁBULO		115	21	29	15	17	33
	vocabolo	17	17				
	vocabulo	98	4	29	15	17	33

No *Quadro 4*, em que figuram todos os restantes vocábulos existentes no *CMQ* de que faz parte a mesma terminação, confirma-se a homogeneidade da segunda sincronia e a consolidação do rumo da mudança que o quadro anterior anunciava.

Difícilmente não se estabelecerão nexos mentais que favoreçam a igualdade das formas a partir da identidade ou da proximidade de conteúdos, que podem contudo ser contrariados por outro tipo de associações, como o carácter recente da importação e a maior capacidade desta de evocar o latim; daí que surjam combinações heterogéneas, como a que se documenta no passo de Barros «de cujo *titolo intitulos* a Cartinha» (1v 11). Barros usa <o> em *titolo*, forma semiculta documentada no séc. XIII que é, de resto, a grafia única que adopta para a palavra (cf. *Quadro 3*), mas <u>, em *intitular*, forma culta documentada no séc. XV.

QUADRO 4: ŮLUS > <u>						
	T	O	B	G	LRT	LRI
CÍRCULO	3			1	2	
CLÁUSULA	20		11		9	
DISCÍPULO	2		2			
FÁBULA	2	1	1			
INCRÉDULO	2				1	1
MAIÚSCULO	1			1		
MÁCULA	1					1
ORÁCULO	3		3			
PARTÍCULA	4				3	1
VÍNCULO	1				1	
VÍRGULA	3		1		2	

A quase inevitável consciência da parcial identidade das palavras *título* e *intitulamos*, e a existência em ambas de uma vogal átona, representada no primeiro caso por <o> e no segundo por <u> levanta a questão da intersecção entre relatinização e redução de <o> átono, o que corresponde à formulação da pergunta: quando a tendência crescente é para fazer corresponder a Ů o grafema <u>, grafias como *título* e *titulo* representam ainda respectivamente [o] e [u] átonos, ou deu-se já, ou pelo menos está em curso, a convergência num fonema único, que será necessariamente [u], e a variação tornou-se então puramente gráfica?

O facto de estarem documentados, em período anterior ou em textos coevos, vocábulos em que Ů está representado por <o>, mas que são grafados no CMQ exclusivamente com <u>, traduz o impulso que os gramáticos quinhentistas imprimem à tendência a eliminar a variação gráfica a favor de <u> e a consolidar a identidade entre Ů e <u>. Assim é que, relativamente aos latinismos remotos, ao núcleo derivacional de *omildade* (C. Maia, 1986: 397) <HŮMĪLITAS –ĀTIS e de *humilde* (R. Silva, 1989: 75), baseado em HŮMĪLIS, correspondem *humilde* (1 oc.: B) e *humilmente* (1 oc.: G); a *sepultura* < SEPŮLTURA, documentado no séc. XIII, corresponde *sepultura* (2 ocs.: O, B.) e a *foturo* < FŮTŮRUS, de que ainda há uma ocorrência n'Os *Lusíadas* (III, 132) a par de 9 outras com <u> (A. Cunha, 1980) corresponde uniformemente *futuro* (9 ocs: O: 2; B: 2; LRT:2; LRI: 3), embora o vocábulo seja sempre termo da nomenclatura gramatical.

Dos latinismos reputados recentes, o *CMQ* revela, como largamente dominante, a equivalência entre Ū e <u>, podendo, no caso dos verbos, a vogal ser também tónica, o que não acontece, contudo, nos casos apontados a seguir: *consultar* < CONSŪLTARE (1 oc.: LRT), *discutir* < DĪSCŪTERE (1 oc.: O), *impugnar* < IMPŪGNARE (2 ocs.: LRT), *juntura* < JŪNCTURA (1 oc.: B), *murmuração* < MŪRMŪRATIO – ŌNIS, *ocupar* < OCCŪPARE (6 ocs.: *acup-* ~ *ocup-*: O.: 2; LRT.: 2, LRI.: 2).

Mas provavelmente nenhum domínio será tão esclarecedor como o da prefixação, pela vitalidade que o processo conhece como meio de ampliação do léxico e porque são frequentes os prefixos que assumem duas formas, uma vernácula e outra relatinizante, sobrepondo-se assim dois estratos, conforme a época: o mais antigo, com <o>, e o mais recente com <u>.

Um dos exemplos mais elucidativos é o do prefixo proveniente da preposição e prefixo latino SŪB (cf. Ali 1966: 253), cuja forma vernácula *sô* é usada enquanto preposição por Oliveira - «quero que minhas obras se pubriquem sô o título de seu nome» (3 10-11; outra oc.: 59 10) – e cuja forma actual, relatinizada, *sub*, não está documentada no *CMQ*. Enquanto prefixo, os compostos apresentam variação entre <o> e <u>: internamente à mesma família lexical, como, grafados com <o>: *socessão* < SŪCESSĪO –ŌNIS (1 oc.: O) e *socessor* < SŪCESSOR –ŌRIS (1 oc.: O), mas com <u> *suceder* < SŪCCĒDĒRE (4 oc.: LRT); ou internamente ao vocábulo: *suposto* < SŪPPŌSĪTUS (3 ocs) : *soposto* (2 ocs: B) , *suposto* (1 oc.: LRT.); *subjunctivo* < SŪBJUNCTĪVUS (4 ocs.), que apresenta as variantes *sojuntivo* (1 oc.: O), *sujuntivo* (2 ocs.: B) e *subjunctivo* (1 oc.: LRT); ou ainda *suprir* < SŪPLERE (26 ocs.): *sopr-* : 11 ocs. (O: 1; B: 10); *supr-* : 16 ocs. (B: 6; LRT: 5; LRI: 5), merecendo destaque, uma vez mais, o facto de as duas variantes serem usadas por Barros.

A consciência da dupla possibilidade da composição é expressa por Leão quando, no capítulo em que se ocupa «Das letras que se dobrão nas dições» (LRT 38v e sgs), refere os compostos de «*Sub*, preposição [latina], ou a nossa *sob* [...] como *succorrer* ou *socorrer*», reafirmando implicitamente, mais adiante, a propósito das palavras «que dobrão .c.» a legitimidade das duas grafias «*succorrer* ou *socorrer*» (42r 23). Para alguém que tão firmemente adopta uma

atitude normativa e tão insistentemente pugna pela fixação da língua por aproximação do latim, a aceitação da dupla grafia indicia um acentuado grau de variação que, contudo, noutros casos não deixa de condenar, como veremos mais adiante.

A relatinização tem como consequências:

1º - A recondução ao grafema latino originário de formas resultantes de fenómenos assimilatórios (*custume* → *costume*) ou de analogia (*cubrir* → *cobrir*).

2º - A substituição da equivalência típica de formas semicultas Ũ = <o> pela equivalência própria de formas cultas Ū = <u>, e daí a coexistência, no período em estudo, de abundante variação no âmbito da mesma palavra (*capitolo* ~ *capitulo*) ou de palavras aparentadas (Barros: *titolo*, mas *intitular*).

3º - O reforço, que tende a tornar-se universal, da equivalência Ũ = <u>, e daí a substituição de latinismos remotos (*sepoltura* → *sepultura*) e a representação por <u> do fonema latino nos latinismos recentes (*ocupar*).

4º - A confluência das duas vias (vernácua ou semi-culta e culta) na produção de numerosas grafias duplas ao nível da palavra ou da família de palavras (*soprir* ~ *suprir*; *socessão* mas *suceder*).

5º - A aceitabilidade da dupla grafia em alguns casos (*socorrer* ou *succorer*) que, expressa por Leão, indicia uma amplitude disfuncional da variação.

Pode concluir-se como fortemente provável que, no âmbito da relatinização, nos numerosos casos em que há variação entre os grafemas <o> e <u>, o grafema <o> represente também ou já predominantemente o fonema [u].

3 – Delimitação da mudança <e> > <i>

Não pode confundir-se a normalização por generalização da equivalência Ĩ = <i> que se dá no âmbito da relatinização e que, dada a área sociolinguística que o movimento afecta, atinge principalmente a língua dos cultos, com a tendência, que Duarte Nunes de Leão refere e condena, para a articulação da vogal anterior pretónica como <i>:

Regra XV. Que guardemos a analogia, & ordem nos vocabulos derivados, & que não

variemos nelles. Porque dizem muitos, *rindeiro*, *vindeiro*, *vistido*, não respeitando aos primitivos. Porque se *renda* se escreve com .e. necessariamente, se *há* de screver assi *rendeiro*, que he seu derivado. E se dizemos *veste* & *vestimenta*, assi *vestir* & *vestido* & assi de *venda*, *vendeiro*. E como dizemos, *pelle*, tambem diremos *pelliteiro*, & *pellica*, & não *pillica*, nem *pilliteiro*. E assi como dizemos *pomo*, diremos *pomar*, e não *pumar*, como muitos dizem. E de *gemer*, diremos *gemido*, e não *gimido*. E como dizemos *pedir* de *peço*, diremos *petição*, e não *pitição*, *pedinte*, & não *pidinte*. E de *ferir*, diremos, *ferimento*, & *ferida*, & não *firimêto*, nem *firida*. E de *mealha*, diremos, *mealheiro*, & não *mialheiro*. E de *meço*, *medes*, *medida*, & não *midida*. E de *mento*, *mêtes*, *mêtira*, & não *mintira*: posto que tambem digamos, *minto*, & *mintes*. (LRT 57v 5-22)

Às correcções anteriores somam-se as que constam da lista «Reformação de algũas palavras que a gente vulgar usa & screve mal» (LRT 69v-71v), em que se incluem, além de algumas repetições, *cileiro*, *milhor* e *milhoria*, *tisouro* e as que figuram na errata de LRI, *pirigrinação* e *piriodos*. Nas formas rejeitadas projectam-se fenómenos originariamente diferentes: a evolução da vogal representada por <e> em hiato, no caso de *mialheiro*; a assimilação à palatal, em *milhor* e *milhoria*; a extensão para o nome, de fenómenos de inflexão vocálica que affectam o verbo, como em *vestir* – *visto* – *vistido*, *ferir* – *firo* – *firimento* e outras; mas o facto que engloba a maior parte das palavras é a assimilação da vogal átona representada por <e>, à vogal seguinte, tónica ou átona (*pirigrinação*, *piriodos*, *gimido*, *pillica*, *pilliteiro*, *pitição*, *pidinte*, *midida*, *mintira*). Só não se integram nos factos descritos *rindeiro*, *vindeiro*, *cileiro* e *tisouro* que podem contudo decorrer da expansão de uma tendência que é activada por circunstâncias particulares específicas e que, à excepção de *mealheiro*, *melhor* e *melhoria*, que reflectem fenómenos generalizados na linguagem oral, têm cariz popular. Em termos de geografia linguística, nada indica que a tendência seja geral. Não seria surpreendente que Leão, de origem alentejana, reproduzisse um traço já então, como hoje, próprio da sua região.

Não pode assim ser considerada provada a evolução geral do fone átono, representado por <e>, seja qual for a sua posição, para [i], como tem sido por vezes admitido (Paiva 2008: 203).

4 – <e> chamado mudo

A notação de <e> paragógico em cantigas líricas galego-portuguesas foi estudada por Celso Cunha (1982) que constatou a exclusividade do fenómeno depois de /l/ e /r/ - como na canção de Joan Zorro «El-rey de Portugale / barcas mandou lavrare» (246) - por exigências de carácter rítmico, e o relacionou com traços idiomáticos regionais do «Norte e Centro de Portugal» (261). O inventário do fenómeno por G. Massini-Cagliari no *Cancioneiro da Biblioteca Nacional de Lisboa* alarga a paragoge a /n/ (1998: 174) e corrobora o ponto de vista de C. Cunha de que «e paragógico é, de fato ‘um desenvolvimento da consoante anterior’» (Cunha 1998: 177).

É este o fenómeno que fornece a Barros o exemplo de «Paragoge»: «como se fáz nos rimãces antigos, que por fazerê cõsoante diziã, os que me soẽ *guardáre*, por *guardár*.» (BGR 34v 23-24)

Fenómenos de anaptixe junto de /l/ e /r/ são referidos por Williams (1975 :111), quer a vogal emergente seja *a*, como em *caronica* ou *e*, como em *fevera* e Nunes (1956: 156), que considera que as vogais acrescentadas são «em geral idênticas à que imediatamente as precede ou segue», como *marmelo*, *fevereiro* ou *coronica*.

Não obstante, os fenómenos documentados no CMQ que designamos de vivos, porque não são simples repercussões de fenómenos ocorridos no passado, dão-se apenas com a vogal representada por <e>, à excepção de uma ocorrência de *algaravio*, em Barros (7v 14), que corresponde ao esquema preferencial descrito por J.J.Nunes. C. Maia encontrou um caso de «anaptixe de *e* pretónico», *ferey* = *frey*, num documento português de 1419 (1986: 531) e, em documentos galegos dos séculos XV e XVI, várias ocorrências da 1ª pessoa do plural do infinitivo flexionado «em que a desinência número-pessoal vai precedida da vogal *e* », como *agardaremos*, *defenderemos* (1986: 757); formas deste tipo não ocorrem nos documentos portugueses estudados pela A. (1986: 758), que remete contudo para textos citados por Nunes (1956: 272, N.2), de Fr. Pantaleão de Aveiro e da *Regra de S. Bento*, onde o fenómeno ocorre também nas formas análogas do futuro do conjuntivo; assim, *tomaremos*, *veremos*, facto que o A. explica pela conservação da vogal das formas originárias do imperfeito do conjuntivo (1956: 300, N. 1), interpretação esta de que diverge F.

G. Gondar, que analisa o fenómeno no galego, e que considera que «a vogal e é de carácter paragógico» (Maia 1986: 759).

A hipótese da existência de uma vogal átona do tipo de [i] do português contemporâneo em *pesar* ou *face* é aventada por P. Teyssier em 1966, na escrita de Barros, com base em grafias como *letera* (*letra*), *leyes* (*leis*), ou *sereyes* (*sereis*) (1966: 189), mas não é retida na *História da Língua Portuguesa* (1982), em que situa o aparecimento do fone «na segunda metade do século XVIII» (Paiva 2008: 202-2003). A datação poderá ser recuada para o século XVII, como admite Marquilhas (2000: 261), com base em grafias como *outeras* (*outras*), *anderade* (*andrade*), *foremos* (*formos*) ou *martire* (*mártir*).

Passando ao exame directo do testemunho dos metalinguístas quinhentistas sobre a língua coeva, Oliveira faz menção, a propósito de palavras terminadas em <r>, de uma variante não generalizada, com vogal paragógica: «*alcaçer* por castelo o qual tem a penúltima grande, ainda que alghũs o pronũcião *alcacere. cõ .e. no cabo*» (OGR 36 5-6).

As práticas escriturais documentadas no *CMQ*, em que se encontram em variação formas que se opõem pela presença ou ausência de <e> configuram os conjuntos a seguir especificados:

1. Em Oliveira, em posição medial, depois de <r> e antes do sufixo número-pessoal *-mos*, na totalidade das formas da 1ª pessoa do plural do infinitivo flexionado e do futuro do conjuntivo: assim, as 2 ocs. do infinitivo flexionado (*daremos* e *teremos*) e as 20 ocs. do futuro do conjuntivo (*alcançaremos*, *dixeremos* (3 ocs.), *falaremos*, *ouveremos*, *foremos*, *meteremos*, *olharemos*, *poderemos* (3 ocs.), *poseremos* (2 ocs.), *quiseremos* (4 ocs.), *teveremos*, *trataremos*). Na 2ª pessoa do plural, só o futuro do conjuntivo está representado (5 ocs.), das quais só uma, de carácter metalinguístico, apresenta esse traço, como ilustração de uma forma hipotética originária da forma corrente do futuro, que resultaria da queda da vogal postónica: «Tambem somos amigos de cortar as vozes: onde se escrevem .l. ou .r. quando depois destas letras se avia d"escrever vogal como *syllba* por *syllaba*: & *fezertes* por *fezeredes* » (30 32 – 31 3).

2. Em posição final, nos nomes das letras mais directamente relacionadas com a paragoge de <e>, em Oliveira (*el*, *er* e *err* - 20 11), e Barros (*çle*, *çre*, *çrre* - 41r 3, 48r 16, 48r 18).

3. Em Barros, *letera* < LITTERA, que pode ser um latinismo: num total de 159 ocs. do lema, só 3 apresentam a forma *letra*, uma das quais, como exemplo de um tipo de «barbarismo», a «sincopa»: «como quando dizemos [...] *letra* por *letera*.» (34v 17-20); do mesmo modo, *leterado* (2 ocs).

4. Em Barros, após ditongo final e antes de <s>, em formas verbais de 2ª pessoa do plural: num total de 46 ocs., 40 são grafadas com <e> (86,95 %): indicativo presente: *amáyes* (2 ocs.), *folgayes* (1 oc.); imperfeito do indicativo: *amáveyes*, *lieyes*, *ouvieyes*, *ęreyes* (4 ocs.); mais-que-perfeito do indicativo: *amáreyes* (3 ocs.), *lereyes* (3 ocs.), *ouvireyes* (3 ocs), *foreyes* (3 ocs.) *tivęreyes* (4 ocs.); futuro do indicativo: *amareyes*, *lereyes* (2 ocs.); condicional: *amarieyes*, *lerieyes*, *ouvirieyes*, *serieyes* (4 ocs.); conjuntivo presente: *ameyes* (2 ocs.), *leáyes*, *ouçáyes* (2 ocs.), *sejáyes* (2 ocs.); imperfeito do conjuntivo: *amásseyes*, *lesseyes*, *ouvisseyes*, *fosseyes* (4 ocs.) Não contêm <e> 6 terminações (13, 04 %): indicativo presente : *quereis* (2 ocs.), futuro do indicativo: *avereis*, *ouvireis*, *sereis* (3 ocs.); conjuntivo presente: *leyaes* (1 oc.).

5. Em Barros, na terminação *-uye* da 3ª pessoa do indicativo dos verbos em *-uir*, no total das 6 ocs.: *concluye*, *destruye* (2 ocs.), *deminuye* / *diminuye* (3 ocs.).

6. Em Barros, a notação de plurais nominais por recurso a *-es*, em formas terminadas no singular por <y>: *leyes* alterna com *leis* (1 oc. cada uma), *reyes* (4 ocs.) com *reis* (1 oc.) e *páyes* apresenta uma ocorrência única, metalinguística, enquanto ilustração do plural de *páy*, do mesmo modo que *cõtray*, *contrayes*, nome de tecido. Se é certo que esse tipo de grafia pode ter sido influenciada pelo alomorfe de formação do plural em português, e pelo plural latino correspondente, a identidade entre essa terminação e as formas verbais referidas em 2. enfraquece essa hipótese, do mesmo modo que a grafia do numeral *seyes* < SEX (2 ocs.).

Nos três primeiros grupos, <e> ocorre em contexto com a vibrante simples e, no segundo grupo também com a lateral, prolongando, nas grafias do tipo de *daremos* = *dáremos*, grafias medievais anteriormente referidas. A hipótese de a letra representar o fone [e] acentua a distância entre estas variantes, excepcionais, e as variantes correntes sem <e>, enquanto a hipótese concorrente, a de a letra representar

[i], atenua as diferenças, que podem ser quase imperceptíveis. Embora possa haver diferenças importantes entre o vocalismo átono do português padrão contemporâneo e o do português quinhentista, a variação entre estas formas suscita a comparação com os resultados de testes perceptivos aplicados e descritos por R. Martins sobre pares caracterizados pela «presença / ausência» da vogal [ə] « nos grupos próprios de oclusiva + líquida», como *querer / crer, tabelado / tablado, perece / prece* (2002 : 176): «Dos 576 estímulos apresentados, 101 (17 %) foram substituídos pela palavra oposta » (2002: 180). Tendo em conta que os testes foram realizados em laboratório e que os testados foram estudantes universitários da Faculdade de Letras, em que o grau de consciência da existência de um fonema [i] cuja presença / ausência é distintiva, é amplificado pela sua inserção em pares mínimos e pela informação adquirida desde o início da alfabetização, compreende-se melhor que, quatro séculos atrás, a nível do discurso gramatical, tenha sido difícil identificar esse fone e conceituá-lo, distingui-lo de [e] e de [i] / [j] átonos e, como fez Oliveira, justificar a dificuldade da distinção invocando a proximidade entre essas vogais (Paiva 2008: 205-209).

Nos três últimos grupos, as sequências finais em que aos ditongos tónicos representados por <ay>, <ey> e <uy> se segue <e> suscitam a mesma pergunta, sendo que a articulação de [e] nestes contextos não pode deixar de ser considerada aberrante, enquanto a equivalência entre <e> e [i], se afigura compatível enquanto consequência de um fenómeno de paragoge; mas tal não é possível senão porque <e> tinha passado a representar [i], o fone mais semelhante ao prolongamento do fone anterior, e por isso não identificado; por outras palavras, a letra é, em sentido próprio, «muda». Em termos gerais, adequa-se-lhe, pelo menos em parte, o que A. Andrade considera próprio da vogal no português contemporâneo: « De entre todas as vogais, o ‘e mudo’ é indubitavelmente a mais instável, a mais sujeita a não ter manifestação fonética própria, o que levou à sua definição como um ‘chevá’ (schwa: e.g. Gonçalves Viana 1883)» (Andrade 1996: 303). As grafias apontadas em 6, como *leyes*, ainda que relacionadas com o alomorfe do plural, -es, são foneticamente tão estranhas como as anteriores, visto que afectam os mesmos ditongos. Quanto a *concluye* e afins (5),

ainda que <e> evoque a vogal final dos verbos em *-ir*, constitui, como nos casos anteriores um meio de prolongar o ditongo.

A representação de um fone que as grafias coevas habituais não registam acusa a influência da língua oral e sugere uma articulação particularmente lenta, o que confirmaria o juízo de Oliveira, é certo que formulado como uma das virtudes da língua portuguesa, de que «falamos com grande repouso como homens assentados» (5 2-3).

Do exame do conjunto das ocorrências ressalta a elevada frequência na associação entre conteúdos gramaticais e sequências gráficas que incluem este <e>, o que revela a estabilidade na relação entre o plano de conteúdo e o plano da grafia; o facto de esse reconhecimento se realizar no âmbito do verbo, campo extenso fortemente estruturado que favorece a captação de identidades, e o elevado grau de sistematicidade das grafias, que decorre do aspecto anterior, reforçam o valor comprovativo destas grafias. Deve ainda ter-se em conta que as formas de 2ª pessoa do plural reunidas em 4 figuram em percentagem muito elevada em paradigmas, a que corresponde tipicamente o grau máximo de consciência na escrita.

Consciência escassa, ou melhor, paraconsciência, como temos vindo a constatar, dadas as restrições de carácter heurístico (cf. Paiva 2005: 212) que se opõem ao reconhecimento desse elemento acústico nunca identificado por um sinal visual, vogal que é a que mais se parece com a inexistência de vogal, quer porque se confunde com o prolongamento de outra vogal ou semivogal, quer porque funciona como elemento de apoio a uma consoante; daí a variação que se estabelece entre <e> e um vazio gráfico, que por isso mesmo é informativo. Por isso Oliveira pôde escrever: «As letras consoantes aspiradas que são .ch. .lh. nh. não tem propria figura ainda até" gora: os nomes dellas são .che. lhe. nhe. os quaes sabidos são sabidas as pronúciações» 22 6-9). Ou seja, as escritas com <e> não representam nem [e], nem [i], mas sim [i], porque só assim «nome» e «pronúnciação» quase coincidem.

5 – Elevação de [o] átono

A elevação de [o] átono transparece em vários dos passos descritivos analisados (Paiva 2005: 207-214), é preconfigurada na direcção

convergente da variação <o> ~ <u> que desagua na língua quinhentista, cujo rumo a relatinização reforça tornando prevaletentes as variantes com <u> (cf. § 2.2.); finalmente é de esperar que do panorama variacional descrito emirjam efeitos valorativos e destes, os positivos incidam na articulação de <o> átono como [u]. É este quadro que o inventário das grafias do *CMQ* deve desenvolver. Na tipificação que estabelecemos a seguir, factos anteriormente referidos são incorporados sinteticamente numa perspectiva conclusiva e são aduzidos outros que os complementam ou se integram na exploração de pistas novas.

1. <o> arcaizante – Resultados etimológicos decorrentes da equivalência $\check{U} = <o>$, a conservação de grafias que não reflectem fenómenos fonéticos ou analógicos, originam, além de casos já relatados, outros factos de variação gráfica, em que subsistirá a variante com <u>, sendo que ou coexistiram as articulações [o] ~ [u], simétricas das grafias, ou a articulação de [u] precedeu a uniformização gráfica, de acordo com o característico atraso da grafia sobre a fonia. Assim, *fogir* ~ *fugir* < FŮGĚRE , de que há em Barros 2 ocs. não metalinguísticas de *fogir* (11r 23, 45v 15), a par de uma oc. de *fugir*, decorrente de inflexão vocálica em *fujo* e extensão analógica ao infinitivo e desencadeada pelo contexto («Compoëdo hum v̄rbo cõ outro dizemos mordefuje, de morder e *fugir*» - 9v 1-3). É ainda o caso de *logar* < LŮCĀLE-, que desde o português medieval alterna com *lugar* e que ilustra talvez uma tendência que remonta ao latim (C. Maia 1986: 406), de que há, em Barros 3 ocs. (BCAD 4 2, BGR 44R 16, BGR 44r 16), a par de 29 ocs. com <u>.

2. <u> inovador – Em palavras também documentadas no português medieval, resultantes de fenómenos assimilatórios ou analógicos e que será posteriormente substituído por formas restauradas: *custume*, *acustumamos*, *cubrir*, *cuberto*, *descuberto*, *encuberto*, anteriormente assinaladas, restauração que não atinge provavelmente o plano fónico, havendo assim continuidade parcial ou total entre a fase pré-relatinizante com variação entre [o] e [u] e a fase moderna e contemporânea, o que quer dizer que a <o> corresponde verosimilmente [u].

3. <u> - Em latinismos recentes ou em palavras cultas, em que pode contudo ter havido acção assimilatória da vogal da sílaba seguinte: *compustura* (B. 56v 2) a par de *compostura* (B: 3 ocs.); *intrudução* (B. 26r 11), a par de *introdução* (B. : 2 ocs.), *interlucuturia*, «emendada» por Leão para *interlocutoria* (LRT 70v 31), na «Reformação de algũas palavras que a gente vulgar usa & screve mal»; *Viturino*, o gramático latino (B. 40r 18, oc. única); *discurria* (B. 21r 12, oc. única). Também neste caso a restauração terá sido apenas gráfica, mantendo-se a articulação [u], como equivalente a <o>.

4. <u> + <m> heterossilábico ou homossilábico – Antes de <m> heterossilábico, num total de 479 ocorrências da palavra *como* em Oliveira, há uma ocorrência de *cumo* <QUÖMÖDÖ, lat. vg. QUOMO, que atesta a atonicidade da vogal: «& assi *cumo* os gregos tem isto» (35 28).

Em contexto com <m> homossilábico em representação da nasalidade, a variação está mais frequentemente atestada. Em Oliveira, num total de 254 ocs. do vocábulo, há 1 oc. de *cĩ*. A hipótese de se tratar de um latinismo involuntário neste defensor da vernaculidade é anulada pelo contexto: «til com .i. faz os nomes masculinos: como patim: & jardim & com .o. também como som & tom: *cĩ* .u. também sam masculinos: como hum, alghum, nenhum, & mais jejum & debrũ.» (65 27-29). É a contiguidade com <u> e palavras terminadas em «um» que desencadeia a forma, mas tal só é possível porque ela tem existência virtual, enquanto significante memorizado de usos orais com [u], relacionáveis com a variante *co* (22 ocs.), que apresenta desnasalação, e em que a vogal é, com elevada probabilidade, pelo menos em alguns contextos, [u]. *Corromper* < CORRŪMPĚRE apresenta <o> constante no total das 10 ocs. distribuídas regularmente ao longo do *corpus*, mas do adjectivo, as 2 ocs., em Barros, apresentam alternância: «e dizem que a lingua Hebræa, nam ę mais que Caldeu *corrupido*» (52r 19-20), «E ainda se póde crer, que estas vózes com antiguidade já devem ser *corrupidas*» (53v 14-15). Das 9 ocs. de *converter* < CONVĚRTĚRE (B: 5, LRT: 4), 8 são grafadas com <o>; a única que apresenta <u> figura em Barros, em alternância com <o>, em formas de infinitivo, usadas em definições: «Vęrbo autivo, ę aquelle que se póde *cĩverter* ao modo passivo» (18r 25-18v 1); «Vęrbo neutro (ę nós-

sa linguágẽ) será aquelle que se nã póde *cõverter* ao módo passivo» (18v 20-21).

5. <u> em derivados – A relação entre um lexema originário com <o> tónico ou átono e um lexema derivado com <u> átono constitui uma das confirmações mais claras da elevação da vogal.

«*Muella*» [«de ave»] é um das palavras incluídas na lista «Dos vocabulos que os Portugueses tem seus nativos» (LRI 107 19), embora a relação derivacional com *moer* não possa ser posta em dúvida.

Fugareiro é corrigido por *Fogareiro* (LRT 70v 26).

O carácter geral da redução da vogal é reconhecido por Leão, como uma regra de derivação, que não deixa no entanto de condenar: “E assi como dizemos *pomo*, diremos *pomar*, & não *pumar*, como muitos dizem.” (LRT 57v 13- 15).

No primeiro e terceiro caso, grafias com <u> estão documentadas em períodos anteriores: *muela*, no *Cancioneiro Geral* (Machado, 1990), 1516, *pumares* em documentos portugueses dos séc. XIV e XV (C. Maia 1986: 404). Quanto ao segundo, a palavra aparentada *fugueira* alterna com *fogueira* (1 oc. de cada uma), nos *Diálogos de São Gregório*, séc. XIV (Silva 1989:75).

6. <o> por <u>: a equivalência reconhecida e ultrapassada.

A grafia *continoa* = *continua*, 3ª pessoa do indicativo presente de *continuar*, que ocorre em Oliveira, anteriormente interpretada (Paiva 2008: 209-210), constitui, no *corpus* em análise, um caso não só único, mas aberrante, visto que [ú] é representado por <o>, mas por isso mesmo mais informativo porque, além do contexto, só uma equivalência generalizada entre <o> e [u] átono torna possível a ocorrência.

De facto, a redução de [o] átono tem como consequência o aumento da distância entre fonia e grafia, uma vez que a conservação das formas gráficas não é apenas o resultado da tendência ao imobilismo que caracteriza a escrita, mas praticar e promover uma escrita correcta é indissociável dos objectivos da *Gramática*, definida por Barros como «hũ módo çerto & justo de falár & escrever» (GR 2r 8-9); o conceito de «correcto», variável com os autores e o tempo, está longe de envolver, generalizada e coerentemente, a aproximação da realidade oral. Por isso, a interiorização da equivalência, em posição

átona de <o> a [u], necessária à leitura, tem como consequência que a imagem acústica virtual de que faz parte [u] átono, possa corresponder, no plano da escrita, <u>, como o atestam os pontos 2, 3, 4 e 5. Mas, ainda no plano da escrita, a estabilidade da equivalência tem também como consequência que, se [u] átono é representado por <u> e por <o>, a par de nexos associativos no âmbito do léxico e da gramática mais ou menos fortemente actuantes, só a memorização das formas gráficas permite adoptar a forma gráfica baseada no uso. Daí que sejam fortemente informativas grafias correntemente designadas de hipercorrectas e que são na realidade decorrentes do reconhecimento de uma regra não universal, a equivalência entre [u] átono e <o>, e da aplicação dessa regra para além dos limites determinados diacronicamente e fixados na sincronia coeva de quem escreve.

Curioso < CŪRIŌSUS apresenta o resultado regular [u], visto que proveniente de Ū, e como tal é grafado em 5 das 6 ocs. do vocábulo. A ocorrência restante, *coriósos*, figura em Barros (26r 10), alternando com *curiósos* (57v 4) e prolongando um uso documentado na *Crónica da Ordem dos Frades Menores* (1209-1285), em manuscrito do séc. XV (Machado 1990).

Paralelamente, o vocábulo só usado por Barros, *curiosidade* < CŪRIŌSITAS, apresenta regularmente <u> em 3 das 4 ocorrências, sendo a restante *coriosidádes* (4v 24). O vocábulo está documentado com <u> nos *Inéditos de Alcobaça* (Machado 1990), séc. XIV-XV.

Difícilmente alguma grafia poderá ser mais esclarecedora do que a do derivado de *curto*, *corteza*, que ocorre em LRI - «E nós por *corteza* da lingua a tudo chamamos emprestar» (42 7) e é corrigido, na «Errata», para *curteza*.

Como consequência da convergência em <u> de diversos factos, às vezes não é possível saber se o uso de <o> por <u> é um fenómeno deste tipo ou se resulta da equivalência ū > <o>, posteriormente substituído por <u>. Assim *guloso* < GŪLŌSUS, forma culta documentada em Gil Vicente como *golloso* (Machado 1990), mas cuja ocorrência em LRI (46 13), pode ilustrar antes a equivalência entre <o> e [u], hipótese que recebe apoio na relação com *gula* <GŪLA, documentada no séc. XIV (Cunha 1996).

Fogir pode ser também uma sobrevivência da forma proveniente de FŪĜĚRE anteriormente referida e ainda grafada com <o> ou a

representação de [u] por <o>, que Leão considera inadequada, visto que na lista «Reformação de algũas palavras que a gente vulgar usa & escreve mal» (LRI 70 22) preconiza a forma grafada com <u>, o que indica que o *bom uso* está desse lado.

Se, em aparente contradição, na mesma lista, Leão coloca *mulher* (71r 14) na coluna das palavras «erradas» e *molher* na das «emendadas» (única ocorrência com <u> dum vocábulo grafado com <o> nas restantes 25 ocs. do CMQ), é porque a inclui na continuidade das grafias em que se projectou a equivalência Ũ > <o>; a grafia com <u> é contudo um forte indício da redução da átona, reputada característica da «gente vulgar», como nos anteriormente citados *fugareiro* (LRT 70v 26), e *pumar* (LRT 57v 14) «como muitos dizem» em que Leão reconhece uma tendência generalizada que não deixa contudo de condenar.

6 – Conclusão

Os factos assinalados, conhecidos através da descrição, das práticas escriturais e de juízos normativos, revelam que a dinâmica geralmente situada nos séculos XVII e XVIII está em curso no séc. XVI.

O grau de generalização da redução de [e] átono a [i] e de [o] átono a [u] pode ser avaliado pela projecção destes fenómenos nas práticas escriturais não só de Oliveira, arguto observador do momento linguístico presente que valoriza na sua genuinidade, mas também de João de Barros, conservador, arcaizante e já propulsor da vaga relatinizante; a diferença de meios sociais a que pertencem comprova que não se trata já de variação confinada aos limites de grupos sociais. Finalmente, a projecção, 30 e 60 anos depois, na obra de Leão, da redução de [o] átono a [u], revela que o fenómeno desencadeia reacções de tipo normativo, o que, embora inerente à atitude do Autor, traduz uma reacção consciente à variante inovadora. Tanto no caso da elevação de [e] como na de [o] átonos, textos menos marcados pela normatividade inerente a gramáticas e ortografias podem ser mais informativos. Não obstante, a elevação de [o], mais facilmente detectável que a de [e], descreve um percurso que, analisado ao longo de um tempo extenso, confirma, no essencial, as fases e os factores condicionantes da expansão da mudança descritos por Labov

(1984: 178-80): a variação [o] ~ [u] dependente do contexto fonético, largamente documentada no português medieval e já orientada para a prevalência de [u], que fenómenos de analogia expandem, recebe, no século XVI, o influxo de outra corrente de origem culta que, acentuando a semelhança entre o português e o latim, por um lado reconduz grafias variáveis com <o> ~ <u> a um <o> etimológico muito provavelmente já representativo de [u] e por outro lado, ao estabelecer a equivalência Ū = <u> provoca num primeiro momento a intensificação da variação que, num segundo momento se reduz pela frequência crescente de <u>, que tende a tornar-se exclusivo. A expansão das variantes com <u>, que tem na origem um juízo de valor positivo e a prevalência dessas variantes que marcam usos prestigiantes, conflui na variação ancestral, favorecendo as variantes com <u>, anteriormente neutras ou conotadas negativamente.

REFERÊNCIAS

- Ali, M. S. 1966. *Gramática Histórica da Língua Portuguesa*. 6ª ed. Rio de Janeiro: Edições Melhoramentos.
- Andrade, A. 1996. Reflexões sobre o 'E mudo' em português. *Actas do Congresso Internacional sobre o Português*. Lisboa: Edições Colibri, II, 303-344.
- Cunha, A. G. da. 1980. *Índice Analítico do Vocabulário de Os Lusíadas*. 2ª ed. Rio de Janeiro: Presença.
- Cunha, A. G. da. 1996. *Dicionário Etimológico Nova Fronteira da Língua Portuguesa*. 2ª ed., reimp. Rio de Janeiro: Editora Nova Fronteira.
- Cunha, C. F. da. 1982. Sobre o eparagógico na épica e na lírica. *Estudos de Versificação Portuguesa (séculos XIII a XVI)*. Paris: Centro Cultural Português.
- Labov, W. 1984. *Sociolinguistic Patterns*. 3ª ed. Philadelphia: University of Pennsylvania Press. [1ªed.: 1972].
- Machado, J. P. 1990. *Dicionário Etimológico da Língua Portuguesa*. 6ª ed. Lisboa: Livros Horizonte.
- Maia, C. de A. 1986. *História do Galego-Português. Estado linguístico da Galiza e do Noroeste de Portugal desde o século XIII ao século XVI*. Coimbra: INIC.
- Marquilhas, R. 2000. *A Faculdade das Letras. Leitura e Escrita em Portugal no séc. XVII*. Lisboa: Imprensa Nacional.
- Mateus, M. H. M.; Martins, M. R. D. 2002. Contribuição para o estudo das vogais átonas [ə] e [u] no português europeu. *Biblos*. **LVIII**: 111-128. Reproduzido em: Martins, M. R. D. 2002. *Fonética do Português. Trinta Anos de Investigação*. Lisboa: Caminho, 169-186.

- Massini-Cagliari, G. 1999. A paragoge rítmica na lírica profana galego-portuguesa. *Actas do XIV Encontro Nacional da Associação Portuguesa de Linguística (1998)*. Lisboa: Colibri/APL. II, 169-182.
- Nunes, J. J. 1956. *Compêndio de Gramática Histórica Portuguesa (Fonética e Morfologia)*. 5ª ed. Lisboa: Livraria Clássica Editora [1ª ed.:1919].
- Paiva, M. H. 2002. *Os Gramáticos Portugueses Quinhentistas e a Fixação do Padrão Linguístico. Contribuição da Informática para o estudo das relações entre funcionamento, variação e mudança*. Vol. I: *Objecto e Método*; Vol. II: *Pré-edições*; Vol. III : *Índice Geral de Vocábulos, Índices Alfabéticos de Formas de Outras Línguas*; Vol. IV: *Conclusões*. Dissertação de doutoramento. Porto: Faculdade de Letras da Universidade do Porto.
- Paiva, M. H. 2008. A descrição do vocalismo átono quinhentista: linhas e entrelinhas nos textos metalinguísticos coevos. *Linguística*. **3(1)**: 197-221.
- Paiva, M. H. No prelo [Fernando Oliveira] *A língua da Gramática à luz do outros textos*. In: C. Morais (Org.) Fernando Oliveira, *Um Humanista Genial*. Aveiro: Universidade de Aveiro.
- Silva, R. V. M. e 1989. *Estruturas Trecentistas. Elementos para uma gramática do Português Arcaico*. Lisboa: Imprensa Nacional.
- Teyssier, P. 1966. La prononciation des voyelles portugaises au XVI siècle d'après le système orthographique de João de Barros». *Annali* [Nápoles: Istituto Universitario Orientale]. **VII(1)**: 127-198.
- Teyssier, P. 1982. *História da Língua Portuguesa*. Tradução de Celso Cunha. Lisboa: Sá da Costa [1ª ed. francesa: Paris: P.U.F., 1980]
- Williams, E. B. 1975. *Do Latim ao Português*. Trad. 3ª ed.: Rio de Janeiro: Tempo Brasileiro [1ª ed ingl.: 1938].

Descrição do corpus

Primeira sincronia (tratada exaustivamente)

1. Fernão d'Oliveira

Grammatica da linguagem portuguesa (1536) – (OGR)....23 538 oc.....35,7 %

2. João de Barros

Texto contínuo da “*Cartinha*” (1539) – (BCA)

Grammatica da lingua portuguesa (1540) – (BGR)

Diálogo em louvor da nossa linguagem” (1540) – (BDL) ...24 549 oc.....37,3 %

Total da primeira sincronia48 087 oc.....73,1%

Segunda sincronia (tratada por amostra aleatória de ¼ de cada texto)

3. Pêro de Magalhães de Gândavo

Regras que ensinam a maneira de escrever e

Orthographia da lingua Portuguesa [...] (1574) – (GRE)

“*Dialogo em defesa da lingua Portuguesa*” (1574) – (GDD)... 2 097 oc.....3,1 %

4. Duarte Nunes de Leão

Orthographia da Lingoa Portuguesa (1576) – (LRT).....8 698 oc.3,5 %

5. Duarte Nunes de Leão

Origem da Lingoa Portuguesa (1606) – (LRI) 6 684 oc....10,1 %

Total da segunda sincronia17 679 oc.26,8%

Total do corpus65 766 oc.

Nota – As citações das obras constantes do CMQ, são transcritas de acordo com as pré-edições, aligeirando-se os critérios então retidos (Paiva, 2002, I, p. 70-79). As principais alterações são as seguintes: 1) uso de <i> e <u>, <j> e <v> de acordo com as convenções gráficas actuais; 2) substituição de <ç> por <c> antes de <e> e <i>; 3) substituição de *s longo* por *s de dupla curva*, mas conservação da sua função de demarcação da palavra; 4) introdução de acentos gráficos apenas quanto estes são indispensáveis à inteligibilidade do texto, casos em que a vogal alterada aparece em itálico; 5) A abreviatura que pode ser dobrada como *pera* ou *para*, variantes que alternam em Oliveira, é representada como p^a. De entre os traços originários conservados, destaco: 1) o uso do til, generalizadamente considerado explicitamente pelos vários gramáticos-ortógrafos como símbolo da nasalidade das vogais, mais adequado do que <m> e <n>, letras que representam propriamente consoantes; 2) a geral conservação da palavra gráfica, sendo a alteração das delimitações originárias assinalada pelas convenções seguintes: [] – reunião de formas separadas: *a][proveitão*;][- separação de formas reunidas: *o][qual*; " - separação de formas reunidas quando há elisão: *d"agua*»; 3) a geral não intervenção em matéria de pontuação.

O primado da escrita¹

António Emiliano

ah.emiliano@fcsh.unl.pt

Universidade Nova de Lisboa (Portugal)

RESUMO: O artigo aborda a questão do 'primado funcional' da língua escrita como chave para o correcto entendimento do estatuto e papel da escrita e da ortografia em sociedades complexas como a portuguesa. Argumenta-se que a manutenção do "paradigma aristotélico" impede esse correcto entendimento. Discute-se o problema da relação entre escrituralidade e oralidade. O estilo geral do artigo é ensaístico, apologético e "prelectivo", com utilização mínima de notas, citações e bibliografia de referência.

PALAVRAS-CHAVE: literacia, escrituralidade, ortografia, Grafemática

ABSTRACT: This article submits that the consideration of the 'functional primacy' of written language is the key to understanding the status and role of writing and spelling in complex societies. It argues that the persistence of the "Aristotelian paradigm" makes that understanding impossible. The issue of the relation between writing and orality is addressed. The general style of the paper is argumentative and lecturelike with a sparse use of footnotes, quotations and references.

KEYWORDS: literacy, written language, orthography, spelling, Graphemics

More than any other single invention, writing has transformed human consciousness.

(Ong 1982: 78)

Linguística - Revista de Estudos Linguísticos da Universidade do Porto - Vol. 4 - 2009, pp. 111-131

¹ Este texto parte, através de total e completa refundição, de um pequeno artigo de opinião publicado, com semelhante título na publicação electrónica luso-brasileira *Revista Autor* (<http://www.revistaautor.com/>) em 1/8/2008. O objectivo principal é o mesmo, argumentar o primado da escrita, embora no texto que aqui se apresenta os argumentos tenham sido significativamente aumentados e acrescentados e diversas secções de carácter opinativo e estilo polemista sobre o Acordo Ortográfico de 1990 tenham sido removidas. Também neste texto se incluem, muito modificados e compactados, alguns excertos de uma secção do Capítulo 1 do meu livro *Fonética do Português Europeu: Descrição e Transcrição* (Lisboa: Guimarães Editores, Guimarães Universitária 1, 2009), dedicada à língua escrita.

1 – Nota introdutória

Este artigo não versa sobre a reforma ortográfica que o *Acordo Ortográfico da Língua Portuguesa de 1990* contém. No entanto, o *Acordo* e a controvérsia ortográfica que se tem desenrolado em Portugal desde 1986 nos *media* serviram de pretexto para as reflexões que seguem, dado que se tem lido e ouvido tanto de detractores como de promotores da nova ortografia portuguesa (alguns dos quais com formação universitária em Linguística) declarações que exprimem uma visão pouco informada ou actualizada do papel e estatuto da língua escrita em sociedades altamente textualizadas como a portuguesa. Muitos linguistas, e não apenas nacionais, persistem na afirmação da primazia da oralidade, referindo-se à língua escrita como se fosse um subproduto e como se a língua oral fosse a primordial e fundamental manifestação do saber linguístico dos falantes. No contexto da discussão em torno dos Acordos Ortográficos de 1986 e de 1990, por vezes conduzida de forma pouco rigorosa ou através posições que não relevam do plano linguístico, esse tipo de afirmações tem pretendido 1) mostrar que as mudanças introduzidas na língua portuguesa escrita são de pouca monta e 2) desacreditar cientificamente aqueles que se opõem a mudanças na ortografia de 1945/1973. Do lado dos opositores, tópicos como 1) a insistência nos estragos causados à língua e cultura portuguesas por causa da supressão das chamadas “consoantes mudas”, as quais materializariam a memória da origem latina da língua e o parentesco com outras línguas, ou 2) a oposição irreductível à reforma por razões de bairrismo nacionalista, também denotam fraca compreensão do que é uma ortografia no contexto alargado da comunicação linguística escrita. A ortografia nacional não é coisa de somenos importância, queira-se ou não alterá-la, pelo que vale a pena discuti-la em termos adequados, começando-se pelo *estatuto funcional da língua escrita* numa sociedade complexa como é a sociedade portuguesa contemporânea.

O longo parágrafo precedente serve apenas para mostrar como posições cientificamente pouco sólidas podem ser usadas para se debater uma reforma que terá, se/quando entrar em vigor, um impacto profundo na vida quotidiana de quem escreve em português europeu.

A perspectiva que se tem da escrita não é necessariamente, ou em todos os casos, uma questão do simples foro académico ou da especulação intelectual: pode ter efeitos reais e práticos na comunidade e no grafolecto da comunidade (a nível didáctico e a nível sistémico). Mas, porque este artigo não é, de facto, sobre o *Acordo Ortográfico da Língua Portuguesa de 1990*, as referências ao mesmo ficam por aqui.

2 – Primado funcional da escrita

A escrita é uma invenção, uma tecnologia (cujo uso depende de outras tecnologias):

Writing (and especially alphabetic writing) is a technology, calling for the use of tools and other equipment: styli or brushes or pens, carefully prepared surfaces such as paper, animal skins, strips of wood, as well as inks or paints, and much more. [...] By contrast with natural, oral speech, writing is completely artificial. There is no way to write 'naturally'.

(Ong 1982: 81-2)

A Humanidade falou durante muitas dezenas de milhares de anos antes de escrever. Os milénios que decorreram desde o surgimento dos primeiros sistemas de escrita são uma parcela ínfima da história da espécie humana. Historicamente, a fala precedeu a escrita e não vale a pena gastar mais espaço com o apuramento da factualidade desta proposição.

Também cada um de nós falou antes de escrever. A aquisição da linguagem resulta da exposição a enunciados orais e manifesta-se na produção de vocalizações e enunciados orais. Não há aqui também controvérsia possível.

A precedência onto- e filogenética da escrita relativamente à fala é facto assente, indiscutido e indiscutível.

O problema é apreciar-se a língua escrita apenas na dimensão linear da sua evolução ou aquisição, ignorando-se ou distorcendo-se a natureza da sua relação com a fala e, sobretudo, escamoteando a

sua dimensão funcional própria. Tal perspectiva limita muito, se não impede mesmo, uma abordagem adequada da comunicação escrita enquanto comunicação linguística.

É hoje consensual entre os especialistas que a escrituralidade 1) tem características substantivamente distintas da oralidade, as quais não resultam de, ou não se reduzem a, um putativo estatuto secundário ou subsidiário da escrita em relação à fala, e 2) não é uma transcrição da oralidade — na realidade, a escrita não é, nem pode ser, uma transcrição da fala, apesar de ser uma representação da língua. Sobre esta questão escreveu lapidariamente David Olson, naquela que é certamente uma das obras mais influentes das últimas décadas de investigação sobre língua escrita:

[...] writing is not the transcription of speech but rather provides a conceptual model for that speech. [...] Writing systems [...] do represent speech. But not in the way that is conventionally held. Writing systems create the categories in terms of which we become conscious of speech. [...] we introspect our languages along lines laid down by our scripts.

(Olson 1994: 89-90)

Esta perspectiva, absolutamente fundamental para o adequado estudo da língua escrita, e a que volto um pouco mais abaixo, inverte completamente os termos em que tradicionalmente se tem equacionado a relação entre escrita e fala e obriga a repensar a questão da “primazia” da língua oral sobre a escrita.

Se a escrita é uma tecnologia no seu aspecto material, uma coisa artificial cuja utilização depende de outras tecnologias e artefactos, como foi acima dito, é uma tecnologia radicalmente distinta de outras, pois possibilita a comunicação linguística através de um *medium* totalmente alheio à fala. É uma tecnologia de comunicação que, como proclamou Walter Ong, reestrutura a consciência e o pensamento. Ao fazê-lo, a escrita altera radicalmente a forma como uma língua é percebida pelos falantes alfabetizados, criando uma metalinguística (ou seja, possibilitando a introspecção linguística) e altera também a sociedade. Melhor: a escrita proporciona o surgimento

de modos novos de organização social assentes em relações e papéis sociais dependentes da prova escrita e da comunicação diferida.

Do ponto de vista do desenvolvimento do falante, a aprendizagem da escrita acarreta 1) aquisição (ou construção, como se preferir) de competências novas — as quais incluem, entre outras, a manipulação de instrumentos de escrita —, 2) estimulação de áreas do cérebro e estruturação sináptica que não são necessárias à “simples” aquisição da linguagem² e 3) despertar de consciência fonológica e lexical avançadas.

Este último aspecto é particularmente relevante para a compreensão do funcionamento dos sistemas de base alfabética, ditos imprópriamente “fonográficos”. É a aquisição da escrita que despoleta a consciência fonológica do indivíduo (muito reduzida à partida) e não o contrário. A própria noção de segmento, na qual assenta toda a análise linguística contemporânea, parece ser um resultado da escrita alfabética:

Our modes of thought are so strongly conditioned by our literate education that nothing seems more ‘natural’ than the apparently ‘real’ phonetic segments of speech, correlated as they manifestly are with the alphabetic symbols of the written language. Yet the view that such segmentation is mostly an imposed analysis, and not the outcome of discovering natural time-boundaries in the speech-continuum, is a view that deserves the strongest insistence.

(Laver 1994: 101)

Esta afirmação de um eminente foneticista não pode deixar de evocar uma proposição fundamental da investigação de Olson sobre a relação entre escrita e oralidade:

[...] because writing is ‘read’, it comes to be taken as a model for speech; whatever is represented in the script becomes an object of knowledge or awareness to the person literate in that script. Thus, to the alphabetically literate

² Como a investigação em Neurologia, Neurolinguística, Linguística Clínica e Psicolinguística tem vindo a mostrar, o domínio de um sistema de escrita resulta na activação de áreas do cérebro que, em falantes não alfabetizados, permanecem dormentes. Essa activação tem consequências cognitivas que ultrapassam o domínio e uso da língua escrita (mas não cabe aqui discutir essa interessantíssima questão).

person, speech is seen as composed of a string of phonemes represented by the letters of the alphabet.

(Olson 1994: 91)

Os dados e análise de vários estudos sobre o desenvolvimento da consciência fonológica durante a aprendizagem da lectoescrita, apresentados por Goswami & Bryant (1990), parecem sustentar de forma inequívoca esta asserção: a discriminação dos segmentos surge tarde no processo de aprendizagem da escrita e não parece desempenhar qualquer papel relevante nos primeiros estádios de leitura. Só em fase posterior e transitória as crianças desenvolvem e usam consistentemente estratégias de descodificação grafofonémica e, finalmente, quando atingem o estágio de leitores-escreventes fluentes, é a leitura holística (i.e. grafolexémica ou logográfica) que predomina. Se bem que este seja ainda tópico de discussão e investigação, diversos autores têm chegado a conclusões semelhantes em estudos sobre o processo de aprendizagem da escrita e da leitura em crianças em idade escolar de várias zonas do mundo ou sobre o processamento da escrita em indivíduos com perturbações da linguagem (escrita ou oral) ou com lesões cerebrais.

Porque a escrita foi uma invenção e é uma tecnologia, a língua escrita é adquirida de forma radicalmente distinta do processo natural de aquisição da linguagem, o qual tem lugar nos primeiros anos de vida do indivíduo (e em certos aspectos se inicia já *in utero*): a aquisição da língua escrita integra diversas aprendizagens de carácter perceptivo, cognitivo e motor, que nada têm a ver com a produção de fala.

Além disso, a literacia tem profundas implicações não só na integração social dos indivíduos, algumas das quais se destaca neste artigo, mas também na própria estruturação das comunidades, como já referi. Além disso, a escrita altera radicalmente também o estatuto das línguas nas comunidades em que se inscreve. Neste respeito, a existência de uma ortografia desempenha um papel fundamental: parte crucial da standardização e elaboração funcional de uma variedade linguística regional ou da textualização de uma língua nacional é o processo de formação, estabilização, difusão e imposição de uma ortografia.

Quando se estuda a escrita, seja numa perspectiva etológica, ecológica, antropológica ou linguística — e felizmente temos já décadas de saber acumulado sobre estas matérias, que muitos linguistas e especialistas em educação parecem ignorar — o que conta não é a sua dimensão histórica ou o seu carácter de “significante de segunda geração”, mas sim o seu *carácter funcional*.

Em termos simples, o estatuto que à escrita é atribuído numa sociedade altamente alfabetizada e textualizada — e, concomitantemente, o papel que ela desempenha numa sociedade desse tipo — é a real medida da sua importância. Do seu *primado*.

Se é indisputadamente certo que milhares de gerações de antepassados nossos falaram antes de haver sistemas de escrita e que todos falámos antes de saber escrever e ler, mais certo é que quando um falante aprende a escrever (bem) e a ler (bem) entra num ‘admirável mundo novo contruído com letras’ — aquilo que David Olson designou «the world on paper», expressão que dá, aliás, título a uma das suas mais conhecidas obras (Olson 1994).

O acesso ao ‘mundo vertido em papel’ pressupõe, como disse acima, alterações no cérebro e aquisição de capacidades motoras e cognitivas novas, seja para segurar num cálamo, numa pena ou num lápis, seja para “teclar” numa máquina de escrever, num computador ou num telemóvel, seja para reconhecer holisticamente padrões gráficomicos numa folha de papel ou num ecrã através de movimentos oculares rápidos e precisos. Entre um humano alfabetizado e um não alfabetizado há um fosso cognitivo abissal, sabe-se hoje.

A escrita é a ‘expansão externa da memória’ de cada um dos membros de uma comunidade alfabetizada e da comunidade inteira, como descreveu longa e detalhadamente Merlin Donald na sua obra fundamental *Origins of the Modern Mind*³ (Donald 1993): com a implantação da escrita e da literacia numa comunidade (mesmo que de

³ Trata-se de uma obra de referência que analisa o papel da escrita no desenvolvimento da mente humana, que foi assassinada numa inqualificável tradução portuguesa intitulada *Origens do Pensamento Moderno*, publicada incompreensivelmente pela Fundação Calouste Gulbenkian.

forma restrita), o acervo cultural (mítico, religioso, histórico, jurídico, tradicional, experiencial) da mesma já não tem que ser transmitido de boca a orelha de geração em geração, já não tem que ser constantemente aprendido e recriado no processo de memorização. Com a escrita, a memória da comunidade é virtualmente infinita e está sempre disponível, de forma fixa e imutável.

Esta dimensão da escrita, que a distingue irredutivelmente da fala (efémera e contingente por natureza), é a chave para a compreensão do real estatuto que àquela se deve atribuir no âmbito das disciplinas científicas que a ela se dedicam.

No seu tratado *Da Interpretação* (περί ἑρμηνείας), Aristóteles escreveu, nas considerações iniciais, que «as palavras faladas são símbolos da experiência mental e as palavras escritas são os símbolos das palavras faladas» e que «as experiências mentais que os sons da língua directamente simbolizam são idênticas para todos os homens, como são também aquelas coisas de que as experiências são imagens».

A perspectiva aristotélica condicionou, de forma negativa, o início da reflexão científica sobre a escrita no século XX. Ainda hoje subsiste esta visão simplista e datada do estatuto funcional e representacional da escrita, a qual impede uma abordagem conceptual e empiricamente adequada (e avançada) da escrita e dos seus problemas. Se queremos entender a escrita na sua especificidade estrutural, sistémica e funcional, temos necessariamente e definitivamente de abandonar o paradigma aristotélico. Se queremos compreender o que é e como “funciona”, de facto, uma ortografia, seja ela de base alfabética, silábica ou lexémica, é obrigatório e urgente abandonar a premissa aristotélica de que as formas escritas são simplesmente significantes de segunda geração ou de segundo grau.

3 – Nenhuma ortografia existe no vazio

A língua escrita não é, de facto, apenas um sistema simbólico de segunda ordem, como cria Aristóteles: a língua escrita, é, antes

de tudo, *língua*, e nenhuma tradição de escrita e nenhuma ortografia existe no vazio.

Muitos estudiosos da linguagem, nacionais e estrangeiros, alguns respeitáveis, ignoraram e ignoram que um enunciado escrito é, acima de tudo e antes de qualquer consideração, um *enunciado*, uma produção linguística. Reconheça-se e compreenda-se que diversos linguistas e filólogos trabalharam em época anterior ao desenvolvimento do estudo científico da escrita e que a sua concepção não fez mais do que reflectir o *status quaestionis* da sua época e o *Zeitgeist* vigente à data nos estudos linguísticos.

Os linguistas e filólogos da segunda metade do século XX e de inícios do século XXI (para não falar dos pedagogos e dos especialistas em didáctica e educação), esses, já não têm pretexto nem justificação para persistirem na visão aristotélica da escrita, a qual, posta a questão em termos chãos, nos conduz a um beco sem saída: se a escrita é simples subproduto ou representação da fala, então estude-se a fala e só a fala, porque tudo o que está na escrita decorre da fala e tudo o que possa ser apontado como específico da escrita não releva da esfera do linguístico. Se a escrita é subproduto da fala, nada há a dizer sobre a escrita que se não possa ou se não deva dizer em primeiro lugar sobre a fala.

Este é, efectivamente, um beco sem saída conceptual e epistemológico que impede a compreensão 1) da aquisição da escrita na sua especificidade cognitiva, 2) das características estruturais, sistémicas e comunicacionais próprias de um qualquer sistema de escrita, e 3) da relação complexa que existe entre escrituralidade e oralidade.

Algumas consequências práticas deste *cul-de-sac* são o não entendimento adequado 1) do surgimento das *scriptae* românicas autónomas na Idade Média (a partir da escrita romana da Antiguidade Tardia), 2) do estatuto da variação scriptolinguística em comunidades linguísticas e textuais desprovidas de um grafolecto standardizado e codificado (como foi Portugal até ao século XVIII), 3) da relação entre unidades do plano escritural (grafemas) e unidades do plano

expressivo (fonemas, morfemas ou lexemas), 4) do problema do erro ortográfico em ambiente de escrita normalizada e codificada, 5) do real impacto de uma qualquer reforma ortográfica na comunidade e nos escreventes individuais, 6) da estrutura própria e desenvolvimento histórico diferenciado da língua escrita face à língua oral em qualquer comunidade.

Antropologicamente, a literacia é um saber local (associado a determinadas práticas e *settings* comunitários) que se insere sempre num contexto social e cultural e que tem uma história, i.e., uma dimensão tradicional. Não há, nestes termos, um conceito universal de literacia, o que significa que é sempre problemática a comparação entre práticas localizadas de literacia, como é inconsequente a avaliação/análise de uma prática local à luz dos padrões de literacia vigentes noutra comunidade ou noutro tempo. Esta questão assume particular relevância para o estudo dos textos medievais latino-portugueses e antigo-portugueses, área em que tenho desenvolvido algum trabalho.

A aquisição da literacia altera de forma permanente e vinculada a inscrição dos falantes na comunidade e no mundo e acarreta uma mudança radical na estrutura das comunidades, nas relações sociais (presenciais ou diferidas) e no próprio mundo.

O ‘mundo em papel’ não é menos substancial ou ‘real’ do que o mundo natural: é um plano ontológico que condiciona a percepção e até a estruturação do mundo natural. De facto, sem literacia não há Estado, não há Direito, não há Civilização, não há Nação, não há Filosofia, não há Ciência, não há tecnologia industrial e não há memória social e cultural de longo prazo. Não há Humanidade como a entendemos hoje.

Sem escrita, o mundo que habitamos seria infinitamente mais pequeno, pois a troca de informação reduzir-se-ia à esfera das relações interpessoais presenciais. Sem escrita não haveria religiões do Livro nem haveria discurso científico e filosófico. Assim, reconhecer o primado da escrita é essencial para entendermos o mundo que

habitamos, que, para lá da realidade física e tri-dimensional em que nos movemos, é um mundo feito de letras, papel e *bytes*.

Ora, a ortografia fornece um sistema normalizado de codificação da memória externa da comunidade linguística: a sua estabilidade é um factor importante de coesão social e cultural, pelo que é “natural” (no sentido de expectável) que, em sociedades que atinjam um determinado grau de complexidade, surjam ortografias e normas linguísticas cultas que tendem a ser estáveis, para não dizer rígidas.

É nesse sentido que se pode com propriedade dizer que uma reforma ortográfica de grande impacto pode acarretar “lesões” para uma língua standardizada, sobretudo por duas razões: 1) porque a ortografia é parte integrante da língua (escrita), 2) porque uma alteração profunda da ortografia introduz descontinuidade cultural e instabilidade no sistema educativo. As lesões serão tanto mais profundas quanto maior for o grau de alfabetização e textualização da comunidade. Numa sociedade urbanizada e industrializada como a portuguesa, com um lastro gigantesco de produção textual e documental e com presença da escrita em todas as esferas da vida quotidiana dos cidadãos (como sucede em qualquer sociedade ocidental avançada), uma reforma ortográfica de grande impacto é inelutavelmente um “terramoto” cultural, social e educacional, independentemente das características intrínsecas da reforma.

* * *

A língua não é primeiro oralidade, e só depois escrituralidade, como defende a “premissa representacional”, chamemos-lhe assim, do paradigma aristotélico. Escrituralidade e oralidade são, na realidade, *media* autónomos de manifestação do saber linguístico.

Mais. Numa sociedade como a nossa, com omnipresença da escrita — vivemos rodeados de letras e de texto, nos espaços públicos e privados, nos objectos mais insignificantes que usamos ou possuímos, na roupa que vestimos, nos contratos que celebramos, no dinheiro que gastamos, etc. — a escrituralidade influencia, condiciona e pre-

cede a oralidade. Quando fomos para a escola aprender a “língua materna” não fomos reaprender algo que já trazíamos de casa (a nossa língua funcional), fomos aprender uma realidade linguística nova, a realidade do grafolecto e da norma padrão ou norma culta, a qual é elemento identitário cultural e nacional da maior importância.

Historicamente, um grafolecto resulta da transmutação, através da escrita, de uma variedade linguística regional, por razões de ordem política e social: um dialecto transmutado em grafolecto adquire o estatuto de língua trans- e supradialectal, com poder e impacto comunicacionais de ordem substantivamente distinta dos dialectos não desenvolvidos funcionalmente. Com a disseminação em massa de textos, proporcionada pelo advento da imprensa, os grafolectos impressos do Ocidente tornaram-se variedades linguísticas de natureza profundamente distinta das variedades que não ultrapassaram o nível da oralidade ou mesmo da escrita quirográfica. São fenómenos linguísticos sem paralelo na história multimilenar da comunicação linguística escrita.

Sincronicamente, é o grafolecto que permite a existência e subsistência de uma sociedade como a nossa. A nossa dependência, completa e quotidiana, de textos e de actos de escrita, traduz-se na impossibilidade de factos simples, como transacções financeiras, actos jurídicos e políticos, actos médicos, aquisição de informação, bens, medicamentos e serviços, produção de prova sobre toda a sorte de coisas (desde a nossa identidade até a um mero aviso de recepção de uma carta ou recibo de despesa), etc., sem recurso à língua escrita normalizada em vigor. Se nada disto é possível nos nossos dias e no nosso mundo sem língua escrita, sem ortografia — garante incontornável da estabilidade do grafolecto e do acervo comunitário de informação que o mesmo codifica — não há continuidade cultural intergeracional estável. Uma mudança ortográfica ou a substituição de uma ortografia implicam sempre um sobressalto ou interrupção no *continuum* geracional. Nada de semelhante se pode apontar à diversidade linguística que se manifesta na fala ou pela fala de indivíduos provenientes de regiões ou estratos sociais distintos.

4 – Falando da ortografia portuguesa

Falar da ortografia portuguesa, um bem que levou 700 anos a estabilizar-se, como se fosse coisa pouca (a estabilização da nossa fronteira política continental levou cerca de 100 anos), e falar de uma mudança ortográfica como uma simples alteração cosmética do sistema linguístico padronizado de uma nação multissecular dotada de um património literário e textual imenso, é simplesmente não se saber do que se está a falar.

Um excelente exemplo de incompreensão aplicada vem-nos da boca de Carlos Reis (autor e co-autor de diversos documentos “estratégicos” para a “promoção” da língua portuguesa) em entrevista de imprensa recente. É um exemplo que se justifica aqui comentar, não apenas porque o epíteto de “filólogo” é frequentemente associado a Reis, ou porque Reis se tem pronunciado publicamente sobre determinadas características que vê como negativas da “pronúncia” lisboeta, ou mesmo porque Reis é um dos principais defensores e promotores do *Acordo Ortográfico da Língua Portuguesa de 1990*; o exemplo em apreço é relevante porque reflecte uma perspectiva “pré-copernicana” sobre a língua escrita defendida por alguém que é, supostamente, um especialista.⁴

Na senda de comentários publicados noutras ocasiões e noutros lugares, Reis vituperou algumas características do português europeu em entrevista que deu ao semanário *Expresso*:

Há aspectos do português do Brasil em que leva vantagem sobre nós. Um deles é a articulação: um filme português passado no Brasil precisa de ser legendado e um brasileiro em Portugal não. Tão simples como isto. Porque os portugueses tendem a obscurecer a língua do ponto de vista articulatório, fonológico, de pronúncia. Engolem as palavras [sic]. O português do Brasil valoriza mais as vogais, os fonemas vocálicos, o que é uma vantagem para o bom entendimento. Temos de fazer um esforço de recuperar coisas que se

⁴ Carlos Reis, professor catedrático da Faculdade de Letras da Universidade de Coimbra e actualmente Reitor da Universidade Aberta, é um ilustre especialista em literatura portuguesa, nomeadamente, Eça de Queiroz.

perderam e isso só pode ser feito na escola, lendo expressivamente, obrigando a pronunciar bem as palavras todas. Mas a importância do Brasil neste cenário é evidente desde algum tempo.

(Carlos Reis, entrevista *Expresso*, 5/7/2008; o 'sic' é meu)

Deixando de lado a terminologia e formulação muito deficientes, o que importa notar é que Reis propõe o raciocínio seguinte:

1) o português europeu tem um *deficit* de vogais;

2) o português veracruziano é melhor do que o de Portugal porque tem mais vogais;

ergo,

3) deve-se restaurar as vogais perdidas do português europeu forçando as crianças portuguesas a ler “expressivamente”, ou seja, a soletrar, para se melhorar a língua em Portugal.

O que tem isto a ver com ortografia ou com língua escrita? É simples.

O nosso sistema ortográfico assenta em convenções grafonémicas que se fixaram a partir do século XIII. Desde o século XIII, época da emergência do ‘portuguez lingua escripta’ (expressão feliz do grande Adolpho Coelho, fundador da linguística portuguesa), mudanças portentosas afectaram a fonologia da língua. No entanto, a escrita, conservadora por natureza e por necessidade, não registou praticamente nada do que aconteceu desde essa época remota. Ora, precisamente dois dos fenómenos mais marcantes da fonologia do português europeu, e que a ortografia não assinala, são 1) o processo de elevação e recuo do vocalismo átono (através do qual as vogais das sílabas inacentuadas perdem abertura e duração) e 2) o processo subsequente de supressão das vogais átonas não-avançadas altas [i] e [u] (o qual resulta na supressão de sílabas, com conseqüente redução do corpo das palavras e alteração profunda do ritmo da língua). O início do chamado ‘processo do vocalismo átono’ é antigo, pois afectou de

forma semelhante as vogais finais do português europeu e do português veracruziano. Na realidade, pode-se fazer remontar o processo a época muito anterior à fundação da Nacionalidade, dado que as sete vogais do latim vulgar tardio ([i], [e], [ɛ], [a], [u], [o], [ɔ]) se reduziram, como se sabe, em posição átona não final a cinco ([i], [e], [a], [u], [o]) e em posição final a três ([e], [a], [o]). Enquanto as variedades veracruzianas confinaram o fechamento extremo às vogais finais, conservando em posição pretónica e postónica não final o vocalismo do latim vulgar (não se diferenciando neste respeito do galego e do castelhano), o português europeu, na sua deriva própria, estendeu o fechamento (elevação) a todas as vogais de sílabas inacentuadas.

No português cis- e transatlântico, as vogais finais de *bata*, *bate* e *batô* são mais fechadas do que as vogais tónicas de *bata*, *berço*, *bela*, *bolo* e *bola*, grafadas com as mesmas letras A, E e O.⁵ No entanto, o português europeu e o veracruziano têm vogais muito distintas nas sílabas átonas iniciais de *falar*, *levar* e *tomar*. O facto de não ter havido necessidade de ajustar a ortografia europeia à redução das vogais átonas, mantendo-se as mesmas formas ortográficas dos dois lados do Oceano — apesar de corresponderem a formas fonéticas divergentes — mostra bem a inexistência de um “princípio fonético” na ortografia portuguesa.

O que Carlos Reis propôs no seu depoimento não é mais do que uma regressão fonológica que apagaria quatrocentos ou quinhentos anos de mudança linguística do português europeu. Ora, como sabemos, não há línguas mais evoluídas nem línguas mais primitivas, línguas mais simples ou línguas mais complexas, línguas intrinsecamente melhores ou piores. Não há mudanças linguísticas más nem boas, como não há sistemas vocálicos bons, maus, vantajosos ou desvantajosos. Todas as línguas mudam a cada momento e de forma sistémica: não evoluem nem involuem, e não melhoram nem pioram. “Reabrir” as vogais fechadas do português, admitindo que tal seria possível e

⁵ Na pronúncia coloquial europeia (lisboeta, sobretudo) as formas *bate* e *batô* são produzidas com vogais finais surdas ([ˈbatʃ], [ˈbatʃ]) ou com supressão das mesmas ([ˈbat], [ˈbat]).

vantajoso, implicaria obrigar os falantes europeus a usar um sistema fonológico arcaico que lhes seria totalmente estranho.

A proposta absurda de Reis ilustra, apesar de tudo, o facto de a imagem gráfica das palavras estar “gravada a ferro e fogo” na mente e condicionar a percepção da língua falada, mesmo por um “especialista”.

Dito de outra forma, a aquisição da literacia envolve, entre outras coisas, a construção de representações grafémicas mentais, as quais, por estarem associadas às representações lexicais e fonémicas mentais, impedem o sujeito linguístico não treinado de estabelecer com sucesso a distinção explícita entre o escrito e o oral. É o que sucede com o falante Reis. De facto, não é por se ter, em português, palavras numerosas grafadas com -A, -E e -O finais que se justifica que se altere o modo de prolação das vogais átonas finais. A língua oral e a ortografia têm ritmos distintos de mudança, porque são códigos de distinta natureza e carácter, o que resulta numa relação complexa entre escrituralidade e oralidade, a qual não é facilmente acessível por introspecção ou sem treino metalinguístico.

5 – Relação entre unidades gráficas e fonéticas

As imagens ortográficas das palavras (representações grafémicas) são *in primis* representações mentais “impressas” na mente dos falantes alfabetizados a partir dos estádios iniciais de aquisição da literacia. Assim, dada a natureza grafofonémica de uma ortografia de base alfabética, é fácil ter-se a ilusão de que a escrita representa os “sons da fala”.

Para se entender adequadamente uma ortografia de base alfabética é necessário distinguir alguns conceitos e níveis de análise. Em primeiro lugar, importa distinguir o plano gráfico do *plano grafémico*.⁶

⁶ Para uma discussão detalhada e ilustrada destes termos e conceitos e dos que se seguem nesta secção v. Emiliano (2009), esp. Capítulo 1.

Em termos graféticos, a escrita é um conjunto de caracteres, um *alfabeto*. Nesta perspectiva não é necessário considerar a relação que os caracteres têm com as unidades fónicas de um sistema linguístico particular. É possível analisar e descrever um conjunto de caracteres sem referência a uma língua ou a uma ortografia particular. Um *carácter* é um símbolo gráfico de escrita considerado abstractamente, independentemente de uma manifestação gráfica concreta: é uma *forma ideal* sem propriedades concretas para além das que a definem como absolutamente distinta de todas as outras entidades do mesmo tipo. Um carácter pode ter uma multiplicidade de manifestações, *i.e.* de *glifos*. Um glifo é um modelo específico ou pauta para a materialização de um carácter num suporte de escrita; é uma entidade abstracta que contém informação precisa sobre características da realização gráfica. Uma *letra* do alfabeto português é uma *classe de caracteres*: a letra A, por exemplo, é o conjunto dos caracteres ‘A maiúsculo ou versal’, ‘A versalete’, ‘A minúsculo’ e ‘A sobrescrito’. A materialização da escrita num suporte por um escrevente ocorre no plano da *implementação gráfica* (ou grafética). Corresponde a eventos espaço-temporalmente localizados, logo, únicos e irrepetíveis, em que os glifos são materializados por grafos: como se vê, a distância conceptual entre ‘letra’ e ‘grafo’ é considerável, e a indistinção entre estes e outros conceitos acima referidos é fonte de alguns equívocos a respeito da estrutura, uso e funcionamento de uma ortografia (ou de uma *scripta*, no contexto da escrituralidade medieval).

O plano grafémico da escrita (objecto da Grafemática) exige, por seu lado, a consideração da relação entre unidades de escrita e unidades linguísticas, ou seja, é um plano de análise (cripto)linguística: se a afirmação ‘A maiúsculo é um carácter do alfabeto romano tipográfico moderno’ é verdadeira, a afirmação ‘A é um grafema’ só adquire conteúdo de verdade ou falsidade em relação a um determinado sistema linguístico. Por outras palavras, uma proposição como ‘X é um grafema’ é desprovida de sentido, é irracional. ‘Grafema’, tal como ‘fonema’, é um conceito relacional, pelo que as unidades de tipo “-émico” só se podem definir e caracterizar em função de, ou em contraste com, outras unidades do mesmo tipo. Os grafemas são as

unidades mínimas de um sistema de escrita⁷ e podem ser constituídos por um ou mais caracteres. Tal como um fonema se pode realizar através de fones distintos (alofones) de acordo com o contexto fonológico, também um grafema se pode manifestar em *grafias* ou *grafos* distintos (alógrafos), embora os factores que condicionam a alografia sejam distintos dos factores de alofonia. Na realidade, há *dois tipos de alografia*: tipo 1) um grafema pode apresentar alógrafos constituídos por caracteres distintos (*alografia profunda*), como é o caso das maiúsculas e minúsculas; tipo 2) os alógrafos são manifestações do mesmo carácter, i.e., são apenas glifos graficamente similares, permutáveis ou dependentes de contexto (*alografia superficial*), como sucedia na escrita medieval e sucede ainda hoje em alguns usos quirográficos particulares. Este último tipo de alografia é formalmente semelhante à alofonia da língua falada.

* * *

Uma ortografia é um sistema de escrita normalizado e codificado explicitamente: legiferado ou não, é necessariamente um sistema rígido imposto, perante o qual os escreventes não têm possibilidade de escolha. A ortografia assenta nos conceitos de norma, correcção e erro, e exclui liminarmente a possibilidade de usos gráficos particulares ou regionais. No âmbito de uma ortografia, um desvio (uma forma heterográfica), é sempre um erro.⁸

Que relação existe entre língua escrita e língua oral no âmbito de uma ortografia de base alfabética? É uma questão complexa, ou melhor, é uma questão simples que envolve resposta(s) complexa(s). Não há ainda consenso, apesar da abundante investigação, muita

⁷ «Entende-se [...] por sistema de escrita o conjunto de elementos gráficos que, associado a regras de concatenação dos mesmos e a regras de mapeamento entre aqueles e determinados elementos linguísticos (fonemas, sequências de fonemas, sílabas, morfemas, lexemas, prosodemas, etc.), permite a representação de um sistema linguístico num suporte de escrita e, logo, a comunicação linguística escrita.» (Emiliano 2009: 69)

⁸ Em Portugal, só a partir de finais do século XVIII se pode falar do surgimento de uma ortografia: a constituição de uma ortografia, processo associado à codificação de uma norma padrão, é normalmente um processo de longa duração (exceptuam-se os casos de ortografias criadas nos séculos XIX–XX por planificação linguística explícita).

produzida por portugueses (em Portugal e alhures), sobre o modelo mais adequado das relações complexas entre escrituralidade e oralidade na mente. Vários modelos têm sido propostos e todos se caracterizam por uma arquitectura modular complexa: muitas discrepâncias resultam do número de módulos postulados e do tipo de ligação entre os mesmos.

Questões fulcrais da investigação actual sobre literacia, e para a quais existe uma quantidade apreciável de respostas (ou seja, modelos), são, por exemplo, como se faz o acesso lexical? Há um *interface* fonológico ou não há qualquer processamento fonológico de uma forma escrita por parte de um leitor fluente? Como se caracteriza exactamente a consciência fonológica de crianças pré-alfabetizadas e de adultos analfabetos? Como evolui essa consciência com a aprendizagem da escrita? Que unidades linguísticas são relevantes para caracterizar a consciência fonológica dos aprendentes? Segmentos, sílabas, rimas silábicas, terminações lexicais, agrupamentos fonotácticos? Etc.

O que é hoje consensual é que — dando como adquirido que a escrita alfabética assenta numa relação entre unidades de escrita (grafemas, que não é o mesmo que ‘letras’) e unidades fonológicas (fonemas, que não é o mesmo que ‘sons’), as quais se manifestam através dos segmentos fonéticos (fones) — nenhuma ortografia de base alfabética foi, é ou será um sistema de transcrição fonética. Por exemplo, a ortografia portuguesa, mesmo depois da grande reforma ortográfica de 1911 e dos ajustamentos de 1945/1973, não reflecte explicitamente (com duas notáveis excepções⁹), nenhuma das grandes mudanças fonético-fonológicas dos últimos doze ou treze séculos de história da língua (!).

A ortografia é, portanto, algo de radicalmente distinto da transcrição fonética: não representa, nem pode representar, o nível da prola-

⁹ 1) desaparecimento, provavelmente no século XV, de hiatos vocálicos (em formas como ‘veer’, ‘teer’, ‘caaveira’, ‘coor’); 2) surgimento de ditongos nasais, provavelmente também no século XV, aos quais correspondem (na grafia de algumas palavras) os dígrafos ãE, ãO e ÕE.

ção dos enunciados, pois não assenta numa relação biunívoca entre unidades da escrita e unidades fonéticas, ao contrário do que sucede num protocolo de transcrição fonética. A relação entre unidades de escrita e unidades fónicas no âmbito de um sistema de escrita alfabético estabelece-se num nível mais abstracto que o fonético, que é o nível fonémico ou fonológico. A relação entre unidades gráficas e unidades do plano expressivo é, numa escrita alfabética, uma relação entre grafemas e fonemas, e até, nalguns casos, morfemas (unidades mínimas de estrutura interna das palavras). Essa relação permite economia nas unidades mínimas de um sistema ortográfico, pois a correspondência entre grafemas e fones pode ser deduzida do contexto grafémico.

Mas é necessário considerar também a natureza conservadora da ortografia, através da qual se mantém intacta, em alguns casos, a representação grafémica de distinções fonémicas que já desapareceram da língua oral. As ortografias têm história — mudam, tal como os sistemas fonológicos — mas a forma como mudam e o ritmo em que mudam não acompanha necessariamente a “evolução” do sistema fonológico: o desaparecimento de determinado fonema ou de determinada oposição fonémica não acarreta simplificação concomitante e imediata da ortografia, por razões fáceis de entender.

A língua escrita é algo de externo aos falantes, pressupõe o uso de uma tecnologia — a escrita — e não apenas de capacidades neurofisiológicas como a produção da fala. As representações grafémicas e as regras de correspondência grafofonémica são, também, categorias mentais, ou melhor, representações mentais (como as representações fonémicas e fonéticas da língua). Manifestam-se, no entanto, num suporte físico exterior ao falante; esse suporte, ao contrário do que sucede na língua oral, assegura a permanência dos actos de língua escrita para lá do momento em que são produzidos. A escrita fixa intemporal e externamente a comunicação linguística, permitindo o registo e transmissão de actos discursivos e de enunciados através do *continuum* intergeracional. A preservação estável do acervo textual da comunidade ao longo do tempo requer, portanto, estabilidade ortográfica, logo, conservadorismo grafémico.

6 – Conclusão

O primado da escrita é condição *sine qua non* para se entender o papel e o estatuto da língua escrita em sociedades avançadas como a nossa. A desvalorização de mudanças na língua escrita através, por exemplo, de intervenção de Estado na ortografia, só pode resultar na desvalorização de um pilar da nossa sociedade. Quando se usa expressões como ‘a língua de Camões’, ‘a nossa magna lingua portuguesa’ (Fernando Pessoa), ‘história da língua portuguesa’, ‘ensino do português língua materna’, etc., está-se a apontar o grafolecto português europeu. Ninguém vai para escola para aprender a falar (seja dialecto regional ou língua nacional supra-regional) mas para aprender a ler, escrever e a discorrer metalinguisticamente (através do ensino da Gramática). Sem ortografia estável não há grafolecto estável. Sem grafolecto não há sociedade portuguesa viável. O mundo, tal como o conhecemos hoje, não é possível sem escrita. Reconheça-se o seu primado e celebre-se a maior e mais portentosa criação do génio humano.

Este artigo é respeitosamente dedicado a Maria da Graça Castro Pinto, vulto da Linguística Portuguesa e da Psicolinguística, como penhor de profunda admiração pessoal e intelectual. Que a sua leitura lhe seja leve.

REFERÊNCIAS

- Donald, M. 1993. *Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition*. Cambridge MA: Harvard University Press.
- Emiliano, A. 2009. *Fonética do Português Europeu: Descrição e Transcrição*. Lisboa: Guimarães Editores (Guimarães Universitária, 1).
- Goswami, U; Bryant, P. 1990. *Phonological Skills and learning to Read*. Hove/New York: Psychology Press.
- Laver, J. 1994. *Principles of Phonetics*. Cambridge: Cambridge University Press.
- Olson, D. 1994. *The World on Paper*. Cambridge: Cambridge University Press.
- Ong, W. 1982. *Orality and Literacy: The Technologizing of the Word*. London/New York: Routledge.

Nomes em *-dor*, *-douro*, *-deiro*, *-dora*, *-doura* e *-deira*: uma abordagem de acordo com o modelo de RFPs em interfaces

Alexandra Soares Rodrigues
xanafs@mail.telepac.pt
*Escola Superior de Educação –
Instituto Politécnico de Bragança (Portugal)*

RESUMO. Este trabalho foca o comportamento dos sufixos *-dor*, *-dora*, *-deiro*, *-deira*, *-douro* e *-doura* na formação de nomes deverbais do Português. Partindo do modelo de “RFPs em interfaces”, baseado na coindexação dos traços semânticos provenientes dos afixos e das estruturas eventiva e léxico-conceptual das bases verbais, pretende-se evidenciar o contributo do afixo para o semantismo do produto lexical. Apesar de se tratar de afixos operadores na construção de nomes de ‘causa’, ‘locativos’ e ‘instrumentos’, verifica-se que o produto de cada afixo tende para determinados semantismos. Essas tendências são explicáveis pela existência de traços semânticos no afixo que, em coindexação com os traços das bases, produz matizes semânticos nos produtos.

PALAVRAS-CHAVE. Regras de formação de palavras; nomes deverbais; coindexação semântica; afixação.

ABSTRACT. This work focus on the behaviour of the suffixes *-dor*, *-dora*, *-deiro*, *-deira*, *-douro* and *-doura* on the formation of Portuguese deverbal nouns. We follow the “Word formation Rules in interfaces” model, which is based on the coindexation of semantic features from the affixes and from eventive and lexical-conceptual structures of the verbal bases. We intend to give evidence of the contribution of each affix to the final meaning of the product. Although all these affixes operate in the construction of nouns of ‘cause’, ‘place’ and ‘instrument’, we observe that the products of each affix have a tendency for certain meanings. Those tendencies are explainable by the existence of semantic features in the affix that, in coindexation with the features of the base, produce semantic nuances in the products.

KEY-WORDS. Word formation rules; deverbal nouns; semantic coindexation; affixation.

1 – Introdução

O presente trabalho tem como principal objectivo mostrar alguns aspectos, nomeadamente semânticos, da formação de nomes deverbais em *-dor*, *-dora*, *-deiro*, *-deira*, *-douro* e *-doura*, de acordo com o modelo genolexical de “RFPs em interfaces”. Começaremos por fazer uma breve apresentação deste modelo, comparativamente com modelos anteriores (§ 2). No § 3, apresentaremos as significações principais que ocorrem nestes produtos. No § 4, observaremos a distribuição numérica dessas significações pelos produtos de cada afixo. No § 5, procederemos a uma explicação dessas significações, com base no modelo teórico adoptado.

2 – O modelo de formação de palavras “RFPs em interfaces”

O modelo de formação de palavras que utilizamos para descrever a construção dos nomes deverbais em *-dor*, *-dora*, *-deiro*, *-deira*, *-douro* e *-doura*, designado por “modelo de RFPs em interfaces”, foi desenvolvido em Rodrigues (2008). O modelo parte da conjugação dos modelos orientados para o *input* (e.g. Beard 1995; Corbin 1987; 1991; Rio-Torto 1993; 1998) com os modelos orientados para o *output* (Plag 1999; 2004).

Nos modelos orientados para o *input*, concebe-se que cada Regra de Formação de Palavras é caracterizada por uma relação entre um tipo categorial e semântico de bases e um tipo categorial e semântico de produtos. Os operadores afixais são actuantes numa RFP de acordo com o tipo de base a que se juntam e o tipo de produto que constroem. Assim, nomes como *abridor*, *esborralhadouro*, *assadeiro* são considerados como produtos da mesma RFP, na medida em que têm por base verbos (*abrir*, *respirar*, *assar*) e são nomes que designam ‘instrumentos’.

Um modelo orientado para o *input* apresenta a vantagem de evidenciar as semelhanças existentes entre produtos de operadores afixais distintos, mas que partilham o mesmo tipo de base, a mesma categoria lexical e o mesmo semantismo (e. g. $V \rightarrow N$ ‘acção de V’: *vingança*, *enlouquecimento*, *moagem*, *avaliação*). No entanto, o mesmo modelo apresenta a desvantagem de não explicar a ocorrência de produtos com o mesmo operador afixal, mas que partem de bases

categorialmente distintas, e que apresentam semantismos divergentes. É este o caso de produtos como *carrão*, *empurrão* e *chorão*. *Carrão* tem por base o nome *carro*; *empurrão* e *chorão* têm por bases verbos (*empurrar* e *chorar*, respectivamente). Semanticamente, *carrão* designa ‘carro grande, caro’. Dado que as bases de *empurrão* e de *chorão* não são nomes, não é aplicável o mesmo tipo de semantismo a estes produtos. *Empurrão* designa o ‘evento súbito e impulsivo de V’; *chorão* designa ‘aquele que V muito’. Não obstante estas diferenças semânticas, que são justificadas pelas diferenças entre as bases *carro*, *empurrar* e *chorar* (cf. Rodrigues 2008: 53-60), é inegável a coincidência de um semantismo de ‘intensidade’ nos três produtos.

O problema dos modelos orientados para o *input* é que não explicam as semelhanças semânticas entre produtos que ostentam o mesmo afixo. Isto deve-se ao facto de, nas versões mais ortodoxas destes modelos (Beard 1995; Corbin 1987), se conceber que um afixo pode apenas operar numa única RFP.¹ Consequentemente, estes modelos recorrem à multiplicação em homónimos de afixos como *-ão*, sempre que se observa que o afixo opera em RFPs distintas. O afixo é descurado no seu contributo semântico para a construção do produto e, mais, na sua capacidade de construção de relações paradigmáticas entre lexemas.

Visão oposta é oferecida pelos modelos de formação de palavras orientados para o *output*, como é aquele proposto por Plag (1999; 2004). Nestes modelos, o eixo central dos paradigmas genolexicais reside no operador afixal e não na relação semântico-categorial entre as bases e os produtos. Como tal, perante produtos como *carrão*, *empurrão* e *chorão*, dá-se importância à unidade semântica acarretada pelo afixo *-ão* e descuram-se as diferenças semânticas decorrentes das diferenças semântico-categoriais das bases. Para além disso,

¹ Corbin apresenta sucessivas alterações ao modelo original de 1987, por forma a resolver estes problemas que decorrem de uma visão centrada na relação entre a base e o produto na definição de cada RFP. Vejam-se, por exemplo, Corbin (1991; 2004) e Corbin & Corbin (1991). Em nenhuma destas reformulações se abandona, no entanto, o pendor de orientação para o *input*. Também Rio-Torto se destaca da visão mais ortodoxa do modelo orientado para o *input*. Em Rio-Torto (1993; 1998), a autora modaliza o modelo de modo a explicar produtos que se assemelham semanticamente devido à partilha do mesmo afixo. Neste trabalho não procederemos a uma análise detalhada dessas propostas. Veja-se Rodrigues (2008: 25-60) para esse propósito.

esta perspectiva descara ainda as semelhanças que se encontram entre produtos que, não partilhando o mesmo afixo, se enquadram no mesmo paradigma semântico-categorial, como *carrão*, *bocarra*, *mulheraça*, por partilharem a mesma relação semântico-categorial entre base e produto.

Observemos produtos como *livraria*, *berraria*, *fanfarronaria*, *moagem*, *vassalagem*, *folhagem*.

Um modelo orientado para o *input* procederá aos seguintes conjuntos:

Conjunto A	Conjunto B	Conjunto C
<i>fanfarronaria</i> <i>vassalagem</i>	<i>berraria</i> <i>moagem</i>	<i>livraria</i> <i>folhagem</i>

O conjunto A é definido pela partilha da base, adjetivo e/ou nome que funcione como predicativo, bem como da saída, que é um nome de qualidade, ou essivo, (Rio-Torto 1998: 122-123). O conjunto B é caracterizado por base verbal e saída nome de evento. O conjunto C é constituído por lexemas nominais que designam ‘quantidade’ que têm por base um nome concreto.

Um modelo orientado para o *output* construirá os seguintes conjuntos:

Conjunto 1	Conjunto 2
<i>fanfarronaria</i> <i>berraria</i> <i>livraria</i>	<i>vassalagem</i> <i>moagem</i> <i>folhagem</i>

O conjunto 1 engloba os produtos do sufixo *-aria*. O conjunto 2 engloba os produtos do sufixo *-agem*. Observe-se que a identificação de cada afixo não se baseia apenas na sua estrutura fonológica, mas no complexo fonológico-semântico. É a apresentação de uma constância semântica entre os produtos que permite a identificação do afixo como único.

Observadas de modo breve as vantagens e as desvantagens dos modelos orientados para o *input* e para o *output*, passamos a apresentar também em traços breves o modelo que desenvolvemos e que designámos por “modelo de RFPs em interfaces”.

Tal como é observável pelos exemplos acima demonstrados, não é viável uma descrição satisfatória do domínio de formação de palavras se nos ativermos univocamente a uma visão direccionada apenas para o *input*, ou a uma visão direccionada apenas para o *output*. Ambas as visões concebem de modo excessivamente estanque o modo de operar quer dos afixos, quer das relações entre bases e produtos.

A análise de um *corpus* de 8414 nomes deverbais em Rodrigues (2008) providenciou-nos exemplos suficientes para questionarmos a validade dos dois modelos acima referidos e para construirmos um modelo alternativo que nos parece estar mais de acordo com os dados empíricos.

O modelo de RFPs em interfaces parte duma visão mais dinâmica do próprio léxico e prevê que o mesmo afixo possa operar em RFPs distintas. Deste modo, não se descarta a constância semântico-categorial observável entre bases e produtos duma determinada RFP, mas põe-se em simultâneo em evidência o contributo semântico do operador afixal na construção do produto, bem como a sua capacidade de se anexar a diferentes bases definidas semântico-categorialmente.

O nosso modelo construirá 5 conjuntos simultâneos a partir dos lexemas *livraria, berraria, fanfarronaria, moagem, vassalagem e folhagem*.

	Conjunto 1		Conjunto 2
Conjunto A	<i>fanfarronaria</i>		<i>vassalagem</i>
Conjunto B	<i>berraria</i>		<i>moagem</i>
Conjunto C	<i>livraria</i>		<i>folhagem</i>

As RFPs são concebidas, no modelo de RFPs em interfaces, como relações dinâmicas entre tipos de bases e tipos de produtos relacionáveis semântico-categorialmente através de operadores afixais, ou de outros mecanismos.² As RFPs mantêm relações também

² Referimo-nos, por exemplo, aos mecanismos de conversão responsáveis pela produção de nomes deverbais como *abraço* de *abraçar* e de verbos denominais, como *ancorar* de *âncora*. Acerca da direccionalidade derivacional destes processos, vejam-se Rodrigues (2001; 2002).

dinâmicas entre si. Neste modelo, os operadores afixais e/ou os mecanismos responsáveis pelas formações não se atêm à laboração em apenas uma RFP. A capacidade de operarem em várias RFPs em simultâneo resulta da conjugabilidade entre os caracteres semânticos do afixo e aqueles que relacionam a base e o produto de cada RFP. Isto conduz a uma permeabilidade entre as RFPs, visível através dos contactos semânticos entre produtos de RFPs distintas, mas que partilham o mesmo afixo.

O modelo de RFPs em interface considera a dimensão semântica como decisiva na actuação genolexical. O modelo prevê um mecanismo de ‘coindexação’ entre traços semânticos do afixo e traços semânticos da base que se revela responsável pela obtenção do semantismo final do produto. O conceito de ‘coindexação’, apresentado em Lieber (2004), sofre no nosso modelo modificações devido à necessidade de adequarmos o modelo teórico aos dados lexicais observados.³ O que o *corpus* analisado em Rodrigues (2008) revela é que num primeiro nível do domínio da formação de palavras, concebido como constituído por fiadas de diversas estruturas, o nível semântico é aquele que actua na selecção entre operador afixal e base.

O referido *corpus* é constituído por 8414 nomes deverbais com os afixos *-agem, -aria, -ão, -ção, -mento, -dura, -ncia, -nça, -nço, no caso dos nomes de ‘evento’, e -ão, -dor, -dora, -douro, -doura, -deiro, -deira, -nte, -vel, -tório, -ório, -ória e -tória, no caso dos nomes de ‘indivíduo’*⁴. O *corpus* foi construído a partir do *Dicionário da Língua Portuguesa*, da Porto Editora, Bluteau (1712-1728) e Domingos Vieira (1871-1874). Recorreu-se ainda ao *Dicionário da Língua Portuguesa Contemporânea* (2001) da Academia das Ciências de Lisboa e ao *Dicionário Houaiss da Língua Portuguesa* (2001-2002) para completar informação semântica. Da análise desse *corpus* foi possível formular traços semânticos centrais para cada um dos afixos que, em coindexação com determinados traços semânticos das bases a que se

³ Não entraremos neste trabalho em detalhes sobre as diferenças entre o conceito de *coindexação* apresentado por Lieber (2004) e aquele que está desenvolvido no nosso modelo. Veja-se Rodrigues (2008: 60-71) para esse efeito.

⁴ ‘Indivíduo’ corresponde a uma entidade delimitada espacialmente e opõe-se a ‘evento’. ‘Indivíduo’ não especifica um ser animado. Pode tratar-se de um ‘locativo’, de um ‘instrumento’, etc.

Significações não assinaladas por representarem variante feminina dos lexemas em *-dor*.

agregam, geram os semantismos dos produtos respectivos (Rodrigues 2008: 227-274).

Os traços semânticos dos afixos permitem compreender por que motivo produtos duma mesma RFP, que, como tal, partilham a mesma significação genérica, apresentam entre si diferentes matizes semânticos. No caso dos nomes de evento, esses matizes vão desde os semantismos de ‘acção súbita’, ‘processo’, ‘estado’, ‘acção reiterada’, entre outros. No caso dos nomes de indivíduo, os matizes são exemplificados por semantismos de ‘causa humana’, ‘causa animal’, ‘instrumento’, ‘locativo’, ‘recipiente’, etc. (Rodrigues 2008: 248-255).

Neste trabalho, centrar-nos-emos nos produtos dos sufixos *-dor, -dora, -deiro, -deira, -douro e -doura*. Apesar de se tratar de produtos de indivíduo que apresentam significações de ‘causa humana’, ‘causa animal’, ‘causa’, ‘locativo’, ‘instrumento’, ‘recipiente’, ‘parte de corpo’, verificam-se preferências de cada um dos afixos em relação a cada uma destas significações. Tentaremos explicitar neste trabalho o modo de operação semântica de cada um dos afixos para a obtenção dos semantismos dos produtos em análise.

As significações apontadas para os lexemas são, salvo indicação contrária, retiradas do *Dicionário da Língua Portuguesa* da Porto Editora, versão *online*.

3 – As significações

As principais significações encontradas nos nomes deverbais construídos através dos sufixos em análise foram as seguintes:

-‘causa’: instigador ou actuante directo não-animado de um evento, parafraseável por ‘que efectua V’ (e. g. *flutuator* ‘o que flutua; o que se conserva à superfície de um líquido’, *ecoador* ‘aquilo que ecoa’);

- ‘causa animal’: instigador ou actuante directo de um evento, que é um animal (e. g. *trepadeira* ‘pássaro sedentário e comum em Portugal, que tem por hábito subir ao longo do tronco das árvores’, *lavadeira* ‘alvéola’);

-‘causa humana’: instigador ou actuante directo humano de um evento (e.g. *canalizador, bordadeira*);

- 'causa substância': instigador ou actuante directo de um evento, que é uma substância (e. g. *catalisador* 'substância que modifica a velocidade das reacções químicas, apresentando-se inalterada no fim da reacção', *despolarizador* 'substância que se coloca numa pilha eléctrica para eliminar ou diminuir a polarização');

- 'causa v': instigador ou actuante directo de um evento, que é um vegetal (e. g. *dormideira* 'planta herbácea, lactescente, da família das Papaveráceas, espontânea em Portugal, que tem propriedades sedativas e narcóticas, e da qual se extrai ópio', *estancadeira* 'planta herbácea, da fam. das Plumbagináceas, aplicada para sustar hemorragias');

- 'instrumento mecanizado': artefacto mecanizado utilizado por causa animada para a efectuação de um evento (e. g. *escavadora*, *aquecedor*);

- 'instrumento não-mecanizado': artefacto não-mecanizado utilizado por causa animada para a efectuação de um evento (e. g. *tapadoura* 'tampa', *abridor*);

- 'locativo': espaço fixo com 3 dimensões (e. g. *pousadeiro* 'lugar onde a ave pousa', *desovadouro* 'lugar onde um animal desova');

- 'locativo causa': espaço fixo com 3 dimensões que instiga ou actua (n)um evento (e. g. *germinadouro* 'lugar onde se faz germinar a cevada para o fabrico da cerveja', *resvaladeiro* 'sítio ou terreno por onde se resvala facilmente');

- 'locativo c': espaço com 3 dimensões coincidente com um segmento ou parte de um corpo animado ou não-animado (e. g. *encabadouro* 'abertura onde entra o cabo de qualquer utensílio', *nascedouro* 'orifício do útero');

- 'locativo c causa': espaço com 3 dimensões coincidente com um segmento ou parte de um corpo animado ou não-animado que instiga ou actua (n)o evento (e. g. *respiradouro* 'orifício destinado a deixar entrar e sair o ar', *coagulador* 'última e principal cavidade do estômago dos ruminantes, também denominada abomaso e coalheira');

- 'locativo p': espaço com 3 dimensões, móvel, que serve de continente/recipiente (e. g. *cozedeira* 'peça de barro em que se leva

comida ao fogo’, *pingadeira* ‘recipiente para aparar os pingos da carne que se assa no forno ou no grelhador’);

-‘locativo p causa’: espaço com 3 dimensões, móvel, que serve de continente/recipiente, que instiga ou actua (n)o evento (*congelador*, *chocadeira* ‘incubadora’);

-evento: acontecimento na dimensão temporal (e. g. *chiadouro* ‘acção frequente de chiar’, *fervedouro* ‘movimento como o de um líquido a ferver’).

4 – Distribuição das significações por produtos

Os produtos em *-dora*, *-deira* e *-doura* elencados no *corpus* já mencionado são aqueles constantes em produtos lexicográficos, isto é, que manifestam significações lexicalizadas, semanticamente distintas daquelas que surgem nos actuais ou hipotéticos lexemas masculinos. Por esse motivo, a inexistência de semantismos de ‘causa’, de ‘causa animal’, ou o escasso número de ‘causa humana’ nos produtos em *-dora* não revelam que estas significações sejam inexistentes em lexemas com este operador afixal. Demonstram apenas que se trata de significações não lexicalizadas. Por exemplo, *avaliadora* representa a forma feminina de *avaliador*, enquanto *escavadora* ‘máquina’ não representa a forma feminina de *escavador*. É apenas contabilizada, para os produtos em *-dora*, *escavadora* e não *avaliadora*. Observe-se, no entanto, que as significações de ‘causa humana’ e de ‘causa animal’ estão ausentes dos produtos em *-doura* na mesma medida em que estão ausentes dos produtos em *-douro*. Não se trata, pois, da mesma situação observável para *-dora*.

Os números percentuais apresentados dizem respeito aos valores absolutos das significações em análise neste trabalho e não daquelas que foram analisadas em trabalhos anteriores (cf. Rodrigues 2008), onde eram comparados mais afixos para além dos que estão em análise neste trabalho. A tabela 1 mostra os valores absolutos das significações em análise para os afixos *-dor*, *-dora*, *-deiro*, *-deira*, *-douro* e *-doura*.

Dos dados explicitados nas tabelas 1 e 2, destacam-se os seguintes:

- os produtos em *-dor* são aqueles que apresentam maior número de significações de ‘causa humana’ (67,05%). Seguem-se os produtos em *-deira* (28,57%), *-deiro* (24,59%), *-dora* (17,50%);

- os produtos em *-douro* e *-doura* não apresentam o semantismo de ‘causa humana’, ‘causa animal’, ‘causa s’ ou ‘causa v’; a única ‘causa’ é a genérica;

- nos produtos em *-dora*, o semantismo mais significativo é o de ‘instrumento mecanizado’, que ocupa 75,00% das significações disponibilizadas pelos lexemas lexicalizados com este sufixo. Nos produtos em *-dor*, a significação de ‘instrumento mecanizado’ apresenta-se com menor percentagem (6,09%) do que a significação de ‘instrumento não-mecanizado’ (8,76%), que é baixa nos produtos em *-dora* (5,00%);

- nos produtos em *-deiro*, encontra-se ausente a significação de ‘instrumento mecanizado’, mas esta está presente nos produtos em *-deira* com 14,29%;

- não ocorrem semantismos de ‘instrumento mecanizado’ nos produtos em *-douro* e em *-doura*;

- a significação de ‘instrumento não-mecanizado’ tem mais representatividade nos produtos em *-doura* (77,27%), seguindo-se os produtos em *-deira* (27,70%) e em *-douro* (16,34%);

- a significação com maior percentagem nos produtos em *-doura* é a de ‘instrumento não-mecanizado’ (77,27%);

- a significação de ‘locativo’ tem maior representação nos produtos em *-douro* (37,62%) e *-deiro* (14,75%);

- a significação de ‘locativo causa’ tem maior representação nos produtos em *-douro* (19,67%) e *-deiro* (12,87%);

- a significação de ‘locativo c’ tem representação apenas nos produtos em *-douro* (4,46%);

- a significação de ‘locativo p’ tem maior representação nos produtos em *-douro* (9,09%), *-doura* (7,92%) e *-deiro* (4,92%);

- os semantismos de ‘evento’ só ocorrem nos produtos em *-douro* (12,38%), *-deira* (7,29%), e *doura* (4,55%).

Pode concluir-se que:

- os produtos em *-douro* e *-doura* não são representativos de ‘causa’ especificada como ‘humana’ ou ‘animal’;
- os produtos em *-douro* e *-doura* são representativos de ‘instrumentos não-mecanizados’;
- os ‘instrumentos mecanizados’ estão a cargo de *-deira* e *-dora*;
- as significações de ‘causa humana’ e de ‘causa animal’ são mais representadas por *-dor*, *-deiro* e *-deira*;
- o semantismo de ‘locativo p’ é mais representado por *-deiro*, *-deira*, *-douro* e *-doura*; mas a representatividade pelos sufixos *-douro* e *-doura* decresce se a significação estiver aliada a ‘causa’ (‘locativo p causa’);
- o semantismo de ‘locativo c’ é representado por *-douro*, mas essa representatividade decresce se esta significação estiver aliada a ‘causa’ (‘locativo c causa’);
- os locativos estão a cargo de *-douro* e *-deiro*; se se tratar de ‘locativo causa’ o sufixo *-douro* perde representatividade;
- ocorrem semantismos de ‘evento’ em *-douro*, *-deiro*, *-doura*.

5 – Interpretação dos resultados

O modelo de formação de palavras que desenvolvemos apresenta a vantagem de, para além de descrever os semantismos de cada produto, explicar esses mesmos semantismos. A partir dos dados acima descritos levantam-se as seguintes questões:

a) por que razão os produtos em *-douro*, *-doura* são alheios aos semantismos de ‘causa humana’ e de ‘causa animal’? Repare-se que, nos domínios semânticos em que estes afixos têm grande peso (‘locativos c’, ‘locativos p’ e ‘locativos’), estes operadores perdem visibilidade se estes semantismos estiverem aliados ao semantismo de ‘causa’.

b) por que razão os sufixos *-dor* e *-dora* apresentam escassos locativos, mostrando preferência por causa humana (*-dor*) e por instrumento mecanizado (*-dora*)?

c) por que razão os sufixos *-deiro* e *-deira* parecem situar-se numa zona intermédia entre *-dor/ -dora* e *-douro/ -doura*, mostrando significações de ‘causa’, ‘instrumento’ e ‘locativo’?

Na introdução que fizemos ao modelo de RFPs em interfaces focámos que o modelo prevê um mecanismo de coindexação (Rodrigues 2008: 60-71). Esse mecanismo é o responsável por estas características semânticas assinaladas nas questões anteriores. O modelo desenvolvido prevê uma composicionalidade semântica dos lexemas em diversas fiadas, muitas vezes não acessíveis directamente. Não se trata, pois, de componentes semânticos que se possam encontrar descritos num dicionário, ou a que o falante possa aceder explicitamente. Dos componentes que intervêm na formação destes produtos temos em consideração o contributo da base e o contributo do afixo. Da base foram delimitados componentes semânticos da estrutura eventiva e da estrutura léxico-conceptual. A estrutura eventiva diz respeito a subcomponentes eventivos *i. e.*, situados no eixo do tempo, com carácter lexical e não sintáctico. Esses subcomponentes são, por exemplo, [ponto de chegada], [ponto de partida], [pontual], [durativo], etc. (Rodrigues 2008: 202-204). A estrutura léxico-conceptual lida com participantes no evento léxico-conceptualmente inscritos, como [causa], [objecto], [locativo], etc. (Rodrigues 2008: 242-248). A estrutura léxico-conceptual não coincide com a estrutura argumental e, portanto, não deve ser tomada por esta.⁵

O sufixo contribui semanticamente para o produto final com um traço semântico (Rodrigues 2008: 235-242). Esse traço semântico foi delimitado para cada sufixo através da análise em contraste de todos os produtos com aquele operador afixal (*e. g. avaliador, canalizador, corredor, abridor*) e dos produtos com a mesma base e com afixo diferente (*bebedeira, bebedouro, bebedor*). O traço do sufixo não é visível directamente em cada produto, mas sim através de semantismos particulares, por vezes secundários, dos produtos do mesmo sufixo, selecção de bases, etc. (Rodrigues 2008: 236, nota 197).⁶

⁵ Neste momento não entraremos em detalhe sobre a descrição dos componentes destas estruturas, nem quanto ao modo de coindexação entre si. Veja-se Rodrigues (2008:94-113) para uma descrição aturada destes mecanismos e componentes.

⁶ Este dado deve ser tido em conta na leitura dos traços que aqui apresentamos de forma breve. Remetemos para Rodrigues (2008: caps. VI e VII) para a verificação dos motivos de estabelecimento de cada traço particular de cada afixo.

Temos assim os possíveis traços a coindexar entre si: um traço do sufixo em acção; um ou mais traços em simultâneo da estrutura léxico-conceptual da base; um ou mais traços em simultâneo da estrutura eventiva da base. A coindexação entre os traços faz-se de acordo com conjugabilidades semânticas entre os traços activados (Rodrigues 2008: 255-267).

Para os sufixos *-dor* e *-dora* foi delimitado o traço [que tem a função de]. Para os sufixos *-deiro* e *-deira* o traço [que tem a funcionalidade de]. Para *-douro* e *-doura* o traço é [propício a/próprio para]. Descreveremos brevemente estes traços para depois tecermos considerações acerca do seu papel na obtenção dos semantismos dos produtos. Tenha-se em atenção que para uma compreensão cabal destes traços é necessário considerar a totalidade dos afixos operadores de 'indivíduo', na medida em que a identificação de cada traço se faz por contraste. Por motivos de espaço, remetemos para Rodrigues (2008: 240-242) para a motivação de cada traço apontado.

O traço [que tem a função de], caracterizador dos sufixos *-dor* e *-dora*, indicia um indivíduo que exerce a função instanciada pela base verbal, sem que se preveja se existe ou não capacidade para o exercício da mesma função (Rodrigues 2008: 346, nota 243).

O traço [que tem a funcionalidade de], presente em *-deira* e *-deiro*, mostra um indivíduo que possui a capacidade de exercer uma função.

O traço [propício a/próprio para], assinalado para os sufixos *-douro/-doura*, mostra um indivíduo adequado à ocorrência do evento significado pela base verbal, sem que esse indivíduo seja actuante ou instigador do evento, ao contrário do que ocorre com os traços [que tem a função de] e [que tem a funcionalidade de]. Este traço é revelador da versatilidade semântica que *-douro* e *-doura* disponibilizam.

5.1 – Produtos em *-dor*

Dos produtos em *-dor* destacam-se aqueles que significam 'causa', especialmente 'causa humana'. Apesar de não apresentar número relevante de semantismos locativos, é de destacar que estes surgem ainda em menor número quando não possuem o componente de 'causa', à excepção do 'locativo', para o qual se encontraram apenas

dois exemplos: *toucador* (Domingos Vieira: «A casa onde alguém touca a cabeça.») e *obrador*⁷ (Domingos Vieira: «Dá-se este nome, nas terras de lanificio da Serra da Estrella, á casa ordinariamente terrea, ou loja, onde cardam a lã, e mesmo onde teem algum tear.»)). Em todo o caso, a soma dos ‘locativos’ aliados a ‘causa’ é muito superior à soma dos ‘locativos’ sem ‘causa’, como se verifica recorrendo às tabelas 1 e 2.

Esta tendência explica-se pela actuação do traço [que tem a função de]. Este traço, coindexado com o traço [causa] da base verbal, e com um traço eventivo dependente de cada base verbal, restringe o ‘indivíduo’ como aquele que é instigador/actuante do/no evento. Atente-se: o traço [que tem a função de] coindexa com um traço eventivo da base verbal, o que provê o produto de um semantismo de ‘que tem a função de Evento’. Ora, este semantismo adequa-se ao instigador/actuante do/no evento. Daí que a coindexação com um traço da estrutura léxico-ceneptual da base verbal se faça tendencialmente com o traço [causa]. Os semantismos finais obtidos são, deste modo, o de ‘causa’, especialmente com a especificação de ‘causa humana’ (e. g. *controlador, desenhador, atirador*), e o de ‘instrumento’ (e. g. *desentupidor, radiador, espanador*).

Isto explica igualmente que os poucos locativos nos produtos em *-dor* estejam relacionados com ‘causa’. Assim, os locativos são sobretudo ‘locativos p causa’, ou seja, recipientes, que exercem eles próprios a ‘função de Evento’. É este o caso de *congelador, refrigerador, assador*, etc. O facto de existirem produtos em *-dor* sem a componente ‘causa’ prova que este carácter semântico não é inerente ao sufixo, antes decorre da base verbal. Um *apontador* ‘livro para apontamentos’, ou um *copiador* ‘livro onde se copia a correspondência expedida’ são ‘locativos p’, que ‘têm a função de Evento’, mas não são ‘causa’, na medida em que não instigam ou efectuam o evento.

⁷ *Toucador* é, segundo Domingos Vieira, «A casa onde alguém touca a cabeça.». Em relação a *Obrador*, Domingos Vieira fornece a seguinte indicação: «Dá-se este nome, nas terras de lanificio da Serra da Estrella, á casa ordinariamente terrea, ou loja, onde cardam a lã, e mesmo onde teem algum tear.».

5.2 – Os produtos em *-dora*

Recordamos que nas formas em *-dora* se mostram apenas aquelas que se encontram lexicalizadas e não as que correspondem à variante feminina do genérico masculino. Esta variação é própria de semantismos de ‘causa humana’ e ‘causa animal’. As significações de ‘instrumentos’ e as de ‘locativos’ não apresentam variação em género; estão antes lexicalizadas ou na forma feminina ou na masculina.

Quanto ao facto de *-dora*, em relação a *-dor*, mostrar maior número de ‘instrumentos mecanizados’, dever-se-á dever à lexicalização maior a que se prestam as formas marcadas morfológicamente, isto é, as formas do feminino. Assim, parece haver uma correlação entre as formas morfológicamente marcadas (femininas) e os semantismos que indiciam maior grau de lexicalização do traço [que tem a função de] - os semantismos técnico-profissionais, visíveis em designações de profissões e em ‘instrumentos mecanizados’. As designações de profissões estão, no entanto, atreitas a variação de género, o que não acontece com os ‘instrumentos mecanizados’.

Recusamos a hipótese de a elisão do nome *máquina* estar na origem dos nomes de ‘instrumentos mecanizados’ como formas tendencialmente femininas, como defendido por Grossmann (1998: 387). Essa elisão poderia ser relativamente a *mecanismo*, *engenho*, *aparelho*, o que providenciaria género masculino aos produtos.

Não existe correlação entre feminino e ‘máquina’, mas antes entre forma marcada que se presta a maior lexicalização e semantismo mais lexicalizado.

5.3 – Produtos em *-deiro*

Nos produtos em *-deiro* destacam-se numericamente os semantismos de ‘causa’, ‘causa humana’, ‘locativo causa’ e ‘locativo’. Para esta tendência concorre o traço do sufixo, que especificámos como [que tem a funcionalidade de], ou seja, ‘que tem a capacidade para o cumprimento de uma função’. São assim comuns as designações de ‘causas humanas’ que indicam uma profissão (e. g. *fiadeiro*, *cevadeiro*, *vindimadeiro*) ou que implicam um juízo de valor (e. g. *rezadeiro*, *benzedeiro*), partindo ambas do grau de ‘frequência’ com que o instigador/actuante efectua o evento.

As ‘causa animais’ corroboram os subcomponentes de ‘frequente’ e de ‘capacidade’ deste sufixo. Os animais designados por estes produtos, de que são exemplo *piadeiro* ‘ave: peto’, *lavadeiro* ‘ratinho da América’, bem como os congêneres femininos, como *trepadeira* ‘ave’, *subideira* ‘ave’, mostram o carácter de ‘frequente’ e de ‘capacidade’ que explicitámos para o contributo semântico de *-deiro/-deira*.

O subcomponente semântico de ‘capacidade’ de ‘funcionalidade’ aparece de forma mais visível nos produtos que designam ‘locativos’. Exemplos como *deslizadeiro*, *resvaladeiro*, *singradeiro*, *atascadeiro*, *atoladeiro*, *despenhadeiro* mostram lugares ‘que têm a capacidade de instigar Evento’. A versão de ‘funcionalidade’ destaca-se mais em exemplos como *apeadeiro*, *cremadeiro*, *malhadeiro*. Observe-se que a ‘funcionalidade’ é inscrita como capacidade que não requer voluntariedade ou controlo por parte do detentor da mesma. Os ‘locativos’ correspondem a espaços aproveitados pela sua funcionalidade para determinado Evento.

5.4 – Produtos em *-deira*

Nestes produtos, o traço [que tem a funcionalidade de] é particularmente visível em lexemas como *poedeira* e *parideira*. *Parideira* significa ‘fêmea em idade (com capacidade) de parir e que é fértil [traço ‘frequente’]’. *Poedeira* aplica-se a ovíparas segundo os mesmos traços (Rodrigues 2008: 240).

Em paralelo com o que se verificou para os produtos em *-dora*, as formas femininas *-deira* mostram mais semantismos de ‘instrumento’ do que os congêneres masculinos. Estes instrumentos, mesmo quando mecanizados, possuem um carácter tradicional para os parâmetros actuais, como *empacotadeira* ‘máquina agrícola para enfardar palha ou feno’, *enfardadeira* ‘máquina agrícola que serve para enfeixar e comprimir a palha ou o feno’. Exemplos de ‘instrumentos não-mecanizados’ são *bulideira* ‘pá com que se separam os pães no forno’, *lançadeira* ‘peça do tear, em forma de naveta, com um pequeno cilindro ao meio (canela) em que se enleia o fio que os tecelões ou teceadeiras fazem passar pela urdidura’, *arrasadeira* ‘pau cilíndrico que serve para arrasar medidas’, etc. Em relação aos semantismos de ‘instrumento’ dos produtos em *-dora*, os em *-deira* mostram carácter

mais tradicional. Os produtos em *-dora* possuem carácter mais tecnológico.

O mesmo carácter tradicional está presente nas designações de profissões, como *bordadeira*, *engarrafadeira*, *assedadeira*, *brunideira*, *cerzideira*, *dobradeira*, *encartadeira*, *lavadeira*, *pregueadeira*, *fiadeira*, etc.

5.4 – Produtos em *-douro*

Nestes produtos destaca-se a ausência de ‘causa humana’ e de ‘causa animal’, mas verifica-se a presença de ‘causa’. Isto mostra que o matiz semântico em ausência é aquele que se relaciona com um instigador/actuante de Evento que tenha a possibilidade de o controlar ou de o efectuar voluntariamente. Recordamos que o traço explicitado para este sufixo é o de [propício a/próprio para], de que está ausente um subcomponente instigador/actuante de Evento.

Como tal, este sufixo mostra-se apropriado para designar ‘causa’ não voluntária (*defumadouro* ‘substância com que se defuma’, *apertadouro* ‘coisa que aperta’, *suadouro* ‘bebida ou remédio para fazer suar’), instrumentos (*aradouro* ‘arado’, *cerradouro* ‘cordão de abrir e fechar bolsas, sacos, etc.’), ‘locativos c’ (*sugadouro* ‘prolongamento que certas plantas parasitas introduzem nas plantas parasitadas para as sugarem’, *sangradouro* ‘parte do braço onde, através de incisão numa veia, se pratica a sangria’) e ‘locativos’ (*assomadouro*, *surgidouro*, *desaguadouro*).

5.6 – Produtos em *-doura*

O formato *-doura* corrobora aquilo que foi dito em relação às formas femininas: em comparação com os pares masculinos, estas evidenciam preferência pelos semantismos de ‘instrumentos’ (*rapadoura* ‘instrumento próprio para rapar; pequena pá com que se rapa a massa nas masseiras’, *roçadoura* ‘foice’). Também corrobora o que dissemos a propósito de *-douro*: ausência de ‘causa humana’ e de ‘causa animal’, de acordo com a adequação ao traço [propício a/próprio para].

5.7 – Semantismos de ‘evento’

Os semantismos de ‘evento’ em produtos de sufixos que constroem prototipicamente designações de ‘indivíduo’ devem-se a extensões de subcomponentes dos traços semânticos dos sufixos. Essas extensões dizem respeito ao semantismo de ‘frequente’. Os sufixos que disponibilizam o semantismo de ‘evento’ são *-deira* (*chiadeira*, *choradeira*, *fungadeira*, *ganadeira*, *gritadeira*), *-douro* (*tremedouro*, *chiadouro*, *gemedouro*) e *-doura* (*varredoura* ‘grande morticínio ou destruição’).

6 – Conclusões

A distinção entre *-dor*, *-deiro* e *-douro* resulta da diminuição da extensão de ‘causa’ entre os três sufixos. Essa extensão relaciona-se com a adequação que cada traço de cada sufixo mostra relativamente ao componente ‘causa’. Essa diminuição acarreta valores numéricos distintos dos vários tipos de semantismos. *-dor* produz semantismos de indivíduo que cumprem a função de Evento. Como tal, são numericamente salientes os produtos em *-dor* com semantismos de ‘causa’, sejam ‘locativos’ ou ‘causa humana/animal’. *-douro* mostra indivíduos que são propícios ao Evento, mas não especificamente instigadores do mesmo. Assim, a correlação com ‘causa’ é ténue, o que explica a ausência de ‘causa humana’ e a presença de ‘locativos’ sem ‘causa’ e de ‘instrumentos’. *-deiro* situa-se numa zona intermédia entre *-dor* e *-douro*. O traço [que tem a funcionalidade de] possibilita semantismos situados nos vários domínios semânticos, ainda que com tendência para a extensão de ‘causa’.

O modelo de RFPs em interfaces, que prevê uma organização em fiadas da estrutura semântica genolexical, concebe que cada afixo contém um semantismo que é decomponível em subcomponentes. Esses subcomponentes são despertados por subcomponentes semânticos das bases. São esses subcomponentes que são responsáveis por matizes semânticos observáveis em produtos da mesma RFP, mas formados por operadores afixais distintos. Neste modelo, provê-se o operador afixal e a base de igual importância para o produto final.

REFERÊNCIAS

- Academia das Ciências de Lisboa e Instituto de Lexicologia e Lexicografia 2001. *Dicionário da língua portuguesa contemporânea*. Lisboa: Editorial Verbo, 2 vols.
- Beard, R. 1995. *Lexeme-morpheme base morphology: a general theory of inflection and word formation*. New York: State University of New York Press.
- Bluteau, R. 1712-1728. *Vocabulário português e latino*. Coimbra: Collegio das Artes da Companhia de Jesus, 8 vols. e 2 suplementos.
- Corbin, D. 1987. *Morphologie dérivationnelle et structuration du lexique*. Vol. 1. Tübingen, Max Niemeyer Verlag.
- Corbin, D. 1991. Introduction. La formation des mots: structures et interprétations. *Lexique*. **10**: 7-30.
- Corbin, D. 2004. Français (Indo-européen: Roman). In: G. Booij; C. Lehmann; J. Mugdan, ; S. Skopeteas (Eds.). *Morphologie/Morphology Ein internationales Handbuch zur Flexion und Wortbildung/ An international handbook on inflection and derivation*. Vol. 2. Berlin; New York: Walter de Gruyter.
- Corbin, D.; Corbin, P. 1991. Un traitement unifié du suffixe *-ier(e)*. *Lexique*. **10**: 61-145.
- Dicionário da língua portuguesa*. Porto: Porto Editora. www.infopedia.pt.
- Grossmann, M. 1998. Formazione dei nomi di agente, strumento e luogo in catalano. *Atti del XXI Congresso Internazionale di Linguistica e Filologia Romanza*. Tübingen: Max Niemeyer Verlag, 383-392.
- Houaiss, A.; Villar, M.; Franco, F. (Eds.) 2002-2003. *Dicionário Houaiss da língua portuguesa*. Lisboa: Círculo de Leitores, 6 vols.
- Lieber, R. 2004. *Morphology and lexical semantics*. Cambridge: Cambridge University Press.
- Plag, I. 1999. *Morphological productivity. Structural constraints in English derivation*. Berlin/ New York: Mouton de Gruyter.
- Plag, I. 2004. Syntactic category information and the semantics of derivational morphological rules. *Folia Linguistica*. **38(3-4)**: 193-225.
- Rio-Torto, G. 1993. Formação de palavras em português. Aspectos da construção de avaliativos. Dissertação de doutoramento apresentada à Universidade de Coimbra. Ms Coimbra.
- Rio-Torto, G. 1998. *Morfologia derivacional. Teoria e aplicação ao português*. Porto: Porto Editora.
- Rodrigues, A. 2001. - *A construção de postverbaís em português*. Porto: Granito Editores e Livreiros.
- Rodrigues, A. 2002. Para compreender o mecanismo de formação dos chamados “derivados regressivos”. In: I. Duarte et al (Eds.). *Encontro Comemorativo do 25º Aniversário do Centro de Linguística da Universidade do Porto* Vol. 1. Porto: CLUP/FCT/FLUP: 9-19.
- Rodrigues, A. 2008. *Formação de substantivos deverbaís sufixados em português*. München: Lincom.
- Vieira, D. 1871-1874. *Grande dicionário Portuguez ou Thesouro da Lingoa Portugueza*. Porto: Editores Ernesto Chardron e Bartholomeu H. de Moraes, 5 vols.

Recensões

Georgeta CISLARU, Frédéric PUGNIERE-SAAVEDRA,
Frédérique SITRI (Coord.). *Analyse de discours et demande
sociale. Le cas des écrits de signalement*. Paris: Presses
Sorbonne Nouvelle. Les Carnets du Cediscor, n° 10. 2008.

172 pp.

ISBN 978-2-87854-430-5

Isabel Margarida Duarte
Faculdade de Letras da Universidade do Porto –
Centro de Linguística da Universidade do Porto (Portugal)
iduarte@letras.up.pt

Esta publicação dá conta de resultados de investigação de um grupo conceituado no campo da Análise do Discurso, o CEDISCOR (a saber, *Centre de recherche sur les discours ordinaires et spécialisés*), fundado por Sophie Moirand e acolhendo investigadores de várias universidades, com o intuito de estudar um *corpus* de textos variados, à luz do enquadramento teórico de uma linguística do discurso. O grupo de pesquisa pretende mostrar, nesse *corpus*, a partir de elementos linguísticos, como o sentido se inscreve na materialidade dos textos, descrevendo também os seus funcionamentos.

O livro em apreço é o número 10 de *Les carnets du Cediscor*, publicação não periódica em que os investigadores do centro reúnem, tematicamente, os resultados da sua investigação levada a cabo, neste caso particular, a partir da observação de práticas redaccionais dos trabalhadores sociais que lidam com crianças em risco e respectivo contexto institucional.

O volume é composto por um «Avant-propos» em que se apresentam os diferentes textos e os respectivos objectivos de investigação, três partes, um apartado de referências bibliográficas comuns aos sete textos publicados, bem organizado, minucioso e actualizado, um

Glossário das siglas utilizadas, dois índices (de noções e de autores citados), resumos em francês e inglês, informações sobre os autores dos trabalhos e sobre o Centro de investigação. Em síntese, o livro apresenta um conjunto de informações complementares úteis e bem organizadas.

Poderíamos considerar a pesquisa de que se dá conta neste livro sob a designação genérica de Linguística Aplicada, uma vez que ela é fruto da aplicação de conhecimentos linguísticos à análise de textos reais vindos de um contexto profissional muito próprio: trata-se de analisar relatórios de sinalização que vários serviços sociais fazem a respeito de crianças em perigo. Os estudos reunidos configuram, portanto, um diálogo entre a Análise de Discurso e a sociedade, quer quanto à constituição do *corpus* e à formulação de hipóteses, quer quanto à escolha dos observáveis, quer quanto àquilo a que os autores chamam «gesto interpretativo», quer ainda quanto à colaboração com profissionais da área da intervenção social cujos relatórios são trabalhados. Configura-se, portanto, uma evidente relação do discurso com um «exterior» histórico, político e social, condicionado por tempos e actores concretos. A interacção linguistas / sociedade está patente nos diferentes textos, inscrevendo-se num movimento mais global de aproximação entre pesquisa e mundo profissional: « on voit des linguistes se préoccuper des applications sociales de leurs travaux ou des acteurs sociaux solliciter les linguistes dans des domaines aussi variés que la didactique [...], le traitement automatique des langues [...], le fonctionnement du langage en situation de travail [...]» (p.12).

A investigação destacou, no *corpus* de relatórios estudados, três tipos de observáveis: (1) a categorização dos actores e dos factos; (2) as marcas linguísticas das expectativas sociais do autor do relatório; (3) o posicionamento enunciativo de quem escreve.

A primeira parte do volume intitula-se «Contraintes et particularités des écrits professionnels dans le domaine social» e os dois textos que a compõem problematizam quer o objecto de pesquisa (os escritos de sinalização de crianças em risco) quer questões metodológicas levantadas pela investigação. A segunda parte, igualmente constituída por dois textos, chama-se «L'évaluation: aspects lexicaux et syntaxiques» e reúne estudos que pretendem fazer a articulação entre

formas linguísticas concretas e estratégias discursivas dos autores dos textos analisados, tendo sempre em conta que a finalidade pragmática dos relatórios estudados é avaliar uma situação precisa. Por fim, na terceira e última parte, «Les représentations entre subjectivité et intersubjectivation», estão agrupados três textos que se debruçam sobre a representação dos dizeres e das emoções.

O primeiro texto, «Quelques régularités des écrits de signalement» (p.21-36), da autoria de Frédéric Pugnière-Saavedra, pretende analisar o respectivo *corpus* a partir da intenção pragmática dos textos estudados, das suas regularidades composicionais e das representações linguísticas dos diferentes actores. Começa por distinguir escritos profissionais, com uma intencionalidade eminentemente prática e escritos informais, administrativos e religiosos. Dentro do hipergénero dos «escritos profissionais», selecciona o género «escritos profissionais do domínio social» e o subgénero «escritos de sinalização», cuja função é avaliar uma dada situação para proteger um menor em risco. Depois de proceder a esta categorização, o estudo procura encontrar regularidades composicionais nos relatórios que se prendem, quer com as expectativas sociais dos profissionais que os redigem, quer com a estrutura e as regras a que devem obedecer tais relatórios, quer ainda com as representações linguísticas dos actores em causa, sublinhando as estratégias de designação (designação da criança sinalizada por um termo genérico, com a ajuda de um sufixo diminutivo, por recategorização jurídica, por recategorização através de um artigo indefinido ou, na parte final dos dossiês, por recategorização por meio de um discurso «psicologizante»). São ainda estudadas outras marcas linguísticas dos actores: o «nós» equivalente à instituição que avalia a situação do menor (configurando o apagamento de um enunciador preciso), os verbos de «demande de faire» situados na parte final dos textos, que incitam os juízes à acção, as estruturas passivas que assinalam a instituição administrativa ou judicial como a mais adequada para dar resposta à situação de risco diagnosticada, o uso do lexema «disfonctionnement» enquanto recategorização metafórica de elementos anteriormente reportados e do lexema «situation» quer como recategorização dos relatos de estados que se vão degradando, quer como forma de caracterização das famílias em causa. Este último lexema é seguido, geralmente, de elementos negativos, como

«situação *violenta*», «situação *que se degrada*», etc. Os dois lexemas que o texto estuda aparecem sobretudo na parte final dos relatórios analisados e configuram aquilo a que o autor chama expressivamente «*épaisseur synthétique*», interiorizada pelo trabalhador social que redige os documentos. Pugnière-Saavedra salienta a forte orientação argumentativa do subgénero discursivo estudado.

No segundo texto, são focalizadas as estratégias linguísticas e as precauções enunciativas de quem redige os relatórios dirigidos aos juizes. A preferência pela palavra «rencontre» é explicada, bem como o já referido apagamento do enunciador por trás de um «nós» que refere a instituição. Patrick Rousseau, cuja tese de doutoramento este trabalho, em parte, retoma, analisa ainda o significado do domínio da palavra «aide» em detrimento do nome «contrôle», o recurso ao subentendido quando o conteúdo do relatório é dificilmente comunicável, a heterogeneidade dos diferentes discursos relatados, várias formas de modalização típicas deste subgénero discursivo, sobretudo a modalização como arte do compromisso e como sinal de honestidade intelectual (não ter a certeza seria uma marca de honestidade), usada como forma de proteger quem escreve de eventuais reponsabilizações profissionais posteriores.

O primeiro texto da segunda parte «Ecrire «ce qui ne va pas» dans le champ de l'enfance en danger: les mots *problème(s)* et *difficulté(s)*» (p. 57-77), situa-se dentro duma perspectiva semântica-discursiva que articula a análise lexicométrica com a análise discursiva. Procura problematizar a relação do locutor com o seu objecto, no caso em apreço a partir da captação do sentido dos dois lexemas (*problema* e *dificuldade*) quando actualizados no discurso, tentando explicar diferenças e contextos de uso, a partir da constatação do carácter vago dos respectivos significados: «un sens référentiel faible et un grand pouvoir de capture extensionnelle» (p. 58). Os lexemas analisados são utilizados como anáforas resumativas mas com especificidades decorrentes de uma elaboração sociodiscursiva própria deste tipo de escritos e, mais frequentemente, nos momentos de análise do que em sequências de relato de factos. Há mais ocorrências da palavra «dificuldade» do que de «problema», preponderância que Marie Veniard explica, não por qualquer hierarquização lexical entre os termos, mas a partir da noção de *habitus discursif*, ou seja,

o funcionamento discursivo habitual de uma dada palavra, o seu contexto léxico-sintáctico e discursivo, analisando a autora fenómenos linguísticos específicos como colocações, regências preposicionais, modalidades (afirmativa ou negativa) e a heterogeneidade discursiva. Os lexemas são analisados e diferenciados a partir dos seguintes critérios: a dimensão tipológica dos sintagmas de que fazem parte; a quantificação; a relação com a pessoa e, finalmente, a expressão da negação. «Dificuldade» usa-se sobretudo para referir entaves interiores; «problema» reporta-se a impedimentos exteriores, colocando em foco a relação do sujeito com uma norma social. A ocorrência dos lexemas em causa em relatos de discurso não é inocente e a autora estuda quer a nominalização no discurso relatado, quer no discurso indirecto canónico com completiva, quer num discurso indirecto com complementação nominal de tipo resumo. A palavra «problema», no discurso relatado, tem sempre por locutor a família da criança; o verbo «avoir» aparece como introdutor e «problema» ocorre com a modalidade negativa ou uma modalidades des-realizante como o uso de tempos do passado. Em resumo: «dificuldade» estaria mais do lado da concordância dos actores, «problema» mais do lado da recusa da família em colaborar com os técnicos.

«L'évaluation dans les rapports de signalement» (p. 79-91), da autoria de Sylvie Garnier, estuda sobretudo a estrutura concessiva e a sua participação na construção das propriedades dos objectos discursivos, mostrando como os marcadores concessivos convertem o narrativo em argumentativo, conduzindo a opinião de quem lê, o que configura uma finalidade prática da instituição que produz os relatórios.

O texto que abre a terceira parte do volume, «Observer et évaluer dans les rapports éducatifs: de la représentation d'un dire singulier à la description de la situation» (pp. 95-116), debruça-se *grosso modo* sobre o relato de discurso, aqui designado por RDA (Représentation du Discours Autre), de forma a perceber a imbricação dos enunciados que remetem para acontecimentos de palavra, os diferentes modos, mais ou menos marcados, de reportar os ditos de outros enunciadores pelo autor dos relatórios, os segmentos introdutores do relato, sobretudo os verbos de palavra e os epistémicos que introduzem uma modalização da asserção. Os textos analisados são contextualizados

no que concerne ao respectivo dispositivo enunciativo e à finalidade pragmática que é comum a todos. Importa considerar, nesta contextualização, a rubrica do relatório em que aparecem os relatos de palavras, e que são sobretudo os apartados onde se resumem as interações entre técnicos e famílias. Nesses momentos do relatório, as sequências relatadas podem ser bem delimitadas ou meras sequências descritivas ou narrativas interpretáveis como a representação de um acontecimento de palavras. No que tange ao enunciado relatado marcado, é preponderante o discurso indirecto, por ser aquela forma de relato que melhor permite contar ou resumir os ditos, por vezes com palavras e expressões entre aspas, e até com marcas de oralização. Os enunciados em que o relato não é marcado revelam mais a presença do relator, com mais nominalizações, categorizações decorrentes de um interdiscurso psicologizante e socioeducativo próprio dos técnicos em causa. Sitri chama a atenção para o facto de a construção nominal revelar um grau mais elevado de reformulação por parte do relator do que a construção com completiva ou infinitiva previsíveis no discurso indirecto. Por seu lado, o complemento nominal do verbo *dicendi* é frequentemente negativo (do ponto de vista do lexema usado, do emprego de sufixos de privação, da presença de quantificadores que restringem o espectro de um nome de valoração positiva). O artigo termina com uma terceira parte muito interessante sobre o emprego do tempo presente e os seus vários valores nos relatórios em apreço. A atenção a aspectos linguísticos muito finos mostra o trabalho difícil dos redactores a braços com um destinatário heterogéneo (as famílias visadas têm acesso aos relatórios) e uma situação humana complexa que os implica e exige uma actuação por vezes urgente.

O segundo texto, «L'intersubjectivation des émotions comme source de sens: expression et description de la peur dans les écrits de signalement» (pp. 117-136), baseia-se nos pressupostos de Damásio (1995) de que as emoções são necessárias para a tomada de decisões e de Plantin (2004), que advoga a inseparabilidade, na argumentação, da emoção e do raciocínio. A emoção terá, segundo Cislaru, uma função discursiva nos textos escrutinados e as suas expressão e descrição configuram semanticamente os textos. Assim, a autora estudo o campo lexical do medo e identifica a localização dos respectivos elementos nos textos de sinalização. Analisa as escolhas

discursivas e lexicais dos autores dos relatórios e as configurações enunciativas e sintácticas dos textos. Por fim, tira conclusões sobre a função pragmática da expressão do medo, que seria, sobretudo, a de condicionar o posicionamento da instituição em relação a uma dada situação considerada preocupante. O texto conclui com uma reflexão sobre os conceitos de empatia e intersubjectivização, mostrando como a mediação emocional pode ser um instrumento de coesão textual.

O último artigo do volume, «Les mots qui blessent» (pp.137-143), estuda sobretudo a violência sentida pelas famílias quando lêem os relatórios que lhes dizem respeito e o conseqüente uso de eufemismos por parte dos técnicos que querem informar e fazer agir, sem ferir.

O conjunto dos sete estudos, precedido do «Avant-propos» e dos documentos já referidos acima revela-se ainda mais interessante do que a soma das suas partes, uma vez que mostra como a Linguística (neste caso a Análise do Discurso) pode ser aplicada, socialmente útil, iluminando e problematizando o trabalho de outros profissionais, por um lado; mas é também um interessante exemplo de como os textos (neste caso, os escritos de sinalização de crianças em risco) são mais eficazmente lidos e compreendidos quando se utilizam instrumentos teóricos rigorosos e objectivos.

REFERÊNCIAS

Damáσιο, A. 1995. *O erro de Descartes*. Lisboa: Europa-América.

Plantin, C. 2004. On the inseparability of emotion and reason in argumentation. In: E. Weigard (Ed.) *Emotions in Dialogic Interactions*. Amsterdam: John Benjamins, 265-276.

Robert SPENCE. *A Functional Approach to Translation Studies: New systemic linguistic challenges in empirically informed didactics*. Berlin: dissertation.de. 2004. 286pp.
ISBN 3-89825-777-0

Thomas J.C. Hüsgen
Faculdade de Letras da Universidade do Porto –
Centro de Linguística da Universidade do Porto (Portugal)
thomsgen@sapo.pt

A monografia de Robert Spence, com o título *A Functional Approach to Translation Studies: New systemic linguistic challenges in empirically informed didactics*, representa a publicação da sua tese de doutoramento na área dos Estudos da Tradução apresentada na Universidade de Leipzig (Alemanha). O autor, docente de Teoria da Tradução nesta mesma universidade, aplica neste estudo o paradigma proposto pelo contextualismo britânico ao texto traduzido com o intuito declarado de provar a aplicabilidade desta perspectiva aos estudos da tradução. Em resposta ao estruturalismo americano e europeu, esta abordagem desenvolvida por J.R. Firth procurou ultrapassar a análise estática e isolada da língua definindo o significado de unidades formais a partir de situações autênticas e contextualizadas. Na sequência desta proposta, na tradição das teorias semânticas pragmáticas de cariz wittgensteineano, Halliday, McIntosh e Strevens (1964: 18ss.) desenvolveram um modelo descritivo estratificado, distinguindo o nível da substância (fonética e fonologia), dos níveis da forma (gramática e léxico) e do contexto (semântica e situação extra-linguística), e lançando o fundamento para os estudos que se propõem descrever o significado em termos estritamente linguísticos. E é justamente este objectivo programático que permite estabelecer, de uma forma inequívoca e declarada, objectivos comuns entre os estudos linguísticos

e os estudos da tradução. Tendo surgido no mundo ocidental no pós Segunda Grande Guerra como sub-disciplina linguística, os estudos de tradução procuraram ganhar a sua autonomia no final dos anos setenta, enquanto área do conhecimento independente. Para isso, procederam a uma mudança de paradigma em que a comparação e a contrastação entre as diferentes línguas em contacto eram substituídas na sua importância por perspectivas centradas em questões de ordem cultural e cognitiva. Recentemente, e com o ressurgimento da linguística computacional no âmbito dos estudos de tradução, tem-se observado uma reaproximação entre estas duas áreas do conhecimento. O estudo aqui apresentado é um bom exemplo para essa “reformulação estratégica” destas duas áreas afins.

Desde já terá que ser sublinhado que não se deve confundir na presente obra esta abordagem com os modelos funcionalistas da *translação*, baseados nas publicações de K. Reiss e H.J. Vermeer. Trata-se de uma tentativa, a nosso ver, bem conseguida, de reaproximar a análise linguística aos estudos da tradução, de certa forma interrompida pela escola da teoria do *escopo*, como ficou conhecido o modelo funcionalista alemão. De resto, o autor faz questão em afirmar, logo no início do seu estudo, o intuito de refazer a ponte entre a linguística aplicada e os estudos da tradução:

Although the present work is concerned with problems of translation, it is intended primarily as a contribution to applied linguistics and not as a contribution to translation studies. This distinction implies no criticism of the latter discipline, nor should the making of it be construed as a principled rejection of interdisciplinarity: it is precisely the aim of the present work to build a bridge between the two disciplines.[...] If at times it appears that the tensions between the opposing [...] perspectives of applied linguistics and translation studies are being deliberately exarcebated, the reader may rest assured that this police is being pursued with a sole desire: that the reconciliation of the two disciplines be swift, precise, and sustainable.

(p.1)

Seguindo uma estrutura clara, por vezes um pouco fragmentada, o autor propõe-se aplicar o modelo sistémico-funcional de M. A. Halliday ao estudo de um *corpus* linguístico produzido por estudantes

de tradução da Universidade de Leipzig. O *corpus* é composto por 52 textos de partida em língua alemã e 1232 textos de chegada em inglês, dos quais 1165 foram produzidos por 220 estudantes de língua materna alemã e outros 67 textos produzidos por 8 estudantes nativos do inglês. O objectivo principal da constituição deste *corpus* prende-se com o desenvolvimento de métodos científicos para a avaliação objectiva de traduções recorrendo ao conceito de interferência na explicação da maioria dos erros tradutivos.

Depois de apresentar no primeiro capítulo a justificação das suas opções teóricas, no segundo capítulo o autor explora minuciosamente os vários pontos de contacto entre a linguística aplicada e a teoria da tradução. Reitera a necessidade de um modelo analítico, no sentido duma didáctica da tradução, que oferece um método científico que não permita apenas descrever e explicar o processo tradutivo de uma forma adequada, mas que, para além disso, ofereça critérios objectivos e empiricamente sustentados para uma avaliação da qualidade de soluções tradutivas. Optando para esse efeito pela teoria sistémico-funcional, Spence explica no terceiro capítulo os principais conceitos do modelo, não deixando de apontar algumas das desvantagens e insuficiências desta abordagem na avaliação de textos traduzidos. No quarto capítulo, que funciona como introdução à parte empírica do trabalho, após uma apresentação mais detalhada do *corpus* analisado, o autor procura responder a possíveis dúvidas por parte do leitor acerca da aplicabilidade de um modelo analítico de cariz linguístico nos estudos da tradução, demonstrando que o modelo sistémico-funcional, em conjugação com ferramentas disponibilizadas pela linguística computacional, oferece enormes vantagens para o estudo da tradução. São estas vantagens que o autor tenta explorar nos capítulos quinto e sexto, onde procede a uma análise exaustiva do *corpus* apresentado. Finalmente, no capítulo sétimo, Robert Spence apresenta propostas para a aplicação didáctica do modelo por si desenvolvido na formação de tradutores.

Não há dúvida que este trabalho oferece uma contribuição válida na recolocação da análise linguística nos estudos da tradução que, desde os tempos da “viragem cultural” iniciada pelos trabalhos de Katharina Reiss e Hans J. Vermeer, passou a ter uma posição demasiadamente secundarizada. Na tentativa de criar a sua

autonomia em relação aos estudos linguísticos, a teoria da tradução tem vindo a menosprezar a perspectiva linguística nos seus modelos explicativos e só recentemente, e com a (re)emergência da linguística computacional, se tem assistido à “reabilitação” da análise linguística nos estudos da tradução. É à luz desta evolução que este trabalho aqui apresentado deve ser enquadrado – por um lado, no lançar de novas pistas para a avaliação de traduções, e, por outro, na proposta de aplicação de um modelo analítico eficaz no ensino da tradução, hoje, por vezes, demasiado “desverbalizado”.

Partindo de uma perspectiva muito influenciada pela análise de erros dos anos 70 e 80, apoiada em autores como König (1972), Leisi (1972), Nickel (1973), Legenhausen (1975), Clyne (1987), Wilss (1989) e Neubert (1989), para mencionar apenas alguns, Spence tenta suplantar uma heurística intuitiva, muito enraizada, a seu ver, na avaliação tradutiva, e elabora um esquema de classificação de erros que, no nosso entender, só peca por apresentar uma visão excessivamente centrada na frase como unidade de tradução. Apesar disso, o autor é bem sucedido na demonstração de que o paradigma sistémico-funcional fornece uma ferramenta de análise linguística válida e produtiva aplicada ao processo tradutivo, e que, apesar de focalizar em demasia na unidade de tradução ao nível da frase, abre novas perspectivas na avaliação e explicação de soluções tradutivas.

REFERÊNCIAS

- Clyne, M. 1987. Cultural Differences in the Organisation of Academic Texts. *Journal of Pragmatics*. **41**: 217-247.
- Firth, J. R. 1956. Descriptive Linguistics and the Study of English. In: F. R. Palmer (Ed.). *Selected Papers of J. R. Firth 1952-59*. London: Longman, 96-113.
- Halliday, M.A.K.; McIntosh, A.; Strevens, P. 1964. *The Linguistic Sciences and Language Teaching*. London: Longman.
- König, E. 1972. Fehleranalyse und Fehlertherapie im lexikalischen Bereich. In: G. Nickel (Ed.). *Fehlerkunde: Beiträge zur Fehleranalyse, Fehlerbewertung und Fehlertherapie*. Berlin: Cornelsen-Velhagen & Klasing, 73-77.
- Legenhausen, L. 1975. *Fehleranalyse und Fehlerbewertung*. Berlin: Cornelsen-Velhagen & Klasing.
- Leisi, E. 1972. Theoretische Grundlagen der Fehlerbewertung. In: G. Nickel (Ed.). *Fehlerkunde*. Berlin: Cornelsen-Velhagen & Klasing, 25-37.

- Wilss, W. 1989. Interferenzerscheinungen beim Übersetzen. Fremdsprache-Grundsprache / Vorschläge zu einer prozeduralen Analyse. In: H. Schmidt (Ed.). *Interferenz in der Translation*. Leipzig: Verlag Enzyklopädie, 7-18.
- Neubert, A. 1989. Interference between Languages and between Texts. In: H. Schmidt (Ed.). *Interferenz in der Translation*. Leipzig: Verlag Enzyklopädie, 56-64.

João de Morais Madureira FEIJÓ. *Ortografia, ou Arte de escrever e pronunciar com acerto a Língua Portuguesa – 1734 –*.
Edição semidiplomática com índice de todas as formas
[Telmo Verdelho, João Paulo Silvestre e Isabel Prates, eds.].
Aveiro: Universidade de Aveiro. 2008.
395 pp. ISBN: 978-972-789-270-9

Rogelio Ponce de León Romeo
Faculdade de Letras da Universidade do Porto (Portugal)
rromeo@letras.up.pt

Durante as últimas décadas, têm vindo a ser divulgados, em forma de edições, gramáticas, vocabulários e outras obras de índole metalinguística; no que toca à edição de textos metaortográficos, é mister salientar a edição fac-similada e a *leitura crítica* – mais do que discutível – que Maria Leonor Carvalhão Buescu realizou, respectivamente, das *Regras que ensinam a maneira de escrever e a ortografia da língua portuguesa* (1981[1574]) de Pêro de Magalhães de Gândavo e da *Ortografia e origem da língua portuguesa* (1983[1576]¹) de Duarte Nunes de Leão; das *Regras* de Gândavo, foi, ainda, levada a cabo uma edição crítica por Rolf Nagel (1969). Na esteira, precisamente, dos trabalhos crítico-textuais – em certos casos, penosos, ingratos e nem sempre bem sucedidos... – que acabamos de mencionar, foi editada em 2008 a *Ortografia, ou Arte de escrever e pronunciar com acerto a Língua Portuguesa* (Lisboa 1734), de João de Morais Madureira Feijó (1688-1741), ao cuidado de um grupo de

Linguística - Revista de Estudos Linguísticos da Universidade do Porto - Vol. 4 - 2009, pp. 169-173.

¹ 1576 é a data da primeira edição da *Ortografia*; a *Origem da língua portuguesa*, como é bem sabido, foi publicada pela primeira vez em 1606, da qual há uma edição anotada – que não nos foi possível consultar – por José Pedro Machado (1945).

especialistas na historiografia linguística portuguesa da Universidade de Aveiro; referimo-nos a Telmo Verdelho, João Paulo Silvestre e Isabel Prates, investigadores que estão a realizar um trabalho de excelência, nomeadamente no campo da história da lexicografia do Português, concretizado no projecto *Corpus Lexicográfico do Português* “que trabalha sobre o texto antigo, particularmente sobre o texto dicionarístico, promovendo a sua edição e o tratamento em base de dados”².

O aparecimento editorial da *Ortografia* – da qual se publicou um ensaio pioneiro da autoria de Filomena Gonçalves (1992) e que tem sido objecto de atenção nas monografias dedicadas ao tema, mostrando o seu impacto na produção metaortográfica posterior (Kemmler 2001: 205-219; Gonçalves 2003: 44-50) – de Feijó parece-nos duplamente relevante: pela obra e pelo autor. Por um lado, a obra gozou de uma grande difusão, materializada no número considerável de edições até 1861. João Paulo Silvestre – autor da breve introdução –, citando Rolf Kemmler (2001: 300-307) – um dos maiores especialistas na história das ideias ortográficas do Português –, contabiliza “12 edições oficiais e 5 impressões clandestinas³, descrevendo [Rolf Kemmler] pelo menos 6 versões da segunda edição [Coimbra, 1739] e 2 da oitava [Lisboa, 1814]” (p. 14). Quanto ao autor, pode afirmar-se que o trabalho gramatical transcendeu o labor metaortográfico; aliás, parece que a sua *Ortografia* constitui sobretudo um complemento – para certos investigadores, mais do que um mero complemento⁴ – à sua gramática – oficialmente, uma explicação em língua portuguesa à gramática latina de Manuel Álvares –, cuja primeira edição se intitulou *Explicationes in omnes partes totius Artis R. P. Emmanuelis Alvari* (Lisboa 1729), ao passo que, nas seguintes edições, saiu com o título de *Arte explicada*.

² Tomamos a citação da página web do *Corpus Lexicográfico do Português* (<http://clp.dlc.ua.pt/Inicio.aspx>).

³ As obras linguísticas de Feijó foram proibidas, como é bem conhecido, com outras gramáticas – entre as quais cabe destacar os *De institutione grammatica libri tres* (Lisboa 1572) de Manuel Álvares –, dicionários e cartapácios redigidos por autores jesuítas, no Alvará régio de 28 de Junho de 1759 (Verdelho 1982: 357).

⁴ Rolf Kemmler afirma que “a *Orthographia* de Feijó não é uma obra totalmente autónoma: trata-se do quarto volume de um comentário da obra *De institutione grammatica libri tres* do jesuíta Manuel Álvares (1572), chamado *Arte explicada*” (2001: 207).

Neste sentido, importa, na *Ortografia*, pôr em relevo um aspecto que não parece ter suscitado o interesse dos investigadores; a saber, a prescrição de tipo exclusivamente gramatical – e não ortográfico – em certos passos da obra; em concreto, na *Terceira parte*, dedicada aos *Erros do vulgo, e emendas da orthografia no escrever, e pronunciar* (pp. 112-124), na qual se enquadra, ocupando quase todo o capítulo, um tratado sobre morfologia verbal, intitulado *Advertencia necessaria para a conjugação dos verbos* (pp. 115-124), que Feijó justifica da seguinte forma:

Como a maior parte de erros que andão introduzidos na pronunciaçãõ, e locuaçãõ do vulgo, nasce de não saberem conjugar os verbos, nem differençarem as suas linguagens, ou diversos modos de significar por *Tempos, numeros, e pessoas*; he preciso darmos aqui huma breve noticia dos *Verbos*, e das suas *Conjugaçoens*; e conjugar alguns, que nos sirvaõ de exemplares para huns, e de excepçãõ para outros (p. 115).

Ora, se noutro estudo púnhamos em relevo que a descrição gramatical que Feijó faz da classificação verbal nas suas *Explicationes / Arte explicada* obedece a uma aproximação da cacterização do sistema latino ao português (Ponce de León 2005: I, 454), na *Advertencia necessaria*, Feijó esboça um breve capítulo centrado de forma exclusiva na explicação do verbo português, extremamente interessante – em nossa opinião – para o estudo das ideias gramaticais no Portugal do século XVIII⁵.

No que toca à edição propriamente dita, os investigadores acima referidos levaram a cabo um trabalho muito cuidado de fixação do texto, com base num exemplar da edição príncipe (localizado na Biblioteca Nacional de Portugal, cota L. 5049 A), procedendo a uma leitura semidiplomática: “Com o objectivo de preservar o testemunho local, bem como a coerência entre o discurso metaortográfico do autor

⁵ A introdução de capítulos exclusivamente gramaticais nos tratados metaortográficos não é inédita na história das ideias linguísticas em Portugal; aparece, por exemplo, um tratado, intitulado *Das partes da vulgar lingua*, sobre as classes de palavra em Português na *Ortografia da lingua portugueza* (Lisboa 1671) de João Franco Barreto (Ponce de León 2006); dezenas de anos mais tarde, Manuel Coelho de Sousa incluirá na obra morfológica do latim – a *Explicação das partes da oração, com todas as circumstâncias, etymologias e intelligencias* (Lisboa 1721), um capítulo dedicado as classes de palavra em Português – a *Explicação das oyto partes da oração da lingua Portugueza* –.

e os exemplos, optou-se por uma edição semidiplomática, mantendo-se as grafias originais, a pontuação e os diacríticos” (p. 14). Parece claro que o objectivo dos editores foi reproduzir de forma exclusiva o texto da edição de 1734... Talvez tivesse sido oportuno, por forma a determinar as divergências textuais – e, por ventura, a evolução do pensamento metaortográfico do autor –, confrontar as edições publicadas em vida de Feijó – isto é, Lisboa 1734 e Lisboa 1739 –, com os problemas crítico-textuais com que os editores se podiam inevitavelmente deparar – a saber: as impressões clandestinas que têm como data de edição 1739, mas que terão sido posteriores a esse ano e, portanto, não pertinentes para determinar a eventual evolução das propostas ortográficas de Feijó –. Teria sido igualmente oportuno que os editores tivessem elaborado índices de nomes e matérias; contudo, é necessário realçar a apresentação de um exaustivo índice de formas (pp. 305-395), extraordinariamente valioso para os investigadores na história da ortografia e na história da lexicografia do Português. Por outro lado, julgamos extremamente acertada o critério dos editores, no tratado intitulado *Erros communs da pronunciaçam do vulgo, com as suas emendas em cada letra*, no sentido de introduzir “entradas remissivas com ortografia actualizada, assinaladas entre parênteses rectos, sempre que as grafias originais impliquem uma localização diferente na ordenação alfabética” (p. 15).

Por tudo o que foi até ao momento referido, devemos saudar a iniciativa da divulgação de um texto sobre a língua portuguesa, fundamental para história das ideias ortográficas, mas também para a história da lexicografia, na medida em que “a extensa lista de formas correctas e erradas, organizada por ordem alfabética, foi precursora dos dicionários práticos da língua portuguesa” (p. 7); por outro lado, alguns capítulos da ortografia têm grande interesse para os especialistas na gramaticografia do Português. Esperamos, por conseguinte, que os editores da *Ortografia* de João de Moraes Madureira Feijó, Telmo Verdelho, João Paulo Silvestre e Isabel Prates, continuem a oferecer-nos os valiosos frutos da sua investigação.

REFERÊNCIAS

- Gândavo, P. de M. 1969[1576]. Die Orthographieregeln des Pêro Magalhães de Gândavo (Rolf Nagel, ed.). *Aufsätze zur Portugiesischen Kulturgeschichte*. **9**: 111-135.
- Gândavo, P. de M. 1981[1574]. *Regras que ensinam a maneira de escrever e a ortografia da língua portuguesa: com o diálogo que adiante se segue em defesa da mesma língua* (Maria Leonor Carvalhão Buescu, ed.). Lisboa: Biblioteca Nacional.
- Gonçalves, M. F. 1992. *Madureira Feijó, Ortografista do Século XVII. Para uma História da Ortografia Portuguesa*. Lisboa: Ministério da Educação/Instituto de Cultura e Língua Portuguesa.
- Gonçalves, M. F. 2003. *As ideias ortográficas em Portugal. De Madureira Feijó a Gonçalves Viana (1734-1911)*. Lisboa: Fundação Calouste Gulbenkian/Fundação para a Ciência e a Tecnologia.
- Kemmler, R. 2001. Para uma história da ortografia portuguesa: o texto metaortográfico e a sua periodização do século XVI até à reforma ortográfica de 1911. *Lusorama. Zeitschrift für Lusitanistik. Revista de Estudos sobre os Países de Língua Portuguesa*. 47-48: 128-319.
- Leão, Duarte Nunes de. 1945[1606]. *Origem da língua portuguesa* (José Pedro Machado, ed.). Lisboa: s. n.
- Leão, Duarte Nunes de. 1983[1576]. *Ortografia e origem da língua portuguesa* (Maria Leonor Carvalhão Buescu, ed.). Lisboa: Imprensa Nacional-Casa da Moeda.
- Ponce de León, Rogelio. 2005. Os verbos em confronto: considerações sobre a tipologia verbal nas Artes gramaticais portuguesas setecentistas (1699-1758). *Gramática e Humanismo. Actas do Colóquio de Homenagem a Amadeu Torres* (Miguel Gonçalves, Augusto Soares da Silva, Jorge Coutinho, José Cândido Martins, Maria José Ferreira, eds.). Braga: Publicações da Faculdade de Filosofia/ Universidade Católica Portuguesa. I, 449-464.
- Ponce de León, Rogelio. 2006. A gramática na ortografia: o caso da *Ortografia da língua portuguesa* (Lisboa 1671) de João Franco Barreto. *Lusorama. Zeitschrift für Lusitanistik. Revista de Estudos sobre os Países de Língua Portuguesa*. **65-66**: 47-63.
- Verdelho, Telmo. 1982. Historiografia linguística e reforma do ensino. A propósito de três centenários: Manuel Álvares, Bento Pereira e Marquês de Pombal. *Brigantia. Revista de Cultura*. **2.4**: 347-383.

Johannes KABATEK (Ed.). *Sintaxis histórica del español y cambio lingüístico: Nuevas perspectivas desde las Tradiciones Discursivas*. Frankfurt am Main/Madrid: Vervuert/Iberoamericana — Serie Lingüística Iberoamericana 31. 2008. 275 pp.
ISBN: 9788484893455

Clara Barros
Faculdade de Letras da Universidade do Porto
– Centro de Linguística da Universidade do Porto (Portugal)
mbarros@letras.up.pt

A obra em análise é constituída por uma introdução, da autoria do editor, e dez artigos de autoria individual. Na origem do conjunto de trabalhos agora publicados esteve um encontro de especialistas de sintaxe histórica do espanhol, que teve lugar em Tübingen, no ano de 2005. São retomadas e desenvolvidas no presente volume as inovadoras concepções teóricas e metodológicas apresentadas numa obra anterior, da mesma colecção (Jacob & Kabatek, Eds., 2001), igualmente colectiva e igualmente resultante de um colóquio de hispanistas (Universidade Humboldt de Berlim, 1999).

Na *Introdução* (pp.7-16), Joannes Kabatek define a metodologia e suas implicações teóricas e apresenta muito sucintamente o teor dos dez trabalhos reunidos neste volume que constituem contributos de investigadores de diversas universidades (figuram, no final do livro, indicações precisas sobre os contactos de todos os autores).

Começa por fazer uma breve definição do conceito de *tradições discursivas* que teve o seu início na romanística alemã, a partir da teoria proposta por E. Coseriu. Sublinha a necessidade de integrar este conceito na teoria linguística, apresentando uma “hipótese forte” da

teoria das tradições discursivas: a história de uma língua varia também de acordo com as tradições dos textos e estas podem condicionar a selecção de elementos linguísticos adoptados a partir de outros sistemas. Defende, portanto, que deve considerar-se redutora uma linguística histórica que não tenha em conta a diversidade textual. Faz ainda aproximações deste conceito ao conceito de género e a conceitos propostos por classificações tipológicas de textos. O objectivo principal é a ampliação da teoria linguística pela inclusão da tradição textual nos estudos linguísticos, contemplando a tradição de textos concretos, a tradição de formas textuais, de elementos, lugares e tópicos. Kabatek propõe mesmo a integração do conceito de tradições discursivas não só na linguística histórica mas também na sincrónica. Chama a atenção para o facto de na história das línguas só existirem textos e tradições textuais e aponta a particular relevância das tradições discursivas para a análise dos percursos de gramaticalização da sintaxe histórica, em particular o movimento de integração de fenómenos periféricos.

Os artigos compilados centram-se na análise de diversos fenómenos que ilustram a necessidade e a utilidade da inclusão do conceito de *tradições discursivas* na linguística histórica em geral e em particular nas questões de sintaxe histórica, não considerada isoladamente mas em ligação com fenómenos de evolução morfológica e lexical. No seu conjunto, os textos dos diferentes autores defendem que, como refere Joannes Kabatek na *Introdução*, a tomada em conta da tradição dos textos nos estudos de sintaxe histórica ou até de sintaxe em geral deveria tornar-se um elemento fundamental da investigação linguística e não 'uma ideia exótica'. A corrente de linguística histórica que integra o conceito de tradições discursivas estuda a actividade linguística de diferentes sujeitos e seus resultados nos textos e na própria língua em que são produzidos. Surgem novas formas de organização dos enunciados, com novos meios para exprimir relações e uma consequente ampliação dos meios de conexão intra e interfrásica que apresentam também uma maior precisão semântica. No uso dos diferentes textos de tradições diversas, não só se vai enriquecendo progressivamente o inventário de morfemas com funções relacionais, mas também as suas possibilidades combinatórias.

Os dois primeiros artigos abordam claramente aspectos teóricos da questão das tradições discursivas dentro da teoria da linguagem.

No primeiro – *Gramaticalización, Género discursivo y otras variables en la difusión del cambio sintáctico* (pp.17- 52) – , da autoria de Concepción Company Company, a A. analisa os novos valores de certas formas e o caminho que percorrem a partir do uso em contextos muito marcados ou específicos para a progressiva perda de restrições de emprego. Faz notar que nesses processos de gramaticalização – processos históricos pelos quais se criam novas formas ou novos mecanismos de expressão gramatical – não há uma imediata anulação dos usos tradicionais de acordo com a propriedade de estratificação de Hopper (que refere na p.18). Fundamentalmente, a A. determina neste trabalho as seis variáveis que podem incidir num percurso de gramaticalização:

1. Profundidade histórica ou antiguidade da mudança em questão.
2. Aspecto fónico da forma ou da construção inovadora.
3. Frequência de emprego das formas ou construções conservadora e inovadora.
4. Tipo de categoria que experimentou a mudança.
5. Tipo de sociedade que usa a língua.
6. Género textual ou tradição discursiva da manifestação do fenómeno.

As quatro primeiras variáveis são definidas como internas ao sistema enquanto que as duas últimas são consideradas externas ao mesmo. Na determinação do carácter de fenómeno interno ou externo ao sistema, a A. defende que o processo de gramaticalização é interno, mas a sua expansão depende do estrato social dos grupos que usam as novas formas. Fazendo apelo à necessidade de incorporar a dimensão social à gramaticalização, já sublinhada por Jacob em 2003, a A. abre o âmbito de estudo destes fenómenos de mudança numa direcção que foi encetada por numerosos trabalhos que se reclamam de uma sociolinguística histórica. Não só estuda a relação entre tradição discursiva e sintaxe histórica, mas discute esta questão no âmbito, e dentro dos parâmetros, de uma teoria integradora da mudança que a leva a identificar as já referidas seis variáveis centrais, sendo a última,

como se viu, a tradição dos textos. O artigo exemplifica cada uma das variáveis e evidencia a complexidade dos fenómenos de mudança histórica. Incita à reflexão sobre a questão da pertinência relativa das diferentes variáveis agora definidas, questionando a necessidade de estabelecer uma ordem de prioridades. Parece deixar à reflexão certas questões como a de saber se a 'tradição discursiva' será só mais um factor na mudança ou se terá um lugar fixo na hierarquia ou uma importância variável conforme o fenómeno de mudança. Este artigo dá sem dúvida algumas respostas sobre o processo de difusão da mudança linguística e debate aspectos centrais da integração do conceito de tradições discursivas na teoria linguística geral.

Também o artigo de Peter Koch – *Tradiciones discursivas y cambio lingüístico: el ejemplo del tratamiento “vuestra merced” en español* (pp. 53-88) – está direccionado para a reflexão teórica sobre a importância do estudo das tradições discursivas na mudança linguística, discutindo o lugar do conceito teórico de tradição discursiva na teoria da mudança linguística e da linguagem em geral. A transposição (“o transplante” pp.66-68) de determinadas regras discursivas para o romance podia fazer-se, na opinião do A. segundo dois percursos. No primeiro, o romance penetra numa tradição discursiva até então reservada ao latim como língua de distância. Por consequência, o romance procede por “imitação” dos procedimentos da tradição discursiva alheia. Fala-se, neste caso, de uma situação de contacto linguístico. Na segunda hipótese, o procedimento em questão passa da tradição discursiva original a outra tradição afim já aberta ao romance. Segundo o A., na história do castelhano e dos idiomas ibéricos em geral, desenham-se ambos os percursos como tenta mostrar aplicando os conceitos definidos à história das formas de tratamento em espanhol, desde o latim até à forma “usted”, que constitui um novo paradigma, fazendo uma transição em que a tradição discursiva e a filiação diplomática ocidental estabelecem um longo caminho entre a inovação esporádica e a generalização. Este trabalho analisa abundante exemplificação textual, de variada tipologia, de tradição literária e não literária, com alguma incidência em textos de natureza epistolar particularmente adequados ao estudo de formas de tratamento; utiliza trabalhos prévios realizados neste campo nomeadamente os de Wright, Primorac e Eberenz.

Os restantes oito artigos deste volume abordam prioritariamente questões metodológicas e aspectos concretos que tentam identificar tradições discursivas particulares ou relacionar as evoluções de certos fenómenos linguísticos com certas tradições discursivas.

Rafael Cano Aguilar, no terceiro artigo – *Los gramáticos españoles del Siglo de Oro: ¿Tradición discursiva, lengua especial?* (pp. 89-108)–, estuda uma tradição concreta que identifica como um “estilo particular” da gramaticografia espanhola a partir de Nebrija, cujas características incluem elementos léxicos, formulaicos, a tradição terminológica e também certos fenómenos sintácticos que, para além do estilo pessoal de cada autor e das tendências gerais observadas na língua da época, são identificáveis como pertencentes a essa tradição com forma linguística própria, inserida por sua vez noutra tradição afim desenvolvida na época.

No quarto artigo – *Apuntes para una caracterización de la morfosintaxis de los textos bíblicos medievales en castellano* (pp.109 - 126) –, Andres Enrique-Arias procura fazer um levantamento tipológico de traços sintácticos característicos dos romanceamentos bíblicos do castelhano medieval. Baseando-se na análise de uma série de versões de “El Libro de Isaías”, descreve certos traços sintácticos hebraizantes e outros arcaizantes que marcam a tradição bíblica (profética, neste caso) como claramente diferenciada de outras tradições medievais. Analisa num primeiro momento as estruturas gerais que ocorrem com certa frequência, destacando a elisão de cópula, as perguntas retóricas introduzidas por *si*, *cierto*, *por ventura* ou sem introdutor. E analisa também a tradução do participio activo semítico ou do participio nominal por construções com o participio presente, substantivos de verbais ou construções relativas. Estuda ainda estruturas particulares que ocorrem com menor frequência. O A. sublinha a importância da análise quantitativa dos dados, que poderá permitir uma avaliação do alcance de cada estrutura em cada texto. Os resultados deste estudo permitem um levantamento dos recursos linguísticos que se encontravam disponíveis na língua medieval revelados nas diferentes soluções encontradas pelos tradutores. A análise do estilo bíblico na língua românica medieval detecta a dependência de uma série de línguas de origem e a presença significativa de uma tradição, com a consequente dificuldade de adaptação a uma outra actualidade realidade.

O projecto explicitado neste trabalho está incluído num outro, mais amplo, de tratamento da linguagem bíblica, sediado na universidade de Tübingen; beneficia assim de colaboração na aplicação de análises quantitativas aos textos bíblicos.

Mario Barra Jover, no artigo *Tradición discursiva, creación y difusión de innovaciones sintácticas: la cohesión de los argumentos nominales a partir del siglo XIII* (pp.127-150), analisa a estrutura dos argumentos nominais em diversas línguas românicas no período medieval. Parte da hipótese de que no momento em que o latim fica nitidamente separado do romance, numa época que, embora diferindo um pouco de língua para língua, se encontra definitivamente estabilizada no período posterior ao século XIII, o romance pode utilizar o latim como adstrato e adoptar diferentes tradições discursivas que são, entre outras, fonte de inovações sintácticas. O A. descreve alguns fenómenos de coesão nominal que permitem formar linhas de coerência por relações coreferenciais ao longo de um texto de dimensões variáveis. Estes meios constituem mesmo um recurso que permite a organização de um texto extenso. Os elementos de coesão analisados procedem de uma tradição escrita, mas o seu uso difundido pode até penetrar na tradição oral, como acontece no caso de *o qual* e *o dito*.

Rolf Eberenz, no artigo "*Ninguno quiere del agua turbia beber*": *sobre construcciones partitivas y su representación en algunos géneros textuales del español preclásico* (pp.151-172), demonstra que a história de uma mudança — o uso de construções como "beber del agua" perdido no espanhol médio (também designado como o outono da Idade Média) para expressões não específicas — não corresponde a uma evolução linear mas depende em grande medida dos textos particulares e das respectivas tradições onde aparecem as construções partitivas. O A. critica as teorias da mudança linguística que observam de fora a evolução de elementos isolados sem diferenciar as tradições discursivas, preconiza a utilização de outras fontes de informação e afirma que o estudo da mudança linguística poderá sempre encontrar outra visão dos fenómenos se tiver em conta a diversidade textual.

O sétimo artigo desta obra — *Tradiciones discursivas y gramaticalización del discurso referido en el "Rinaldo de Palacio" y las "Crónicas" del Canciller Ayala* (pp. 173 -196) — é da autoria de José Luis Girón Alconchel. Debate a questão teórica do lugar das tradições

discursivas no âmbito da teoria da gramaticalização e, por outro lado, analisa concretamente o exemplo da gramaticalização do discurso relatado, partindo do discurso directo num contínuo entre oração recta, oração mista e oração oblíqua. Salaria que existe grande dependência da selecção dos meios não tanto em relação a factores diacrónicos mas sobretudo em relação à competência de distinguir entre diferentes tradições discursivas, o que o A. demonstra analisando duas obras de índole diversa escritas pelo mesmo autor.

O artigo seguinte — *El peso de la tradición discursiva en un proceso de textualización: un ejemplo en la Edad Media castellana* (pp.197- 224) —, é da autoria de Lola Pons Rodríguez . A A. analisa esquemas de junção representativos de diferentes tradições discursivas que se traduzem em processos de textualização. Utilizando a metodologia do grupo de Tübingen, alia no seu trabalho uma análise filológica rigorosa e uma análise linguística inovadora. Demonstra, neste trabalho, que a inclusão de uma característica de determinada tradição discursiva num outro tipo de texto se pode tornar marca específica da nova forma de textualização.

O penúltimo artigo do volume — *Dinámica de estructuras acanciales en los Siglos de Oro: el ejemplo del verbo encabalar* (pp. 225-248) —, escrito por Wulf Oesterreicher, analisa obras do século XVI, de carácter metalinguístico, e observa a evolução do verbo ‘encabalar’ em que se verifica uma alteração de uso e de estrutura argumental. O estudo parte da análise detalhada dos textos e do contexto cultural que os envolve, única garantia para o estabelecimento dos sucessivos momentos do percurso de mudança. O A. reflecte sobre a relação entre variedades e tradições discursivas e critica certas análises quantitativas de dados que classifica de cegas (esquecendo, a meu ver, algumas virtudes inegáveis dessas análises).

O último artigo deste livro, da autoria de Salvador Pons Boodería, intitula-se *Gramaticalización por tradiciones discursivas: el caso de esto es* (pp. 249-274). A partir da análise de textos jurídicos, em que observa algumas das mais antigas ocorrências de ‘esto es’, o A. mostra como esta estrutura parte do valor textual preexistente, o do latim *id est*, transitando para outros textos dentro do mesmo género discursivo como, por exemplo, os da legislação de Afonso X, e mesmo para outros géneros como a prosa científica e as traduções.

O A. demonstra que em finais do séc. XV a construção analisada já se encontra numa ampla variedade de textos, convergindo na mesma função que 'a saber', até estabilizar no moderno e corrente introdutor de reformulação.

No seu conjunto, os artigos reunidos neste volume perseguem dois objectivos fundamentais: a apresentação de diversos aspectos da concepção de uma nova perspectiva sobre a sintaxe histórica do espanhol a partir das tradições discursivas; e a sugestão da utilidade da aplicação deste conceito à história de outras línguas, românicas e não românicas, generalizando assim esta aproximação teórico-metodológica.

Os resultados aqui apresentados no âmbito do estudo da sintaxe histórica do espanhol inserem-se num projecto de estudos romanísticos da universidade de Tübingen que se propõe analisar a estrutura de dados linguísticos, as tradições discursivas das línguas românicas e realizar uma análise quantitativa e pluridimensional de *corpora* históricos. Trata-se um projecto que alia, entre outros aspectos, o trabalho de ordem filológica e a análise computacional. Disponibiliza ainda informações acerca do projecto TRADISC, programa informático de tratamento semi-automático de dados, que permite diferentes tipos de análises com possibilidade de extrair conclusões acerca dos textos e das tradições em que estão inseridos. Disponibiliza diversas ligações informáticas a esse projecto, sediado na universidade de Tübingen.

É de sublinhar o facto de serem abordados conjuntamente, nesta obra, os aspectos teóricos da questão das tradições discursivas, as questões metodológicas e os aspectos concretos que tentam ou identificar tradições discursivas particulares ou relacionar as evoluções de certos fenómenos linguísticos com certas tradições discursivas.

Para além dos contributos gerais que as implicações teóricas do conceito de tradições discursivas trazem à teoria da linguagem, esta obra interessa de um modo particular aos investigadores que estudam temas relacionados com a mudança linguística, a história das línguas românicas, a história da língua espanhola, a gramática histórica em geral. É sobretudo nestes domínios que as diversas propostas teóricas e metodológicas aqui apresentadas se mostram inovadoras pelo tratamento pluridimensional de *corpora* históricos e pela renovada con-

jugação de uma tradição filológica de investigação dos textos com as mais recentes correntes da ciência linguística.

REFERÊNCIA

Jacob, D.; Kabatek J. (Eds.). 2001. *Lengua medieval y tradiciones discursivas en la Península Ibérica: Descripción gramatical - pragmática histórica – metodología*. Frankfurt am Main/Madrid: Vervuert Iberoamericana.

Guy COOK. *The Discourse of Advertising*. London:
Routledge. 2001. 256 pp. Second Edition.
ISBN 0-415-23455-7 (Paperback)

Alexandra Pinto
alexandrapinto@net.sapo.pt
Faculdade de Letras da Universidade do Porto (Portugal)

Esta segunda edição do livro de Guy Cook sobre o Discurso Publicitário apresenta alterações substanciais relativamente à edição original do livro, datada de 1992, nomeadamente, consagrando vários apartados a formas de comunicação publicitária emergentes, como é o caso da Internet, e actualizando muitos dos anúncios em análise.

Assumindo como princípio que os anúncios constituem um género discursivo autónomo, o trabalho propõe-se construir um *framework* para a descrição e análise de cada anúncio em particular e da publicidade em geral, utilizando, simultaneamente, este género discursivo para apresentar e discutir conceitos da Linguística, da Semiótica e dos Estudos Literários.

O trabalho assume, como princípio metodológico de análise, a heterogeneidade e a multimodalidade dos anúncios, fazendo, assim, uma opção de base por uma corrente de Análise do Discurso em que o texto e o contexto são perspectivados *holisticamente*.

Ao longo do livro, vários *tokens* de publicidade são analisados, desde *outdoors* a anúncios na Internet, anúncios de imprensa, de rádio ou televisão, tomando sempre em consideração a complexa interacção que a linguagem verbal estabelece com os outros processos de comunicação que co-ocorrem com ela: a música, as imagens, os *layouts* gráficos e, também, outros elementos do contexto, como os destinadores e destinatários, ou o próprio meio de circulação da mensagem.

Cada capítulo contém, igualmente, indicações bibliográficas suplementares para o tema específico desenvolvido, bem como exercícios práticos.

Começando, num capítulo introdutório, por tentar definir o anúncio publicitário como um género discursivo, Guy Cook estabelece a prioridade de assumir a multimodalidade dos anúncios como objecto de estudo, recusando o isolamento da linguagem dos restantes processos de comunicação co-presentes. De forma a estabelecer delimitações conceptuais, o autor indica as acepções em que utiliza alguns dos termos mais importantes do seu texto: contexto, substância, paralinguagem, situação, co-texto, participantes, função, discurso entre outros. Focaliza também a atenção no conceito de género discursivo, adoptando a definição de Swales (1990: 58):

a class of communicative events which share some set of communicative purposes. These purposes [...] constitute the rationale for the genre. This rationale shapes the schematic structure of the discourse and influences and constrains choice of content and style (...) In addition to purpose, exemplars of a genre exhibit various patterns of similarity in terms of structure, style, content and intended audience.

(Swales 1990: 58)

A definição do anúncio publicitário como um género discursivo autónomo não pode resultar apenas da sua função, uma vez que, para além da função persuasiva mais comumente aceite para esta classe de objectos discursivos, os anúncios cumprem outras funções, muitas vezes, não intencionadas pelo emissor, tais como informar, desinformar, divertir, alertar. A teoria dos protótipos (Rosch 1977) constitui, para Guy Cook, um bom enquadramento para a definição do discurso publicitário como um género discursivo, permitindo, inclusivamente, lidar com a faceta mais perversiva da publicidade de constantemente cruzar a fronteira do género, criando *tokens* que não respeitam as características prototípicas.

Guy Cook alerta, por exemplo, para que, no caso do anúncio publicitário, o contexto de inserção é tão importante para o reconhecimento do género/protótipo como as características intrínsecas do próprio anúncio.

Ainda no capítulo introdutório, o autor passa em revista as diversas tipologias de anúncios disponíveis, criticando os critérios que lhes servem de base - meio de circulação dos anúncios; produtos/serviços anunciados; técnicas de construção do anúncio; consumidor ou *target* - para concluir que apenas uma tipologia em que vários critérios se cruzem pode descrever adequadamente o universo dos anúncios.

Após a parte introdutória, o livro divide-se em 3 partes: Parte I - Materiais, Parte II - Textos e Parte III - Pessoas, partes que, por sua vez, se subdividem em diversos capítulos.

A Parte I, tal como o título indica, examina a componente material dos anúncios, abordando questões como as implicações que a multimodalidade produz nas mensagens publicitárias e a influência do contexto nas mesmas. São passadas em revista várias combinações possíveis de processos de comunicação, como a música, as imagens, o discurso escrito e o discurso falado e comentados casos práticos. A paralinguagem é também objecto de análise através do comentário de casos de exploração intencional da grafologia e da fonologia. Neste apartado, o autor compara a exploração da paralinguagem na Literatura e na Publicidade, referindo, especificamente, o caso da Poesia Concreta como um exemplo da iconicidade das palavras na Literatura.

A Parte II concentra-se na componente textual dos anúncios. Começando por recuperar alguns conceitos de Semântica e Pragmática, o autor analisa o sentido de palavras e slogans publicitários no seu contexto de enunciação, remetendo para diversas técnicas de ordem morfológica, gráfica, fonética e semântica, que potenciam os jogos de sentido nos anúncios. Assim, o autor assinala que a exploração da conotação é central nas mensagens publicitárias, apresentando análises de vários casos de duas categorias de anúncios de produto: Perfumes e Carros.

Um outro aspecto analisado na Parte II do livro é a importância da prosódia no discurso publicitário. Guy Cook salienta mesmo que esta é uma das características opostas, embora não privativa, deste género discursivo face a outros géneros em que a prosódia não é proeminente. Comenta vários anúncios em que o fenómeno da prosódia é explorado, sugerindo, aliás, uma pequena tipologia para

estas manifestações: “ad-poems, borrowed poems, jingles, borrowed songs and prosodic ads.” (Cook 2001: 126).

Outras formas de paralelismo são igualmente importantes no discurso publicitário: para além do paralelismo gráfico e fonológico, o paralelismo lexical, o estrutural, o semântico e o discursivo. Frequentemente todas estas formas de paralelismo se cruzam num mesmo texto publicitário.

Os recursos da coesão discursiva nos anúncios publicitários são um dos aspectos tratados nesta Parte dedicada ao Texto. Num discurso onde a organização retórica assume uma centralidade notável, todas as manifestações de paralelismo formal na publicidade são um mecanismo de coesão discursiva importante. Outros processos como a elipse sintáctica, várias formas de anáfora e catáfora, o valor dos conectores ou ainda os pronomes mais usados, são analisados em contextos específicos.

Na Parte III do trabalho, Guy Cook concentra-se nas “vozes” em confronto no discurso publicitário, nomeadamente nas do emissor e do receptor do discurso, avaliando o tipo de relações entre participantes promovidas por este tipo de discurso. Entre os mecanismos linguísticos revistos pelo autor para analisar este sub-tema, contam-se a elipse sintáctica, a distribuição da informação nova e da informação dada nos enunciados e as pressuposições e implícitos das mensagens. O autor utiliza a teoria de Bakhtin para analisar o confronto de vozes nos anúncios (Volosinov, 1986), aplicando-a, nomeadamente, à análise da intertextualidade neste tipo de discurso – intertextualidade intra-género e intertextualidade inter-géneros (Cook 2001: 193, 194).

Para terminar, Guy Cook passa ainda em revista a forma como diferentes tipos de públicos e discursos – artistas, escritores, académicos, opinião pública - perspectivam a publicidade e as suas funções na sociedade.

Guy Cook termina o seu livro com a compilação das características que, ao longo do livro, permitiram desenhar o discurso publicitário como um género discursivo autónomo, chamando a atenção para o carácter versátil, instável e mutante deste género que faz a sua prototipicidade evoluir constantemente.

REFERÊNCIAS

- Cook, G. 1992. *The Discourse of Advertising*. London: Routledge.
- Cook, G. 2001. *The Discourse of Advertising*. Second Edition. London: Routledge.
- Rosch, E. 1977. Human categorization. In: N. Warren (Ed.). *Advances in Cross-Cultural Psychology*. New York: Academic Press, I, 1-49.
- Swales, J. 1990. *Genre Analysis*. Cambridge: Cambridge University Press.
- Volosinov, V. N. [nome usado por M. M. Bakhtin]. 1986. *Marxism and the Philosophy of Language*. Cambridge MA: Harvard University Press [versão russa original: 1929].

