

PROPOSTA DE NOVOS CONCEITOS E UMA NOVA NOTAÇÃO NA FORMULAÇÃO DE PROPOSIÇÕES E DISCUSSÕES ETIMOLÓGICAS

Mário Eduardo VIARO*
Aldo Luiz BIZZOCCHI**

- **RESUMO:** O presente artigo visa a apresentar a metodologia de trabalho e os principais postulados teóricos que norteiam a elaboração do DELPo (Dicionário Etimológico da Língua Portuguesa), a cargo do NEHiLP-USP (Núcleo de Apoio à Pesquisa em Etimologia e História da Língua Portuguesa da Universidade de São Paulo), visto tratar-se de projeto inovador tanto em seus aspectos teóricos quanto empírico-operacionais. Pretende-se aqui apresentar as inovações conceituais, notacionais e terminológicas propostas pelos autores, que embasam a elaboração do *DELPo*. Para tanto, baseamo-nos em referenciais teóricos sobre neologia (ALVES, 2007; BARBOSA, 1993, 1996), lexicogênese (BIZZOCCHI, 1998) e etimologia (VIARO, 2011), bem como propomos uma atualização na simbologia utilizada na formulação de proposições etimológicas, que ao mesmo tempo dirima ambiguidades e inconsistências da notação tradicional e dê conta das inovações conceituais aqui introduzidas. Pretende-se que tanto os processos etimológicos descritos quanto os símbolos a eles correspondentes se tornem padrão na pesquisa em etimologia e na lexicografia etimológica.
- **PALAVRAS-CHAVE:** Etimologia. Lexicografia. Dicionários etimológicos. Proposições etimológicas. Discussão etimológica. Simbologia etimológica. Notação linguística.

Introdução: o NEHiLP

O Núcleo de Apoio à Pesquisa em Etimologia e História da Língua Portuguesa (NEHiLP), vinculado à Pró-Reitoria de Pesquisa da Universidade de São Paulo, visa à produção e divulgação de pesquisas científicas sobre Linguística Histórica, Filologia e Etimologia. Para tal, congrega especialistas de várias áreas que se dedicam diretamente a tais estudos. Muitos pesquisadores desse núcleo interdisciplinar são especialistas em

* USP – Universidade de São Paulo - Departamento de Letras Clássicas e Vernáculas. São Paulo – SP – Brasil. 05508-900 – maeviaro@usp.br

** Membro do Grupo de Pesquisa do Núcleo de Apoio à Pesquisa em Etimologia e História da Língua Portuguesa - NEHiLP. USP – Universidade de São Paulo - Departamento de Letras Clássicas e Vernáculas. São Paulo – SP – Brasil. 05508-900 – aldo@aldobizzocchi.com.br

história antiga e moderna, estruturas linguísticas e teorias de reconstrução linguística. O método empregado é o da pesquisa em documentos antigos e atuais, com vistas à organização da informação linguística que permita gerar dados de qualidade para a consulta tanto de especialistas em linguística como de outros setores da sociedade que se interessem por etimologia (sobretudo jornalistas e cientistas das mais variadas áreas). Além do fichamento das primeiras datações de vocábulos e de suas acepções, o NEHiLP disponibiliza informação histórica classificada segundo suas características sociolinguísticas e estilísticas, associadas a informação sobre frequência de uso.

Dentre os seus projetos estão o projeto *Retrodatação* e o *DELPO*.

O projeto *Retrodatação*

Uma das tarefas do etimólogo é coletar contextos em determinadas obras e associá-los à data de publicação da mesma obra, que, necessariamente, deve ser uma edição confiável. Por meio dessas datações, é possível estabelecer o *terminus a quo* da palavra e estabelecer etimologias. Tradicionalmente, o etimólogo é entendido como uma pessoa extremamente erudita, que sabe onde localizar as melhores fontes e não raro conhece detalhes de conteúdo das obras que consulta/cita. Obviamente, ter erudição e ser organizado são qualidades sempre desejáveis para o pesquisador desse trabalho, mas não são o ponto principal. O ideal é que o menor número de erros seja cometido. Erros que dependam apenas das propensões humanas sempre existirão, mas há alguns tipos de erros que podem ser evitados, como os motivados por cansaço e pela falta de atenção, dado o manuseio de um grande número de informações. Nesse ponto, acredita-se que um programa que faça a comparação automática de um determinado *terminus a quo* com a data da obra analisada auxiliaria muito as pesquisas.

As vantagens desse programa sobre a pesquisa manual seriam várias:

(a) numa pesquisa manual, o pesquisador não tem disponíveis em sua erudição todos os *termini a quo* de todas as palavras de uma língua, de modo que suas descobertas se pautam, na maioria das vezes, apenas pela “sensação” de que a palavra “não deveria estar naquele texto”, por ele supor (baseado exclusivamente na sua experiência de falante e/ou pesquisador) que a palavra seja mais recente do que a data do texto investigado;

(b) para sanar a impossibilidade de uma investigação exaustiva, o pesquisador acaba por se especializar, seja na forma (palavras com um determinado sufixo, por exemplo), seja no significado (palavras de um determinado campo semântico), seja numa suposta origem (por exemplo, palavras de origem africana). Desse modo, o texto precisaria ser revisto por uma quantidade indefinida de pesquisadores para dele serem extraídas todas as informações que possam ser interessantes para os estudos etimológicos.

Um programa que tivesse idealmente a lista de todas as palavras da língua portuguesa com seus respectivos *termini a quo* poderia teoricamente apresentar ao pesquisador todas as ocorrências em que a data do texto processado for menor que a do *terminus a quo*, sem que ele recorresse à sua erudição, e sanaria os problemas do

item (a) acima, ao mesmo tempo que faria uma varredura integral, anulando também a parcialidade indesejável de (b).

Para isso foi idealizado pelo NEHiLP o programa computacional *Moedor*, cujo projeto inicial foi encaminhado à Pró-Reitoria de Pesquisa da Universidade de São Paulo em 2012, juntamente com a criação do Núcleo, e que foi desenvolvido pelo NEHiLP em colaboração com o Instituto de Matemática e Estatística da USP entre 2013 e 2014.

O projeto *Dicionário Etimológico da Língua Portuguesa (DELPO)*

O projeto *Retrodatação* visa a algo mais amplo que a coleta dos dados. Diferentemente do que ocorre com a maior parte das línguas europeias (inglês, francês, espanhol e italiano), a informação etimológica presente nos dicionários da língua portuguesa é bastante falha no tocante aos seus dados etimológicos. Neles:

- confundem-se derivação sufixal/prefixal e etimologia;
- confundem-se étimo da palavra e sua origem remota;
- não há cuidado suficiente com étimos de línguas ágrafas;
- desconhece-se muito da influência árabe;
- há total arbitrariedade no que toca a étimos de origem indígena e africana;
- há abundância de étimos fantasiosos que descaracterizam o estudo etimológico como um trabalho científico.

Acresça-se ainda que:

- há poucos dicionários etimológicos que contam com datas fiáveis para seus *termini a quo*;
- não há até hoje uma metodologia para o trabalho do *terminus ad quem*;
- a língua portuguesa está longe de ter hipóteses etimológicas e *termini a quo* para aceções, pois o que existe são majoritariamente datações de lemas.

Todos os pesquisadores que se dedicam a aspectos históricos e à diacronia da língua portuguesa sentem falta de material especializado à altura do dicionário etimológico *Oxford* para a língua inglesa, do *Le Robert* para o francês, do de Cortellazzo & Zolli para o italiano ou do de Corominas para o espanhol. As obras mais completas que temos em língua portuguesa são as várias publicações de Antônio Geraldo da Cunha, o dicionário de Houaiss & Villar e o dicionário de José Pedro Machado. São ausentes ou muito falhos os estudos dos séculos XVII, XVIII e XX.

Assim sendo, as propostas etimológicas encontram-se pouco desenvolvidas em língua portuguesa, pois as mais elaboradas e integradas numa visão românica datam

apenas do final do século XIX até meados da década de 20 do século XX. Desde então abundam etimologias *ad hoc* e étimos desconhecidos.

Como produto final da pesquisa do NEHiLP, desenvolver-se-á o *DELPO* (*Dicionário Etimológico da Língua Portuguesa*), inicialmente *on-line*, mas com possíveis versões impressas. Além da desejável versão integral do *DELPO*, imagine-se que também será possível fazer publicações parciais, por exemplo, por áreas específicas de conhecimento.

Para a elaboração do *DELPO* há grupos especializados de pesquisadores em:

- (a) palavras de base latina ou associadas a substratos antigos (antigos helenismos, palavras celtas, ibéricas, etc.) e de superstrato germânico;
- (b) palavras de origem ameríndia;
- (c) palavras de origem africana;
- (d) palavras de origem árabe;
- (e) palavras de origem indiana, chinesa e japonesa;
- (f) palavras de origem desconhecida;
- (g) palavras internacionais europeias surgidas a partir do século XVII (de várias origens, sobretudo italiano, espanhol, francês e inglês);

Inovações conceituais e notacionais introduzidas pelo *DELPO*

Além do aspecto inovador da metodologia utilizada para a constituição do *corpus* sobre o qual o dicionário será elaborado, também foram introduzidas inovações conceituais no que respeita à descrição dos fenômenos evolutivos que afetam as palavras ao longo de sua história, com repercussões na formulação das chamadas *proposições etimológicas*, o que ensejou o desenvolvimento paralelo de uma nova notação descritiva desses fenômenos.

Muitos dos conceitos aqui apresentados são de uso corrente na Etimologia; outros já haviam sido propostos anteriormente (BIZZOCCHI, 1998, 2013; VIARO, 2011), ao passo que alguns, ainda, estão sendo apresentados agora pela primeira vez. À época, esses estudos aliaram a pesquisa etimológica a técnicas quantitativas de lexicometria (análise estatística informatizada de léxico), tendo consistido numa etapa de revisão bibliográfica e reelaboração das teorias e conhecimentos etimológicos disponíveis e numa etapa de aplicação do novo modelo teórico assim desenvolvido a um *corpus* de textos acadêmicos e jornalísticos em seis línguas europeias, a saber, português, espanhol, francês, italiano, inglês e alemão.

O motor dessa pesquisa fora a constatação de que, quando se tomam traduções de um mesmo texto para várias línguas europeias (por exemplo, um manual de instruções

redigido em vários idiomas), nota-se que, onde o francês e o inglês empregam a palavra latina *instruction* (do latim *instructionem*), em que só a desinência foi adaptada, o português substitui o sufixo *-tionem* pelo vernáculo *-ção* (*instrução*) e suprime o *c* do radical, enquanto o italiano elimina o *n* e o *c* e adapta a ortografia (*istruzione*), e o alemão cria um decalque a partir do latim (*Anweisung*). Ou seja, diante da necessidade de importar uma palavra latina, cada língua a incorpora segundo um processo diferente.

Especificamente com respeito ao português, constatou-se que os dicionários etimológicos disponíveis não fazem distinção clara entre os itens lexicais herdados do latim em época pré-histórica (léxico hereditário), os criados a partir de material vernáculo em época histórica (criações intralinguísticas) e os resultantes de empréstimo, limitando-se as mais das vezes a apontar o étimo e a apresentar abonações. Vários deles classificam vocábulos cultos, sobretudo termos técnicos (*termômetro*, *helicóptero*, etc.), como resultantes de composição entre elementos gregos e/ou latinos como se tal processo tivesse ocorrido simultaneamente em português e em outras línguas, isto é, como se tais termos não tivessem sido criados num determinado idioma (quase nunca o português) e só então exportados aos demais.

Além disso, muitos trabalhos de Etimologia, Filologia e Linguística Histórica dividem o léxico em palavras cultas e vulgares, desconsiderando a existência de palavras semicultas ou híbridas, bem como as de classificação problemática; outras obras reconhecem os semicultismos, mas classificam como cultos vocábulos que, a rigor, deveriam ser tidos como semicultos; outras, ainda, tratam decalques de termos estrangeiros como palavras vernáculas, ignorando, ainda, que o decalque pode ser total ou parcial, sintagmático ou semântico. Finalmente, alguns autores, como Bechara (1998), confundem vulgar e culto com hereditário e empréstimo, como se todas as palavras vulgares tivessem sido herdadas (desconsiderando-se, pois, as composições e derivações modernas dessas palavras) e todos os empréstimos fossem cultos, isto é, provenientes do grego ou latim sem nenhum metaplasmo (desconsiderando, agora, empréstimos como o do inglês *futebol*, que nada tem de culto, ou semicultismos como *artigo* e *cabido*).

Outro aspecto totalmente ignorado pela literatura em questão é que, ao criar e renovar seu léxico, as línguas fazem escolhas entre as várias possibilidades que lhes são dadas – criação intralinguística, empréstimo, decalque, etc. – e que essas escolhas, por serem feitas de forma sistemática, tornando-se mesmo injunções em certos casos, revelam algo da visão de mundo dos falantes em relação à própria língua. Além disso, como tais escolhas mudam com o tempo, pode-se reconstituir a evolução histórica dessa visão de mundo pela análise sincrônica do espectro etimológico do léxico, seja no momento presente ou em sincronias passadas.

O projeto *DELPO* também incorpora e amplia as inovações notacionais (VIARO, 2011) em grande parte para dar conta do próprio enriquecimento conceitual acarretado pelo projeto, bem como para eliminar ou ao menos minimizar inconsistências que a notação etimológica tradicional comporta.

Como em todo trabalho lexicográfico, o *DELPO* será constituído a partir do preenchimento de fichas, totalmente informatizadas e *on-line*, das quais as proposições etimológicas e a discussão etimológica são campos a ser preenchidos pelos pesquisadores. As notações e conceituações aqui apresentadas são definidas como norma a ser seguida por esses pesquisadores e constam em manual próprio, disponível no site do Núcleo (www.nehilp.org).

A formulação de proposições etimológicas segundo a metodologia adotada no *DELPO*

Como primeira inovação notacional, o *DELPO* faz distinção entre os símbolos * (dado reconstruído) e ★ (dado inexistente), antepostos à forma linguística em análise. Este último símbolo é usado, conforme Viaro (2011), para indicar formas inexistentes ou impossíveis em vez do asterisco chomskyano. Já o símbolo * é estritamente reservado na sua interpretação schleicheriana, mais antiga, como “forma reconstruída” (portanto supostamente existente, ao menos em teoria). A razão disso é que o uso contraditório do asterisco cria aporias no estudo diacrônico.

Uma etimologia pode ser expressa de maneira sucinta por uma ou mais *proposições etimológicas*, que são afirmações acerca das transformações ocorridas sobre dados de uma mesma sincronia ou de duas sincronias contíguas.

Nessas transformações, temos o *dado original* (x) cronologicamente anterior ao *dado derivado* (x') separado por um símbolo entre eles.

Ambos podem ser *dados atestados* (x), *dados inexistentes* (★x) ou *dados reconstruídos* (*x). Os dados atestados encontram-se em obras, os dados inexistentes e os dados reconstruídos não se encontram em obra alguma; no entanto, os dados reconstruídos obedecem a uma sequência previsível de transformações diacrônicas, ao passo que os dados inexistentes são meros exemplos conjecturais da aplicação de regularidades analógicas, que não têm nenhuma frequência de uso em uma determinada sincronia e em um determinado sistema linguístico.

Também podemos falar de *dados desusados* (†x), quando, em uma determinada sincronia, a frequência de uso for baixa ou nula, comparada a sincronias anteriores. O conceito de desuso está ligado ao estudo do *terminus ad quem*.

Os *dados*, portanto, pertencem a *sistemas* linguísticos, que, por sua vez, estão vinculados a *línguas* (definidas também politicamente). Antes dos dados, portanto, deve aparecer o nome da língua da qual é proveniente.

São oito os símbolos utilizados na descrição das sincronias:

$x \rightarrow y$ ou $y \leftarrow x$	o dado linguístico y é derivado morfológico de x
$x \Rightarrow y$ ou $y \Leftarrow x$	o dado linguístico y é afetado por analogia com x
$x \blacktriangleright y$ ou $y \blacktriangleleft x$	o empréstimo x transforma-se em y
$x \equiv y$	x e y são homófonos
$x \cong y$	x é cognato de y
$x \approx y$	x é uma variante de y
$x \sim y$	x e y são flexões do mesmo paradigma
$x + y \rightarrow z$	z é uma composição de x e de y

São quatro os símbolos vinculados à descrição diacrônica:

$x > y$ ou $y < x$	x é étimo/origem de y
$x \gg y$ ou $y \ll x$	o significado x se tornou y
$x \geq y$ ou $y \leq x$	y é um decalque de x
$x \Rightarrow y$ ou $y \Leftarrow x$	x foi substituído por y

Os elementos x ou y sempre terão o formato “língua *dado*”, exceto se a língua for o português (nesse caso teríamos apenas “*dado*”), exceção feita se não se trata de uma palavra usada em toda a língua portuguesa; nesse caso é preciso especificar: “português europeu”, “português brasileiro”, “português angolano”, “português cearense”. Não se usa porém a expressão “português arcaico” (nem “português medieval”, “português renascentista”, etc.) uma vez que cada sincronia deve ser suficientemente clara por meio de uma numeração superscrita anteposta ao dado.

Uma proposição etimológica tem *autoria* (que é apresentada por meio de uma indicação bibliográfica) e também *graus de certeza*.¹ Define-se o grau de certeza por meio de um número. Para tal é preciso confirmar a aplicação regular das leis fonéticas em cada sincronia. Para checar a regularidade da incidência de leis fonéticas, foi desenvolvido pelo NEHiLP o programa *Metaplasgador*, disponível para uso público em seu site (aba “Programas/Metaplasgador”). Os graus de certeza são:

- **[1] Certo:**
quando houver aplicação regular das leis fonéticas entre dados não reconstruídos, além de manutenção de significado.
- **[2] Provável:**
quando houver aplicação regular de leis fonéticas entre dados reconstruídos, além de manutenção do significado;
quando houver irregularidade nas leis fonéticas em apenas um *locus* da etimologia em dados não reconstruídos, além de manutenção do significado.

¹ Segundo Jespersen (1954, p. 307, nota 1): *It is of course, impossible to say how great a proportion of the etymologies given in dictionaries should strictly be classed under each of the following heads: (1) certain, (2) probable, (3) possible, (4) improbable, (5) impossible – but I am afraid the first two classes would be the least numerous.*

- **[3] Possível:**
quando houver aplicação regular das leis fonéticas entre dados não reconstruídos, mas não houver manutenção do significado;
quando houver aplicação regular de leis fonéticas entre dados reconstruídos, mas não houver manutenção do significado;
quando houver irregularidade nas leis fonéticas em apenas um *locus* da etimologia em dados reconstruídos, mas manutenção do significado.
- **[4] Improvável:**
quando houver irregularidade nas leis fonéticas em apenas um *locus* da etimologia em dados reconstruídos ou não, além de não-manutenção do significado;
quando houver irregularidades nas leis fonéticas em mais de um *locus* da etimologia em dados reconstruídos ou não, com manutenção do significado.
- **[5] Impossível:**
quando houver irregularidade nas leis fonéticas em mais de um *locus* da etimologia em dados reconstruídos ou não e, além disso, não houver manutenção do significado.

Resumidamente (sendo i = número de *loci* irregulares):

Tabela 1 – Atribuição de graus de certeza a partir de índices da etimologia proposta.

Regularidade das leis fonéticas ($i \leq 1$)	<i>sim</i>		<i>sim</i>		<i>sim</i>		<i>sim</i>		<i>não</i>		<i>não</i>		<i>não</i>		<i>não</i>	
Reconstrução	<i>sim</i>		<i>sim</i>		<i>não</i>		<i>não</i>		<i>sim</i>		<i>sim</i>		<i>não</i>		<i>não</i>	
Manutenção do significado	<i>sim</i>		<i>não</i>													
Número de <i>loci</i> irregulares	0	1	0	1	0	1	0	1	> 1	> 1	> 1	> 1	> 1	> 1	> 1	> 1
Grau de certeza	[2]	[3]	[3]	[4]	[1]	[2]	[3]	[4]	[4]	[5]	[4]	[5]	[4]	[5]	[4]	[5]

Fonte: Elaboração própria.

Esse protocolo é válido para palavras herdadas. Nos casos de analogia, empréstimos, substituições e decalques, outros critérios devem ser adotados. Os graus de certeza poderiam ser acrescidos de + quando houvesse uma atuação analógica. Por exemplo, se uma proposição etimológica de grau 4 não segue as leis fonéticas porque necessita de uma explicação de cunho analógico, sua notação seria [4+]. Os empréstimos evidentemente sofrem adaptação ao sistema fonológico da língua-alvo na sincronia em que ocorrem; no entanto, nem sempre esse sistema é conhecido, tendo em vista a falta de dados suficientes para a sua reconstrução, bem como a natureza eventual do empréstimo. Questões como a riqueza da documentação influenciam necessariamente no julgamento. O mesmo se pode dizer das substituições e dos decalques.

A discussão etimológica

A discussão etimológica se detém na argumentação de cada proposição etimológica, na variação (diatópica, diafásica, diastrática, diacrônica) da forma e do significado da palavra e nos cognatos interlinguísticos (e eventualmente nas suas disparidades formais ou semânticas). São considerados pertinentes os comentários fundamentados sobre a atualidade da palavra ou sobre sua raridade, seu grau de especialização, sobre seu uso ou desuso.

Também são pertinentes questões sintáticas que envolvam aspectos morfossintáticos de concordância (como gênero, número ou regências específicas), assim como da participação da palavra em lexias.

Nesse campo do verbete, plenamente desenvolvido numa acepção principal, discorre-se sobre as diferentes acepções que lhe são subordinadas.

Por exemplo, suponhamos que *estrela*₁, “corpo celeste”, seja vocábulo herdado, mas *estrela*₂, “atriz famosa”, seja decalque do inglês *star* e não uma derivação direta.

A solução para isso está na acepção principal *estrela*₁, dentro de cujo campo *Etimologia* estará toda a história da palavra *estrela*.

Numerados estarão, dentro desse campo, primeiro, a proposição “latim *stellam* > *estrela* ⇔ latim *astrum*” e toda a discussão etimológica e seus cognatos e, em seguida, a proposição “inglês *star* ≥ *estrela*”, com seus próprios comentários e cognatos. Assim sendo, o comentário de “inglês *star* ≥ *estrela*” não estará na ficha “*estrela*₂”, onde haverá apenas uma remissão à acepção principal “*estrela*₁”. Sendo assim o campo *Etimologia* corretamente preenchido, essa discussão reaparecerá no lema.

Nas discussões tradicionais de Etimologia, classificam-se os vocábulos em cultos ou eruditos (*cultismos*) e vulgares ou populares (*vulgarismos*). Essas categorias dizem respeito à origem do vocábulo e não ao seu uso: há palavras herdadas, como *escorreito*, que só ocorrem no registro ultraformal, assim como há cultismos, como *operário*, que são de amplo uso em todos os níveis de linguagem, do ultraformal ao informal e popular. É importante distinguir, portanto, ao estudar uma acepção, a origem culta ou popular de seu uso.

No entanto, há gradações importantes entre esses dois extremos. Podemos dizer que há *semicultismos* quando há hibridismo de elementos cultos e vulgares, seja por combinação ou por vulgarização da forma greco-latina. Também há *semivulgarismos*, ou seja, vocábulos existentes na língua desde seus primórdios, surgidos no período chamado de “latim cristão”, quase sempre ligados à Igreja, disseminados na fala popular, mas de baixa frequência de uso, assim como restritos a certos gêneros discursivos em que a fala era mais monitorada. Somente uma investigação detalhada das sincronias pretéritas (feitas com pesquisa e com a ajuda de ferramentas como o *Metaplasgador*) poderá dizer exatamente em que momentos o latim clássico atuou com maior ou menor força para exercer um papel na etimologia de uma palavra.

Metaplasmos irregulares, como no caso de *escola*, *cabido*, *cônego* e *missa* apontam para os chamados *semivulgarismos*.

Por fim, segundo essa distinção entre palavras cultas e populares, tornam-se inclassificáveis os neologismos fonológicos (*poperô, tilim-tilim*),² as derivações de siglas (*ufologia, petista, aidético*) ou de nomes próprios (*amperímetro, kantiano*), bem como as recomposições (*reprografia, informática, metrô, minissaia, showmício*).

Considerando-se ainda que uma palavra pode ter origem numa língua sem qualquer contribuição, direta ou indireta, de outra, ou pode ter-se formado a partir de elementos mórficos externos ao sistema ou mesmo com elementos autóctones mas segundo um modelo estrutural estrangeiro, reconhecem-se dois processos básicos de formação lexical:

- os autogenéticos (hereditariedade ou criação intralinguística);
- os alogenéticos (empréstimo ou criação a partir de elementos alógenos).

Pode-se dizer que são autogenéticas as palavras que não contenham nenhum elemento proveniente de outro sistema linguístico, quer no plano da expressão quer no do conteúdo, e que são alogenéticas as que contenham pelo menos um desses elementos, seja ele de natureza morfológica, seja semântica.

Assim sendo, fará parte da seção “Discussão etimológica” do campo *Etimologia* a seguinte terminologia para as acepções (BIZZOCCHI, 1998, 2009, 2013):

1. Processos autogenéticos (*autogenia*):
 - a. *herança* (vocábulo vernáculos ou hereditários);
 - b. *neologia fonológica* (neologismo fonológico);
 - c. *ressemantização* (de palavras autogenéticas);
 - d. *composição* ou *derivação* (a partir de palavras autogenéticas).
2. Processos alogenéticos (*alogenia*):
 - a. *empréstimo* de palavra estrangeira (“estrangeirismo”);
 - b. *ressemantização* (de palavra autogenética com significado importado);
 - c. *ressemantização* (de palavra alogenética com significado vernáculo);
 - d. *composição* ou *derivação* (a partir de palavras alogenéticas).

Os cultismos e semicultismos enquadram-se na categoria alogenética, já que resultam de elementos importados do grego ou latim. No exemplo citado mais acima, *estrela*₁ é vocábulo autogenético, ao passo que *estrela*₂ é alogenético.

É preciso considerar ainda que a fronteira entre as duas famílias de processos é fluida, visto que palavras vernáculos podem receber significados estrangeiros, bem

² Muitas dessas formações são consideradas onomatopáicas (imitação do som do significado: *plim-plim, miau, zunzum, ziguezague, tique-taque, tilim*, e seus derivados *miar, zunzunar, ziguezaguear, tilintar*). Discute-se muito se essas palavras são de fato neologismos *ex nihilo* (criação de um novo significante a partir do nada: *poperô, chinfrim, pirlimpimpim, zureta*), mas há muito que estudar ainda sobre a expressividade fonológica (VIARO; FERREIRA; GUIMARÃES-FILHO, 2013, p. 58-105).

como ser restauradas, isto é, ter seu significante herdado substituído pelo significante latino que lhe deu origem. Por exemplo:

mosteiro ⇒ *monastério* ◀ latim *monastērĭum* ◀ grego μοναστήριον.

Nesse aspecto, é discutível o estatuto dos *epônimos*, nomes próprios que se tornam comuns. Muitos deles homenageiam o criador do objeto que nomeiam. Por exemplo, *ampère*, *macadame*, *gilete* e *zeppelin* são epônimos dos sobrenomes *Ampère*, *Gillette*, *MacAdam* e *Zeppelin*, respectivamente. No caso de *gilete*, o substantivo comum não derivou diretamente do nome do inventor da lâmina de barbear e sim da marca do produto, originalmente fabricado pelo próprio inventor. Outro caso de marca registrada que se tornou nome comum é *maisena* (< *Maizena*®). Dentre os compostos e derivados desse tipo temos: português *abreugrafia*, francês *voltmètre*, inglês *Newtonian*, alemão *Kantismus*, e, de modo geral, todos os derivados de nomes de pessoas e países, como *hitlerista* e *zimbabuês*.

Siglas e acrônimos são às vezes passíveis de lexicalização e chegam a ter uma pronúncia silábica, enquanto outras permanecem pronunciadas de modo soletrado (BARBOSA, 1993, 1996). De qualquer maneira, elementos formados por *acronímia* como *USP*, *AIDS*, *PT*, e seus derivados *uspiano*, *aidético*, *anti-AIDS*, *petista* também formam grupos especiais a ser estudados.

Detalham-se a seguir os processos etimológicos que envolvem casos acima descritos.

Herança

São classificados como *vernáculos* ou *hereditários* os vocábulos herdados diretamente do sistema linguístico numa sincronia imediatamente anterior, sem a intervenção de outro sistema qualquer. No caso de línguas românicas como o português, é vernáculo tudo quanto já existia no latim vulgar e continuou a existir, de modo que, quando adquiriram o *status* de línguas distintas do latim, esse acervo de material lexical e gramatical passou a constituir o léxico de base dessas línguas. A representação de uma herança é $x > y$.³ As formas nominais latinas, nesse caso, virão no *acusativo* (sem apócope do *-m*).

É preciso ressaltar que, devido ao intercâmbio de palavras entre o latim e o germânico durante os últimos séculos do Império Romano e início da Idade Média (séculos I a VII d.C.), há também palavras vernáculas nas línguas românicas que são de origem germânica (por exemplo, português *guardar* < latim vulgar **guardare* ◀ frâncico *wardan*), assim como palavras vernáculas nas línguas germânicas que são de origem

³ O símbolo $>$ restringe-se a mudanças do significante; assim sendo, será usado pelo NEHiLP para mudanças não só de caráter fonético ou fonológico, mas também gráfico.

românica (por exemplo, inglês *dish* < germânico comum **diskaz* ◀ latim *discus*). Visto que esses intercâmbios se deram entre o latim e uma língua germânica, essas palavras, dependendo da sincronia, ou são empréstimos ou são herança.

Diferentemente da notação tradicional, é possível haver notações como $x > y < z$, que significam que o elemento y é uma forma com dupla origem. Por exemplo:

latim *pro* > *por* < *por* ~ †*per* < latim *per*.

Empréstimo

O termo *empréstimo* é aplicado a tudo aquilo que não é criado em sincronia nem é herdado, ou seja, uma palavra proveniente de um outro sistema numa determinada sincronia. O termo *estrangueirismo* costuma ser usado como sinônimo de *empréstimo*, mas é preciso fazer algumas distinções, pois há os considerados “estrangueirismos nacionalizados” e os “não nacionalizados”. O termo *estrangueirismo* não é, portanto, um termo técnico adotado pelo NEHiLP na seção de “Discussão etimológica” do campo *Etimologia*. Faz-se tradicionalmente exceção às palavras vindas do grego ou do latim clássicos. Salvos uns poucos casos (*status*, *habitat*), palavras do grego e do latim já ingressam plenamente “nacionalizadas”, ao passo que alguns estrangueirismos ou estão nacionalizados na grafia e pronúncia (*futebol*, *abacaxi*) ou não (*pizza*, *bonbonnière*). Diferentemente do uso tradicional, não se fará distinção entre esses dois grupos de palavras (isto é, as vindas do grego ou do latim clássico e as vindas de outras línguas) e em ambos os casos se fará a sua representação como $x \blacktriangleright y$ (as formas nominais latinas, nesse caso, virão no *nominativo*). Isso é válido tanto para os cultismos diretos quanto para os indiretos (quando uma palavra latina ou grega, por exemplo, entra no português não diretamente do latim ou do grego, mas por meio de uma língua moderna intermediária, como o francês, o italiano, o espanhol ou o inglês).

Exemplos de propostas etimológicas:

latim *temperātūra* ▶ francês *température* ▶ *temperatura*;
grego θέρμανση ▶ latim *theatrum* ▶ francês *théâtre* ▶ *teatro*.

Algumas vezes, uma palavra de um sistema A é admitida pelo sistema B e, tempos depois, retorna para A. Não raro esse retorno ocorre quando a palavra já havia caído em desuso em A ou se havia modificado formal e/ou semanticamente. Outras vezes, a forma recém-chegada convive como alótopa da outra, mais antiga. Esse caso específico é chamado de *retroversão*. Por exemplo, o vocábulo português *fetice* é empréstimo do francês *fétiche*, por sua vez importado do português *feitiço* e ressemantizado. Simbolicamente:

fetice ◀ francês *fétiche* (“feitiço” >> “fetice”) ◀ *feitiço*.

Nesse caso, representam-se as formas alótropas como cognatas: *feitiço* ≅ *fetichê*. Também pode ocorrer retroversão de formas cultas, como se vê no francês †*parformer* ► inglês *perform* ► francês *performer*. O lat. *humor*, “líquido”, passou nessa acepção ao fr. medieval *humour* (atual *humeur*) por tradução do sufixo. Este passou por empréstimo ao inglês *humour*, que mais tarde adquiriu a acepção de “senso de humor, graça, humorismo”. Esta acepção retornou ao francês como *humour*, que hoje convive com *humeur* (formas divergentes). Igualmente, o francês *entrevue*, “entrevista”, passou ao inglês *interview* por restituição⁴ do prefixo *e*, posteriormente, voltou ao francês como *interview*. Hoje, ambas as formas convivem em francês com significados diferentes: *entrevue* é uma entrevista para tratar de negócios (uma oferta de emprego, por exemplo) e *interview* é a entrevista jornalística.

Alguns cultismos, por serem indiretos, entram na língua com a forma fonética do idioma a partir do qual foram importados. Isso produz uma divergência entre a forma assumida por esse cultismo e a que seria esperada se ele proviesse diretamente do grego ou latim.

Por exemplo:

grego φρένησις ► latim *phrenēsis* ► francês *frénésie* ► *frenesi*;
grego Σειρήν ► latim *Sīrēn* ► francês *sirène* ► *sirene*;
latim *domīnō* ► francês *domino* ► *dominó*.

Quando uma língua importa um semicultismo, este normalmente mantém seu caráter semiculto. Por exemplo, o francês *nécessaire* provém do latim *necessarius* (com adaptação da terminação), o qual passou ao português *nécessaire* na acepção de “bolsa ou estojo feminino para utensílios de toalete”. Da mesma forma, o latim *sociētās* gerou o francês *société* e, a seguir, esta forma passou ao inglês *society*.

Normalmente, os cultismos entram nas línguas vulgares por via escrita e só depois chegam à fala. Por isso, as línguas importadoras costumam conservar a grafia greco-latina (salvo adaptações ortográficas obrigatórias), ainda que em detrimento da pronúncia. Entretanto, às vezes, pode haver a transmissão oral de um cultismo indireto, isto é, a partir de outra língua que não o grego ou latim. Nesse caso, pode ser reproduzida a pronúncia dessa língua com adaptação da grafia. Por exemplo:

francês *estime* ► inglês *esteem* [is'ti:m] (e não ★*estime* [is'tajm]).

Assim como acontece com palavras vulgares, cultismos também podem ser emprestados acompanhados de desinências, que se incorporam ao radical da palavra na nova língua. Os cultismos franceses *privé* e *habitué* passaram sem alteração ao port. *privé* e *habitué*, respectivamente, portanto com manutenção da desinência de particípio -é [e] e não ★[ɛ]. Por ser um elemento francês e não latino, ele confere aos termos portugueses um caráter híbrido, isto é, semiculto.

⁴ Para a definição de restituição, ver o item *Decalque*, mais adiante.

Decalque

Tradicionalmente chama-se de *decalque*, *calque* ou *clipping* a tradução de um vocábulo (ou dos elementos que o compõem) de um outro sistema diferente do estudado por equivalentes do sistema em questão. Sua representação é $x \geq y$.

Exemplos:

inglês *skyscraper* \geq *arranha-céu*;

inglês *hot-dog* \geq *cachorro-quente*;

francês *chou-fleur* \geq *couve-flor*.

O NEHiLP faz um uso ampliado do conceito de *decalque*, pois também o aplica ao caso de substituições parciais. Muitos cultismos, em determinadas sincronias, fizeram uma espécie de “tradução” de sufixos, que são inexplicáveis do ponto de vista fonético. Por exemplo, a palavra *relação* não foi herdada do latim (portanto, não poderíamos indicá-la, como se faz normalmente nos livros de filologia românica e linguística histórica, como latim *relātiō* $>$ *relação*, muito menos como se fosse uma palavra herdada, como latim *relātiōnem* $>$ *relação* ou latim *relātiōne-* $>$ *relação*), tampouco foi um mero empréstimo (e seria impreciso indicá-la como latim *relātiō* \blacktriangleright *relação*, porque não se trata de uma mudança fonética regular, mas sistemática). A forma mais adequada de indicá-la seria latim *relātiō* \geq *relação*. Ou seja, o símbolo \geq usado para decalque também se aplicará para tradução parcial:

francês †*étrangier* \geq *estrangeiro*;

inglês *goalkeeper* \geq *goleiro*;

espanhol *cañón* \geq *canhão*.

Outro caso particular de decalque é o chamado *empréstimo semântico* (também conhecido como *extensão semântica* ou *loanshift*), como no já mencionado inglês *star* (“corpo celeste” \gg “atriz famosa”) \geq *estrela* (ambas as acepções) ou em inglês *mouse* (“animal” \gg “periférico de informática”) \geq português europeu *rato* (ambas as acepções). O empréstimo semântico também ocorre com palavras cultas:

latim *nuclĕus* (“caroço” \gg “núcleo de uma célula ou de um átomo”) \geq alemão *Kern* (ambas as acepções, cf. *Kernphysik*, “Física Nuclear”).

Também se incluem como subtipos de decalques os chamados *empréstimos de restituição* (BIZZOCCHI, 1998, p. 124), nos quais parte de um vulgarismo ou semicultismo estrangeiro é traduzido por elementos cultos gregos ou latinos. Por exemplo:

ing. *feed back* \geq *retroalimentar*.

O mesmo pode ocorrer com partes de uma palavra como o sufixo *-el*, do francês que se torna *-al* em português em casos como francês *opérationnel* ≥ *operacional*.

Pode-se dizer que houve uma restituição parcial quando algumas partes semicultas ou vulgares são traduzidas por correspondentes cultas e outras não, como em:

inglês *starship* ≥ francês *astronef* ≥ *astronave*.

O inverso também pode acontecer. Quando um decalque ocorre por meio da substituição de elementos gregos ou latinos por equivalentes vernáculos (*tradução de cultismo*). Nesse caso, ocorre a substituição de parte da palavra culta por outra de natureza semiculta, semivulgar ou vulgar, com significado equivalente. Isso pode ocorrer de forma total ou parcial. Exemplos:

latim *superpōnĕre* ≥ *sobrepôr*;

latim *interrumpĕre* ≥ *interromper*;

latim *perfectus* ≥ *perfeito*;

latim *commōtiō* ≥ *comoção*.

O mesmo pode ocorrer para étimos semicultos ou vulgares:

francês *désordre* ≥ *desordem*.

A pesquisa etimológica dos empréstimos e decalques deve ser metódica, pois nada garante que haja conexão etimológica em palavras que podem ter surgido de maneira independente e motivadas pelo próprio conceito e não uma pela outra. Por exemplo, não se pode afirmar com segurança que *saca-rolhas* seja uma tradução do francês *tire-bouchons*, nem que *acendedor de cigarros* seja tradução do inglês *cigarette lighter*. Como o significante é fortemente motivado pelo significado (assim como *laranjeira* é motivado por *laranja*, o que leva praticamente todas as línguas a nomearem a árvore a partir do nome da fruta), é muito difícil afirmar que uma língua tenha influenciado outra na escolha da designação. Todavia, tampouco se pode afastar essa hipótese: quando um artefato de origem estrangeira é introduzido numa sociedade, é natural que, junto com a coisa, venha o nome. Portanto, quem cunhou *saca-rolhas* ou *acendedor de cigarros* em português certamente não desconhecia as denominações originais desses objetos. Estes são, pois, casos em que a explicação exata da etimologia do vocábulo depende de dados empíricos raramente disponíveis, o que pode levar esses itens lexicais a permanecer indefinidamente como de classificação problemática ou insolúvel, diminuindo, por assim dizer, o grau de certeza da proposição etimológica.

Derivação

O símbolo $x \rightarrow y$ designa amplamente uma palavra y formada em sincronia dentro de um sistema por diversos processos, como *prefixação*, *sufixação*, *parassíntese*, *derivação regressiva* ou *derivação imprópria* (também conhecida como *conversão*), seja a partir de radical vulgar, seja herdado ou estrangeiro. Exemplos:

cabeça \rightarrow *cabecear*;
mesa \rightarrow *mesário*;
saudoso \rightarrow *saudosismo*;
transação \rightarrow *transar*.

É preciso muito cuidado, pois muitas palavras são erroneamente classificadas nos dicionários etimológicos como derivadas, quando são, na verdade, traduções. Exemplo:

desafiar \leq francês *défier* (e não “*desafiar* \leftarrow francês *défier*”).

Um outro caso em que o símbolo \rightarrow é utilizado é o caso conhecido como *neologismo flexional* (ALVES, 2007). Quando o português cria *soldada* a partir de *soldado*, *professora* de *professor* e *presidenta* de *presidente*, trata-se de um tipo de derivação semelhante à que ocorre em espanhol e italiano quando derivam *banano* (“bananeira”) de *banana* por analogia com *pero/pera* ($<$ lat. *pirum/pira*, de *pirus/pirum*), isto é, a derivação ocorre por mudança de vogal temática (pois não há flexão de gênero em substantivos). Trata-se de um caso bem distinto da mera flexão. Uma forma como *cantamos* não deriva de *cantar* nem de nenhuma outra flexão, pois no sistema não existe hierarquia nem, propriamente dizendo, uma flexão mais antiga que a outra. Casos como esses são simplesmente indicados da seguinte forma:

cantamos \sim *cantar*,

apontando que as duas formas flexionais se encontram dentro do mesmo paradigma.

Há derivações que são formadas por meio de radicais cultos e afixos igualmente cultos. Nesse caso, não se trata de empréstimos, pois o étimo não pertence a sincronia alguma e a sistema algum. Por exemplo, o português *iniciativa* é um empréstimo do francês *initiative*, que foi criado a partir de um inexistente latim \star *iniciativa*, derivado de *initiare*, ou seja, utilizaram-se elementos do latim mas a palavra latina propriamente dita não existe. A indicação dessa proposta etimológica seria, porém, como se a palavra latina estivesse na mesma sincronia do francês, pois foi a fonte de inspiração do neologismo culto:

latim *initiare* \rightarrow francês *initiative* \blacktriangleright *iniciativa*.

Outros exemplos:

latim *fractus* → inglês *fractal* ► *fractal*;

grego γένος → alemão *Gen* ► inglês *gene* → inglês *genome* ► *genoma*.

Chama-se *truncção* (ALVES, 2007, p. 68) a eliminação de uma parte, geralmente a final, da sequência lexical. O resultado desse processo é parte de um elemento de composição ou a sequência formada por elemento e parte de outro.⁵ Representa-se a parte truncada entre barras verticais. Exemplos em português:

preju|ízo | → *preju*;

micro|computador | → *micro*;

vice|-presidente | → *vice*;

ex|-marido | → *ex*;

francês *métro|politain* | → francês *métro* ► *metrô*;

francês *photo|graphie* | → francês *photo* ► *foto*;

francês *cinéma|tographie* | → francês *cinéma* ► *cinema*.

Por vezes, a derivação ocorre a partir de radical culto com afixo semiculto ou vulgar (port. *agricultável*, *deseducar*) ou a partir de radical semiculto com afixos de qualquer natureza (fr. *désordre*). Nesta categoria entram também os casos da chamada *falsa derivação* (ou *pseudoderivação*). Ocorre que, muitas vezes, ao importar um vocábulo greco-latino, uma língua lhe acrescenta sufixos vulgares que nada alteram no significado da palavra, ao mesmo tempo que a forma sem o sufixo não existe na língua (ou existe, mas a forma sufixada não deriva dela). Trata-se, neste caso, de uma *pseudossufixação*, e o elemento acrescentado é, na verdade, um *pseudossufixo*, uma vez que inexistente a palavra primitiva de onde a suposta derivação originária. Por exemplo:

latim *commodus* ► inglês *commodious* (não há inglês ★ *commod*);

francês *photographe* ► inglês *photographer* (pois inglês *photograph* é derivação regressiva);

latim *philosōphus* ► inglês *philosopher* (não há inglês ★ *philosoph*);

latim *litterārius* ► alemão *literarisch*;

latim *physīcus* ► alemão *physikalisch*;

latim *mūsīcus* ► francês *musicien*.

⁵ Esses elementos são conhecidos como *prefixoides*, *sufixoides* ou, mais genericamente, *quase-morfemas* e *fractormorfemas*, dentre outros nomes. Composições feitas com esses elementos são chamadas de *palavras-valise* e o processo, por vezes, é conhecido como *recomposição* (LINO, 1990, p. 30-31; ALVES, 2007, p. 69).

Composição

Uma palavra composta é formada pela composição entre dois ou mais radicais (vernáculos ou estrangeiros) em uma dada sincronia. Só entram nesta categoria as palavras que tiverem efetivamente sido criadas por esse processo na própria língua que se está analisando. Assim, *puxa-saco*, *pernilongo* e *cabisbaixo* são compostos vulgares legitimamente portugueses. Já *cachorro-quente*, citado no item anterior, não é resultado de composição em português, mas, como vimos, tradução do ing. *hot-dog*. A notação da composição é $x + y \rightarrow z$.⁶

Grande parte dos cultismos existentes atualmente não existia em grego ou latim, tendo sido criada nas próprias línguas europeias modernas, especialmente no francês e no inglês e, secundariamente, no italiano e alemão, por serem línguas mais influentes culturalmente em determinadas sincronias. Como boa parcela desses novos cultismos são termos técnico-científicos, é natural que surjam nas línguas dos países onde ocorre a maior parte das inovações científicas e tecnológicas. Esses cultismos são criados basicamente de duas maneiras: por composição entre radicais cultos ou por derivação a partir de um radical culto com afixos igualmente cultos.

Por exemplo:

grego ἕλιξ + grego πτερόν → francês *hélicoptère* ► *helicóptero*;
francês *social* + francês *démocrate* → francês *social-démocrate* ► *social-democrata*.

A composição também pode ocorrer entre elementos que sofreram truncação. Exemplos:

show + *|co|mício* → *showmício*;
brasi|leiro + *|para|guaio* → *brasiguaiio*;
portu|guês + *|espa|nhol* → *portunhol*;
francês *inform|ation* + francês *|auto|matique* → francês *informatique* ► *informática*;
inglês *repro|duction* + inglês *|photo|graphy* > inglês *reprography* ► *reprografia*;
inglês *auto|mobile* + inglês *part* → inglês *autopart* ≥ *autopeça*;
inglês *mini|ature* + inglês *skirt* → inglês *miniskirt* ≥ *minissaia*.

Analogia

O fenômeno da analogia⁷ é uma transformação que ocorre em sincronia e dentro de um mesmo sistema. Na analogia, uma palavra ou um grupo de palavras (um *molde*)

⁶ A denominação *composição híbrida* é por vezes dada àquela forma construída com um radical culto e um semiculto, com um radical culto e um vulgar, com um radical semiculto e um vulgar ou, ainda, com dois radicais semicultos (por exemplo, port. *auriverde*, *rubro-negro*, *bafômetro*).

⁷ O produto de muitas analogias é às vezes conhecido como *etimologia popular*.

atua sobre outras formando um terceiro elemento. A analogia, propriamente dita, não está na mesma *dimensão* do fenômeno diacrônico, mas deveria ser representada *perpendicularmente* a ele. Como essa notação é complexa, coloca-se a seta especial que representa o fenômeno analógico (\Rightarrow) na direção oposta à da *transmissão* (por herança ou por empréstimo). Assim sendo, se x influencia y de tal modo que z reflete essa influência, o étimo será marcado como $z \Rightarrow y < x$ ou como $x > y \Leftarrow z$ (outros símbolos podem ocorrer no lugar de $>$ dependendo do caso), onde dizemos que x é o étimo, y é o produto analógico e z é o molde analógico.

Exemplo:

latim **foresta* $>$ *floresta* \Leftarrow *flor* (e não \star *foresta*);

latim *consecrare* \blacktriangleright francês *consacrer* \Leftarrow francês *sacrer* (e não francês \star *consécrer*).

Em alguns casos, um cultismo pode ser emprestado e dispor dos mesmos elementos cultos, porém combinados de modo incompatível com o sistema fonológico do grego ou do latim. Por exemplo, o espanhol fez corresponder aos vocábulos latinos *dīminuĕre*, *immortālis* e *commōtīō* as formas *disminuir*, *inmortal* e *conmoción*, ou seja, reabilitou formas arcaicas do latim, anteriores à assimilação (embora **dīminuĕre*, **immortālis* e **commōtīō* não sejam atestadas, mas dedutíveis). No entanto, o surgimento dessas formas no espanhol não deve ter-se dado por um desejo de reconstruir a forma primitiva dessas palavras, mas por efeito da analogia com outras palavras (por exemplo, *disponer*, *intenso* e *contracción*). Outras vezes, uma palavra culta sofre mudanças fonéticas, típicas de palavras vulgares, embora haja problemas com respeito às sincronias envolvidas. Esse fenômeno, por vezes, é conhecido como *metamorfismo* (BIZZOCCHI, 1998, p. 104). Às vezes, essas mudanças são obrigatórias. A investigação das sincronias em que tais mudanças ocorreram é um trabalho urgente para o NEHiLP, pois caracterizar tais formas como resultados de analogia só pode ser feito por meio da determinação do vocabulário das sincronias pretéritas em que ocorreram. Exemplos:

latim *secta* \blacktriangleright *seita* \Leftarrow -IT-

latim *doctor* \blacktriangleright *doutor* \Leftarrow -UT-

latim *conceptus* \blacktriangleright *conceito* \Leftarrow -IT-

latim *statua* \blacktriangleright *estátua* \Leftarrow E-

Nesse caso também podemos incluir os *semivulgarismos*, que não podem ser considerados empréstimos, mas, ao mesmo tempo, não têm exatamente o mesmo estatuto das palavras herdadas, uma vez que seu uso era inicialmente restrito a ambientes de fala monitorada, como a igreja e a universidade, por exemplo, o que resguardou tais palavras de sofrer todos os metaplasmos regulares a que estavam sujeitas as palavras vernáculas, resultando daí formas híbridas, como lat. *canonicus*, *clericus*, *capitulum* e *articulus*, de que resultaram no português as palavras *cônego*, *clérigo*, *cabido* e *artigo*. A categoria dos semivulgarismos é a única que não está sujeita a ganhar novas

unidades, visto que o fenômeno que as produziu ocorreu uma só vez, na passagem de uma língua histórica para outra.

A atuação analógica pode gerar vulgarismos, semicultismos e, teoricamente, até cultismos. Exemplos:

latim *dromedārius* > francês *dromedier* ► alemão *Trampeltier* ⇐ alemão *trampeln* + alemão *Tier*;

latim *corporalis* ► italiano *caporale* ⇐ italiano *capo*;

latim *impediō* > *impeço* ⇐ *peço* < latim **petiō* ⇐ latim *petō*;

latim *impressa* > *imprensa* ⇐ *prensa* < latim *prehēnsa*).

Substituição

O fenômeno da substituição envolve, normalmente, duas transmissões distintas que se cruzam e é representado por $x \Rightarrow y$.

Tal símbolo é particularmente útil para entender o rearranjo diacrônico pelo qual passam os sistemas. Assim, um subconjunto dos advérbios interrogativos envolve os significados “onde”, “de onde”, “por onde” e “aonde”, que eram expressos em latim respectivamente por *ubi*, *unde*, *qua* e *quo*. Podemos dizer que latim *unde* > *onde*, mas latim *qua* \Rightarrow *por onde*.⁸

Esse é o caso, por exemplo, do fenômeno (comum em diversas sincronias) da *restauração* (também chamada de *refecção*) de uma forma popular (que cai em desuso) por outra culta, de mesma origem (MAURER JR., 1951, p. 62), ou seja, algo como $z > x \Rightarrow x' \blacktriangleleft z$. Essa substituição pode ser parcial ou total. Por exemplo:

latim *silentium* > †*seenço* \Rightarrow *silêncio* \blacktriangleleft latim *silentium*;

latim *flōrem* > †*chor* \Rightarrow *flor* \blacktriangleleft latim *flōs*, *-ris*;⁹

latim *monastērium* > *mosteiro* \Rightarrow *monastério* \blacktriangleleft latim *monastērium*.

Por vezes a restauração é parcial:

latim *inimicum* > †*ẽemigo* \Rightarrow *inimigo* \leq latim *inimicus* (e não ★*inimico*);

latim *fēlicem* > †*fiz* \Rightarrow *feliz* \leq latim *fēlix*, *-icis* (e não ★*felice*).

A restauração, por vezes, também ocorre apenas na grafia, por exemplo:

⁸ Este caso, particularmente, é bastante complexo, pois podemos dizer que numa determinada sincronia: *onde* (< latim *unde*) \sim *u* (< latim *ubi*) e depois, com o desuso de *u*, a palavra *unde* > *onde* (“de onde” >> “onde”), ao passo que, mais modernamente, *a + onde* > *aonde* (“aonde” >> “onde”) como, bem mais antigamente, espanhol *donde* (“de onde” >> “onde”) < *de + onde*.

⁹ A forma refeita pode, por vezes, sofrer novo metamorfismo; é o que ocorre com *flor* > †*frol*.

latim *nīdum* > francês *ni* ⇒ francês *nid* ◀ latim *nīdus*;
latim *salūtem* > francês *salu* ⇒ francês *salut* ◀ latim *salūs, -ūtis*.

Nesses casos, só afeta a grafia, não a pronúncia, mas há casos em que a pronúncia se modifica por causa da grafia, como em latim *sub* > *so* ⇒ *sob* ⇔ latim *sub* ou latim *nāscēre* > *nacer* ⇒ *nascere* ◀ latim *nāscēre*, que no português europeu se pronuncia [nɐʃˈser].

Por vezes, a palavra vulgar supostamente refeita não foi documentada e deve ser reconstruída, por exemplo:

latim *sōlum* > **soo* ⇒ *solo* ◀ latim *sōlum*;
latim *altum* > **outo* ⇒ *alto* ◀ latim *altus*;
latim *clārum* > **charo* ⇒ *craro* ~ *claro* ◀ latim *clārus*.

Algumas restaurações ocorrem por meio de semicultismos. Esse fenômeno é bastante comum em algumas sincronias, como, por exemplo, após a Renascença. Algumas etimologias (simplificadas):

latim *elēphas, -antis* ► †*alifante* ⇒ *elefante* ◀ latim *elēphas, -antis*;
latim *astronomīa* ► †*astrolomia* ⇒ *astronomia* ◀ latim *astronomīa*;
latim *informāre* ► †*enformar* ⇒ *informar* ◀ latim *informare*.

Também se incluem aqui os casos denominados como *transcrição*, nos quais um cultismo é criado para substituir uma parte de um termo culto já existente. Por exemplo:

latim *signātārīus* ► italiano †*segnatario* ⇒ italiano *firmatario* ← italiano *firma*;
latim *jūdaismus* ► italiano *giudaismo* ⇒ italiano *ebraismo* ← italiano *ebreo*.

Conclusão

O presente artigo procurou demonstrar sucintamente como o NEHiLP pretende, a partir de tecnologia computacional inovadora e uma grande equipe de pesquisadores do Brasil e exterior, elaborar um dicionário etimológico da língua portuguesa que se equipare em extensão, qualidade e fidedignidade dos dados às melhores obras do gênero disponíveis no mundo para outros idiomas. Por sua plataforma computacional, o *DELPO* será não só obra de consulta por especialistas e público em geral como também permitirá pesquisas por tema (palavras de determinada época, palavras de determinada origem, palavras com determinado prefixo ou radical, etc.), além da geração de gráficos, estatísticas, relatórios e muitos outros recursos que poderão vir a posicionar o Brasil e a língua portuguesa como futuras referências no campo da ciência etimológica.

Como vimos, um dos aspectos inovadores do projeto *DELPo* é a adoção de novos conceitos e categorias classificatórias dos fenômenos etimológicos, com a consequente introdução de uma nova notação simbólica que dê conta dos mesmos. Espera-se que tais conceitos e símbolos se tornem, progressivamente, correntes na literatura etimológica e filológica, o que será mais uma contribuição brasileira ao avanço científico na área.

De todo modo, como nenhuma teoria científica é pronta e acabada, a própria práxis da pesquisa e elaboração do dicionário revelará se tais inovações são suficientes ou se novos conceitos e símbolos precisarão ser introduzidos, ou mesmo se os atuais precisarão ser modificados ou corrigidos. Sugestões e contribuições serão sempre bem-vindas.

VIARO, M.; BIZZOCCHI, A. Proposal of new concepts and a new notation in formulating etymological propositions and discussions. *Alfa*, São Paulo, v.60, n.3, p.579-601, 2016.

- *ABSTRACT: This paper aims to present the work methodology and the main theoretical postulates that guide the preparation of DELPo (Etymological Dictionary of the Portuguese Language), under the responsibility of NEHiLP (Center for Support to the Research in Etymology and History of the Portuguese Language), as it comes to an innovative project both in its theoretical and empirical-operational aspects. It is intended here to present the conceptual, notational and terminological innovations proposed by the authors, that support the preparation of DELPo. For this, we rely on theoretical references on neology (ALVES, 2007; BARBOSA, 1993), lexicogenesis (BIZZOCCHI, 1998), and etymology (VIARO, 2011) and propose an update on the symbols used in the formulation of etymological propositions, which can at the same time clear up ambiguities and inconsistencies of traditional notation and answer for the conceptual innovations introduced here. It is intended that both the described etymological approach and the symbols corresponding to them become standard in the research in etymology and etymological lexicography.*
- *KEYWORDS: Etymology. Lexicography. Etymological dictionaries. Etymological propositions. Etymological discussion. Etymological symbolism. Linguistic notation.*

REFERÊNCIAS

ALVES, I. M. **Neologismo**: criação lexical. 3.ed. São Paulo: Ática, 2007.

BARBOSA, M. A. **Léxico, produção e criatividade**: processos do neologismo. 3.ed. São Paulo: Plêiade, 1996.

BARBOSA, M. A. Acrograma e sigla: estatuto semântico-sintático e tratamento na obra lexicográfica. In: REUNIÃO ANUAL DA SBPC, 45., 1993. **Anais...** Recife: Sociedade Brasileira para o Progresso da Ciência, 1993. p. 477.

BECHARA, E. As palavras também escondem a idade. In: ELIA, S. et al. (Org.). **Na ponta da língua 1**. Rio de Janeiro: Liceu Literário Português; Lucerna, 1998.

BIZZOCCHI, A. **Léxico e ideologia na Europa ocidental**. São Paulo: Annablume, 1998.

BIZZOCCHI, A. A ideologia das raízes. **Língua Portuguesa**, São Paulo, v.40, p. 60-63, Feb. 2009.

BIZZOCCHI, A. Processos de formação lexical das línguas românicas e germânicas: uma nova perspectiva teórica. **Domínios de Linguagem: Revista Eletrônica de Linguística**, v.7, n.1, p. 9-39, July 2013.

JESPERSEN, O. **Language: its nature, development and origin**. London: Allen & Unwin, 1954.

LINO, M. T. R. da F. Observatório do português contemporâneo. In: COLÓQUIO DE LEXICOLOGIA E LEXICOGRAFIA, 1., 1990. **Actas**. Lisbon: Universidade Nova de Lisboa, 1990. p. 28-33.

MAURER JR., T. H. **A unidade da România ocidental**. São Paulo: FFCL-USP, 1951.

UNIVERSIDADE DE SÃO PAULO. **Programa Metaplasmodor**. São Paulo, [2016]. Disponível em: <www.usp.br/nehilp/infos/metaplasmodor-gh-pages/index.php>. Acesso em: 20 jul. 2016.

UNIVERSIDADE DE SÃO PAULO. Núcleo de Apoio à Pesquisa em Etimologia e História da Língua Portuguesa. **Manual do NEHiLP**. São Paulo, 2015. Versão 2.1. Disponível em: <www.usp.br/nehilp/infos/manual.pdf>. Acesso em: 20 jul. 2016.

VIARO, M. E. **Etimologia**. São Paulo: Contexto, 2011.

VIARO, M. E.; FERREIRA, M. J.; GUIMARÃES-FILHO, Z. O. Derivação ou terminação: limites para a semântica, lexicologia e morfologia históricas. In: VIARO, M. E. (Org.). **Morfologia histórica**. São Paulo: Cortez, 2013. p. 58-105.

Recebido em março de 2016

Aprovado em julho de 2016

