

Efeitos da (nova) ortografia no léxico do português

mecanismos gramaticais na grafia de algumas palavras e resultados no uso

Enilde Faulstich

SciELO Books / SciELO Livros / SciELO Libros

FAULSTICH, E. Efeitos da (nova) ortografia no léxico do português: mecanismos gramaticais na grafia de algumas palavras e resultados no uso. In LOBO, T., CARNEIRO, Z., SOLEDADE, J., ALMEIDA, A., and RIBEIRO, S., orgs. *Rosae: linguística histórica, história das línguas e outras histórias* [online]. Salvador: EDUFBA, 2012, pp. 363-380. ISBN 978-85-232-1230-8. Available from SciELO Books <<http://books.scielo.org>>.



All the contents of this chapter, except where otherwise noted, is licensed under a Creative Commons Attribution-Non Commercial-ShareAlike 3.0 Unported.

Todo o conteúdo deste capítulo, exceto quando houver ressalva, é publicado sob a licença Creative Commons Atribuição - Uso Não Comercial - Partilha nos Mesmos Termos 3.0 Não adaptada.

Todo el contenido de este capítulo, excepto donde se indique lo contrario, está bajo licencia de la licencia Creative Commons Reconocimiento-NoComercial-CompartirIgual 3.0 Unported.



Efeitos da (nova) ortografia no léxico do português: mecanismos gramaticais na grafia de algumas palavras e resultados no uso

Enilde FAULSTICH
Universidade de Brasília

Introdução

Este trabalho objetiva apresentar algumas reflexões de natureza linguística, subjacentes à nova ortografia da Língua Portuguesa. O princípio e a motivação para o estudo centram-se no fato de que consideramos que as operações da mudança linguística se dão em um lugar em que novas regras podem ser geradas, para que a língua funcione com propriedade. Este lugar é o *fundo lexical*, que, para os fins deste estudo, adaptamos da teoria funcionalista de S. Dik, primordialmente, da Gramática Funcional (GF), de 1978.

A justificação para o estudo é o interesse pelos estudos da ortografia de Língua Portuguesa, desenvolvidos na linha de pesquisa em Léxico e Terminologia, que coordenamos no Centro de Estudos Lexicais e Terminológicos da UnB, especialmente no projeto *Política linguística em correlação com a história da língua portuguesa*.

Para compor o conteúdo aplicado, decidimos estudar somente alguns casos de emprego do hífen, em vista da extensão do assunto no texto do (novo) *Acordo* e da extensão deste artigo.

Sob essa perspectiva, na seção 1, apresentaremos um pouco da história do português; na seção 2, exporemos a metodologia que serve de base ao estudo desenvolvido; na seção 3, apresentaremos alguns casos, com o fim de ilustrar regras do fundo lexical e, na seção 4, discutiremos a funcionalidade do hífen mediante o fundo lexical.

I Entre a pronúncia e a grafia: um pouco de história

No decorrer da história, atos legislativos sobre a língua do Brasil pontuam a feição que a Língua Portuguesa recebeu do lado de cá do Atlântico. Uma das curiosidades, nas distintas sincronias do português do Brasil que vale a pena aqui retomar, diz respeito às diversas denominações, que, registradas nos documentos, a língua recebeu, tais como língua nacional, idioma nacional, língua pátria, língua vernácula, títulos que vão aparecer, principalmente, nos manuais escolares. É a carta Constitucional de 1988, a nova Constituição Federal, promulgada em 5 de outubro, que declara, no Título II, Capítulo III, Art. 13: "A língua portuguesa é o idioma oficial da República Federativa do Brasil". O termo "oficial" denota certo avanço político-linguístico dos legisladores, pois, se o termo não abarca toda a diversidade linguística do país, ao menos demonstra sensibilidade, ao acabar com a impressão de que a língua nacional é o português sem variação, uma mesma língua em todo o território nacional, como se não houvesse uma realidade multilíngue e multidialetal no país. Por outro lado, a fixação do nome língua portuguesa atenuou os eufemismos denominativos.

O Brasil, federação composta, atualmente, de 26 Estados e um Distrito Federal, com uma população estimada, em 2008, pelo Instituto Brasileiro de Geografia e Estatística (IBGE), em 193 milhões de habitantes, se firmou como uma comunidade linguística bastante representativa, separada política e geograficamente de Portugal desde 1822. Por essa época, o falar ibérico, aqui na América, já deixava marcas de que duas variedades do português, a europeia e a sul-americana, caminhavam em sentidos diferentes. Essa diversidade encontra justificação nos procedimentos de implantação e de difusão da língua no Brasil e está descrita em muitas obras que, na contemporaneidade, reconhecem as duas normas. Um dos focos dessa bifurcação é a ortografia.

Até o século XVIII, o falar brasileiro se manteve fiel à mesma pronúncia de Portugal, porém, em decorrência do contato, a língua, aqui, assimilou influências das línguas indígenas e africanas. Mais tarde, a imigração europeia maciça, instalada principalmente no Centro-Sul, também contribuiu para os fatores de diversificação do português, no Brasil. Mas é no século seguinte, com o movimento modernista, que se instaura o espírito nacionalista nas artes e na língua.

Em 1922, a Semana de Arte Moderna teve por objetivo, entre outros, romper com os modelos tradicionais de Portugal e privilegiar as singularidades do falar brasileiro, com ênfase ao tom coloquial da linguagem. A propósito, Edith Pimentel Pinto, em seu livro *O português do Brasil*, esclarece que:

O nacionalismo característico dos anos 20-45 manifesta-se, em assuntos lingüísticos, não só diretamente, pela campanha em favor da autonomia da variante brasileira, mas também obliquamente, em várias frentes de atuação, pelo esforço de firmar e comprovar essa autonomia. Nessas condições, dar-lhe uma forma gráfica mais ajustada à prosódia brasileira era uma necessidade básica; e outra a emergência de uma expressão literária *sui generis* –

mais evidente no campo do léxico, quer pelo recurso ao regional, quer pela experimentação neológica. (PINTO, 1981, p. XLII-III)

No percurso de construção da identidade nacional, em 1958, nasceu o projeto da Nomenclatura Gramatical Brasileira-NGB, resultado dos trabalhos de uma comissão formada por filólogos e linguistas brasileiros. Em Portaria ministerial nº 36, de 28 janeiro de 1959, o Ministro da Educação e da Cultura recomenda, no Art. 1º:

“a adoção da Nomenclatura Gramatical Brasileira no ensino programático da língua portuguesa e nas atividades que visem à verificação do aprendizado, nos estabelecimentos de ensino”. E no Art. 2º aconselha que entre em vigor: “a) para o ensino programático e atividades dele decorrentes, a partir do início do primeiro período do ano letivo de 1959; b) para os exames de admissão, adaptação, habilitação, seleção e do art. 91, a partir dos que se realizarem em primeira época para o período letivo de 1960.” (GUÉRIOS, 1960)

Nos entremeios da fixação de uma norma escrita brasileira, foi publicada, em 1975, a primeira edição do *Novo dicionário AURÉLIO da língua portuguesa*, um marco de tecnologia linguística do português no mundo. O professor Aurélio Buarque de Hollanda Ferreira e uma equipe de colaboradores escreveram a grande obra lexicográfica da língua portuguesa corrente. O dicionário *Aurélio*, na terceira edição, apresenta um *corpus* atualizado, revisto e ampliado, com 435 mil verbetes, locuções e definições. Esta edição de 2004, disponível também em CD-ROM, permite-nos fazer buscas cruzadas de itens lexicais e de formantes, de maneira acelerada.

Em 2001, apareceu o *Dicionário HOUAISS da língua portuguesa*, elaborado no Instituto Antônio Houaiss de Lexicografia e Banco de Dados da Língua Portuguesa, no Rio de Janeiro. Pouco mais de um ano antes de ser publicado o dicionário, o lexicógrafo Houaiss faleceu, e os trabalhos foram concluídos pela equipe encabeçada por Mauro de Salles Villar. É notório que, em 2009, saiu a primeira edição “com a nova ortografia da língua portuguesa”, em formato impresso e em CD-ROM, resultante das mudanças indicadas no texto do *Novo acordo ortográfico da língua portuguesa*. Nessa edição de 2009, o *Dicionário eletrônico HOUAISS* possibilita, como condição de consulta, selecionar o dicionário desejado, se o *Dicionário da língua portuguesa*, se o *Dicionário de elementos mórficos*, sendo que este último organiza as entradas pelas partes, denominadas de desinência, elemento de composição, prefixo e sufixo.

As mudanças na ortografia da língua portuguesa passaram a valer, no Brasil, a partir de 1 de janeiro de 2009. Há, contudo, um histórico de tentativas de unificação ortográfica, que pode ser resumido na sequência seguinte.

Em 1911, foi adotada em Portugal a primeira grande reforma ortográfica, que não fora extensiva ao Brasil; por sua vez, em 1912, a Academia Brasileira de Letras aprovou uma lei que privilegiava a escrita baseada na pronúncia, a chamada escrita fonética.

Em 1931, foi aprovado o primeiro acordo ortográfico entre Portugal e Brasil, mas, como não produziu o efeito desejado – o da unificação –, provocou uma *Convenção ortográfica*, que se realizou em 1943. Antes disso, em 1934, a Nova Constituição brasileira

revogou leis e decretos anteriores e retornou à ortografia que vigia em 1891. Mais tarde, em 1938, o Brasil recuperou o acordo de 1931, com ligeiras alterações.

Em 1943, a *Convenção ortográfica* procurou sanar as divergências entre Brasil e Portugal. Mas a Convenção não surtiu o efeito desejado. Em 1945, Getúlio Vargas nomeou uma Comissão para realizar uma reforma ortográfica, porém, com a queda de Vargas, desencadeou-se uma série de ofensivas contra o referido Acordo. Mesmo sem acordo, em 1947, saiu, nos dois países, a publicação do *Vocabulário ortográfico resumido*.

Em 1971, no Brasil, a Lei nº 5675, de 18 de dezembro de 1971 “aprova alterações na ortografia da língua portuguesa e dá outras providências”, sob o governo de Emílio G. Médici. Por sua vez, em 1973, por meio do Decreto-Lei n.º 32/73, de 6 de fevereiro, em Portugal, foram suprimidos os acentos que marcavam a sílaba subtônica nos vocábulos derivados com o sufixo *-mente* ou iniciados por *-z-*, como já havia sido feito no Brasil. Em 1975, animaram-se novamente os dois países, e, assim, a Academia Brasileira de Letras e a Academia de Ciências de Lisboa elaboraram um projeto de acordo, o qual não foi aprovado oficialmente.

Em 1986, as *Bases analíticas da ortografia simplificada da língua portuguesa, de 1945, renegociadas em 1975 e consolidadas em 1986*, foram discutidas, no Rio de Janeiro, pelos representantes da Comunidade dos Países de Língua Portuguesa (CPLP), mas nunca foram implementadas.

Em 1990, as delegações da CPLP, reunidas em Lisboa, deliberaram sobre um novo *Acordo ortográfico da língua portuguesa*, com a previsão de entrada em vigor em 1 de janeiro de 1994. Mas no Brasil, somente em 1995, precisamente em abril, no Governo de Fernando Henrique Cardoso, o Senado Federal aprovou o texto do *Acordo ortográfico da língua portuguesa*, pelo Decreto Legislativo nº 54 do mesmo ano. Em 1998, um *Protocolo modificativo ao acordo ortográfico da língua portuguesa* foi elaborado e assinado, na cidade de Praia, Cabo Verde; esse primeiro Protocolo retirou do texto original a data de entrada em vigor do [novo] *Acordo ortográfico*, aprovado em 1990.

Em julho de 2004, um *Segundo protocolo modificativo ao acordo ortográfico* foi aprovado em São Tomé e Príncipe, com a meta de que o *Acordo ortográfico* de 1990 fosse ratificado por três países membros, para que entrasse em vigor nesses países, e não por todos os países. Em 2008, o Brasil ratificou o *Protocolo modificativo* e, em decorrência, em 29 de setembro de 2008, o Presidente Luís Inácio Lula da Silva assinou as mudanças da ortografia da língua portuguesa usada no Brasil, que passaram a valer a partir de 1 de janeiro de 2009.

A adoção da NGB, a publicação de obras que a adotaram, principalmente, gramáticas, e a publicação de dicionários elaborados por equipes brasileiras fixaram a utilização de uma norma própria do português do Brasil. É o *standard* do português sul-americano que passou a prevalecer na produção da escrita brasileira. Esses passos foram definitivos para a normalização da língua portuguesa *do e no* Brasil (FAULSTICH, 2001, p. 112-113).

Antes de prosseguir, ressaltamos que a síntese apresentada serve para reavivar momentos em que a ortografia, sob a interpretação de documentos históricos, é um divisor

de variedades da língua portuguesa, como se somente a forma de grafar palavras determinasse uma língua de cá e uma língua de lá. Observemos, então, pontos que merecem reflexão, como i) a escrita de uma língua natural segue regras de acordo com sua origem, ii) a forma de ortografar pode ser alterada por legisladores que veem o código como um registro extralinguístico, iii) as regras de ortografia não são, necessariamente, regras que decorrem do sistema de uma língua, mas que só funcionam se forem sistêmicas. Visto desse modo, a redação deste artigo é motivada pelo desejo de experimentar a aplicação de regras funcionais a questões de ortografia, para tentar explicar por que a escrita de algumas palavras deve ser de uma ou de outra forma. Com a intenção de estudar o assunto, passaremos à seção seguinte.

2 A metodologia

A fim de desenvolver o argumento que sustenta nossa compreensão de que, de fato, alguns aspectos da mudança na ortografia do português seguem princípios de análise teórica e prática que nos permitem interpretar os fenômenos linguísticos, de acordo com os fatos da língua portuguesa escrita, adotamos os procedimentos metodológicos registrados nesta seção 2.

2.1 Os dados

Para o desenvolvimento do tema, extraímos dados significativos de documentos e de obras, a saber:

- i) do texto do *Acordo ortográfico da língua portuguesa* de 1990, ratificado no Brasil, em 29 de setembro de 2008, e em vigor desde 1 de janeiro de 2009. Serviram a nossos propósitos alguns exemplos e algumas informações de natureza explicativa.
- ii) do *Dicionário HOUAISS da língua portuguesa*, com nova ortografia, Rio de Janeiro, Objetiva, 2009, 1ª edição, do qual retiramos palavras escritas com hífen e os respectivos significados contidos na definição.

2.2 A teoria

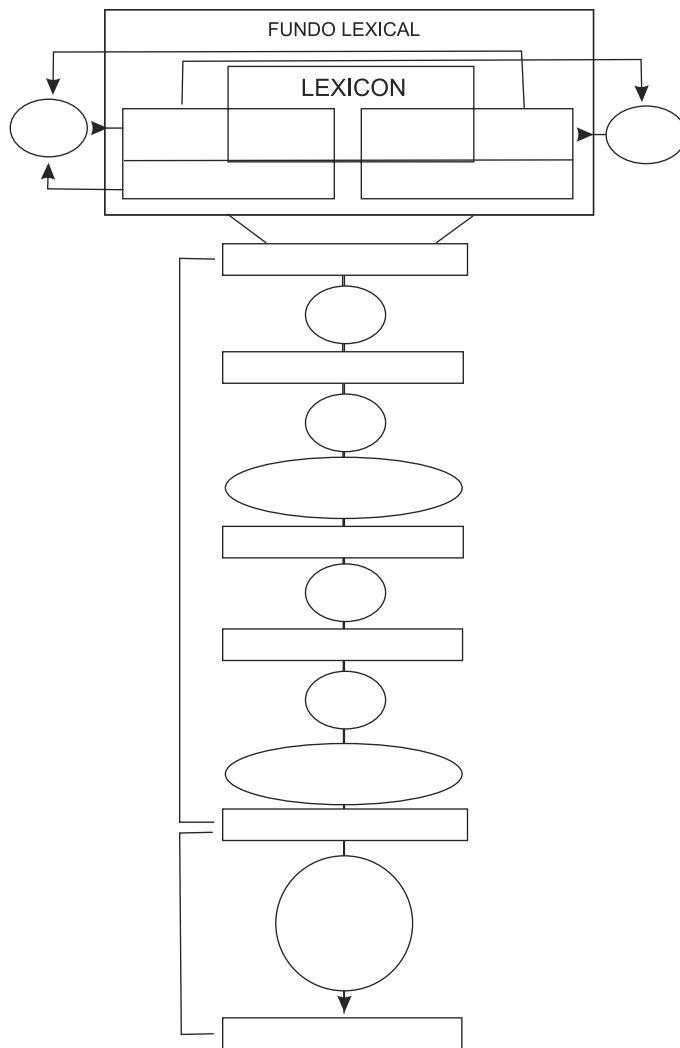
A fundamentação teórica que subsidia nossas reflexões está centrada no funcionalismo de S. Dik (1978) – no paradigma funcional –, de onde extraímos as ideias básicas sobre fundo lexical e dali ampliamos em vista da aplicação. A interpretação de fundo lexical tem sido uma constante nos estudos desenvolvidos no Centro Lexterm, primordialmente em dissertações de Mestrado e em teses de Doutorado.

Como base teórica para a composição do pensamento desta exposição, faremos uma breve discussão sobre FUNDO LEXICAL, na seção 3.

3 O fundo lexical

Toda língua possui um fundo lexical, que constitui um dos fundamentos da expressão linguística e que representa o “dicionário mental” do indivíduo. O fundo lexical é, por conseguinte, um componente no qual se acumulam todos os elementos léxicos de uma língua – predicados e palavras –, assim como as regras, por meio das quais é possível criar novas entidades de um modo produtivo. Na figura a seguir, apresentamos uma imagem “vazia” do esquema criado por Dik (1978); aparecem na imagem dois focos de representação da gramática funcional, o fundo lexical e o *lexicon*, que emolduram o desempenho das entidades que aparecerão na imagem “cheia”, adiante.

Figura 1: A organização da Gramática Funcional



Nesta breve exposição, dizemos que, na organização geral da Gramática Funcional (GF), as representações de FUNDO estabelecem a correspondência entre ontologia e estrutura, como postula a GF dikiana. Na composição do *fundo*, ontologia refere o estudo

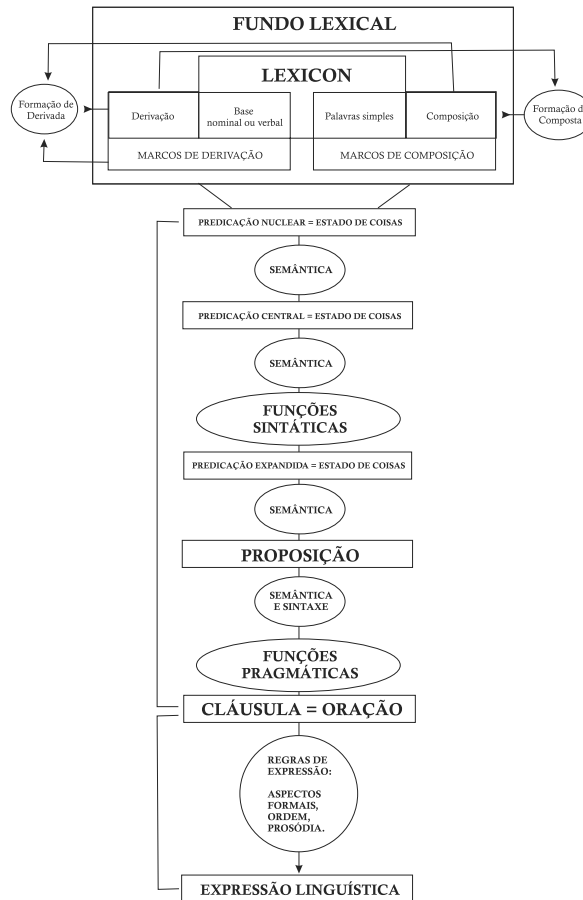
das propriedades gerais, ao mesmo em tempo que as propriedades particulares do ser, e, na nossa visão, dos seres em movimento funcional, em que os estados temporais formam “uma série sistêmica de fases associadas”, de “eventos”, “de ações interconectadas”. Séries dessa natureza comportam processos, os quais podem ser biológicos, mentais e históricos, além de entrelaçarem-se com os momentos políticos e estratégicos (cf. CASTRO, 47-48), enquanto, por sua vez, a estrutura contempla a gramática de uma língua, observada por meio da descrição e da análise, do ponto de vista de sua função, no ato da comunicação.

As bases, em seu conjunto, estruturam o subcomponente do *fundo*, denominado **Lexicon**. O *lexicon* é formado por regras abstratas, que, segundo nosso entendimento, estão de acordo com a origem da língua. No caso do português, essas regras de fundo são estruturadas desde o latim. Para Neves (2002, p. 79-80), a organização da gramática tradicional baseia-se na tradição grega e latina, na qual as classes de palavras foram instauradas como partes do discurso e vistas, basicamente, no funcionamento sintagmático. É esse o ponto principal das reflexões teóricas que nos motiva a descrever como funcionam os mecanismos gramaticais na grafia de algumas palavras e que efeitos semânticos, no uso, resultam dessa aproximação. Veremos a aplicação mais adiante, na seção 4.

Dando continuidade à exposição da figura 1, na imagem “cheia” (ver Figura 2), o *fundo lexical*, que é alimentado pelo *lexicon*, contém as bases que estruturam uma língua. A organização hierárquica das entidades serve para demonstrar como uma língua funciona, por meio de cálculos operacionais semânticos, sintáticos e pragmáticos.

Destacam-se, na Figura 2, a seguir, as funções de estrutura ontológica e gramatical, que emolduram a compreensão do ‘estado de coisas’, como as vemos no mundo. Na figura, aparecem duas funções: as sintáticas e as pragmáticas. No entanto, mesmo não sendo denominada como tal, a função semântica é um domínio com papel fundamental na FG, em vista de ser, de fato, uma função de relação, no escopo representacional. No conjunto, as funções sintáticas, as pragmáticas e as semânticas produzem as regras do cenário para a expressão linguística, que, é por fim, o componente de saída, no evento discursivo.

Figura 2: Adaptação da organização geral do Fundo Lexical ¹



Vistas separadamente, as **funções semânticas** especificam o papel que desempenha uma entidade no evento ou situação. Existe correspondência entre estas funções e o Estado de Coisas e entre este e as funções sintáticas e pragmáticas. É esta relação de dependência que cria as condições para o ato de fala realizar-se na cláusula e ser expresso na língua. Ressalvemos que o **estado de coisas** é uma entidade que se define como “a concepção de algo que pode ser entendido em algum mundo”(DIK, 1978, p. 105)

As **funções sintáticas** especificam a perspectiva de apresentação de um evento ou estado e as **funções pragmáticas** especificam o caráter informativo que recebem os constituintes que organizam uma expressão linguística.

Como já referido, vale lembrar que os princípios metodológicos que organizam a GF de Dik se dividem em três componentes fundamentais, a saber: *fundo*; *expressão da cláusula*; *expressão linguística*.

O *fundo* contém as estruturas linguísticas que satisfazem a duas funções comunicativas básicas, que são **predicar** e **referir**. Predicar é indicar propriedades ou estabelecer relações entre entidades. Referir é guiar o interlocutor no processo de identificação de uma entidade específica. Todas as línguas possuem fundo lexical. O *fundo*, por ter como papel

¹ Esta figura é uma adaptação feita por nós para os fins de nossas pesquisas. Traduzimos e adaptamos do inglês para o português as expressões que a preenchem. Veja a figura original no final deste artigo.

representar o “dicionário mental” do indivíduo, acumula todos os elementos léxicos de uma língua, que são predicados e palavras, assim como as regras por meio das quais é possível criar novos elementos de um modo produtivo. A cláusula é a oração em que a expressão é usada, e a expressão linguística é a denominação da forma final das palavras (lexemas) e dos elementos gramaticais até o momento em que seja necessário expressá-los.

3.1 Algumas ilustrações do funcionamento das regras

Na tentativa de mostrar como funcionam os mecanismos gramaticais na escrita, apresentaremos ilustrações, com o movimento de regras, no português. O ponto de partida é o de que no *fundo* e no *lexicon* se armazenam as regras léxicas, por meio das quais se podem derivar e compor novos elementos de uma maneira produtiva. As bases nominais e as verbais da língua estão no *fundo* e são apreendidas pelo indivíduo durante o processo de aquisição de sua língua. Ora, derivam-se palavras acrescentando à base ou um prefixo, ou um sufixo, ou os dois ao mesmo tempo, formando, assim, uma palavra derivada. Por sua vez, para compor palavras, compõem-se palavras ao adicionar bases primitivas ou derivadas a outras bases primitivas ou derivadas, no processo, cujo resultado é a criação de novas palavras. Como derivação e composição seguem regras de atuação da gramática no léxico, consideramos que derivação e composição são processos lexicais regidos por **marcos**. Marco é padrão básico ou *input* de processos morfossintáticos derivativos e posicionais. Vejamos algumas ilustrações nas seções a seguir relacionadas.

3.1.1 Estudo de caso 1: ‘carrinho’

No aprendizado de línguas, *input* é o conjunto de dados que o indivíduo recebe ao ouvir a língua sendo usada ao seu redor. Na pronúncia da palavra [caiNnho] (carrinho), uma criança de 2 anos e meio tem como marco de derivação a base primitiva **carro** mais o derivacional **-inho**. A operação lexical se faz com o apagamento da vogal temática **-o** da base primitiva [carr0]; dá-se o acréscimo do sufixo **inh-** [carrinh-], mais a reposição da vogal temática **-o** [carrinho]. O estado de coisas está pronto. No entanto, uma regra de expressão, de natureza da prosódia, atua na fonética, porque, no uso da linguagem infantil, a regra produz restrição ao uso do /R/ múltiplo, apagando-o da fala, por ainda não estar finalizado no *input* fonológico infantil. O efeito desse *input* aparecerá, logo depois, na fala mais amadurecida e deverá ser treinado na escrita. Esta é uma tarefa que é muito mais de registro de código do que, efetivamente, de uso na língua.

Nesse processo de morfologia derivativa, o acréscimo de um sufixo a uma base disparou regras léxicas, proporcionou a formação de nova palavra, própria da faixa etária do usuário, e efetuou a comunicação de maneira produtiva.

3.1.2 Estudo de caso 2 : ‘veto consensualizado’

O uso da palavra **veto**, por um adulto que usa linguagem técnica, numa situação de fala como ‘**veto consensualizado**’ se estrutura no marco da derivação deverbal. Na base primitiva **vetar**, opera-se o apagamento do flexional de infinitivo **-r** [veta0-], e, depois,

da vogal temática **-a [vet0-]**; dá-se o acréscimo da vogal temática **-o [veto]**. Nos marcos dessa derivação, há restrições à regra pragmática: o uso de ‘veto’ segue as mesmas regras impostas por ‘vetar’, um verbo de ação com agente concreto e complementado por nome, como em “O presidente vetou a Lei ...” Sendo ‘veto’ um nome abstrato de ação, é predicado ou pela preposição de+nome, como em “O veto do presidente...;” ou pela preposição a+nome abstrato, como em “X sustentou seu veto à entrada de menores...”

Mas, como entender a composição de **veto consensualizado**? Nos marcos da composição de **veto consensualizado**, verificamos que, no uso da palavra criada ***consensualizar**, a operação se inicia na derivada **consensual**, por meio do acréscimo do sufixo **-izar**; dá-se a supressão do flexional de infinitivo **-r [consensualiza0]** e, depois, o acréscimo da forma derivacional **-d(o)/(a) [consensualizad/o/a]**, de participio ou adjetivo. A restrição semântica que atua na regra de composição de **veto consensualizado** pode ser interpretada como: a palavra simples ‘veto’ existe no fundo lexical, por regras naturais da LP; a derivada **consensualizado** entra no *lexicon*, por regras de morfologia derivacional que a LP permite por ter no fundo regras com marcas existentes. A restrição à regra se dá na operação derivacional, em que ao adjetivo ‘consensual’ acrescenta-se o derivativo **-iz [consensualiz-]**, mais a vogal temática **-a [consensualiza-]**, mais a marca de infinitivo **-r [consensualizar]**. Assim, a terminação **-izar** em consensualizar/consensualizado segue uma regra que deriva verbo e nominalização, com a semântica de qualificador de ‘consenso’, que é um nome abstrato. Nos marcos de composição, podemos observar que o princípio aplicado seguiu mecanismos de seleção da morfologia derivativa e de combinação sintagmática. A aplicação dessas regras léxicas proporcionou a criação de nova palavra, de maneira produtiva, segundo a necessidade de expressão.

Neste ensaio de regras, procuramos encontrar subsídios gramaticais e semânticos que comprovem que há casos que, nas chamadas ‘Bases do Acordo’, onde estão registradas as mudanças ortográficas de uso ou de não uso do hífen, não são apenas um jogo de imposições extralinguísticas, mas têm fundamentos linguísticos para o registro das palavras da forma como se apresenta. Com essa percepção, passaremos a interpretar algumas ‘bases’ do (novo) *Acordo ortográfico da língua portuguesa*, principalmente as ‘Bases XV, XVI, XVII – uso do hífen’.

4 A funcionalidade do hífen mediante o fundo lexical

Mudanças propostas no texto do *Acordo ortográfico* em vigor nos motivam, neste estudo, a refletir sobre a atuação de algumas decisões, que se encontram nas ‘Bases do Acordo’ e sobre o resultado dessas decisões no léxico português, numa dimensão contínua, que vai da sílaba ao lexema e às junturas morfossintáticas, tendo como cenário o fundo lexical, primordialmente o ponto de criação de unidades lexicais. Nossa atenção se volta para as terminologias científicas e técnicas, como expressão linguística. As terminologias, constituídas pelo conjunto de termos, são expressões linguísticas, que refletem ontologias.

Por sua vez, as ontologias estão contidas em algum estado de coisas, que revela os constituintes semânticos e sintáticos num evento discursivo. Observamos, por conseguinte, que os termos científicos e técnicos são expressões linguísticas usadas em áreas do conhecimento, em eventos discursivos que referenciam especificações. Uma pista para esse reconhecimento é a forma como são apresentadas nos dicionários de língua, com as marca de especialidade, como *Fís.* (física), *Mat.* (matemática), *Cul.* (culinária), *Med.* (medicina), entre outras. Os dicionários de língua, que relacionam, em ordem alfabética, os termos, cumprem o papel de denotar as terminologias como quaisquer outras palavras, no entanto é preciso reconhecer que um termo científico ou técnico é entidade do discurso desde que esteja em uma cláusula e obedeça às regras de expressão.

O que acabamos de dizer vem ao encontro de um dos legados do *Acordo ortográfico* quando declara que “os Estados signatários tomarão, através das instituições e órgãos competentes, as providências necessárias com vista à elaboração de um vocabulário ortográfico comum da língua portuguesa, tão completo quanto desejável e tão normalizador quanto possível, no que se refere às terminologias científicas e técnicas.” Este declarado interesse dos subscritores do *Acordo* nos motiva, à luz dos estudos linguísticos, a observar a oscilação na grafia de termos derivados e compostos com prefixos e pseudoprefixos em que o hífen ora é usado, ora não é usado, fato que nos induz a questionar que mecanismos serão aplicados na intenção de o vocabulário resultar “tão normalizador quanto possível”.

4.1 Interpretação de alguns dados sob a perspectiva do fundo

A oscilação na grafia de palavras derivadas e compostas está refletida, no *Acordo*, no emprego de estruturas formais, como prefixos e pseudoprefixos. Se prefixo é um conjunto de poucas letras, ou estrutura com uma ou duas sílabas, afixado à esquerda de uma forma pronta, com função semântica autônoma, como em ‘infeliz’, ‘recomeçar’, um pseudoprefixo deve ser entendido como um ‘falso’ prefixo, mas que segue a mesma ordem e tem a mesma função que um prefixo. Serve de ilustração ‘auto-’, que ora aparece registrado nas gramáticas e nos dicionários como prefixo, ora não, mas que, conforme nossa interpretação, desempenha, no estado de coisas, papel de prefixo e tem função semântica autônoma. Entendemos que prefixos e pseudos- são, de fato, lexemas, uma vez que têm lugar no “estado de coisas”, com predicções plenas, respaldados pela entrada autônoma nos dicionários e pela possibilidade de serem ou de não serem hifenizados, na forma da expressão linguística.

O hífen, como é sabido, aparece na palavra escrita e, por isso, torna-se difícil descrever o estatuto teórico desse traço de união, como entidade linguística. No entanto, o uso do hífen pode ser observado no escopo morfológico em que palavras, formadas por derivação ou por composição, estabelecem entre si marcas lexicais que estão mais de acordo com os aspectos formais, de ordem ou de prosódia, segundo a vizinhança do lexema, do que com as exigências da estrutura da língua. A esse respeito, encontramos em Marquilha (1987, p. 108) a seguinte observação:

... o hífen no interior das formas compostas [...] nunca foi muito interrogado na sua forma e função próprias, e lá foi aparecendo e desaparecendo conforme os hábitos tipográficos de cada época.

E mais adiante registra que só em 1739, na *Orthographia* de Madureira Feijó, há o testemunho de o hífen ter sido morfológicamente assimilado à linha de seguimento que, desde os copistas medievais, se traçava no fim da regra (-) (MARQUILHAS, 1987, p. 113)

Em seguida, chama atenção para a regra criada por Barbosa Leão, em 1875, nas *Considerações sobre a orthographia portugueza*, em que o autor afirma, à página 66:

[Há] palavras compostas que se devem escrever de modo que a sua composição fique patente, não só para que isso influa na maneira de dividir algumas, mas também para evitar equívoco no modo de pronunciar-as todas. Ao que podem servir d'exemplo *bemfeito*, *comtudo* e *bemaventurado* [...] Por isso n'estes e análogos casos entendemos que a composição deve ser marcada na orthographia, escrevendo-se *bem-feito*, *bem-aventurado*, etc. (MARQUILHAS, 1987, p. 114)

Em continuidade à interpretação do emprego do hífen, diz a autora, à página 114, que:

E onze anos depois já estava formada uma teoria sobre o assunto, em tudo semelhante à do Acordo de 1986: dois autores, numa *Orthographia Portugueza* de 1886 sugeriram, em relação às palavras compostas, a união “n'uma só sem precisão de hyphen, quando a primeira é invariável” – *guardaroupa*, *guardalouça*, *paraquedas*, *bemvindo*; e a manutenção do sinal em “guarda-marinha, ferro-velho, gentil-homem, etc., em que as primeiras são variáveis, pois no plural se diz ferros-velhos, guardas-marinhas, gentis-homens.²

Podemos ver, nas citações anteriores, que o emprego do hífen na formação e na criação de palavras se mantém instável. Advertimos que nosso propósito neste estudo não é o de descrever uma regra que seja palavra final, mas é interpretar o emprego do hífen, em algumas situações de uso, à luz da função que tenha ou não na extensão sintagmática. Para isso, passamos a alguns comentários pontuais.

4.2 A forma auto-

Dividimos o lexema **auto-** em grupos, de acordo com o significado que apresenta na estrutura sintagmática. Uma estrutura formada por auto- +base é considerada, na tradição, derivação prefixal. No nosso modo de interpretar, o escopo de auto- é composicional, uma vez que é um lexema prefixo que acompanha outro lexema verbal ou nominal.

Apresentaremos, a partir de então, 4 grupos em que auto- se divide por propriedade e por função, como está registrado, atualmente, no *Dicionário HOUAISS*, com a nova ortografia, mas o arranjo é nosso.

² Os dois autores citados por Marquilhas e relacionados na bibliografia da autora são VALENTE/ALMEIDA, Santos / Francisco de, *Orthographia portugueza*, Lisboa 1886. (sic)

4.2.1 Auto- com propriedade ‘si mesmo’ e como agente reflexivo:

(i) autorretrato ‘retrato feito por si e de si’

4.2.2 Auto- com propriedade ‘próprio’, como agente:

(ii) autosserviço ‘serviço executado pelo próprio’

4.2.3 Auto- com propriedade ‘próprio’, como paciente:

(iii) auto-hemoterapia ‘tratamento de si com sangue de si retirado e injetado por outra pessoa’

4.2.4 Auto- com propriedade ‘mesmo’, como agente recíproco:

(iv) automóvel ‘movimento de si produzido por mecanismo de si’

Os termos de (i), (ii), (iii) e (iv) estão no léxico da língua portuguesa e seguem marcos de composição de lexema prefixo+palavra simples. No âmbito da predicação nuclear, a lexicalização de auto+palavra vai ser decidida pela semântica, porque auto- modifica o significado mediante o estado de coisas em que passa a habitar. As propriedades sintáticas e semânticas atuam nas bases, de tal forma que o lexema ganha o significado preciso no estado de coisas pelas funções comunicativas de predicar e referir. Dessa ordenação, qualquer proposição de uso é regulada pela semântica e pela sintaxe, para que, no dicionário mental e, por consequência, na cláusula se complete, e a expressão linguística detenha as propriedades funcionais de serem comunicadas e compreendidas no ato de fala.

Na ordem sintagmática de palavras compostas, o emprego de auto-, ou de outro prefixo lexical, terminado pela mesma vogal com que se inicia o lexema seguinte – como em ‘micro-ondas’ e ‘arqui-inimigo’ – passa a ser, pelo (novo) *Acordo ortográfico*, hifenizado; entendemos que esse papel do hífen tem por finalidade evitar, na escrita, a crase dos sons contíguos idênticos, sem que perturbe o significado que temos das expressões.

Interessa-nos, então, verificar em que lugar do fundo lexical a regra do hífen é exigida pela expressão linguística. Poderemos, para isso, elaborar algumas hipóteses.

Primeira hipótese. Se a inserção do hífen for estruturada no âmbito das funções sintáticas, então o foco de referencialidade, de ordem linguística, é o prefixo lexical, que passa à categoria de adjetivo e deixa de ser pseudoprefixo; passa a nome com semântica própria. Significa dizer que com hífen ou sem hífen a semântica do adjetivo se mantém, portanto, em autosserviço ou auto-serviço, tanto quanto em micro-ondas ou microondas a referência do objeto se mantém, em qualquer estado de coisas, em qualquer ato de fala, sem qualquer restrição de regras, apenas com a exigência gráfica de incluir o -s- no lugar do traço de união para atender a uma regra de expressão de natureza prosódica; pronuncia-se o /s/ surdo no lugar do sonoro e, para isso, o registro gráfico tem de ser feito com -ss-. O resultado é acréscimo de letra, que mal pode ser considerado epêntese, a não ser pela exigência fonética. Em micro-ondas, apesar de a interpretação ser semelhante com relação à autonomia referencial do objeto, o hífen prevê a inviduação dos sons contíguos

próximos, em vista de manter a pronúncia de “o-o”. No caso, a regra de expressão de natureza prosódica é fraca, porque a tendência é o grupo de força, na pronúncia, atenuar um dos sons idênticos e até mesmo reduzir a composta micro-ondas para ‘micro’, com tem sido chamado o objeto, que nem por isso perde o significado e a referência, já que a cláusula dá conta do evento de fala, e “micro” se mantém no estado de coisas como “forno de micro-ondas” ou “forno de microondas” ou, quem sabe, “forno de microndas”.

Segunda hipótese. Se a inserção do hífen for estruturada no âmbito das funções pragmáticas, então o foco é de natureza extralinguística, quer dizer que alguém, em algum momento, decidiu que registrar “auto-hemoterapia” é melhor que ‘autohemoterapia’ ou ‘autoemoterapia’, mesmo que a função semântica se mantenha intata. Com hífen ou sem, a referencialidade de ‘auto-hemoterapia’ se resolve no estado de coisas, uma vez que o termo é técnico e é usado somente em ato de fala em que os interlocutores sabem o que estão dizendo, para quem, quando e onde. No caso, o usuário de linguagem técnica, fixou o conceito operacional de que **X** retira o sangue de **Y** e **X** injeta o sangue de **Y** em **Y**. Eis que a proposição pronta fornece os procedimentos para a estruturação da cláusula. Como o hífen não tem amparo na prosódia, então esse traço de união se torna desnecessário.

Terceira hipótese. Se a inserção do hífen estiver fora dos padrões estruturais da língua e tiver o uso determinado por regras que determinam fatos gramaticais, então a representação desse traço, na grafia das palavras compostas, responde aos anseios de uma política de língua, que pode ser ou não atendida pelos usuários. Essa hipótese bloqueia todas as regras do fenômeno de criação de palavras porque não tem lugar no fundo lexical e não é, por conseguinte, funcional. É o caso de ‘auto-hemoterapia’,³ que se resolve fora da estrutura, e de ‘auto-ônibus’ cujo sentido de ‘auto-’ está fora do paradigma da semântica de ‘próprio’, que contempla, por regra de fundo, o significado genérico de ‘si mesmo’.

No *Dicionário HOUAISS* (2009), os autores incluíram o texto completo do *Acordo ortográfico da língua portuguesa* de 1990, na versão impressa, e, na versão eletrônica, há o *Dicionário de elementos mórficos*, do qual extraímos as entradas que seguem:

aut(o)-

elemento de composição

antepositivo do gr. *autós,ê,ó* ‘(eu) mesmo, (tu) mesmo, (ele) mesmo, (si) mesmo’ (representado em lat. por um reduzido número de helenismos em que *auto-* ocorre como pref.); nessa qualidade, depreendido prov. de *autômato*, começa a aparecer em umas poucas form. mais ou menos internacionais datadas do Renascimento, expandindo-se a partir do sXIX na terminologia científica, de tal modo que hoje tende quase a autonomizar-se como voc.; o V.O. registra, entre *autarcia* e *autozincografia*, cerca de 650 palavras com este pref.; ver *auto*, por sua autonomização vocabular e subsequente função de el.comp. antepositivo

auto-

elemento de composição

antepositivo, com o valor de ‘automóvel’ (1866, no fr. *automobile* ‘que se move por si mesmo’), oriundo de *aut(o)-*, ver, por braq. mais ou menos internacional, como o nosso *auto*; com esse sentido particular aparece em um número expressivo de pal. ger. do sXX: *autobomba*, *autobonde*, *autobus*, *autocaminhão/autocamião*, *autocanhão*, *autocarga*, *autocarro*,

3 Tratamento de certas doenças pela retirada e nova injeção do sangue do próprio paciente (HOUAISS, 2009).

autódromo, autoescola, autoestrada, autogasogênio, autogasógeno, autolotação, automobilismo, automobilista, automobilístico, automobilização, automobilizar, automotor, automotriz, automóvel, auto-ônibus, autopeça, autopipa, autoplano, autorama, autosserviço, autotanque, autotransporte, autotrém, autoveículo, autovia, autoviação.

O verbete auto- ‘elemento de composição’ mostra que grande parte das palavras listadas é termo, e que, por isso, esses termos são usados em discursos especializados; em menor número, aparecem termos usados no dia a dia e, ainda, outros que não são usados, no Brasil, em nenhum evento de fala, como, ‘autobonde’, ‘autocanhão’, ‘autotrém’ e que, sequer, aparecem relacionados, como entradas, no *Dicionário HOUAISS* (2009).

Ao rever a Base XV do *Acordo – Do hífen em compostos, locuções e encadeamentos vocabulares*, encontramos o exemplo ‘paraquedas’, escrito da mesma forma como já decidido por Valente/Almeida, em 1886, conforme citamos anteriormente. No texto do *Acordo*, em Base XV, 1º, Obs., a justificção para a escrita de ‘paraquedas’, ‘paraquedista’, ‘mandachuva’ é que “certos compostos, em relação aos quais se perdeu, em certa medida, a noção de composição, grafam-se aglutinadamente”. Por sua vez, Valente/Almeida defendem “a união n’uma só sem precisão de hyphen, quando a primeira é invariável, como ‘guardalouça’” (cf. MARQUILHAS, 1987, p. 114). No paradigma ‘guardalouça’ poderemos incluir ‘guardachuva’, da mesma forma como ‘mandachuva’, que já parece com essa escrita na Base XV do (novo) *Acordo*. Porém ‘guarda-chuva’ e guarda-louça aparecem hifenizados, enquanto o registro recomendado de ‘manda-chuva’ é sem hífen, como demonstramos anteriormente.

Do ponto de vista da lexicologia funcionalista, a expressão ‘guarda-chuva’ está alocada no mesmo marco composicional de ‘mandachuva’, ‘paraquedas’ e ‘guardalouça’, em que base+base, ou palavra simples+palavra simples, formam uma expressão linguística autônoma, pois as regras semânticas são modificadas pelo estado de coisas, no âmbito da predicação nuclear. Quer dizer que ‘manda’ [manda-r], ‘para’ [para-r] e guarda [guarda-r] passam por derivação deverbal e sofrem nominalização, entrando no léxico por meio de regras sintáticas que possibilitam as estruturas ‘o manda X’, ‘o para X’, ‘o guarda X’ em que as bases ativas, por combinação sintagmática, predicam argumentos exigidos pela primeira base. Perdem, no entanto, o vigor semântico, à medida que ‘mandachuva’ é, por metáfora com ‘ordenar’, “indivíduo com influência”, ‘paraquedas’ é, por antonímia com ‘tombar’, “artefato de pano ou de náilon” e ‘guarda-louça’ é, por identidade com ‘acondicionar’, “armário”. Esses novos significados são frutos de terminologização, que, por consequência, criam unidades terminológicas complexas (UTCs). As UTCs resultam do movimento das funções pragmáticas, em conformidade com as regras de expressão, tais como, ‘valvuloplastia aórtica percutânea com balão’ (cirurgia), ‘bactéria de vida livre’ (biologia) e ‘aguoa de frol cõ almísquer’ (culinária arcaica), ‘clorofluorcarboneto’ (química), entre outras, são termos regidos por regras sintáticas, semânticas e pragmáticas cujos formativos predicam e referem conceitos e organizam uma gramática própria das

linguagens de especialidade e mais distante da tradição da língua comum (FAULSTICH, 2003, p. 12 e ss.).

Nessas composições, a ordem é responsável pela expressão linguística, na cláusula pragmática. O evento de fala denota que os significados implicados se expandiram no fundo lexical e, pela nova estrutura, passaram a ocupar outro lugar na ontologia.

Conclusão

A análise dos dados, com o suporte de uma interpretação de regras funcionalistas, demonstrou que

1 – o hífen é um elemento gráfico, com repercussão na escrita das palavras compostas ou derivadas, mas não é responsável pela lexicalização ou pela terminologização de palavras;

2 – a lexicalização é um processo que resulta de operação de regras, nos marcos derivacionais e composicionais de dois ou mais lexemas, no escopo da lexicologia funcionalista;

3 – a terminologização, do mesmo modo, resulta de operações no escopo do léxico especializado ou terminologia funcionalista;

4 – as unidades terminológicas complexas (UTCs) prescindem de hífen porque as funções semânticas e sintáticas seguem as exigências das regras de expressão - como forma, ordem e prosódia – comandadas pelas funções pragmáticas da cláusula;

5 – o hífen não faz parte do fundo lexical de qualquer língua, porque desempenha o papel de notação ou traço, e para esse papel não há regra de estrutura consistente;

6 – o hífen é uma representação gráfica de superfície, por isso, na escrita da palavra, pode estar presente ou não;

7 – o hífen, na representação gráfica, zela mais pelos aspectos visuais do que estruturais da palavra, como, por exemplo, ‘contra-almirante’ e ‘micro-onda’.

Referências

CASTRO, Suzana de (2008) *Ontologia*. Rio de Janeiro: Jorge Zahar.

DIK, Simon C. (1978) *Functional grammar*. Amsterdam: North-Holland.

FAULSTICH, Enilde (2001). CPLP: um lugar de falas múltiplas. In: SARAIVA, José Flávio S. (Org.). *Comunidade dos Países de Língua Portuguesa (CPLP): solidariedade e ação política*. Brasília: IBRI. p. 105-143.

FAULSTICH, Enilde (2003). Formação de termos: do constructo e das regras às evidências empíricas. In: FAULSTICH, Enilde; ABREU, Sabrina Pereira de (Org.). *Linguística aplicada à terminologia e à lexicologia*. Porto Alegre: UFRGS/IL/NEC. p. 11-31.

GUÉRIOS, R. F. Mansur (1960). *A nomenclatura gramatical brasileira definida e exemplificada*. São Paulo: Saraiva.

HOUAISS, Antônio; VILLAR, Mauro de Salles (2009). *Dicionário HOUAISS da língua portuguesa*. Rio de Janeiro: Objetiva.

MARQUILHAS, Rita (1987). O acento, o hífen e as consoantes mudas nas Ortografias antigas portuguesas. In: CASTRO, Ivo; DUARTE, Inês; LEIRIA, Isabel (Org.). *A demanda da ortografia portuguesa*. Lisboa: Edições João Sá da Costa. p. 103-116.

PINTO, Edith Pimentel (1981). *O português do Brasil: textos críticos e teóricos, 2: 1920/1945: fontes para a teoria e a história*. Seleção e apresentação. Rio de Janeiro: Livros Técnicos e Científicos; São Paulo: Ed. da Universidade de São Paulo.

Figura 3: Overall model of FG

