

Problemas de morfologia e semântica histórica do sufixo –eiro

Mário Eduardo Viaro¹

Departamento de Letras Clássicas e Vernáculas – Universidade de São Paulo (USP)
Caixa Postal 31258 – 01309-970 – São Paulo – SP – Brasil
maeviaro@usp.br

Abstract. *This paper discusses the importance of clear definition of presuppositions in studies of the so-called suffix derivation. A large amount of intuition is often used, which is also mixed with traditional points of view. Tests could easily reveal some lack of scientific behavior and an urgent need of inclusion of history.*

Keywords. *Morphology; Semantics; Portuguese Language; Derivation.*

Resumo. *Este artigo discute a importância de definições claras dos pressupostos em estudos da assim chamada derivação sufixal. Costuma-se utilizar uma grande quantidade de intuição, misturada com pontos de vista tradicionais. Testes podem facilmente revelar uma certa falta de postura científica e uma necessidade urgente de inclusão da história.*

Palavras-chave. *Morfologia; Semântica; Língua Portuguesa; Sufixos.*

1. Introdução

O capítulo da Morfologia referente à formação de palavras tem raízes profundas na historiografia, pois desde a gramática de Dionísio, o Trácio, houve preocupação em classificar e subclassificar classes morfológicas. Entre os nomes, uma separação das mais tradicionais foi a que distinguia palavras simples das derivadas. De forma quase intuitiva aprendemos que *sapato* é uma palavra simples e *sapateiro*, derivada. Com a necessidade de fundamentar um objeto de estudo claro que diferenciasse do da Psicologia, da História, da Sociologia, Antropologia, da Filosofia ou mesmo da Gramática Tradicional, a Lingüística do Pós-Guerra acabou firmando-se cada vez mais em determinados pressupostos que não conseguiu manter de maneira radical. Um dos mais divulgados entende o falante nativo e sua respectiva competência como um árbitro fiel de decisões como a gramaticalidade. Essa posição de juiz, antes apenas intuitiva no Estruturalismo, acabou sendo legitimada, no Gerativismo, e, posteriormente, na Fonologia e na Morfologia, embora os semanticistas, sempre divididos em inúmeras correntes, vissem isso com tácito ceticismo. Como decorrência do papel de julgador do falante nativo, nasceu a convicção, fundamentada em parâmetros epistemológicos, que cedo se tornou um segundo pressuposto: o de que a língua que interessava ao lingüista era a mesma que interessava ao falante, ou seja, sua etapa sincrônica atual, pois o falante nativo *grosso modo* desconhece as etapas anteriores da língua que culminaram nas formas atuais. Dessa forma, surgiu uma estéril discussão entre uma nova geração que preconizava uma primazia dos estudos sincrônicos de uma língua (como sistema) e uma outra, considerada ultrapassada, associada quer à Gramática Tradicional quer à Filologia ou aos Estudos Clássicos, que não abria mão de seus conhecimentos de latim e

da visão diacrônica da língua (vista como instituição). Felizmente a assim chamada guerra entre Sincronia e Diacronia foi superada pelos estudos funcionalistas atuais. Hoje há maturidade suficiente para entender que a língua é sistema *e* instituição (RIO-TORTO 1998, 2004).

Mesmo alguém que, hoje em dia, preconize ainda a primazia do elemento sincrônico, não poderia negar que, ao obter, por meio de comutação, um lexema {*sapat*} e um morfema {*eir*}, a partir de *sapateiro*, chamar isso de *derivação* não deixa de ser paradoxal. Pois *derivar* pressupunha, na visão tradicional de onde vem o termo, que uma palavra provinha da outra ou que vinha *antes* da outra: *sapato* surgiu primeiro e daí veio *sapateiro*. Mesmo quem postula a falta de necessidade diacrônica pensa assim hoje. A diacronia está, portanto, implícita na derivação, pois as duas palavras não surgiram ao mesmo tempo. Essa sutil distinção esbarra, muitas vezes, na questão da competência do falante nativo, que poderíamos chamar de “intuição”. Essa intuição, por sua vez, é formada de regras e de explicações *ad hoc* que jamais poderiam ser idênticas às de uma explicação científica. Essas regras têm conformação dogmática, pois diante de *separar* e *separação*, é comum dizer que a segunda palavra vem da segunda. Formase, assim, uma RFP (regra de formação de palavras) do tipo $x]_{ar} \rightarrow x]_{ação}$, que é totalmente intuitiva e, contudo, válida para grande número de palavras. No entanto, um lingüista sabe que o infinitivo é apenas uma das formas de flexão verbal, a usada nos dicionários de português (que poderiam ter usado outra, como a primeira pessoa do presente do infinitivo nos de latim ou grego ou a terceira pessoa do pretérito, nos de árabe). A “primitividade” do infinitivo é, portanto, algo cultural, extralingüístico, aprendido nas instituições e divulgado oralmente e pela escrita. Em si nada há para afirmar que a forma simples seja a do infinitivo. Um bom exemplo disso é apontado historicamente: com base nessa intuição, ninguém discordaria que *colar* $\rightarrow$ *colação*, mas do ponto de vista histórico, *colação* é muito mais antigo, pois vem do latim *collatio*, derivado de *collatus*, particípio de *conferre*. O verbo *colar*, criado analogicamente, surgiu depois. Dessa forma, a RFP é exatamente o inverso da postulada acima. A afirmação reacionária a esse dado, que declara valer apenas a intuição e não o dado histórico é argumento de cunho emocional e jamais científico, embora não inútil. Se se permite um paralelo, seria o mesmo que um biólogo declarar que as focas, os tubarões, as tainhas, os manatis e as baleias têm aspecto parecido e é o que basta para classificá-los todos como peixes. Se hoje são considerados respectivamente pinípedes, condrícteos, osteícteos, sirênios e cetáceos, classificações totalmente diferentes, isso não se deve a seu aspecto externo, pois entendem que a forma é apenas uma epifania: ela surgiu, pela evolução, pressionada pelo meio que é comum a todos: o ambiente aquático. Se na biologia aceitamos que as baleias são evolutivamente mamíferos e não peixes, a ponto de taxar de ignorante quem pense o contrário e, em lingüística, há relutância para enxergar paralelos que envolvam fatos de idêntica dimensão histórica, só existe uma razão: os dados lingüísticos pertencem a uma dimensão emocional, sobretudo a língua materna, a lembrarmos Bakhtin: “a palavra nativa é percebida como um irmão, como uma roupa familiar, ou melhor, como a atmosfera na qual habitualmente se vive e se respira. Ela não apresenta nenhum mistério. Só pode apresentar algum, na boca de um estrangeiro” (1988:100). Falar algo dela que não tenha correspondência imediata é quase profaná-la. Dessa forma, a língua como um todo se tabuíza. Também o corpo humano, tão íntimo a nós quanto nossas próprias palavras, era objeto de tabuização e, por isso, a Medicina se desenvolveu muito mais lentamente do que os estudos da Biologia. Não estamos sós, portanto.

Mas ao longo de pesquisas históricas, o número de exceções aumenta, de modo que é incômodo pensar em soluções *ad hoc*: parece-nos natural que *coagir* → *coaço*, pois *reagir* → *reaço*, mas apenas a segunda derivação está correta do ponto de vista diacrônico e não a primeira: o latim já tinha a forma *coactio*, da qual se gerou *coaço*, mas o verbo *coagir* veio posteriormente, invalidando a regra da primitividade do infinitivo. O mesmo se pode dizer de prefixos: intuitivamente, *cobrar* → *recobrar*, pois uma forma prefixada dogmaticamente surge a partir de uma não-prefixada, mas historicamente *recobrar* é mais antigo, pois vem de *recuperare*, a forma sem prefixo é posterior. Elencar um ou dois casos e generalizar é próprio da natureza comunicativa humana, não de um estudo científico que vise à descrição dos fatos. Talvez as regras fundamentadas no “erro” intuitivo dos três exemplos acima (para não citar inúmeros outros das derivações regressivas e impróprias) não sejam úteis para estudos de descrição, mas o seja para a Pragmática e outros estudos de comunicação. Também a semelhança da forma dos animais citados acima não é levada em conta pelo taxonomista, mas o é para ecólogos, etólogos, fisiologistas.

O trabalho aqui apresentado faz parte de reflexões de um Grupo de Pesquisa, intitulado Morfologia Histórica do Português, cadastrado no CNPq, do qual sou líder. Os verbetes foram levantados a partir das informações colhidas no dicionário Houaiss da Língua Portuguesa pelas mestrandas Amáble Bianca Nogueira e Nilsa Areán García e pela orientanda de Iniciação Científica Maria Paula Roncaglia. Agradecimentos especiais a Zwinglio Guimarães e Leandro Mariano, doutorandos do Instituto de Física da USP pela manipulação e tratamento de dados e pelo auxílio técnico indispensável que vêm prestando, e para o prof. Dr. Waldemar Ferreira Netto, da USP, que gentilmente cedeu cópia de uma lista por ele preparada, em que constavam as 14002 palavras mais frequentes (com número igual ou maior a 100 ocorrências) extraídas a partir de um imenso *corpus* de textos, dos mais diversos, por ele coletados e digitados (o número total de ocorrência dessas palavras somadas remonta a 15.958.907).

2. Produtividade do sufixo –eiro.

Sob o rótulo “produtividade” surgem na verdade dois problemas: (a) um elemento é dito produtivo quando é possível obter muitas palavras por meio dele. Nesse caso chamaremos o sufixo de *prolífico*. Assim sendo, *-eba* não é um sufixo muito prolífico (*decoreba*, *natureba*, *mistureba*...) ao passo de *-ção* é bastante prolífico. A prolificidade é, portanto, o resultado da produtividade; (b) um elemento também é dito produtivo quando dele se obtém grande número de neologismos, ao passo que não o é se sua produção já estagnou. Nesse caso, manteremos o rótulo *produtivo*. Assim, *anti-* é um prefixo bastante produtivo ainda hoje, ao passo que *cis-* é praticamente um sufixo desusado, exceto talvez no âmbito da Química. A neologia é o grande ponto de avaliação da produtividade de um elemento.

Um sufixo prolífico não é, necessariamente, produtivo (pode ter deixado de formar neologismos após criar muitas palavras) e um não-prolífico pode trazer uma força produtiva subestimada se nos pautamos somente nos dicionários. A língua atual, vista em sua abstrata sincronicidade, possui tendências que não são fáceis de enxergar, o que nos faz legitimar a especulação de como ela seja, baseados em nossos poderes de juiz e na nossa experiência, isto é, nos nossos conhecimentos restritos aos meios por onde circulamos. A língua, porém, não é isso, assim como os animais para um Biólogo não são apenas os que ele, como indivíduo, conhece ou conheceu a partir de sua experiência. Esse cientista, que é ao mesmo tempo observador e juiz, mais uma vez, é

um comportamento herdado da Gramática Tradicional. Uma vez que as ciências nascem ou da observação dos dados ou por meio de soluções apriorísticas plenamente aplicáveis, vemos que os estudos de formação de palavras não se encaixam nem na primeira postura científica, indutiva (pois não sabe depreender os fatos sem se pautar em regras eivadas de exceções), nem na segunda, dedutiva (pois o conhecimento dos fatos é dogmaticamente restrito ao conhecimento do observador).

Objetivamente, o linguísta que não se pauta somente em sua intuição necessita de *corpora* para fazer suas afirmações, que não são leis exatas, e para responder a suas perguntas. O Dicionário Houaiss, por exemplo, é um *corpus* de tamanho bastante extenso, com suas quase 200.000 palavras. Pesquisas em buscadores da Internet como o do site www.google.com, são também fontes bastante exaustivas para tira-teimas, mas de uma coisa não se deve esquecer: eles não são a *langue*, o sistema.

Qual a prolificidade de *-eiro* e de suas flexões (*-eira*, *-eiros*, *-eiras*)? No Houaiss temos, sem homônimos, 4482 palavras (excluindo o próprio sufixo e sete alomorfes) terminadas nessa seqüência de grafemas; com homônimos, o número aumenta para 4667. Excluindo os casos de justaposições não-analisáveis (como *atum-verdadeiro*, *bicho-barbeiro*, ou seja, que têm elementos repetidos no *corpus*, mas não outros como *roupa-velheiro*, *são-joaneira*, que não têm), os casos de composições óbvias (como *bioengenheiro*, *aeropioneiro*), prefixações facilmente parafraseáveis (são excluídos casos como *antibrasileiro*, *antecordilheira*, mas não *desempenadeira*, *despenhadeiro*), redobros (*cheira-cheira*), bem como as variantes, ortográficas e regionais da mesma palavra (*samaumeira*, pois já há *sumaumeira*), além dos casos de pseudo-homônimos (*cortadeira*, como redução de *formiga-cortadeira*) chega-se, numa primeira avaliação, que ainda não exclui erros de avaliação, a 3718 ocorrências. É grande o número de palavras que fogem à nossa experiência, que jamais empregamos ou vimos escritas.

Por outro lado, a partir do *corpus* de textos, exatamente 120 palavras figuram entre as mais freqüentes: *primeiro* (mais de 20.000 ocorrências); *brasileiro* (mais de 9000 ocorrências); *janeiro*; *maneira*; *dinheiro* (mais de 5000 ocorrências); *companheiro*; *terceiro*; *estrangeiro*; *inteiro*; *verdadeiro*; *cadeira* (mais de 2000 ocorrências); *guerreiro*; *pereira*; *cavaleiro*; *madeira*; *financeiro*; *conselheiro*; *carreira*; *oliveira*; *bandeira* (mais de 1000 ocorrências); *fevereiro*; *fronteira* (mais de 900 ocorrências); *ligeiro*; *cheiro*; *ribeiro*; *beira*; *Vieira*; *mineiro* (mais de 800 ocorrências); *cavalheiro*; *herdeiro*; *enfermeiro*; *Ferreira* (mais de 700 ocorrências) etc. Tampouco o *corpus* de textos exclui problemas como seleção e tratamento de textos, erros de digitação, palavras obsessivamente usadas em alguns textos e totalmente ausente em outras (chama a atenção, por exemplo, palavras como *joaneira*, *soeiro*, *reposteiro*, *outeiro* e *serapilheira* figurarem entre as mais freqüentes).

Dessa forma, chega-se à conclusão que os *corpora* são insuficientes: o dicionário apresenta palavras demais e os textos, palavras de menos. A neologia, fonte da produtividade, evidentemente pode ter uma abonação num dicionário, mas arduamente figura em *corpora* de textos, menos ainda numa lista de palavras mais freqüentes, pois neologismos são, por definição, palavras raras.

3. Segmentação

As mais de três mil palavras obtidas no Houaiss e outras que apareceram nos textos trazem imediatamente a questão se são apenas palavras de fato com o sufixo *-eiro* ou se possuem apenas uma seqüência grafemática homófona. Para tais situações, o

Estruturalismo sempre se valeu do método da comutação, que trazia, ao lado de excelentes soluções, outros tantos problemas (MALKIEL 1964). O primeiro problema que se levanta diz respeito à natureza do significado dos sufixos. Não raro, os sufixos foram e são considerados como destituídos de significado, uma vez que há uma certa irregularidade na formação das palavras. De *sapato* deriva-se *sapateiro*, assim como de *sal* sai *saleiro*. Referencialmente nada haveria de comum entre uma coisa e outra, pois *sapato* é um objeto e *sapateiro* uma pessoa, e, apenas a partir do radical *sapat-* ou *sal-*, teríamos, com o auxílio do sufixo, uma associação obscura de atos ou fatos de alguma forma relacionados, ainda que vagamente. Entretanto, não é incomum o radical encontrar-se opaco. Apesar de *carpinteiro* e *marceneiro* terem radicais pouco claros, ninguém negaria que se trata de profissões. Há, por meio dos mesmos elementos, palavras como *carpintaria* e *marcenaria*, que justificariam a sua segmentação do ponto de vista dos métodos que privilegiam apenas o eixo sincrônico, mas que significam *carpint-* e *marcen-*? Indagando melhor: a que objetos ou idéias eles se referem? Também o sufixo de *chiqueiro* tem a mesma função do de *galinheiro*, sem que fique claro ao falante o que quer dizer o radical *chic-*: o eterno problema dos morfemas “*cranberry*” (BLOOMFIELD 1958). A partir desse raciocínio podemos também inferir que é o sufixo que carrega a maior quantidade de significado e isso parece ser confirmado fonologicamente pelo acento tônico (VILLALBA 2000).

A par disso, observa-se que uma palavra derivada tem três graus de significação. Uma palavra como *barbeiro* significando “profissão” reside na junção do radical da palavra-base *barba* e o mesmo *-eiro* indicador de profissão, no entanto já não se pode falar que *-eiro* signifique “inseto” quando *barbeiro* é o transmissor da doença de Chagas, tampouco significa “que (faz algo) mal”, quando *barbeiro* é o mau condutor de veículos. Há portanto que se diferenciar o *significado da base* (muitas vezes apagado com o tempo), o *significado do sufixo* (que, quando perde a prolificidade, também pode se tornar irreconhecível) e *significado total da palavra derivada* (que é, muitas vezes, imprevisível). De fato, *pedreiro* tem a raiz de *pedr-*, mas um pedreiro não trabalha apenas com pedras, assim como uma *leiteira* pode servir para ferver água e não leite. O momento da criação neológica do vocábulo, normalmente perdido, envolve também fatores sociolingüísticos de natureza variada. Uma vez criada a palavra, a atuação de metáforas é inevitável. Real empecilho para uma segmentação tranqüila oferecem os interfixos, estranhos signos com significante, mas sem significado. Entre os interfixos integram-se vogais e consoantes de ligação bem como sílabas ou seqüências de sons, pertencentes a sílabas distintas. De *chá* nasce *chaleira*, de *café*, *cafeteira*. A solução neoplatônica da alomorfia a partir de um morfema ideal costuma resolver esse problema e gerar outro, pois não há consenso se a alomorfia é do radical (*chal-*, *cafet-*) ou do sufixo (*-leir-*, *-teir-*). Outros exemplos: de *canavial* surge um interfixo *-avi-*; de *beberrão*, temos *-err-*; de *comilão*, *-il-*; de *dorminhoco*, *-inh-*; de *sabichão*, *-ich-*.

4. Frequência de uso

As palavras estão inseridas num discurso, que, por sua vez, tem, como dito, uma dimensão histórica e outra sociolingüística. Como no caso da produtividade, que discute a dimensão histórica (neológica) da derivação, o termo “frequência”, associado à dimensão sociolingüística, é normalmente usado numa indesejável polissemia: (a) Subjetivamente, uma palavra é “frequente” quando é conhecida ou supostamente conhecida por muitos ou todos os falantes de uma determinada língua. Isso se baseia preponderantemente na experiência individual: eu posso achar que *verdolengo* seja uma

palavra conhecida por qualquer pessoa, mas pequenos testes comprovariam se estou certo ou errado. Como é impossível rastrear os atos de fala, palavras efetivamente conhecidas por todos se impõem para a formação da *koiné* no caso de fragmentação dialetal (como ocorre atualmente com o alemão suíço), neutralizando traços “excessivamente marcados”, mas não é muito diferente nos casos de menor fragmentação dialetal, como no Português Brasileiro, apenas muito mais sutil. Em vez de “frequência”, essa característica subjetivamente detectada seria mais bem denominada como *usualidade*; (b) Objetivamente, uma palavra é “frequente” se, baseando-nos em *corpora*, constatarmos que aparece muitas vezes. Essa frequência absoluta é normalmente artificial e variável, mas tem a vantagem de não ser impressionista. Continuemos a chamá-la de *frequência*; (c) Também objetivamente, um *corpus* que separasse seus textos por meio de uma tipologia (hoje os critérios são ainda muito distintos), constataria que algumas palavras ocorrem em todos os tipos de texto numa dada língua (os artigos e as preposições, por exemplo), enquanto outras apareceriam frequentemente em muitos tipos, outras já estariam restritas a um determinado tipo de texto ou ainda outras seriam raras, aparecendo em alguns ou em todos os textos. Esse trânsito intertextual de algumas palavras poderia permitir que as denominássemos de “frequentes”, mas estamos diante de um outro comportamento, mais interessante, embora ainda indisponível, para as afirmações da Linguística. Preferimos chamar essa característica de *pantextualidade*. Igualmente etiquetas dadas a algumas palavras, problema comum dos lexicógrafos, que precisam chamá-las de *familiares*, *populares*, *chulas*, *regionais*, *poéticas*, *arcaicas* entre outros rótulos, estão associadas ao problema da tipologia textual, uma vez que se percebe, pela intuição sobre sua frequência usual, que essas palavras não têm frequência pantextual. O problema da frequência só será resolvido quando se houver uniformidade na questão tipológica dos textos. Sem dados seguros e objetivos sobre a frequência das palavras, dificilmente se poderá criar regras seguras para a derivação, antes se prosseguirá com o uso de regras tradicionais como a da primitividade do infinitivo já aventada.

5. Consciência do falante

Excluídos o estudo histórico e o estudo da pantextualidade acima apresentado, qualquer teste que invoque a consciência ou a competência do falante nativo se pauta única- e exclusivamente numa intuição, isto é, na usualidade das palavras. Abandonado o idealismo da norma culta preconizado pela Gramática Tradicional, o indivíduo observador (entenda-se, portanto, o lingüista, como representante de todos os falantes de uma língua, cultos ou incultos), por meio de sua introspecção e de seu julgamento, toma decisões sobre a gramaticalidade. É comum, no entanto, que pessoas cultas, flagradas usando formas que crêem não utilizar, insistam que não as utilizam, quando não ficam ofendidas por isso. Isso é comuníssimo quando brasileiros não realizam vogais pós-tônicas como [u] ou [i], pois crêem que só os portugueses apaguem vogais e admitam, no máximo, que “pessoas incultas” o façam numa “fala rápida”. A norma culta (e o modelo reacionário e ideal da escrita) retorna, portanto, instaurando um fenomenal paradoxo, em todos seus aspectos. Numa sociedade ocidental, em que contam a tradição platônica, aristotélica e o monoteísmo, isso não é de se espantar, mas o subjetivismo surpreende a vontade do objetivismo.

Das 120 palavras coletadas no *corpus* de textos, num teste informal, um informante de nível superior, doutor em Letras, afirmou não conseguir entender ou detectar a base em 38 casos (31,7%). Em alguns casos, a derivação assume posições

interessantes: afirmou que *gelo* → *geladeira*. De fato, a memória depende da frequência usual e, para o informante, *gelo* é mais associável, por estar mais rapidamente presente na memória, do que *gelar*, verbo defendido, por exemplo, por um linguista que postularia a derivação deverbal a partir do particípio *gelado* para esses casos em *-eiro*. Outras situações que causaram dúvida no informante foram as seguintes derivações: *engenho* → *engenheiro*, *cruz* → *cruzeiro*, *enfermo* → *enfermeiro*, *carta* → *carteira*, *barro* → *barreira*, *solto* → *solteiro*, *cocho* → *cocheiro*, *travessa* → *travesseiro*, *tesoura* → *tesoureiro*, *carne* → *carneiro*, *caldo* → *caldeira*, *trincar* → *trincheira*. De fato, outros informantes com certeza não teriam dúvidas em alguns (como *enfermo* → *enfermeiro*), ou não veriam casos de derivação em outros (como em *cruzeiro*), ou proporiam outras palavras simples em outros (talvez *coche* → *cocheiro*, *tesouro* → *tesoureiro*). Subjetivismos semelhantes se instauram em propostas teóricas, em que testes semelhantes são aplicados a si mesmo. Prosseguindo nosso teste, o informante encontrou ainda situações, que apresentava como palavras simples outras palavras derivadas ou ainda palavras que necessitariam muitas mudanças fonéticas para justificar a derivação: *companhia* → *companheiro*, *três* → *terceiro*, *mar* → *marinheiro*, *carpir* → *carpinteiro*, *brigar* → *brigadeiro*, *correr* → *carreira*, *cheirar* → *cheiro*, *rio* → *ribeiro*, *beirada* → *beira*, *cavalheirismo* → *cavalheiro*, *herança* → *herdeiro*, *palmito* → *palmeira*, *par* → *parceiro*, *ribeirão* → *ribeira*, *frei* → *freira*. Isso mostra que a palavra-base que vem para a consciência do falante nativo é a mais proeminente na memória, definida, entre outras coisas, com base em sua usualidade.

6. Diacronia

Do ponto de vista diacrônico, a produtividade de *-eiro* teria advindo do sufixo latino *-arius*. Surge o problema da continuidade diacrônica: como o sufixo sobreviveria *como o mesmo* dentro de dois sistemas distintos? Essa contradição eliminaria a distinção entre o sistema do português e o do latim, criando, aparentemente, um absurdo. No entanto, quando se fala do português arcaico e do português moderno, se vêem duas etapas da mesma e não duas etapas distintas. De qualquer forma, cria-se, introduzindo a diacronia, contudo, um novo critério, diferente do da consciência do falante para selecionar, entre as seqüências finais grafemáticas do tipo *-eiro*, quais seriam sufixos verdadeiros e quais seriam falsos. Seriam verdadeiras todas as palavras que viessem do *-arius* latino ou que fossem baseadas numa base portuguesa (em uso atual ou não) *+eiro*. Entre as mais de três mil palavras que sobraram do Houaiss, 80 delas não passam rigorosamente sob esse crivo: *cathedram* > *cadeira*, *materiam* > *madeira* etc. ou têm origem obscura (*sueira*, *queira*, *adeira* etc.), número que se pode considerar muito baixo (2,15%), mas não totalmente desprezível para problemas de Estilística. Das 38 ocorrências da indefinição da palavra-base pelo falante no teste, excluindo os 5 casos já citados de palavras raras, cujo significado total não foi compreendido, constatou-se que 13 casos tampouco se justificavam diacronicamente (*inteiro*, *cadeira*, *madeira*, *mosteiro*, *Vieira*, *feira*, *esteira*, *mosteiro*, *capoeira* e os dias da semana), um é problemático (*algibeira*), dois remontam a étimos hipotéticos (*cordeiro*, *derradeiro*), ao todo, 39,4 % da lista, outros 5 casos eram de estrangeirismos (*maneira*, *estrangeiro*, *bandeira*, *fronteira*, *ligeiro*) e os restantes 12 casos eram antigas palavras formadas no latim ou em português antigo (*janeiro*, *fevereiro*, *dinheiro*, *primeiro*, *caveira*, *ladeira*, *cachoeira* e sobrenomes: *Silveira*, *Moreira*, *Teixeira*, *Medeiros*, *Siqueira*). Interessante é observar que nos 55 casos em que o informante não teve problema para decidir qual era a palavra-base, 100% de fato eram formadas no português. As demais palavras, cuja

base não era nem muito opaca, nem muito transparente, eram solucionadas por meio do critério da usualidade, já exemplificadas acima.

Disso se conclui que palavras transparentes geram determinado grau de certeza na determinação/ reconstrução da base, certeza essa que é transferida para as bases mais ou menos opacas. Ao mesmo tempo, gera a certeza de que o sufixo não está presente nas bases totalmente opacas, embora isso não condiga com o percurso diacrônico do sufixo. Mais uma vez observamos que a existência do sufixo, numa perspectiva sincrônica, depende do reconhecimento arbitrário do radical: quanto mais opaco, tanto maior a certeza de que o sufixo não está ali. Paralelamente à biologia, as nadadeiras dos peixes, as asas dos morcegos e as patas de outros mamíferos são homólogas do ponto de vista evolutivo. Não é por que um ser vivo faz uso dela para nadar e outros para voar, caminhar ou pegar coisas, que não seja possível estabelecer uma evolução comum. A funcionalidade do sufixo, embora atrofiada em *janeiro*, aponta percurso idêntico na linha do tempo ao de *sapateiro*. Por outro lado, as asas dos insetos e das aves não têm um ancestral comum, da mesma forma que *madeira* e *sapateiro* não têm: uma coincidência, no entanto, que é importante para a criação do *molde fônico* *-eiro*, este sim, importante sincronicamente, pois está diretamente ligado à produtividade neológica, que não tem necessariamente vinculações diacrônicas. Assim, *macaxeira* não tem nada a ver com *-arius*, mas por uma coincidência, a palavra tupi *makaxéra* tinha semelhança com palavras derivadas que indicam “nomes de planta”, como *roseira*, *palmeira*, *mangueira* etc., tanto que graficamente um *-i-* etimológico (raramente pronunciado) foi acrescentado ao ser escrita. Um verdadeiro caso de “evolução convergente” ou de “mimetismo de Bates” lingüístico. A diacronia também enfrenta seus problemas: a reconstrução de sua forma e/ou de sua frequência pantextual é apenas uma delas. Um bom étimo, além disso, deve estar conforme as transformações esperadas e deve ter datação, que nos estudos do português, é ainda incipiente, malgrado os enormes esforços de Antônio G. Cunha, apresentados parcialmente no dicionário Houaiss. Uma data, contudo, raramente é segura: não se pode ignorar o fato de que a palavra surja na fala antes (não raro, séculos antes) de aparecer testemunhada na escrita, mas, com certeza, uma data exclui a ilusão de que ela é antiga ou recente, baseada na intuição.

7. Classificação e genealogia dos significados

A classificação semântica é um trabalho mais árduo a se estabelecer. Há dois tipos de classificação de sufixo: uma que estabelece grandes categorias abstratas (tremendamente criticadas por autores como WIERZBICKA, 1996) que diria que em *mangueira* temos um “locativo” tanto quanto em *galinheiro* e outra, mais concreta, associada a paráfrases: “x]_{eiro/a} é uma árvore que produz x]_n”. A extensa polissemia de *-eiro* se explica pelo menos em cinco etapas a serem confirmadas pelo extenso trabalho do Grupo de Morfologia Histórica (VIARO 2003, *prelo*): (a) inicialmente x]_{eiro/a} era um sufixo denominativo e relacional, formador de adjetivos, que, posteriormente, adquiriram função substantiva. Herdara essa função do *-arius* latino. Como heranças dessa fase temos o caso de *dinheiro*, *primeiro*, *janeiro* entre outros, não-reconhecíveis pelo informante como portadoras de significado. Contudo, esse uso não perdeu sua produtividade; (b) x]_{eiro/a} passa a indicar substantivos e a relação estabelecida é a profissão já no período tardio do latim e início da Idade Média, tendo aumentado sobremaneira sua produtividade no século XVI, quando começou a concorrer outros agentivos, sobretudo com *-ista* de origem grega. Paralelamente, x]_{eiro/a} se vinculou ao

nome de árvores frutíferas, algumas já no latim vulgar, outras presentes apenas no latim ibérico e sobretudo no português após o século XVI. Outros nomes de planta se vincularam (*roseira*, *espinheiro*), independente de estar associado a árvores frutíferas. Mais tardio (após séc.XVII) é o uso do *-eiro* gentílico (*brasileiro*, *campineiro*), menos prolífico, por causa dos concorrentes *-ense* e *-ano*; (c) Por meio de uma metaforização, o $x]_{eiro/a}$ de profissões passa a atribuir funções de objetos (*saleiro*, *charuteira*, *geladeira* etc.). Dessa função de guardar, presente em muitos objetos, passa-se facilmente para locais (*galinheiro*); (d) Por meio de uma metonimização, o $x]_{eiro/a}$ enfatizaria o conteúdo e, por conseguinte, a quantidade do que está sendo guardado, gerando um valor coletivo (*formigueiro*, *vespeiro*, *cabeleira*, *montoeira*). Por uma generalização quantidade migra para intensidade (HEINE & KUTEVA, 2002) e passa a ser um intensificador (*nevoeiro*, *poeira*); (e) Por fim, o $x]_{eiro/a}$ adquire um traço multicausal de pejoratividade (ULLMANN 1973) que é associável tanto a substantivos concretos (*pianeiro*) quanto a abstratos (*bandalheira*, *ladroeira*, *choradeira*, *asneira*), sobretudo a doenças e outras indisposições (*leseira*, *canseira*, *gagueira*, *cegueira*, *pasmaceira*).

8. Conclusões

O presente trabalho não pretendeu criar nova teoria ou modelo, mas chamar a atenção para que o método intuitivo da competência do falante nativo é por demais impreciso para auxiliar em problemas básicos como o da segmentação e separação dos falsos prefixos de seqüências fônicas homófonas. Estudos que envolvam diacronia e a tipologia textual, sobre a qual se possa falar de uma frequência pantextual, são imprescindíveis, sem os quais se poderá cair em uma infinidade de soluções *ad hoc*. Um bom exemplo disso ocorre com a palavra *chuveiro*, datada do séc. XV, como “forte pancada de chuva” (portanto com valor intensificador, a partir de *chuva*). Como esse sentido é obsoleto hoje em dia, prevalecendo uma acepção secundária, formada a partir de uma metáfora sobre o significado total da palavra desde o séc. XIX (talvez decalque do *shower* inglês), pode-se pensar que *-eiro* signifique vagamente algo como “objeto” e colocá-lo ao lado de *saleiro*, *assadeira* etc. No entanto não há paráfrase possível com sua base *chuv(a)*, como ocorre com esses objetos (“objeto em que se guarda $x]_n$ ” ou “objeto em que se $x]_v$ ”) e mesmo que se idealize alguma (algo como “objeto que se parece com $x]_n$ ”), não será produtiva. Esse é um dilema que surge quando a base não é opaca, mas a derivação sufixal se mostra um tanto estranha (comportamento que já valeu aos sufixos rótulos como “sem significado” ou “assistemáticos”). Não é preciso insistir mais sobre quão indesejáveis são essas soluções *ad hoc* que tentam excluir dados diacrônicos, rejeitando significados que não são freqüentes segundo a usualidade. Longe de querermos levantar bandeiras chauvinistas em prol da importância dos estudos diacrônicos, pensamos que, nesse caso particular, no mínimo, seria prudente classificar *chuveiro* como uma palavra portadora de um *-eiro* intensificador, embora o significado total, atual e usual da palavra não seja esse. Cumpre lembrar que nenhuma ciência se pauta na intuição pura e simplesmente e assim nasceu a nova Astronomia. A não agirmos assim, num novo paralelo com a taxonomia biológica, estaríamos afirmando que algo como “um urso-panda não é um carnívoro porque não come carne”. Intuitivamente, de posse da informação que esse animal só coma bambu, diríamos que está correto e que ele é um herbívoro. Para o zoólogo, todavia, não há uma ordem especial na classe dos mamíferos para os animais herbívoros: isso é uma questão etológica e não taxonômica. O nome “carnívoro” é um rótulo antigo, atribuível a determinados animais; não é intuitivo, mas convencional, com uma definição muito

rigorosa, pautada na reprodução, na dentição do animal, numa ancestralidade comum e não na observação direta ou na intuição com base no “uso coloquial” do termo etc. Por outro lado, o fato de os pandas não comerem carne se deveu a acidentes de percurso evolutivo. Se o rótulo parece “mal empregado” para um falante qualquer, isso se deve talvez ao desconhecimento da existência do panda ao se cunhar o termo. Mas é interessante observar que o termo não foi remendado por causa dessa descoberta. Os biólogos se valem do *princípio da antigüidade terminológica*, que os livra do encargo da onisciência e os poupa de inúmeros dissabores terminológicos presenciados na Lingüística. Para eles, um rótulo é um significante qualquer com uma definição rigorosa e não precisa ter relação direta com a referência: equívocos interpretativos iniciais não são suficientemente fortes para refazerem-se os rótulos. Nós, lingüistas, embora conscientíssimos da arbitrariedade do signo, contudo, às vezes nos esquecemos disso e inflacionamos esterilmente nossas terminologias que precisam, “por fetichismo”, ter relação com a realidade (OGDEN & RICHARDS 1976) ou trabalhamos com classes muito vagas não definidas suficientemente.

Referências bibliográficas:

- BAKHTIN, Mikhail. *Marxismo e filosofia da linguagem*. São Paulo: Hucitec, 1988.
- BLOOMFIELD. *Language*. London: George Allen & Unwin, 1958.
- HEINE, Bernd; KUTEVA, Tania. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press, 2002.
- HOUAISS, Antônio; VILLAR, Mauro (org.) *Dicionário Houaiss da língua portuguesa*. Rio de Janeiro: Objetiva, 2001.
- MALKIEL, Yakov. Generic analysis in word formation, 1964. In: SEBEOK, Thomas A. (ed.) *Current trends in linguistics*. Paris: Mouton, 1970. p. 305-364.
- OGDEN, C. K.; RICHARDS, I. A. *O significado de significado*. Rio de Janeiro: Zahar, 1976.
- RIO-TORTO, Graça M. *Morfologia derivacional: teoria e aplicação ao português*. Porto: Porto Ed., 1998.
- _____. (org) *Verbos e nomes em português*. Lisboa: Almedina, 2004.
- ULLMANN, Stephan. *Semântica: uma introdução à ciência do significado*. Lisboa: Calouste Gulbenkian, 1973.
- VIARO, Mário E. *Os sufixos portugueses numa visão diacrônica*. (no prelo)
- _____. Para um estudo de semântica sincrônica dos sufixos derivacionais em português do séc. XIII. *Grupo de Estudos Lingüísticos do Estado de São Paulo*. Taubaté: Unitau, 2003. CD-ROM (com. 95).
- VILLALBA, Alina. *Estruturas morfológicas do português: unidades e hierarquias nas palavras do português*. Lisboa: Calouste Gulbenkian/ Fundação para a Ciência e Tecnologia, 2000.
- WIERZBICKA, Anna. *Semantics: primes and universals*. Oxford/New York: Oxford University Press, 1996.