



VIARO, Mário Eduardo. **A morfologia histórica e os estudos etimológicos da língua portuguesa.** *Revista Diadorim / Revista de Estudos Linguísticos e Literários do Programa de Pós-Graduação em Letras Vernáculas da Universidade Federal do Rio de Janeiro.* Número especial 2013. [<http://www.revistadiadorim.letas.ufrj.br>]

A MORFOLOGIA HISTÓRICA E OS ESTUDOS ETIMOLÓGICOS DA LÍNGUA PORTUGUESA

Mário Eduardo Viaro^{10*}

RESUMO

A partir do momento em que a Linguística passou a dedicar-se à segmentação de palavras, já no final do século XVIII, uma nova perspectiva se abriu com o estudo especializado da Morfologia, a qual se encontrava mesclada à Lexicologia, desde a sistematização das classes de palavras pela Gramática ocidental no período alexandrino. As hipóteses acerca da origem de sufixos, prefixos e desinências, no entanto, durante o período do desenvolvimento da Gramática Histórico-comparativa, exerceram um papel importante na compreensão de certos universalismos, confirmados apenas recentemente pela Linguística Cognitiva. No entanto, a descrição do desenvolvimento diacrônico do componente semântico dos mesmos elementos de composição, apesar de adequada, nem sempre era rigorosa com relação às sincronias pretéritas em que os mesmos elementos estudados se integravam. Paralelamente a isso, o estudo etimológico continuava a enfatizar apenas o léxico.

Em período de fecunda discussão sobre os mecanismos de transmissão etimológica, tipologia linguística, variação diatópica e estilística, a criação de um novo método etimológico foi interrompido pelas convulsões sociais do século XX.

Com o advento da Linguística Moderna e o amadurecimento teórico do Estruturalismo e do Gerativismo nas descrições e explicações teóricas para a sincronia atual, originou-se uma teorização do fenômeno diacrônico, aplicável ao setor morfológico. Retomado o estudo morfológico de viés diacrônico, após o maior conhecimento das relações intersistêmicas, somou-se ao conhecimento das transformações formais e semânticas também a necessidade de coerência intersistêmica nos modelos, mediante um método etimológico mais adequado aos dias de hoje.

PALAVRAS-CHAVE: Diacronia; Etimologia; Língua Portuguesa; Morfologia; Teoria morfológica.

10. Professor livre-docente da área de Filologia e Língua Portuguesa, do Departamento de Letras Clássicas e Vernáculas, da Faculdade de Filosofia, Letras e Ciência Humanas da Universidade de São Paulo, coordenador do GMHP (www.usp.br/gmhp) e do NEHiLP (www.usp.br/nehilp). E-mail: maeviaro@usp.br.

ABSTRACT

With the advent of Modern Linguistics and the theoretical maturing of the structuralist and generative theory in descriptions and explanations for the current synchrony, a theorization of the diachronic phenomenon was applied to the morphological sector. After the greatest knowledge of the interlinguistic relationships, the morphological study of diachronic bias gained new dimensions: in addition to the knowledge of the formal and semantic transformations, also the need of consistency in the intersystem models has to be taken into account.

KEYWORDS: Diachronic studies; Etymology; Portuguese; Morphology; Morphological Theory.

Ciência e cientificidade

Qual é a essência da cientificidade que se procura para um formalismo de um estudo tão complexo quanto o da Linguística Diacrônica? Se pudermos responder a essa pergunta, resolveremos de uma só vez os problemas que enfrentam a Etimologia científica e todas as suas aplicações, como é o caso da Morfologia Histórica. Já dizia Hegenberg (1969: 209), com referência às por ele chamadas Ciências Sociais, entre as quais insere a Linguística:

é raro o acôrdo entre investigadores quanto a fatos, explicações e métodos em suas áreas de pesquisas. Entre especialistas, manifestam-se profundas divergências que os separam não apenas em questões metodológicas, mas em questões substantivas (...) O que importa, realmente, é reconhecer que embora as generalizações a propósito de fenômenos sociais, tais como se apresentam, sejam de restrito alcance, formulação nem sempre satisfatória, aceitabilidade fatural dependente de grande número de restrições e exceções, elas não diferem muito das generalizações que têm sido obtidas no campo das ciências naturais.

Parece-me que os pressupostos sobre os quais se deve alicerçar um estudo não-especulativo da história da língua deve partir de noções compartilhadas por todos. Da obviedade dos pressupostos nasceriam saltos de raciocínio que caracterizam as hipóteses de uma teoria. No entanto, esse elemento óbvio resvala-se na tradição de pensamento corrente, contra a qual dificilmente resistimos. Este texto, portanto, tenta discutir no nível dos pressupostos assumidos a partir de alguns anos de pesquisa com dados pelo Grupo de Morfologia Histórica (GMHP) e, mais recentemente, pelo Núcleo de pesquisa em Etimologia e História da Língua Portuguesa (NEHiLP), ambos da Universidade de São Paulo.]

A tradição filosófica ocidental se mesclou à religiosa ao final do Império Romano. Conhecê-la talvez seja crucial para repensar a Linguística do século XXI. Essa afirmação se justifica pois a

Gramática Tradicional caminhou paralela à Filosofia e dela pouco se distanciou. O caráter dedutivo e neoplatônico nunca foi abandonado pelas investigações da Linguagem, exceto talvez nos estudos de Edótica e de Dialetoлогия, que retomaram a indução. Mesmo a Etimologia, que deveria ser puramente empírica, só o foi em raros momentos, pois desde Platão até hoje é revestida de uma essência estranhamente dedutiva (VIARO, 2011). Durante o tempo de Francis Bacon, vimos a indução ser reivindicada, ao desmembrar-se de uma tradição filosófica ao ponto de negá-la. Depois de Locke, não faltaram as reações contrárias nos debates promovidos por Leibniz e Berkeley. Teríamos de esperar até Hume para a busca de uma síntese. Um discurso espetacularmente conciliador de correntes em Linguística ocorreu somente na época dos Neogramáticos do século XIX. Inversamente, a reação pró-indução vinda de Schuchardt serviu para colocar em cheque alguns pressupostos, que começaram a ser desenvolvidos pelos autores do movimento *Wörter und Sachen*, no início do século XX. Esses novos parâmetros, não plenamente desenvolvidos, acabaram por ser esquecidos por conta das mudanças promovidas pelas guerras mundiais e pela renovação dedutiva promovida pelo Estruturalismo e pelo Gerativismo. Novo alento indutivo se anteverá, no final do mesmo século, em alguns autores da Sociolinguística e, posteriormente, pela renovação da Crioulística e por universalistas como Heine & Kuteva (2002).

À guisa de pressupostos

1. A língua-ferramenta

Se a essência da realidade parece ser um tema da Filosofia, havendo mesmo quem a negue, verdade é que, de uma perspectiva que chamaremos *ingênua*, todos os falantes têm consciência de que há algum tipo de distinção entre língua e realidade. Isso é válido mesmo para culturas que privilegiem a magia como papel importante de suas vidas. Até o mais isolado dos povos da terra entende, de modo intuitivo, que há palavras para representar coisas, reais ou não, e regras para que essas palavras sejam utilizadas. Entre as palavras escondem-se abstrações, operações lógicas e elementos que remetem também a algo para além do mero raciocínio. Apesar de nem sempre ser constante e nítido na mente dos falantes, há consciência de que palavras não se confundem (ou não devem confundir-se sempre) com os seres nomeados. Palavras geram coisas e coisas produzem palavras. Palavras e coisas não se confundem, portanto, pelo menos não deveriam confundir-se num nível ideal. Uma teoria deveria desvencilhá-los, mas essa talvez seja a tarefa mais difícil de toda a Humanidade. Os tabuísmos, por exemplo, mostram-nos que essa relação entre significado e referente não é algo tão simples para o falante quanto o signo de duas faces adotado pelo Estruturalismo.

Além das *palavras* e das *coisas*, falamos de *línguas* como entidades óbvias. Mas em que tipo de referência se insere uma língua? Seria uma coisa do mundo real, uma abstração ou uma ficção? Desde o primeiro contato intertribal da história da Humanidade em tempos imemoriais, sabemos que algumas coisas têm mais de um nome dependendo da comunidade. Os gregos chamavam qualquer expressão não-helênica de bárbara. Podemos concluir que a gênese da noção de língua não provém da observação de nosso *êxito comunicativo* mas da consciência da *dificuldade* ou mesmo do *impedimento de compreensão* da língua alheia.

O contato interpessoal permitiu também observarmos que, mesmo quando essa intercompreensão se efetua de modo razoável, há variação na expressão e no significado de algumas palavras. Mais profunda é a atribuição interpessoal dos significados, pautada em valores e experiências. A mesma língua tem uma natureza múltipla, objetiva e subjetivamente. Objetivamente, basta que observemos variações regionais, diferenças de idade, de sexo, da classe social, da religião entre os falantes. Subjetivamente, basta vermos os mal-entendidos que permeiam nossas vidas, a ponto de acreditarem alguns, por vezes, que quanto maior o contato, maior o abismo entre as pessoas.

A essa língua, flagrada no seu aspecto vivo, chamaremos de *língua-ferramenta*. Serve-nos como instrumento de comunicação, no entanto, é de pouca utilidade como metalinguagem de qualquer ciência que tome a própria língua como objeto de estudo. Muitos, na ciência e na filosofia, já perceberam que abundam ambiguidades e até mesmo contradições semânticas no conteúdo de seus signos.

Essa superabundância semântica raramente serve para a *descrição* e para *explicação* de fatos. No nível puramente teórico, gera interpretações desnecessárias, praticamente desconhecidas pelas ciências exatas. Paradoxalmente, não contribui para nenhum acúmulo de conhecimento a exploração da potencialidade máxima de uma palavra, tal como utilizada de modo mais ou menos consciente por filósofos como Hegel, Nietzsche, Wittgenstein, Heidegger e outros, os quais multiplicam o sentido dos termos e buscam sua mensagem numa nebulosa semântica artificialmente construída. Inócua ou prejudicial, essa linguagem, mesmo quando interpretada como inútil para as postulações linguísticas, infiltra-se sorrateiramente sob a forma de ceticismo ou niilismo (SOKAL & BRICMONT, 1998). Na verdade, há tempos, devido à imensidão do estudo das palavras, confunde-se a evidente dificuldade de destrinchar todos seus aspectos (que requereria método, trabalho conjunto e organização) com a impossibilidade de revelar sua essência. Uma solução alternativa costumeira é o *recorte*, caracterizado em Linguística pela figura do *corpus*, a qual evidencia o problema de raramente apontar para um norte, dado o cipoal de *métodos* atualmente utilizados, os quais são confundidos com *teorias*.

Para resolver isso, uma linguagem simbólica parece ser mais eficiente, embora não prescindida também de uma metalinguagem e, portanto, de pressupostos alicerçados epistemologicamente numa realidade social, caindo, no final das contas, no mesmo problema da língua-ferramenta.

2. A língua ideal

Dessa língua-ferramenta jamais teria nascido o conceito de *língua ideal* sobre a qual se pautam as gramáticas tradicionais. Foi necessário um salto tecnológico para isso, que foi a *escrita*. As grafias de motivação fonológica talvez tenham sido o ponto de partida da sua formação, o qual foi reforçado ainda mais pelo caminho paralelo de diversas tradições filosóficas, religiosas e gramaticais no mundo.

Nesse sentido, o que se interpreta como a “língua real” não é aquela que sempre foi usada na comunicação (à qual chamamos *língua-ferramenta*), mas uma convenção adotada pela maioria, portanto um *recorte*, não raramente artificial. A origem dessa convenção se fundamenta em tentativas de imitação, motivadas direta ou indiretamente por relações de poder (uma personagem histórica, uma região mais influente ou idealizada ou mesmo uma imposição). Algumas vezes a língua ideal se distanciou muito da língua falada, outras vezes ambas se misturaram em maior ou menor grau, gerando um *dégradé* muitíssimo complexo. Na verdade, o que chamamos aqui de *ideal* tem a ver com um componente essencial de construtos teóricos desde o movimento eleata até o surgimento do Platonismo. Nesse sentido, a idealização é apenas o resultado de modelos artificiais para a apreensão e regulamentação da realidade fugidia caracterizada pelo *pánta rheî* heraclítico que, por algum motivo ainda não suficientemente compreendido à luz da psicologia, deva ser substituída e evitada (KIRK, RAVEN & SCHOFIELD, 1983).

Essa idealização deve, portanto, ser diferenciada daquele elemento da cognição humana básico e instintivo, do qual dependem inúmeras atividades mentais. A criação de um sistema de unidades mentais conhecidas como “fonemas”, por exemplo, depende de uma abstração de frequências sonoras na mente da criança antes mesmo da sua expressão oral em sua aquisição da língua (cf. DEL RÉ et alii 2006). A tendência à construção dos significados genéricos a partir da experiência individual com um número limitado de seres é algo inato. A imensa maioria das pessoas lida com essa realidade de modo extremamente caricatural e não têm uma percepção prodigiosa (nem uma memória) semelhante, por exemplo, à de Stephen Wiltshire (SACKS, 1995: 188-243). Chamamos a essa idealização natural e caricatural, ou seja, essa capacidade inata de lidar com o *pánta rheî* da realidade de modo socialmente qualificado como adequado, sem qualquer tipo aprendido, de *sematogênese*.

Diferenciar os seres ideais criados naturalmente pela sematogênese que normalmente impelem a criança à linguagem dos seres ideais criados por teorias e por outras convenções sociais mais complexas parece crucial para qualquer teoria que envolva o estudo da Cognição e da Linguagem. Mas isso nem sempre se faz de modo claro. Por exemplo, modelos que mencionam o problema, conhecidos como *motherese* e *UG* não se confundem com a sematogênese, que é um fato. É mais fácil refutar uma teoria do que negar um fato, a menos que se parta para os padrões berkeleyanos de argumentação filosófica (AYERS, 1975).

A distinção entre *fato* e *teoria* nem sempre é óbvia na teorização linguística. Supomos que uma língua é “um conjunto de signos”, os quais têm, por conseguinte, uma estrutura essencialmente idêntica. Trata-se de algo que deduzimos da leitura do *Cours de linguistique générale*, por exemplo, mas, aparentemente, essa interpretação se reveste de um dogmatismo. Há signos e signos e não há por que imaginar que suas estruturas sejam sempre idênticas. Há, por exemplo, signos cuja existência independe de uma estrutura linguística. Parece ser o caso da *negação de predicados*, que existe em todas as línguas do mundo.

Obviamente, para *expressar* a negação (e, portanto, valer-se de uma estrutura) muitas línguas dispõem de uma palavra isolada (um advérbio de negação como o *não* do português ou o *nicht* do alemão), outras dispõem de várias palavras, pois possuem uma espécie de concordância com o tempo verbal (como o árabe clássico, que tem *lâ* para o presente, *mâ* com verbos no passado e *lan* no futuro), outras, ainda, expressam a negação por meio de um afixo (como as diversas formas de expressá-la em japonês). Seja como for, a negação é algo universal.

Mas seria essa afirmação algo que denota alguma cegueira do pensamento eurocêntrico? Se línguas ocidentais costumam ter uma única negação de predicado, então somente a partir desse ponto de vista enviesado, poderíamos afirmar que a pluralidade é uma decorrência da unidade ideal? Sim, se confundirmos o nosso *não* ocidental, elemento de uma classe de palavras conhecida como advérbio, à *não*, função, portanto, significado “negação de predicado”.

Mas onde estará esse significado universal? No signo, encarcerado num sistema ou já na mente prelinguística? Aparentemente ninguém ensina a uma criança a noção de “negação de predicado” de forma analítica, pois ela parece reconhecê-la facilmente, muito antes de aprender a falar de modo bastante satisfatório. Tampouco essa noção pode vir de maneira indutiva. Um conceito universal como esse preexiste, como um expediente mental, de complexa formação dentro da evolução das espécies, tal como os instintos (LORENZ, 1978).

Não podemos misturar, portanto, conceitos como *negação de predicados* com um conceito qualquer (tal como *cadeira*, *amor*, *interstício*, *língua*). Nem podemos confundir essa negação de predicados (que nasce, aparentemente, de contextos volitivos “não quero” para caminhar rapidamente para contextos definitórios “não é” e descritivos “não vou”) com outros tipos de negação, como a *negação imperativa* (“não faça isso!”) ou com mecanismos mais sofisticados, como a *negação nominal* (“não-violência”, “in-satisfação”, “in-comum”).

Também é difícil imaginar como não-universais palavras que denotam partes do corpo humano (mesmo que digamos que uma língua como o português tenha uma palavra para *mão* e outra para *braço*, enquanto no russo, *ruka* equivale às duas coisas indiferenciadamente). Hoje parece evidente afirmarmos que a crença na universalidade dos signos se trata de algo ingênuo. Com frequência apare-

ce, porém, como pressuposto teórico, nem sempre declarado, antes do Estruturalismo. Por outro lado, a consciência gerada pelos estruturalistas não pode ser generalizada filosoficamente para seu outro extremo sem revestir-se de dogmatismos (por exemplo, em afirmações que *nenhum signo é traduzível perfeitamente de uma língua para outra pois reflete apenas um recorte social exclusivo de uma realidade intangível*). Usando esse mesmo tipo de argumento simplista, nega-se a possibilidade de haver dois sinônimos perfeitos em uma língua. É preciso retornarmos ao *hèn oída hóti oudèn oída* de Sócrates e reavaliarmos se essas certezas passam de fato por um teste rigoroso.

Em suma, um pressuposto aceitável para modelo deve levar em conta os dados e não as projeções que lançamos sobre os dados. Devem levar em conta a língua-ferramenta e não a língua ideal. Isso parece um problema recorrente da visão dedutiva em Linguística Moderna, que abandonou os dados para ir diretamente à reflexão.

Tradição gramatical e desenvolvimento científico

Mas saberemos se analisamos razoavelmente os dados? Há um modo correto ou isso seria uma pretensão? No tocante à Morfologia, essa não é uma pergunta ociosa. Por exemplo, fala-se de *deverbais* para palavras como *mandão* e *fechadura*. No entanto, podemos estabelecer claramente que *mandão* vem de *mandar*, já em *fechadura* temos por base a palavra *fechado*, particípio de *fechar*. Se as formas derivadas provêm das simples e é uma abstração usar o infinitivo para representar todo o verbo (assim como usamos o masculino singular para representar todo o paradigma adjetival), não é estranho que no caso de *fechadura* tenhamos claramente uma flexão real (o particípio) e não uma flexão simbólica (como o infinitivo)? Além disso, não é um acidente histórico dizer que o particípio é uma flexão verbal, visto que isso nem sempre foi algo inquestionável, como revela a própria historiografia? Em Dionísio Trácio, particípios e verbos são classes distintas e há razões morfológicas para separá-las, uma vez que as flexões dos demais tempos verbais seguem um padrão amalgamado de número e pessoa e não um padrão discreto de gênero e número.

Por fim, não é fácil determinar se uma forma como *alfinetada* é um denominal (de *alfinete*), um deverbais (de *alfinetar*) ou uma mudança semântica provocada pela flexão feminina do particípio. Que fazer nesses casos? Adotar uma postura *ad hoc* diminui o caráter científico das premissas subsequentes de uma argumentação. Não parece um bom começo.

A questão da *tradição gramatical* interfere mais nos julgamentos da Linguística do que se pode imaginar. Por exemplo, sob a perspectiva estruturalista, algumas classes de palavras tradicionais, como as interjeições, até mesmo entram em contradição com a definição de palavra apresentada pela própria teoria. Se o significante de um signo deve compor-se de fonemas previstos num sistema e se toda língua tem um repertório fechado de fonemas, definido por meio de oposições, então veremos que

algumas interjeições usuais do português, apesar de terem significado, não podem ser consideradas signos no sentido utilizado pela Linguística, pois são produzidas adequadamente mediante a utilização de sons não previstos, cliques, tons etc. Em suma, se se valem de traços que não são de modo algum integrantes do sistema de oposições dos fonemas que integram língua portuguesa, como serão um signo linguístico? Uma negação na definição já se encontra na base, impedindo as premissas seguintes. Aparentemente, para resolver esse problema, seria necessário aceitar que os signos utilizados pela língua *podem ou não* seguir o sistema fonológico. Mais que isso: que esses *signos* presentes nos sintagmas e paradigmas de uma língua não são somente *símbolos* linguísticos, como a grande maioria, mas às vezes também *ícones* e *índices*, pois nada garante que as interjeições sejam uma classe coesa e, de fato, algumas são verdadeiras palavras enquanto outras são se aproximam das onomatopeias, já outras não são nem uma coisa nem outra. E, de fato, alguns pseudomorfemas perfeitamente depreensíveis em segmentações (como as vogais temáticas) parecem apenas apontar para um paradigma e não significar algo, como se espera de um símbolo linguístico.

Desmontando-se, por meio das observações, as tradições milenares da Gramática gera-se um resultado pouco sistemático ao final e isso é exageradamente sentido como perigoso para a Linguística, que anseia por ser uma ciência rigorosa. No entanto, desde o tempo de Francis Bacon observa-se que a tentação de um dedutivismo é uma ansiedade que pode ser reprimida. A remontagem dos dados inicialmente descosidos é paulatina, mas não impossível para o pensamento coerente e persistente. Toda ciência se constrói paulatinamente e não somente pelo surgimento dos *insights* dos gênios. Buscar corroborações para teses inicialmente absurdas, ignorando exceções evidentes, aparentemente é um dos passos do êxito do pensamento dedutivo (como, por exemplo, as leis de Newton, surgidas em plena época de desprestígio do dedutivismo, revitalizando-o). Crê-se que haja sempre variáveis que explique as exceções. Obviamente as haverá se a teoria for realmente bem construída. Mas apenas gênios chegaram ao ponto de “imaginar” do nada uma teoria dedutiva que funciona bem, a imensa maioria anônimos. A figura do gênio foi muito apreciada no Romantismo ainda não é uma meta de todo abolida na pesquisa acadêmica. Hoje, aparentemente, só o trabalho de equipes coordenadas podem apresentar resultados similares aos das grandes intuições. A indução é peça-chave para as grandes soluções dedutivas e faltam-nas em Linguística, malgrado a passionalidade das correntes.

O julgamento dos dados

A língua ideal, tal como apresentada acima, é construída pelo poder local ou pelas gramáticas. Utilizada nas teorias linguísticas, é apenas uma idealização arbitrária, criada sobre a reflexão metalinguística e a serviço de um determinado recorte teórico com fins quer à normatividade, quer à explicação. Igualmente, podemos falar que conceitos como *falante-ouvinte ideal*, *de comunidade homogênea*, *de conhe-*

cimento total da língua materna, de capacidade ilimitada de atenção e memória (tais como apresentado em Chomsky 1965:3) são construtos a serviço de um modelo dedutivo.

Em Química, por exemplo, não há um hidrogênio ideal, embora podemos falar que haja três tipos de hidrogênio, isto é seus isótopos (prótio, deutério e trítio), mas todos são considerados hidrogênio porque por definição é o número de prótons que diz de qual elemento estamos falando. Alguém por capricho poderia questionar o caráter definidor dos prótons e dizer que cada isótopo é um elemento distinto (e não faltariam argumentos para isso), transformando a Tabela Periódica de Mendeleev, com pouco mais de 100 elementos, em algo bem mais complexo, com mais de 3000 elementos, mas há perda substantiva de elegância no modelo nessa multiplicação desnecessária. O caráter ideal da definição baseada no número de prótons é melhor do que, por exemplo, uma definição que levaria em conta o número de núcleons (elementos no núcleo atômico, que incluiria prótons e nêutrons). Também o critério da abundância não é o melhor. O deutério, poderia, usando o critério da abundância poderia ser pensado como o “hidrogênio ideal” para quem imaginasse que no núcleo atômico deva existir sempre o mesmo número de prótons e nêutrons (o que é evidentemente falso), no entanto, o prótio representa 99% de todo o hidrogênio existente e em seu núcleo apenas há um próton e nenhum nêutron.

Da mesma forma, não há palavra ideal, uma vez retirados os elementos “que não deviam estar ali” ou que “não estão frequentemente ali”. A verdade científica não é obtida por meio de médias, pois mesmo os desvios para mais ou para menos a partir de uma expectativa devem ter uma explicação no próprio sistema que gerou a teoria, para não se recorrer ao recurso *ex machina*. Negar a realidade tem sido um critério muito aplicado, desde o neoplatonismo e extensamente aceito em Teologia, mas não é um recurso razoável para o pensamento científico. Mesmo ciências que se valem da “verdade média”, apressam-se para entender os desvios. Estamos longe disso.

Paradigma e analogia

Distinguindo adequadamente a língua-ferramenta da língua ideal, torna-se fácil compreender a língua como *conjunto de paradigmas*, essencial para a Morfologia. Para além da sua representação visual, um paradigma, do ponto de vista cognitivo, nada mais é que uma rede associativa (ou subconjunto), cujos elementos são afins quer pelo significado, quer pelo significante. Tradicionalmente, em Morfologia, o significante é o mais importante para o estabelecimento de um paradigma, no entanto, a semântica não é abandonada, de modo que se pode falar de alomorfia ou mesmo de supletivismo morfemático, dependendo da distância entre um morfema específico e o ‘morfema ideal’. Por outro lado, a alossemia de uma mesma forma pode gerar grandes discussões em torno da existência de uma polissemia ou homonímia de sufixos, por exemplo, de modo que esses conceitos requerem uma teoria, a qual, por sua

vez, requer unidades operacionais, uma lógica e, conseqüentemente, pressupostos que sustentem seus corolários, teoremas e outras afirmações ousadas.

A capacidade de criar paradigmas é algo também inato à linguagem (*paradigmogênese*) e bastante diversificada nas línguas, porém não se deve confundir com a artificialidade das soluções teóricas. Um paradigma natural apenas lembra vagamente um paradigma artificial criado por um construto teórico (por exemplo, uma classe de palavras em uma gramática qualquer), embora nesse caso, a intuição funcione, aparentemente, melhor do que com o próprio conceito de signo. O exemplo dos participípios acima mencionado mostra que paradigmas naturais nem sempre são perfeitamente equivalentes a paradigmas obtidos teoricamente e devem ser distinguidos deles. Essencialmente a oposição entre paradigma e sintagma, apesar de bastante operacional, não é algo que se sustente, a não ser visualmente. De fato, do ponto de vista da sua essência, o sintagma se trata apenas da pura sequencialidade temporal da produção vocal com o tempo.

Na produção vocal, a Semântica (cognitiva e preverbal) antecede uma Morfologia e um Léxico (obtidas pela aquisição de linguagem), que se expressam por meio do sintagma, a única realidade concreta. A partir dele, por abstração, criam-se:

- (i) um paradigma de sequências permitidas nesse idioma (ou seja, uma Sintaxe), assim como
- (ii) um paradigma de palavras (um Léxico), que pode ser subdividido em:
 - (iii) um paradigma de unidades sem significado (uma Fonologia), por sua vez, subdividida em paradigmas de moldes fônicos (Viaro 2011) e de fonemas, e
 - (iv) em um paradigma de unidades com significado (uma Morfologia teórica), que pode subdividir-se em:
 - (v) um paradigma de elementos puramente relacionais (*pure-relational* Sapir 1921), como as vogais temáticas
 - (vi) e em outro, de elementos referenciais (uma Semântica teórica), onde podemos falar de relações como sinonímia ou hiperonímia.

Desse modo dizemos que a expressão se opõe à análise. A primeira é cultural (*théseis*), uma vez que não há expressão sem os elementos da língua, obtidos por meio aquisição de linguagem, enquanto a segunda é biológica (*phýseis*), pois a análise é inata à espécie humana: a língua em si é a notável síntese procurada desde os estoicos. O resultado da análise, porém, é necessariamente um construto teórico que retroalimenta a expressão.

Mais complexo é o problema da *analogia*. Largamente estudado em Linguística desde o século XIX, a analogia, inicialmente oposta pelos estoicos à anomalia, passa a opor-se às leis fonéticas aceitas então, funcionando conjuntamente com elas num modelo que prescinde do conceito de anomalia.

Suponhamos que numa sincronia S¹ tenhamos, numa flexão verbal, as formas *vestir* e *vesti* (respectivamente infinitivo impessoal e primeira pessoa do singular do pretérito perfeito do indicativo). Nessa sincronia observamos claramente que a sílaba final de ambas palavras se diferencia porque, no primeiro caso, há uma coda silábica (representada por *-r*) e, na segunda não. Suponhamos que numa sincronia subsequente S² haja, entre outras transformações, a aplicação de uma regra de supressão (no caso, apócope), de tal modo que promova a transformação *-r > ∅*. O resultado coincidente de ambos os signos seria, portanto, [vis'ti]. Pois bem, pode-se dizer que se trata de uma homonímia, de modo que poderemos falar de [vis'ti]¹ e [vis'ti]². Suponhamos ainda que a regra atue sobre todo o sistema e não especificamente sobre essa palavra (o que também poderia acontecer, como provam os dados históricos), ou seja, tem o caráter de lei fonética. Nesse caso, seria correto afirmar que há uma equivalência formal, embora o significado (ou a função, dependendo da teoria empregada) seja mantido intocado diacronicamente. Observa-se que o mesmo não ocorre com verbos da primeira conjugação (por exemplo, *falar* e *falei*), nem com os da segunda (*vender* e *vendi*), nem entre *pôr* e *pus*.

Do ponto de vista de um modelo que preconize a existência de um modelo analógico, não basta que descrevamos que 'o infinitivo dos verbos da terceira conjugação em S² equivale formalmente à primeira pessoa do singular do pretérito perfeito do indicativo' ou vice-versa, mas que se aponte a direção da atuação de uma palavra (ou flexão) sobre a outra. Claro está, porém, a direção da analogia não prescinde de uma hipótese.

Suponhamos que se estabeleça a hipótese de que 'a direção deve vir da forma não-alterada para a modificada', ou seja, a analogia seria exercida do pretérito perfeito para a forma apocopada do infinitivo. A favor da regra está o fato de haver atuação analógica evidente com *vir* e *vim*, uma vez que [vĩ] é um possível resultado, mas não ★[vi], que suporia que uma direção contrária (da alterada para a não-alterada). Contra a explicação analógica está a argumentação de que [vĩ] é forma herdada do português antigo *vīir* < latim *venire*, no entanto é preciso haver argumentos que legitimariam a apócope do *-r* em estágio anterior a S² (VIARO, 2011).

Uma hipótese não pode ter validade *ad hoc* se pretende ter valor dedutivo. A hipótese analógica de que ‘a direção deve vir da forma não-alterada para a modificada’ deveria explicar também outros tipos de convergência. Mas é complicado aplicá-la irrestritamente. Por exemplo, sabe-se que, em certos subsistemas do português, o presente do subjuntivo é substituído por uma forma que se assemelha ao infinitivo apocopado. Nesse caso, não cabem contra-argumentações puramente fonéticas com os chamados verbos irregulares. Assim, expressões como *se eu cantar*, *se eu vender*, *se eu vestir* são inegavelmente próximas em S^1 aos infinitivos *cantar*, *vender*, *vestir*, mas a equivalência formal não ocorre entre *se eu vir*, *se eu intervier*, *se eu fizer* e os infinitivos *ver*, *intervir*, *fazer*. Mas mesmo em S^2 , onde se diz *se eu ver*, *se eu intervir*, *se eu fazer*, há formas mais renitentes à mudança, como *se eu for*, *se eu tiver* (e nunca ★*se eu ser*, ★*se eu ter*) etc.

Como explicar essa irregularidade que impede a atuação irrestrita da analogia? Fora do âmbito da linguística poderíamos citar o seguinte paralelo: a água ferve, por definição a 100°C. Mas muitas vezes isso não é verdade. O valor para a temperatura de ebulição da água como 100°C parece aplicar-se apenas ao nível do mar, pois subindo uma montanha, o valor diminui. Isso só se explica por meio de uma operação com uma variável que leve em conta não só a definição, mas outro fator (no caso, a altitude). No caso de uma hipótese linguística, quantas variáveis são necessárias para que as exceções se integrem na regra geral?

Variáveis

Sejam as seguintes definições:

- (1) *paradigma* é um conjunto não-unitário de palavras reunidas por pelo menos alguma característica comum, seja do ponto de vista da forma, seja do significado, em uma sincronia S qualquer;
- (2) chama-se *molde fônico* a um subconjunto não-vazio de palavras dentro de um paradigma;
- (3) se S^1 é uma sincronia imediatamente anterior a S^2 , diz-se que uma palavra de S^1 sofreu uma mudança formal se há uma outra palavra em S^2 , que se supõe relacionada por *transmissão contínua*. Nesses casos, a palavra de S^1 é chamada de *étimo* da palavra de S^2 (VIARO, 2011);

(4) numa *mudança formal regular*, cada elemento que compõe a palavra em S^1 equivale a outro elemento em S^2 , conforme regras específicas obtidas por indução (*leis fonéticas*). Uma mudança formal que não siga tais regras é chamada de *mudança formal irregular*;

(5) chama-se *analogia* a uma hipótese de mudança formal irregular de um étimo por imitação de um molde fônico de S^1 ;

(6) *produto analógico* é a palavra em S^2 que se supõe relacionada por transmissão contínua de um étimo que sofreu analogia;

Todo molde fônico pertencente a um paradigma possui um potencial analógico. Desse modo, qualquer molde fônico poderia atuar sobre qualquer palavra em S^1 . Por que somente algumas o fazem? Quais são as variáveis que estão presentes para que um molde fônico seja um forte candidato a um étimo analógico? Tal resposta se mostra imprescindível para diversas questões de Etimologia mas também para entender a questão da mudança morfológica de todo o sistema.

Tomemos a seguinte hipótese: *o potencial analógico de um molde fônico é diretamente proporcional à sua frequência de uso*. De fato, a frequência de uso parece ser uma dessas variáveis. Uma palavra que afete outra deve ser frequente. Como num paradigma, todas as palavras poderiam, teoricamente, atuar analogicamente, uma sobre as outras, podemos concluir que uma palavra somente atua analogicamente sobre outra palavra num mesmo paradigma de menor frequência de uso. A atuação da analogia não é uma lei, pois pode ocorrer ou não. Por que não ocorre sempre? Outras variáveis estão em questão, como veremos. Seria estranho, no primeiro momento, que uma palavra rara, desusada ou arcaica pudesse alterar por analogia qualquer outra num paradigma. Da mesma forma, subconjuntos mais frequentes (conhecidos normalmente como *regulares*) parece ser uma fonte constante de atuação analógica. Mas não só a *frequência de uso*, mas a *dimensão* dos moldes (isto é, o número de elementos em seu conjunto) também parece atuar de maneira direta no potencial analógico.

Portanto, igualmente poderíamos dizer que: *o potencial analógico de um molde fônico é diretamente proporcional à quantidade de seus elementos*. Novamente, um molde composto de uma única palavra muito frequente pode ser eficaz no aumento do potencial analógico. Da mesma forma, um molde bastante numeroso tem o mesmo efeito. Se o molde é numeroso e frequente, estaremos bem próximo de um bom candidato à analogia.

Resumidamente: *sendo S um sistema de regras e palavras, pertencentes a n subconjuntos, chamados paradigmas $P^1, P^2... P^n$ e sendo M um subconjunto do paradigma P^n , chamado molde fônico, o*

potencial analógico (A) de M é, portanto, uma função (sobrejetora). Na formulação da teoria dos conjuntos diríamos que:

$$\begin{array}{ll} |S| = \aleph_0 & M \subseteq \subseteq P^n \subseteq \subseteq S \\ F(M) = \mu & A(M) = \mu \cdot |M| \end{array}$$

Como essas hipóteses têm sido criadas aqui de maneira intuitiva (e também dedutiva), é necessário que vejamos alguns exemplos: a frequência de *ser*, *ter* é muitíssimo maior do que a dos demais verbos. Eles costumam atuar diacronicamente como moldes fônicos? Em línguas flexivas, como o português e o russo, a alta frequência costuma associar-se com a propagação da irregularidade. O verbo *ser* tem um paradigma irregular herdado do *esse* latino. Entre as irregularidades há as chamadas supleções. Mas nem sempre uma alegada supleção não tem uma raiz comum: *eram* tem o mesmo radical que *esse*, mas as razões da diferença entre *er-* e *es-* são diacrônicas (no primeiro caso, por exemplo, houve um rotacismo de um *-s- original em posição intervocálica). Verdadeira supleção se encontra entre *esse* e *fui*, por exemplo. O verbo *ire* também sofreu supleções, algumas provenientes do verbo *esse*. Em português *fui* é a primeira pessoa do singular do pretérito perfeito do indicativo tanto do verbo *ser* quanto do verbo *ir*. Podemos falar de analogia nesse caso? Aparentemente não. Algumas construções latinas com *fui* substituíram o antigo *ii* em período muito antigo (*fui in funus*, por exemplo, em vez de *ii in funus*). A equivalência semântica entre *ii* e *fui* nesses casos, nascida do sintagma (isto é, por meio de uma semântica contextual), migrou para a morfologia e razões de ordem fonética reduziram a frequência de uso de *ii* em detrimento de *fui*. O verbo **andare* em várias línguas românicas também substituiu flexões de *ire* pelo mesmo sentido semântico e fonético.

Igualmente a primeira pessoa do singular do presente do subjuntivo *sim* foi substituído por *sedeam*, transformando-se em *seja* por razão de desgaste fônico provavelmente por causa de uma ampliação semântica de *sedere*, que atuava esporadicamente como sinônimo de *esse* (por exemplo, também o infinitivo *ser* em português provém de *sedere* como prova a forma antiga intermediária *seer*). Podemos dizer então que, a despeito da sua alta frequência, o desgaste fonético (motivador de palavras curtas demais ou de ambiguidades indesejadas) pode ser até mais importante na substituição de uma forma do que a frequência de uso. Mas seria uma substituição lexical como as acima apresentadas algo que poderíamos acrescentar como exemplo de analogia ou de mudança formal? Aparentemente, teríamos três tipos de mudança formal: a mudança fonética, a analogia e a supleção. É preciso estabelecer símbolos distintos para as três, respectivamente (Viaro 2011):

- $a > b$: a se transforma em b mediante regras fonéticas;
 $c \Rightarrow a > b$: a se transforma em b mediante analogia com c ;
 $a \square b$: a é substituído por b mediante sinonímia e/ou desgaste fônico;

Esses três casos de mudança formal precisam ser distinguidos de $a \rightarrow b$, que significa simplesmente que b é a atuação (em sincronia) de uma regra morfológica produtiva. Também o neologismo por decalque, indicado como $a \geq b$ indica que a foi substituído por um sinônimo b , no entanto, b busca motivações semânticas em a para fazê-lo. Podemos considerar esse tipo de transformação algo realizado também em sincronia. Se representarmos cada sincronia por meio de parênteses, essas fórmulas ficarão mais claras, a saber:

- $(a) > (b)$: *maturum > maduro*;
 $(c \Rightarrow a) > (b)$: (*quatro* \Rightarrow *cinque*) $>$ *cinco* ;
 $(a) \square (b)$: *cepit* \square **capuit*;
 $(a \rightarrow b)$: *caixa* \rightarrow *caixão*;
 $(a \geq b)$: *hot dog* \geq *cachorro-quente*;

Com relação especificamente à analogia, verificamos que, por vezes, não é apenas uma palavra de altíssima frequência que tem esse poder, mas todo um grupo de palavras dentro do paradigma. Nesse caso, o molde fônico contém mais elementos, todos de razoável frequência. Sendo 3CONJ a coincidência fonética causada pela apócope do *-r* entre infinitivo e a primeira pessoa do singular do pretérito perfeito do indicativo dos verbos da terceira conjugação, poderíamos representar o caso da nasalização de *vir* acima concisamente da seguinte forma:

$$(3CONJ \Rightarrow vir) > [vĩ];$$

Semiogênese

A *semiogênese* pode ser definida, no tocante ao signo linguístico, como a capacidade de parte de uma palavra adquirir um significado inexistente na sincronia anterior, também pode ser explicada pela analogia. Sequências de fonemas, isto é, elementos não-significativos do signo, às vezes, adquirem um significado e, em seguida, alguma produtividade.

Sabe-se, por exemplo, que a partir de neologismos como *decoreba*, surge um sufixo *-eba*, donde nascem palavras como *natureba*, *mistureba* e outros. Qual o significado de *-eba*? Aparentemente

natureba é uma forma mais coloquial de *naturista*, carregada de um certo valor ou pejorativo ou jocoso. Também *mistureba* nada mais é que uma *mistura*, talvez de coisas insólitas, talvez não, mas com o mesmo valor pejorativo ou jocoso. A frequência de uso de um vocábulo expressivo como *decoreba* dentro de um certo universo de discursos parece ser a alavanca para as neologias subsequentes. Não parece ter sido distinto o que aconteceu com outros sufixos hoje altamente produtivos. Por exemplo, o sufixo internacional *-ista* em português antigo era muito pouco produtivo (praticamente se circunscrevia a palavras como *batista* e *evangelista*), mas hoje possui toda uma gama de formas e serve de base para outras formações, como em *-ístico*, *-ística* etc. (AREÁN-GARCÍA, 2007, 2011).

Parece, portanto, que são dois os caminhos da formação de sufixos: a aglutinação (como no caso de *-mente* em português) e a semiogênese (a grande maioria das vezes, algumas muito recentes, outras já herdadas). No caso da semiogênese, muitíssimo mais frequente nas línguas flexivas, o papel do molde fônico analógico é extremamente importante.

Por vezes, há o trânsito de um caminho de formação para o outro. Por meio de aglutinação de radicais eruditos formavam-se palavras em *-dromo*. Alguns eram tão antigos, que a noção de “correr” do verbo grego *drómō* já se encontrava diluída, como em *pródromo* < gr *pródromos* “precursor” (já atestado no século XVII com outros sentidos como “sintoma”, “preâmbulo”). Provavelmente da restauração da palavra grega *palíndromos* (a palavra portuguesa *palíndromo* ocorre só no final do século XIX), fica evidente que a terminação *-dromo* era útil para expressar sinteticamente a ideia de “correr”. Criaram-se muitos neologismos (ou ressuscitaram-se outras palavras, de sentidos aproximados), sobretudo na área de Ictiologia e de Morfologia Botânica, desde a primeira metade do século XIX: *acródromo*, *actinódromo*, *anádro*, *anfídromo*, *anfísdro*, *antídromo*, *axonódromo*, *broquidódromo*, *campilódromo*, *camptódromo*, *catádromo*, *craspedódromo*, *diádromo*, *dictiódro*, *hifódromo*, *paralelódromo*, *plasmódromo*. A maioria dessas construções são empréstimos de alguma outra língua (latim científico, francês, inglês e alemão são sempre as mais recorrentes).

Desde meados do século XIX, a Arquitetura se valeu desse tipo de construção para descrever espaços (como *perídro*), muitos deles voltados a esportes onde há corridas (a partir de *hipódromo*, helenismo recuperado já no século XVIII). São abundantes esses casos no século XX: *autódromo*, *canódromo*, *cartódromo*, *motódromo*, *velódromo*. Nesse sentido, também na mesma época, o significado do segmento *-dromo* deixa de ser ‘local onde X corre’, para ‘local onde X V’, sendo V um verbo qualquer (talvez a base para esse novo sentido seja *aeródromo*, do início do século XX): *aeromodelódromo*, *bumbódromo*, *camelódromo*, *cosmódromo*, *sambódromo*, *tauródromo*.

É fácil de perceber que a polissemia do elemento de composição é diretamente proporcional à produtividade. Nesse caso, *-dromo* sempre foi sentido como um elemento de composição, mas a radical mudança não difere da atribuição de sentido a uma sequência previamente assemântica. Isso ocorre

porque muitas das criações anteriores ao sentido mais recente tinham radical erudito, portanto, opaco para a maioria dos falantes. De fato, *hipódromo* tem uma base *hip-* que aparece em *hípica*, *hipismo*, o que não o faz totalmente opaco para o significado “cavalo” para muitos falantes cultos, no entanto, é preciso um conhecimento ainda maior para entender palavras raras como *hipófago*, *hipomante*, *hipurria*, *hipiatro* ou mesmo uma palavra mais frequente como *hipopótamo*. No entanto, *motódromo*, *camelódromo*, *sambódromo* têm radicais muito mais transparentes. Deduz-se, portanto, que o potencial de produtividade está diretamente ligado ao reconhecimento dos radicais.

Mas que vem a ser o reconhecimento? Em *sambódromo*, reconhece-se *samb-* como a própria palavra *samba*. Esse reconhecimento ocorre em sincronia e não tem nada a ver com a história da palavra. Talvez não seja fácil, mesmo para um falante culto de português perceber que *hipocampo* tenha o *hipo-* ligado a “cavalo” enquanto *hipotálamo* tenha a ver com *hipo-* “sob”. Mas embora isso não seja fácil, não é impossível e tem a ver com a própria aquisição da linguagem ao longo da vida. Para um falante do inglês, o conhecimento etimológico, ainda que parcial, é requisito para que essas palavras sejam grafadas corretamente: *hippocampus* e *hypothalamus*.

Portanto, se há um reconhecimento de radicais e afixos, que o motiva senão uma “impressão”, que só é possível numa determinada sincronia? E sendo essa impressão variada, como falar de uma única tendência de reconhecimento? Vemos que *-dromo* foi mais frequente em determinadas subáreas da Biologia e da Arquitetura, portanto, é de se esperar que nessas áreas, o reconhecimento seja maior que em outras. Ou deveríamos apenas nos ater à linguagem média? Se sim, onde está ela? Não seria a ideia de que há um “conhecimento médio” sobre o qual se pauta a gramaticalidade e a verdade dedutiva uma ilusão? Não leva em seu bojo pressupostos politicamente tendenciosos? Ou, na melhor das hipóteses, não seria apenas uma visão ingênua de língua o que norteia vários campos da Linguística, inclusive o da Morfologia? Que testes experimentais provam a existência do “significado médio” que justifica a atuação do falante como juiz de sua própria língua?

Produtividade

Um sufixo neológico, portanto, é um tipo de resultado analógico cuja identificação do molde é objeto de estudo da morfologia de cunho diacrônico. Aliada à visão de língua ideal e metalinguística, a consciência de que existem paradigmas fez surgir, desde o tempo dos gregos (explicitamente na gramática alexandrina de Dionísio Trácio), o conceito de *classes de palavras*. Mais tarde, no Ocidente, por inspiração nas gramáticas indianas, o conceito de paradigma se ampliou para unidades significantes menores que as palavras. Nasce aí, em meados do século XIX, a noção de uma Morfologia *stricto sensu* em oposição à Fonética e à Sintaxe. Uma única palavra, mais tarde se dirá, pode compor-se de vários

morfemas, isto é, de vários signos linguísticos. No entanto, na rígida dicotomia saussureana que difere o eixo sintagmático do paradigmático, falta espaço para a atuação analógica.

Ora, em diacronia não há sistema, como preconiza o *Cours de linguistique générale* e a sincronia nada mais é que um recorte arbitrário dentro da existência heraclitiana de uma língua. Desse modo, o molde analógico se forma da própria substância do conteúdo e adquire forma devido a uma espécie de falsa segmentação e de uma semiogênese. Surge a dúvida de como isso é possível, mas as etimologias, embora disponham de graus de probabilidade, mostram inúmeros exemplos, muitos deles de difícil refutação.

Diremos, portanto, que cada item lexical é associado a uma variável de frequência de uso. Se dissermos que há apenas quatro graus de frequência de uso (*frequentíssimo, comum, raro, inexistente*), dentro de um determinado grupo de falantes, haverá uma média entre os usuários que determinaria, nessas quatro faixas, qual é a frequência de uso de uma palavra qualquer. Antes de perseguirmos esse raciocínio, precisamos lembrar que essa frequência de uso pode variar individualmente num mesmo grupo, dadas as condições biográficas do falante. Também dependendo da heterogeneidade discursiva escolhida num conjunto de menor ou maior extensão de falantes pode haver distorções numa pesquisa que pretenda descrever essas mesmas frequências.

Seja como for, a analogia só terá um modo de apresentar-se. Primeiramente, a analogia só ocorre em sincronia. Seria estranho imaginar que palavras em contiguidade temporal afetem ou sejam afetadas. Passada uma sincronia, a herança da sincronia subsequente é composta integralmente de resultados e processos cuja extinção ainda não é definitiva.

Desse modo, podemos definir *uma sincronia é um conjunto de palavras herdadas e de processos igualmente herdados que geram novas palavras*. Somente em línguas mortas ou em risco de extinção não existem processos que gerarão novas palavras para os futuros falantes. Não há, portanto, como imaginar uma sincronia-zero, a primeira de todas e, aparentemente, ela nunca existiu, uma vez que hoje em dia se questiona até mesmo se a linguagem é um desenvolvimento tecnológico do *Homo sapiens*, uma vez que nada impede que a vejamos como seja um fenômeno contínuo herdado de homínídeos ainda mais antigos (DAWKINS, 2004). Uma sincronia é, antes de tudo, um recorte temporal artificial, à disposição para a reflexão linguística.

Se numa sincronia há um conjunto de palavras herdadas e cada palavra possui uma diferente frequência de uso, calcada na experiência individual e compartilhada em maior ou menor grau com os demais falantes, podemos postular que *as palavras mais frequentes podem ser as mais propícias para a criação de um molde analógico*. Parece razoável. Uma palavra rara não alavancaria uma mudança, exceto numa única circunstância. Portanto, dizer que um determinado sufixo é produtivo é um modelo que pretende descrever uma sincronia futura. O que podemos dizer é que dispomos de uma hipótese

na qual determinado sufixo é considerado produtivo porque é prolífico (VIARO, 2010). No entanto, observamos que forma prolíficas podem não ser produtivas e, inversamente, formas pouco prolíficas podem tornar-se produtivas, se fazemos o histórico das sincronias pretéritas. Deduziremos que:

A *prolificidade* se associa a cada \mathbb{N}_0 de cada sincronia S; já a *produtividade* se associa à projeção de uma sincronia futura. A *expectativa de produtividade* (E) é *diretamente proporcional à expressividade do molde fônico* (ϵE).

Quanto maior o potencial analógico, maior a expectativa de produtividade, mas se a expressividade for zero, a expectativa de produtividade também será zero.

$$\begin{aligned} A(M) &= \mu \cdot |M| \\ E(M) &= \epsilon \cdot A(M) \therefore \epsilon = \frac{E(M)}{A(M)} \therefore \epsilon = \frac{E(M)}{\mu \cdot |M|} \therefore \epsilon = \frac{E(M)}{A(M)} \therefore \epsilon = \frac{E(M)}{\mu \cdot |M|} \therefore \epsilon = \frac{E(M)}{F(M) \cdot |M|} \\ \therefore \epsilon &= \frac{E(M)}{F(M) \cdot |M|} \end{aligned}$$

Como mensurar a expectativa da produtividade para calcularmos sua expressividade? Como calcular a frequência sem valermos da *huge chart*? Sem esses dados, como responderíamos como se formam os *moldes fônicos que atuaram na analogia e que atuarão na produtividade*? A resposta não se encontra apenas no abandonado Simbolismo Fonético, mas também na pouco desenvolvida Semântica (que depende dos contextos históricos).

De fato, na produção da fala há muitas vezes aquilo que a Estilística costuma chamar de “expressividade”. O conceito vago deve dispor de uma definição rigorosa. Apesar de ser um conceito refratário a definições, portanto, pouco operacional (e, diriam muitos, pouco científico), é inegável que alguns signos não dispõem apenas de um significado neutro. Mas em que distingue o significado normal do significado expressivo? Palavras chulas, vocábulos elevados, neologismos, gírias: que têm além de um significado e um significante? Normalmente remete-se à *parole*, ao discurso e a tantos outros elementos *ad hoc* aquilo que caracteriza o expressivo, de modo que a linguagem real parece confundir-se com aquele ideal setecentista do *inutilia truncat*.

Diz-se, inclusive, de modo confuso, que significado não é a mesma coisa que sentido. O significado na interpretação do signo do *Cours de linguistique générale*, contudo, é uma dicotomia e se opõe ao significante. Um traço pejorativo atribuído a uma palavra como *amigo* por uma entonação qualquer (numa frase como “*esse meu amigo me pôs em maus lençóis*”) é evocada como que sobrevoando o significado ideal da palavra *amigo*. Como o sintagma tem a capacidade de mudar o signo ao seu

bel-prazer não é claro. O recurso *ad hoc* do ‘contexto’ é por demais vago e acientífico. No entanto, um recurso como a ironia, presente nesse exemplo (em que *amigo* equivale a *mau amigo*) não é possível com qualquer palavra, por exemplo, com *relógio*. Podemos dizer que há dois tipos de significado: um descritivo e um avaliativo e é possível sobrepor-los com facilidade.

Por vezes o sentido avaliativo requer algo mais que uma mera intonação e exige algo que poderíamos chamar de *massa fonológica*. O recurso mais comum em português são os sufixos de diminutivo: *esse seu amiguinho me pôs em maus lençóis*. Nesse caso, a ironia é ainda mais evidente. Por vezes é uma palavra cuja frequência de uso é pervertida. Uma pergunta como *onde a senhorita esteve?* não é uma mera interrogação, mas transmite de modo eficiente ironia e irritação por causa do uso de um elemento normalmente *in absentia*, isto é, com baixa frequência de uso.

Voltando, portanto, à nossa afirmação anterior de que *as palavras mais frequentes são as mais propícias para a criação de um molde analógico*, podemos encontrar uma refutação no fato de que há palavras pouco frequentes e que podem atuar com igual força num molde analógico. No caso do sufixo *-eba* não há, para a criação do molde analógico *decoreba*, nenhuma palavra que possamos dizer que seja extremamente frequente (*jurubeba*, *tatu-peba*, *pereba* etc), no entanto, a ironia somada de todas elas (simplesmente pelo fato de serem palavras populares) pode fazer exsurgir esse valor e impingir o caráter coloquial e informal. Como prová-lo para além da introspecção e da sensação de falante, que quase sempre se revela pouco científica?

De qualquer forma, é razoável pensar que uma palavra ou um paradigma que sirvam para a criação de um molde analógico não pode ter baixíssima frequência de uso ou, ao menos, não funcionaria apropriadamente como reconstrução, no caso de uma proposta etimológica. Desse modo, podemos concluir que *quanto mais frequente uma palavra e quanto mais expressiva tanto mais provável será a sua candidatura a um molde fônico*. Nem uma palavra pouquíssimo frequente, nem uma pouco expressiva tem capacidade para alavancar a neologia. O mesmo vale para os estudos tanto para os estudos lexicais quanto para os morfológicos. Nesse sentido, as palavras têm o mesmo comportamento dos elementos que as compõem, por exemplo os sufixos, quando falamos de baixa ou alta produtividade sufixal.

A Morfologia e a Linguística

Diacronicamente, porém, os elementos que são construídos pela analogia se acumulam aos elementos herdados, de modo que não são poucos os chamados morfemas de prolificidade baixa (como *-ongo* em *japorongo*), para não falar de terminações únicas, com um ou dois exemplos (como *-ujo* em *marujo*), às vezes dificilmente agrupáveis semanticamente. Normalmente o conjunto de elementos não aproveitáveis para construções futuras, por terem produtividade inexpressiva, é conhecido como *lé-*

xico, opondo-se à própria *morfologia* da língua, no sentido aronoffiano, no entanto essa divisão só faz sentido do ponto de vista cognitivo, que, de modo algum, é o mais importante para uma teorização morfológica. Por mero acidente histórico, uma palavra do léxico pode adquirir frequência de uso e consequente expressividade, atuando analogicamente sobre outras, ressuscitando morfemas já pouco produtivos ou mesmo criando novos por falsa segmentação. A oposição entre léxico e morfologia com base na produtividade é rígida demais e gera mais exceções e explicações *ad hoc* quanto a oposição que podemos fazer entre elementos frequentes e pouco frequentes.

Mas como atingir a *huge chart* bloomfieldiana para afirmarmos com certeza que um elemento seja de fato frequente? Obviamente, isso é impossível (BLOOMFIELD, 1933:46-47). Portanto, entre o recurso da intuição e o do recorte, uma postura indutiva julga o segundo um pouco mais científico que o primeiro. Servirá como *corpus* o universo de textos da *internet*, acessável por meio de buscadores como o *Google*? Ou nos valeremos de *corpora* montados sobre rígidos princípios, de modo que, assim, nos salvaguardaríamos de fazer afirmações sobre seleções viciadas? Aparentemente qualquer afirmação sobre a frequência de uso é algo de natureza estatística.

Como fazem outras ciências? Para a Zoologia, um animal coletado e descrito como espécie nova é conservado num museu ou numa outra instituição. Ele é conhecido como *holótipo* e é preservado para o estudo de outros pesquisadores. Por motivo de destruição, pode-se designar um substituto, chamado *neótipo*. Palavras são coletadas há tempos para exemplificar línguas, dialetos, gírias, no entanto, a sua localização nas biografias e a organização desses dados, mesmo depois da invenção da *media* magnética e eletrônica não é rigorosa, uma vez que há a tradição do *testemunho*. Se se baseasse somente em testemunhos, a zoologia teria tido problemas em avançar teoricamente rumo a modelos dedutivos como o da Teoria da Evolução, pois ainda estaria tentando resolver problemas criptozoológicos que envolvessem dragões, unicórnios ou o Ipujiara. Não se submetendo jamais a uma segunda opinião, o dado linguístico, por vezes coletado, por vezes introspectado, é fugidio e só é recuperável quando é escrito (e mesmo assim é por vezes recuperado por meio da memória) ou quando é gravado.

Com o tempo, os exemplares guardados em museus (holótipos, neótipos etc) passaram a ser entendidos não mais como representantes de uma espécie, mas como representantes de uma população. Como não é possível fazer um paralelo exato desse procedimento com os estudos linguísticos, imagina-se que é algo incomparável. No entanto, a ideia da reprodução em Biologia assemelha-se bastante à da comunicação em Linguística. A ideia da comunicação como elemento norteador fez com que se imagine que um elemento lexical é um elemento de uma língua incomparável ao de outra. Toda semelhança física é quase tomada como um acidente que, se não é coincidente, é pouco interessante ao estudo linguístico, pois cada sistema é uma caixa fechada em si mesmo. Diacronicamente, isso é obviamente errôneo. Uma palavra como português *dentista* e o francês *dentiste* não foram formadas por

coincidência, mas há uma dependência entre elas de caráter histórico. Se o signo lexical é arbitrário, a combinação, os valores semânticos decorrentes e a escolha paradigmática não são coincidências. Há algo de dogmático quando se pensa que não interessa ao estudo da comunicação, afinal um falante do português achará mais fácil aprender a palavra *dentiste* do que a palavra alemã *Zahnarzt*. Aparentemente esse dogmatismo nasce da política, de questões nacionalistas derivadas dos limites geográficos das nações, e não têm nada a ver com a Linguística. Se indivíduos pertencentes a distribuições regionais distintas pertencem a um *pool* genético que permite o seu cruzamento e procriação, estamos diante de uma mesma espécie, por mais distinta que ela se apresente. A criação de híbridos estéreis mostram que são espécies muito próximas. Em linguagem, a eficiência comunicativa, total ou parcial, de uma palavra, mesmo se de línguas distintas, evidencia, por exemplo, que *dentista* e *dentiste* não são palavras distintas, mas a mesma palavra, diacronicamente falando. No entanto, para a Diacronia, o elemento da compreensibilidade não é crucial, uma vez que um decalque pode ser também considerado a mesma palavra. Por fim, seja qual a relação entre dois indivíduos linguísticos, por mais distantes que sejam do ponto de vista formal ou semântico, se é possível uma relação entre os dois, ainda que hipotética, estamos trabalhando com a possibilidade de serem o *mesmo* elemento sob a pressão de mudanças.

Em Física e em Química, um fato só é verdadeiro se reproduzível por experimento, mas em Linguística, sabe-se, um estímulo qualquer, por mais bem elaborado que seja, não gerará a repetição do fato linguístico que se busca, mas algum elemento de uma rede de associações à qual o fato está associado. A maioria das palavras interessantes para uma análise se extinguem na própria produção de fala, restando apenas a excepcionalidade do testemunho escrito ou gravado. Sobre esses testemunhos reconstruímos sincronias pretéritas. No caso da sincronia presente, valemo-nos também da lembrança de expressões vivenciadas em momentos passados (nesse caso, obviamente não filogeneticamente, mas ontogeneticamente), ou das sensações de produtividade, único elo que temos com o futuro.

Eis portanto a natureza do testemunho linguístico:

- (a) palavras vinculadas a sincronias pretéritas, localizáveis em textos específicos. Neste caso também estão testemunhos gravados. Falta uma sistematização desses dados, de modo a facilmente encontrá-los. Falta também uma reconstrução da sua frequência de uso, bem como da sua abrangência;
- (b) palavras vinculadas à sincronia presente, vinculadas à ontogenia da testemunha, com sensação da sua frequência de uso, abrangência e produtividade;

Não deixa de ser espantoso percebermos que o que chamamos de *sistema* é ou uma reconstrução ou uma sensação. Seria portanto a Linguística uma espécie de manifestação do estágio precientífico de uma Ciência da Linguagem futura? Ou busca algo inalcançável? Da mesma forma, o holótipo não é a espécie, mas um representante dela. Substituído pelo conceito de *população*, o conceito de espécie transcende o atomismo dos espécimes individuais. Dessa forma, o holótipo é uma amostra da *huge chart* da espécie, escolhido por acaso. Não é distinto o papel do dado linguístico. Um novo espécime coletado na mesma população ou em outra pode alterar sobremaneira o entendimento daquilo que se entende por espécie. Pode até mesmo anulá-la pelo mecanismo da sinonímia, hoje muito mais desenvolvida devido à Genética e à Cladística. Seria o papel da Linguística tratar de seus fatos por meio de modelos, como quando a Astronomia trata de galáxias que estão a anos-luz de distância?

Os métodos em Biologia alteraram o conceito de *espécie*, tornando-a inclusive em algo fluido o bastante para percebermos que beira o recorte arbitrário sobre o movimento heraclitiano dos seres. Seriam as línguas diferentes? O encerramento do signo e o recorte da realidade nos dá mais segurança do atomismo? Se a resposta for positiva, aparentemente há muito de ingênuo nas definições da Linguística. Em Química, as propriedades dos elementos e dos compostos seguem direções difíceis de percorrer. Basta um próton e as propriedades de um elemento se alteram completamente. Não é assim com as combinações de fonemas? Nos átomos, porém, a mesma matéria de que se compõem os prótons é uma só. Diferenças quantitativas fazem um átomo de hidrogênio diferenciar-se de um de hélio e um átomo de oxigênio diferenciar-se de uma molécula de oxigênio. Nas línguas, as unidades mínimas fonológicas já são semelhantes a átomos formados e só veríamos alguma unidade nos elementos acústicos, uma vez que o elemento articulatório que os produz de difícil redução a um único elemento.

Se não há paralelo perfeito entre o objeto de estudo da Linguística e o da Zoologia, da Física ou da Química, haverá então na Lógica? Muitos apostam que sim. Reduzindo a linguagem à Lógica, atinge-se, segundo alguns, o seu cerne (ao menos do significado). No entanto, a Lógica, por mais requintada que seja, é uma espécie de língua ideal. Está ao mesmo tempo, muito além da linguagem, pois expressa o inexprimível pelas línguas, e também muito aquém dela, pois não expressa as coisas mais básicas para o ser humano, como o tempo ou a valoração. A Lógica, como uma paralinguagem, portanto não é o melhor modo de fazer nossas reduções, que devem provir de universais. No entanto, o que é, de fato, universal em linguagem é de uma banalidade inoperante. Por exemplo, é universal que as línguas têm um número finito de fonemas, que têm mais substantivos que conectivos, que dispõem de uma sintaxe etc. Tudo que se arroga universal para além do trivial, na verdade não o é, pois convive com a exceção e precisa, na falta de variáveis inquestionáveis, do *ad hoc* explicativo, que resvala na Retórica. Dizer que a linguagem não pode ser explicada pela Lógica não é o mesmo que dizer que os modelos que expliquem seus fenômenos não devem pautar-se numa espécie de Lógica.

Um modelo de linguagem

Na falta de uma *huge chart*, os elementos coletados para nossa análise devem fazer parte de um *continuum* reconstruído, seja em Sincronia, seja em Diacronia. Essa reconstrução deve ser algo que possa ser provado. Nesse sentido, a Linguística se assemelha mais à Paleontologia. Nessa ciência, por exemplo, há previsão do que podemos encontrar num determinado estrato, dadas as exatidões das confirmações da teoria da Evolução. Não esperamos encontrar peixes antes do período Ordoviciano, por exemplo. Nas línguas, o mesmo acontece, mas apenas com relação às palavras que designam descobertas tecnológicas e elementos culturais específicos. *Grosso modo*, não se pode dizer da grande maioria das palavras, que vão e vêm, graças à perenidade que lhes é conferida pela escrita.

Em Morfologia Histórica valemo-nos dos dados obtidos pela Filologia, das sistematizações das Gramáticas descritivas, do ordenamento ainda incipiente da Lexicologia e das conclusões da Etimologia e da Linguística Histórica. Metodologicamente, é necessário cuidado extra com as interpolações, com as edições utilizadas e com as cópias, que podem deslocar dados de sincronias adequadas para outras menos prováveis. Tudo isso refletirá na reconstrução da sincronia, tal como uma interpretação errônea ou induzida pela falta material em Paleontologia pode criar uma tradição de informação dificilmente corrigível do ponto de vista da divulgação. Por mais divulgado que seja um fato, após uma refutação, não terá sobrevida em Ciência. Mesmo valendo-se de dados filologicamente trabalhados, um dado num aparato crítico não é menos importante do que o da versão escolhida para compor o texto editado. Conjecturas paleográficas na leitura de um texto de escrita complexa não pode ter o peso de um dado. Conjecturas filológicas para recompor um texto, por mais bem feitas que sejam, não são dados reais.

Um método etimológico mais adequado aos dias de hoje não poderá deixar de levar em conta isso. Sobre os dados filológicos nasce a Etimologia e sobre ambos, a Linguística Histórica. Talvez devêssemos, à luz de tantas discussões e reflexões entabuladas há dois séculos, apresentar uma solução e isso deveria iniciar pelo próprio instrumental dessas ciências. A definição de conceitos, tão pulverizados em correntes, deveriam ser decididos em congressos internacionais que não favorecessem as especializações, tal como o IPA é universalmente usado por todas as correntes. A Morfologia poderia dar um passo nesse sentido. A vitória do relativismo nas Humanidades em nada tem colaborado para isso há mais de um século. Esses mesmos conceitos poderiam valer-se de uma Terminologia internacional que sugerisse sua tradução nas mais diversas línguas do globo, algo de que nem mesmo o IPA dispõe (Viaro 2011). Os símbolos utilizados pela Linguística Diacrônica carecem de especificação mais rigorosa. O asterisco schleicheriano, por exemplo, para indicar forma hipotética criou polissemia com o gerativismo, no qual significa forma agramatical: uma síntese de ambas as perspectivas não deve ignorar isso. Além disso, a forma hipotética não segue a sugestão de Jespersen de marcar os graus de

certeza (Jespersen 1954). Por fim, os símbolos mais utilizados em diacronia (a saber, > e <) misturam transformações fonéticas e morfológicas, assim como não deixam clara a distância dos saltos intersin-crônicos que preveem.

Corrigidos esses problemas instrumentais, a própria noção de Língua deveria ser revista, de modo a não confundir-se com a língua-ferramenta ou com a língua-ideal de que falamos inicialmente. A própria noção de sistema deve estar atrelada apenas à comunicação, que é uma faceta da linguagem distinta da transmissão. Uma delas não é mais importante do que a outra, contudo. Não é possível entender a comunicação e a analogia sem o conhecimento da Teoria da Evolução e dos modelos desenvolvidos hoje na Neurologia e na Psicologia acerca do funcionamento do cérebro. Além disso, o método indutivo da História aliado à perfeita compreensão da Aquisição da Linguagem são imprescindíveis para compreender o fenômeno da Transmissão. O estudo de línguas crioulas em formação pode auxiliar-nos nessa tarefa.

Por fim, o estudo diacrônico depende da descrição de sistemas (em diversas sincronias), da definição de expressividade e de informações relativas às frequências de uso. Esses dados são tingíveis apenas por meio de métodos e não de convenções dogmáticas. Além disso, depende também de reconstruções, as quais sabidamente não são fatos, mas modelos, passíveis de refutação a qualquer momento, no sentido popperiano do termo. Paradoxalmente, só a indução pode conferir cientificidade à Ciência da Linguagem e à Morfologia em particular. Nesse transitar entre teorias e acúmulos de dados é que nos encontramos atualmente.

Referências

AREÁN-GARCÍA, Nilsa. *Estudo comparativo de aspectos semânticos do sufixo -ista no português e no galego*. 2v. Diss. de mestrado. São Paulo: FFLCH/USP, 2007.

AREÁN-GARCÍA, Nilsa. *Aspectos sincrônicos e diacrônicos do sufixo -ístico(a) no português e no galego*. 2v. Tese de doutorado. São Paulo: FFLCH/USP, 2011.

AYERS, Michael R. *George Berkeley - philosophical works including the works on vision*. London: Everyman/ J. M. Dent/ C. E. Tuttle.

BLOOMFIELD, Leonard. *Language*. New York/ Chicago/ San Francisco/ Toronto: Holt, Rinehardt and Winston, 1933.

CHOMSKY, Noam. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press, 1965.

DAWKINS, Richard. *The ancestors' tale: a pilgrimage to the dawn of life*. London: Weidenfeld & Nicholson, 2004.

DEL RÉ, Alessandra et alii. *Aquisição da linguagem: uma abordagem psicolinguística*. São Paulo: Contexto, 2006.

HEGENBERG, Leônidas. *Explicações científicas: introdução á filosofia da ciência*. São Paulo: Herder/Edusp, 1969.

HEINE, Bernd & KUTEVA, Tania. *World lexicon of grammaticalization*. Cambridge: CUP, 2002.

JESPERSEN, Otto. *Language: its nature, development and origin*. London: George Allen & Unwin, 1922 (1954¹⁰).

KIRK, G. S.; RAVEN, J. E.; SCHOFIELD, M. (ed). *The presocratic philosophers: a critical history with a selection of texts*. Cambridge: CUP, 1983.

LORENZ, Konrad. *Vergleichende Verhaltensforschung: Grundlagen der Ethologie*. Wien/ New York: Springer, 1978.

SACKS, Oliver. *An anthropologist on Mars: seven paradoxical tales*. New York: Vintage Books, 1998.

SAPIR, Edward. *Language: an introduction to the study of speech*. New York: Harcourt, Brace & World, 1921.

SOKAL, Alain; Bricmont, Jean. *Fashionable nonsense: postmodern intellectuals' abuse of science*. New York: Picador, 1998.

VIARO, Mário Eduardo. Sobre a inclusão do elemento diacrônico na teoria morfológica: uma abordagem epistemológica. *Estudos de lingüística galega*. Santiago de Compostela, 2: 173-190, 2010.

VIARO, Mário Eduardo. *Etimologia*. São Paulo: Contexto, 2011.