

CHUNKS EM MOLDES: HIPÓTESE DE DIÁLOGO ENTRE BYBEE E BOOIJ¹

Janderson Lemos de SOUZA²
(Universidade Federal do Estado de São Paulo)

Resumo: Neste artigo, proponho a aproximação entre duas propostas teóricas: a que Bybee (2010) formula com base na concepção de *chunking* como um mecanismo de processamento sensível à frequência de uso e aquela em que Booij (2010) aprofunda a concepção de construções morfológicas como esquemas. Uso a formação da palavra “entorno”, objeto de uma das minhas orientações de iniciação científica com foco no comportamento semântico, como evidência de que, entre os *chunks* previstos pela autora, os que correspondem a palavras recaem num dos moldes previstos pelo autor.

Palavras-chave: *Morfologia; Semântica; Linguística Cognitiva.*

INTRODUÇÃO

A linguística cognitiva tem privilegiado duas grandes abordagens às unidades de uma língua: a que atenta para os produtos, entendidos segundo alguma concepção de construção, em remissão a Langacker (1987) se o limite for o próprio quadro teórico ou a Bloomfield (1933) se o limite for a pequena história da linguística tal como entendida por Robins (1997), e a que atenta para os processos, um dos quais seria o *chunking*, ou formação de *chunks*, unidades de organização da memória formadas a partir de *chunks* anteriores como efeito da repetição.

Na abordagem construcional, convivem diferentes definições de *construção*: “(...) *invoking the construction concept does not guarantee unequivocal usage of the term*” (Lampert & Lampert, 2010: 33). Afora as especificidades de cada autor, que este artigo não se dispõe a esmiuçar, permanece válido o pressuposto de que:

“If constructional schemas may be considered 'templates for the construction and evaluation of novel expressions' (Langacker 2005a: 170), the constructional schema (at whatever level) becomes the basic site for any studies of word-formation in Cognitive Linguistics (...)” (LAMPERT & LAMPERT, 2010: 38)

Nas próximas seções, exploro essa relação entre a identificação de construções gramaticais, em geral, e morfológicas, em particular, e a formação de novas unidades, tendo em vista retomarduas questões tradicionais, que discuti no XVII Congresso Internacional da ALFAL como parte do projeto *A morfologia e suas interfaces*: (i) o que se entende por processos de formação de palavras e (ii) os mecanismos cognitivos que estruturam a polissemia. Ao final, pretendo demonstrar a validade da relação na descrição da palavra “entorno” no português brasileiro.

¹ Título da conferência que apresentei na abertura do V Seminário do Núcleo de Estudos Morfossemânticos do Português (NEMP) a convite dos líderes do grupo de pesquisa e organizadores do evento, Maria Lucia Leitão de Almeida e Carlos Alexandre Gonçalves, em novembro de 2014.

² Doutor em Letras Vernáculas pela UFRJ. Professor Adjunto da UNIFESP.

Chunks em moldes: hipótese de diálogo entre Bybee e Booij

1. “PROCESSOS? FORMAÇÃO? PALAVRAS? QUE SENTIDO ISSO TEM EM LINGUÍSTICA COGNITIVA?”³

A expressão “processos de formação de palavras” foi incorporada pela descrição e pelo ensino da língua portuguesa, a tal ponto que parece não abrigar convicções teóricas. O lexicalismo gerativo (Chomsky, 1970) consolidou o entendimento da expressão como indicativa da incidência de algoritmos na criatividade lexical, organizada por regras de redundância morfológica e regras de redundância semântica (Jackendoff, 1975) ou regras de formação de palavras e regras de análise estrutural (Basílio, 1980) e sensível ao bloqueio (Aronoff, 1976).

“A competência lexical corresponde não (apenas) ao conhecimento das palavras da língua, mas (também) ao conhecimento de um conjunto de regras, que representa a capacidade do falante de relacionar itens lexicais uns aos outros, analisar a estrutura interna das palavras e formar novas palavras.

A perspectiva gerativa deu um grande impulso nos estudos lexicais, na medida em que focaliza o léxico como conhecimento, em oposição à visão tradicional do léxico como vocabulário. A noção de competência lexical, abarcando as regras que determinam as possibilidades de formação, introduz no conceito de produtividade, já utilizado por autores clássicos, uma relevância crucial, já que a produtividade de regras de formação deixa de ser um comentário sobre a vitalidade de um processo específico para se transformar no fator determinante da constituição da classe potencial de palavras de uma língua.” (BASÍLIO, 2010: 2)

A mencionada produtividade se caracteriza como um conceito formal e pressupõe que o falante ideal conhece a totalidade das palavras reais de sua língua, como um *software* instalado e disponível à incidência de regras de recuperação ou acesso (*retrieval*).

“Um dos problemas relativos à utilização de regras para a descrição de processos lexicais nesta abordagem está na imprecisão do conceito. Conforme observado em Basílio (1998), o conceito de regra não é explicitado de modo uniforme em seu uso na literatura de cunho lexicalista sobre formação de palavras. Esta é uma situação desconcertante, pois a idéia da previsibilidade estabelecida por regras permeia todo o desenvolvimento da gramática gerativa, sendo a gramaticalidade determinada pelo conjunto de regras que define a língua correspondente. Assim, é deveras problemática a imprecisão do termo no gerativismo lexicalista.” (BASÍLIO, 2010: 3)

Nesse quadro teórico, o termo “processo” remete ao modelo de análise adotado:

“In an influential article of the mid 1950s, Hockett pinpointed three **models** of grammatical analysis in general – three different ‘frames of reference’ (to adapt his words) within which an analyst might ‘approach the grammatical description of a language and state the results of his investigation’ (first sentence of HOCKETT, ‘Models’). In the terms which we are using, these are particular sets of formal principles. Of Hockett’s three, one, which he called the ‘Word and Paradigm’ model, evidently referred to the traditional description of the older European languages. Another, which he labelled ‘Item and Arrangement’, is a model in which morphemes are the basic units of meaning and in which they were arranged linearly. The third (‘Item and Process’) is one in which the structure of a word is specified by a series of operations (...)” (Matthews, 1991: 21-22)

Trata-se, portanto, de processo como termo técnico que remete ao modelo de análise adotado pelo quadro teórico, e não como palavra da linguagem ordinária. O compromisso com uma cognição modular que organiza uma gramática baseada em regras exige que o termo seja revisto ou abandonado por um modelo que se compromete com uma cognição conexionista que organiza uma baseada no uso.

³ Comunicação apresentada no XVII Congresso Internacional da ALFAL, em julho de 2014, na UFPB.

Janderson Lemos de Souza

Já o termo “formação” integra a expressão em sua interpretação ativa:

“De fato, podemos observar que ‘formação’ tem duas interpretações: uma interpretação ativa, em que o termo se refere ao processo de formar palavras; e uma interpretação mais passiva, em que o termo se refere à maneira como as palavras estão constituídas.”(BASILIO, 1987: 16)

Isso porque, como vimos, esse quadro teórico se interessa pela “(...) classe potencial de palavras de uma língua”, e não pelo acervo das palavras reais. Essas importam como conhecimento prévio ao julgamento do falante quanto à necessidade de criar uma nova palavra.

Essa forma de exprimir a distinção entre o a formar e o já formado é posterior à que se notabilizou:

“Nesta seção demonstraremos como abordar simultaneamente novas criações lexicais e construções fossilizadas no léxico, através da distinção entre Regras de Formação de Palavras (RFP's) e Regras de Análise Estrutural (RAE's).” (BASILIO, 1980, 49)

Em reflexão posterior, a autora enfoca a interpretação passiva de “formação” e conclui que “(...) a expressão ‘Regras de Formação de Palavras’ deve ser substituída, ou, pelo menos, entendida como ‘padrões frequentes de configuração estrutural de unidades lexicais’” (Basilio, 1998: 252).

O enfoque na dimensão passiva é o que me parece levar à identificação de “padrões frequentes” e são esses padrões frequentes que são concebidos como construções gramaticais tanto na linguística gerativa quanto na linguística cognitiva. Já o enfoque na dimensão ativa preserva na linguística gerativa o compromisso como processos e regras, enquanto alimenta na linguística cognitiva o compromisso com a frequência de uso. Guardadas as especificidades de cada teoria, é a alternância do olhar para uma ou outra dimensão do léxico o que sobressai como recorrente na história da linguística, em geral, e da morfologia, em particular.

Como veremos, Bybee (2010) enfoca a mudança linguística como sucessivas reorganizações cognitivas (processo), enquanto Booij (2010) destaca os moldes das construções existentes (produto). O espaço para a diacronia em Bybee (2010) se deve ao rompimento com o princípio da sincronia, típico de modelos baseados no uso.

A propósito, é a frequência de uso que produz o que Langacker (1987) chama *status de unidade*, ou seja, ser uma unidade da língua depende de sua *rotinização*, e não de alguma propriedade formal da unidade. Também é a frequência de uso que produz o que Bybee (2010) entende por *chunking*:

“The principal experience that triggers chunking is repetition. If two or more smaller chunks occur together with some degree of frequency, a larger chunk containing the smaller ones is formed. Chunking is of course a property of both production and perception and contributes significantly to fluency and ease in both modes. The longer the string that can be accessed together, the more fluent the execution and the more easily comprehension will occur.” (p. 34)

Para concluir a questão sobre o que se entende por “processo de formação de palavras”, uma vez abandonado o compromisso com algoritmos representado pelo termo “processo” e mantida a alternância de foco ora na dimensão ativa ora na dimensão passiva do léxico, resta explicitar o que se entende por palavra na vertente da linguística cognitiva aqui adotada:

Chunks em moldes: hipótese de diálogo entre Bybee e Booij

“In cognitive terms, a word can be described as a symbolic label of mental categories referring to (in)animate objects, to states, actions, conditions and qualities as they are perceived by and conventionally construed in the human mind in interaction with the social and natural environment. According to this definition, words can also label complex mental categories. For example, the word *lover* expresses an agentive role (encapsulated semantically in the suffix *-er*) that is tied to the state of love. This simple example indicates the general creative capability of the human mind to construct and label new concepts, also by combining existing mental categories. In the history of linguistic research, such creative processes have come to be referred to as processes of word formation.” (ONYSKO & MICHEL, 2010: 2)

Pensando assim, situo o *chunking* de Bybee (2010) como parte da “capacidade criativa da mente humana de construir e rotular novos conceitos” (dimensão ativa) e os moldes de Booij (2010) como os “padrões frequentes” dos “rótulos simbólicos” das “categorias mentais existentes” (dimensão passiva). Nesses termos, a expressão ‘processos de formação de palavras’ deve ser substituída, ou, pelo menos entendida como ‘a influência dos padrões frequentes na combinação de rótulos simbólicos existentes para construir e rotular novos conceitos’.

2. “HETEROSSEMIA NA INTERFACE SEMÂNTICA-MORFOLOGIA”⁴

Uma vez explicitado o que considero como processos de formação de palavras, passo a defender que se tome a heterossemia como um dos mecanismos que organizam a polissemia. Mas por que tratar da polissemia neste artigo? Porque a linguística cognitiva se caracteriza, dentre outros pressupostos, pela defesa do foco na polissemia em detrimento do foco na homonímia:

“(…) a polissemia é um fenómeno demotivação, que introduz uma certa redundância no léxico mental, a passo que a homonímia é um fenómeno acidental. O que daqui se pode concluir é que polissemia e homonímia não constituem uma dicotomia estrita, mas antes fazem parte de um *continuum* de relação de sentidos. E metodologicamente, como evidenciaremos na parte descritiva deste estudo, uma análise *polissémica* será preferível a uma análise *homonímica* sempre que se encontrarem factores de coerência semântica num complexo de sentidos associados a uma mesma forma.” (SOARES da SILVA, 2006: 49)

Por sua vez, a polissemia é concebida não somente como a associação de mais de um significado a uma forma, mas sobretudo como uma relação cognitivamente estruturada entre os significados: “*Polysemy occurs when a single word has more than one meaning – and when those meanings are systematically related. Systematic relationship is crucial here*” (Lakoff, 1987: 316). Em outras palavras, a polissemia consiste na categorização de significados por meio de padrões de figuratividade, nomeadamente a metáfora e a metonímia:

“Hoje, sabemos que metáfora e metonímia são fenómenos conceptuais por natureza, processos e modelos cognitivos, constitutivos do nosso sistema conceptual, modos naturais de pensar e de falar, tanto na linguagem corrente como no discurso científico, radicados na experiência humana e responsáveis quer pela estruturação do pensamento, da linguagem e da acção, quer pela inovação conceptual. Esta deslocação para o plano do sistema conceptual de fenómenos tradicionalmente identificados na linguagem e relegados para um nível anormal e este reconhecimento da naturalidade e ubiquidade

⁴ Comunicação apresentada no XVII Congresso Internacional da ALFAL, em julho de 2014, na UFPB.

Janderson Lemos de Souza

do pensamento metafórico e metonímico enformam a teoria cognitiva contemporânea da metáfora e da metonímia (...)” (SOARES da SILVA, 2006: 111)

É por isso que:

“Cognitive Linguistics approaches to word-formation have investigated (and will most likely continue to do so) metaphoric and metonymic concept extensions (...), figure-ground alignment (...), schematization (...), conceptual fusion or integration in terms of the blending framework (...), and form-meaning iconicity (...), to mention some persistent topics. In any case, the ultimate goal of the proposed semanticization of word-formation research under the aegis of Cognitive Linguistics is ‘to provide more comprehensive and consistent descriptions of individual word-formation phenomena’ (Ungerer 2007: 651).” (LAMPERT & LAMPERT, 2010: 31)

Dado que o foco na polissemia é constitutivo do quadro teórico e que sua concepção envolve diretamente a metáfora e a metonímia, defendo que a heterossemia também seja considerada um mecanismo organizadora da polissemia.

Antes de passar a essa defesa, retomo algumas indagações formuladas por Onysko & Michel (2010) tendo em vista apresentar um esboço de resposta afirmativa a uma delas:

“(...) two immediate questions arise as soon as word formation knocks on the door of research on language and cognition. First of all, are the theoretical frameworks that structure the field of cognitive linguistics capable of investigating processes of word internal structuring? Secondly, and related to the first question, what kinds of attempts have been made so far to explain word formation from a cognitive vision of language?” (p. 9)

Em resposta à segunda indagação, Lemos de Souza (2010, 2012) consiste em uma das tentativas de explicar a formação de palavras da perspectiva cognitivista. Dedicado à nominalização de verbos, tema clássico desde Chomsky (1970), o trabalho chega a uma generalização muito próxima à que esses autores propõem na citação anterior.

Basicamente, defendo que a formação de palavras deixe de ser concebida como um fenômeno morfológico sensível ao fator semântico e passe a ser concebida como um fenômeno semântico com repercussão morfológica. O fato de que palavras são rótulos simbólicos de categorias mentais unido ao fato de que tais rótulos são associados a um complexo de sentidos que apresenta coerência semântica graças a padrões de figuratividade me levaram a rever a relação base + afixo: a base não como uma forma, mas sim como um dos significados associados a uma forma polissêmica, e o afixo como um indicador de qual dos significados associados à forma foi escolhido pelo conceptualizador para construir um novo conceito.

Vejam alguns dados que ilustram a proposta. Em itálico, a forma. Em azul, o padrão de figuratividade incidente. Em vermelho, o(s) significado(s) com que o afixo se relaciona e o novo rótulo formado a partir dessa relação:

	término, fim	⇒	[a[<i>cabo</i>]ar]
	↑ metáfora		
(1) <i>cabo</i> – s.m.:	cabeça, ponta		
	↓ metáfora		
	cargo militar		
	↓ metáfora		
	comandante, chefe		

Cadernos do NEMP, n. 6, v. 1, 2015, p. 67-81.

Chunks em moldes: hipótese de diálogo entre Bybee e Booij

- (2) *claro* – adj.: propriedade física ⇒ [[claro]idade] [[claro]ear]
 ↓ metáfora
 propriedade abstrata ⇒ [[claro]eza] [es[claro]ecer]
- (3) *executar* – v.: agir, empreender ⇒ [[executar]ivo]
 ↓ metáfora
 matar ⇒ [[executar]or]
- (4) *balada* – s.f.: música
 ↓ metonímia
 diversão noturna ⇒ [[balada]eiro]
- (5) *lixo* – s.m.: coisa ⇒ [[lixo]eiro]
 ↓ metonímia
 recipiente
- (6) *jornal* – s.m.: coisa ⇒ [[jornal]eiro]
 ↓ metonímia
 instituição ⇒ [[jornal]ista]

Os dados de (1) a (6) devem ser tomados como evidência de que nem todos os significados associados a uma forma são selecionados para se combinarem a um afixo. Essa disponibilidade de significados aponta para a possibilidade de novas combinações, assim como a disponibilidade dos afixos, sem o entitativismo que busca afinidade semântica entre formas.

Os dados também revelam que mais de um significado associado a uma forma pode ser selecionado e cada um dar origem a um conceito e um rótulo novos. Se os afixos se combinassem com as formas, incidiria o bloqueio, concebido por Aronoff (1976) como uma restrição formal. O dado em (6), por exemplo, permite afirmar, do ponto de vista da linguística cognitiva, que o jornal que está em *jornalista* não é o jornal que está em *jornaleiro*. Longe de ser uma afirmação de homonímia – ou novamente a relação seria formal –, trata-se de significados diferentes, organizados pela metonímia que incide graças à contiguidade entre a coisa jornal e a instituição jornal.

Produtividade e recursividade, conceitos caros ao lexicalismo gerativo, são redefinidos, consequentemente. O primeiro passa a ser concebido como essa disponibilidade de significados associados a uma forma e, portanto, permanece no terreno na potencialidade, distinto do da produção, concernente à rentabilidade das combinações. O segundo deixa de ser concebido como a ativação infinita de algoritmos a partir da presença de uma categoria formal como *input* e como *output* num processo na acepção técnica já discutida e passa a ser concebido como a possibilidade infinita de seleção de um ou mais de um dos significados associados a uma forma polissêmica para a formação de uma nova unidade simbólica. A infinitude que caracteriza a recursividade é preservada graças ao fato de que a nova unidade simbólica também goza da forte vocação à polissemia que a linguagem ordinária lhe confere.

Janderson Lemos de Souza

De volta à polissemia e aos mecanismos que a organizam, o foco inicial nos substantivos deverbais me levou a perceber que a polissemia pode apresentar um caráter predeterminado, que não se pode atribuir nem à metáfora nem à metonímia. Basilio (2004) entende que se trata de “polissemia sistemática”:

“Chamamos de polissemia sistemática a multiplicidade de interpretações de caráter pré-determinado numa forma lingüística. Assim, nas formações lexicais, a polissemia sistemática é uma estratégia valiosa para a utilização de um determinado elemento no exercício de várias funções interligadas. No caso das nominalizações deverbais, a polissemia sistemática corresponde a diferentes instâncias de uso freqüente e necessário, todas relacionadas a um processo global de desverbalização do verbo.” (p. 56)

A autora se refere ao fato de que substantivos deverbais exibem um significado verbal e um significado nominal, conforme já havia identificado em trabalho anterior:

“(…) tanto nomes morfológicamente básicos quanto nomes deverbais podem apresentar tanto uma interpretação nominal quanto uma interpretação verbal. As várias interpretações de formas nominalizadas foram devidamente enfatizadas na literatura por constituírem um problema crucial para a hipótese transformacionista. Assim, o fato de que formas nominalizadas de verbos podem apresentar várias extensões de significado não constitui novidade alguma.” (BASILIO, 1980: 81)

No quadro gerativista, o fato poderia ser explicado por serem as categorias formais V e N as envolvidas no processo. No quadro cognitivista, entendo que o fato possa ser explicado por serem as categorias semânticas verbo e substantivo as envolvidas no processo, pressupondo que ser verbo morfológicamente seja consequência de ser verbo semanticamente, assim como ser substantivo morfológicamente seja consequência de ser substantivo semanticamente. Nesse quadro, a metáfora e a metonímia delineiam o percurso da figuratividade na composição da “coerência semântica num complexo de sentidos associados a uma mesma forma” mas não predeterminam as categorias dos significados porque incidem depois de eleitas as categorias pelo próprio “processo de formação de palavras”.

Vialli (2013), ao tratar da reduplicação, que concebo como a via icônica da nominalização de verbos dado que a iteratividade semântica se reflete na repetição morfológica, reencontra as “interpretações de caráter pré-determinado” percebidas por Basilio (2004) e as atribui ao fenômeno da heterossemia, tal como identificado por Booij (2010):

“De acordo com Lichtenberk (1991), a heterossemia é um fenômeno no qual um dado elemento morfológico apresenta significados diferentes, porém relacionados, que se distinguem a depender do contexto gramatical em que são aplicados. No caso da reduplicação de base verbal, os nomes resultantes do processo possuem significados diferentes, mas não são homônimos, já que têm a mesma origem.

Logo, a heterossemia se caracteriza, além da polissemia, pela diversidade não apenas semântica, mas também sintática. No caso da reduplicação de base verbal, como se trata de um processo nominalizador, os argumentos verbais serão apagados tanto nos nomes que expressam evento quanto naqueles que expressam o sentido de coisa. Contudo, a distribuição sintática dos nomes (evento e coisa) resultantes do processo de reduplicação é diferente, sobretudo no sentido de que: (a) as formas com ideia de evento podem apresentar papéis temáticos evocados pela construção do verbo-base do processo e que podem aparecer na expressão nominal sob a forma de adjuntos adnominais e/ou complementos nominais, evidenciando a existência de posição argumental, que exige papel de participante; (b) na aceção de coisa, a nominalização é tamanha que os papéis temáticos não serão evocados de forma alguma.” (p. 102-103)

Chunks em moldes: hipótese de diálogo entre Bybee e Booij

Para Booij (2010), as construções atuam como moldes que predeterminam as interpretações dos itens polissêmicos. Para Vialli (2013), os modos de escaneamento, dinâmico e estático, propostos por Langacker (1987) predeterminam que as interpretações das reduplicações sejam de evento e de coisa. Para Lemos de Souza (2010), a conversão de escaneamento dinâmico em escaneamento estático é a motivação semântica da nominalização de verbos, e não o fator que distribui as interpretações verbais e nominais dos substantivos deverbais, como em Vialli (2013). Para aquele, essa distribuição se caracteriza como polissemia sistemática, enquanto, para esta, é obra da heterossemia.

Mas qual a definição de heterossemia que está na base desta discussão? Segundo Lichtenberk (1991), a heterossemia se aplica a “(...) casos (dentro de uma dada língua) em que dois ou mais significados ou funções que são historicamente relacionados, no sentido de derivarem da mesma fonte última, são sustentados por reflexos dessa fonte que se encontram em diferentes categorias morfossintáticas” (p. 476). Na leitura cognitivista do conceito, os significados predeterminados, ao mesmo tempo em que são sincrônica e diacronicamente associados, distribuem os substantivos deverbais em diferentes categorias morfossintáticas.

Sendo assim, mantenho a posição de que a conversão de um escaneamento em outro é a motivação semântica de nominalizar verbos e entendo que, assim como a metáfora e a metonímia, a heterossemia organiza a polissemia, já concebida por Lakoff (1987) como categorização de significados⁵, só que dentro dos limites estabelecidos pelas categorias envolvidas pela formação em questão. Dessa forma, a heterossemia passa a ser vista como o mecanismo que produz a polissemia sistemática, distinta da polissemia tal como apresentada por Lakoff (1987) e Soares da Silva (2006) por ter uma sistematicidade que reflete as categorias semânticas envolvidas no processo.

Como a nominalização consiste em “um processo global de desverbalização do verbo” e a reduplicação é, junto com a derivação sufixal e a derivação regressiva, uma das vias da nominalização, não acompanho o entendimento de que se dê “(...) a seleção de uma forma que coincide com o tema verbal, a terceira pessoa do singular do presente do indicativo, base para o processo de cópia que a reduplicação demanda” (Vialli, 2013: 117). Essa posição acompanha a de Gonçalves & Almeida (2008): “Em português, a reduplicação é um fenômeno responsável pela nominalização: a repetição de uma forma verbal na 3ª pessoa do singular do presente do indicativo (...)” (p. 40).

Entendo que o elemento duplicado seja o próprio tema por ser a porção formal isenta de marcas. Afinal, porque “(...) categorias como tempo, modo e número-pessoa, assim como transitividade, são propriedades fundamentais do verbo e não do substantivo, a mudança de classe de verbo para substantivo desfaz a obrigatoriedade de especificar tempo e modo verbais e explicitar agentes e objetos ou beneficiários do processo verbal.” Assim, “(...) a nominalização do verbo nos permite expressar a noção verbal em si, sem as amarras dos requisitos gramaticais do verbo” (Basilio, 2009: 42).

A discussão sobre qual é o elemento copiado na reduplicação é bem mais complexa do que posso contemplar aqui. Certamente, existem argumentos a favor de ser a terceira pessoa.

Como forma de corroborar a adequação de incluir a heterossemia como mecanismo que organiza a polissemia, dois fatos já apontados pela literatura passam a poder ser tratados de modo coerente com a linguística cognitiva, em geral, e com a concepção de formação de palavras aqui exposta, em particular. O primeiro é que:

⁵ “(...) polysemy appears to be a special case of prototype-based categorization, where the senses of the word are the members of a category. The application of prototype theory to the study of word meaning brings order into an area where before there was only chaos.” (p. 378)

Janderson Lemos de Souza

“A formação de substantivos a partir de verbos surge também da motivação gramatical de possibilitar o uso em estruturas que sintaticamente exigem um substantivo. Um exemplo bem característico desse tipo de motivação é o das nominalizações em expressões com verbos de suporte, conforme ilustrado abaixo:

- (7)
- a. Maria deu um grito e desmaiou.
 - b. João vai fazer uma declaração de voto.
 - c. Maria deu uma saída rápida.
 - d. A diretora fez uma observação interessante.

Em todos os casos de (7), usa-se uma expressão com um verbo de suporte em lugar da forma verbal. A construção perifrástica exige um verbo de suporte e a forma nominalizada do verbo correspondente. Temos, portanto, um requisito gramatical da estrutura da expressão perifrástica.” (BASILIO, 2009: 41)

A releitura dessa proposição à luz de outro quadro teórico exige uma série de cuidados. Sem simplificar suas especificidades, a releitura consiste em dizer que o substantivo deverbal com interpretação verbal atua na sentença como o verbo semântico. No entanto, a condição morfológica de substantivo leva à condição sintática de ocupante de argumento, o que exige a convocação de um verbo morfossintático com argumento para o verbo semântico ocupar. Ser um verbo estritamente por manifestar as categorias modo, tempo, número e pessoa e por atuar na sentença em conjunto com um verbo semântico nominalizado é o que a autora chama verbo de suporte.

O outro fato se extrai da análise de dois dados:

“O caso de *passadeira* já é um pouco diferente; podemos prever o significado ‘mulher que tem a profissão de passar a ferro’, se tomarmos o verbo *passar* com esta acepção; mas o significado do ‘tapete relativamente comprido que se coloca em lugares de passagem’, embora relacionado à base *passa-*, radical de *passar*, é imprevisível. O caso de *passadeira* é, portanto, um caso intermediário entre a total regularidade e a total irregularidade de conexão semântica entre o significado ou uso que sua estruturação morfológica faria prever ou esperar.

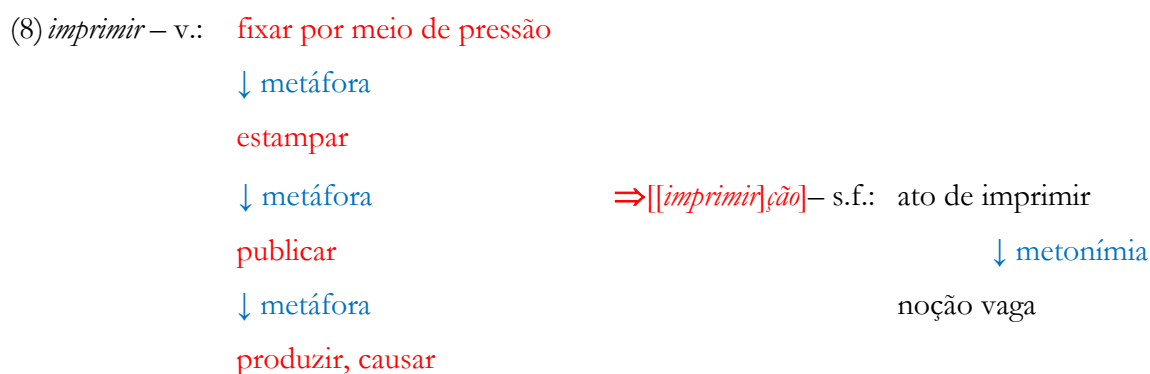
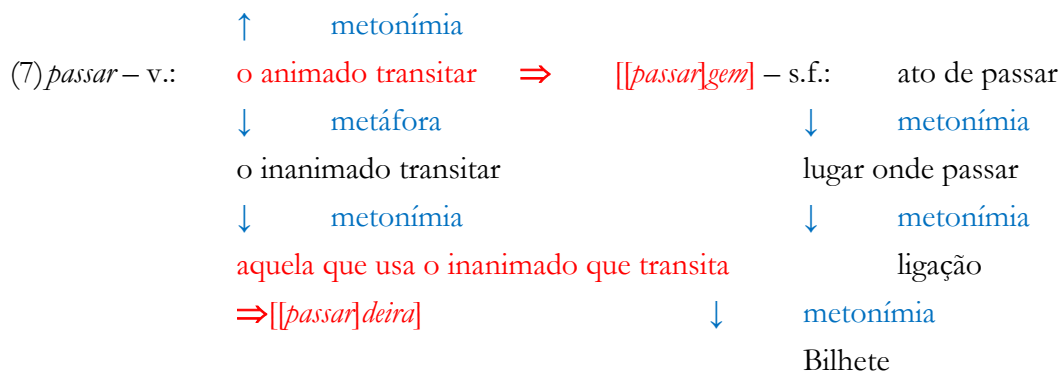
Um exemplo deste último caso seria o de *impressão*, onde o significado ‘noção vaga’ nada tem a ver com o significado da base verbal *imprimir*.” (BASILIO, 1987: 20-21)

Novamente se impõe o cuidado de rever uma análise realizada há quase três décadas dentro do estado em que se encontrava o lexicalismo gerativo. O objetivo incidental continua sendo elucidar como temas clássicos são tratados mediante o compromisso com a linguística cognitiva de modo a evitar que o tratamento estabilizado na literatura interfira na leitura deste artigo.

A análise de *passadeira* e *impressão* aponta para graus de previsibilidade entre a estrutura morfológica e certos significados. Dispenseemos a ambos os dados o tratamento proposto por Lemos de Souza (2010, 2012), ressalvado, assim como para os dados de (1) a (6), que a identificação de cada padrão de figuratividade mereceria fundamentação, o que foge aos limites deste artigo:

Chunks em moldes: hipótese de diálogo entre Bybee e Booij

tapete que fica onde o animado transita ⇒ *[[passar]deira]*



Os dados em (7) e (8) ilustram o fato já mencionado de a nova unidade simbólica (*passagem* em relação a *passar*, *impressão* em relação a *imprimir*) também ser polissêmica, o que abre espaço para a recursividade. O que há de específico em (8) é todos os significados associados à forma serem selecionados pelo mesmo afixo para a formação da nova unidade simbólica: *impressão*.

Basilio (1987) situa um dos significados associados a *passadeira* e um dos significados associados a *impressão* em diferentes graus de regularidade. O quadro gerativo se caracteriza como imanentista por pretender isolar o significado literal do significado figurado e o conhecimento linguístico do conhecimento enciclopédico (cf. Lemos de Souza, 2010, cap. 4). Como se pode ver acima, os significados considerados não totalmente regulares resultam da incidência de padrões de figuratividade, o que causa estranheza à concepção de formação de palavras como fenômeno morfológico num quadro imanentista como a linguística gerativa mas não à concepção de formação de palavras como um fenômeno semântico num quadro não imanentista como a linguística cognitiva. Em outras palavras, o que pode surpreender são os significados específicos, devidos à participação do conhecimento de mundo, e não a incidência dos padrões de figuratividade.

Afora as condições específicas de cada fato, a evidência que pretendo apontar com a menção a eles é de que “a heterossemia se caracteriza, além da polissemia, pela diversidade não apenas semântica, mas também sintática”. Os significados verbais e nominais associados aos substantivos deverbais caracterizam uma polissemia restrita pela heterossemia às categorias semânticas envolvidas na conversão de verbos em substantivos, enquanto a distribuição dos significados deflagra diferentes configurações sintáticas e semânticas: a atuação de verbos de suporte associada aos significados verbais, e a atuação da metonímia associada aos significados nominais a serviço da nomeação – neste caso, do tapete em *passadeira* e da noção vaga em *impressão*.

3. O CASO DE *ENTORNO*

O estudo sobre a palavra “entorno” se encontra em Rossinholo (2013). Conforme anunciado como objetivo deste artigo, retomo-o para apontar a atuação do *chunking* previsto por Bybee (2010), abordagem que considero enfocada na dimensão ativa do léxico – a dos processos –, e a atuação de um dos moldes previstos por Booij (2010), abordagem que considero enfocada na dimensão passiva do léxico – a dos produtos.

Para Bybee (2010):

“The underlying cognitive basis for morphosyntax and its hierarchical organization is the chunking of sequential experiences that occurs with repetition (...). Chunking has been identified as a process influencing all cognitive systems, based on the general organization of memory. As Newell 1990 put it:

A chunk of memory organization, formed by bringing together a set of already formed chunks in memory and welding them into a larger unit. Chunking implies the ability to build up such structures recursively, thus leading to a hierarchical organization of memory. Chunking appears to be a ubiquitous feature of human memory (p.7).”(p. 34)

Com isso, a autora retoma um conceito que já frequentava a descrição gramatical em língua inglesa, conceito que remete não somente a outros (Cf. “*chunk*” em Crystal 1997 e Richards, Platt & Platt 1992), como também ao campo do processamento da linguagem.

Segundo Bybee (2010), os *chunks* podem abrigar unidades que ocorrem também em outras construções e unidades que ocorrem somente nos *chunks*. Em português, me parecem exemplos do primeiro tipo *dar corda*, *à vista* e *a prazo*, entre tantos, e exemplos do segundo *dar trela* (somente em PE, *trela* é usado livremente, como sinônimo de *coleira* em PB), *à beça*, *por um triz* e *à toa*, como o *in spite of* descrito pela autora.

O que sobressai dos exemplos de *chunks* indicados pela autora é que não há nenhuma propriedade formal que os identifique. O que os caracteriza é o fato de serem unidades rotinizadas pelo uso e armazenadas como tais em prol da própria organização da memória. A hipótese de diálogo entre este modelo e o de Booij (2010) se esboça a partir da constatação de que *chunks* podem ser, não só mas também, palavras, e são exatamente esses os que se enquadram nos moldes previstos pelo autor.

Assim como o conceito de *chunk* frequentava a descrição gramatical em língua inglesa antes de Bybee (2010), a percepção de que unidades menores se organizam em unidades maiores está na base do próprio conceito de geratividade, que dá nome à linguística gerativa, e já estava presente no estruturalismo. Novamente, o que afasta a abordagem proposta por Bybee (2010) dessas outras é a combinação de motivação semântica com frequência de uso como pressuposto. Reconhecer a anterioridade de descrições que reconheçam o fenômeno não reduz o ineditismo da descrição aqui adotada simplesmente porque correspondem a epistemologias distintas.

A título de tal reconhecimento, a definição de *chunk* em Bybee (2010) também remete à de *lexia* no estruturalismo:

“Reconhece Pottier que a palavra deve ser conceituada e reconhecida, de acordo com o conjunto seqüencial de combinações, em cada língua. Trata-se de uma unidade transitória até chegar ao nível superior: o nível *funcional*.

A *unidade funcional*, caracterizada pela relação entre o elemento e a classe, denomina-se *lexia*. É uma “*unidade lexical memorizada*”.

Chunks em moldes: hipótese de diálogo entre Bybee e Booij

Observe-se que *banana-da-terra*, *lexia* (unidade lexical memorizada), traduz a idéia de um tipo de banana. Não se pensa nos elementos separadamente.

Uma sequência chega a uma *lexia* por um processo lento de lexicalização. É o resultado de um hábito associativo.”(FREITAS, 1997: 74)

Novamente, o papel da memória, como em Bybee (2010). No entanto, as *lexias* se restringem aos tipos de sintagma, o que ancora a memorização em propriedades formais, o que não ocorre na proposta da autora.

Desde Langacker (1987), o *status de unidade* é atribuído à rotinização, tomada como responsável por tornar opaca a estrutura interna das unidades. Os estudos desenvolvidos pelo NEMP sobre processos não concatenativos de formação de palavras, aos quais remeto o leitor dada a impossibilidade de recapitulá-los todos aqui, podem ser invocados como fundamento para afirmar que especialmente o cruzamento vocabular e a substituição lexical só são viáveis exatamente porque o falante acede à estrutura interna das unidades.

Bybee (2010) retoma a discussão acerca do grau de acesso do falante à estrutura das unidades, que, já em Langacker (1987), corresponde ao conceito de *analísabilidade*, e a ele atribui a propriedade *gradiência*, o que dissolve a oposição transparência/opacidade tomada em termos absolutos:

“Analysability, according to Langacker 1987: 292, is the ‘recognition of the contribution that each component makes to the composite conceptualization’. Analysability would include the language user’s recognition of the individual words and morphemes of an expression as well as its morphosyntactic structure. This measure is also gradient and would relate to the extent to which the parts of an expression activate the representations of these parts.” (Bybee, 2010: 45)

Como evidência da atenção dos linguistas ora para a dimensão ativa ora para a dimensão passiva do léxico, a autora, em trabalho anterior, havia proposto dois tipos de esquema, o que me levou a estabelecer a relação entre sua proposta e a de Booij (2010):

“Pinker and Ullman (2002: 458), major proponents of the Words-and-Rules theory, claim that their approach ‘is compatible with constraint – and construction-based theories of language, as long as they allow for variables and combinatorial operations’ – as does Booij’s constructional morphology (...). But for other construction grammarians, especially for those following Langacker, constructional schemas are definitely not rules but symbolic units themselves.

Bybee (1995, 2001: 126-128), on the other hand, introduces two types of schemas (with both types being ‘templatic’ in nature, source-oriented schemas and product-oriented schemas: While the former are obviously the equivalents of the classical generative ‘rules’, the latter rather correspond to Langacker’s constructional schemas, and they are also the prevalent ones in Bybee’s (1995: 428).” (LAMPERT & LAMPERT, 2010: 39)

Afora a assunção, nada trivial, de que tanto a cognição modular quanto a cognição conexionista favorecem gramáticas organizadas por construções, os esquemas de Booij (2010) me parecem coincidir com os “product-oriented schemas” de Bybee (1995, 2001).

Dentre os moldes propostos pelo autor, um deles é o da prefixação, assim formalizado: [X[Y]_y]_y. O molde prevê rótulos categoriais para a construção menor (o “y” interno) e para a construção maior (o “y” externo), o que atende à descrição de unidades como *anteprojeito* [ante[projeto]_s]_s.

No entanto, substantivos deverbais que apresentam um prefixo mas são formados por derivação regressiva, como *retorno* e *contorno*, já não recaem no molde, uma vez que a forma *torno*

Janderson Lemos de Souza

em questão não corresponde ao substantivo *torno*, e sim a uma sequência que apresenta a marca *-o* de formação a partir de *retornar* e *contornar*, respectivamente:

“Observem, por exemplo, os pares *apertar/aperto*, *ameaçar/ameaça* e *cortar/corte*. Nesses pares, vemos que a vogal pela qual termina o nome pode ser *-a*, *-e* ou *-o*. Assim, se considerarmos que esses verbos são de derivação regressiva, pelo menos teremos que considerar que se trata de um caso misto, pois também ocorre o acréscimo das vogais.” (BASILIO, 1987: 39)

O caráter misto a que se refere a autora diz respeito à possibilidade de descrever os substantivos deverbais como sufixados. No entanto, essa opção descritiva apagaria a distinção, por exemplo, entre *espera* e *esperança* a partir de *esperar*, visto que ambos os sufixos deverbais são descritos como formados por sufixação, igualando-se a, por exemplo, *salvação* e *salvamento* a partir de *salvar*.

Embora essa seja uma opção defensável, a indistinção entre derivação regressiva e derivação sufixal me parece uma perda descritiva em favor da preservação dos instrumentos de análise. A alteração dos moldes me parece a opção mais adequada aos dados.

Nesse caso, o substantivo *entornose* distingue porque nele a sequênci*atorno*corresponde ao substantivo *torno*, a forma *analísável* à esquerda é que não é um prefixo. No entanto, seu grau de analisabilidade é tão transparente quanto o de *retorno* e *contorno*, o que me leva a propor que o rótulo categorial interno no molde seja eliminado: [X[Y]]_y. Ressalte-se que *entorno* poderia ser a nominalização de *entornar*, caso em que teríamos, assim como *retorno* e *contorno*, produtos de regras sincrônicas. Contudo, trata-se de um sintagma preposicional (*em torno*) convertido em palavra (*entorno*), um *chunk* sintático convertido em *chunk* morfológico, caso em que temos um processo diacrônico, o *chunking*. Afinal, “*There is no fundamental distinction between morphological and syntactic constructions, which are fully parallel in all immediately relevant aspects*” (Langacker, 1987: 82).

Dessa forma, teríamos [re[torno]]_s, [con[torno]]_s e [en[torno]]_s, numa descrição que concilia o *chunking* como processo de formação de *entorno* com os moldes revelados a partir da análise estrutural das unidades simbólicas da língua. Essa proposta de descrição se baseia na concepção de *entorno* como um *chunk* que corresponde a uma construção morfológica tomada como molde, considerada a lexicalização do SPrep “em torno” na unidade simbólica “entorno” como a rotinização de um *chunk* da sintaxe como um *chunk* do léxico.

CONSIDERAÇÕES FINAIS

Este artigo se dedica à oportunidade de cotejar processos diacrônicos com produtos sincrônicos por meio de dois modelos: o proposto por Bybee (2010), que destaca o *chunking* como um desses processos, e o proposto por Booij (2010), que propõe moldes para as construções gramaticais, que seriam as reais unidades de análise da morfologia. Dentre os *chunks* previstos pela autora, estão unidades que correspondem às construções previstas pelo autor, nomeadamente as que costumam ser reconhecidas como palavras.

Por sua vez, o conceito é discutido por linguistas de diferentes convicções teóricas, que reconhecem a necessidade de distinguir entre palavras que compõem o léxico de uma língua e palavras que podem vir a compor mediante a ativação de diferentes processos. O processo de formação de verbos a partir de substantivos como objeto de pesquisa do NEMP foi destacado como ilustrativo do papel da polissemia na linguística cognitiva, o que abriu espaço para a função da heterossemia.

Chunks em moldes: hipótese de diálogo entre Bybee e Booij

Por fim, a formação *entorno* foi apresentada como unidade que pode, diacronicamente, ser descrita como resultado de *chunking* e, sincronicamente, ser descrita como construção gramatical, o que, por sua vez, exigiria a adaptação de um dos moldes.

REFERÊNCIAS

- ALMEIDA, Maria Lucia Leitão de. Aqui é hoje: a dimensão espaço-temporal em português – o caso das preposições, advérbios e conjunções. In: Jürgen Heye (Org.). *Flores verbais*. Rio de Janeiro: Editora 34, 1995.
- ARONOFF, Mark. *Word formation in generative grammar*. Cambridge/Massachusetts: The MIT Press, 1976.
- BASILIO, Margarida. *Estruturas lexicais do português*. Petrópolis: Vozes, 1980.
- BASILIO, Margarida. *Teoria lexical*. São Paulo: Ática, 1987.
- BASILIO, Margarida. Observações sobre a conceituação de “formação”, “regra” e “palavra” na expressão “regras de formação de palavras”. In: André Valente (Org.). *Língua, linguística e literatura: uma integração para o ensino*. Rio de Janeiro: EdUERJ, 1998.
- BASILIO, Margarida. Polissemia sistemática em substantivos deverbais. In: Roberta Pires de Oliveira & Apóstolo Nicolacópulos (Org.). *Semantics: lexicon, grammar and use*, nº 47. Florianópolis: Editora da UFSC, p. 49-71, 2004.
- BASILIO, Margarida. *Formação e classes de palavras no português do Brasil*. São Paulo: Contexto, 2009.
- BASILIO, Margarida. Abordagem gerativa e abordagem cognitiva na formação de palavras: considerações preliminares. In: *Linguística*, vol. 6, nº 2, dezembro de 2010.
- BLOOMFIELD, Leonard. *Language*. Chicago: The University of Chicago Press, 1933.
- BOOIJ, Geert. *Construction morphology*. Oxford: Oxford University Press, 2010.
- BYBEE, Joan. *Morphology: a study of the relation between meaning and form*. Amsterdam/Philadelphia: John Benjamins, 1985.
- BYBEE, Joan. *Phonology and language use*. Cambridge: Cambridge University Press.
- BYBEE, Joan. *Language, usage, and cognition*. Cambridge: Cambridge University Press, 2010.
- CHOMSKY, Noam. Remarks on nominalization. In: *Readings in English transformational grammar*. Waltham: Ginn & Co, 1970.
- CRYSTAL, David. *Dictionary of linguistics and phonetics*. 4 ed. Massachusetts: Blackwell, 1997.
- FREITAS, Horácio Rolim de. *Princípios de morfologia*. 4 ed. Rio de Janeiro: Oficina do Autor, 1997.
- GONÇALVES, Carlos Alexandre & ALMEIDA, Maria Lucia Leitão de. Das relações entre forma e conteúdo nas estruturas morfológicas do português. In: Carlos Alexandre Gonçalves & Maria Lucia Leitão de Almeida (Org.). *Diadorim*, vol. 4. Rio de Janeiro: UFRJ, 2008.
- JACKENDOFF, Ray. Morphological and semantic regularities in the lexicon. In: William Bright et alii (Ed.). *Language: Journal of the Linguistic Society of America*, vol. 51, nº 3. Baltimore: Waverly Press Incorporate, p. 639-671, 1975.
- LAKOFF, George. *Women, fire, and dangerous things: what categories reveal about the mind*. Chicago: The University of Chicago Press, 1987.

Janderson Lemos de Souza

LAMPERT, Martina & LAMPERT, Günther. Word-formation or word formation? The formation of complex words in cognitive linguistics, In: Alexander Onysko & Sascha Michel (Ed.). *Cognitive perspectives on word formation*. Berlim/Nova York: Walter de Gruyter, 2010.

LANGACKER, Ronald. *Foundations of cognitive grammar: theoretical prerequisites*. Vol. I. Stanford: Stanford University Press, 1987.

LEMONS DE SOUZA, Janderson. *A distribuição semântica dos substantivos deverbais em -ção e -mento no português do Brasil: uma abordagem cognitiva*. UFRJ, tese de doutorado, 2010.

LEMONS DE SOUZA, Janderson. Alinhamento, balanceamento, direção e suspensão: constatações sobre os substantivos deverbais no português brasileiro. In: *Linguística*, vol. 8, nº 1, junho de 2012.

ONYSKO, Alexander & MICHEL, Sascha. Unravelling the cognitive in word formation. In: Alexander Onysko & Sascha Michel (Ed.). *Cognitive perspectives on word formation*. Berlim/Nova York: Walter de Gruyter, 2010.

RICHARDS, Jack; PLATT, John & PLATT, Heidi. *Dictionary of language teaching & applied linguistics*. 2 ed. Essex: Longman, 1992.

ROBINS, Robert Henry. *A short history of linguistics*. 4 ed. London: Longman, 1997.

ROSSINHOLE, Giovanna. As ocorrências de "em torno" e "entorno" no português brasileiro a partir de uma perspectiva cognitivista. In: Carlos Alexandre Gonçalves (Org.). *Anais do II Colóquio Brasileiro de Morfologia*. Rio de Janeiro: UFRJ, 2013.

SOARES DA SILVA, Augusto. *O mundo dos sentidos em português: polissemia, semântica e cognição*. Coimbra: Almedina, 2006.

VIALLI, Luciana. *Reduplicação de base verbal: uma análise pela morfologia construcional*. UFRJ, tese de doutorado, 2013.

CHUNKS IN TEMPLATES: POSSIBLE DIALOGUE BETWEEN BYBEE AND BOOIJ

Abstract: *In this article, I propose the intersection of two theoretical paths: the one that Bybee (2010) formulates on the assumption that chunking is a usage-sensitive cognitive process, and the one where Booij (2010) re-enforces the assumption that morphological constructions are schemas. Then I take the formation of "entorno" in Brazilian Portuguese, studied by one of my undergraduate supervisees from the semantic perspective, for evidence that among all sorts of chunks described by Bybee those that correspond to words fall under one of the schemas described by Booij.*

Key-words: *Morphology; Semantics; Cognitive Linguistics.*