

Artigos

Em torno da palavra como unidade lexical: Palavras e composições

Margarida Basilio*

Abstract

In this work the concept of word is discussed as a lexical and as a structural unit and clarified by analyses of borderline cases of compounding and prefixation. The concept of compounding is discussed in the description of Brazilian Portuguese and the V+S process of compounding is analyzed in detail, in order to illustrate the structural aspects involved in the notion of compounding.



conceito de palavra é de grande dificuldade em morfologia, dadas as múltiplas dimensões em que esta unidade pode ser enfocada, as quais nem sempre coincidem, além de

* Pontifícia Universidade Católica do Rio de Janeiro.

apresentarem diferentes graus de relevância. Na palavra, entendida como uma unidade lexical, uma sequência fônica se associa de modo relativamente estável a (a) um significado ou conjunto de significados; (b) um conjunto de propriedades sintáticas; (c) um conjunto de propriedades morfológicas; e (d) um conjunto de determinações de uso.¹

Na Gramática Tradicional, o conceito de palavra não é colocado em questão, o que é natural, pois no modelo clássico de descrição gramatical, posteriormente denominado Palavra e Paradigma, a palavra é a unidade mínima de análise linguística.²

No estruturalismo, no entanto, dado o surgimento teórico e a ênfase do morfema como unidade básica da morfologia, não apenas a palavra deixou de ser a unidade mínima de análise linguística; mais do que isso, deixou de ser uma unidade relevante da estrutura da língua. Com efeito, a partir da adoção do modelo Elemento e Arranjo, a identificação de morfemas (elementos) passa a ser o objetivo da análise morfológica, permanecendo as combinações de elementos como objeto de estudo da sintaxe. Perde-se, pois, a clareza sobre a razão de ser da palavra como unidade estrutural. Esta é uma situação complexa, na medida em que o âmbito da morfologia é delimitado pela palavra enquanto unidade estrutural. Do mesmo modo, dependem deste conceito as divisões internas da morfologia (derivação, flexão e composição).

A definição de palavra de Bloomfield, bastante difundida, é uma tentativa de resolver o problema criado no estruturalismo a partir do conceito de morfema, o qual coloca a palavra numa posição inconfortável, entre o morfema e o sintagma oracional. Segundo Bloomfield, a palavra é a forma livre mínima: uma forma que pode ocorrer isoladamente, por si só constituindo um enunciado, e não podendo ser totalmente subdividida em formas livres.³ Assim, sendo livre, a palavra se distingue dos morfemas presos, radicais, afixos ou clíticos; sendo mínima como forma livre, distingue-se dos sintagmas oracionais, que podem conter mais de uma forma livre.

Esta definição focaliza a palavra como unidade da estrutura do enunciado: dado um enunciado, podemos segmentá-lo em palavras a partir desta definição de potencialidades. Assim, não são abarcados, dentre outros, a dificuldade prática de reconhecer palavras eventualmente pronunciadas em sequência não interrompida no decorrer do ato de fala, nem o problema teórico de se distinguir palavras de suas diferentes manifestações de caráter flexional.

Dentro destes limites, a definição é eficiente; mas naufraga, no momento em que focalizamos a questão da composição. Como dizer que uma palavra composta não é totalmente divisível em formas livres?

Devemos observar, em favor de Bloomfield, que no estruturalismo a composição é definida não como uma construção lexical baseada em duas

1 Dadas as enormes dificuldades na definição de palavra, a caracterização acima tem apenas o objetivo de esclarecer o objeto que estamos considerando neste trabalho; a possibilidade de uma real definição do termo está longe de constituir um projeto viável a curto prazo.

2 V. Hockett, 1954. Neste modelo, a flexão ajuda a delimitar a palavra; por outro lado, apenas a partir do séc. XIX a morfologia começa a abarcar também a derivação.

3 V. Bloomfield, 1926.

ou mais palavras, como encontramos nas gramáticas normativas, mas como uma construção que contém dois ou mais radicais. Assim, poderíamos considerar que, enquanto formadores de palavras compostas, os radicais correspondentes seriam presos. Por outro lado, existe no inglês, acompanhando as construções compostas, todo um fenômeno de acentuação que interfere na situação da palavra composta.⁴

Entretanto, é patente a não aplicabilidade da definição para os compostos do português. Tomemos um caso como *sofá-cama*. Podemos realmente dizer que *sofá-cama* é uma forma mínima como forma livre? Ou seja, podemos dizer que *sofá-cama* não se subdivide em duas formas livres, *sofá* e *cama*? Na medida em que um *sofá-cama* (e *navio-escola*, *carta-bilhete*, *carta-bomba*, *bomba-relógio*, etc.) designa algo caracterizado ao mesmo tempo na primeira parte e na segunda parte da composição, não podemos dizer que a forma, enquanto parte da composição, se caracteriza de modo diferente, como presa, ou que apresenta uma mudança dramática de significado.

A questão seguinte é a de se o problema é da definição de palavra, ou do conceito de palavra composta. Do ponto de vista histórico, o problema maior deriva provavelmente do descompasso entre a definição tradicional de composto, em que o conceito de palavra é pressuposto, e a necessidade de um critério definidor de palavra gramatical dentro do quadro estruturalista. Aparentemente, a razão de podermos conviver com relativo conforto neste descompasso se relaciona com o fato de que palavra é não apenas uma unidade morfológica, mas sobretudo como unidade lexical: o léxico é via de regra definido como o conjunto de palavras de uma língua. Assim, conceptualizamos os compostos como conjuntos de palavras que funcionam lexicalmente como uma palavra só.

Neste trabalho pretendo abordar alguns casos problemáticos em torno da palavra como unidade lexical e observar a necessidade de distinguirmos as dimensões morfológica e lexical da palavra, além da dimensão gráfica, que não podemos deixar de mencionar, dada a importância quase exclusiva dada à língua escrita pela gramática normativa.

Consideremos, em primeiro lugar, a conhecida discussão da gramática normativa sobre a situação da prefixação. A maior parte dos gramáticos normativos considera a prefixação como parte da derivação, embora sempre com alguma marca de hesitação, por causa da origem histórica preposicional de muitos prefixos.⁵ Assim, além da emergência da questão sincronia/diacronia, temos dubiedade pelo fato de preposições serem, entre outras coisas, palavras gráficas – embora obviamente não sejam radicais.

Ainda na prefixação, bastante complexa é a situação de elementos marcadores de dimensão, tais como mini-, multi-, maxi-, mega-, etc.; tais elementos são, na realidade, difíceis de nomear, dada a incerteza sobre se seriam prefixos ou radicais. Neste caso se verificam problemas derivados do descompasso entre uma visão tradicional e uma visão estruturalista de fenômenos morfológicos.

4 *Idem*.

5 Para uma análise mais extensa desta discussão, v. Basílio, 1989.

Consideremos, por exemplo, a situação de *mini-*, que aparece em construções como *mini-saia*, *mini-loja*, etc. A rigor, temos a raiz *min-*, presente em formas como *menos*, *menor*, *minoria*, *minuto*, *minúcia*, *mínimo*, *diminuir*, etc.. Mas *mini-*, na forma usada hoje em dia, é desconhecido em latim; parece provir de uma redução de *mínimo* ou *miniatura*, e é uma formação recente, de certo modo internacionalizada; formas como *mini-saia* aparecem a partir da década de sessenta.⁶ Trata-se de um formativo que só ocorre prefixado, o que nos levaria a considerá-lo como prefixo; no entanto, o fato de conter o radical *min-* dá margem a diferentes análises. Cunhado mais ou menos ao mesmo tempo em oposição a *mini-*, o formativo *maxi-* apresenta características análogas.

Um terceiro problema envolvendo a prefixação é a situação de prefixos como *pre-* e *pos-*, que apresentam uma versão acentuada *pré-*, *pós-*, como em *prefixo* e *pospor*, em oposição a *rendimentos pré-* e *pós-fixados*. Embora tenhamos idêntica sequência segmental relacionada a idêntico significado, observamos ao mesmo tempo uma diferença de acentuação, relacionada a uma diferença de comportamento: enquanto *pre-* e *pos-* com a vogal média átona fechada são realmente formas presas, *pré-* e *pós-* com a vogal média aberta e acentuada, já figuram numa área acinzentada, pois, além de conservarem a tonicidade, permitem a coordenação com omissão da base na primeira forma (*pré-* e *pós-fixado*), o que os desqualifica como formas presas, embora, certamente, não os defina como formas livres.

Um quarto problema ainda nesta área é a situação do elemento *não* que modifica substantivos e adjetivos. Alguns autores o consideram como prefixo; outros incluem formações com *não-* entre os compostos.⁷

Existe uma diferença fonológica entre o *não* utilizado na negação do processo verbal e o *não* usado na modificação de verbos e adjetivos na fala coloquial mais rápida, como vemos em frases como *num quero*, *num precisa*, etc. pois a mudança *não*→*num* jamais ocorre com o uso lexical do *não*. Esta diferença, entretanto, não incide sobre a situação morfológica do *não* como modificador de substantivos e adjetivos, já que a variação fonética se limita a usos pré-verbais.

A situação em que é mais fácil caracterizar *não-* como formativo vocabular é a situação em que este elemento se combina com adjetivos, na formação de adjetivos de classificação binária, como em *não-alinhados* (países), *não-saturadas* (gorduras), *não-contáveis* (substantivos), *não-euclidiana* (geometria) *não-arredondadas* (vogais), etc. Nestes casos, o elemento *não-* se junta ao adjetivo para adicionar-lhe a alteração semântica da negação, de modo que o adjetivo passa a ser negativo. Assim, por exemplo, *substantivos não-contáveis* são substantivos subcategorizados por sua propriedade de não poderem ser contados; e assim por diante. Ou seja, o que define o estatuto de formativo vocabular do *não* é, fundamentalmente, o escopo: não se trata de predicar negativamente uma qualidade a um referente, mas de atribuir um caráter negativo ao próprio significado do adjetivo.

6 Mas algumas formas aplicadas a transporte já aparecem com este elemento em inglês no sec. XIX.V. The Random House Dictionary of the English Language. No Aurélio, registra-se que *mini-* viria de *míni(mo)*.

7 V., por ex., Alves (1987) e Cunha&Cintra (1985).

Um segundo caso em que a questão do escopo também é clara, definindo a situação vocabular do elemento *não-*, é aquele em que *não-* se combina com substantivos que têm referente humano, agente ou afetado, de modo que o todo apresenta uma função designadora de seres, como em *não-fumantes, não-sócios, não-pagantes, não-agressor*, etc.

Existem casos, por outro lado, em que o uso do *não* é mais adequadamente analisado como sintático, apesar do escopo ser nominal. Considerem-se os exemplos abaixo:

- (1) Convencidos da não pertinência da questão, os associados votaram contra sua inclusão na pauta do dia.
- (2) A não participação dos Estados Unidos enfraqueceu o impacto da proposta.
- (3) Dada a não operacionalidade do sistema, a inauguração foi transferida.

Casos como os acima sugerem que devemos considerar que o escopo do advérbio *não* inclui não apenas verbos, mas também nominalizações de verbos; dada a possibilidade de ocorrência aberta de casos de negação do fato verbal em forma nominalizada, assim como o caráter predicativo de tais ocorrências, a análise de *não* nesses casos como prefixo ou formador de composto não se sustenta.

Uma outra área vizinha à da prefixação é a de expressões de valor adverbial como *a pé, a caráter, a nado, a cavalo; de manhã, de tarde, de noite; de repente, de súbito, de frente, de lado, de costas*; e assim por diante. Nestas expressões, denominadas tradicionalmente locuções, temos uma forma dependente ligada ao nome, passando o conjunto a ter uma função adverbial de expressão de modo ou circunstância. Do ponto de vista fonológico, o elemento preposicional se junta ao elemento portador de tonicidade, formando uma sílaba átona pré-tônica no vocábulo fonológico então constituído; e, do ponto de vista sintático, o todo assim formado não possibilita intercalação de qualquer outro elemento.⁸ Assim, embora tais expressões sejam consideradas como locuções, esta análise só é plenamente adequada se nos circunscrevermos à língua escrita. Do ponto de vista estrutural, essas expressões adverbiais se configuram como palavras: o elemento considerado preposicional não pode ser descolado do outro, igualando-se a um prefixo em suas características distribucionais.

Caso semelhante, mas tomado na direção contrária, é o que temos com as formações adverbiais em *-mente*. Um advérbio *X-mente* é uma palavra gráfica, e, portanto, o formativo adverbializador *-mente* é considerado como sufixo. Advérbios *X-mente*, no entanto, se comportam em diferentes ângulos como expressões não sufixais. Em primeiro lugar, advérbios em *-mente* (por ex. *juntamente, francamente, lentamente*, etc.) são formados a partir da forma feminina de adjetivos biformes, isto é, não obedecem à generalização segundo a qual formas flexionadas não são derivantes. Em segundo lugar, advérbios

8 Para uma análise detalhada dessas expressões, v. Souza (em preparação).

em *-mente* (por ex. *cegamente, abertamente, acintosamente*, etc.) furtam-se ao processo fonológico segundo o qual vogais médias abertas passam a fechadas em posição átona pré-tônica, mantendo a vogal aberta no primeiro elemento apesar da vogal tônica do segundo. Finalmente, advérbios *X-mente* permitem a ocorrência coordenada com o elemento *-mente* aparecendo apenas uma vez numa sequência, em posição final, como em *livre e espontaneamente, cuidadosa e vagarosamente*, etc.⁹ O conjunto desses fenômenos sugere que, embora tenha havido uma relativa soldadura entre o elemento adjetivo e o adverbializador, não temos suficiente evidência para uma caracterização vocabular. É, portanto, questionável a análise de construções *X-mente* como palavras derivadas por sufixação.

Os dois últimos casos, o das locuções *de X, a X*, etc. e o dos advérbios *X-mente*, ilustram bem o problema do descompasso entre unidades estruturais e unidades gráficas: no primeiro caso, dois vocábulos gráficos correspondem a um vocábulo estrutural; no segundo, uma palavra gráfica corresponde a dois vocábulos estruturais.

Todos os casos apontados acima estão relacionados, de um modo ou outro, à afixação. Passamos, agora, a tratar da composição propriamente dita. Como vimos, a composição é definida como uma formação em que temos duas ou mais palavras ou radicais. Esta definição tradicional deixa duas áreas de indefinição: por um lado, a questão das construções existentes, em oposição aos processos de formação; por outro, a caracterização do composto enquanto palavra.

A primeira questão envolve construções às vezes arroladas pelas gramáticas como compostos. Por ex., *sempre-viva, Maria-vai-com-as-outras*, e *bentevi* são construções que apresentam mais de uma palavra ou radical, e, segundo a definição tradicional, seriam consideradas compostas. Entretanto, não correspondem a um processo lexical produtivo de composição.

É necessário ressaltar a distinção em português entre, por um lado, a composição como um processo sistemático de expansão lexical que consiste na combinação semântica de duas palavras inseridas num modelo estrutural pré-determinado para fins de formação de novas palavras designadoras de seres, eventos ou propriedades;¹⁰ e o composto que corresponde à eventual lexicalização de conjuntos de palavras, por motivos históricos de diferentes ordens. Exemplos como *guarda-roupa, sofá-cama, olho-de-sogra* e *luso-brasileiro* ilustram o primeiro caso; *sempre-viva* e *Maria-vai-com-as-outras* ilustram o segundo.

O segundo problema, a caracterização do composto como palavra, envolve a diferenciação entre palavra composta e estruturas sintáticas outras, ou seja, sintagmas não vocabulares de diversos tipos. Mais exatamente, embora a palavra composta possa ser claramente distinta da palavra simples pelo fato de conter mais de uma palavra em sua formação, ainda é necessário

9 Para uma análise detalhada desses fatores, v. Basilio 1998.

10 V. Basilio 1987.

11 Dado que este problema é circunscrito aos compostos de palavras, não levaremos em conta nesta discussão os compostos de radicais ou bases presas. Tais compostos (cf. *agricultura, radiologia*, etc.) apresentam estruturas peculiares, que não se confundem com sintagmas oracionais.

estabelecer o que distinguiria a palavra composta de outras unidades linguísticas que contenham duas ou mais palavras.¹¹ Ou seja, é necessário caracterizar a palavra composta não apenas no fato de conter mais de uma palavra em sua formação, mas também em suas propriedades estruturais de palavra, que a distinguiriam de outras unidades.

Para que possamos depreender essa estrutura, abordaremos um caso mais nítido, o das formações V+S, em que S representa o objeto de V e o todo corresponde a um ser designado pela ação de V sobre S: *guarda-roupa*, *mata-mosquito*, *porta-guardanapo*, *beija-flor*, etc. Estas construções ilustram o mecanismo da composição como processo de expansão lexical no português, dado o fato de que o significado do todo, inteiramente previsível, vai além do significado das partes, sendo uma função do processo de composição.¹²

Mais especificamente, temos um processo de composição que forma estruturas do tipo [[X]v[Y]s]s, em que o produto é um substantivo que designa um agente ou instrumento caracterizado pela ação do verbo sobre o objeto. Assim, por exemplo, um *mata-mosquito* é um profissional designado pela função do verbo *matar* sobre o objeto *mosquito*; um *guarda-roupa* é um objeto caracterizado pela função representada pelo processo verbal *guardar* em relação ao objeto *roupa*; e assim por diante.¹³

Tendo em vista que o aproveitamento da estrutura sintática V-OD é um processo lexical estável, formando construções que não apenas apresentam um significado distinto do de suas partes, mas também um comportamento sintático diferente do da estrutura sintática apenas enquanto tal, podemos registrar com certa facilidade o processo de composição e seus produtos. Assim, por exemplo, a frase (4) abaixo,

(4) *Compramos um guarda-roupa usado*

é estruturalmente absurda se não interpretarmos *guarda-roupa* como composto – isto é, a sequência em (5)

(5) V– Det – V- S(fem) – adj(masc)

é totalmente agramatical.

Vejamos agora a questão do ponto de vista das características estruturais dos produtos. Do ponto de vista fonológico, é clara a diferença de ritmo, acentuação e autonomia fonológica dos elementos envolvidos conforme se trate do composto ou da mesma sequência de palavras numa situação de não composição, conforme pode ser verificado pela leitura esclarecida dos exemplos abaixo,

(6) a João mata mosquito no verão para ajudar o prefeito, ...
b João, mata-mosquito no verão para ajudar o prefeito,

12 Nos demais casos, que também abundam em formações novas, as colocações são menos nítidas, porque surge a pergunta de se são, realmente, compostos.

13 Observe-se que a função do composto continua transparente mesmo nos casos de interferência metafórica, como *beija-flor* ou *louva-deus*.

embora uma situação realmente de oposição seja impossível, dada a diferença de classe entre substantivo e verbo, que corresponde obrigatoriamente a diferenças de distribuição.

Do ponto de vista morfológico, constata-se na formação do composto o congelamento do verbo na terceira pessoa do presente do Indicativo (ou a utilização do tema verbal), de modo que é nítida a distinção entre o composto e o não composto, em que este apresenta flexão, como vemos em (7):

- (7) a João e Maria matam mosquito no verão para ajudar o prefeito,...
- b João matou mosquito no verão passado...
- c João e Maria, mata-mosquitos no verão....

Do ponto de vista sintático, como já tínhamos observado, o conjunto estruturado V-S não se comporta como um sintagma verbal, mas como um substantivo. Por exemplo, a coordenação é impossível nas sentenças b:

- (8) a João mata mosquito e barata no verão para ajudar o prefeito...
- b João, mata-mosquito (*e barata) no verão

Do mesmo modo, não podemos intercalar algum elemento entre o elemento verbal e o elemento nominal:

- (9) a João mata muito mosquito no verão para ajudar o prefeito...
- b João mata todo o tipo de mosquito no verão ...
- c João, mata- (*muito) mosquito no verão....
- d João, mata- (*todo o tipo de) mosquito no verão...

e assim por diante.

Temos, portanto, motivos de ordem fonológica, morfológica e sintática para considerar essas construções V+S como distintas das eventuais sequências idênticas V+S correspondentes a sintagmas verbais. A estes se adiciona o fator semântico, que especifica essas construções como nomes de agente, como parte do processo de composição. Como estamos diante de uma construção que se comporta como uma palavra e que contém mais de uma palavra em sua construção, temos todos os motivos para afirmar que estamos diante de uma palavra composta.

A questão que se coloca agora é se podemos dizer que os outros casos de palavras compostas que normalmente consideramos podem realmente ser considerados como compostos do ponto de vista morfológico ou se, ao contrário, são compostos apenas do ponto de vista lexical.

Tomemos, por exemplo, um caso como *óculos escuros*. Este caso se distingue de outros como *carta branca* ou *amor perfeito*, em que o significado do todo é inteiramente diferente do significado das partes componentes. No caso de *óculos escuros*, o significado do todo está relacionado ao significado das partes, embora não se trate apenas de um adjetivo qualificando um substantivo, conforme vemos a seguir:

- (10) a os óculos devem ser mais escuros com vestidos claros, e claros com vestidos mais escuros.
b Estes óculos estão ficando escuros;
c Eu prefiro óculos escuros mais claros, para que eu possa usar até 6 h.

Os exemplos têm a finalidade de ilustrar o fato de que a expressão *óculos escuros* designa um tipo de objeto específico, apesar da transparência da função semântica dos dois elementos componentes. Em a. e b, chamamos atenção para o fato de que o adjetivo *escuro* qualifica *óculos*, sendo o resultado diferente do obtido na expressão focalizada; em c., mostramos que o adjetivo *claro* pode qualificar *óculos escuros* sem se criar uma anomalia semântica. Assim, podemos dizer que *óculos escuros* é um composto do ponto de vista lexical.

Ou seja, apesar da relação semântica entre os componentes e o todo, temos aspectos do comportamento do adjetivo *escuros* em relação ao substantivo *óculos* que mostram que existe lexicalmente a expressão autônoma *óculos escuros*, independente da eventual qualificação de escuros para quaisquer óculos. Isto, no entanto, não nos autoriza a considerarmos *óculos escuros* como um composto do ponto de vista morfológico. Um dos problemas que temos é o fato de que a concordância de gênero e número permanece. No caso, como se trata de uma palavra pluralia tantum, o adjetivo vai para o plural. Ou seja, ao contrário de mata-mosquito, em que verificamos a invariabilidade do elemento verbal, temos em *óculos escuros* concordância de número. Um segundo problema é o fato de que na língua popular existe a variante *óculos pretos*. Ou seja, embora o fator designação de um tipo específico de objeto nos leve a considerar uma possível autonomia lexical da expressão *óculos escuros*, a evidência do mecanismo de concordância e da possibilidade de substituição de um dos elementos por outro semanticamente equivalente nos desautoriza a considerar a sequência como um composto morfológico.

Vemos, portanto, que a questão da análise de sequências S+Adj como compostos ou não está longe de estar resolvida de modo satisfatório. Temos sequências que constituem claramente compostos lexicais, tais como *carta-branca*, *amor-perfeito*, etc., em que o significado do todo não é previsível do significado das partes; e temos sequências como *óculos escuros*, em que o significado do todo é parcialmente previsível do significado das partes, mas que não apresentam comportamento indubitável de um único elemento lexical, embora apresentem função global de designação. E, dada a questão da concordância, em ambos os casos não temos evidência de uma formação composta do ponto de vista morfológico. É necessária, portanto, a análise cuidadosa das várias sequências normalmente consideradas como compostos, antes que possamos determinar quais são realmente os processos de composição na língua portuguesa.

Referências Bibliográficas

- ALVES, I. *A produtividade do prefixo não- no português contemporâneo*. Ciência e Cultura 39 (11):1026-8, 1987.
- BASILIO, M. *Teoria Lexical*. Ed. Atica, São Paulo, 1987.
- BASILIO, M. *Prefixos: a controvérsia derivação/composição*. Cadernos de Linguística e Língua Portuguesa, 1:1-13. PUC-RIO, 1989.
- BASILIO, M. *Morfológica e Castilhamente: um Estudo das Construções X-mente no Português do Brasil*. DELTA 14, n. especial:15-25
- BLOOMFIELD, L. *A Set of Postulates for the Science of Language*. Language 2, 153-64., 1926. (In: Readings in Linguistics, 26-31)
- CUNHA, C. e CINTRA, L. *Nova Gramática do Português Contemporâneo*. Rio de Janeiro, Nova Fronteira, 1985
- HOCKETT, C. *Two Models of Grammatical Description* Word 10:210-31, 1954. (In: Reading in Linguistics, 26-31).
- SOUZA, J. L. *Palavra e Locução nos Estudos Linguísticos*. Tese de Mestrado, PUC-RIO (em preparação).
- The Random House Dictionary of the English Language*. 2a Edição. Random House, N.York, 1987.