

André ELISEU
Departamento de Linguística
Faculdade de Letras de Lisboa

Alina VILLALVA
Departamento de Linguística
Faculdade de Letras de Lisboa

TIRA - TEIMAS : ENTRE MORFOLOGIA E SINTAXE

1. Introdução

No presente trabalho procuramos consolidar a hipótese de existência, em Português europeu, de um processo de formação de palavras que consiste na reanálise de expressões sintáticas complexas. Vejam-se os seguintes exemplos de compostos formados por este processo ¹:

- (1) a. abre-latas
- b. corta-papéis
- c. corta-aramé
- d. limpa-chaminés

Nos exemplos apresentados, as palavras possuem uma estrutura interna que podemos representar como $[vN]_N$. Todos os compostos deste tipo têm uma estrutura binária $[v^*(X)]_N$ cujo primeiro constituinte tem uma natureza verbal. Os exemplos (2) ilustram os casos de estruturas atestadas em que o segundo constituinte do composto não é um nome:

- (2). a. fala-barato
- b. bota-fora
- c. faz-tudo

Na nossa análise consideraremos especialmente os compostos de verbo e nome, que

representam o caso típico desta classe de palavras ².

2. Hipótese da base sintáctica

A hipótese de que estes compostos são gerados por um processo de formação de palavras que consiste na reanálise de expressões sintácticas complexas parte de um certo número de observações:

(i) O primeiro constituinte é uma forma verbal flexionada na terceira pessoa singular do presente do indicativo, que é formalmente idêntica à do tema verbal do infinitivo na primeira e na segunda conjugações, mas distinta na terceira e nos chamados verbos irregulares:

(3)	a.	amola-tesouras	1ª conjugação
	b.	lambe-botas	2ª conjugação
	c.	abre-latas	3ª conjugação
	d.	faz-tudo	verbo irregular

Como se sabe, os processos derivacionais de formação de deverbais não seleccionam bases flexionadas, mas sim formas do radical (*arranhão, estremeção*), do tema do infinitivo (*brincadeira, tremedeira, frigideira*), do tema do particípio passado (*organização, perdição, demolição*) e do tema do particípio presente (*abundância, ascendência, advertência*).

(ii) Estas construções não têm estrutura morfológica³: o constituinte da direita não é o núcleo morfológico, ou seja, as propriedades morfossintácticas do composto (categoria sintáctica (cf. 4. e, f, g), género (cf. 4. b, d) e número (cf. 4. c, d)) não são determinadas pelo constituinte da direita, ainda que ocasionalmente possam ser coincidentes (cf. 4a).

- (4) a. $\left[\left[\text{corta} \right] \left[\text{arame} \right]_{N[+masc, +sing]} \right]_{N[+masc, +sing]}$
 b. $\left[\left[\text{guarda} \right] \left[\text{chuva} \right]_{N[-masc, +sing]} \right]_{N[+masc, +sing]}$
 c. $\left[\left[\text{busca} \right] \left[\text{pólos} \right]_{N[+masc, -sing]} \right]_{N[+masc, +sing]}$
 d. $\left[\left[\text{abre} \right] \left[\text{latas} \right]_{N[-masc, -sing]} \right]_{N[+masc, +sing]}$
 e. $\left[\left[\text{faz} \right] \left[\text{tudo} \right]_{PRON} \right]_{N[+masc, +sing]}$
 f. $\left[\left[\text{fala} \right] \left[\text{barato} \right]_{ADJ} \right]_{N[+masc, +sing]}$
 g. $\left[\left[\text{bota} \right] \left[\text{fora} \right]_{ADV} \right]_{N[+masc, +sing]}$

Por outro lado, estas construções violam a relação ISA (=É UM) (cf. Allen 1978), segundo a qual o significado do composto está contido no significado do constituinte da direita:

- (5) a. Floricultura É UMA cultura
 b. * Corta-arame É UM arame

(iii) Os compostos de (1) podem ser relacionados com frases em que ocorrem os nominais (*latas, papéis, arame, chaminés*) e os verbos que entram na formação destas palavras (*abrir, cortar, limpar*). Vejam-se, como exemplo, os seguintes casos de compostos com o verbo *cortar*⁴:

- (6) a. Um corta-papéis cf. * Isto corta papéis
 b. * Um corta-fitas cf. Ele corta fitas
 c. # Um corta-manteiga cf. # Isto corta manteiga
 d. * Um corta-estrelas cf. * Ele/isto corta estrelas (int. literal)

(iv) O constituinte nominal do composto tem a mesma função temática que o argumento interno do verbo base, isto é, respeita as propriedades de selecção categorial e semântica do verbo; assim, a aceitabilidade dos compostos contendo uma dada forma corresponde à

aceitabilidade de frases com o mesmo verbo, como se pode ver em (6). Compare-se a aceitabilidade de (6a) e (6b) com a inaceitabilidade de (6d), e veja-se como em (6c) a aceitabilidade da palavra não atestada varia nas mesmas condições da frase. Tipicamente, o constituinte nominal é um argumento com a função- θ *tema*, na maioria dos casos com o traço semântico [- animado] ⁵.

(v) O processo de formação requer que o composto contenha o nominal que é complemento directo do verbo, excluindo a possibilidade de formar palavras pela combinação do sujeito e do verbo. Apenas se refere a inaceitabilidade da formação a partir dos nominais sujeito, visto que nos casos relevantes analisados apenas ocorrem verbos transitivos, i.e., verbos definidos por uma grelha temática bi-argumental com um argumento externo e um argumento interno (cf. Williams 1981b). Note-se que a nossa análise prevê a impossibilidade de formar compostos com nominais diferentes de objecto directo. Vejam-se os seguintes exemplos, que mostram como apenas pode ocorrer como membro da estrutura da palavra o nome que é argumento interno do verbo:

- (7) a. X abre latas (X=pessoa / instrumento / objecto)
 a'. Abre latas
 a". * Abre-X
 b. X afia lápis (X= aluno / desenhador)
 b'. Afia-lápis
 b". * Afia-X

(vi) A estrutura destes compostos obedece a condições de boa-formação sintáctica respeitando, nomeadamente, a ordenação canónica do Núcleo e Complementos: os elementos ocorrem agrupados em sequências $V + \left\{ \begin{array}{l} \text{Compl} \\ \text{Obl} \end{array} \right\}$. Comparem-se os exemplos de (8) com os apresentados em (1) e (2):

- (8)
- | | |
|--------------------|-----------------|
| a.* latas-abre | d.* barato-fala |
| b.* papéis-corta | e.* fora-bota |
| c.* chaminés-limpa | f.* tudo-faz |

(vii) A estrutura interna destas palavras mostra que elas são formadas por sequências de unidades lexicais, que podem ocorrer autonomamente noutros contextos, e que não é plausível considerar como constituindo entradas independentes no Léxico. Assim, por exemplo, não parece justificável considerar as duas ocorrências de *latas* (ou de *abre*) como correspondendo a itens lexicais distintos no seguinte caso:

- (9) Um abre-latas que não abre latas não presta!

(viii) A interpretação do composto obedece aos mesmos princípios que a interpretação da frase correspondente: em (10a) a interpretação do predicado *cortar papéis* é literal no composto e na frase, enquanto (10b) admite duas leituras possíveis, uma literal e outra não literal, em ambos os casos. Os casos de leitura literal da frase levam à interpretação composicional da palavra, o que não ocorre nos casos em que a frase não pode ser interpretada literalmente (considere-se, como exemplo desta situação, a forma *arranha-céus*).

- (10)
- | | | | |
|------|-------------------|-------|--|
| a. | Um corta-papéis | cf. | Isto corta papéis |
| b. | Um corta-fitas | cf. | Ele corta fitas |
| c. # | Um corta-manteiga | cf. # | Isto corta manteiga |
| d. * | Um corta-estrelas | cf. * | Ele/isto corta estrelas (int. literal) |

(ix) O composto reflecte as restrições semânticas sobre o argumento externo do verbo base. Esta propriedade, que é especialmente evidente no caso das palavras que têm interpretação composicional e a estrutura $\{ V N \}_N$, pode ser observada através de uma paráfrase envolvendo uma relativa que retoma os elementos constitutivos da palavra, como em *um*

pesa-cartas é um instrumento / um objecto que pesa cartas / que serve para pesar cartas.

Como se pode observar, o antecedente da frase relativa permite explicitar as restrições semânticas do argumento externo:

- (11) a. * Um corta-papéis é um instrumento que corta papéis
b. % Um corta-fitas é um objecto que corta fitas
b' % Um corta-fitas é uma pessoa / um presidente que corta fitas
c. %* Um arranha-céus é um prédio que arranha céus
c'. * Um arranha-céus é uma pessoa que arranha céus

3. Reanálise

Na secção anterior apresentámos os argumentos em que se baseia a hipótese de que estes compostos têm uma base sintáctica. Veremos em seguida que, no entanto, estes compostos têm as propriedades típicas das palavras:

(i) Apesar de derivarem de uma estrutura sintáctica, estes compostos são sintacticamente opacos⁶, ou seja nenhum processo sintáctico pode inserir, mover ou substituir os seus constituintes:

- (12) a. * tira-nódoas de gordura
b. * nódoas-tira
c. * tira-as

(ii) Estes compostos têm as propriedades típicas dos nomes: possuem traços de género e número, estão submetidos às restrições semânticas que afectam os nomes, têm a distribuição dos nomes, podem ser objecto de processos sintácticos de modificação de nominais, etc. Assim, por exemplo, a forma *abre-latas* tem as propriedades sintácticas e semânticas próprias de um Nome comum, não animado, contável:

- (13) a. Comprei um /o / este *abre-latas*.
 a'. Comprei um *abre-latas* que não funciona.
 a". Comprei um óptimo *abre-latas* / uns / três óptimos *abre-latas*.
 b. O teu *abre-latas* foi comprado na feira do Relógio.
 c. * O *abre-latas* fugiu.

(iii) A flexão do verbo está bloqueada. A forma verbal comporta-se como uma forma invariável.

- (14) a. * *tiras-teimas*
 b. * *tiram-teimas*
 c. * *tirou-teimas*

(iv) Há uma relação evidente entre este tipo de compostos e os nomes-sujeito⁷ de formação derivacional:

- | | | | |
|------|----|-----------------------|-------------------|
| (15) | a. | <i>afia-lápis</i> | <i>afiador</i> |
| | b. | <i>amola-tesouras</i> | <i>amolador</i> |
| | c. | <i>guarda-portão</i> | <i>porteiro</i> |
| | d. | <i>limpa-botas</i> | <i>engraxador</i> |

A interpretação dos nomes-sujeito depende crucialmente da função temática associada ao sujeito do verbo base. Quando o sujeito tem a função de Agente, Dressler (1986: 527) sugere que a sua interpretação pode variar de acordo com uma hierarquia de valores entre [+animado] e [-animado], esquematizada do seguinte modo:

- (16) agente humano > agente animal > planta que produz frutos > agente impessoal > instrumento > condições locais de eventos ou estados (locativas ou de origem)

Paralelamente, verifica-se que estes compostos correspondem maioritariamente a nomes de agente inanimado (instrumentos, objectos, substâncias) e a nomes de agente animado, que designam geralmente profissões ou uma espécie de caricatura ou de avaliação depreciativa de tipos de pessoas. São ainda formados por este processo um número menor de nomes de animais e de plantas e, finalmente, de nomes de acção⁸.

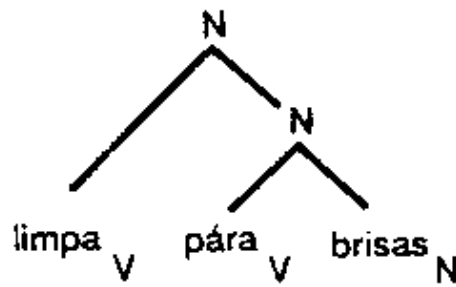
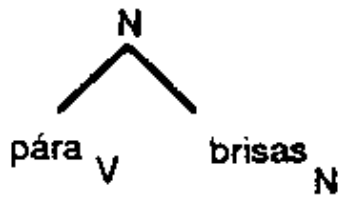
(v) A constituição interna destes compostos obedece às condições de boa-formação sobre a interpretação das palavras. Como se sabe, enquanto unidades do nível lexical, os nomes apenas têm uma referência "genérica"⁹, no sentido em que a sua interpretação carece de ser fixada através de processos sintácticos como, por exemplo, a determinação.

No caso destes compostos, podemos interpretar as restrições sobre a forma do constituinte verbal e nominal como uma estratégia de satisfação do requisito sobre a interpretação dos nomes: ambos os constituintes ocorrem nas formas que permitem a interpretação genérica nas respectivas categorias sintácticas: o verbo na terceira pessoa do singular do presente do indicativo¹⁰ e o nome como um plural sem determinante (bare plural)¹¹. Há casos em que nomes contáveis ocorrem no singular. Trata-se, geralmente, de casos em que o plural é inadequado por razões não-linguísticas (cf. *beija-mão*) ou de formas semanticamente não composicionais, frequentemente geradas pela variante do português do Brasil, e que designam espécies animais ou vegetais (cf. *beija-flor*)¹².

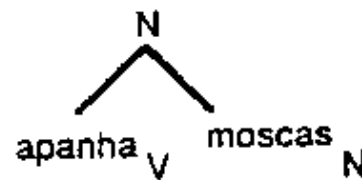
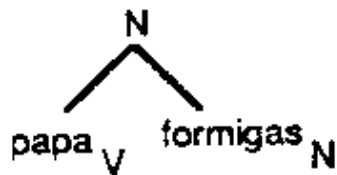
(v) Estes compostos podem ser seleccionados por este mesmo processo de formação de palavras, ou por outros processos, nomeadamente derivacionais:

(17)

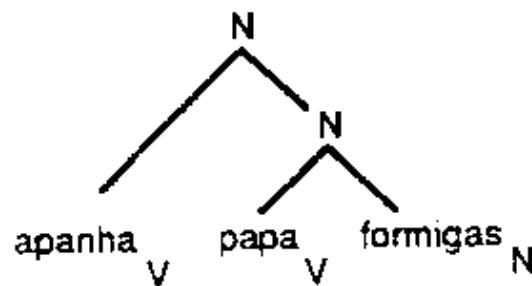
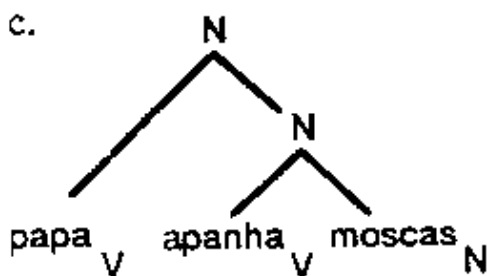
a.



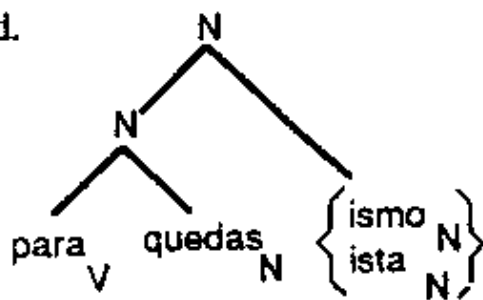
b.



c.



d.



Tendo em conta estas observações parece plausível analisar os compostos do tipo de *tira-teimas* como uma unidade morfológica gerada por um processo de reanálise a partir de uma estrutura sintáctica.

A análise de outros compostos e de expressões lexicalizadas cuja estrutura interna mostra uma base sintáctica, requer igualmente uma explicação segundo um processo de reanálise.

Vejam-se os casos dos nomes apresentados em (18) e (19):

- (18) a. baixo-relevo
 b. cofre-forte
 c. relógio de sol
- (19) a. sobe-e-desce
 b. # desce-e-sobe
 c. * sobe-desce-e
 d. * e-sobe-desce
 e. #* sobe-e- $\left\{ \begin{array}{l} \text{lata} \\ \text{bem} \end{array} \right\}$

Em (19) pode observar-se como a estrutura interna dos compostos obedece às regras gerais da sintaxe em Português, nomeadamente quanto à ordenação dos constituintes coordenados (cf. (19b) versus (19c,d)), e a sua compatibilidade categorial (cf.(19e)).

Note-se, por fim, que compostos com as características gerais acima assinaladas ocorrem em outras línguas para além do Português, nomeadamente noutras línguas românicas. Vejam-se os seguintes exemplos:

- | | | | |
|---------|----------------------|----|-----------------------|
| (20) a. | arranha-céus (port) | b. | porta-aviões (port.) |
| | rasca-cielos (cast.) | | porta-aviones (cast.) |
| | grate-cels (cat.) | | porte- avions (fr.) |
| | gratte-ciel (fr.) | | portaerei (it.) |
| | grattacielo (it.) | | |

Mas também existem construções equivalentes em línguas tão distantes do Português como o Vietnamita (cf. Spencer 1991) e o Mandarim (Anderson 1985):

- (21) a. nhà dĩa ly
 'pessoa' 'geografia' = 'geógrafo'

- b. nhà li.ch su
 ‘pessoa’ ‘história’ = ‘historiador’

- (22) dong shì
 ‘controlar’ ‘negócios’ = ‘membro da direcção’
 lit. pessoa que controla os negócios

Em Japonês (cf. Shibatani e Kageyama 1988), o composto tem como base um sintagma gerado livremente e não lexicalizado:

- (23) amerika o (AC) hoomon
 amerika-hoomon
 ‘visita americana’

Em Inglês (cf. Selkirk 1982), há dois tipos de estruturas $[V N]_N$. Uma é equivalente à das línguas românicas:

- (24) a. pickpocket
 ‘apanhar’ ‘bolso’ = ‘carteirista’
 b. sawbones
 ‘serrar’ ‘ossos’ = ‘cirurgião’
 c. cutthroat
 ‘cortar’ ‘garganta’ = ‘assassino’
 lit. pessoa que corta gargantas

Na outra, o nome é núcleo do composto, mas não satisfaz a estrutura argumental do verbo:

- (25) a. swearword
 ‘praguejar’ ‘palavra’ = ‘palavrão’
 b. whetstone
 ‘afiar’ ‘pedra’ = ‘pedra de afiar’

c. hovercraft

'pairar' 'capacidade' = veículo que se desloca sobre uma almofada de ar

Estes dois tipos de compostos são marginais em Inglês, mas há um terceiro tipo que produz nomes-sujeito compostos, equivalentes aos do português:

- (26) a. abre-latas can opener
 b. arranha-céus skyscraper
 c. porta-aviões aircraft-carrier

Por outro lado, a hipótese da existência na gramática de operações ou processos que afectam as especificações categoriais de constituintes sintácticos é necessária independentemente da análise do caso destes compostos. Geralmente aceita-se que certas construções sintácticas obrigam a uma análise que admite operações de recategorização: estão neste caso as estruturas analisadas em termos de reanálise — por exemplo, as construções de incorporação, como as construções causativas, ou as sequências Verbo-partícula em Inglês. Outro caso passível de ser considerado como objecto de recategorização é o das frases que ocorrem em posições estruturais típicas de constituintes nominais (por exemplo, frases em posição de sujeito ou de objecto directo), na medida em que se admita que a inserção na estrutura superior leva à identificação da categoria da estrutura encaixada a partir da etiqueta do nó de inserção¹³.

Como é evidente, esta análise levanta um grande número de questões, quer do ponto de vista da descrição da sua estrutura interna, quer quanto à caracterização do mecanismo envolvido na sua derivação, nomeadamente no que diz respeito à relação entre as componentes morfológica e sintáctica da Gramática.

4. Descrição do processo

A descrição acima apresentada contém argumentos que fundamentam empiricamente a hipótese de que a formação destes compostos envolve a derivação de uma unidade morfológica a partir de uma estrutura sintáctica, e que esse processo implica a recategorização da expressão. A explicação do processo de formação destes compostos por reanálise de uma estrutura sintáctica deverá determinar a natureza categorial da expressão que é objecto de reanálise, estabelecer as condições - internas e externas - a que obedece o processo, e caracterizar o seu funcionamento.

Poderemos descrever o funcionamento do processo de formação destes compostos a partir da sequência de operações representada em (27) : neutralização de propriedades da estrutura interna (27 b) da expressão sintáctica de origem (parcialmente apresentada em (27 a)); obtenção de uma forma à qual são aplicáveis as regras morfológicas relevantes (27 c); e finalmente, reinterpretação estrutural e reinterpretação categorial, com a atribuição da categoria Nome (27 d).

- (27) a.
$$\left[\text{FLEX} \begin{bmatrix} +3^{\circ} \text{ pess} \\ +\text{sing} \\ +\text{pres} \\ +\text{ind} \end{bmatrix} \left[\text{VMAX} \left[\text{N}'' \emptyset \right] \left[\text{V}' \left[\text{V} \text{ abrir} \right] \left[\text{N}'' \left[\text{DET} \emptyset \right] \right] \left[\text{N}^{\text{latas}} \right] \right] \right] \right]$$
- b.
$$\% \left[\text{VMAX} \left[\text{N}'' \emptyset \right] \left[\text{V}' \left[\text{V} \text{ abre} \right] \left[\text{N}'' \left[\text{DET} \emptyset \right] \right] \left[\text{N}^{\text{latas}} \right] \right] \right]$$
- c.
$$\% \left[? \left[\text{V} \text{ abre} \right] \left[\text{N}^{\text{latas}} \right] \right]$$
- d.
$$\left[\text{abre latas} \right]_{\text{N}}$$

A forma de partida (27 a) e a forma de chegada (27 d) devem satisfazer condições de boa-formação, respectivamente, nos domínios da Sintaxe e da Morfologia. O processo de reanálise envolvido na formação destes compostos pode, assim, ser entendido como um sistema de projecção (mapping) da Sintaxe na Morfologia.

Como atrás foi referido, o constituinte verbal do composto ocorre sob uma forma "bloqueada" do presente do indicativo, o que indica a perda das propriedades próprias do constituinte FLEX da estrutura sintáctica. Dado que FLEX constitui o núcleo sintáctico das estruturas frásicas, se interpretarmos o bloqueio da flexão do constituinte verbal do composto como um indício da inexistência de um constituinte sintáctico FLEX, podemos colocar a hipótese de que a reanálise opera a partir da eliminação do núcleo sintáctico, entendendo a recategorização como uma forma de satisfazer a boa-formação da expressão (27 c), o que apenas pode ocorrer na Morfologia. Note-se que a forma do constituinte verbal do composto mostra que a reanálise opera sobre uma sequência {V+Flex}, isto é, que a estrutura que serve de entrada à reanálise pertence a um nível gramatical sequente à incorporação (ou afixação) do verbo. Assim, parece ser plausível considerar que a reanálise opera sobre uma estrutura posterior à Estrutura-S (Forma Fonética).

Note-se que estas expressões em Português (e em outras línguas de Sujeito Nulo) são ambíguas categorialmente: na ausência de contexto sintáctico, podem ser interpretadas como uma expressão sintáctica ou como uma palavra; veja-se, como exemplo, *abre-latas* — a expressão fonética pode ser interpretada como uma frase sem sujeito expresso (denotando um dado estado de coisas) ou como um palavra (denotando uma entidade)¹⁴. Esta ambiguidade é específica destes compostos, como se pode observar através da comparação com outras palavras formadas por outros processos morfológicos ou com expressões sintácticas:

- (28) a. [N/F abre latas]
 b. [N'/*F abridor de latas]
 c. [*N/F abriu latas]

A operação de reanálise cujos efeitos categoriais foram acima descritos tem, evidentemente, consequências do ponto de vista da estrutura temática. Nomeadamente,

existe uma equivalência entre a frase e o composto enquanto domínios de realização das propriedades temáticas dos predicadores verbais. Como foi observado em (2.iv, vii, viii) os compostos têm uma interpretação argumental equivalente à das frases correlatas. Como se viu atrás (cf. 27 a,b), a nossa hipótese considera que a forma sintáctica sobre a qual se aplicam directamente as regras da morfologia é uma expressão equivalente ao SV (= V^{MAX}) da frase de origem, isto é, o domínio sintáctico de realização das propriedades de selecção do núcleo verbal, assumindo a hipótese de que o "argumento externo" é realizado basicamente na posição de especificador de SV (cf. Sportiche 1989). A recategorização da expressão implica a reatribuição do papel temático atribuído à posição à esquerda do verbo. Consideramos que de um ponto de vista temático a reanálise consiste na absorção do papel θ pela palavra composta. Esta hipótese é consistente com as análises dos chamados argumentos implícitos de Roeper (1987) e Roberts (1987) (ver também: Keyser & Roeper 1984; Baker, Johnson & Roberts (1989)) e é adequada à interpretação destas palavras referida em (2. viii). Note-se, finalmente que este mecanismo satisfaz o Princípio de Projecção na medida em que permite a realização das propriedades do predicador num domínio por si definido. A representação seguinte pretende apresentar as operações envolvidas na reanálise temática subjacente à formação destes compostos. (29a) contém a representação das posições temáticas da forma reanalisada (cf (27c)); e (29b) ilustra a reatribuição do papel- θ 1. Note-se que, enquanto argumento realizado do verbo, o nome requer neste estágio da derivação uma função temática, pelo que a operação descrita em (29b) apenas pode incidir sobre o papel- θ 1. De acordo com a análise dos nomes proposta por Roberts (1987: 240 sgs), existe uma assimetria entre a atribuição do papel- θ externo e do papel- θ interno, uma vez que a atribuição do segundo é opcional e a do primeiro é obrigatória.

Finalmente, (29c) representa a forma final, em que o valor da função temática associada à categoria é uma variável a ser instanciada no processo de inserção lexical; isto significa que

o composto é opaco do ponto de vista sintáctico, tendo as propriedades típicas dos nomes.

- (29) a. $((\text{---})_{\theta 1} \text{ verbo (nome)}_{\theta 2})$
 b. $((\text{verbo (nome)}_{\theta 2})_{x=(\text{---})_{\theta 1}})$
 c. $[\text{verbo nome}]_{N(X)}$

A reinterpretação categorial da sequência [V N] como um nome exige uma atribuição de valores às categorias morfossintácticas relevantes para os nomes, nomeadamente género e número. Como vimos, esses valores não podem ser calculados morfológicamente (por percolação), uma vez que a sequência é exocêntrica (cf (30)=(4)):

- (30) a. $[[\text{corta}][\text{arame}]_{N[+\text{masc}, +\text{sing}]}]_{N[+\text{masc}, +\text{sing}]}$
 b. $[[\text{guarda}][\text{chuva}]_{N[-\text{masc}, +\text{sing}]}]_{N[+\text{masc}, +\text{sing}]}$
 c. $[[\text{busca}][\text{pólos}]_{N[+\text{masc}, -\text{sing}]}]_{N[+\text{masc}, +\text{sing}]}$
 d. $[[\text{abre}][\text{latas}]_{N[-\text{masc}, -\text{sing}]}]_{N[+\text{masc}, +\text{sing}]}$
 e. $[[\text{faz}][\text{tudo}]_{\text{PRON}}]_{N[+\text{masc}, +\text{sing}]}$
 f. $[[\text{fala}][\text{barato}]_{\text{ADJ}}]_{N[+\text{masc}, +\text{sing}]}$
 g. $[[\text{bota}][\text{fora}]_{\text{ADV}}]_{N[+\text{masc}, +\text{sing}]}$

Verifica-se, no entanto, que os valores de género (masculino) e número (singular) seleccionados são, de algum modo, os valores não-marcados para cada uma das categorias. Por outro lado, se se considerar a flexão plural destas estruturas ou a afixação avaliativa, constata-se que elas ocorrem na sua posição normal, ou seja, na posição final à direita:

- (31) a. os corta-árames * os cortas-arame * os cortam-arame
 b. os guarda-chuvas * os guardas-chuva * os guardam-chuva
 c. os fala-baratos * os falas-barato * os falam-barato
 d. o abre-latazinho * o abrinho-latas

Assim, pode inferir-se que, em presença de uma estrutura exocêntrica, os processos morfológicos subsequentes operam igualmente de modo não-marcado.

As formas resultantes exibem uma tendência para adoptar o comportamento regular de palavras simples (ie. sem estrutura interna), o que se verifica, por exemplo, na perda de tonicidade no constituinte verbal, na transformação do morfema de flexão verbal numa vogal de ligação e na 'regularização' da flexão plural do constituinte nominal:

- (32) a. guardanapo
 b. mascão
 c. parapeito
 d. corrimão (corre-mão)
 e. pl.corrimãos / corrimões

5. Considerações finais

Apresentámos neste trabalho alguns argumentos a favor da hipótese segundo a qual um tipo particular de compostos em Português tem uma base sintáctica. Segundo a proposta aqui apresentada, tais palavras são resultantes de um processo de nominalização que consiste na reanálise das estruturas sintácticas de partida, através de operações de recategorização e identificação de constituintes que afectam apenas os elementos sintácticos lexicalmente realizados.

De um ponto de vista empírico, estes compostos exibem um conjunto bem definido e

característico de propriedades morfológicas e temáticas que permite a sua identificação e delimitação, não só em Português como em diversas línguas.

De um ponto de vista teórico, a análise destes compostos coloca alguns problemas à concepção de gramática comumente adoptada nas várias versões da Gramática Generativa que desenvolvem o modelo apresentado nas *Lectures on Government and Binding* (Chomsky 1981)¹⁵, e que é constituído por módulos autónomos (componentes da gramática: Estrutura-P, Estrutura-S, Forma Lógica e Forma Fonética e sistemas de princípios: Teoria do Caso, Teoria Temática, etc).

Como se sabe, este modelo de gramática é apresentado nas *Lectures* de forma esquemática e alguns aspectos são deixados intencionalmente pouco definidos¹⁶; no entanto, parece poder considerar-se existir alguma contradição entre aquilo que é definido como o funcionamento modular da gramática e a forma sequencial como é concebida a relação entre componentes: o modelo é semi-modular, visto que embora FF e FL sejam componentes paralelos, a sua relação com os restantes componentes e a destes entre si é sequencial.

Tal como é geralmente assumida, esta perspectiva implica que os elementos lexicais são sucessivamente projectados nos vários níveis de representação a partir de uma representação inicial e que a informação gramatical contida num dado nível é necessariamente maior que a informação contida nos níveis anteriores (tanto quanto a sequencialidade se mantém).

Ora, o caso dos compostos aqui analisado sugere a necessidade de rever esta concepção do modo de funcionamento da gramática. Como vimos, trata-se de estruturas sintácticas que ocorrem em posições típicas dos Nomes por terem sido objecto de uma recategorização — este processo é evidentemente incompatível com uma concepção sequencial da relação entre os componentes da gramática, requerendo a possibilidade de acesso à informação relevante em qualquer ponto da representação gramatical, isto é, a um modelo inteiramente modular e interactivo.

No modelo das *Lectures*, a Morfologia não é um componente da Gramática, considerando-se que os processos derivacionais operam no Léxico, e que os processos da morfologia flexional se manifestam na Forma Fonética. Esta perspectiva é posta em causa por autores que adoptam a Hipótese Lexicalista Forte¹⁷, ou seja, que defendem um tratamento unificado da Morfologia, baseando a sua posição, fundamentalmente, no facto de tanto a derivação como a flexão recorrerem a processos formais idênticos (nomeadamente a afixação), e na constatação de que nenhum processo sintáctico pode operar no interior do domínio da palavra. DS&W (1987: 4) propõem mesmo uma distinção entre Léxico (caracterizado como um conjunto não-estruturado de objectos semi-gramaticais, que podem ser palavras, sintagmas ou frases) e a Morfologia, que é uma sub-teoria da Gramática, tal como a Sintaxe. Na sequência de DS&W (1987), mas citando também Baker (1988c), Borer (1988) e Shibatani & Kageyama (1988), Spencer (1991: 455) sugere que a Morfologia constitui um módulo da Gramática, autónomo e não ordenado relativamente aos vários níveis de representação, e que contém um conjunto de princípios sobre a boa-formação fonológica e morfológica da estrutura das palavras. Operando em paralelo com a restante derivação gramatical, o módulo de Morfologia pode, assim, interagir com todos os outros níveis de representação, o que parece resolver a questão levantada pela análise dos compostos do tipo de *tira-teimas*.

- 1 A investigação desenvolvida com vista ao presente trabalho envolveu uma etapa de recolha e classificação de dados a partir da consulta de dicionários. Este procedimento, que foi aplicado sem preocupações de natureza estatística, justifica-se pela natureza do objecto de estudo, uma vez que a listagem de palavras pareceu ser a forma mais económica de aceder aos dados pertinentes. Apesar de a recolha não poder ser considerada como uma amostra representativa, certos dados quantitativos parecem relevantes e serão mencionados quando tal se justificar.

Note-se que a existência de neologismos mostra que este processo de formação de compostos se mantém produtivo em Português: *lança-chamas*, *lança-granadas*, *lavatudo*, *para-quedas*, *porta-aviões*.

- 2 Os casos de compostos com nominais representam cerca de 94% do total de compostos analisados (13 em 228).
- 3 A caracterização da estrutura morfológica assenta, fundamentalmente, na identificação do núcleo morfológico. O núcleo morfológico (cf. Williams 1981a) corresponde sempre a uma posição estrutural periférica, que é aquela que transmite as suas propriedades (nomeadamente o género e o número) ao nó superior. Em português, a estrutura morfológica das palavras complexas é uma estrutura de núcleo à direita:
- $$\begin{array}{ccccccc} [organ] & [iza] & & [ção] & & & \\ N & TV & TV & N & N' & & \end{array}$$

A existência de constituintes finais, como os sufixos avaliativos e os sufixos de flexão, incapazes de transmitir à palavra propriedades determinantes como a categoria sintáctica, motivou a reformulação do conceito de núcleo de palavra. A noção de núcleo relativizado (cf. DS&W 1987) prevê que a palavra pode ter vários núcleos (categoria, género, número, etc) e que o núcleo relativamente a um dado traço corresponde ao constituinte mais à direita que contém informação quanto a esse traço.

- 4 Para indicar as diferenças de aceitabilidade dos exemplos, optou-se pela seguinte convenção: para além dos casos agramaticais e duvidosos assinalados por * e ?, como é habitual, são assinalados com # os casos de formas não atestadas, e com % os casos cuja aceitabilidade depende de variações de interpretação ou variantes estilísticas.
- 5 Na quase totalidade dos casos atestados, o nominal tem o traço [-animado]. Como é evidente, a caracterização semântica do argumento está relacionada com as propriedades de selecção do verbo e um composto como *apanha-moscas* exemplifica um caso de estruturas com um nominal [+animado]. O único caso de composto formado com um nominal [+humano] é *caça-rapazes*, cuja interpretação não é composicional.
- 6 A opacidade sintáctica é uma propriedade de todas as expressões linguísticas sintacticamente inanalísáveis, qualquer que seja o seu grau de complexidade interna. Estes objectos linguísticos são átomos sintácticos, ou seja, são unidades cuja estrutura interna e processos de formação não são relevantes em sintaxe, onde só é visível a informação associada ao nó superior.
- 7 Este tipo de nomes é tradicionalmente designado como 'agentivo'. Sabe-se, no entanto, que o nome derivado nem sempre estabelece uma relação de agente relativamente ao verbo derivante, mas que se verifica uma relação sistemática entre o nome derivado e a função temática do sujeito do verbo derivante: *atacar* -> *atacante* (agente), *sofrer* -> *sofredor* (experienciador), *ser convidado* -> *convidado* (tema). Por esta razão, Booij (1986: 507) propõe a designação de nomes-sujeito, uma vez que o sufixo herda a função

temática associada à posição de sujeito do verbo derivante. Esta designação será aqui adoptada.

8 (i) AGENTE INANIMADO

abre-latas	afia-lápis	apanha-migalhas	apanha-moscas	apara-lápis
arranca-raízes	arranha-céus	arrebenta-diabos	busca-fundo	busca-pólos
busca-vida(s)	caça-minas	cata-vento	chupa-tinta	conta-giros
conta-gotas	conta-passos	conta-quilómetros	conta-rotações	conta-voitas
corta-aramé	corta-charutos	corta-fios	corta-fogo	corta-mar
corta-palha	corta-papel	corta-rio	corta-vento	estanca-rios
estanca-sangues	fura-bolos	fura-gelos	ganha-pão	gira-discos
guarda-chuva	guarda-comida	guarda-costas	guarda-fatos	guarda-fogo
guarda-jóias	guarda-lama	guarda-louça	guarda-lume	guarda-mancebos
guarda-pó	guarda-pratas	guarda-queadas	guarda-raios	guarda-roupa
guarda-sol	guarda-vento	guarda-vestidos	guardanapo	lança-bombas
lança-chamas	lança-granadas	lança-torpedos	lava-tudo	limpa-calhas
limpa-chaminés	limpa-pára-brisas	limpa-pratas	limpa-vidros	mata-borrão
mata-cães	mata-moscas	mata-piolhos	mata-pulga	mata-ratos
monta-cargas	parapeito	pára-águas	pára-brisas	pára-choque
pára-chuva	pára-fogo	pára-lama	pára-queadas	pára-raios
pára-sol	pára-vento	pesa-papéis	pica-ponto	porta-aviões
porta-bagagem	porta-cartas	porta-chapéus	porta-chaves	porta-escovas
porta-jóias	porta-lápis	porta-livros	porta-luvas	porta-maías
porta-moedas	porta-penas	porta-retratos	porta-revistas	quebra-gelo
quebra-luz	quebra-mar	quebra-nozes	saca-molas	saca-roilhas
salva-vidas	talha-mar	tira-cápsulas	tira-leite	tira-linhas
tira-nódoas	tira-teimas			

(ii) AGENTE ANIMADO

amola-tesouras	arranca-dentes	arranca-pinheiros	arranca-pinheiros	arrasa-montanhas
arrepia-cabelo	beija-flor	busca-vida	caça-fantasmas	caga-baixinho
caga-lume	caga-na-saquinha	caga-na-telha	caga-raiva	caga-regras
cata-vento	chupa-flor	chupa-jantares	chupa-mel	chupa-sangue
corta-pau	corta-vento	cruza-bico	desmancha-prazeres	espalha-brasas
espirra-canivetes	espreita-marés	estraga-albardas	fala-barato	faz-tudo
fura-bolo	fura-greves	fura-paredes	fura-vidas	ganha-dinheiro
guarda-costas	guarda-fios	guarda-freio	guarda-linha	guarda-livros

guarda-portão lambe-botas	guarda-redes lambe-cu	guarda-rios limpa-botas	guarda-rios limpa-calhas	guarda-vidas limpa- candeeiros
limpa-chaminés papa-açorda papa-jantares papa-santos pica-peixe pisa-mansinho quebra-esquinas	manda-chuva papa-figos papa-léguas passa-culpas pilha-galinhas porta-bandeira quebra-nozes	mata-cães papa-formigas papa-missas pica-flor pinga-amor porta-estandarte saca-molas	mata-mouros papa-gente papa-moscas pica-osso pinta-monos porta-novas salta- pocinhas traga-mouros trinca-pintos	mata-sete papa-hóstias papa-moscas pica-pau pisa-flores porta-voz salva-vidas
sega-vidas trinca-espinhas vira-casaca(s)	suga-flores trinca-fortes vira-lata	topa-a-tudo trinca-nozes		tranca-ruas troca-tintas
(iii) PLANTAS apanha-moscas mata-cavalo	estanca-sangue mata-cão	fura-paredes mata-pulga	girassol papa-moscas	lava-pé
(iv) NOMES DE ACÇÃO beija-mão escalda-rabo	beija-pé finca-pé	bota-fora ganha-pão	corta-mato lava-pés	escalda-pés mata-bicho

- 9 Segundo DS&W (1987), todas as palavras têm uma propriedade comum: uma interpretação genérica, que é diferente da que as frases podem receber. Esta diferença resulta, por hipótese, da existência de referência temporal nas frases e da sua inexistência no interior das palavras. Também não existe referência pronominal no interior das palavras. É esta opacidade referencial das categorias X0 que lhes confere um carácter genérico.
- 10 A terceira pessoa singular do presente do indicativo é em todo o paradigma de flexão verbal, a forma menos marcada: não contém qualquer morfema de pessoa, número, tempo ou modo.
- 11 A expressão da leitura genérica no nominal do composto varia de acordo com as propriedades do nome isolado. Assim, quando este constituinte é um nome contável, a sua forma é geralmente plural: *abre-latas, amola-tesouras, apanha-moscas, arranca-dentes, conta-gotas, guarda-jóias, limpa-chaminés, papa-jantares, porta-aviões*.
- 12 Noutros casos, registam-se as duas formas, ainda que, provavelmente, cada falante utilize apenas uma delas (o dicionário Lello, por exemplo regista apenas a forma assinalada a escuro): *arranha-céu / arranha-céus; guarda-fato / guarda-fatos; guarda-lama / guarda-lamas; pára-brisa / pára-brisas; pára-choque / pára-choques; porta-bagagem / porta-bagagens; quebra-cabeça / quebra-cabeças; saca-rolha / saca-rolhas; tira-teima*

/ *tira-teimas*.

Por último, registam-se casos em que a alternância singular / plural corresponde a diferentes interpretações: *fura-bolo*=pessoa metedicha / *fura-bolos*=dedo indicador / *estanca-sangue*=planta que tem propriedades de coagulação / *estanca-sangues* = rosário colocado na cabeça com o fim de suster hemorragias nasais.

- 13 Veja-se, por exemplo, a análise de Raposo (1987), em que se assume o estatuto nominal das frases em determinadas posições sintáticas.
- 14 Poder-se-ia argumentar que no exemplo dado nos encontramos perante uma forma lexicalizada, pelo que não seria adequada a análise aqui proposta. Embora se admita a existência de compostos deste tipo lexicalizados, deve notar-se que casos como *corta-manteiga* (um neologismo aqui criado) que, apenas podem ser descritos como resultado de aplicação de reanálise possuem exactamente as mesmas propriedades. Assim, parece irrelevante o facto de os compostos poderem ocorrer como elementos lexicalizados (através de um processo necessariamente subsequente à sua formação nos termos aqui referidos).
- 15 Referimos aqui de forma simplificada o modelo apresentado em Chomsky (1981) por ser referência comum no quadro da Gramática Generativa. Como se sabe, diversos autores propuseram alternativas mais ou menos radicais a este modelo inicial (cf. Riemsdijk e Williams 1981, 1986; Koster 1987, Duarte 1987, entre outros).
- 16 Como exemplo de aspectos deixados em suspenso, veja-se a questão da concepção derivacional ou representacional da gramática.
- 17 Cf. Halle (1973), Jackendoff (1975) e Williams (1981a), entre outros.

Referências

S. R. ANDERSON

- 1985 Typological distinctions in word formation In T. Shopen (ed) *Language Typology and Syntactic Description III* Cambridge: Cambridge University Press (3-56)

M. ALLEN

- 1978 *Morphological Investigations* Universidade de Connecticut: Dissertação de PhD

M. BAKER

- 1988a *Incorporation: A Theory of Grammatical Function Changing* Chicago: The University of Chicago Press
1988b Review: Di Sciullo and Williams (1987) *Yearbook of Morphology* 1 (259-284)
1988c Morphology and syntax: an interlocking independence in Everaert, Evers, Huybregts e Trommelen (eds), *Morphology and Modularity* Dordrecht: Foris (9-32)

- M. BAKER, K. JOHNSON I. ROBERTS
1989 Passive Arguments Raised *Linguistic Inquiry* 20. 2 (219-251)
- Geert BOOIJ
1986 Form and meaning in morphology: the case of Dutch 'agent nouns' *Linguistics* 24 (503-518)
- H. BORER
1988 On the parallelism between compounds and constructs *Yearbook of Morphology* 1. (45-66)
- N. CHOMSKY
1981 *Lectures on Government and Binding* Dordrecht: Foris
- J. Almeida COSTA e A. Sampaio e MELO (org.)
Dicionário da Língua Portuguesa 5ª edição Porto: Porto Editora
- Anna Maria DI SCIULLO e Edwin WILLIAMS
1987 *On the Definition of Word* Cambridge, Massachusetts: The MIT Press
- Wolfgang U. DRESSLER
1986 Explanation in natural morphology, illustrated with comparative and agent noun formation *Linguistics* 24 (519-548)
- I. DUARTE
1987 A construção de topicalização na gramática do português: Regência, Ligação e Condições sobre o movimento. Diss. Doutorado, Universidade de Lisboa.
- Aurélio Buarque de Holanda FERREIRA
1986 *Novo Dicionário da Língua Portuguesa*. 2ª edição Rio de Janeiro: Nova Fronteira
- Morris HALLE
1973 Prolegomena to a theory of word-formation *Linguistic Inquiry* 4 (3-16)
- Ray JACKENDOFF
1975 Morphological and semantic regularities in the lexicon *Language* 51 (639-671)
- S. J. KEYSER e T. ROEPER
1984 On the Middle and Ergative Constructions in English *Linguistic Inquiry* 15 (381-416)
- J. KOSTER
1987 *Domains and Dynasties: The Radical Autonomy of Syntax* Dordrecht: Foris
- E. RAPOSO
1987 Case Theory and Infl-to-Comp: The Inflected Infinitive in European Portuguese *Linguistic Inquiry* 18 (85-109)
- H. van RIEMSDIJK e E. WILLIAMS
1981 NP Structure *The Linguistic Review* 1 (171-218)
1986 *Introduction to the Theory of Grammar* Cambridge, Massachusetts: The MIT Press
- I. ROBERTS
1987 *The Representation of Implicit and Dehematized Subjects* Dordrecht: Foris

- T. ROEPER**
 1987 *Implicit Arguments and the Head-Complement Relation* *Linguistic Inquiry* 18 (267-310)
- E. SELKIRK**
 1982 *The Syntax of Words* Cambridge, Massachusetts: The MIT Press
- M. SHIBATANI e T. KAGEYAMA**
 1988 *Word formation in a modular theory of grammar: a case of post-syntactic compounds in Japanese* *Language* 64 (451-484)
- Andrew SPENCER**
 1991 *Morphological Theory* Cambridge: Blackwell
- D. SPORTICHE**
 1989 *Le mouvement syntaxique: Contraintes et Paramètres* *Langages* 95 (35-80)
- Alina VILLALVA**
 1991 *Portuguese compounds* *Rivista di Linguistica* (em publicação)
- Edwin WILLIAMS**
 1981a *On the notions 'lexically related' and 'head of a word'* *Linguistic Inquiry* 12 (245-274)
 1981b *Argument Structure and morphology* *The Linguistic Review* 1 (81-114)
- s/ autor
 Lello Universal 2 volumes Porto: Lello & Irmão