

A formação dos compostos no português

Cláudio Moreno (UFRGS)

ABSTRACT – After discussing the different kinds of compounds in Portuguese, the paper stands up for the idea that all compounds are formed in the syntax and, as they enter the lexicon, they pass through different stages of lexicalization.

RESUMO – Depois de discutir os diferentes tipos de compostos, em português, o artigo expõe a idéia de que todos os compostos são formados na sintaxe e que, ao entrarem no léxico, passam por diferentes estágios de lexicalização.

O presente artigo é parte de minha tese – *Morfologia Nominal do Português: um estudo de Fonologia Lexical*¹ –, em que examino a morfologia do vocábulo nominal do Português Brasileiro de acordo com os pressupostos da Fonologia Lexical. Nele proponho uma organização do nosso léxico seguindo o modelo de Borowsky (1993), que defende a tradicional divisão em dois níveis – o Nível 1, ou *Nível do Radical*, e o Nível 2, ou *Nível do Vocábulo* –, com a importante inovação de que *toda a fonologia do Nível do Vocábulo precede a morfologia deste mesmo nível*. Dentro deste quadro, proponho situar a *derivação sufixal* no *Nível do Radical*, e a *derivação produtiva* (com *-inhV* e *-íssimo*) e a *prefixação* no *Nível do Vocábulo*. Os compostos serão considerados como formações exclusivamente *pós-lexicais*, admitindo-se uma tendência de alçamento de compos-

¹ Apresentada para obtenção do título de Doutor em Letras, em 1998, na PUCRS. O texto integral pode ser solicitado no endereço eletrônico cmoreno@zaz.com.br

tos cristalizados para o *Nível do Vocábulo*, num processo gradiente de lexicalização; estou inclinado a incluir, neste grupo, as formações com *-zinhV* e *-mente*.

1 Vocábulos simples, compostos e frases

Todos os que entram no território dos compostos sentem que estão pisando em terreno movediço. Sua delimitação, sua flexão, tudo parece impreciso; às vezes, temos dúvida até mesmo quanto a sua existência como *vocábulo* (e não como *frase*). Essa indecisão, que não se restringe aos compostos do Português, nasce do território incerto em que eles são formados, exatamente no limite entre a morfologia e a sintaxe. De um lado, mais próximo dos contornos da morfologia, há casos limítrofes entre o vocábulo *simples* e o *composto* (como demonstrei (1977) para os elementos *-íssimo*, *-zinho* e *-mente*); do outro lado da escala, nas proximidades da sintaxe, mais imprecisa ainda é a distinção entre *vocábulo* e *frase*: a construção dos compostos, como veremos, é tão assemelhada às construções da sintaxe que se torna difícil – se não impossível – determinar claramente o limite entre *compostos* e *sintagmas*. Paradoxalmente, entretanto, o fato de falarmos em *vocábulo composto* (em todas as línguas) parece apontar para a idéia obrigatória de que eles constituem parte do *léxico*. Isso vai nos obrigar a definir, dentro do modelo da Fonologia Lexical, em qual componente eles são formados – se no componente lexical ou fora dele. Há basicamente duas hipóteses: na primeira, sua formação se daria em algum lugar *antes* da sintaxe (o que levaria ao problema de determinar em que nível se localiza este tipo de formação vocabular – se no estrato cíclico ou pós-cíclico – e definir sua posição relativa para com a afixação, a flexão e outros processos morfológicos); na segunda hipótese, que me parece a correta, sua formação acontece no *componente sintático*; neste caso, é necessário explicar de que maneira eles ingressam no léxico, depois de formados.

Na verdade, não acredito que exista, no Português, mecanismo *morfológico* para a *formação* de compostos. O output de nossas regras morfológicas é sempre um vocábulo morfológico, ao passo que todos os compostos são formados de *dois* (raramente mais do que isso). Dois vocábulos prontos, que já passaram por toda a derivação, começam a ser usados na forma de uma expressão habitual; esta expressão, por força de regras *semânticas*, passa a ser analisada pelo falante como um todo e é lexicalizada. Como esses vocábulos estavam ligados por uma relação sintática, é natural que a expressão que veio a se tornar um composto reflita as relações sintá-

táticas de onde ela proveio: ou *N+modificador*, ou *V+complemento*, ou *modificador E modificador* (coordenação). É esta origem sintática, aliás, o fator que dificulta a delimitação precisa entre o que já é vocábulo composto e o que ainda é uma simples estrutura frasal.

2 Objetos morfológicos e palavras sintáticas

Essa vem a ser a mesma posição de Di Sciullo e Williams (1987): ao estudar os compostos no Inglês, eles distinguem (1) os que constituem *objetos morfológicos* e (2) os que são *palavras sintáticas*, que funcionam como *frase* na morfologia e como *palavra* na sintaxe, podendo ser criados pela regra que denominam *Nonmorphological Word-Creation Rule*:

(1) Y > X P

onde Y representa a categoria lexical e XP representa as categorias lexicais máximas (VP, NP, AP, PP).² Esta regra prediz que qualquer unidade sintática pode ser *reanalisada* como uma palavra. Segundo os autores, o critério para distinguir entre os dois tipos é a *posição do núcleo*: nos compostos que são *objetos morfológicos*, o núcleo fica à direita, como nas derivações; nos compostos que são *palavras sintáticas*, o núcleo fica à esquerda:

(2) [[wolf]_N [children]_{N,pl}]_{N,pl} mas [[break]_V [down]_P]_V

Para as línguas românicas, entretanto, Di Sciullo e Williams não postulam a existência desses dois tipos encontráveis no Inglês; nelas, todos os compostos são *palavras sintáticas*. É o que defendemos no presente trabalho, para o Português Brasileiro: todos são sintaticamente transparentes, mas funcionam como *vocábulo*, já que podem ser inseridos na posição X^o. Nesta posição, os compostos têm interpretações diferentes da *simples frase* (em *pé de moleque*, *pé* não é uma parte do corpo, como na sintaxe). Como veremos adiante, o trabalho de Lee (1995) não aceita a afirmação de que todos os compostos do PB sejam frases reanalisadas; para ele, haveria também, à semelhança do Inglês, compostos que são *objetos morfológicos*.

² Aplicando esta regra para o PB, Lee (1995) diz que Y = [+N], i.e., N ou A.

3 A posição do núcleo nos compostos no PB

Antes de refutar sua proposta, vamos examinar os compostos do PB quanto à existência e posição do seu núcleo. Como vimos, o critério para distinguir entre os dois tipos de compostos do Inglês seria a *posição do núcleo*: à direita nos compostos que são *objetos morfológicos* e à esquerda nos compostos que são *palavras sintáticas*. *Esquerda e direita* devem ser entendidas aqui à luz da estrutura do sintagma nominal do Inglês: nas relações de modificação, o modificador fica sempre à *esquerda*. Se fôssemos aplicar este critério para o PB, teríamos de inverter as combinações: nos *objetos morfológicos*, o núcleo ficaria à esquerda; nas *palavras sintáticas*, o núcleo ficaria à direita. Entretanto, esse critério da posição do núcleo não é válido para o PB, que segue outros parâmetros, já que os compostos ou (1) apresentam núcleo à esquerda, ou (2) apresentam núcleo à direita (menos comuns), ou (3) não têm núcleo.

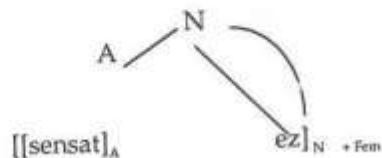
3.1 Compostos com núcleo à direita

A Convenção IV de Lieber, aplicável ao Inglês, determina que, se dois radicais são irmãos (isto é, formam um composto), traços do radical à direita infiltram para o nó que domina esses radicais. No PB, isso se verifica em casos pouco numerosos:

- | | | |
|-----|----------------------|-----------|
| (3) | ponta pé (m.) | (f. + m.) |
| | caféi cultura (f.) | (m. + f.) |
| | vaso constrição (f.) | (m. + f.) |
| | vara pau (m.) | (m. + f.) |
| | água pé (m.) | (f. + m.) |

Nestes compostos, a categoria do todo é determinada pelo nome à direita, que é o *núcleo*, dentro do postulado pela Conv. IV. Seria realmente sedutor defender esta parametrização para o PB, pois ela estaria também de acordo com a estrutura básica das palavras *derivadas* do PB, em que a categoria e os traços do afixo mais à direita rotulam o nó mais alto, como em

(4)



A estrutura presente nos compostos em (3), contudo, como afirmamos, não é muito freqüente no PB. E. C. Pereira afirma que esta é a forma mais comum no Alemão, no Inglês e no Latim; no Português, os vocábulos que seguem este processo são, em sua maioria, formações eruditas, de origem latina e grega.

Os compostos mais comuns com núcleo à direita são os do tipo A + N, como em

- | | | |
|-----|------------------------|--------------------|
| (5) | baixo relevo (m.) | alto relevo (m.) |
| | primeiro ministro (m.) | van glória (f.) |
| | segunda feira (f.) | salvo conduto (m.) |

Os compostos assim formados são *nomes*, com todos os traços correspondentes, e não *adjetivos*. É que aqui atua uma relação de *modificação* entre os seus membros e, neste caso, o núcleo pode ser identificado sem ser necessário fixar parâmetros em qualquer direção. Cedeño propõe para o Espanhol, em casos semelhantes, uma modificação da Convenção IV de Lieber, que também é válida para o PB, já que ambas as línguas não têm posição fixa, como o Inglês, para o modificador:

(6)

Numa configuração de modificação, os traços de N infiltram para o nó da árvore, independente da posição que ocupa no composto.³

3.2 Compostos com núcleo à esquerda

Nos compostos em que há relação de modificação entre seus membros - N+N ou N+A -, a configuração mais produtiva no PB é *núcleo + modificador*, seja este um *adjetivo* ou um *nome* com valor de adjetivo:

- | | | | |
|-----|-----|---------------|--------------|
| (7) | (a) | N + Adj | |
| | | amor perfeito | sangue frio |
| | | senso comum | estado maior |
| | (b) | N + N | |
| | | homem rã | sofá cama |
| | | navio escola | beira mar |

Em (b), é muito fácil determinar o núcleo: os três primeiros são *masculinos*, embora sejam formados com elemento *masculino* à

³ Cedeño, 593.

esquerda e feminino à direita. Em *beira mar*, dá-se exatamente o inverso. Os do tipo (a) são verdadeiros compostos *endocêntricos*: *sangue frio* designa um subconjunto do significado dos elementos contidos ou designados pelo substantivo *sangue* (cf. Selkirk, 1983). Quando falamos de *sangue frio*, há o sentido de que *sangue*, e não algo ou alguém fora do composto, adquiriu a característica de *frieza*; portanto, o elemento qualificado *sangue* deve ser o núcleo do composto. O mesmo se pode dizer de *navio escola*, que é um *navio* que funciona como *escola*; o 2º elemento define o 1º, assim determinando, semanticamente, que o núcleo é *navio*.

3.3 Compostos sem núcleo

Como vimos acima, não basta examinar as categorias dos membros que participam da composição; é necessário também levar em consideração a *categoria do vocábulo assim formado*. As combinações que geram *adjetivos* (geralmente A+A; às vezes, N + A) são *exocêntricas*, como o Ing. *redhead*. *Redhead* não é uma *head* que é *red*, mas *alguém* que tem estas características. *Boca mole* não é uma *boca* que é *mole*, mas *alguém* que fala (e age) sem firmeza. Aqui fica bem manifesta a característica fundamental de um composto exocêntrico: como bem explica Bloomfield (236), a construção implica precisamente que o objeto não pertence à mesma espécie que o núcleo (*boca*): significa, na verdade, *alguém* que possui determinado objeto (o 1º membro) de determinada qualidade (o 2º membro). Note-se que, neste caso, numa sequência como *um boca mole*, o gênero do artigo não concorda com o núcleo e sim com um nome ausente no sintagma mas subentendido, “que deve ser tomado como o ocupante do nó vazio”. Lemle acrescenta que este mecanismo que o PB apresenta de criar nomes a partir de adjetivos só existe “porque é possível haver adjetivos ligados a nós nominais vazios sujeitos a interpretação” (Lemle, 105). Da mesma forma, *surdo mudo* é atribuído a alguém que está fora do composto.⁴

Na falta, portanto, de um núcleo claramente distinguível – que, neste caso, está sendo referido por *implicação* ou *metáfora* –, devemos considerar *sem núcleo* estes compostos com funções adjetivas. Segundo Cedeño, esses vocábulos seriam gerados pela regra

- (8) A → { Adj } A
 { N }

Da mesma forma, os compostos V + Comp não têm núcleo à direita ou à esquerda:

- (9) ganha pão porta aviões
 toca discos beija flor
 salva vidas arranha céu

Se postulássemos o núcleo à direita, faríamos predições incorretas para o PB. Por exemplo, *porta aviões* e *guarda chuva* seriam

- (10) (a) (b)
- ```

 N N
 / \ / \
[porta] .v [aviões] N + PL [guarda] v [chuva] N , FEM

```

Em (a), o nome à direita é plural, e este traço deveria infiltrar até o N; no entanto, é um N singular (*um porta aviões*); em (b), *chuva* faria todo o N ser feminino, quando é indiscutivelmente masculino. Se, por outro lado, considerássemos como núcleo o membro à esquerda, os compostos acima receberiam equivocadamente o traço de V.

Para Cedeño, o composto V + Comp, o mais produtivo do Espanhol (e do PB, sem dúvida), seria assim estruturado:

- (11)
- ```

      NP
     /  \
    NP   VP
   /  \  /  \
 Det  N' V   NP
  |    |  |   |
  o    e  quebra  nozes

```

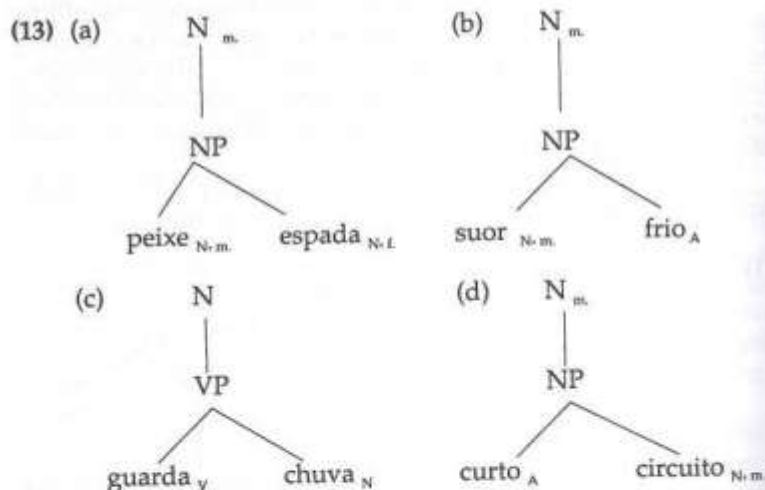
A presença aqui de um sujeito NULL (*e*) é determinada pelo próprio sistema flexional do Português e licenciada pelo determinativo *o* (artigo). Nesta análise, então, os compostos exocêntricos V + Comp devem ser parametrizados para receber os traços do núcleo à esquerda de uma NP vazia – o que é feito pela regra acima. Vil-

⁴ Cuidado: em *surdo mudo*, é necessário distinguir entre (1) o *adjetivo* composto, exocêntrico – *alguém* que é *surdo* e *mudo* – e (2) o *substantivo* composto, endocêntrico, com núcleo à esquerda: *o surdo* que é *mudo*. Por não perceber esta distinção, como vimos, a gramática escolar considerou esse vocábulo uma exceção; Lee, pela mesma razão, caiu em equívoco.

lavra conclui que este tipo de composto é *sem núcleo*, pois interpreta sua estrutura como



Como, para ela, o traço *default* dos compostos do Português é [+N] masculino, a categoria lexical e o gênero masculino dos compostos *V+ Comp* são totalmente previsíveis. Tudo nos leva, portanto, a considerar adequada para nossos compostos a descrição de Villalva, que, ao examinar o PEuropeu, chega à conclusão de que o composto pode ter o núcleo à direita, à esquerda ou simplesmente não ter núcleo algum:



Todos os exemplos acima são justificados por Villalva pela regra da formação de palavra não-morfológica, já que o núcleo da palavra sintática infiltra para o composto resultante. No caso de (c), como não existe núcleo, o composto assume o traço default, que é [+N, +m.]. Não existem, portanto, compostos do tipo obje-

tos morfológicos no Português.⁵ A posição diversificada do núcleo em nossos compostos não é traço determinante de uma distinção como a que Di Sciullo e Williams fazem para o Inglês. Confirmando o que aqueles autores sugeriram para as línguas românicas em geral, Villalva conclui que todos os compostos do Port. são palavras sintáticas reanalisadas e, como tal, formadas no componente sintático. Como veremos adiante, estas criações da sintaxe vão, a partir de determinado momento, ingressar no léxico.

4 A análise de Lee

Em seu estudo sobre os compostos do PB, Lee (1995) não aceita a posição de Villalva; para ele, aceitar que o processo de composição tenha como resultado apenas palavras sintáticas não daria conta da opacidade para a descrição e a operação sintática, nos casos que envolvem compostos que teriam as características de verdadeiros objetos morfológicos. Como explicar que determinados compostos funcionam como "unidade única"? Para justificar essa diferença de comportamento, Lee postula a existência de *compostos verdadeiros* (que chama de *compostos lexicais*, correspondentes aos *objeto morfológicos*) e de *pseudocompostos* (estes sim, e só estes, correspondentes às palavras sintáticas reanalisadas de Di Sciullo e Williams).

A partir desta distinção, Lee pretende explicar vários fatos que perturbam a descrição dos compostos no PB:

- 1) Podem ter dois acentos (*tóca discos*)
- 2) Podem flexionar-se entre os constituintes (*garotas propaganda*)
- 3) Só podem ser vocábulo [+N] (N, A; mas *V)
- 4) Podem formar DIM entre constituintes (*guardinha noturno*)
- 5) Podem flexionar-se mais de uma vez (*homens rãs*).

Para ele, os *compostos lexicais* formam-se no léxico e são *sintaticamente opacos*: (a) comportam-se como uma *palavra única*; (b) não permitem flexão interna; (c) não admitem derivação no 1º constituinte; (d) não mantêm concordância entre os seus componentes; (e) podem, como palavras comuns, servir de base para derivações.

⁵ Na verdade, Villalva aponta o caminho para negar a existência de objetos morfológicos em qualquer língua. A diferença quanto à posição do núcleo, usada como critério por Di Sciullo e Williams, pode ser explicada pela sintaxe: seria a estrutura interna da NP de cada língua o que determinaria a posição do núcleo. No Inglês, o modificador fica sempre à esquerda do núcleo, enquanto no Português o modificador pode ficar à direita ou à esquerda.

Da outra parte, os *compostos pós-lexicais* formam-se no componente pós-lexical (sintaxe) e são *sintaticamente transparentes*: (a) permitem flexão interna; (b) admitem derivações no 1º constituinte; (c) mantêm a concordância entre seus componentes. Só estes últimos é que seriam as palavras sintáticas reanalisadas de que fala Villalva, nascidas da atuação de regras de formação de palavras não-morfológicas, propostas por Di Sciullo e Williams (1987). Segundo ele, essas seriam as configurações dos *pós-lexicais*:

(14) N + (prep.) + N -	sofá cama	homem rã
	trem bala	garota propaganda
	fim de semana	pé de moleque
N + A -	bóia fria	carro forte
	mesa redonda	pão duro
A + N -	curto circuito	primeiro ministro, boa vida
A + A -	surdo mudo	

Em todos estes compostos percebe-se a estrutura sintática que lhes deu origem. Note-se que os dois primeiros grupos têm o núcleo à esquerda; o terceiro tem o núcleo à direita. Em todos eles existe a relação de *modificação*. O último exemplo, *surdo mudo*, foi mal classificado por Lee, como veremos adiante.⁶

Para Lee, existiriam *compostos lexicais*, formados dentro do léxico, correspondentes aos objetos morfológicos de Di Sciullo e Williams. Segundo ele, este grupo de compostos apresentaria as seguintes formações:

(15) N + N -	auto peça
	cine clube
	ferro via
	rádio táxi
A + A -	italo brasileiro
	médico cirúrgico
	econômico social
	sócio cultural
V + N -	guarda chuva
	porta voz
	pára quedas

⁶ Não há razão para colocar em dois grupos distintos, como faz Lee, *surdo mudo* e *médico cirúrgico*. Ambos são compostos A + A, com o mesmo comportamento de seus constituintes. Lee, como veremos, incorreu num velho equívoco de nossas gramáticas descritivas: confundiu *Nomes* e *Adjetivos* compostos.

Qual o critério para afirmar que estes compostos são diferentes dos pós-lexicais? Como vimos acima, Villalva (e Cedeño, para o Espanhol) comprovaram que não é possível fixar uma parametrização rigorosa para a posição do núcleo nos compostos do PB. Para distingui-los, então, Lee lança mão de critérios principalmente *morfológicos*: a possibilidade de flexão (ou não) do 1º membro; a ocorrência de derivação interna; a afixação do DIM ao 1º elemento; e a *concordância* entre os dois elementos. Dentre as evidências que Lee apresenta para se opor à posição de Villalva, duas, de imediato, não podem ser levada em consideração: o *plural* e a *concordância*.

4.1 A flexão interna dos compostos

Lee afirma que nos compostos *lexicais* só há flexão à direita (*rádio táxis, ferro vias, italo brasileiros, guarda chuvas*); nos *pós-lexicais*, pode haver flexão no 1º constituinte: *primeiros ministros, boas vidas, fins de semana, trens bala*, etc.

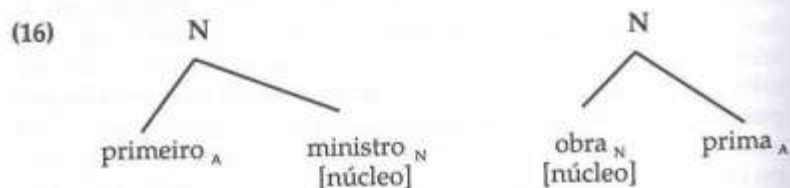
Essa afirmação sugere que sua análise se baseou em um *corpus* muito limitado. Certamente o plural não pode ser usado como evidência segura para uma análise dos compostos no PB, tal é a hesitação que os falantes, como os gramáticos, demonstram ao tentar determinar quais as formas aceitáveis. A razão dessa incerteza é, sem dúvida, o caráter sintático dessas formações, que permite interpretações diferentes das relações entre os constituintes e possibilita aplicações divergentes da regra do plural. Além disso, atua aqui a possibilidade de reanálise de uma geração para outra: no composto em que uma geração percebe X + Y, outra pode perceber só o conjunto (Contreras). Isso fica bem evidenciado nas intermináveis discussões, na tradição gramatical, sobre qual seria o plural de *guarda marinha* (*guardas marinha, guardas marinhas* ou *guarda marinhas*); *navio escola* (*navios escola, navios escolas, navio escolas*); *decreto lei* (se *decretos lei* ou *decretos leis*; por que não *decreto leis*?); entre muitos outros.

4.2 Os compostos N+N

Lee distingue dois tipos de compostos com esta estrutura: (a) *ferro via, auto peça* (lexicais, em que só o 2º flexiona), e (b) *sofá cama, homem rã* (pós-lexicais, em que a flexão ocorre no 1º). Como vimos, essa possibilidade de variação do 1º elemento é usada por Lee como pedra de toque para identificar os pós-lexicais – a flexão interna do composto elimina a possibilidade de considerá-lo como um objeto morfológico uno. Parece-me uma posição absolutamente

correta; no entanto, inadequado é concluir que a *não-variação* do 1º elemento dos compostos de (a) seja suficiente para situar sua formação dentro do léxico e não na sintaxe, como os outros. Em outras palavras: o fato de *homem rã* formar o plural no seu 1º membro (*homens rãs*, ou *homens rã*) e *auto peça* não, é apenas uma evidência positiva a favor da classificação de *homem rã* como pós-lexical, mas não uma evidência negativa que nos leve à conclusão de que *auto peça* seja lexical. Essa diferença de comportamento não nos parece suficiente para colocar os compostos de (a) no léxico e os de (b) na sintaxe. Em ambos está presente a mesma relação *núcleo – modificador*; em (a) o núcleo está à direita, enquanto em (b) ele está à esquerda. Tanto em (a), como em (b), o núcleo apresenta flexão.

O fato de o modificador no PB não ter posição fixa em relação ao substantivo – podendo vir tanto à direita (preferencial) quanto à esquerda, é o responsável pela existência desses compostos que, apesar de parecerem diferentes, são realizações da mesma estrutura de modificação. Isso fica bem claro nos compostos de *substantivo* e *adjetivo*, que podem ser tanto N + A quanto A + N (embora o núcleo seja sempre o N): *salvo conduto*, *primeiro ministro*; *obra prima*, *onça pintada*.



O modificador, como vemos, mantém com o núcleo a relação de concordância característica dos adjetivos em relação aos substantivos. Nota-se no PB a tendência natural de interpretar os compostos segundo as relações mais usuais dentro do SN. Essa estrutura *Substantivo + Adjetivo*, que é uma relação sistemática de modificação no PB, faz com que, em compostos N + N com núcleo à esquerda, um dos constituintes – o não núcleo – seja acertadamente interpretado como A, podendo até ocorrer a concordância com o núcleo (nos exemplos abaixo, ambas as flexões são usuais, mesmo com pessoas letradas):

- (17) escola modelo – escolas modelo ou escolas modelos
 palavra chave – palavras chave ou palavras chaves
 operário padrão – operários padrão ou operários padrões
 café concerto – cafés concerto ou cafés concertos

Silveira Bueno (202) aponta, inclusive, casos de concordância de gênero do 2º constituinte:

- (18) palha (de) trigo palha triga palhas trigas
 palha (de) milho palha milha palhas milhas
 pedra (de) raio pedra raia pedras raias (meteorito)
 sapo (de) concha sapo concho sapos conchos (cágado)

Leite de Vasconcelos, um dos mais atentos pesquisadores da antiga escola filológica, já registrava essa hesitação, documentando casos como *farinha triga*, *palha centeia*, *pera marmela*, *rosa padra* [!] (de *rosa do padre*).⁷

4.3 Compostos N+N, com núcleo à direita

Como vimos acima, E. Carlos Pereira (GH) afirma que os tipos sintéticos (determinante + determinado) são bem raros na composição vernácula, embora constituam o tipo clássico no Latim e no Grego, bem como no Inglês e no Alemão:

- (19) van glória salvo conduto livre arbítrio
 prea mar baixo relevo alto relevo
 baixa mar livre pensador vara pau
 gentil homem clara bóia bom senso

Quando os compostos têm o núcleo à direita, como nos exemplos de Lee, não ocorre a concordância opcional do N que está servindo de modificador:

- (20) auto peças cine clubes ferro vias

Note-se que a posição do núcleo sempre varia (à esquerda ou à direita). O A, quando fica à esquerda, pode deixar de flexionar (*livres pensadores*, *livre pensadores*, *livre arbítrios*, embora *altos relevos*, *primeiros ministros*). O N modificador (antes ou depois) não deveria variar; no entanto, quando fica na direita, é interpretado como *adjetivo* e ocorre a dupla flexão.

A principal característica de compostos como *homem rã* e *guarda noturno*, como vimos, é que o elemento à direita fica numa relação do tipo *modificador* em relação ao da esquerda. Em outras palavras, o membro da direita semanticamente qualifica o da es-

⁷ *Opusculos*, I, 438 – Citado por S. Bueno.

querda. N+N: *homem rã* é um homem que recebe os atributos de uma rã ou age como uma. N + Adj.: *guarda noturno* refere-se a um guarda que exerce suas atividades à noite. Ou no tipo Adj + N, *alto relevo*, o membro da direita é modificado pelo adjetivo da esquerda: é um *relevo* que é *alto*. Os exemplos acima teriam as seguintes interpretações:

- (21) a. Este é um homem que parece ou age como uma rã.
b. Este é um guarda que é noturno.
c. Este é um relevo que é alto.

Aqui há uma relação de qualificação, tanto morfológica quanto semântica, entre os membros do composto. Desta forma, o núcleo pode ser obtido sem fixar parâmetros em qualquer direção.

4.4 A indefinição entre N e A

A falta de limites precisos, no PB, entre Substantivos e Adjetivos também contribuiu para obscurecer o quadro de nossos compostos. Por exemplo, sobre o plural de *democrata cristão*, Luft, o mais lúcido representante dos gramáticos descritivistas, diz que há duas possibilidades: se o interpretarmos como A + A, nos moldes de *médico cirúrgico*, o plural seguirá [[A + A] + s]: *democratas cristãos*; "a outra hipótese – que soa melhor – com flexão dupla – deve ser interpretada como [substantivo + s] [adjetivo + s], isso é, concordância nominal: os democratas (que são) cristãos" – *democratas cristãos*.

Este exemplo ilustra a interferência de aspectos semânticos e sintáticos na decisão quanto à forma do plural – que, no caso, como em muitos outros semelhantes, depende da definição, por parte do falante, de qual é o núcleo do vocábulo: se for A + A, temos uma relação de simples *adição*, num composto sem núcleo (*alguém que é democrata E cristão*); se, por outro lado, for N + A, o composto passa a ter núcleo à esquerda, e a relação é a de *modificação* (o *democrata que é cristão*).⁸ A nosso ver, Lee atribuiu pouca relevância à natureza sintática dos compostos, fato que o levou a examinar esses vocábulos a partir, apenas, da *categoria lexical de seus constituintes*, sem levar em conta que, como toda a tradição gramatical já

⁸ Veja-se, a respeito, a grande quantidade de casos analisados por Zélio dos Santos Jota, em seu *Dicionário de Dificuldades da Língua Portuguesa* (546-86). A tentativa de sistematização deste autor, que representa o pensamento gramatical tradicional, é infrutífera, mas é um eloquente exemplo da total indecisão dos falantes do PB a respeito do plural dos compostos.

advertiu, N e A, no PB, são categorias sem limites precisos. Lemle (102) considera este um dos casos fronteiros de classificação das palavras no Português, já que todo adjetivo que modifica um N referente a um ser humano pode exercer o papel de nome, "incorporando o conceito de pessoa ao seu próprio sentido, que passa a ser: uma pessoa com a qualidade expressa pelo adjetivo" (os *ricos*, os *maus*, o *avarento*, o *brasileiro*, etc.).⁹ Desta forma, um A que forma com um N um sintagma muito coeso semanticamente e de uso muito freqüente pode ocupar o núcleo de um SN, incorporando o sentido do adjetivo e do nome:

- (22) O *pobre* sofre mais com frio. Os *franceses* têm muito orgulho de sua literatura.

É evidente que isso também ocorre com os compostos. É o caso de *surdo mudo*, que já perturbou muitos gramáticos e veio também atrapalhar a análise de Lee. Este vocábulo não é, como ele afirma, um A + A, pós-lexical, que teria a flexão *surdos mudos*. Aqui vale a lição definitiva de C. P. Luft: quando é *surdo* que ocupa o núcleo do SN *surdo mudo*, sua leitura é "o *surdo* que também é *mudo*", ou seja, é interpretado como N + A (núcleo + modificador), ocorrendo naturalmente a flexão do 1º elemento e a obrigatória concordância do 2º: o *surdo mudo*, a *surda muda*, os *surdos mudos*, as *surdas mudas*. Quando, no entanto, *surdo mudo* é um A (formado de A + A, interpretado como a coordenação de *surdo* e *mudo*; é um composto sem núcleo, exocêntrico – *alguém é surdo e mudo*), passa a comportar-se como qualquer adjetivo composto: menino *surdo-mudo*, menina *surdo-muda*, meninos *surdo-mudos*, meninas *surdo-mudas*. Da mesma forma (exemplifica Luft), (a) os *azuis celestes*, os *vermelhos intensos* (N), contra (b) vestidos *azul celeste*, flores *vermelho intenso* (A). Em (b), temos uma construção sintática tradicional no Português: "nossas genuínas maneiras de dizer fazem-se com o auxílio da preposição *de* ou das locuções *de cor*, *de cor de*, *da cor de*" (Sousa da Silveira, citado por Luft). No exemplo acima, temos "flores de um *vermelho intenso*", ou "flores *da cor de vermelho intenso*", o que explica a invariabilidade.

Na verdade, *surdo mudo*, quando núcleo do SN, tem a mesma construção dos outros compostos N+A ou A+N (e não A + A, co-

⁹ Esse processo ocorre mesmo com certos adjetivos referentes a [- HUM], que acompanham com tal freqüência determinados nomes que "acabam incorporando em si o sentido desse nome, cuja presença física torna-se até dispensável" (Lemle, 102): o *combustível* (material), os *móveis* e *imóveis* (bens), um *longo* (vestido), e assim por diante.

mo Lee interpretou); é na *sintaxe* que reside a explicação desse e de outros exemplos de N (seja um *derivado*, seja um *composto*) que são usados como modificadores:

(23)

- | | |
|--------------------|--|
| (a) o prodígio | menino prodígio |
| os prodígios | meninos prodígio [? meninos prodígios] |
| (b) o puro sangue | cavalo puro sangue |
| os puros sangues | cavalos puro sangue |
| | [?* cavalos puros sangues] |
| o pele vermelha | índio pele vermelha |
| os peles vermelhas | índios pele vermelha |
| | [?* índios peles vermelhas] |

Quando usados como modificador, sua invariabilidade é idêntica a de qualquer substantivo usado na mesma posição sintática: *meninos prodígio* (ou, contra a norma escolar, **meninos prodígios*); *funcionários fantasma* (ou **funcionários fantasmas*). No entanto, como vimos, esta é uma leitura educada; é cada vez mais freqüente a reanálise desse tipo de construção, dando ao N da direita a mesma concordância que qualquer modificador assumiria, ocorrendo a flexão também do 2º elemento: *palavras chaves*, *horas aulas*, *cor de rosas*, etc. e, *ipso facto*, *cavalos puros sangues*, *cavalos puro sangues*, *índios peles vermelhas*, *índios pele vermelhas*.

4.5 Os compostos A+A (x E y)

Lee situa os compostos do tipo A + A entre os compostos *lexicais*; uma das evidências, diz ele, é a "falta de concordância" entre os constituintes. Enquanto nos *pós-lexicais* do tipo A + A, A + N e N + A "o gênero e o número de seus constituintes sempre coincidem" (65): *surdos mudos*, *mesas redondas*, *boas vidas*, nos compostos *lexicais* *italo brasileiro*, *italo brasileira*, *italo brasileiros*; *judeu americano*, *judeu americana* esta concordância não ocorre. Aqui também parece ter havido equívoco nos dados de que o autor dispunha: os compostos A + A (sem entrar no mérito de serem *lexicais* ou *pós-lexicais*) sempre terão a flexão apenas no 2º constituinte. Isso não é razão suficiente para considerar que, neste caso, os elementos flexionais são afixados ao composto inteiro, não apenas ao 2º constituinte. O 1º membro fica sempre numa forma fixa, *invariável*; o 2º membro é que apresenta flexão:

- | | |
|---------------------------------|---|
| (24) jantar político partidário | reunião político partidária |
| jantares político partidários | reuniões político partidárias ¹⁰ |

A gramática escolar afirma que a regra geral para a flexão dos compostos A + A é a seguinte: o 1º elemento ficará sempre neutro (forma não marcada: masc./sing), enquanto o 2º recebe a flexão:

- | | | |
|----------------|-------------------------------|-------------|
| (25) conceitos | político _{m. sing} | partidários |
| complexos | industrial _{m. sing} | militares |

Não nos parece ser o que acontece, já que se percebe, muitas vezes, uma verdadeira *redução* posterior do 1º elemento do composto: há a supressão do sufixo, igual ou similar ao que está presente no 2º elemento. Este fenômeno já foi observado para o Italiano (Scalise; Thornton):¹¹

- | | | |
|-------------------------|---|---------------------|
| austriaco-ungarico | → | austro-ungarico |
| orticolo + frutticolo | → | ortofruticcolo |
| agrario + minerario | → | agro-minerario |
| anarquista + socialista | → | anarco - socialista |
| vegetale + minerale | → | vegeto-minerale |

No Português, no lugar do sufixo, também aparece, quase invariavelmente, o Elemento Terminal -o:

- | | | |
|-------------------------|---|--------------------|
| (26) infantil + juvenil | > | infanto juvenil |
| verbal + nominal | > | verbo nominal |
| cerebral + visceral | > | cérebro visceral |
| labial + dental | > | lábio dental |
| muscular + membranoso | > | músculo membranoso |

Isso poderia erroneamente sugerir a hipótese de que não há, na verdade, supressão do sufixo do 1º elemento, mas apenas o acréscimo do sufixo ao elemento da direita; o adjetivo teria a seguinte derivação:

¹⁰ Poderíamos, seguindo a linha de Luft, distinguir uma reunião *político partidária* de uma reunião *política partidária*: *reunião [política [e] partidária]* vs *[reunião política] partidária*.

¹¹ Há casos, contudo, em que a regra não se aplica. *Nazi-Fascista* não recebe este -o. Além disso, não se pode dizer que este vocábulo tenha sido formado por apagamento do sufixo, porque o /i/ é claramente uma parte do sufixo -ista - o que seria inexplicável se fosse o sufixo que tivesse sofrido o apagamento. Se dissermos, contudo, que uma regra da Morfologia Prosódica - que Thornton denomina de *accorciamento* - reduziu o 1º membro ao tamanho do voc. mínimo, o /i/ é explicado.

[cérebro + víscera] al]_A
[[lábio + dente] al]_A

A posição do acento seria, desta forma, mais facilmente explicada; o sufixo formador do adjetivo seria acrescido ao composto *uma só vez, à direita*. Aparentemente, isso favoreceria a idéia de Lee de que estes compostos se formam no léxico. No entanto, essa hipótese não se sustenta com exemplos como *materno infantil*, resultante da coordenação de *maternal* e *infantil* (sendo ambos adjetivos, com sufixo), e não de *materno* e *infante* (adjetivo e substantivo), porque compostos coordenados serão sempre [X e Y]. Defendemos que estes compostos sofrem a perda do sufixo do primeiro membro *depois de formados*. Em fenômeno similar ao de uma fatoração, o 1º radical não precisa apresentar o sufixo; este só aparece no membro à direita, mas vale para os dois. Além disso, colaborando com a idéia de que há uma redução *posterior* do composto, persiste sempre a possibilidade de não se formar o composto, deixando transparente a estrutura sintática deste tipo de coordenação: *labial e dental, maternal e infantil*

Este tipo de construção em que o 1º membro, geralmente de origem latina, é reduzido teve amplo emprego literário, erudito, para formar adjetivos indicando relações entre diferentes nacionalidades (Câmara Jr., 1975, p. 216):

(27) franco prussiano	indo americano	nipo chinês
anglo americano	teuto brasileiro	sino japonês
luso brasileiro	nipo germânico	euro africano
afro brasileiro	teuto nipônico	afro europeu

Os gentílicos reduzidos (*teuto, anglo, sino, franco*, etc.) vêm do Latim (*teutoni, angli, sinae, francus*) e estão, por isso mesmo, presentes em vocábulos de todas as línguas ocidentais. Podemos acrescentar ao grupo outros elementos fixos de composição, também de uso próprio da comunidade científica: *euro, infero, súpero, gênito, vésico*, etc.¹² São, portanto, itens lexicais que, à semelhança dos vocábulos comuns, saem prontos do Nível 1 e podem participar da formação de compostos.

Note-se que a ordem dos elementos não é fixa – a aliança entre a Alemanha e o Japão poderia ser denominada de *nipo germânica*, ou *nipo alemã*, ou *teuto nipônica*, ou *teuto japonesa*. Além disso, é

¹² O Formulário Ortográfico que abre o PVOLP (1943) chama *anglo, greco, infero, póstero, súpero*, etc., curiosamente, de “prefixos que representam formas adjetivas”.

restrito o número de adjetivos gentílicos que têm essa forma reduzida: *argentino, boliviano, uruguaio*, p. ex., não têm; para *brasileiro*, já se experimentou *braso, brasílio e brasílico*, certamente por analogia, mas seu emprego não passou de tentativas isoladas. Emprega-se, com naturalidade, “amizade *brasileiro paraguaia*”, “relação *argentino chilena*”, seguindo-se a tendência atual de forma compostos de A + A com ambos os elementos intactos:

(28) técnico burocrático	médico cirúrgico
político partidário	cubano soviético
cívico eleitoral	econômico financeiro
médico odontológico	árabe israelense

Havendo, como afirmamos, *redução* (ou *supressão*), seria este um fortíssimo argumento para a hipótese da formação dos compostos *fora do léxico*: os sufixos já se ligaram a ambos os radicais que entram na combinação; só depois que os vocábulos *prontos* coordenaram-se para formar o novo vocábulo é que ocorre a redução. Isso, aliás, é possível só no *campopós-lexical*, pois o processo morfológico do PB não admite regra de supressão de morfema.

Podemos assumir que a *supressão do sufixo se dá na sintaxe*; o composto se *lexicaliza* já com a sua nova forma, e os dois componentes, portanto, são reanalisados no Nível 2, onde o 1º membro recebe o ET “-o” e volta a receber o acento automático.

Também é necessário justificar por que em vários vocábulos o acento do 1º é proparoxítono:

(29) vegetal animal	>	végeto animal
literário recreativo	>	lítero recreativo
inferiór anterior	>	ínfero anterior
posteriór palatal	>	póstero palatal

Essa posição marcada do acento, na antepenúltima sílaba, condiz com o caráter alatinado, erudito, dessas formações autônomas de nosso léxico, próprias do vocabulário científico. Podemos considerá-los casos especiais, fora do sistema do PB – tanto que existem, como vimos acima, com a mesma acentuação, em outras línguas (Ing. *infero lateral, gênito urinary, vésico* – ver OED; Fr. *gênito-urinaire, vésico-rectal*; etc.), sendo utilizados nas mesmas circunstâncias.

Seja reduzido, seja integral, o adjetivo composto, uma vez formado, passa a flexionar apenas *à direita*. O que explica essa ausência de flexão do 1º membro? Onde houve redução, ainda se

poderia apontá-la como a razão: causas *sociais e econômicas*, mas causas *sócio-econômicas* (*sócio*, que é a redução de *social*, fica invariável). Entretanto, nos compostos em que os dois adjetivos estão meramente coordenados, ambos com seus sufixos, qual seria o motivo? Por que “conseqüências *econômicas e financeiras*”, mas “*econômico financeiras*”? Se postularmos que houve um alçamento destes compostos para o Nível 2, a presença da marca de flexão apenas na borda direita do composto está de acordo com nosso sistema flexional.

Este é um tipo de composto extremamente instável, que pode ser desfeito a qualquer momento sem que haja *descomposição semântica*. Neles, é quase impossível distinguir o que é *frase* e o que é *vocabulo*. *Econômicas e financeiras* ou *econômico financeiras* – a escolha entre as duas alternativas não parece alterar nada do significado da expressão, como ocorreria com *beija flor*, em que, apesar da persistência do significado dos componentes (já que é uma ave que *beija as flores*), o composto é facilmente distinguível de uma seqüência “x *beija flor*”. Uma reunião *política e partidária* é o mesmo que uma reunião *político partidária*. Sandmann (139) já havia observado que os dois constituintes estão rigorosamente no mesmo nível – não há Determinante, nem Determinado; a prova é que se pode inserir uma conjunção (E) entre eles, como vimos. Além disso, pode-se formar um composto com *vários* componentes: *franco luso brasileiro, sócio econômico culturais*. Acrescento a ausência de uma ordem fixa entre os componentes:

- (30) relação *palestino israelense* ou *israelense palestina*¹³
causas *sócio políticas* ou *político sociais*

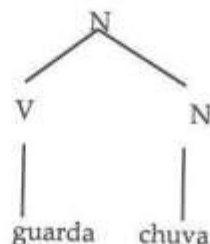
Por ironia, foi este composto A + A um dos exemplos usados por Lee como argumento para postular a existência de *compostos lexicais* – exatamente estes que parecem mais próximos da sintaxe que os demais. Como afirmamos acima, a *presença* de flexão interna é um sinal seguro do caráter *pós-lexical* de qualquer composto; contudo, a *ausência* da flexão nesta posição *não* é indício necessário de que o composto tenha sido formado antes da sintaxe (portanto, como quer Lee, no componente *lexical* – seja *cíclico*, seja *pós-cíclico*).

¹³ Nos conflitos ocorridos em Jerusalém, em setembro de 1996, o jornal ZH usou, em manchete, a forma “*israelo - árabe*”.

4.6 Os compostos V+Comp

A divisão defendida por Lee, necessária para a justificativa dos dois tipos de compostos, esbarra também na errônea interpretação dos compostos de V+N. Ao analisá-los, Lee baseia-se em Câmara Jr. para dizer que o elemento verbal é representado, nesses casos, pelo *radical verbal + vogal temática* – ou seja, “não carrega traços de modo, tempo, pessoa ou número” (Lee, 70). Isso lhe permite incluí-los entre os *compostos lexicais*, já que uma das características principais deste grupo é a ausência de flexão entre os constituintes. Dessa forma, estariam situados, segundo o modelo adotado por Lee, no *nível a*, onde coloca “todos os processos derivacionais, a flexão irregular e alguns processos de composição, aos quais se podem acrescentar os sufixos derivacionais” (12). Por essa razão, em *guarda chuva* Lee encontra apenas 1 vocábulo morfológico, com a seguinte estrutura:

(31)



Já em *vaivém* – composto V + V – reconhece que as formas verbais *vai* e *vem* deixam entrever a presença da flexão. Comparando (a) *vaivém* com (b) *quebra quebra*, argumenta que em (a) isso não traria problema, pois a flexão verbal irregular acontece no mesmo nível em que ocorre a composição (*nível a*); no entanto, mesmo assim Lee reluta em admitir que o V em *vaivém* tenha traços de tempo/modo, pessoa/número, pois isso implicaria também admitir que o mesmo acontece com formas verbais como *quebra quebra* – o que prejudica sua análise, que situa a flexão *regular* no *nível β*. Para preservar seu modelo, admite (aliás, acertadamente) que a solução seria classificar este tipo de composto entre os *compostos lexicalizados*.

Sua análise, no entanto, está em desacordo com a compreensão que os falantes têm dos vocábulos V + Comp. Nos compostos V+N, há uma estrutura sintática usual de *Vtrans. + complemento*. O constituinte verbal, neste caso, é a 3ª *pessoa do singular* – e não, como Câmara Jr. afirmava, o simples tema do verbo. Gramáticos

tradicionais discutiam se essa forma verbal era Imperativo ou Indicativo, uma vez que na maioria dos casos essas duas formas são idênticas. Embora haja exemplos de imperativo no PB, a esmagadora maioria é sentida como Presente do Indicativo. O que não há dúvida é de que se trata de uma forma verbal com traços de flexão:

- (32) (a) vem-cá (caranguejo)
salve-rainha
não-me-toques (arbusto)
- (b) foge-foge (correria)
põe-mesa (inseto)
cospe-cospe (peixe)
vai-não-vai (indecisão)
vai-da-valsa (ir no vai-da-valsa: ao léu)
- (c) lança-perfume
toca-discos
mata-borrão
quebra-nozes
vira-lata
conta-gotas

Em (a), temos o Imperativo; em (b), o verbo está claramente conjugado no Presente, 3ª pessoa do sing. (cf. foge, fugir; põe, pôr; cospe, cuspir; vai, ir); tudo nos permite pensar que acontece o mesmo em (c). Aliás, essa é a interpretação natural desses compostos:

- (33) (aparelho que) toca discos
(papel que) mata borrão
(cachorro que) vira lata
(utensílio que) quebra nozes

Estes fatos vêm reforçar nossa posição de que os compostos V+N, como também os demais, são formados como estruturas sintáticas que se lexicalizam posteriormente.

4.7 Derivação dos compostos

Na teoria de Lee, uma das características mais marcantes dos compostos que ele chama de *lexicais* é que, ao contrário dos *pós-lexicais*, deles podem formar-se derivados por sufixação:

(34)

[[fotograf]N ar]V [[puxa saco]A]ismo]N [[rádio táx]N ista]N

Para Lee, os compostos que chama de *pós-lexicais* não podem fazê-lo, a não ser por raros prefixos: ex[homem rã], super[primeiro ministro]. Como já observamos, Lee pensa que aí reside uma das peculiaridades que justificam a postulação de que um grupo de compostos formado dentro do léxico (os objetos morfológicos) constituem a base para processos derivacionais:

- (35) pão duro - pão durismo pão duragem
dedo duro - dedurismo dedurar
rádio táxi - rádio taxista
puxa saco - puxa saquismo
pára quedas - pára quedismo

Por não levar em consideração o quadro fluido do plural dos compostos no PB, Lee aqui se depara com a o problema de *pão duro*, *dedo duro*. Segundo seu critério, estes vocábulos seriam compostos *pós-lexicais* por admitirem morfema plural entre seus constituintes, mas, ao contrário do que seria de se esperar em sua análise, também podem servir de base para derivações como *pão durismo*, *dedurar* (por haplogogia de *dedodurar*). Para justificar esses contra-exemplos, propõe considerá-los *compostos lexicalizados*, que ele equipara a expressões idiomáticas, "estruturas lexicalizadas, passando a integrar o léxico da língua como se fossem itens *lexicais*" – do tipo *conto do vigário*: ao derivar *pão durismo*, a entrada seria *pão duro*, de tal maneira que não acontece **pães durismos*. Contudo, essa explicação, apesar de sua simplicidade, é rejeitada por ele, que a julga circular, uma vez que implicaria a existência de três entradas *lexicais* diferentes: *pão*, *duro* e *pão duro*. Sem aceitar essa conclusão – que nos parece a única satisfatória – Lee conclui que é melhor deixar em aberto a questão dos compostos lexicalizados.¹⁴ Sua análise não percebeu que problema aqui é o mesmo de qualquer derivação a partir de compostos, como *guarda chuvada*, *porra louquice*, *puxa saquismo*, etc. – parece não ser possível evitar a inclusão do composto como *entrada lexical*, como pretende Lee.

¹⁴ A flexão que ele registra como significativa, *pães duros*, e sobre a qual baseia seu raciocínio, não é a usual no PB, que prefere *pão duros*, *dedo duros* (no feminino, *pão dura*, *pão duras*).

Esta capacidade de formar derivações – que Lee aponta como uma característica que separa os compostos lexicais dos demais – é o ponto crucial da análise dos compostos, não para classificá-los em grupos distintos, mas para ajudar-nos a determinar a posição desse tipo de vocábulo na estrutura do léxico do PB. Como acontece com qualquer palavra comum, certos compostos, ao se lexicalizarem, alçando-se ao Nível 2, podem formar palavras novas através de afixação:

- | | | | |
|------|-----------------|---|-------------------------|
| (36) | porra louca | – | porra louquice |
| | guarda chuva | – | guarda chuvada |
| | bossa nova | – | bossa novista |
| | sem vergonha | – | sem vergonhice |
| | fã clube | – | fã clubista |
| | dom juan | – | dom juanismo |
| | primeiro mundo | – | primeiro mundista |
| | meio campo | – | meio campista |
| | estado novo | – | estado novista |
| | sem cerimônia | – | sem cerimonioso |
| | cara dura | – | cara durismo |
| | cinema novo | – | cinema novista |
| | puxa puxa | – | puxa puxeiro |
| | quinta essência | – | quinta essencial |
| | pai d'égua | – | paidegual, apaideguado. |

Uma solução seria postular que é o 2º elemento que sofre a derivação, não o composto todo. Os dois radicais constituintes não passam por todo o ciclo. Shaw diz que cada constituinte do composto passa por suas derivações normais, e só depois os componentes se reúnem para formar um só vocábulo (esta unidade não é assegurada por critérios morfológicos ou fonológicos, mas sintáticos e, acima de tudo, semânticos). No entanto, os dados do PB sugerem que essa solução, apesar de adequada, necessita incluir o fato de que o composto, depois de formado, foi lexicalizado, e a derivação olha agora para o composto como um todo. Por exemplo, do composto *bossa nova* (/bOssa nOva/) é possível derivar *bossa novista* (/bOssa novista/). O sufixo *-ista* liga-se comumente a nomes de movimentos do mundo artístico ou do mundo intelectual, para indicar os seus adeptos ou seguidores (*impressionista, marxista, budista*). Como vemos, é o composto *bossa nova* (importante movimento de renovação da música popular brasileira dos anos 60), e não apenas o adjetivo *nova*, que vai servir de base para a derivação, que se processa com a sufixação regular à direita – ou seja, com a

afixação apenas ao 2º elemento, que será reexaminado pela regra do acento e sofrerá a mudança automática da vogal média /O/ → /o/. No 1º membro, *bossa*, o acento atribuído continua, podendo manter, *ipso facto*, a vogal /O/. Para que ele não sofra também a regra do acento, será necessário postular uma regra que mantenha o limite do vocábulo prosódico – o que conseguimos facilmente admitindo que o composto foi incluído no Nível do Vocábulo por alçamento. O composto, uma vez formado, passa a ser também entrada lexical e, como tal, pode servir de base para as derivações. O mecanismo dessa derivação vai depender do grau de estabilidade do composto, mas, de qualquer forma, ela só se aplica ao membro da direita, porque o composto está sendo visto como um todo.

Uma forte evidência de que estamos diante de um novo vocábulo são as formações sufixais que se estabelecem: as derivações *saquismo* (puxa saquismo), *durismo* (cara durismo, pão durismo), *louquice* (bicha louquice), *vergonhice* (sem vergonhice), por exemplo, não existem como fomas livres no léxico. Nestes exemplos, os sufixos *-ismo* e *-ice* estão usados como pejorativos, exatamente como em *consumismo, machismo, alarmismo, ou criancice, gramatiquice, bacharellice* – em consonância com a carga pejorativa dos compostos *puxa saco, cara dura, sem vergonha*, que serviram de base para a derivação:

- (37) [[cara] [dura]] ismo
 [[sem] [vergonha]] ice

No mesmo sentido, em *guarda chuvada*, o sufixo *-ada*, usado em formações que significam “golpe dado com”, foi adicionado ao objeto *guarda chuva*:

- (38) [[guarda] [chuva]] ada

Também é significativo que o composto, quando é Adjetivo, admite a formação especial com *-íssimo*:

- (39) pão duro – pão duríssimo grã fino – grã finíssimo

Se nestes casos é possível alegar que o sufixo se acrescenta normalmente ao adjetivo que está à direita – o que seria normal, no quadro do PB –, em *sem vergonhíssimo*, por exemplo, é o Substantivo *vergonha* que recebe *-íssimo* (que só se liga a adjetivos), confirmando o fato de que o composto já entrou no léxico e passou a ser interpretado como uma só unidade lexical: a derivação está olhando para o todo. Se a base para a derivação é o composto, resta ex-

plicar de que maneira só o 2º membro é considerado nas operações morfológicas e fonológicas. Depois de entrar no léxico, os compostos – em variadas gradações – parecem perder o limite *morfológico* e são vistos como palavra *simples* (embora conservem os limites prosódicos), o que explica a derivação normal à direita.

O que resta indiscutível é que a maioria deles continua apresentando *dois* vocábulos prosódicos (o que fica evidente pela manutenção dos acentos, a não ocorrência da neutralização das vogais, etc.).

Admitindo-se o movimento de regras de um componente para outro, segundo Kiparsky, 1982; Kaisse e Shaw, 1985; Halle e Mohanan, 1985; Pulleyblank, 1985; Rubach, 1985, entre outros, a evolução típica de um composto do PB parece passar por estágios que se ajustam à proposta da Fonologia Lexical:

- **estágio 1** – O germe do novo composto, formado na sintaxe, constitui um sintagma comum que estende sua aplicação, com certa regularidade, no nível pós-lexical.
- **estágio 2** – O composto entra no léxico, no nível 2 (o do *Vocábulo*) como duas palavras que mantêm seus domínios prosódicos, com *Dim* e flexão individuais. (Alguns terminam seu processo lexical aqui).
- **estágio 3** – O composto é alçado para o nível 1 (o do *Radical*), onde se dá a passagem dos dois domínios prosódicos a um só, ajustando-se à Morfologia deste nível. O *Dim* e a flexão são transferidos para a direita.

Este modelo aponta para os vários pontos de imprecisão que encontramos na análise dos compostos na maioria das línguas. Por exemplo, a passagem de um estágio para outro reflete a fronteira imprecisa entre *sintagma* e *vocábulo* (*papel almaço*, *papel cuchê*, *papel acetinado* são vocábulos?), assim como hesitações já registradas quanto ao plural dos compostos (*livres pensadores* ou *livre pensadores*; *sofá caminha* ou *sofazinho cama*?). Quando o composto entra no terceiro estágio, ocorre, muitas vezes, a perda da consciência da composição, uma vez que a perda do acento do 1º elemento dificulta o seu reconhecimento como radical isolado.¹⁵

Esta hipótese reflete a inegável importância da *história* individual de cada composto e parece ser a justificativa para a antiga distinção entre *justaposição* e *aglutinação*; esta última vai situar-se, na verdade, no 3º estágio de nosso modelo. Câmara Jr. (1969) já

tinha observado que “a aglutinação é a morte de uma justaposição na história da língua”. Mateus é mais específica: “A distinção entre *justaposição* e *aglutinação*, que se baseia no fato de a estrutura fonológica dos compostos conter ou não fronteiras de palavra, resulta, provavelmente, do fato de essa transformação ser progressiva e gradual e, por hipótese, dependente da evolução do valor referencial do composto” (390).

Vocábulos como *pernalta*, *pernilongo*, *planalto*, *pontiagudo*, *boquiaberto* representariam exemplos do último estágio dessa evolução:

- (40) 1º [pláno]_N + [áto]_A
 2º [[pláno] [áto]]_N
 3º [planálto]_N

Vemos, assim, que a grande variedade do quadro dos compostos não se deve aos diferentes níveis em que seriam formados; não há compostos *lexicais* e *pós-lexicais*, como quer Lee. *Todos os compostos são formados no nível pós-lexical*, e sua diversidade se deve aos elementos que integram cada um deles e, mais ainda, ao *estágio de lexicalização* em que o composto se encontra. Isso é que explica as diferentes análises que deles faz o falante – com as variantes, as gradações intermediárias, as hesitações – e deixa prejudicada qualquer tentativa de uma análise puramente sincrônica, como tentou Lee.

4.8 Diminutivo dos compostos

Pelo que expusemos na seção anterior, também ficam prejudicadas as conclusões de Lee sobre a formação do DIM. A formação do DIM no 1º ou no 2º membro não nos autoriza a distinguir compostos lexicais de pós-lexicais (para ele, os primeiros se comportam como elemento único: *guarda roupinha*; os pós-lexicais têm um núcleo que pode conservar seu estatuto de palavra independente: *homenzinho rã*, *horinha extra*). Poderíamos postular que o DIM vai-se aplicar sempre ao *núcleo*, quando houver; nos compostos *sem núcleo* (como V+Comp), o DIM só aparece à direita. Contudo, a tendência, como vimos acima, de lexicalizar os compostos, alçando-os para o Nível 2, faz com que as formas de (37) coexistam: a diferença entre (a) e (b) é apenas o nível de lexicalização que se está atribuindo aos compostos listados. Os exemplos em (b) são tratados como vocábulos já pertencentes ao Nível 2, o Nível do *Vocábulo*:

¹⁵ Com certeza não são todos os falantes que reconhecem em *corrimão* ou *planalto* os radicais que os constituem (*correr* + *mão*, *plano* + *alto*).

(41)		(a)		(b)	
	sofá cama	-	sofazinho cama	ou	sofá caminha
	obra prima	-	obrinha prima	ou	obra priminha
	peixe rei	-	peixinho rei	ou	peixe reizinho
	pão de ló	-	pãozinho de ló	ou	pão de lozinho
	homem rã	-	homenzinho rã	ou	homem rãzinha

Quanto à formação de domínios morfológicos e domínios prosódicos, Lee propõe a seguinte taxonomia para os compostos:

(a) <i>composição I</i> (ferrovia, espaçonave)	1 morfológico 1 prosódico (só 1 acento)
(b) <i>composição II</i> (puxa saco)	1 morfológico 2 prosódicos (2 acentos)
(c) <i>pós-lexical</i> (primeiro ministro)	2 morfológicos 2 prosódicos (2 acentos)

Se concordarmos com Villalva, que afirma que *todos* os compostos são *palavras sintáticas* reanalisadas, não vemos razão para aplicar ao PB essa distinção entre *compostos lexicais* (objetos morfológicos) e *compostos pós-lexicais* (palavras sintáticas). No PB, não há duas instâncias diferentes para a criação de compostos, uma no léxico e outra fora dele. *Todos os compostos se formam na sintaxe*. O quadro acima não corresponde, como é a intenção de seu autor, aos níveis de formação dos compostos; ele apenas reflete os diferentes estágios em que os compostos se encontram: os dois tipos diferentes que Lee encontrou no PB e classificou segundo a distinção de Williams não se devem ao fato de ter sido *criados* em níveis diferentes, mas sim ao fato de se encontrarem em diferentes *estágios* dessa lexicalização progressiva.

5 Conclusão

Todos os compostos nascem como palavras sintáticas – ou melhor, como sintagmas; todos vão-se lexicalizando; todos apresentam, inicialmente, a marca de sua origem: dois vocábulos morfológicos e fonológicos prontos, que se uniram semanticamente. Não se trata, portanto, de dois radicais que se concatenam por uma regra WRF, mas de dois vocábulos, com elementos terminais e flexão próprios.

Parece que há um momento em que determinada *seqüência* passa a ser encarada como um *todo* – e, se falamos em *momento*, entramos obrigatoriamente na dimensão temporal: é um processo; há uma *passagem* da sintaxe para o léxico, o que nos permite supor que diferentes gerações ou mesmo falantes individuais façam análises diferentes de uma mesma seqüência de vocábulos. Uma vez formado o composto, passa a fazer parte do léxico e pode ele mesmo servir de base para derivações, embora na maioria dos casos continue guardando os limites dos vocábulos morfológicos que o constituem. Acrescente-se que esta derivação a partir de vocábulos compostos é pouco produtiva, à exceção dos adjetivos gentilícios derivados de topônimos compostos. Nota-se uma direção de tendência bem clara: a progressiva estabilização do item assim formado, em direção à *unidade*, até chegar ao ponto em que se perca a consciência da composição e todos os processos morfológicos e fonológicos individuais do 1º elemento deixem de se manifestar; qualquer aplicação de regra, a partir daí, vai se efetuar sobre o todo indivisível.

Este processo de lexicalização do composto pode ser percebido em vários exemplos em que os falantes fazem diferentes interpretações do vocábulo já formado, atribuindo maior ou menor transparência aos seus constituintes, pronunciando-os ora como um todo, ora mantendo a individualidade de seus constituintes.

No PB, no dialeto do Sul, por exemplo, temos compostos como

(a)	tele fône	(b)	téle atriz
	tele patía		téle novela
	tele visão		téle tema
	tele grâma		téle jornalismo

Os vocábulos do grupo (a) são compostos que já atingiram o *último estágio* do processo. Constituem um único vocábulo morfológico e, sem dúvida, um único vocábulo prosódico, como se pode perceber pela existência de um só acento e pelo fechamento da vogal média do radical /tEle/. A partir da década de 70, contudo, este radical passou a ser usado com outro significado além do original (“distância”), em vocábulos em que significa, de maneira abreviada, *televisão*. Nos vocábulos do grupo (b), a composição ainda está transparente, com o acento duplo e a manutenção da vogal aberta. Mais recentemente, a oferta de serviços por telefone produziu um novo grupo de compostos com *tele*, já aqui signifi-

cando *telefone*, com as mesmas características dos compostos do grupo (b): /têle chavéiro/, /têle serviço/, /têle amigos/¹⁶

Outra evidência dessa verdadeira flutuação da análise dos compostos pode ser encontrada na ortografia, principalmente no discutido emprego do hífen. Como já observei anteriormente (Moreno, 1977), não existe um critério unitário quanto ao seu emprego; os ortógrafos divergem entre si e do que ficou estabelecido no PVOLP de 1943. Câmara Jr. (1969) afirma que "o emprego deste sinal gráfico é incoerente e confuso". No dicionário Aurélio Eletrônico, por exemplo, encontramos

(a) pedra angular	(b) pedra - mármore
pedra filosofal	pedra - pomes
pedra lascada	pedra - imã
pedra de toque	pedra - sabão

É evidente que, em muitos casos, o hífen está sendo usado para distinguir o *composto* do simples *sintagma*, principalmente quando existe uma clara individualidade semântica: *cachorro-quente*, *pé-de-moleque*. Contudo, escrevemos *sanguessuga*, *girassol*, *passatempo*, *vai-vém*, mas *pára - sol*, *sangue - frio*, *passa - pé*, *vai - volta*, sem nenhuma motivação semântica. Ao lado de *vai-vém*, encontramos a forma paralela *vai-e-vem*; escreve-se *anteontem* e *antes de ontem*.¹⁷

É inútil tentar ver os compostos como o produto de uma regra de formação de palavra (RFP) que tenha a propriedade de unir dois radicais. Não haveria, assim, como explicar por que alguns radicais se combinam e outros não. Na verdade, os constituintes do composto são *vocábulos prontos* (já flexionados, ou com Elemento Terminal – em suma, prontos: o *output do componente lexical*) que se unem por razões semânticas, geralmente mantendo a estrutura sintática que lhes deu origem. Com o passar do tempo – portanto, diacronicamente –, vão assumindo traços de unidade morfológica e fonológica, à medida que desaparece a consciência de sua construção. O todo é reanalisado como uma coisa só, e começam as *perdas* dos traços individuais de cada componente, especialmente

¹⁶ Numa sala de aula com 100 alunos, solicitei que lessem em voz alta 4 vocábulos escritos em quadro-negro. Todos, em uníssono, leram *telefone*, *telepatia*, mas *têle-avêla*, *têlechavéiro*. Da mesma forma, ao ditar *oleoduto* como /oleoduto/, a maioria pediu confirmação de que eu estava ditando o que eles chamam de /Óleo ditto/.

¹⁷ Este é o motivo por que deliberadamente evito usar hífen ao grafar os compostos examinados no presente trabalho.

do elemento da esquerda (mesmo quando este era o núcleo), em vista de sua posição de 1º membro de uma formação do Nível 2. Muda o acento, muda o vocalismo, cai o Elemento Terminal, etc.

Lembro que essa possibilidade de ingressarem itens pós-lexicais no léxico já foi sugerida por muitos autores (Kiparsky, 1982; Kaisse e Shaw, 1985; Halle e Mohanan, 1985; Pulleyblank, 1985; Rubach, 1985; entre outros), que sentiram a necessidade de admitir alguma recursividade do nível sintático para a morfologia, quase sempre – e não por acaso – para lidar com problemas relativos aos compostos.

Bibliografia básica

- CÂMARA JR., J. M. *Estrutura da Língua Portuguesa*. Petrópolis: Vozes, 1970.
- . *História e estrutura da Língua Portuguesa*. Rio de Janeiro: Padrão, 1975.
- CEDENO, Rafael. Headship assignment resolution in Spanish compounds. In: CAMPOS, H.; MARTÍNEZ-GIL, F. (eds.). *Current studies in Spanish Linguistics*. Washington: Georgetown University, 1985. p. 573-598.
- DI SCIULLO, A.; WILLIAMS, Edwin. *On the definition of word*. Cambridge: MIT, 1987. 118 p.
- LEE, Seung-Hwa. *Morfologia e Fonologia Lexical do Português do Brasil*. Campinas, 1995. 190 p. (Tese de doutoramento)
- LEMLE, Miriam. *Análise Sintática; teoria geral e descrição do Português*. São Paulo: Ática, 1984.
- MATEUS, Maria Helena Mira. *Aspectos da Fonologia Portuguesa*. Lisboa: Instituto Normal de Investigação Científica, 1982.
- MORENO, C. *Os diminutivos em -inho e -zinho e a delimitação do vocábulo nominal em Português*. Porto Alegre, UFRGS, 1977. (Dissertação de Mestrado)
- NESPOR, Marina. The Phonological Word in Italian. In: *Advances in Nonlinear Phonology*. Harry van der Hulst & Narval Smith, ed. Foris Publications, 1985.
- PEREIRA, Eduardo Carlos. *Grammatica expositiva*. 47. ed. São Paulo: Nacional, 1926. 390 p.
- . *Grammatica historica*. São Paulo: Weiszflog, 1916. 604 p.
- THORNTON, Anna M. On some phenomena of prosodic morphology in Italian: accorciamento, hypocoristics and prosodic delimitation. *Probus*, n. 8, p. 81-112. 1996
- VILLALVA, Alina *Estruturas morfológicas: unidades e hierarquias nas palavras do Português*. Lisboa, 1994. (tese de doutoramento)
- . *Compounding in Portuguese*. Instituto de Linguística Teórica e Computacional. *Working Papers*, n. 2, Apr. 1990.