

## A FORMAÇÃO DE PALAVRAS: COMPONENTE INDEPENDENTE OU APENAS SUBCOMPONENTE? \*

0. A designação portuguesa «formação de palavras» (como aliás as suas correspondências em diferentes línguas: fr. *formation des mots*, ingl. *word formation*, ai. *Wortbildung*) é ambígua, uma vez que indica tanto o processo (ou o sistema de processos) formativo como o resultado desse processo (isto é, o conjunto de lexemas morfológicamente complexos resultantes desse processo) e ainda como o ramo da linguística que estuda as estruturas em que a língua forma novas unidades léxicas (para este último aspecto possui o alemão uma expressão própria: *Wortbildungslehre*).

A formação de palavras, quer como possibilidade sistemática de formação numa língua particular, quer como produto dessa possibilidade, terá de ser enquadrada numa descrição global. Contudo, perante a complexidade do fenómeno linguístico «formação de palavras», complexidade resultante das regularidades semânticas, sintáticas e morfológico-fonológicas e das semi-regularidades ou mesmo das irregularidades totais, este domínio aparece como uma prova-teste para a avaliação (em sentido positivo ou negativo) das teorias linguísticas e gramaticais.

As abordagens ao estudo da formação de palavras resumem-se a dois tópicos que se contrapõem e compõem: *análise* e *síntese*. Aliás, estes dois tópicos representam uma caracterização metodológica não apenas para a formação de palavras, mas também para o estudo da linguagem<sup>1</sup>, e os princípios englobados no ponto de

---

\* Este trabalho é uma remodelação (e versão em português) de uma comunicação preparada para o «XVI Congresso de Linguística e Filologia Românica», realizado em Palma de Maiorca, em 1982.

<sup>1</sup> Cfr. GABEUENTZ, G. v.d. — *Die Sprachwissenschaft*, 2.<sup>a</sup> ed., Leipzig, 1901, pp. 85 e 479, COSERIU, E. — *Semantik, innere Sprachform und Tiefenstruktur*, in «Folia Linguistica», 4, 1971, pp. 53-63, sobretudo pp. 61-2.

vista da *síntese* (ou princípios generativos) foram aplicados à descrição da formação de palavras muito antes de ter surgido a gramática transformacional generativa<sup>2</sup>. Ao opormos os dois pontos de vista da descrição-análise e síntese-não queremos insinuar que esses dois princípios se devam excluir, mas antes se pretende sublinhar que se o ponto de partida privilegiar a análise, o resultado é (ou será) a descrição morfológica e/ou semântica das unidades léxicas morfológicamente complexas numa língua e que se privilegiar o aspecto regular e produtivo do processo de formação (= derivação das estruturas formativas a partir de estruturas mais simples), o resultado é (ou será) uma explicação-descrição generativo-transformacional. Estes dois pontos de vista surgem actualmente com diferentes nomes, como «Wortbildung» e «Wortgebildetheit» (=formação como processo/formação como resultado do processo)<sup>3</sup>, hipótese lexicalista e hipótese transformacionista, hipótese propriamente sintáctica<sup>4</sup> e, finalmente, hipótese lexicológica (integração da formação das palavras na componente «léxico»). Isto é, a discussão à volta da formação de palavras centra-se na distinção feita entre a criatividade linguística como «rule-governed behaviour» enquadrada na competência (linguística do falante/ouvinte) e as propriedades específicas da formação de palavras, a que se acrescenta um terceiro ponto, que é o de saber qual a relação entre a formação de palavras e as componentes da gramática. Este último ponto é particularmente importante, uma vez que, na descrição numa língua, quer se tome como base a sintaxe ou o léxico, a morfologia ou a semântica, encontrar-se-ão sempre fenómenos que têm a ver com a formação de palavras e que, além disso, não podem ser convenientemente explicitados apenas com os instrumentários sintácticos, lexicais, morfológicos ou semânticos. O que deve ser salientado, antes de mais, é a diferença entre o método de análise da língua e a própria língua (como, por exemplo, há uma

<sup>2</sup> Cfr. BREKLE, H. E.; KASTOVSKY, D. — *Wortbildungsforschung: Entwicklung und Positionen*, in BREKLE, H. E.; KASTOVSKY, D., (edd.) — «Perspektiven der Wortbildungsforschung», Bonn, 1977, pp. 7-19.

<sup>3</sup> DOKULIL, M. — *Zum wechselseitigen Verhaeltnis zwischen Wortbildung und Syntax*, in «Travaux linguistiques de Prague», 1, 1964, pp. 215-224 e *Zur Theorie der Wortbildungslehre*, in «Wissenschaftliche Zeitschrift der Karl-Marx-Universitaet Leipzig», Gesellschafts — und Sprachwissenschaftliche Reihe, 17, 1968, pp. 203-211.

<sup>4</sup> MARCHAND, H. — *The Categories and Typès of Present-Day English Word Formation*, 2.<sup>a</sup> ed., Muenchen, C. H. Beck, 1969.

deferença abissal entre poder descrever sintacticamente a formação de palavras e o afirmar-se que as palavras complexas (derivadas) são equivalentes a frases reduzidas).

Um outro ponto que merece ser sublinhado é o que se refere ao dogmatismo de teorias que não leva à discrição da formação de palavras, dos factos linguísticos, mas apenas à descrição das próprias teorias, forçando os factos linguísticos para provar as teorias e não o inverso; desta cegueira dos modelos<sup>5</sup> resultam soluções técnicas e a ignorância consciente ou inconsciente dos factos linguísticos.

## 1. A formação de palavras e a sintaxe

Os sintacticistas <sup>6</sup> afirmam que as formações de palavras são expressões complexas, que apenas se distinguem de outras expressões complexas por determinadas propriedades da estrutura superficial (como, por exemplo, pela configuração morfológica). Deste modo, a descrição sintáctico-semântica das palavras derivadas não necessita de quaisquer meios especiais na teoria gramatical. Em princípio, as formações derivadas serão apenas variantes das estruturas superficiais relativamente às estruturas básicas (sintácticas ou semânticas). A caracterização das estruturas básicas das palavras derivadas, assim como a explicitação das relações entre as estruturas básicas e as estruturas superficiais, são possíveis por meio de determinados mecanismos, mecanismos estes que não são apenas necessários na explicação das palavras derivadas.

É certo que a formação de palavras e a sintaxe têm muito de comum: uma e outra são expressão da criatividade linguística; a

---

<sup>5</sup> É o que W. MOTSCH (*Ein Plaedoyer fuer die Beschreibung von Wortbildungen auf der Grundlage des Lexikons*, in BREKLE, H. E.; KASTOVSKY, D. (edd.), «Perspektiven...», Bonn, 1977, pp. 180-202, 180) designa por «Modellblindheit» (= preocupação excessiva de modelos).

<sup>6</sup> A relação sintaxe-formação de palavras foi já assinalada por W. PORZIG (*Die Leistung der Abstrakta in der Sprache*, in «Blätter fuer Deutsche Philosophie. Zeitschrift der Deutschen Philosophischen Gesellschaft», 4, 1931, pp. 66-77) e por Ch. BAIXY (*Linguistique générale et linguistique française*, 2.<sup>a</sup> ed., Bern, 1944, p. 102) e, de modo mais sistemático, por R. B. LEES (*The Grammar of English Nominalizations*, Bloomington, 1960) e pela chamada hipótese transformacionalista da Gramática Generativa (cujo ponto de partida é: a formação de frases e a formação de palavras procedem dos mesmos princípios).

sintaxe permite ao falante, com meios finitos (um inventário finito de elementos e de regras) produzir um conjunto não finito de combinações e a formação de palavras permite ao falante, com base em elementos existentes, criar designações léxicas correspondentes a novas situações. Além disso, a sintaxe e a formação de palavras utilizam as mesmas categorias: uma frase como *alguém conduz um carro* é analisável em Sujeito-Predicado-Objecto e na relação Agente--Predicado-Objecto (afectado); as mesmas categorias são usadas na descrição do sintagma derivado: *condutor de carros*.

Contudo, tão evidentes como as semelhanças entre a formação de palavras e a sintaxe são as diferenças: a sintaxe não pressupõe um conjunto de combinações memorizadas (ou memorizáveis) numa comunidade linguística, a não ser como modelos abstractos ou como combinações fixas (discurso repetido <sup>7</sup>), enquanto a formação de palavras tem como produto unidades léxicas complexas que fazem parte (em diferentes graus) do conhecimento (activo-passivo) da comunidade linguística, no sentido de que integram o seu vocabulário e pertencem ao léxico disponível como algo «ready-made». Por outro lado, a regularidade, a produtividade, estão presentes em grau e modo diferentes na sintaxe e na formação de palavras: a formação de palavras (deriuatio uoluntaria) é caracterizada por restrições da norma e por condicionamentos idiossincráticos em grau muito maior do que a sintaxe.

Passamos a concretizar mais de perto, com exemplos do português, as semelhanças e diferenças entre sintaxe e formação de palavras. A formação de palavras e as combinações sintácticas são produzidas de acordo com certas regras e estas regras podem ser combinadas e esta combinação está ainda sujeita a regras. Tanto na sintaxe como na formação de palavras podemos combinar unidades produzidas por diferentes regras:

sintaxe: uma cláusula relativa pode combinar-se com uma outra cláusula relativa:

*Os que se manifestam contra o programa que o governo apresentou e não oferecem ao povo uma alternativa viável..., são sempre os mesmos.*

---

<sup>7</sup> Para o conceito de «discurso repetido» (Wiederholte Rede) cfr. COSERIU, E., *Probleme der strukturellen Semantik*, Tuebingen, 1973, pp. 35 e ss.

## A FORMAÇÃO DE PALAVRAS

formação de palavras: de um deverbais substantivo pode formar-se um adjetivo:

*pôr -> posição -> posicionar*

No entanto, se considerarmos as possibilidades de expansão de uma unidade básica, estas diminuem da sintaxe para a formação de palavras, assim como também diminuem da formação de palavras para a flexão:

flexão: *mãe/mães*

formação de palavras:

MATRE { *mae*  
*mãezinha*  
*materno*  
*maternal*  
*madre*  
*madrinha*  
*língua-mãe, terra-mãe*

sintaxe:

*Os utentes do metro — devem primeiro comprar bilhete*  
*— estão atentos à partida*  
*— guardam-se dos ladrões*  
*— lêem os jornais*  
*— compram bombons*  
*ect., etc.*

A recursividade sintáctica é quase indelimitável:

*O homem que encontrei e me falou, que conheço desde há muito e a quem estimo...*

**MÁRIO VILELA** A recursividade não

tem o mesmo poder na derivação:

*morder -> mordiscar -> mordiscadela-> mordiscadelazinha  
-> mordedura -> mordedurazinha -> mordedela ->  
mordedelazinha -> remorder -> remardiscar*

*centro -> central -> centralizar-> centralização ->  
centralizável -> -> descentralizar->  
redescentralizar*

Além disso, há factos na formação de palavras que não encontram explicação nem paralelo na sintaxe. Examinemos os factos seguintes:

- 1) dos verbos factitivos em -tar é possível a formação de substantivos em -ção:

*ornamentar -> ornamentação  
fragmentar -> fragmentação  
sedimentar -> sedimentação  
aumentar -> ? atormentar ->  
? comentar-> ? (comentário)*

- 2) de designações dum instrumento formam-se verbos:

*crivo -> crivar, pincel -> pincelar, escova -> escovar, etc.*

Contudo, se para *vassoura* existe o verbo correspondente *vassourar* (que é usado com certas restrições) não é o usual, pois normalmente usa-se *varrer* (para a acção equivalente); a «rada não corresponde qualquer verbo \**aradar* mas sim *arar* (usado com certas restrições) e sobretudo *lavrar*,

## A FORMAÇÃO DE PALAVRAS

Ora a sintaxe não explica nem as lacunas <sup>8</sup>, nem as preferências <sup>9</sup> permitidas e impostas pela língua. É que uma formação possível, segundo um determinado modelo, pode ser inaceitável na norma da língua: por vezes é inaceitável porque o respectivo significado já está preenchido por uma outra palavra (*arado* ~> \**aradar*, mas sim *lavar*).

As formas *tarde* e *cedo*, formas adverbiais de tempo, e pertencentes a um mesmo domínio semântico, comportam-se de modo diferente relativamente à derivação verbal ou adjectival:

*tarde* -> *tardar*  
          -> *tardio*  
          -> *entardecer*  
*cedo* -> ?  
          -> ?  
          -> ?  
          -> *cedinho*

Para *cedo* há a forma adjectiva *serôdio*.

De verbos designativos de certas expressões ou actividades humanas é possível formar frequentativos: *cantar* -> *cantarolar*, *saltar* -> *saltitar*, *morder* -> *mordiscar*; mas não são possíveis as mesmas formações para *ver*, *o/Aar*, *cheirar*.

As designações dos femininos de animais são feitas através de processos derivacionais (*galo* -> *galinha*, *cão* -> *cadela*) ou lexemáticos (*cavalo-égua*, *carneiro-ovelha*), além dos processos regulares (flexionais). Há uma designação genérica para a designação global de certos números: *novena*, *dezena*, *vintena*, *trintena*, *centena* e porque não há essas designações para outros conjuntos de números?

Existe o modelo de formação adjectiva (com base latina) para nomes de parentesco, faltando, contudo, essa formação em muitos casos:

*pai* -> *paternal*  
*mãe* -> *maternal*  
*filho* -> *filial*  
*filha* -> *filial*

---

<sup>8</sup> Para *aumentar* não há \**aumentação* mas *aumento*, nem para *atormentar* há \**atormentação* ou *atormento*.

<sup>9</sup> A *vassoura* corresponde *víz/rér* e a *arado* corresponde, como verbo factitivo, *lavar*.

mas para:

*tio* -> ?  
*tia*  
*avo* -> ?  
*avó*  
*primo* -> ?  
*pnrna*

*sobrinho*  
-> ?  
*sobrinha*

Se para *avô/avó* existem as formações *avoengo* e *ancestral*, estas representam uma especialização de sentidos (variantes de significado) do termo base.

Concluindo: as regularidades na formação de palavras são importantes, mas as irregularidades não o são menos e também são fenómenos da língua. E a sintaxe não está sujeita a essas irregularidades, nem pode dar conta dessas mesmas irregularidades.

## 2. A formação de palavras e a semântica

A formação de palavras tem pontos de contacto com a semântica (semântica lexical), uma vez que as palavras formadas fazem (ou passam a fazer) parte do léxico, independentemente da sua formação ou produção: essas formações estão sujeitas à lexicalização, à idiomatização, etc. Além disso, as formações léxicas, regressadas ao léxico, participam das relações gerais próprias do léxico: relações paradigmáticas (estruturação em campos lexicais e classemáticos) e sintagmáticas (solidariedades lexicais). Contudo, um tal tratamento (apenas lexemático ou semântico) corre o risco de esquecer que a formação de palavras e a sintaxe trabalham com as mesmas categorias (Agente/Sujeito, Objecto afectado ou efectuado/Objecto, Locativo/Determinação adverbial de lugar, etc), a menos que se introduza nessa análise a gramaticalização como factor determinante.

A proposta da semântica generativa dissolve o problema da formação de palavras, considerando todo o léxico, quer o primário, quer o secundário (palavras derivadas), como dirigido pelos mesmos

princípios, aliás os mesmos das frases totais, integrando assim todo o léxico na formação de palavras (e na formação de frases).

Contudo, há certos factos na formação de palavras que escapam a qualquer previsão ou regularidade semântica: existe, por exemplo, em português, um modelo de derivação Substantivo + *-esco*, em que o sufixo insiste numa determinada variante significativa do lexema base, normalmente, superlativando pejorativamente essa variante: *holyoodesco*, *folhetinesco*, *magazinesco*, *livresco*. Essa mesma derivação parece ser possível com nomes de pessoas a acentuar determinado sema (ainda no sentido pejorativo) do nome base: *quixotesco*, *cidesco*, etc. Contudo, o português, neste tipo de derivação, é muito restritivo. Pergunto: o que seria necessário para haver derivações como: *eanesco* (*governo eanesco*), *cunhalesco* (*processo cunhalesco*), *pintassilguesco* (*alternativa pintassilguesca*), *gonçalvesco* (*período gonçalvesco*), *salazaresco* (*partido salazaresco*)! O português actual usa nestes casos, o afixo mais neutro *-ista* (relativamente a Salazar já se tem usado *salazarento*). Como explica a semântica ou a sintaxe tais opções e restrições? Isto é, poderemos afirmar, com a mesma precisão, que «é uma formação léxica possível», como «é, uma frase possível ou uma construção semântica possível», numa determinada língua, sem se ter em conta os factos extralinguísticos?

### 3. A formação de palavras e a morfologia

A relação formação de palavras-morfologia torna-se evidente, uma vez que as novas palavras são formadas com base em afixos, que estão sujeitos aos condicionamentos morfológico-categoriais das palavras-base e dos próprios afixos. A partir deste pressuposto tem-se feito uma descrição da formação com bases puramente formais, como se pode ver pelas distinções prefixação, infixação, sufixação e justaposição. Contudo, este modelo de descrição ignora a inexistência de paralelismo entre os processos formais (prefixação, sufixação, infixação e justaposição) e os tipos de funções presentes nas palavras formadas: diferentes processos materiais podem ter uma mesma função e vice-versa (como, por exemplo, a quantificação pode ser expressa por prefixação (*retomar*) e por sufixação (*saltitar*) ou as derivações *leitor*, *casinha*, *tardio* são funcionalmente heterogéneas).

Efectivamente, parece haver uma proximidade muito grande entre flexão e derivação: tanto uma como outra são, normalmente, de natureza afixai. Contudo, a flexão consiste na constituição de formas dum lexema, enquanto a derivação constitui um novo lexema: formas de lexema, como formas de tempo, modo, aspecto, número e pessoa (nos verbos), género e número (substantivos e adjectivos). Isto é, *trabalhei* é uma forma do verbo *trabalhar*, *casas* é uma forma do substantivo *casa*, *belas*, uma forma do adjectivo *belo*; mas *trabalhador* não é verbo como *trabalhar*, nem substantivo apenas como *trabalho*. Por outras palavras, o afixo *-dor* não é uma forma flexionai de *trabalhar* ou *trabalho*, mas constitui uma nova palavra: os morfemas (ou monemas) tais como *-dor* são morfemas derivacionais ou lexicais, distinguindo-se dos morfemas flexionais pela sua função própria.

Além disso, em comparação com a fixidez de posição da flexão, é maior a possibilidade de combinação na derivação:

*des-militar-iza-ção*

A flexão é previsível, opera dentro de regras, mesmo nos casos irregulares há uma regra; não acontece o mesmo na derivação: não podemos prever todas as palavras possíveis. É certo que os transformacionistas falam de formulação de regras, de esquemas de formação, etc, de lexicalização, de idiomatismos, de irregularidades: não seria melhor falar de especificidade na formação de palavras considerando este domínio como uma componente específica da gramática? Tomemos a série seguinte:

*correr -> corredor ->*  
*corrida ->*  
*correria*

Será possível prever as mesmas derivações para um verbo pertencente ao mesmo domínio semântico como *fugir*? As regras apresentadas pela gramática generativa, neste domínio, são, a coberto da formalização, mal definidas e uma boa definição da regra é condição indispensável para um bom entendimento da distinção «rule-governed creativity» e «rule-changing creativity» e da explicação do aspecto criativo da formação de palavras novas e da sua posterior

## A FORMAÇÃO DE PALAVRAS

hierarquização. Será previsível, por exemplo, uma regra como:  
Adj -> Vb -> Sbst -> Adj?

*moderno -> modernizar-> modernização -> modernizável*  
*-> \*modernizacional*  
*belo -> embelezar -> embelezamento -> \*embelezamento!*

Concluindo: são múltiplos os aspectos incluídos na formação de palavras: morfológicos, sintáticos, semânticos, léxicos. A estes deve ainda acrescentar-se a necessidade de integração de elementos extralinguísticos, se se quiser interpretar de modo conveniente as formações léxicas. Por isso mesmo, o tema «formação de palavras» pode ser considerado como prova-teste das teorias linguísticas.

### 4. A formação de palavras na gramática transformacional generativa

Veremos sem grandes pormenores o modo como a gramática generativa resolve o problema da formação de palavras. A gramática generativa apresentou até hoje três modelos essenciais de descrição:

M1: a base é formada pela descrição sintáctica da estrutura profunda;

M2: a base é formada pelas informações no léxico; M3: a base é formada pela descrição da estrutura semântica.

O modelo preferido e mais desenvolvido é o M1, o da chamada teoria Standard<sup>10</sup>, cuja forma é a seguinte:



<sup>10</sup> CHOMSKY, N.—*Deep Structure, Surface Structure, and Semantic Interpretation*, in STEINBERG, D. D.; JACOBVITS, L. A. (edd.) — «Semantics. An Interdisciplinary Reader in Philosophy, Linguistics and Psychology», Cambridge, Univ. Press, 1971, pp. 183-216.

A componente sintáctica, que forma o núcleo da descrição, é constituída pelas regras de transformação que especificam uma classe de estruturas sintácticas profundas e estas reflectem-se numa classe de estruturas sintácticas superficiais. Às estruturas profundas, que produzem as formações de palavras por meio das regras de transformação, contêm todas as informações na representação sintáctica, informações que são fundamentais para a interpretação semântica. Uma vez que as estruturas superficiais das formações léxicas são de informação muito diminuta para a estrutura do significado, a descoberta das estruturas profundas de determinados tipos de formações léxicas constitui um manancial de problemas teórico-empíricos. A interpretação semântica, neste modelo, depende apenas das estruturas sintácticas profundas das formações léxicas e não necessita das informações das estruturas superficiais das mesmas formações. Contudo, as formações léxicas não contêm muitas vezes as informações que têm de ser representadas na componente semântica do léxico. À componente morfológica pertence a descrição dos pormenores morfémicos das formações léxicas, tais como as regras morfológicas, afixos de derivações, etc.

Uma tal proposta parece ser adequada para uma maior transparência do significado dos elementos de certos tipos de formações léxicas derivadas a partir da estrutura profunda (por exemplo, no composto *abre-latas* (:a forma *abre* em *abre-latas*, contém o morfema de natureza pronominal, aqui representado por um morfema zero) <- «instrumento (*algo*) que (serve para abrir) *abre latas*).

Contudo, em muitas destas explicações, é necessário o recurso a perífrases, que, muitas vezes, não passam de truques técnicos para explicitar certas relações semânticas: por exemplo, *colher de pau* ('colher feita de pau?', 'colher para guardar pau?', etc), *chapéu de chuva* ('chapéu feito de chuva?', 'chapéu para defender da chuva?'), etc. Além disso, expressões lexicalizadas como *cara de pau*, *pé de cabra*, *pele de galinha*, encontram-se no léxico lado a lado com expressões que são produtos normais da formação de palavras. Há ainda o problema da produtividade de um modelo, que ou se confunde com a regularidade, ou (entendida como probabilidade de ocorrência em novas formações) terá de ser considerado como problema do domínio da «performance»; assim como o problema da aceitabilidade de uma formação se situa na «performance».

Perante estas dificuldades, sobretudo as que surgem a partir de formações lexicalizadas, Chomsky desenvolve a «hipótese lexicalista»<sup>11</sup>, que designámos por M2, para o domínio das nominalizações. Jackendoff<sup>12</sup> ampliou a hipótese lexicalista, aplicando-a a todos os compostos nominais. Segundo este modelo a informação necessária à caracterização duma formação léxica é representada no léxico. Todos os aspectos comuns (previsíveis) e regulares das formações léxicas são expressos pelas regras de redundância.

Mas neste modelo ficam por explicar os processos característicos que produzem as formações léxicas, nem mesmo se explica com clareza em que consistem as regras de redundância. Também é verdade que Chomsky não se preocupa muito com as formações léxicas, mas antes com as consequências do modelo lexicalista para a sintaxe, que quer a toda a força que seja a componente básica.

O modelo que designámos por M3 trata muito indirectamente o problema da formação de palavras, e sempre em conexão com os problemas da estrutura semântica, como acontece na análise de *Kill(tó)* como CAUSE TO BECOME NOT ALIVE<sup>13</sup>. Propostas mais concretas feitas dentro desta perspectiva<sup>14</sup>, não caracterizam também o que de particular e específico há na formação de palavras. Se o modelo Standard e a hipótese lexicalista distinguiam as formações léxicas de outras formações complexas apenas por pormenores morfológicos, também a semântica generativa não distingue entre a estrutura subjacente e o produto dessa estrutura.

---

<sup>11</sup> CHOMSKY, N.—*Remarks on Nominalizations*, in JACOBS, R. A.; ROSENBAUM, P. S. (edd.) — «Readings in English transformational Grammar». Mass., Waltham, 1970, pp. 184-221.

<sup>12</sup> JACKENBOFF, R. C.—*Morphological and semantic regularities in the lexicon*, in «Language», 51, 1975, pp. 639-671.

<sup>13</sup> MCCAWLEY, J. D. — *Lexical insertion in a transformational grammar without deep structure*, in «Papers from the 4th regional meeting. Chicago Linguistic Society». Chicago, 1968, pp. 71-80.

<sup>14</sup> Como, por exemplo, a que foi feita por BREKLE, H. E. — *Generative Satzsemantik im System der Englischen Nominalkomposition*, 2.<sup>a</sup> ed., Muenchen, W. Fink, 1976.

## 5. A formação de palavras e a lexicologia

5.1. O carácter especial do léxico é posto em evidência por M. Halle<sup>15</sup>, R. Jackendoff<sup>16</sup>, M. H. Aronoff<sup>17</sup>; W. Motsch<sup>18</sup> considera mesmo a formação de palavras como resultado da força do próprio léxico e da necessidade de representação da realidade extralinguística. O aspecto criativo da língua (língua em uso, ou uso da língua) é comumente aceite pelas diferentes teorias linguísticas. Esse aspecto criativo tem sido colocado apenas na sintaxe ou na semântica (frásica). Contudo, exprimir situações novas exige muitas vezes designações linguísticas novas: objectos novos, factos novos, situações novas, são expressos lexicalmente por meio de expressões léxicas e não apenas por meio de combinações sintácticas. E isto é muito mais verdade se admitirmos (como muitos linguistas admitem) que a representação da realidade extralinguística se faz primariamente no léxico duma língua particular. Veja-se o que acontece actualmente com palavras como *eurocomunismo*, *terceiro-mundismo*, *eurosocialismo*, etc. Perante esta evidência resta-nos saber (ou procurar saber) se estas novas formações terão de ser tratadas por meios sintácticos (na sintaxe), por meios semânticos (na semântica) ou por qualquer outro processo duma outra natureza. Há que ter em conta a força da própria língua para formar palavras novas («Neue Wörter entstehen auf der Grundlage von im Lexikon enthaltenen Wörtern und Informationen»)<sup>19</sup>. As construções sintácticas originam-se de acordo com a fórmula:

unidades elementares + regras -> expressões complexas.

As formações léxicas obedecem à fórmula:

palavras complexas estruturadas -> novas palavras complexas.

<sup>15</sup> HALLE, M. — *Prolegomena to a theory of word formation*, in «Linguistic Inquiry», 4, 1973, pp. 3-16.

<sup>16</sup> JACKENDOFF, R. — *Morphological and semantic regularities in the lexicon*, in «Language», 51, 1975, pp. 639-671.

<sup>17</sup> ARONOFF, N. H. — *Word Formation in Generativ Grammar*, Cambridge, Mass., The MIT Press, 1975.

is MOTSCH, W. — *Ein Plaedoyer fuer die Beschreibung von Wortbildungen auf der Grundlage des Lexikons*, in BREKLE, H. E.; KASTOVSKY, D. (edd.) — «Perspektiven der Wortbildungsforschung», Bonn, 1977, pp. 180-202.

<sup>18</sup> MOSTCH, W. — *Plaedoyer...*, Ibid., p. 193.

Admitir esta proposta será talvez privilegiar a analogia, o que já foi feito pelos neogramáticos, mas leva também à conclusão de que a formação de palavras terá de ser tratada no léxico. E, nesta proposta, caberá a uma descrição da formação de palavras, a fazer no próprio léxico, esclarecer os seguintes pontos:

- a caracterização do significado duma formação, caracterização que deverá constar de duas partes: *a)* a indicação da informação previsível ligada a um determinado tipo de formação); *b)* a indicação da informação não previsível, que é específica da unidade léxica correspondente, como acontece na formação de *livresco* (*educação livresca*) no que tem de especialização em relação ao sintagma subjacente (atente-se nos exemplos seguintes: *comprador de livros* e *comprador livresco*), ou em formações como *escrevinhar*, *cuspinhar*, etc.
- a descrição das propriedades sintáticas duma formação ou o seu aspecto gramatical.
- a descrição das propriedades morfológico-fonológicas das formações.

Haverá um processo ou uma teoria capazes de responder a todos estes quesitos? E se houver um tal processo, qual a relação da formação de palavras com a gramática (supondo que a formação não é um domínio isolado)?

## 5.2. A formação de palavras no ponto de vista funcional

Procurar saber qual o lugar da formação de palavras na linguística, teremos primeiro de procurar saber qual a função deste domínio linguístico ou dessa formação. A função da formação léxica será a de enriquecer o vocabulário pela mudança categorial e/ou semântica das palavras existentes. A formação como disciplina descritiva estuda essas mudanças na medida em que forem estruturas <sup>20</sup>.

---

<sup>20</sup> A este propósito afirma G. STEIN (*The place of word-formation in linguistic description*, in BREKLE, H. E.; KASTOVSKY, D. (edd.), «Perpektiven der Wortbildungsforschung», Bonn, 1977, pp. 219-235, 235): «Word-formation is different from inflection as well as syntax and semantics and that is based on the function of word-formation in human language which consists in changing

É nesta perspectivação que se move E. Coseriu <sup>21</sup> ao integrar a formação de palavras na chamada lexemática. Os pontos básicos da formação de palavras são o conteúdo e a gramaticalização. Ou, mais propriamente, ao tomar o ponto de vista do conteúdo como base, faz corresponder a nova formação a uma gramaticalização do léxico primário e faz ainda corresponder os tipos de formação aos modos e condições desta gramaticalização. A relação gramatical (=relação análoga à relação gramatical) dos produtos de formação resulta das equivalências de significado entre estes produtos e as construções que significativamente lhes correspondem: por exemplo, *beleza* = 'o facto de ser belo/a', tendo esta forma analítica apenas uma função metalinguística, sendo 'facto', na referida fórmula, não a palavra primária *facto*, que é determinada por 'ser belo', mas apenas um nome para a indicação da substantivação; assim como 'ser' é o nome de predicatividade (predicação atributiva) e 'belo/a' o nome da unidade na língua primária (língua-objecto), isto é, um *belo* sem género e número. Por outras palavras, a fórmula significa: *belo(s)-bela(s)* + gramaticalização por meio da predicação atributiva + substantivação. As relações entre as respectivas bases de formação e os seus produtos resultam da comparação significativa entre estas bases e os seus equivalentes nos produtos formados: assim, *beleza* contém relativamente a *belo(s)-bela(s)* as determinações adicionais 'predicatividade' e 'substantividade'. Os produtos de formação não são, no seu significado, equivalentes às suas bases e não são produzidos por meio duma

---

the word class and/or the meaning of existing words in language. Word-formation itself makes up a part of lexicology which studies the structures of the lexicon of a language» e quase o mesmo é também proposto por W. KUERSCHNER (*Generative Transformationsgrammatik und die Wortbildungstheorie von H. Marchand*, in BREKLE, H. E.; KASTOVSKY, D. — «Perspektiven...», pp. 129-39, 137): «Vielleicht müssen die traditionellen Gebiete der Wortbildung auf verschiedene grammatische Komponenten verteilt werden die Transposition z.B. in die Syntax, die semantischen Derivationen in die eigentliche Wortbildungslehre, die dann einen Sonderstatus hätte».

<sup>21</sup> Sirvo-me sobretudo de: COSERIU, E. — *Les structures lexématiques*, in «Zeitschrift fuer französische Sprache und Literatur», Beiheft 1 (Neue Folge), 1968, pp. 3-16; Id. — *Bedeutung und Bezeichnung im Lichte der strukturellen Semantik*, in HARTMANN, P.; VERNAY, H. (edd.) — «Sprachwissenschaft und Uebersetzen», Muenchen, 1970, pp. 104-121; Id. — *Die funktionelle Betrachtung des Wortchatzes*, in «Sprache der Gegenwart-IdS», 39, 1976, pp. 7-25; Id. — *Inhaltliche Wortbildungslehre (am Beispiel des Typs «coupe-papier»)*, in BREKLE, H. E.; KASTOVSKY, D. (edd.) — «Perspektiven...», Bonn, 1977, pp. 48-61.

transformação: os produtos de formação contêm sempre mais do que as suas bases lexicais correspondentes.

Tendo-se como base de descrição o conteúdo e a função respectiva, são dois os dados que constituem o critério de determinação dos tipos de formação de palavras: a gramaticalização e o número de elementos presentes na base. Isto é, se a gramaticalização implícita relaciona dois elementos na base: temos a *composição*; se a gramaticalização opera apenas com um elemento de base: temos a *modificação* e o *desenvolvimento*. Por sua vez, a modificação e o desenvolvimento distinguem-se pelo modo como opera a gramaticalização: na modificação a gramaticalização corresponde a uma função inactual («não análoga à frase»), no desenvolvimento a gramaticalização corresponde a uma função actual («análoga à frase»). Isto é, na modificação e no desenvolvimento trata-se da gramaticalização de um único elemento na base, na composição temos na base dois elementos em relação gramatical entre si. Portanto, na modificação há uma função gramatical inactual (do tipo de género, número), no desenvolvimento há uma função gramatical actual (do tipo das funções sujeito, predicado, complemento do verbo), na composição pode haver uma função actual ou inactual.

5.2.1. Assim, *senhora* —> *senhorita* —> *senhorinha*, *árvore* -> *arvoredo*, *vermelho* -> *vermelhusco*, *cantar* -> *cantarolar*, *beijar* -> *beijocar* são modificações; na modificação a categoria gramatical é a mesma na base e na forma modificada, uma vez que a função gramatical não se altera (na base e na forma modificada), e em que a função gramatical, aqui incluída, é 'inactual', isto é, trata-se duma função que apenas diz respeito ao lexema como tal e não a este lexema como elemento da frase ou como elemento de sintagmas.

Em *esperto* -> *esperteza*, *chegar* ~> *chegada*, *barco* -> *embarcar*, *rico* —> *enriquecer* -> *enriquecimento*, temos desenvolvimentos. No desenvolvimento, parte-se sempre dos lexemas como elementos da frase ou como elementos do sintagma e a categoria gramatical da formação desenvolvida é diferente da da base: adj. -> subst, vb-> -> subst, sbst. ~> vb., adj. -> vb. -> sbst.

A composição pode ser prolexemática ou lexemática: prolexemática se um dos elementos da base for «prolexema» (elemento de natureza pronominal), lexemática se ambos os elementos da base forem lexemas. Por exemplo, «elemento substantivo geral de natureza pronominal» (= algo como: «alguém» ou «algo») + *ler* -> *leitor*,

«alguém ou algo» + *despertar* -> *despertador*, «alguém ou algo» (que vende) + *livro* -> *livreiro*, etc; *boquiaberto*, *mata-frades*, etc, fornecem-nos os diferentes elementos possíveis da composição.

O elemento importante na formação de palavras é a gramaticalização: é a gramaticalização que faz com que a formação de palavras seja formação de palavras e os tipos de formação correspondem essencialmente a modos da gramaticalização. Isto quer dizer que a gramaticalização não está apenas presente em algumas formações, mas em todos os processos formativos. Contudo, deve observar-se que esta gramaticalização diz respeito a uma gramática do léxico, isto é, trata-se duma «gramática» que não pode confundir-se com a gramática sem adjetivos (com a gramática da morfossintaxe). Por isso dever-se-á falar, não de funções gramaticais, mas sim de funções análogas às da gramática. Assim, por exemplo, a formação dos colectivos é uma pluralização, mas não estamos em presença de um simples plural, mas sim em presença de 'uma pluralidade unitária e considerada como pluralidade' (*arvoredos* não é igual a *árvore*), ou *boquiaberto* não é igual a *boca aberta*; *chegada* implica um verbo predicativo, mas sem modo, tempo, número e pessoa; *riqueza* implica um nome predicativo, mas sem diferença de género e de número. Estas formações não reduzem frases concretas como *Pedro chegou/chegara/chegaria*, *que chegasse*, *João é rico* /*Joana é rica*, mas referem-se a uma função predicativa geral de *chegar*, *rico(s)-rica(s)*; ou as formações *toneladas-dia*, *barris-hora*, embora contenham uma função preposicional, não contêm, contudo, qualquer preposição como tal; ou o caso de *acendedor* que contém um elemento pronominal «alguém ou algo», elemento que não é realizado em português. Por isso, a formação de palavras exprime funções gramaticais mais gerais e diferentes daquelas que são usuais nas respectivas línguas.

Por todas estas razões afirma E. Coseriu (*Inhaltliche Wortbildungslehre...*, p. 56): «..., die Wortbildung kann nicht «entweder zur Syntax oder zum Lexikon» gehören. Die Wortbildung ist ein autonomes Gebiet der Sprache, das «Grammatikähnliches» und rein Lexikalisches einschliesst, ein hierarchisch geordnetes Kontinuum von den Haupttypen der wortbildenden Verfahren bis zu den vereinzelt fixierten; und die Wortbildungslehre ist ein autonomer Zweig der funktionellen Semantik, der bei den «grammatikähnlichen» Funktionen der Wortbildungsverfahren beginnen und bis zu den Fixierungen in der Bezeichnung gelangen muss» (=A formação de palavras não pode fazer parte da sintaxe ou do léxico. A formação de palavras é um

## A FORMAÇÃO DE PALAVRAS

domínio autónomo da língua, domínio que integra algo análogo à gramática e puramente lexical, é um contínuo hierarquicamente ordenado de tipos principais de processos formativos que vão até às fixações individuais; e o estudo da formação de palavras é um ramo autónomo da semântica funcional, ramo que terá de começar com as funções análogas às funções gramaticais e terminar nas expressões fixas de designação).

5.2.2. Os tipos principais da formação de palavras subdividem-se de acordo com os modos de existência/presença das funções gramaticais implícitas. Assim, na *modificação*, podemos ter:

mudança de género: rei -> rainha galo -  
» galinha conde ->  
condessa

quantificação:

diminuição: casa -> casinha

-> casota

chorar -> choramingar diminuição com

apreciação aproximativa: velho -> velhote

grande -> grandinho vermelho -> vermelhusco

aumentação: velhaco -> velhacão casa ->

casarão grande -> grandalhão colectivo:

árvore -> arvoredado criado -> criadagem

laranja -> laranjal intensificação: feijão ->

feijoada garfo -> garfada repetição: fazer ->

refazer negação: legal -> ilegal

onerar ~> exonerar fazer -> desfazer

simetria -> dissemetria

parcialização: ver -> entrever ->

antever -> prever

No *desenvolvimento*, entre outras coisas, podemos distinguir as seguintes subdivisões:

a) com base nas funções sintácticas incluídas na base:

desenvolvimento predicativo: chegar -> chegada  
rico -> riqueza

desenvolvimento atributivo: de inverno -> invernical  
dos trópicos -> tropical

desenvolvimento de complementos preposicionais:  
em barco -> embarcar de  
barco -> desembarcar  
em rico -> enriquecer

b) com base em outras determinações gramaticais da base contidas nas formações desenvolvidas:

*vitória* contém a diatese activa de *vencer derrota*  
contém a diatese passiva de *derrotar obstinação*  
contém a reflexividade de *obstinar-se ferida* contém  
o aspecto resultativo de *ferir gabarola* contém o  
aspecto frequentativo de *gabar comestível* a  
eventualidade (passiva) de *comer martelada, paulada*  
contém o aspecto semelfactivo de *martelo/pau*

c) a topicalização dum determinada relação sintáctica da base:

do objecto: peixe -> pescar do

lugar: habitar -> habitação

Uma outra propriedade dos processos de formação de palavras é a de que se podem combinar entre si os diferentes processos:

*monstro* -> *monstruoso* -> *monstruosidade*

(modific.) (desenvolvim.) *jeito* -> *jeitinho*

(modif.) -> *jeitão* (modif.) -> *jeitoso* (desenv.) ->

*jeitosinho* (modif.)

A combinação dos processos formativos é particularmente evidente em casos de composição como *corta-papel*, *saca-rolhas*, *guarda-livros*, *espalha-brasas*, *parabrisas*, *bate-chapas*, *conta-gotas*. Como E. Coseriu demonstrou (*Inhaltliche Wortbildungslehre...*), o primeiro elemento: *corta-*, *saca-*, *guarda-*, *espalha-*, *bate-*, *conta* — é um elemento substantivo que apenas materialmente coincide com a forma verbal do imperativo, pois trata-se da supressão do afixo *-dor* (cortador de, guardador de, batedor de, etc.), como pode ver-se por casos paralelos em que a língua não admite essa redução: *tratador de cavalos* (e não *trata-cavalos*), *contador da luz* (e não *conta-luz*), ou ainda pelos casos em que a língua distingue duas variantes de sentido, uma pela forma reduzida e outra pela forma plena: *guarda-livros*, *guardador de rebanhos*, *guarda-costas*, *guardador da vinha*, ou casos em que as duas formações são possíveis: *guarda* (ou) *guardador de esquinas*.

Isto é, trata-se nos casos de composição referidos (do tipo: *corta-papel*, *quebra-nozes*, etc.) da combinação da composição prolexemática e composição lexemática: *corta* = 'algo que corta' (normalmente realizado pelo afixo — *dor*: *cortador de carnes*), *quebra* (em *quebra-costas*) = 'algo que quebra', *conta* (em *conta gotas*) = 'algo que conta' — composição prolexemática. Há também composição lexemática pela presença de dois elementos na base: *quebra-costas*, *quebra-cabeça*, etc.

## 6. Conclusão

Após a enumeração mais ou menos crítica das diferentes propostas de tratamento de «formação de palavras», parece-nos evidente que se trata de um fenómeno linguístico complexo, e para o qual terá de se ter em conta o que é específico da mesma formação de palavras. A integração deste domínio da língua na gramática não se fará sem problemas, ou então haverá que subdividir a «formação» em subdomínios e inseri-los nas diferentes componentes da gramática. Talvez seja mais coerente e correcto considerar como elemento dinamizador da formação de palavras a «gramaticalização», entendida como análogo ao que se encontra na gramática (sem adjectivos) ou, melhor dito:

*MÁRIO VILELA*

como uma «gramática» do léxico. Além disso, a formação de palavras é condicionada pela realidade extralinguística e determinada pela «norma» da língua: o que equivale a dizer que a palavra «formada» não corresponde (geralmente), no seu conteúdo, ao significado dos seus elementos no seu conjunto. Talvez seja cientificamente mais correcto considerar a «formação de palavras» como algo independente de qualquer outra componente da gramática, ou como domínio pertencente ao próprio léxico.

*Mário Vilela*