

Maria Tereza Camargo Biderman
UNESP

Unidades complexas do léxico

1. Introdução

O léxico de uma língua inclui unidades muito heterogêneas – desde monossílabos e vocábulos simples até seqüências complexas formadas de vários vocábulos e mesmo frases inteiras como é o caso de muitas expressões idiomáticas e provérbios. Por outro lado, não existem critérios teóricos abrangentes e bem estabelecidos para o reconhecimento das unidades complexas de um idioma. Aliás, o fenômeno da lexicalização de combinatórias lexicais (sintagmas discursivos) não se verifica de modo uniforme e reiterado e também logicamente estruturável. Acresce ainda que os falantes muitas vezes discordam sobre o grau de cristalização de tais seqüências. Assim, as fronteiras de demarcação do que já está estocado no tesouro lexical da língua e o que é combinatória discursiva são fluidas. E por isso mesmo,

Convém assinalar que o problema das combinatórias a incluir num dicionário, seja ele um dicionário de língua monolíngue ou plurilíngue, é dos mais difíceis problemas da lexicografia tradicional, precisamente porque a intuição dos falantes não é suficiente para determinar o grau de fixidez das combinatórias, nem a extensão, estabilidade e vitalidade do seu uso. (Nascimento, 1998,185).

A *identificação da unidade léxica* num texto discursivo exige uma fundamentação teórica consistente. Um dos maiores defeitos dos dicionários tradicionais é o fato de não se fundamentarem em uma teoria lexical, sobretudo em um conceito bem estabelecido de unidade lexical, fato esse que pode explicar o porquê de unidades complexas do léxico, especialmente as EIs, estarem assistemática e inadequadamente registradas nos dicionários da Língua Portuguesa.

2. Exame da problemática teórica das unidades complexas do léxico

Em seu estudo sobre “Phrases figées du français” Gross (1982) nota que os lingüistas geralmente atribuíram às expressões cristalizadas (*expressions figées*) um caráter de exceção, de anomalia lingüística e não tentavam propor maneiras de tratamento científico para elas. Contudo, trata-se de um fenômeno de envergadura que a Lingüística ignora pela ausência de estudos sobre essa matéria. A primeira questão que se põe é terminológica: devemos chamá-las de frases, formas ou expressões congeladas (cristalizadas) como Gross? Além disso, existem ainda as expressões chamadas comu-

mente de idiomáticas e proverbiais. E também as locuções de natureza gramatical conforme a tradição das gramáticas. Em sua análise Gross ocupou-se também de metáforas e clichês.

Um problema formal levantado por Gross e muito importante, a meu ver, é o ortográfico. Acha – e esta também a minha opinião – que os substantivos compostos deveriam ser grafados mais apropriadamente com hífen como por ex.: *cul-de-sac*, *pomme-de-terre*. Para nós em português: *escola-de-samba*, *fim-de-semana*, *jogo-do-bicho*.

Começemos por considerar o caso de verbos compostos ocorrentes em muitas expressões idiomáticas. Vejamos um exemplo emprestado de Xatara (1994) que fez um paralelismo com o exemplo de Gross, a saber:

bater as botas

A sintaxe desta seqüência cristalizada não difere da sintaxe de uma seqüência livre. Cf. as variações de tempo e modo possíveis:

Carlos bateu as botas. [= Carlos morreu]

Todos sabiam que Carlos ia bater as botas.

Temo que Carlos vá bater as botas.

Geralmente as variações de tempo/modo são causadas por condições impostas pela concordância.

O sentido da seqüência *bater as botas* não é previsível a partir de *bater* [= dar pancadas; chocar-se com] e de *botas* [= tipo de calçado].

De fato, temos aqui uma combinatória cristalizada da herança cultural registrada na memória coletiva com o significado de < morrer >. E por isso, podemos afirmar que ela faz parte do acervo do léxico e não se trata de uma combinatória discursiva qualquer. Entretanto, o verbo que ela comporta não é analisável segundo as regras que se aplicam a frases superficialmente idênticas. Modificações possíveis e aceitas em frases quaisquer são impossíveis no caso da expressão cristalizada.

As variações possíveis, além daquelas acima referidas, seriam relativas à posição dos advérbios. Cf. *Carlos bateu as botas muito cedo*. *Carlos bateu muito cedo as botas*. *E seu irmão também bateu as botas*.

É possível também fazer uma inserção: *Carlos, desta vez, bateu as botas*.

As seqüências livres são aquelas em que sujeito e complemento têm distribuição livre, sendo as únicas restrições e coerções as determinadas pela semântica. Inversamente, os sintagmas cristalizados são frases do mesmo tipo, porém, em que um ou vários dos actantes são lexicalmente invariáveis. Gross (1982) constatou que parece haver um continuum entre frases livres e cristalizadas. No artigo referido ele menciona 8.000 frases ou expressões cristalizadas em francês nas quais verificou o seguinte:

- 1) havia menos de 600 sujeitos fixos
- 2) havia mais de 1.300 objetos diretos livres
- 3) havia mais de 1.700 objetos indiretos livres
- 4) havia mais de 1.000 complementos de nomes fixos

Por conseguinte, objetos e complementos parecem ser mais cristalizados que os sujeitos.

Diversamente dos provérbios que exibem um caráter de generalidade, as expressões cristalizadas/ idiomáticas só se aplicam a uma situação particular.

Como Gross, considero como frases ou seqüências cristalizadas séries de formas sem verbo. Ver exemplos do francês: *au feu, à l'abordage, à vos marques, en voiture, à vos souhaits, à la santé de Max*, etc. Em português: *de longe, de longe em longe, longe disso, a olho nu, de olho (bem) aberto, de olho em*, etc.

Formas há que desencadeiam a intuição de frases; tais seqüências combinatórias são muito numerosas e não podem ser negligenciadas como formas excepcionais. “Constatamos, pois, que as formas cristalizadas que são tradicionalmente consideradas como exceção, ocupam no léxico um volume comparável àquele das formas livres correspondentes. “(Gross, 1982,181).

De fato, tais seqüências são extremamente numerosas no português como no francês. Os verbos muito freqüentes no português devem sua polissemia, em parte, à constituição de variadas matrizes verbais distintas que se lexicalizaram. Cf. por exemplo:

andar a reboque, (não) *andar* bem das pernas; *dar* baixa, *dar* cabo, *dar* um jeito (jeitinho); *encher* a barriga (o bucho, o estômago), *encher* de orgulho, *encher* o peito, *encher* os bolsos, *encher*-se de brios; *entrar* na rotina, *entrar* no ar, *entrar* no páreo; *fazer* a cabeça, *fazer* água, *fazer* amor, *fazer* as pazes, *fazer* as honras (da casa), *fazer careta*, *fazer* carnaval, *fazer* charme, *fazer* das suas, *passar* a grana, *passar* a limpo, *passar* em branco, *passar* em revista, *passar* fome (sede, frio), *passar* para (prá) trás, etc.

Um estudo minucioso e sistemático dessas combinatórias verbais poderia sanar o empirismo que geralmente se manifesta nas descrições dos verbos encontradas nos dicionários, associando-se assim as diferenças de sentido às variações na forma.

Todavia não é fácil separar os *sintagmas cristalizados* de formas *intuitivamente freqüentes* ou em *vias de lexicalização*.

Na análise destes sintagmas cristalizados ou em vias de lexicalização constatamos, entre outras coisas, que a ausência do artigo pode indicar que se trata de uma abstração e que a palavra assume, nesse caso, um valor metafórico ou geral. É o caso de um bom número de lexicalizações. Quando se reintroduz o artigo, atribuindo à locução seu valor concreto, a expressão se *desagrega*. Opor 1) e 2) e 3) e 4).

- 1) De tanto *levar e trazer recado* o moleque já nem ligava mais.
- 2) O moleque *levava e trazia um recado* do namorado de Glorinha.
- 3) Ninguém aqui vai *levar vantagem*.
- 4) Glorinha é jovem, *levando a vantagem* de ser bonita.

Ademais, a inclusão de um determinante pode mudar completamente o sentido de uma expressão. Cf. em francês:

prendre un parti = se décider

prendre son parti = se résigner

prendre le parti de quelqu'un = se ranger de son côté

- *manter distância de alguém* C *manter uma distância*

- *marcar presença* C *marcar (a) sua presença*

Com respeito ao uso e funcionamento destas unidades complexas é preciso acrescentar outro aspecto. Geralmente, no interior do campo semântico de cada palavra, gravita seu oposto no léxico mental dos indivíduos. Jean Cohen afirma que “a solidariedade dos opostos só atua no nível da língua. Desde que se passa à fala essa solidariedade se rompe. (apud Diaz: 1984, 145) Isso explica porque geralmente as EIs e os sintagmas cristalizados não se enquadram no princípio da oposição e da negação. Explica-se: não estamos no nível do sistema mas da norma; por conseguinte está registrada apenas uma das formas na memória coletiva. Assim, não podemos dizer por exemplo: * *idéias brancas* embora digamos *idéias negras*, * *idéias desfeitas*, mas *idéias feitas*; *amor livre*, mas não * *amor preso* e assim por diante.

Às vezes há a possibilidade de empregar uma ou outra:

(não) levar a mal
(não) levar desaforo para casa
(não) ter coração

Contudo, em geral, é impossível passar do negativo ao afirmativo e inversamente.

não fazer mal a uma mosca -- mas não: * *fazer mal a uma mosca*
levar a ferro e fogo -- mas não: * *não levar a ferro e fogo*

Em parte isso se explica porque na EI o núcleo da significação não é mais redutível ao verbo; estende-se a seus complementos. Esse núcleo funciona como pivô para uma interpretação idiomática da EI.

III. Examinemos de modo mais abrangente as combinatórias lexicais. Para isso vou me servir de um estudo de Ornella Corazzari (1992): “Phraseological Units”. Seguindo a proposta de Corazzari, vamos chamar estas unidades lexicais complexas de *unidades fraseológicas* (UF), embora possamos atribuir-lhes outras denominações. Apenas para simplificar e respeitar o parâmetro da autora.

Começemos por lembrar que a identificação das UF é fundamental para a análise computacional de textos porque elas se comportam irregularmente tanto morfossintaticamente como semanticamente. Ademais, não se trata de um problema lingüístico menor, muito pelo contrário. Tais UFs incluem uma vasta gama de combinatórias que foram estudadas superficialmente pelos lingüistas teóricos.

A autora propõe a seguinte definição:

unidades fraseológicas são seqüências de palavras que têm uma coesão interna do ponto de vista semântico e que possuem propriedades morfossintáticas específicas.

Examinemos mais de perto a definição de UF:

1) são seqüências de, pelo menos, duas palavras separadas por brancos, hífens o apóstrofos:

mercado negro; caixa eletrônico; caixa preta; bóia-fria; decreto-lei; mina d'água

Embora sejam compostas por mais de uma palavra, elas se classificam funcionalmente como uma única categoria léxico-gramatical. Assim *caixa eletrônico* composto

de N+ Adj comporta-se semanticamente como substantivo e *levar a cabo* e *levar grana* comportam-se como verbo.

De um ponto de vista sintático as UF têm graus diferentes de cristalização, isto é, elas resistem a algumas manipulações morfossintáticas (transformações, inserção de modificadores, flexão) e comutações léxicas que são geralmente possíveis com construções equivalentes comuns. (Corazzari, 1992, 5)

Também com respeito à questão da idiomaticidade, as UFs divergem: o significado global da EI pode depender ou não do significado de suas unidades léxicas componentes.

As EIs são expressões semanticamente opacas cujo significado não depende do sentido de cada um de seus componentes. Por outro lado, *colocações* são seqüências semanticamente transparentes, formadas de itens lexicais que geralmente co-ocorrem. Para Cruse (apud Corazzari, 1992) uma *metáfora morta* é uma expressão que adquiriu um significado metafórico estereotipado (por ex. *ter os pés na cova*). Do ponto de vista semântico essas últimas parecem inserir-se entre as expressões totalmente transparentes e totalmente opacas. Essas três classes distintas de UF são definidas principalmente com base na noção de “constituente semântico”, isto é, um fragmento de frase que possui um significado constante qualquer que seja o contexto. As *EIs* e *metáforas mortas* deveriam ser consideradas como unidades léxicas simples e mínimas embora sejam lexicalmente complexas. Inversamente, as *colocações* mantêm sua composicionalidade porque elas têm componentes que são eles próprios constituintes semânticos.

Outros autores consideram outros aspectos da UF: expressões *fixas*, *semi-fixas* e *variáveis*, dependendo do tipo de modificações morfossintáticas que elas admitem. As *semi-fixas* aceitam variações lexicais e flexões de alguns de seus componentes até um determinado ponto, enquanto as *variáveis* admitem a inserção de modificadores (adjetivos, advérbios).

IV. Gross e Danlos analisaram as construções com verbo suporte (CVS) consideradas como um tipo particular de colocação como:

fare un viaggio / fazer uma viagem; to take a walk / commettre une aggression

Cf. também em português: *bancar o palhaço; dar um tempo; entregar o ouro; fazer corpo mole; ganhar tempo; ter cara de pau; ter vergonha na cara*.

O verbo parece semanticamente vazio, ou esvaziado de conteúdo. O significado global e as restrições de seleção são determinadas pelo complemento verbal e não pelo verbo.

Outras classificações contemplam numerosos tipos de estrutura interna destas UFs. Vejamos alguns exemplos para o sintagma nominal:

N **de** N: *bodas de prata / prata da casa*
olho de águia; jogo de cintura

N **a** N: *cara a cara; cheque ao portador*

N **em** N: *salto no escuro; olho no olho*

N Adj: *cesta básica; circuito elétrico; jogo cênico/ jogo duplo*

Adj N: *baixo astral; livre arbítrio/ livre concorrência*

N N: *efeito estufa; seguro desemprego*
 N e N: *parede e meia; unba e carne*
 N Part Pres: *água fervente* [raro]; *conta corrente*
 N Part Pas: *corrente alternada; condomínio fechado; jogo de cartas marcadas*
 V **conj** V: *leva e traz* [raro]; *vai e vem*
 V V: *corre-corre* [raro]
 N Prep Adj N: *centro de alta pressão; centro de baixa pressão; direito de livre defesa; exercício do livre-arbítrio; fio de alta tensão; regime de livre empresa*
 Adj N Prep N: *livre flutuação de preços; livre fluxo de (da) informação; livre negociação de (dos) salários*

Embora as UF sejam estruturalmente complexas, tanto sintática como semanticamente, elas se comportam como verbos, substantivos, advérbios, adjetivos, preposições, etc. Um grande número de expressões que pertencem às categorias sintáticas menores tais como preposição, advérbio, etc. são, em geral, denominadas *locuções* nas gramáticas tradicionais e nos dicionários.

As UF possuem graus diversos de idiomaticidade e de cristalização (que Corazzari chama de *frozenness*). Assim as *lexias complexas*:

dona de casa, jogo de cintura, livre arbítrio, mercado negro

- têm maior grau de lexicalização do que:

centro de alta pressão, jogo amistoso, jogo de cartas, livre fluxo de capitais, tecnologia de ponta.

Se confrontássemos corpora textuais de épocas diferentes da língua seria possível observar as etapas sucessivas de lexicalização por que passa uma combinatória de palavras até se cristalizar como unidade léxica. No processamento automático de um corpus lingüístico, o analista atua sobre um corte sincrônico na diacronia da evolução lingüística – por isso tais diferenças têm que ser consideradas.

V. Critérios para a classificação das UF.

A identificação e definição dos diferentes tipos de UF não é problema simples; longe disso. A aplicação de testes mostra que a linha divisória entre classes diferentes não pode ser claramente determinada. Entre construções totalmente cristalizadas (p. ex. *bilhete azul, tudo azul*) e menos cristalizadas (p. ex. *tomar uma decisão/ tomar uma importante decisão, tomar qualquer decisão*) e entre EI (p. ex. *levar a ferro e fogo*) e não idiomáticas (*levar em conta, levar em consideração*) há muito casos intermediários que mostram graus diferentes de cristalização e de idiomaticidade.

A hipótese é que as UF são expressões irregulares cuja distribuição e interpretação não podem ser explicadas a partir de uma análise composicional e com base nas regras gerais da língua. Portanto, os testes aplicáveis serão diferentes. Além disso, em princípio, os testes a serem usados são monolíngues, ou seja, baseiam-se nas regras específicas da língua em questão, embora possam eventualmente ser aplicados a mais de uma língua.

1º Seqüências sem verbo

Vejam os testes de Gross (1988 apud Corazzari) adequados para seqüências Adj + N:

(uma) mesa redonda [em um simpósio p.ex.] -- não é uma forma da mesa
caixa preta [dos aviões] -- não indica a cor da caixa
mercado negro / olbo gordo / bilhete azul

Nesses casos o adjetivo expressa uma “qualidade ou propriedade” do nome mas constitui uma unidade com ele, não se tratando de predicação nestas UF. Para esse tipo de *lexia complexa*:

a) não é possível a nominalização:

* a gordura do olho

* o pretume da caixa

* o negrume do mercado

b) também a modificação adverbial é impossível

* *olbo muito gordo*

* *caixa muito preta*

* *mercado muito negro*

* *bilhete muito azul*

c) a coordenação é sempre possível com elementos que exercem o mesmo papel sintático; não neste caso. Pode-se dizer por exemplo:

olbo preto e redondo

mas não: * *olbo gordo e redondo*

caixa branca e grande

mas não: * *caixa preta* [a dos aviões] *e grande*

O adjetivo da *lexia complexa* parece ser obrigatório e, às vezes, pode ser até usado isoladamente no decorrer no contexto. Cf.

No negro o preço do dólar era R\$ 2,90.

raios ultravioletas → ultravioletas → Os ultravioletas podem causar danos à pele.

d) flexões

Essas UFs (seqüências sem verbo) têm características diferentes relativamente ao número:

– algumas são usadas só no plural:

bodas de prata, bens de capital, comunhão de bens, encargos sociais, finanças públicas, jogo de búzios, jogo de panelas, jogos olímpicos, mercado de capitais

– outras são usadas só no singular:

jogo do bicho, jogo de vida ou morte, linba dura, mercado negro, queda livre, prata da casa, sistema binário, zona azul

– outras ainda são usadas no sing. e no pl.:

ao pé de/aos pés de; jogo-de-azar/ jogos-de-azar; jogo(s) de palavras; jogo (s) eletrônico(s); levantar vela(s); na ponta do pé (dos pés); questão de terra/ questões de terra; sistema de comunicação/ sistema de comunicações; sistema de esgoto(s)

e) restrições de co-ocorrência

Apenas um adjetivo ou um número restrito de adjetivos pode combinar-se com o nome ou vice-versa. Nesse caso a substituição do nome ou do adjetivo por outro do mesmo campo semântico ou por um sinônimo dá um resultado incorreto. Cf.

bom senso / * ótimo senso
classe trabalhadora / * classe trabalhista
custo real / * custo verdadeiro
jogo amistoso / * jogo amigável
luta livre / * briga, combate livre
queda livre / * caída livre

Estes e outros testes mostram que temos aqui casos de *lexias complexas*. Não se pode dizer tampouco no caso de *luta livre*:

* a luta é livre / *a luta é muito livre
* a liberdade da luta

Cf. ainda: *efeitos especiais*

* os efeitos são especiais
* a especialidade dos efeitos
* efeitos muito especiais
* efeitos especiais e inesperados

Contudo, muitas expressões não são totalmente cristalizadas nem totalmente livres, mas casos intermediários de sintagmas lexicalizados que podem ser flexionados. A modificação adverbial, por exemplo, é possível em apenas uns poucos casos [*amigo íntimo*: *amigo* muito *íntimo*; mas não *vontade férrea*: *vontade* muito *férrea* ou *custo astronômico*: *custo* muito *astronômico*]. Muitas vezes a inserção de expressões parentéticas é possível [ex. *efeitos*, por assim dizer, *especiais*; *um amigo*, por assim dizer, *íntimo*].

A comutação do modificador é possível, às vezes. Ex:

boas idéias / *ótimas idéias*
idéias gerais / *idéias genéricas*
boas maneiras / *finas maneiras*

A comutação do substantivo já é mais rara. Contudo, a rigor, apenas um número restrito de itens lexicais podem substituir o adjetivo ou o substantivo de uma UF.

2º Seqüências com verbos.

Testes para avaliar o grau de cristalização da UF.

a) inserção

levar ao ar / *levar rapidamente ao ar*
levar anos / *levar muitos anos* / *levar anos e anos*
ter a cabeça feita / *ter a cabeça totalmente feita*
ter bossa / *ter muita bossa*

b) a extração de um componente da UF por meio de certas construções sintáticas geralmente não é possível.

- passivação: * os anos são (foram) levados [base: levar anos]
- relativização: * a cabeça que foi feita pela TV [base: ter a cabeça feita]
- c) substituição por uma pro-forma

A substituição de um ou mais elementos da UF por um pronome também não é possível. Cf.: * a TV *a* tinha feito (a cabeça)

- d) modificação (flexão, comutação de determinante, adjetivo)

Em geral não são possíveis operações sintáticas em que os componentes da UF são flexionados ou substituídos.

levar no bico / * levar nos (bicos), num, no seu bico

levar grana / * levar dinheiro, tutu

Os testes evidenciam que se pode distinguir entre expressões totalmente cristalizadas e seqüências freqüentes que podem ser modificadas livremente. Na verdade as expressões totalmente cristalizadas não são muito freqüentes, sendo comuns os casos intermediários, em que são permitidas algumas modificações morfossintáticas. Tudo se passa na língua e no vocabulário como numa pista de corrida – muitos corredores já ultrapassaram a barreira de chegada, outros estão-se aproximando dela e outros vêm chegando de mais longe. Nessa imagem retomada de minha tese de doutoramento, estou simbolizando o pódio de chegada como o término do processo da lexicalização.

Concluindo esse tópico:

- modificações possíveis:
- comutação verbo/ nome/ adjetivo

levar a vida toda / *levar a vida inteira*

levar a sua vida / *levar a sua vidinha*

ter dó / *ter pena*

- apagamento do determinante / da preposição
- *ter um enfarte* / *ter enfarte*
- *ter um (uma) amante* / *ter amante*
- *ter disposição (para)*

Entretanto, muitas vezes, a comutação do determinante ou da preposição acarreta uma mudança no sentido global da expressão.

Em geral a flexão lexical não é possível. Também no caso de UF com verbo há expressões que se usam no plural e outras que se usam apenas no singular. Por exemplo:

singular: *levar calote*, *levar chumbo*, *ter fome*, *ter frio*;

plural: *ter alucinações*, *ter engulbos*; *ter as mãos livres*

Outras expressões, porém, são usadas tanto no singular como no plural e geralmente com o mesmo significado. Cf.

ter cbilique = *ter cbiliques*

ter despesa = *ter despesas*

ter compromisso = *ter compromissos*

ter dúvida = *ter dúvidas*

Não vão ser tratadas aqui as lexias complexas formadas com adjetivos, advérbios, bem como das locuções prepositivas, adverbiais, conjuncionais e pronominais porque isso nos obrigaria a ampliar muito o tamanho deste artigo. Creio que o que foi acima discutido já evidencia a complexidade desta problemática e como podemos analisar essas *combinatórias lexicais*.

VI. Consideremos, por fim, o caso mais extremo de unidade complexa e heterogênea – a expressão idiomática (EI). Esse tipo de seqüência apresenta uma grande fixidez, razão por que deve integrar o acervo lexical da língua, não sendo, portanto, uma seqüência discursiva. Cf. por exemplo:

arrancar o mal pela raiz; botar as barbas de molbo, botar as cartas na mesa, botar (pôr) lenha na fogueira; colocar uma pá de cal em; comer o pão que o diabo amassou.

As EIs de uma língua exibem uma enorme heterogeneidade. Do ponto de vista da sua natureza, esse tipo de *combinatória lexical* é fruto da cultura. Relativamente ao Português Brasileiro, uma parcela das expressões idiomáticas do PB foram herdadas de nossa cultura-mãe, a portuguesa, recebidas juntamente com o patrimônio cultural que é o léxico. Contudo, dada a diversidade do meio ambiente e da sociedade brasileira comparada à portuguesa, bem como a dinâmica da língua, foram-se criando variantes do acervo herdado, ao mesmo tempo que se iam construindo criações novas. É preciso enfatizar que a questão das EIs nos remete ao domínio da norma e não da língua. Assim sendo, as EIs são aprendidas de cor como se aprende o vocabulário do idioma e elas fazem parte do acervo da cultura e não do sistema lingüístico. Por outro lado, sabemos que estas EIs vão sendo armazenadas na memória individual e na memória coletiva e passam a fazer parte do léxico da língua. Entretanto, e infelizmente, raramente elas foram registradas nos dicionários da língua como seria desejável.

Segundo O. Diaz. (1984) as EIs "traduzem hábitos verbais".

Daí resulta que essas seqüências cristalizadas geralmente não possuem correspondência em outra língua. Um exemplo característico seria a expressão *quebrar o (um) galbo* do Português Brasileiro. Cada termo deste sintagma não conserva sua identidade própria e se torna assim não-analisável. Por essa razão, os constituintes de uma EI como essa se tornam indissociáveis, não permitindo a supressão ou acréscimo de um elemento. As EIs são típicas de uma nação e enraizadas na sua cultura. O Português Europeu, por exemplo, não possui esta EI.

De fato, as EIs podem ser consideradas como uma lista de irregularidades fundamentais de um idioma e de uma norma lingüística.

Lembremos ainda que as EIs são sintagmas que não têm paradigmas. No uso corrente, sobretudo da linguagem coloquial e também literária, as EIs podem ser empregadas deformadas – para ironizar, por exemplo.

Concluindo: é necessário que se faça um amplo e exaustivo estudo das unidades complexas do léxico do português para que elas possam ser adequadamente descritas e registradas nos dicionários.

BIBLIOGRAFIA

- BENSON, M.; BENSON, E. e ILSON, R. (1986), *The BBI Combinatory Dictionary of English: a guide to word combination*. Amsterdam, Philadelphia, John Benjamins Publishing Company.
- BIDERMAN, M. T. C. (2001), *Teoria Lingüística*, (Teoria lingüística Lingüística Computacional), São Paulo, Martins Fontes.
- CORAZZARI, O. (1992), *Phraseological Units*, Consiglio Nazionale delle Ricerche. Istituto di Linguistica Computazionale. Network of European Reference Corpora (NERC), serial nº 68, Pisa (manuscrito).
- DANLOS, L. (1981), "La morphosyntaxe des expressions figées", *in Langages*, Paris, nº 63, pp. 53-74.
- DANLOS, L. (1988), "Les expressions figées", *in Langages*, nº 90, Paris.
- DIAZ, O. (1983-84), "Observations sur les expressions lexicalisées", *in Cahiers du Centre Interdisciplinaire des Sciences du Langage*, Toulouse, Université Toulouse-Le-Mirail, nº 5, pp. 139-152.
- GROSS, M. (1988), "Les limites de la phrase figée", *in Langages*, Paris, nº 90, pp. 7-22.
- GROSS, M. (1982), "Une classification des phrases figées du français", *in Revue québécoise de linguistique*, v. 11, n. 2, pp. 151-185.
- LODOVICI, F. M. M. (1989), *Elementos constitutivos dos Idiomatismos no Português do Brasil*, Pontifícia Universidade Católica de São Paulo (PUC), São Paulo (dissertação de mestrado).
- NASCENTES, A. (1966), *Tesouro de Fraseologia Brasileira*, 2ª ed. Rio de Janeiro, Livraria Freitas Bastos.
- NASCIMENTO, M. F. B. (1998), "Dicionário de Combinatórias do Português (variedade europeia)", *in Alfa*, S. Paulo, Edunesp, pp. 183-203.
- XATARA, C. M. (1994), *As expressões idiomáticas de matriz comparativa*, Araraquara, UNESP, Dissertação de Mestrado.
- XATARA, C. M. (1997), *A tradução para o Português de Expressões Idiomáticas em Francês*, Araraquara, UNESP, Tese de doutorado.

