

Problemas de tratamento e sistematicidade na compilação das combinatórias lexicais no Novo Aurélio Século XXI

*Renata Stela Valente, Iovka Bojítlova Tchobánova,
Mário Sérgio Almeida, Sofia Ferreira*
Instituto de Linguística Teórica e Computacional – ILTEC

1. Introdução

Após a análise de cerca de mil verbetes compreendidos entre as letras a, b e c do dicionário eletrônico Novo Aurélio Século XXI, observou-se diversas incorreções na compilação das combinatórias lexicais. É bem verdade que a compilação de combinatórias lexicais nos dicionários de língua parece ser o calcanhar de Aquiles dos lexicógrafos. Entende-se que a dificuldade em se tratar as combinatórias nos dicionários vem da própria dificuldade em se compreender a “natureza semântica” das diferentes combinatórias lexicais que permeiam uma língua. Os diferentes fenômenos semânticos que permeiam as combinatórias lexicais são ainda pouco estudados por lingüistas e lexicólogos da língua portuguesa, onde as análises se baseiam nas noções de composicionalidade ou não da combinatória lexical e na noção ainda pouco clara de co-ocorrência privilegiada. Além disso, os estudos a respeito das combinatórias lexicais em português observam geralmente a frequência de ocorrências de um determinado seguimento lexical ou as estruturas sintagmáticas das diversas possibilidades combinatórias da língua. Entendemos que o resultado desses trabalhos auxilie na extração automática das combinatórias lexicais; no entanto, contribuem pouco à compreensão do semantismo das combinatórias. À dificuldade de se reconhecer a natureza semântica das combinatórias lexicais soma-se, no caso do português, a colocação do hífen em certas co-ocorrências lexicais que adquirem assim a condição de palavras compostas¹ e, por isso, ganham entradas de dicionário, ou seja, elas são compiladas na nomenclatura dos dicionários.

Neste trabalho, além de serem apresentados diversos problemas encontrados na compilação das combinatórias lexicais no Aurélio eletrônico século XXI, discute-se a natureza

¹ Estão excluídos desta observação os nomes de espécie tais como *begónia-de-folha-estreita*, *amendoeira-da-praia*, *andorinha-do-oco-do-pau*, *pomba-legítima*, etc.; os adjetivos gentílicos como *ponte-pretense*, *riograndense*, etc.

das combinatórias lexicais à luz do sistema de Funções Lexicais (FL) da Lexicologia explicativa e combinatória – LEC (Mel'čuk *et al.* 1995) que deriva da Teoria Sentido Texto - TST (Mel'čuk 1981, 1984, 1987, 1997) e identificam-se, por meio das ferramentas de análise que essa teoria oferece, as razões dos problemas de compilação encontrados. Finalmente, apresentam-se argumentos em favor de uma compilação “ergonômica” inspirada no *Dictionnaire explicatif et combinatoire* - DEC (Mel'čuk *et al.* 1984, 1988, 1992, 1999) da TST, e possível graças à tecnologia informática da qual dispomos nos dias atuais.

2. A combinatória lexical no Novo Aurélio Século XXI

Na seção *Como Usar o Aurélio Século XXI* (p. XV), define-se o que é uma locução² e como ela é apresentada:

“A locução é formada por duas ou mais palavras que representam uma só unidade semântica, sendo uma delas a que constitui a cabeça do verbete. As locuções, sempre em **negrito**, estão no fim do verbete, após a marca ◆. Cada locução pode ter uma ou mais definições, como se fosse ela mesma um pequeno verbete, com transcrições fonéticas (no caso de expressão em língua estrangeira), rubricas, exemplos, abonações, achegas, etc. (...)”

Como se pode notar nesta definição, uma locução tem uma ou várias definições, em outras palavras, uma ou várias acepções. Assim, por exemplo, no verbete *baile*, encontra-se a seguinte locução com seus diversos significados:

Dar um baile em. *Bras.*

1. Mexer com; troçar de.

2. Chamar a atenção de; repreender;

3. Exercer domínio absoluto sobre (o adversário); *O Flamengo deu um baile no Fluminense.*

A locução pode também, segundo a orientação de compilação do Novo Aurélio, ter todas as outras informações que constam do verbete de uma palavra, isto é, transcrição fonética, abonações, etc. Note-se que a definição de locução fornecida dá conta da *consciência da autonomia semântica de uma locução por parte dos lexicógrafos*. Essa autonomia semântica da locução leva à pergunta: sendo a locução semanticamente autônoma ela não mereceria uma entrada de dicionário? Neste trabalho somos a favor de uma descrição autônoma de um *certo tipo* de locução, pois há razões práticas, teóricas e tecnológicas que favorecem o *tratamento individual de certas locuções*. Essas diferentes razões serão tratadas nas seções que seguem.

Ainda com respeito à definição de locução, ela traz à tona pelo menos três grandes dificuldades de tratamento pelo lexicógrafo:

² Neste trabalho, os termos *combinatória lexical*, *co-ocorrência lexical* e *locução* são considerados sinônimos.

- 1) o de traçar uma linha demarcatória entre palavra composta e sintagma sem hífen;
- 2) o de conceder ou não a uma combinatória com autonomia semântica uma entrada de dicionário;
- 3) o de fazer uma combinatória lexical vir a ser tratada como uma palavra composta: *chove-não-molha*, *papos-de-aranha*, *ponta-cabeça*, etc.

2.1. Problemas

Problema 1 : tratamento e sistematicidade [razão prática]

Foram encontrados no *Novo Aurélio Século XXI* os seguintes casos na compilação das locuções:

- 1) locuções que vêm acompanhadas das suas definições, por exemplo, *sair de banda* = escapulir-se furtivamente;
- 2) locuções que são definidas por outras locuções sinônimas, *levar uma bandeira* = *levar um fora*;
- 3) locuções que são definidas por locuções sinônimas e estas por palavras simples sinônimas, *dar de língua* = *dar à língua* = *tagarelar*, *parolar*, *taramelar*, *linguajar*;
- 4) locuções que são definidas por palavra simples sinônimas e estas definidas tanto por palavras simples sinônimas quanto por locuções sinônimas, *bafo de boca* = *bafo*, = *conversa fiada*, *bazófia*, *gabolice*, *prosa*;
- 5) locução definida por outra locução seguida da definição, *vir à baila* = *vir a propósito*³ = fazer-se lembrado oportunamente.

Como se observa, a definição das locuções por sinônimos⁴ é uma constante no dicionário. Entende-se que a arte da confecção de um dicionário em formato papel estava e ainda está condicionada às restrições de volume e peso, o que faz com que a definição por sinônimos seja uma constante na técnica lexicográfica. Desse modo, uma locução é definida e as locuções sinônimas lhe são remetidas. Como mostram os exemplos acima, há locuções que são definidas por locuções sinônimas e estas são definidas por outras locuções ou palavras simples sinônimas. Ora, um consulente que procura o significado de uma locução cujo verbete remete a uma outra locução e o verbete desta ainda a outra locução ou palavra simples, criando-se assim uma concatenação de remissivas sem uma definição analítica, vê frustrada sua procura pelo significado da locução se, nessa busca pelo significado nenhuma das locuções sinônimas ou palavras simples apresentadas lhe for familiar, podendo assim inferir o significado da locução que procura. É necessário lembrar que os dicionários de língua contam com o conhecimento linguístico nato do consulente. Um consulente estrangeiro que esteja aprendendo o português se verá em *papos-de-aranha* com tal consulta.

³ Esta combinatória não está compilada nem no verbo *vir* nem no substantivo *PROPÓSITO*.

⁴ Devemos esclarecer um ponto aqui: fazemos um abuso de linguagem quando dizemos "definição por sinônimo": um sinônimo não é uma definição.

Problema 2: combinatórias lexicais diferentes tratadas da mesma forma

Observou-se ainda que as locuções *conversa fiada* e *brincar com fogo*, por exemplo, recebem o mesmo tratamento isto é, são compiladas na seção reservada às locuções⁵ nos verbetes *CONVERSA* e *FOGO* respectivamente. Ora, ainda que classificadas sob uma mesma etiqueta isto é, locuções, essas duas combinatórias são diferentes: em *conversa fiada* um significado se mantém, o de ‘conversa’, em *brincar com fogo* nenhuma das lexias que constroem o sintagma mantém seu significado, isto é, não há composicionalidade. Temos assim dois fenômenos lingüísticos que merecem tratamento lexicográfico distinto, a saber:

- 1) a locução *conversa fiada* tem vínculo semântico com a lexia *conversa*. Deve então ser tratada no verbete *CONVERSA*;
- 2) a locução *brincar com fogo* não tem vínculo semântico nem com a lexia *fogo*, tampouco com a lexia *brincar*. Por essa razão não deve ser tratada nos verbetes *FOGO* OU *BRINCAR*.
- 3) a locução *brincar com fogo* tem um significado autônomo, ou seja, desvinculado dos significados das lexias *brincar* e *fogo*. Por isso ela deve ter uma entrada de dicionário.

Problema 3: o hífen

Observamos ainda a compilação de palavras compostas escritas com hífen como entrada de dicionário e a compilação dentro de verbetes das locuções sem hífen. Assim, *banho-de-cheiro* tem uma entrada de dicionário, enquanto que *banho de ervas* ou *banho de folhas*, seus sinônimos, por não serem hifenizados foram compilados no verbete *banho*. A presença ou não do hífen numa locução leva à distinção entre palavra-composta e locução, fazendo com que certas combinatórias apareçam na nomenclatura do dicionário enquanto outras são compiladas na seção reservada às locuções.

Como se sabe, recomenda-se que os compostos que constituem uma unidade semântica e que mantêm acentuação própria sejam escritos com hífen. Segundo Cunha (1982), o hífen “(...) indica que dois ou mais vocábulos formam uma unidade semântica”. Temos assim *fim-de-semana*, *pôr-do-sol*, *copo-de-leite*, etc. compostos que designam um referente. Esses compostos têm necessariamente uma entrada de dicionário. Outros compostos submetidos a essa regra designam um estado como *papos-de-aranha* ou *chove-não-molha*.

⁵ O Dicionário Aurélio explica o seguinte: “Adotou-se, aqui, em relação à fraseologia, o critério, eminentemente prático, do *Diccionario da Real Academia Espanola*, critério também seguido por Nascentes: se a frase contém substantivo ou palavra substantivada, naquele ou nesta se faz o registro (assim, *ação entre amigos* virá em *ação*; *pôr a mão na consciência*, em *mão*; “O prometido é devido” [provérbio], em *prometido*); seguem-se, na ordem de preferência, o verbo, o adjetivo, o pronome e o advérbio; havendo na expressão mais de um vocábulo da mesma classe gramatical, será o primeiro deles o preferido (p. ex. *tirar a sardinha com a mão do gato* figura em *sardinha*); as palavras *pessoa* e *coisa*, o pronome *alguém*, os verbos de ligação – *ser*, *estar*, *continuar*, *ficar*, etc. –, se não constituírem parte essencial, imutável, da locução, não são levados em conta; nem se levam em conta os verbos auxiliares” (Aurélio Século XXI, prefácios, XI). [diretiva apresentada na primeira edição do Aurélio, 1975].

Note-se que esses compostos são introduzidos por verbo: estar, ver-se, encontrar-se em *papos-de-aranha*, ficar num *chove-não-molha*. Há compostos que designam igualmente um estado, mas por não serem escritos com hífen, não ganham uma entrada de dicionário, tal o caso de (estar) *entre a cruz e a caldeirinha* ou (estar) *entre a cruz e a espada*, locuções que no *Aurélio Século XXI* aparecem compiladas no verbete CRUZ. Pelo exposto temos que:

1. apenas os compostos hifenizados têm uma entrada de dicionário;
2. os compostos que geram uma unidade de sentido, mas escritos sem hífen não têm uma entrada de dicionário;
3. o consulente que procura a definição de uma locução no dicionário precisa saber antecipadamente se ela é escrita com ou sem hífen; se escrita com hífen, ela constará da nomenclatura do dicionário. Do contrário, sem hífen, ela aparecerá na seção final do verbete de uma das lexias que constitui a locução.

Alguns autores argumentam que o hífen em palavras compostas é responsável pela alteração de significados. Segundo eles, *pão-duro* com hífen é a pessoa avarenta e *pão duro* sem hífen é o pão que está endurecido. Encontramos afirmações como “o hífen pode mudar totalmente o significado das palavras”⁶. Não podemos concordar com tal afirmação. Um traço gráfico que aparece na escrita não pode e não deve alterar o significado de um segmento de lexias. Parece-nos que esse tipo de afirmação decorre do processo de *análise*, no qual, o interlocutor, mas neste caso específico, o leitor tomaria em consideração o hífen para decodificar o significado da locução. Ora, sabemos que a locução *pão-duro* significando ‘avarento’ mesmo escrita sem hífen, será em *contexto* decodificada pelo leitor com o significado ‘avarento’. Ele certamente não confundirá o significado da locução *pão-duro* com os lexemas *pão duro*.

No nosso trabalho, dá-se prioridade absoluta ao falante. Assim, nossa análise orienta-se pelo processo de *síntese*, isto é, o falante busca no seu “dicionário mental” as lexias desambiguadas para compor a sua frase⁷. Ele conhece muito bem as lexias que quer empregar para construir a sua frase.

Esses dois processos, o da *síntese* e o da *análise* levam a considerações teóricas diferentes: no processo da *síntese*, o do falante, cada lexia de um vocábulo⁸ tem um e apenas um significado. Pelo contrário, no processo de *análise*, o do interlocutor, uma lexia carrega-se de um determinado significado segundo o contexto no qual ela se insere⁹.

⁶ <http://www.ufv.br/tutoria/portugues/hifen.htm> sob o tema “Coisas do Hífen”.

⁷ Do percurso Representação Semântica até a Representação Fonológica,

⁸ Compreenda-se aqui vocábulo segundo a definição da TST: “Um vocábulo é o conjunto de todas as lexias L_1, L_2, \dots, L_n que satisfazem simultaneamente às duas condições seguintes:

1. os significantes de L_1, L_2, \dots, L_n são idênticos;

2. os significados de duas lexias entre L_1, L_2, \dots, L_n são ligados (diretamente ou indiretamente)” (MeI’çuk *et al.* 1995:159).

⁹ Note-se que o processo de *análise* dá lugar a certas considerações divulgadas na Linguística tais como “uma lexia pode ter vários significados que se alternam dependendo do contexto discursivo.” Somos claramente contrários a essa consideração, pela impossibilidade i) de um modelo linguístico que dê prioridade à actividade linguística do interlocutor e, conseqüentemente, ii) da *alternância* do significado de uma lexia.

Pelo exposto, considera-se i) inapropriada a afirmação que um sinal gráfico, o hífen, distinga significados; ii) inadequado que o hífen determine a apresentação de uma combinatória lexical na nomenclatura de dicionário.

3. Quadro Teórico: Postulados da TST

Como foi dito na introdução deste trabalho, apresentaremos com base no sistema de *Funções Lexicais* (FL) da LEC que deriva da TST, soluções para alguns dos problemas observados na compilação das combinatórias lexicais no Aurélio Século XXI. Desde logo, iniciamos pelos três postulados do modelo de representação lingüística, o *Modelo Sentido-Texto* (MST), e em seguida apresentamos de forma sucinta o sistema de FL que trata as combinatórias da língua e, finalmente, apresentamos algumas propriedades do *Dictionnaire Explicatif et Combinatoire* (DEC) no qual as formulações teóricas da TST e da LEC são postas em prática.

O MST parte do *sentido* em direção ao *texto*¹⁰, o que justifica o nome da teoria, ou seja, do sentido da *síntese* ou da produção da fala e não no sentido da análise ou da compreensão da fala como já foi explicado na seção anterior. É importante salientar que a estrutura geral do MST parte de uma RSem em direção a uma *representação fonética* [=RFon], com as diversas representações que se intercalam entre essas duas. Deste modo:

$$\text{RSem} \rightarrow \text{RSintP} \rightarrow \text{RSintS} \rightarrow \text{RMorfP} \rightarrow \text{RMorfS} \rightarrow \text{RFonP} \rightarrow \text{RFonS}$$

Como tal, são três os postulados da TST:

Postulado 1: A língua como Correspondência “Sentido-Texto”

A língua é um sistema finito de regras que especifica uma CORRESPONDÊNCIA multi-multívoca entre um conjunto finito contável de sentidos e um conjunto infinito contável de textos.

$$(1) \{ \text{RSem}_i \} \leftarrow \text{langue} \Rightarrow \{ \text{RFon}_j \} \quad | \quad 0 < i, j < \infty$$

Note-se que este postulado pressupõe o caráter discreto das representações e, por conseguinte, do modelo.

Postulado 2: O modelo Sentido-Texto (MST) como ferramenta de descrição de línguas

A correspondência (1) [= entre sentidos e textos] deve ser descrita por um DISPOSITIVO LÓGICO constituindo um modelo funcional da língua tipo *Sentido-Texto* e elaborado na direção *Sentido-Texto*¹¹.

¹⁰ Compreenda-se a palavra *texto* aqui como a produção oral do falante.

¹¹ Segundo Mel'cuk (1997 : 5), « D'un point de vue formel, le passage "Sens Texte" et le passage "Texte Sens" sont équivalents. Mais du point de vue de l'élaboration du modèle par des linguistes et de sa présentation, ce n'est pas du tout le cas. (...) ».

Desta forma, o MST deve receber na entrada uma R_{Sem} e produzir na saída uma R_{Fon}, reproduzindo de forma válida a correspondência entre o *sentido* que um locutor quer expressar e o *texto* que veicula esse sentido. O MST disponibiliza a fala de forma a ser a tradução do significado.

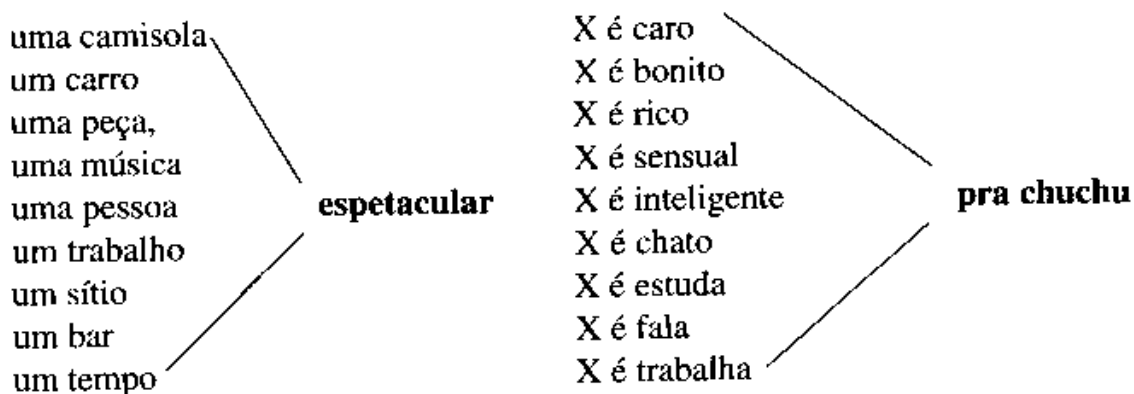
Postulado 3: A Frase e a Palavra como Unidades de Base da Descrição Lingüística

Na descrição da correspondência (1), dois NÍVEIS INTERMEDIÁRIOS de representação dos enunciados são necessários de forma a destacar os fatos lingüísticos pertinentes: a representação SINTÁTICA [= RSint] que corresponde às regularidades específicas da FRASE e a representação MORFOLÓGICA [= RMorf] que corresponde às regularidades específicas da PALAVRA.

O MST considera a frase e a palavra [para sermos mais precisos a *palavra-forma* Mel'čuk (1993)] que são as unidades máxima e mínima da fala, unidades autônomas e universais. No âmbito da *frase*, verifica-se a ordem das palavras, a concordância e o regime, a estrutura comunicativa, a co-ocorrência lexical restrita, etc., enquanto que na *palavra* observa-se a flexão e a derivação, bem como as alternâncias fonêmicas. É precisamente para apreender as características da frase e da palavra que o MST postula os dois níveis intermediários entre sentidos (significados) e textos. Esses níveis oriundos já de tradição são observados por quase todas as escolas lingüísticas.

3.1. A combinatória lexical tratada pela TST [razão teórica]

A *combinatória lexical* é um fenômeno das línguas naturais. A noção de *combinatória lexical* corresponde à associação ou combinação de dois (ou mais lexemas), no qual um lexema é, na nomenclatura da TST, a *palavra-chave* e o outro o *valor*¹². Há combinatórias lexicais livres e combinatórias lexicais restritas. As combinatórias lexicais livres são representadas por lexemas que se combinam com um vasto número de lexemas. Assim, o lexema *espetacular* em português europeu e a locução *pra chuchu* em português brasileiro se combinam com diversos lexemas:



A TST define as combinatórias lexicais as quais chama *frasema* em três tipos:

¹² Hausmann (1979) chama esses lexemas respectivamente de *base* e *colocativo*.

Frasema completo

o resultado da combinação dos seus elementos é diferente do significado de cada elemento que o constitui:

'A' + 'B' = 'C'

brincar com fogo, barata tonta, pedra no sapato, dar um baile

Semi-frasema ou *colocações*

i) o resultado da combinação dos seus elementos é parcialmente diferente do significado de um elemento que o constitui ou seja, um dos elementos mantém o significado e o outro revela um outro significado:

'A' + 'B' = 'AC'

banho de chuva, banho salgado, banho de cheiro, banho turco

ii) o resultado da combinação dos seus elementos mantém o significado de cada elemento:

'A' + 'B' = 'AB'

bandeira a meio pau

Como podemos notar, neste subtipo de semi-frasema, há composicionalidade, ou seja, o significado 'A' mantém-se bem como o significado 'B'. Note-se, no entanto, que há seleção do *colocativo* pela *base*:

Bandeira a meio pau

Bandeira a meio mastro

**Bandeira a meia vara*

Quasi-frasema

o resultado da combinação dos seus elementos mantém o significado de cada elemento e um outro significado é acrescentado aos mesmos.

'A' + 'B' = 'ABC'

bandeira branca : 'bandeira' + 'branca' = 'bandeira' + 'branca' + 'paz'

O fenômeno das combinatórias lexicais é bastante evidente quando se comparam diversas línguas. Evidencia-se o fato de que uma mesma base em línguas variadas seleciona lexemas particulares para expressar o mesmo significado. Para se dizer em português, alemão, russo, húngaro, árabe e chinês,¹³ que alguém recebeu muitos aplausos (um solista, por exemplo, pela sua execução majestosa) se diz, a partir da base *aplausos* o seguinte:

¹³ As combinatórias do alemão, russo, húngaro, árabe e chinês foram extraídas de Mel'čuk (1997: 26-27).

Português

Magn¹⁴ (aplauso) = //¹⁵ ovação

Alemão

Magn(*applaus*) = *tosender* ‘bramante’

Russo

Magn(*aplodismenty*) = *burnye* ‘tempestuoso’, *gromovye* ‘estrondosos’

Húngaro

Magn(*taps*) = *viharos* ‘turbilhante’, *vas-* ‘de ferro’

Árabe

Magn(*tasfiq*) = *harr* ‘quente’

Chinês

Magn(*zhangshēng*) = *léidòng* ‘estrondosos’| posposto.

Como se pode notar, são diversas as formas lexicais que as diferentes línguas utilizam para expressar o significado ‘intenso’ representado pela FL **Magn**. A combinatória resulta assim de uma particularidade linguística própria a cada língua. Por essa razão, elas não são calculáveis, mas devem ser interiorizadas da mesma forma que o léxico de uma língua.

3.2 Função lexical (FL): $f(x) = y$

Há dois tipos de FLs: as FLs padrões e as FLs não-padrões.

3.2.1. Função lexical padrão

O conceito de FL padrão assenta na hipótese que os casos de combinatória lexical restrita baseiam-se num pequeno número de sentidos específicos – abstratos e gerais. Mel’èuk (1997) observa que as FLs padrões são ainda universais, ou seja, existem em qualquer língua e são quase suficientes para descrever de forma *sistemática e formal* o conjunto de colocações de uma determinada língua.

Existem 56 FLs padrões identificadas pela TST¹⁶. Entre as FLs padrões distinguem-se as FLs sintagmáticas e as FLs paradigmáticas que estabelecem relações semântico-sintáticas entre lexemas autônomos.

As FLs se estruturam a partir da função matemática $f(x) = y$ onde f corresponde no caso das FLs padrões a um sentido geral e universal, x é a palavra-chave e y é o valor da função.

Atente-se ao quadro explicativo com alguns exemplos de FLs¹⁷:

¹⁴ **Magn** do latim *Magnus* é uma FL padrão que expressa o significado ‘intenso’.

¹⁵ Símbolo indicando que o valor da FL apresenta significados fusionados numa única lexia.

¹⁶ O leitor encontrará uma apresentação detalhada das 56 FLs padrões em Mel’èuk *et al.* 1995, capítulo III.

¹⁷ Os exemplos apresentados são quase todos próprios à variedade europeia.

Função Lexical	Descrição da FL	Tipo	Palavra-chave (base): x	Valor (colocativo): y
Intensificador [Magn]	Modificador adjectival ou adverbial da palavra-chave que expressa o sentido 'muito' ou 'intenso'	sintagmática	pagar apanhar	couro e cabelo; forte e feio;
Laudativo [Bon]	Modificador adjectival ou adverbial da palavra-chave que expressa o sentido 'bom' enquanto aprovação subjectiva do locutor	sintagmática	comida cantar	de lorde; como um rouxinol;
Confirmador [Ver]	Modificador adjectival ou adverbial da palavra-chave que expressa o sentido 'tal como é preciso'	sintagmática	falar conversa	bem; séria; [AntiVer] fiada, à toa, vazia, mole, oca, sem pés nem cabeça;
Suporte [Oper]	Fornece um verbo semanticamente vazio à palavra-chave	sintagmática	remate	fazer [ART ~];
Real [Real]	Fornece um verbo pleno à palavra-chave	sintagmática	cuspo	pegar [com ~]; colar [com ~];
Função Lexical	Descrição da FL	Tipo	Palavra-chave (base): x	Valor (colocativo): y
Sinônimo exato ou quasi-sinônimo [Syn]	Quando os sentidos das lexias A e B têm componentes semânticos em comum e se substituem em determinados contextos	paradigmática	cabeça fria	calma, paciência
Antônimo exato ou quasi-antônimo [Anti]	Quando os significados das lexias são idênticos a não ser pela negação que se encontra num dos dois	paradigmática	chuva torrencial	[AntiMagn _{pouca}] = chuva miudinha

3.2.2. Função Lexical não-padrão

As FLs não-padrões são combinações que não podem ser descritas pelas FLs padrões visto o seu sentido ser *específico* a uma língua e, portanto, *não generalizável*. Por isso, a sua descrição é uma mini-definição e tende a pertencer a um universo especializado. Assim como para as FLs padrões, distinguem-se para com as FLs não-padrões entre sintagmáticas e paradigmáticas. Vejamos alguns exemplos de FLs não-padrões sintagmáticas:

**'com emblema
de um país'** (bandeira) = nacional
**'com emblema de um país
usada em navio mercante ou
embarcação de recreio'** (bandeira) = mercante

4. Dicionário Explicativo e Combinatório (DEC)

O DEC é um dicionário teórico resultante da reflexão a respeito de princípios, regras e critérios de análise de fenômenos lexicais. Nesse dicionário são testadas diversas elaborações teóricas da TST. Por isso, é conhecido como "bancada de testes". Pertence, pois, a lexicologia experimental não sendo vocacionado para o grande público. É importante notar que as análises feitas no âmbito deste dicionário contribuem também ao refinamento das propostas teóricas. Trata-se assim de uma interação constante, num vai-e-vem entre a lexicologia teoria e a lexicologia experimental.

Segundo Alonso Ramos e Mantha (1996) existem seis propriedades básicas do DEC. São elas:

- 1) elaborado no âmbito da TST;
- 2) direcionado para o texto;
- 3) um verbete do DEC baseia-se da definição da lexia: a sua representação semântica serve de base à descrição de todas as relações paradigmáticas e sintagmáticas com as outras lexias da língua;
- 4) enfatiza a co-ocorrência lexical restrita;
- 5) é formal, logo, todas as informações relacionadas ao sentido bem como à sintaxe são indicadas de maneira precisa e explícita;
- 6) a descrição de cada lexia é exhaustiva.

Com relação ao item 4 que nos interessa particularmente no trabalho apresentado nestas páginas, apresentamos a seguir a descrição da co-ocorrência lexical da lexia *condamner*_{1,3}, apresentada no DEC vol. IV, p. 199:

Syn	∩	: châtier, punir, sanctionner
Anti	∩	: acquitter, blanchir, disculper, innocenter, gracier
Professionnel-S ₁		: juge ; comité de discipline

S ₁	: accusé, défendeur
S ₂	: amende, peine, punition
S ₃	: crime, délit; infraction
S ₄	: jugement, verdict
S _{instr}	: condamnation
S _{res}	: sévèrement
Magn	: avec raison
Ver	: à tort, injustement; par erreur
AntiVer	
W étant	
une exécution, S ₂ Perf	: condamné

5. Proposta de compilação das combinatórias lexicais

A partir do exposto nas páginas anteriores quanto à natureza semântica das combinatórias lexicais e da análise em frasemas da LEC, notadamente os conceitos de *frasema completo* e *semi-frasema*, consideramos justificada a proposta de compilação das combinatórias que preservam um significado, o da palavra-chave, no verbete da palavra-chave. Assim, *conversa fiada* que é um elemento de um conjunto de elementos (*conversa séria, conversa sem pé nem cabeça, conversa pra boi dormir, conversa de embalsar cachorro*, etc.) que representa um *tipo de conversa*, deve ser repertoriada no verbete CONVERSA. Do mesmo modo, *banho-de-cheiro* deve ser repertoriado no verbete BANHO junto das outras combinatórias que representam um *tipo de banho* (*banho de ervas, banho de folhas, banho cheiroso, banho turco*, etc.)

Consideramos igualmente justificada a proposta de compilação dos frasemas completos num verbete autônomo. Assim, *brincar com fogo, pedra no sapato* e *dar um baile* cujas lexias não preservam nenhum significado das lexias que compõem a combinação, devem ter um verbete próprio.

6. Dicionários eletrônicos [razão tecnológica]

Entendemos que a compilação sugerida dos frasemas completos em formato de papel tornaria ainda mais volumosos os dicionários de língua. No entanto, nos dias de hoje, a capacidade de armazenamento de dados em formatos eletrônicos ou seja em CD-roms, torna a proposta de compilação sugerida viável. Assim, graças à informática, a lexicografia “digital” pode aplicar adequadamente as propostas da lexicologia. Além disso, cada *frasema completo* sendo acompanhado de uma definição analítica, otimiza o tempo de aquisição do significado por parte do consulente.

7. Conclusões

Apresentamos neste trabalho diversos problemas de compilação relacionados às combinatórias lexicais encontrados no *Aurélio Eletrônico Século XXI*. Entendemos que esse dicionário não seja o único a apresentar esses problemas. A compilação inadequada das combinatórias é resultado das diferenças semânticas que elas apresentam. Essas dife-

renças são hoje cada vez mais esclarecidas pela lexicologia. Neste trabalho, utilizamos a LEC, *Lexicologie explicative et combinatoire* da TST, *Théorie Sens-Texte* que dá as ferramentas adequadas para a compreensão dos diversos fenômenos semânticos que permeiam as combinatórias lexicais. Após a apresentação dos conceitos de *frasema completo*, *semi-frasema* e *quasi-frasema*, propusemos a compilação das combinatórias lexicais apresentada no DEC, *Dictionnaire explicatif et combinatoire* que se justifica tanto do ponto de vista lexicológico quanto prático, este último tanto mais nos dias atuais em que os recursos informáticos otimizam a capacidade de armazenamento de dados, bem como os meios de acesso à informação contida nos CD-roms.

Referências

- Alonso Ramos, M. E S. Mantha (1996) Description lexicographique des collocatifs dans un Dictionnaire explicatif et combinatoire : articles de dictionnaire autonomes?. *Actualité scientifique*. In Clas, A., P. Thoiron & H. Béjoin (dirs.) Montréal, Aupelf-Uref, pp. 233-253.
- Benson, M., E. Benson & R. Ilson (1986) *The BBI Combinatory Dictionary of English. A Guide to word Combinations*. Amsterdam/ Philadelphia: John Benjamins.
- Cunha, A.G. (1982) *Dicionário etimológico*. Rio de Janeiro: Nova Fronteira.
- Ferreira, A. B. H. (1986) *Novo dicionário da língua portuguesa*. 2a ed., Rio de Janeiro: Nova Fronteira.
- Ferreira, A. B. H. (1999) *Novo Aurélio. Século XXI. O Dicionário da Língua Portuguesa*. Dicionário Eletrônico. Rio de Janeiro: Editora Nova Fronteira & Lexikon Informática.
- Hausmann, F. J. (1979) Un dictionnaire de collocations est-il possible? *Travaux de linguistique et de littérature*, 17(1), pp. 118-129.
- Mel'čuk, I. (1981) Meaning-Text Models : A Recent Trend in Soviet Linguistics. *Annual Review of Anthropology*, 10, pp. 27-62.
- Mel'čuk, I. (1993) *Cours de morphologie générale*, vol. 1.4. Montréal/Paris: Presses de l'Université de Montréal/Éditions du C.N.R.S. vol. 1, 412 pp.
- Mel'čuk, I. (1984) Un nouveau type de dictionnaire: Le dictionnaire explicatif et combinatoire du français contemporain. *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques*, vol. I, Montréal: Les Presses de l'Université de Montréal.
- Mel'čuk, I. (1987) From Meaning to Text : Semantic Representation in the Meaning-Text Linguistic Theory and a New Type of Monolingual Dictionary. *Work Papers [The Summer Institute of Linguistics]*, 31, Grand Forks, ND : University of North Dakota, pp. 73-125.
- Mel'čuk, I. (1997) *Vers une linguistique Sens-Texte*. Paris: Collège de France.
- Mel'čuk, I. et al. (1984, 1988, 1992, 1999) *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques*. vol. I, II, III, IV. Montréal: Les Presses de l'Université de Montréal.

- Mel'čuk, I., A. Clas & A. Polguère (1995) *Introduction à la lexicologie explicative et combinatoire*. Coll. Champs linguistiques/Universités francophones, Louvain-la-Neuve/Paris, Éditions Duculot/AUPELF-UREF.
- Mel'čuk, I. & A. Polguère (1987) A Formal Lexicon in the Meaning-Text Theory (or How to Do Lexica with Words). *Computational linguistics* (Special Issue on the Lexicon), vol. 13, n. 3-4, pp. 261-275.
- Polguère, A. (2003) *Lexicologie et sémantique lexicale, notions fondamentales*. Les Presses de l'Université de Montréal.
- Valente, R (2000) Diferenças e Similaridades Colocacionais entre o Português Brasileiro e o Português Europeu. Estudo baseado na Noção de Função Lexical da Teoria Sentido-Texto. In *Atas do IV Congresso de Lingüística e Filologia da Universidade do Estado do Rio de Janeiro*. Rio de Janeiro, de 27 de setembro a 01 de agosto.