

THESAURUS, DICIONÁRIO ANALÓGICO E DICIONÁRIO IDEOLÓGICO*

Michelle Machado de Oliveira (PPGL/UnB)

RESUMO

O objetivo deste artigo é analisar, com auxílio do método comparativo-descritivo, a estrutura dos: i) *Thesaurus of English Words and Phrases, de Roget* ii) *Diccionario ideológico de la lengua espanola, de Casares*, iii) *Dicionário analógico da língua portuguesa, de Azevedo*; a fim de verificar semelhanças e diferenças ocorrentes. Constatou-se que Azevedo (1974) segue o modelo do Thesaurus de Roget (1852); enquanto Casares (1959) apresenta um outro plano de classificação das idéias. Na elaboração desses tipos de obras onomasiológicas, é primordial destacar o papel que o conceito deve ocupar ao estabelecer as relações entre os itens lexicais.

PALAVRAS-CHAVE: Thesaurus. Dicionários analógico. Ideológico

1. INTRODUÇÃO

Os termos *thesaurus*, dicionário ideológico e dicionário analógico muitas vezes são utilizados para denominar um mesmo tipo de repertório lexicográfico. Há, portanto, uma diferença entre o objeto denominado e objeto definido (BABIBI, 2006, p. 40). Como essas obras são de caráter onomasiológico, ou seja, o ponto de partida do elaborador é o conceito (significado) para chegar à denominação ou às unidades linguísticas (significante), percebeu-se a necessidade de verificar semelhanças e diferenças ocorrentes nelas. Além disso, a influência da estrutura do *thesaurus* foi identificada na elaboração dos dicionários analógicos e ideológicos.

2. CONCEITO E FUNCIONALIDADE DE TESAURO

Um dos objetos de análise deste artigo é o *Thesaurus of the english words and phrases* de Roget¹, publicado em 1852, uma grande referência que serve de base para elaboração de dicionário analógico e de outros tipos de tesouros. O autor explica que, nesse *Thesaurus*, as palavras e as frases da Língua Inglesa não foram classificadas de acordo com os sons da ortografia e não estão em ordem alfabética, mas foram organizadas de acordo com as ideias que expressavam. O objetivo da obra consiste em permitir que se localizem as palavras pelas quais as ideias pudessem ser mais bem expressas em textos.

A partir de 1940, o termo *Tesouro* começou a ser utilizado na esfera da Ciência da Informação, em especial no processo de recuperação da informação, como sendo um instrumento capaz de transportar conceitos e suas relações mútuas, tal como expressos na linguagem dos documentos com controle de sinônimos e estruturas sintáticas simplificadas (DODOBEI, 2002, p. 66).

Segundo Cavalcanti (1978, p. 27), *Tesouro* pode ser definido como “uma lista estruturada de termos associados empregada por analistas de informação e indexadores, para descrever um documento com a desejada especificidade, em nível de entrada, e para permitir aos pesquisadores a recuperação da informação que procura.” *Thesaurus* não é um dicionário, mas sim um instrumento com itens lexicais recolhidos com base em uma estrutura conceitual, a fim de contribuir com a indexação e com a recuperação de informações do domínio de conhecimento, facilitando a busca de conhecimento.

3. COTEJO ENTRE DICIONÁRIO IDEOLÓGICO, DICIONÁRIO ANALÓGICO E TESAURO

Os dicionários ideológicos nasceram na segunda metade do século XIX depois que o *Thesaurus* de Roget foi publicado. Os dicionários analógicos nasceram mais ou menos no mesmo período, mas desta vez na França, em 1862, com Prudence Boissière, autor do *Dictionnaire analogique de la langue française: répertoire complet des mots par les idées, des idées par les mots*.

Geralmente, a macroestrutura desses três repertórios lingüísticos em análise é composta de partes características. Segundo Babini (2006, p. 40), o plano de classificação das

¹ Não foi possível ter acesso à edição original, publicada em 1852. Nesta pesquisa, a edição de 1962 foi utilizada, já que foi a versão que possui a estrutura mais próxima a original em comparação com outras consultadas.

ideias contém as ideias principais da obra e corresponde ao sistema conceitual, depois os tópicos contidos nesse plano são desenvolvidos em quadros sinópticos que contém classes de ideias.

A outra parte é formada por grupos de palavras ligadas por critérios, há cada uma das ideias contidas nos quadros sinópticos. Em cada grupo de palavras são indicados sinônimos da palavra-entrada e seus antônimos, possibilitando buscas por meio desses elementos. O principal critério adotado para a reunião das palavras em grupos é a analogia.

A última parte contém, listadas, em ordem alfabética, as palavras contidas no dicionário. Esse componente estrutural do dicionário varia de uma obra para outra e pode conter, ou não, definições. De fato, em alguns desses dicionários se caracterizam como um índice alfabético remissivo que orienta o leitor sobre onde encontrar a palavra que deseja; em outros, além dessas informações, são apresentadas definições das diferentes acepções que a unidade lexical em questão possui na língua.

De acordo com Babini (2006, p. 40), “*as diferenças entre dicionários analógicos e ideológicos não são muito grandes: os analógicos não contém, em geral, nem um plano de classificação das idéias, nem os quadros sinópticos.*” Entretanto, ao realizar a análise das obras em questão foi detectado que, na prática, o do Azevedo possui essas estruturas mencionadas, o que contradiz a afirmativa de Babini.

O método comparativo-descritivo foi utilizado para constatar as semelhanças e diferenças entre: *Thesaurus of the English words and phrases*; ii) *Dicionário analógico da Língua Portuguesa*; iii) *Diccionario ideológico de la lengua espanola*. A tabela a seguir mostra, de forma sintética, a estrutura da organização macroestrutural das obras mencionadas:

<i>Thesaurus of the English words and phrases</i>	<i>Dicionário analógico da Língua Portuguesa</i>	<i>Diccionario ideológico de la lengua espanola</i>
Plano de classificação	Classificação das palavras	Plano geral de classificação ideológica
Categorias em ordem numeral	Quadro sinóptico de categorias	Texto
Texto	Texto	Parte alfabética (definição)
Índice remissivo	Índice remissivo	

Tabela 1: Estrutura macroestrutural dos dicionários onomasiológicos em análise

Embora Roget tenha considerado que o thesaurus tenha apenas duas partes: o texto e o índice, neste artigo, considerou-se a divisão em quatro partes, por causa da estrutura inicial (plano de classificação e as categorias em ordem numeral), a qual é primordial para a compreensão da maneira como as idéias foram organizadas.

Roget ressaltou “o índice serve para encontrar o lugar que você quer dentro do texto”. No índice remissivo, quando se pesquisa a idéia desejada, encontram-se as palavras relacionadas a essa idéia, as quais têm um número que remeterão ao verbetes numerados em que se pode achar tal palavra. A última parte do Thesaurus é formada por esse índice, o qual possui as palavras agrupadas em ordem alfabética. À frente de cada palavra, tem-se o número da classe em que o item lexical pode ser encontrado, conforme a distribuição feita na classificação das categorias.

No sistema de classificação, as palavras foram organizadas em seis categorias conceituais: “relações abstratas”, “espaço”, “matéria”, “intelecto”, “volição”, “afeições”, as quais são subdividas dentro de seções. Essas categorias foram subdividas em classes, que, por sua vez, em seções, subseções e conceitos isolados. Essas subseções possuem números e por meio desses que as palavras podem ser encontradas.

Dentro dos verbetes, os itens lexicais são organizados de acordo com as categorias gramaticais nas três obras e seguem a seguinte ordem: substantivos, verbos, adjetivos, advérbios, preposições, conjunções e interjeições. No entanto, o dicionário ideológico não realiza tal categorização de maneira tal sistemática e explícita.

Após realizar a comparação entre a obra de Roget e de Azevedo, foi constatado que este copiou o plano de classificação elaborado por aquele, apenas fez a tradução para o Português. Para ilustrar a forma como foi estruturado esse plano, observe o quadro a seguir:

CLASSIFICAÇÃO DAS PALAVRAS		
<i>Classes</i>	<i>Seções</i>	<i>Números</i>
I. RELAÇÕES ABSTRATAS	I. Existência.....	1— 8
	II. Relação.....	9— 24
	III. Quantidade.....	25— 57
	IV. Ordem.....	58— 83
	V. Número.....	84—105
	VI. Tempo.....	106—139
	VII. Mudança.....	140—152
	VIII. Causa.....	153—179
II. ESPAÇO.....	I. Em geral.....	180—191
	II. Dimensões.....	192—239
	III. Forma.....	240—263
	IV. Movimento.....	264—315
III. MATÉRIA.....	I. Em geral.....	316—320
	II. Inorgânica.....	321—356a
	III. Orgânica.....	357—449
IV. INTELECTO.....	I. Formação das idéias.....	450—515
	II. Comunicação das idéias ..	516—599
V. VONTADE.....	I. Individual.....	600—736
	II. Com referência à sociedade	737—819
VI. AFEIÇÕES.....	I. Em geral.....	820—826
	II. Pessoais.....	827—887
	III. Simpáticas.....	888—921
	IV. Morais.....	922—975
	V. Religiosas.....	976—1000

Figura 1: Classificação das Palavras de Azevedo (1974)

A idéia *esporte* foi selecionada com intuito de investigar como aparece nas três obras em análise, assim será possível verificar o tratamento dado à palavra e a maneira de organização dos verbetes em cada uma delas. Ao consultar a parte do texto do *Thesaurus*, não foi encontrada tal idéia, mas ao ir ao índice remissivo, foram observadas as seguintes informações:

Sport, amusement, 840
 gaiety, 836
 wit, 842
 enjoyment, 827
 contentment, 720

Figura 2: Verbetes *sport* no índice remissivo de Roget (1962)

Enquanto no *Thesaurus* de Roget o plano é numerado até 1000, no *Diccionario ideológico de la lengua española*, o plano geral de classificação ideológica é composto por um conjunto de associações de idéias, numeradas de 1 a 38, primeiramente, de maneira

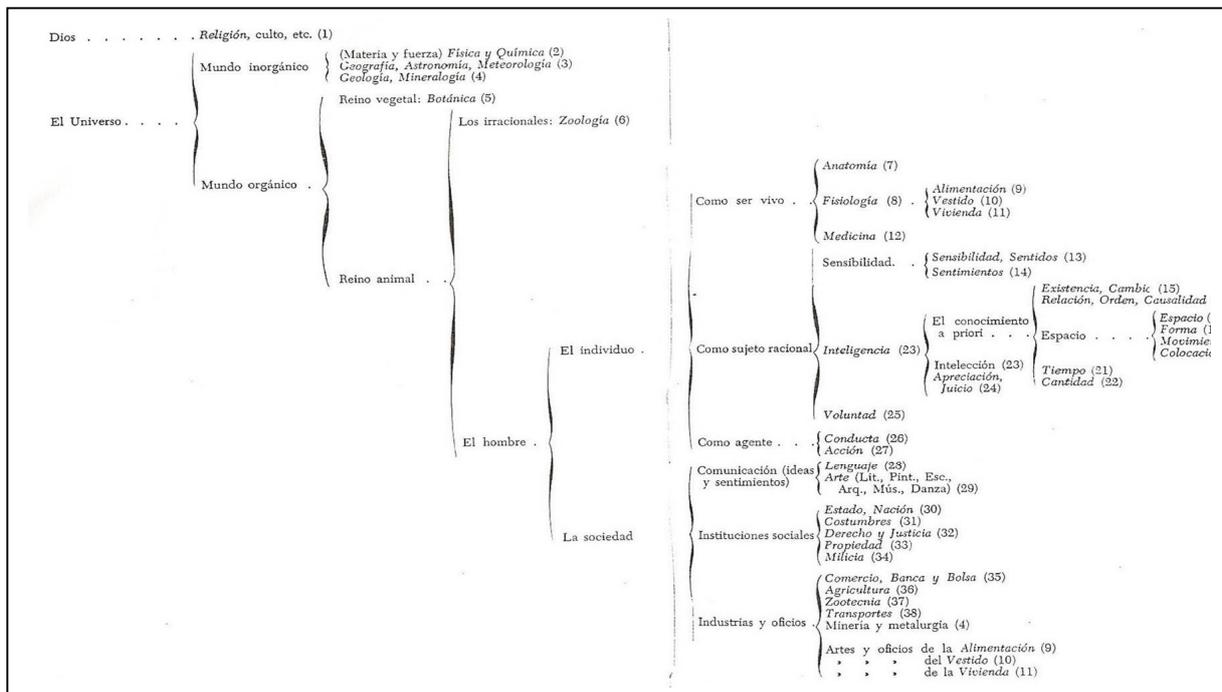


Figura 3: Plano geral de classificação ideológica de Casares (1959)

Ao consultar o prefácio do dicionário de Casares, há o esclarecimento de que os precursores imediatos da classificação ideológica do léxico foram Roget na Inglaterra (1852) e Boissière na França (1862). Por meio dos dados analisados, é possível dizer que a obra de Roget não influenciou esse dicionário ideológico intensamente como ocorreu no dicionário de Azevedo.

No dicionário ideológico, o item lexical equivalente a *esporte* é *deporte*, o qual está inserido na subcategoria *costumbres* (31). Os verbetes subsequentes mostram como tal palavra aparece na parte analógica e na parte alfabética, respectivamente:

DEPORTE (31)

Deporte, Sport, deportismo, profesionalismo, diversión, fiesta, record, marca, entrenamiento, competición, campeonato, fina, semifina, cuartos de fina, eliminatoria, olimpiada, pentatlón, decatlón, copa, carrera, salto, natación, deslizamiento, equitación, rodeo, lucha, atletismo, gimnasia, esgrima, arma, tiro, juego, pelota, vehículo, vectación, gestación, ciclismo, automóvil, navegación, regata, remo, aeronáutica, viaje, turismo, excursionismo, escultismo, alpinismo, montañismo, montaña, patín, caza, cetrería, pesca, campo, estadio, estadiu.

entrenar-se

deportista, deportante, explorador, alpinista, montañero, finalista, campemichón, equipo, deportivo, pedestre.

Figura 4: Verbetes *deporte* na parte analógica de Casares (1959)

deporte m. Recreación, juego, ejercicio físico o diversión al aire libre.

Figura 5: Verbetes *deporte* na parte alfabética de Casares (1959)

A parte organizada em ordem alfabética consiste em um dicionário de língua conciso e com várias acepções de um item lexical. Desse modo, o consulente não terá apenas um índice com uma lista de palavras que talvez conheça vagamente, terá disponível a definição de cada um dos significados que pode ser vista pontualmente. Essa é a explicação dada pelo autor no prefácio. Realmente é uma característica relevante que não tem nas outras duas outras.

No repertório lingüístico de Azevedo, as palavras que remetem a *esporte* não foram encontradas diretamente em nenhuma das partes. Então, ao fazer relação de ideias, o item lexical *divertimento* foi buscado por meio do índice remissivo, o qual se encontra da maneira abaixo para expressar está ideia:

divertimento

drama.....599

diversão..... 840

Figura 6: Verbetes *divertimento* de Azevedo (1974)

Ao ir a seção 840, veja a organização do verbete abaixo:

840. Divertimento “*entretém (pop.), passatempo, diversão, recreação, entretenimento, entretenimento, desenfado, solaz, esparecimento, encanto, feitiço, esporte, desporto, brinco, alegrias, alívio, distração, desenfado, caçoada, graça, gracejo, troça, pilhéria, gargalhada, 838, jocosidade, bufonaria, truanice, arlequinada, funambulismo, espírito 842; zombaria, chocarrice, cahlaça, travessura, filistria = brincadeira, maganeira, maganice, rapaziada, estroinice, escapadela, gazelo, diabrura, traquinada, traquinice, loucura, farra, bandarria, farrancho, noitada, festarola, folia, recreio, pândega, festim, função, bródio, comezaina, dicongo, banzé, banbochata.*

Festa rijia, rasgada, forrobodó, pagode, galhofaria, serão, sarau, serração da velha, trebelho, dança, pastorela, bolero, regadinho, pavana, chico-da-ronda (bras.), fandango, ... Fox-trot, one-step, cake-walk.

(dança dos negros): batuque, xiba (bras.), caxanbu, chica, Cumbe quinbetes, sarambeque, bendenguê, tanglomanglo, etc;

Festival, repasto = banquete, convívio, festança, festa, festejo, ...

Lugar de diversão: teatro, cinema, arena, estádio, jardim, passeio público, logradouro, circo, parque, hipódromo...

Figura 7: Verbete divertimento de Azevedo (1974)

Considerações Finais

É primordial considerar a afirmativa, a seguir, de Campos (2006: 350): “conceito é o ponto de partida para estabelecer as relações conceituais”, pois estas são as bases das obras de caráter onomasiológico. Se essas relações não forem feitas corretamente, o resultado obtido será apenas meras listas de palavras agrupadas por associações aleatórias sem critérios coerentes.

Nos tesouros, dicionários analógicos e ideológicos não se podem cometer o erro de ampliar tanto o significado de uma palavra para que ela não deixe de ter utilidade. Os lexicógrafos e os elaboradores de thesaurus antes de categorizarem as palavras e montarem as relações devem fazer a reflexão: dentre todos possíveis significados da palavra x, o mais razoável e mais útil é atribuir-lhe o significado y; deste modo, poderemos usar a palavra x com rendimento máximo (WILSON, 2005: 36-37). Assim as associações arbitrárias entre as palavras são eliminadas e a análise de conceito surge como uma ferramenta necessária por ser um modo de gerar conscientização e nos ensina a evitar armadilhas da linguagem (WILSON, 2005: 36-37, 39).

O dicionário analógico de Azevedo segue o modelo do *Thesaurus of English Words and Phrases* de Roget, já o dicionário ideológico apresenta um outro plano de classificação das idéias. Ao estabelecer relações analógicas entre os itens lexicais de uma língua, é necessário ter em mente que “significado vai mais fundo que o hábito e o uso, está na base de todo o equipamento conceitual do homem, o qual está enraizado na sua personalidade e nas experiências passadas, assim como Wilson (2005: 137) postulou. Pode-se deduzir que o esquema mental do ser humano possui conceitos que são influenciados pela cultura devido às suas vivências, porém, é necessário ter o cuidado para que o uso habitual de um significado não deturpe o conceito original.

O uso dos tesouros, tanto como instrumento de representação/recuperação de informações quanto de registros lingüísticos precisa de uma inter-relação dos conceitos

emergentes e a conseqüente alteração no significado dos conceitos existentes na área e de suas relações com as outras.

A elaboração de um tesouro implica bases classificatórias e as categorias fundamentais que permitem correto posicionamento dos conceitos nas classes e a organização de domínio; e a definição de cada conceito é o elemento que vai comprovar objetivamente tanto as relações hierárquicas como as partitivas e associativas com outros conceitos. (CAMPOS, 2006, p. 348)

A identificação de conceitos relacionada à sistematização e representação de uma determinada área do conhecimento é de suma importância para os Sistemas de Recuperação da Informação e para os lexicógrafos, pois facilita no processo de busca e seleção da informação ao pesquisador.

THESAURUS, ANALOG DICTIONARY AND IDEOLOGICAL DICTIONARY *

ABSTRACT

The objective of this article is to analyze, using the comparative-descriptive method, the constituent structures of i) the *Roget's Thesaurus of English Words and Phrases*, ii) *Casares's Dicionário Ideológico de la lengua espanhola*, and iii) *Azevedo's Dicionário Analógico da lingua portuguesa*, in order to verify similarities and differences. The study found that Azevedo (1974) follows the model proposed in the *Roget's Thesaurus* (1852), while Casares (1959) presents a different pattern of grouping the ideas. In the elaboration of this kind of onomasiological works, it is essential to highlight the role that the concept must perform when establishing relations among lexical items.

KEYWORDS: Thesaurus, dictionary ideological and analogical

REFERÊNCIAS

AZEVEDO, F. F. dos S. *Dicionário analógico da língua portuguesa: idéias afins*. Brasília: Coordenada, 1974.

BABINI, Maurizio. Do conceito à palavra: os dicionários onomasiológicos. *Ciência e Cultura*. São Paulo, v. 58, n. 2, 2006. Disponível em:

<http://cienciaecultura.bvs.br/scielo.php?script=sci_arttext&pid=S0009-67252006000200015&lng=en&nrm=iso>. Acesso em: 21/05/2008.

- CAMPOS, M. L. A.; GOMES, H. E.. Methodology for construction of conceptual thesaurus: categorization as a theoretical principle. *Perspectivas em Ciência da Informação*. Belo Horizonte, v. 11, n. 3, 2006. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-99362006000300005&lng=en&nrm=iso>. Acesso em: 17/06/2008.
- CASARES, J. *Diccionario ideológico de la lengua espanola*. Barcelona: G Gili, 1959.
- CAVALCANTI, C. R. *Indexação e tesouro: metodologia e técnica*, Brasília: ABDF, 1978.
- DODEBEI, Vera Lucia Doyle Louzada de Mattos. *Tesouro: linguagem de representação da memória documentária*. Niterói: Intertexto, 2002.
- ROGET, Peter Mark; BROWNING, David Clayton. *Everyman's thesaurus of english words and phrases*. London: J M Dent, 1962.
- WILSON, J. *Pensar com conceitos*. Tradução Waldéa Barcellos. 2ª ed.- São Paulo: Martins Fontes, 2005.