

# Da Percepção à Recuperação da Informação<sup>1</sup>

*From Perception to Information retrieval*

Dr. Fabio José Dantas de Melo<sup>2</sup>

**Resumo:** A *práxis* do profissional da informação deve possibilitar o acesso de todo e qualquer usuário às fontes informacionais de que necessitem. Para isto é indispensável munir este profissional de ferramentas que o permitam explorar o multifacetado universo lingüístico do público-alvo de seu trabalho. Segundo conceitos da psicolingüística, semiótica e lingüística cognitiva, e analisando os processos de representação e recuperação da informação, este artigo se propõe a refletir sobre os desafios da busca de informação com aquilo que o usuário tem a sua disposição: sua língua e a experiência de mundo.

**Palavras-chave:** Lingüística cognitiva, Redes semânticas, Recuperação da informação, Organização do conhecimento.

**Abstract:** The practice of information professionals should allow access of any user to information sources they need. For this it is necessary to equip these professionals the tools to exploit the multifaceted linguistic universe of the target audience for their work. According to concepts of psycholinguistics, cognitive linguistics and semiotics, and analyzing the processes of representation and retrieval of information, this article intends to reflect on the challenges of finding information about what the user has at his disposal: the language and experience world.

**Key Words:** Cognitive linguistics, semantic networks, Information retrieval, Knowledge organization.

<sup>1</sup> Texto publicado simultaneamente em MELO, F. J. D. de ; BRASCHÉR, M. (Orgs.) "Fundamentos da Linguística para a formação do profissional de informação" Brasília: Thesaurus, 2011.

<sup>2</sup> Doutor em Lingüística e PhD em Ciência da Informação. Bolsista de PCI do Instituto Brasileiro de Informação em Ciência e Tecnologia – IBICT . E-mail: [fabiojose@unb.br](mailto:fabiojose@unb.br)

*Um dos segredos da comunicação é ter curiosidade a respeito do mapa do outro. Você vai perceber que, mesmo tendo as mesmas peças do quebra-cabeça, surpreendentemente outra pessoa forma um desenho diferente do seu.*

XAVIER GUIX

*Língua e Cultura são duas modalidades paralelas de uma atividade mais fundamental: (...) o espírito humano.*

CLAUDE LÉVI-STRAUSS

Como afirmam Lastres & Albagli (*apud* Araújo; Rocha 2009, p.11), as sociedades encontram-se na Era da Informação e do Conhecimento, “(...) em que as atividades humanas estão centralmente baseadas e organizadas em torno das atividades de geração, **recuperação** e uso de informações e conhecimentos” (grifo meu).

Minha atenção volta-se, este artigo, para a busca de documentos com emprego de termos da linguagem natural presentes no título ou no corpo do texto daqueles, uma vez que – de modo geral – os usuários de bibliotecas desconhecem a existência de vocabulários controlados e, outras vezes, não trazem a indicação de autores para chegarem a livros dos assuntos pretendidos.

A reflexão proposta é que os documentos armazenados em bases de dados precisam ser indexados levando-se em conta o **universo conceitual** e a **diversidade sociolingüística** dos usuários. “A Lingüística Cognitiva”, um dos referenciais teóricos utilizado nas formulações aqui expostas, “assume que **fatores históricos e socioculturais** [dentre outros] são necessários e fundacionais na caracterização da estrutura lingüística” (SILVA *apud* SILVA & GONÇALVES 2004). Como, por exemplo, as *formas variantes de palavras* que, segundo as investigações de Bourne (*apud* Lopes 2002, p.61), “afetam consideravelmente os resultados das buscas”.

Enfatizam Dias-da-Silva e Di Felippo<sup>3</sup> que, do ponto de vista psicolingüístico, existe um processo mental chamado *especificação de conceitos (conceptualização)*, o qual podemos apontar como o responsável pela realização de uma busca. Este processo se desenrola em três momentos:

1. a identificação do objeto pretendido;
2. a seleção de uma representação sintático-semântica do objeto;
3. a codificação dessa representação em termos fonológicos

---

<sup>3</sup> Disponível em <http://www.gel.org.br/estudoslinguisticos/volumes/32/htm/comunica/ci035.htm> (consultado no dia 10 de agosto de 2010).

Consideremos o seguinte artefato:

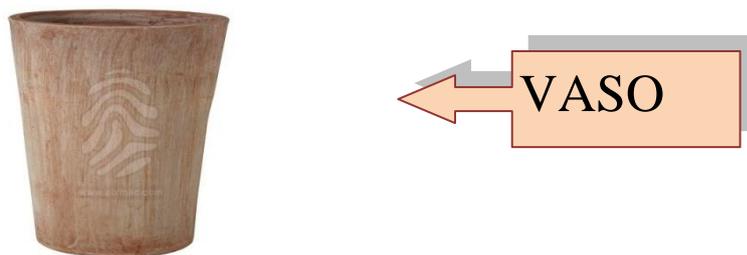


Figura 1 – Referente e seu rótulo lingüístico

Note que tal objeto não causaria muita polêmica quanto a sua identificação como *vaso*. Entretanto, para muitos dos leitores deste artigo, o objeto seguinte traria discordâncias sobre qual rótulo lingüístico que lhe caberia:

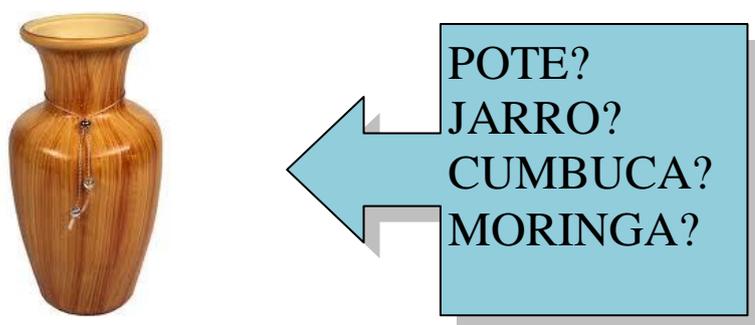


Figura 2 – Referente e suas diferentes denominações regionais

Trata-se de um pote, de um jarro ou de uma cumbuca? Se procurássemos tal objeto num banco de dados de uma instituição situada na região nordeste do país, empregariamos um dos termos anteriores ou *moringa* para recuperá-lo? Afinal, conforme critério funcional, o objeto acima reúne as propriedades de uma *moringa*, que é um regionalismo. Com isso, estamos na segunda fase do processo de conceptualização. Pais (2009), em artigo *Da cognição à semiose*<sup>4</sup>, expõe que

Os seres humanos distinguem-se dos outros animais do planeta, por sua diversidade lingüística, cultural, social e histórica. Cada comunidade lingüística e sociocultural sustenta, assim, nos processos semióticos verbais, não-verbais e sincréticos, uma 'visão do mundo', um *mundo semioticamente construído*. Nessas condições, embora as potencialidades biológicas dos seres humanos sejam as mesmas, as pessoas que nascem numa comunidade são ensinadas a 'pensar o mundo', tal como se constrói na língua, na cultura, na sociedade, nas semióticas não-verbais e sincréticas. Daí resulta que a *percepção biológica é culturalmente filtrada* (PAIS, 2009)

Assim, a conceptualização é parte intrínseca de outro processo natural, denominado *categorização*, que consiste na identificação, classificação e nomeação de diferentes entidades como membros de uma mesma categoria. Tal organização é influenciada em parte pelas

<sup>4</sup> Disponível em <http://www.filologia.org.br/vicnlf/anais/caderno07-09.html> (consultado no dia 14 de agosto de 2010).

experiências do indivíduo em seu ambiente sociocultural de formação e em parte pelas experiências coletivas.

As categorias conceituais diferem de cultura para cultura, como também entre indivíduos de mesma cultura. Porque a rotulagem de uma categoria é, em grande medida, determinada por um dos elementos da relação triádica de signo como formulada por Charles Sanders Peirce:

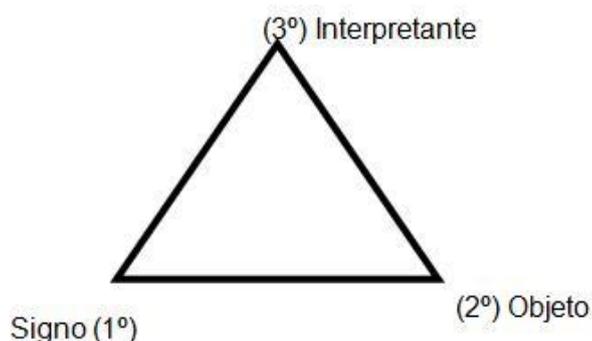


Figura 3 - Representação do signo peirceano

É, de fato, o *interpretante* que está em cena quando se levanta o aspecto da variabilidade de conceituação sobre um mesmo referente (objeto). Segundo Santaella (2005 p. 58-59):

A noção de interpretante não se refere ao intérprete do signo [o indivíduo], mas a um processo relacional que se cria na mente do intérprete. A partir da relação de representação que o signo mantém com seu objeto produz-se na mente interpretadora um outro signo que traduz o significado do primeiro (é o interpretante do primeiro). Portanto, o significado de um signo é outro signo (...) criado na mente por aquele (SANTAELLA, 2005, p.58-59)

Na semântica de um item lexical, declara Silva (1997), há **sentidos prototípicos e periféricos**. Tanto que a sua estrutura tem a forma de uma *network* (rede) <sup>5</sup>, na qual as margens do campo são ocupadas pelos sentidos mais específicos (de contextos mais particulares) e o núcleo, pelo sentido mais geral. Por isso, a pertinência – no exercício de síntese documentária – da noção de *protótipo*, provinda da Linguística Cognitiva. **Protótipos** são os exemplares mais típicos, mais representativos, de uma categoria.

<sup>5</sup> Do artigo *A Linguística Cognitiva: Uma breve introdução a um novo paradigma em Linguística* (1997) disponível no [www.facfil.ucp.pt/lingcognit.htm](http://www.facfil.ucp.pt/lingcognit.htm) (acessado em 16 de agosto de 2010).



Amoretti (2001) afirma que

A representação do conhecimento em rede facilita, pois, a apreensão do conhecimento porque a memória humana reconhece e retém mais rapidamente os exemplares prototípicos, respondendo de maneira mais satisfatória às expectativas de realidade dos leitores, facilitando o processo mental da compreensão. A rede simula aspectos típicos da cognição humana, tendo como característica essencial a flexibilidade na modelagem de fenômenos cognitivos que é a capacidade da rede de sempre completar os conceitos descritos através da associação de novas propriedades aos conceitos básicos (AMORETTI, 2001, p. 50)

A rede, enquanto simulação da cognição humana, pode ser vista como arquitetura do Léxico no ambiente mental dos falantes. E sua configuração se adéqua com as instâncias que a Lingüística Estrutural postulou, no século passado, para estruturação do Léxico, que são os *campos lexicais* e os *campos semânticos*. Para a compreensão destes dois construtos lingüísticos é preciso recorrer ao de *campos conceituais*. O campo conceitual refere-se a domínio cognitivo, ou seja, uma área da experiência sociocultural coberta por um conjunto de termos. A determinação de um campo conceitual é função de uma cosmovisão.

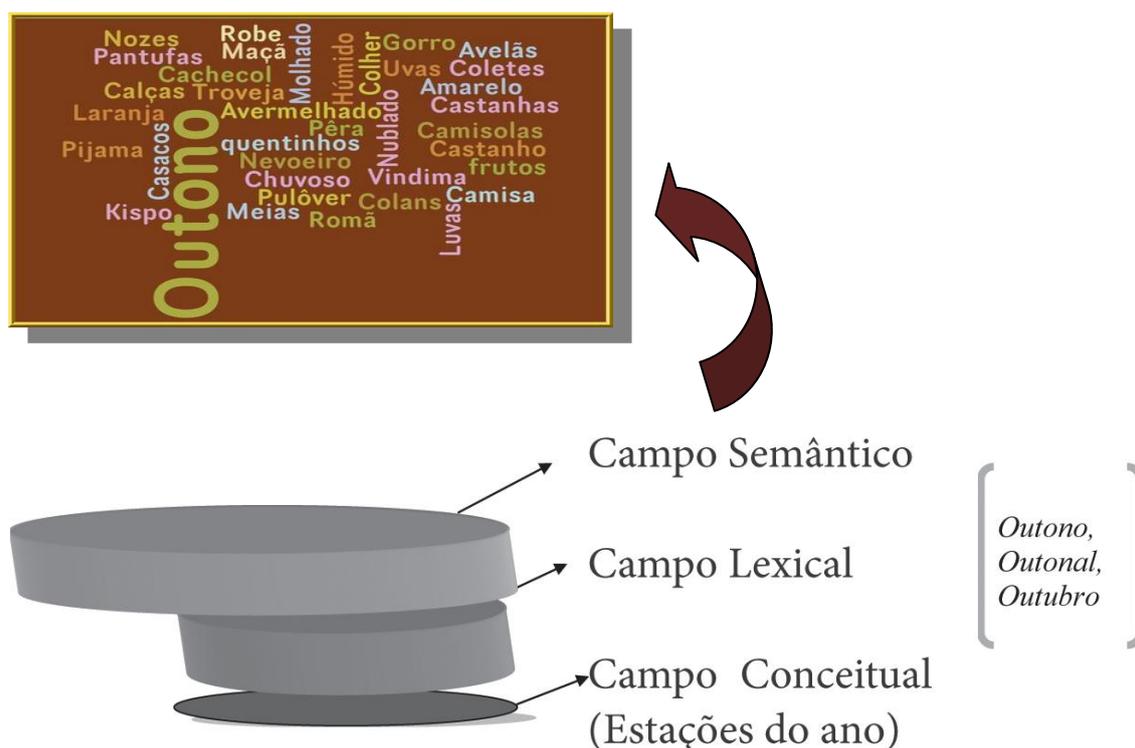
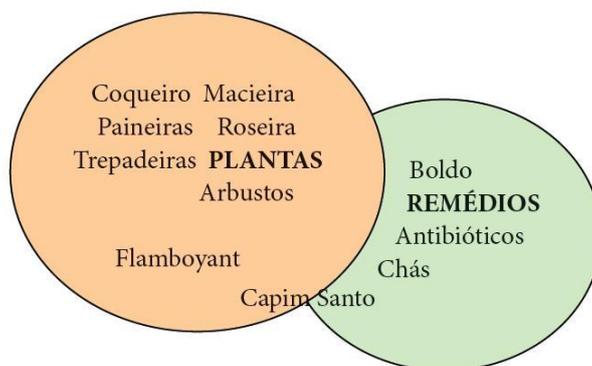


Figura 5 – Campo Semântico de *Outono*, campo lexical e campo conceitual  
(Fonte: <http://img511.imageshack.us>)

Vimos que os membros mais representativos, isto é, aqueles que os falantes primeiro evocam ao escutar ou ver o nome de uma categoria, são os *membros prototípicos* ao redor dos quais os demais (*membros periféricos*) se agrupam. De maneira que os membros mais afastados do centro (*locus* do protótipo) podem fazer parte de outras categorias como mostra o diagrama seguinte:



**Figura 6 – Interpenetração de dois campos lexicais, possibilitando a visualização do elemento periférico (*capim santo*) e dos elementos prototípicos (*antibióticos, arbustos etc.*)**

O esforço da síntese documentária está, assim, na seleção dos conceitos centrais do universo do objeto informacional em análise em detrimento de conceitos marginais, cuja pertinência às categorias do referido universo é duvidosa. E mais: é preciso sensibilidade para empregar descritores (os rótulos lingüísticos) que se ajustem aos modelos cognitivos culturais dos usuários. Para isso, a confecção de mapas conceituais pode se dar previamente à análise documentária na forma de trabalho de campo à semelhança do que fazem os antropólogos para alcançar a *weltanschauung* (visão de mundo) “nativa”. Estes mapas reconstróem uma visão interpretativa da realidade, a qual pode ser capturada nas práticas discursivas dos membros do grupo pesquisado. Como explica Duque:

os dados necessários para desenhar o **mapa de uma categoria** se obtêm do falante, como resultado de diversas **tarefas experimentais** e rejeitam as taxonomias científicas, porque não constituem uma representação adequada do modo como as pessoas organizam e entendem a realidade.

Na formação antropológica é de suma importância a experiência e a elaboração da *alteridade*, uma revolução no “olhar” que nos leva a encarar aquilo que nem se consegue imaginar devido à dificuldade em prestar atenção ao que é habitual, familiar, cotidiano e considerado evidente. Considerar este componente na educação do profissional da informação é despertá-lo para a realidade de que os seres humanos têm em comum a **capacidade de elaborar modos de conhecimento muito diversos**; é fazê-lo captar a experiência do Outro e transpor isto para sua prática de preservação e disseminação de

objetos informacionais, capacitando-o a lidar com uma variável que escapa ao objetivismo das taxonomias científicas, como bem colocou Duque. Como em recente artigo acerca do trabalho de organização de acervos em Casas de Candomblé, em Salvador (BA), Oliveira (2010) assim conclui: “(...) como profissionais da informação [...] precisamos inserir em nossos discursos, em nossas listas de discussão e em nossos eventos, temas tais como **saberes das comunidades tradicionais, etnoconhecimentos (...)**” (grifos meus) (Oliveira, 2010, p. 90).

### **Referências bibliográficas**

AMORETTI, Maria Suzana Marc. Protótipos e estereótipos: aprendizagem de conceitos Mapas Conceituais: experiência em Educação a Distância. **Revista Informática na Educação: Teoria & Prática**, v.4 n.2, PPGIE/UFRGS, dezembro 2001.

ARAÚJO, Evandro Nicomedes; ROCHA, Elisa Maria Pinto da. Trajetória da sociedade da informação no Brasil: proposta de mensuração por meio de um indicador sintético. **Revista Ciência da Informação**, Brasília, v. 38, n. 3, p. 9-20, set/dez., 2009.

DUQUE, Paulo Henrique. Teoria dos Protótipos, Categoria e Sentido Lexical. **Revista Philologu**, Ano 7, n. 21, s.d. Disponível em <http://gelvarj.sites.uol.com.br/prototipicidade.html>. Acessado em 13 de agosto de 2010.

LOPES, Ilza Leite. Estratégia de busca na recuperação da informação: revisão da literatura. **Revista Ciência da Informação**, Brasília, v. 31, n. 2, p. 60-71, maio/ago. 2002.

OLIVEIRA, Andréa Carvalho. Direito à Memória das comunidades tradicionais: organização de acervo nos terreiros de candomblé de Salvador, Bahia. **Revista Ciência da Informação**, Brasília, DF, v. 39 n. 2, p.84-91, maio/ago., 2010

SANTAELLA, Lúcia. **O que é semiótica**. São Paulo: Brasiliense, 2005.

SILVA, Augusto Soares. Linguagem, Cultura e Cognição, ou a Lingüística Cognitiva. *In*: SILVA, Amadeus Torres & GONÇALVES, Miguel (org.) **Linguagem, Cultura e Cognição**: Estudos de Linguística Cognitiva. Coimbra: Almedina, 2004. (vol. 1)