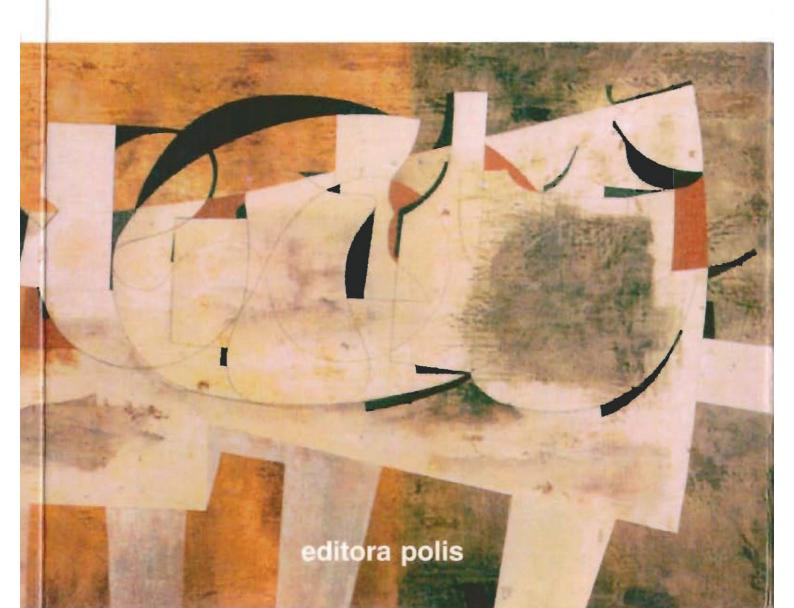
Anna Maria Marques Cintra
Maria de Fátima Gonçalves Moreira Tálamo
Marilda Lopes Ginez de Lara
Nair Yumiko Kobashi

## Para entender as linguagens documentárias



Esta obra apresenta os aspectos fundamentais das linguagens documentárias: sua natureza, sua estrutura e suas funções. Aborda, de maneira especial, as questões lingüísticas e lógicas que fundamentam sua elaboração. Incorpora, nesta segunda edição, um novo item — a Introdução — em que se ressaltam o caráter sistêmico das linguagens documentárias e seus aspectos culturais. Além disso, foi inteiramente revista, de modo a eliminar imperfeições.

Cada vez mais as linguagens documentárias vêem se mostrando como importantes ferramentas de organização e distribuição de informação. Elas já são hoje consideradas imprescindíveis para agregar valor à informação especializada, na medida em que, por seu intermédio, a tarefa de organizar tematicamente a informação torna-se mais consistente.

Mas além de seu caráter organizacional, as linguagens documentárias viabilizam o compartilhamento de informações produzidas por diferentes instituições. Decorre daí o fato de os sistemas ou redes cooperativos de informação não prescindirem de algum tipo de vocabulário controlado para a constituição de seus dispositivos informacionais, sejam eles de uso local ou disponibilizados para públicos amplos, através de redes eletrônicas.

A utilização crescente de linguagens documentárias baseia-se na evidência de que, sem uma linguagem compartilhada, não é possível a comunicação entre serviços de informação e seus usuários. Em vista disso, as linguagens documentárias se oferecem como instrumentos poderosos de socialização da informação. E a informação, na medida em que está diretamente relacionada ao conhecimento, tem papel decisivo na mudança dos destinos da humanidade. Espera-se que sua leitura contribua para aprimorar a formação daqueles que já atuam ou atuarão profissionalmente no vasto campo da organização e transferência da informação.

#### Anna Maria Marques Cintra Maria de Fátima Gonçalves Moreira Tálamo Marilda Lopes Ginez de Lara Nair Yumiko Kobashi

# PARA ENTENDER AS LINGUAGENS DOCUMENTÁRIAS

2ª edição revista e ampliada

editora polis 2002 Copyright © 2002 das autoras

Capa: Ivone Quetzal

Ilustração da capa: Ben Nicholson, Agosto de 1956 (Vaid'Orcia).

Revisão: Nair Yumiko Kobashi e Marcos Frederico

Editoração eletrônica: Editora Polis

#### Ficha catalográfica

CINTRA, Ana Maria et al. Para entender as linguagens documentárias. 2ed rev. e ampl.. – São Paulo I Polis, 2002.

96 p. (Coleção Palavra-Chave, 4)

ISBN 85-7228-0012-X

1. Linguagens documentárias. I. Título.

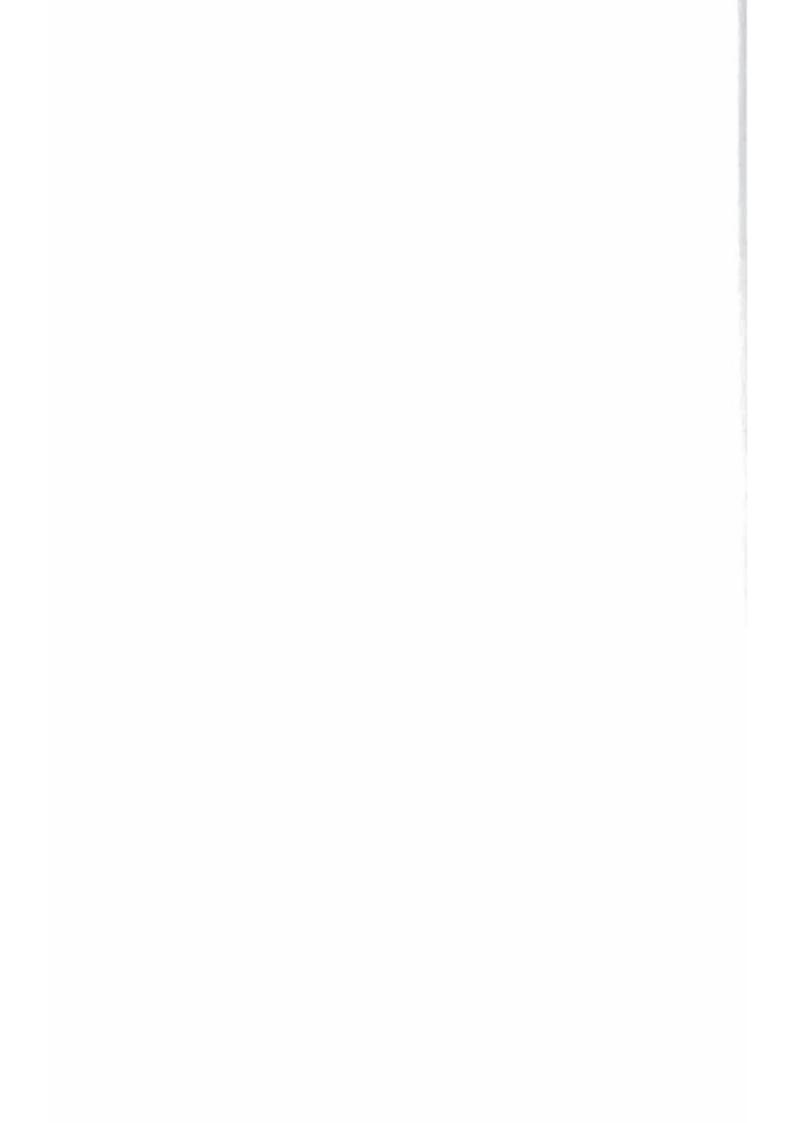
CDD = 025.4 CDU = 025.4

Direitos reservados pela EDITORA POLIS LTDA.

Rua Caramuru, 1196 – Saúde – 04138-002 – São Paulo, SP Tel.: (11)5594-7687 e (11)275-7586 – Fax: (11)275-7586

e-mail: polis@cditorapolis.com.br

O sentido não é nunca uma evidência. Edward Lopes



#### Sumário

Apresentação 9
Introdução 13
<ol> <li>Conhecimento, informação e linguagem</li> <li>Conhecimento e informação</li> <li>Linguagem: características gerais</li> <li>26</li> </ol>
<ul> <li>2. Linguagens documentárias 33</li> <li>2.1 Natureza, especificidade e funções 33</li> <li>2.2 Configuração das linguagens documentárias 42</li> </ul>
3. Sistema nocional 49 3.1 Relações hierárquicas 55 3.1.1 Relação genérica 58 3.1.2 Relação partitiva 61 3.2 Relações não-hierárquicas ou sequenciais 62
<ul> <li>4. Relações lingüísticas e documentação 67</li> <li>4.1 Polissemia e ambigüidade 70</li> <li>4.2 Sinonímia 74</li> <li>4.3 Hiponímia 77</li> </ul>
Bibliografia 87

#### Apresentação

O homem vive entre os campos semiológicos. No seu cotidiano, caminha de um para outro, consciente ou inconscientemente. Enreda-se nas construções arquitetônicas, passa por esculturas (às vezes sem ver), ouve "sons" perdidos de músicas ou de gritos, restos de conversas – agressivas ou carinhosas. Cada um desses campos pelos quais ele transita diariamente tem seu código específico. E ele *trans-ita* no sentido primeiro: vai *através* de (trans) um caminho (ito) que as gerações passadas construíram para ele e que sua própria condição de *humano* lhe permite "receber" de vários modos: em um dos pólos, não "percebendo" a extensão do mundo em que vive; no outro pólo, "percebendo" tal extensão, apropriando-se dele e modificando-o, construindo novo mundo, novos mundos. Re-construindo-se no fazer.

Esses campos semiológicos, com seus códigos próprios, muitos deles não-verbais, entrelaçam-se e manifestam, na verdade, a condição da sociedade naquele momento histórico. Essa inter-relação entre os campos, essa "costura" é realizada pelo código verbal, pelo signo verbal, pela palavra. Ou seja: os campos semiológicos são manifestações sóciocuturais de uma dada sociedade. Embora com suas especiatores de uma dada sociedade.

cificidades, eles revelam a cultura daquela sociedade, naquela etapa de desenvolvimento. E a cultura é transmitida, predominantemente, pela palavra. Por isso, só a palavra tem a condição de penetrar todos eles, de "interpenetrá-los".

Esse é um dos motivos porque se afirma a importância da palavra. Sua condição de plasticidade permite-lhe ser o suporte do conhecimento. Sem conhecimento, o homem permanece sempre muito próximo do pólo dos que não percebem a extensão do mundo em que vivem, em que circulam. E aí está um dos aspectos da importância da *informação*.

Nesse sentido, parece que a informação cumpre papel decisivo na mudança dos destinos da humanidade, uma vez que ela está diretamente ligada ao conhecimento e ao desenvolvimento de cada uma das áreas do saber, já que todo conhecimento começa por algum tipo de informação e se constitui em informação (\*...) E para que o conhecimento da sociedade não se perca e possa ser compartilhado, ele é registrado num dado suporte: livro, imagem, foto, disco etc. passando a se constituir num documento.

A informação não é um dado. Ela se constrói no encontro de duas dinâmicas: a dinâmica de quem "emite", de quem "enuncia" (o enunciador) e a dinâmica de quem "recebe" o enunciado (o enunciatário). Ela ocorre sempre num espaço onde as posições de quem "fala" e de quem "ouve" são intercambiadas, num jogo de forças permanente.

Aí começa a linguagem documentária. Como conseguir que o conhecimento acumulado não se perca, que se tenha acesso a ele, de tal modo que não seja necessário "reinventar a roda" a cada geração? A memória coletiva, a transmissão

oral da cultura são aspectos fundamentais, mas como, nas áreas científicas, fazer conhecer o conteúdo de aproximadamente 60.000 revistas científicas e cerca de um milhão de artigos individuais? Esses dados, citados neste livro, revelam a estimativa de 1960. Certamente a produção científica é muito maior.

Ninguém ousaria pensar que é possível conhecer toda a produção de uma determinada área do saber. Mas é necessário, pelo menos, ter acesso a seus avanços e partir deles na construção do novo conhecimento. Eis aí, de novo, a linguagem documentária.

Os desafios são númerosos. Num mundo em que, ao que parece, o homem "não diz", apenas "é dito" pelas palavras; em que se tem, predominantemente, "a voz do dono" e não o homem como "dono da voz"; em que os discursos de máscara circulam como mercadorias de maior valor, como trabalhar a linguagem documentária? Afinal, ela pressupõe, por um lado, a importância da divulgação da informação para que o homem assuma sua própria voz; por outro lado, ela pressupõe o sujeito que vai "passar" o conhecimento científico, elaborado na linguagem polissêmica, para outra – a linguagem documentária.

A palavra carrega a prática social da sociedade, enfeixa os valores de um determinado momento histórico. É sub-reptícia. Atua, sem que tenhamos consciência de seu papel. Este sujeito que vai "traduzir" o texto científico para a linguagem documentária carrega consigo essa formação. Também o sujeito que elaborou o texto científico. Na condição de sujeito, cada um terá seu universo de valores, que lhe foi transmitido pela cultura. Como evitar os desvios nessa tradução?

Como dizem as autoras, de um sistema de relações que se caracteriza pela virtualidade, a LN (Língua Natural), usada pelo sujeito do texto científico e pelo sujeito que fará a "tradução" (e ainda pelo sujeito que "receberá" a *informação*, é bom não esquecer) passa-se para um sistema de relações não-virtual – a LD (Linguagem Documentária)

Mas, diferentemente da LN, o sistema de relações das LNs não é virtual, bem como seus mecanismos de articulação são extremamente precários, em face daqueles existentes nas línguas, em geral. Bem ao contrário, elementos dessa linguagem específica são selecionados de universos determinados e seu sistema de relações é construído, sendo indispensável, para utilizá-la, a existência de regras explícitas. Por esse motivo, as LNs são linguagens construídas.

O mundo contemporâneo se desnuda em sua complexidade: todos os povos lutam para ter vez e voz no concerto das nações. A constituição de pólos hegemônicos consolidase a partir do conhecimento. E a linguagem documentária joga papel decisivo nessa realidade.

O desafio é grande. As palavras, "suspensas no ar", param sua dança. Mas as autoras desse livro, com ciência e competência, topam o desafio. E vencem. É ler para crer.

Maria Aparecida Baccega Professora Livre-docente da Escola de Comunicações e Actes da Universidade de São Paulo

#### Introdução

Se no passado as diferentes formas de nomear o arco-íris entre povos, ou as relações diversas entre línguas para dizer "eu estou com dor de cabeça" eram percebidas como casos particulares ou idiotismos, hoje, são entendidas como manifestações naturais, geradas pelas diferenças de significante e de substância semântica, uma vez que o significado compõe, indissociavelmente, a unidade numa dada língua.

Com efeito, enquanto na nossa cultura distinguimos sete cores no arco-íris, entre os bretãos e os gauleses esse número cai para quatro, por exemplo, na zona onde distinguimos azul e verde, eles identificam apenas o "glas". E, de forma semelhante à nossa expressão "Eu tenho uma dor de cabeça" correspondem em francês ou em italiano outras relações: "J'ai mal à la tête" ou "Mi duole il capo".

Evidentemente, em muitos casos pode-se ter a ilusão de mera transposição de nomes ou troca de significantes, como em "cão/dog/chien". Entretanto, a esses significantes correspondem significados relacionados a toda experiência cultural dos falantes de cada língua, o que levará, irremediavelmente, a significados diferentes, portanto a signos diferentes.

Pode-se, pois, dizer que cada língua natural – LN – analisa os dados da experiência segundo padrões que dependem da tradição cultural e do momento social do povo que a fala. Isso faz com que possamos dizer que cada LN é, a rigor, uma análise da sociedade, do homem participante de um grupo e de sua cultura.

Martinet (1969) diz: "Uma língua é um instrumento de comunicação segundo o qual de modo variável de comunidade para comunidade se analisa a experiência humana em unidades providas de conteúdo semântico e de expressão fônica e em unidades distintivas sucessivas".

Mas isso não permite dizer que numa mesma comunidade as estruturas lingüísticas sejam homogêneas. Facilmente observa-se que as pessoas de uma mesma comunidade lingüística não falam do mesmo modo. Entretanto, desde que essas diferenças sejam tais que não impeçam a comunicação, dizemos que estamos diante de uma mesma língua.

Seguramente, se fizéssemos uma análise dos sons produzidos pelas pessoas de um grupo culturalmente homogênco e de mesma língua, encontraríamos inúmeras diferenças. No entanto, essas diferenças de timbre, de intensidade, de altura etc., não são, freqüentemente, sentidas nem pelo emissor, nem pelo receptor, nem tampouco como impeditivas de comunicação.

A "posição" relativa do signo explicita a noção de valor. Quando dizemos que uma peça de alabastro vale x, dizemos que ela pode ser trocada por um outro objeto de natureza diferente; dinheiro, ouro etc. Portanto seu poder de troca está condicionado a relações fixas existentes entre ele e objetos da

mesma natureza, como outras peças de alabastro, provenientes de outros lugares.

Na LN, o elemento de troca é o signo lingüístico, que associa um significante (imagem acústica) a um significado (conceito). Seu poder de troca está ligado ao fato de poder servir para designar uma realidade lingüística que lhe é estranha (realidade atingida por intermédio de seu significado, mas que não é seu significado). Mas este poder significativo que constitui o signo é estritamente condicionado pelas relações que o unem aos outros signos da língua, de sorte que não se pode escolhê-lo sem o recolocar numa rede de relações intra-lingüísticas.

Com os limites próprios de uma linguagem construída, as linguagens documentárias – LDs – se valem de quase todos os conceitos apresentados para a LN, constituem sistemas onde as unidades se organizam em relações de dependência.

No entanto, não se pode dizer que sejam signos, uma vez que faltam a suas unidades características básicas de signo: significante e significado articulados segundo padrões sócio-culturais com disponibilidades virtuais de significação.

Também não dependem nem da tradição cultural, nem do momento social e sim de convenções estabelecidas no conjunto do próprio sistema que é, por isso, estático e homogêneo.

Não se processa com essas linguagens uma comunicação no sentido estrito. Processa-se, antes, uma decodificação pura e simples, à maneira de códigos estáticos.

Mas, a utilização de unidades retiradas da LN, dá às LDs um caráter particular que as torna, de certa forma, diferentes dos sistemas estáticos. Na sua utilização há como que uma contaminação da mobilidade da LN, passada via escolhas lexicais que se transformam em unidades documentárias. Assim, as LDs não se livram completamente de interferências culturais que acabam por exigir um trabalho quase permanente de atualização.

O caráter sistêmico fica garantido com a impossibilidade de se ler uma unidade em separado. De fato, cada unidade só pode ser "lida" na sua relação com as demais unidades componentes do sistema.

Por serem sistemas construídos, as LDs são econômicas. Só que não se trata da aplicação do princípio de economia da LN e sim de uma racionalização de escolhas e de procedimentos, que permitam uma utilização eficaz do sistema.

As relações paradigmáticas e sintagmáticas também ocorrem, só que de forma bastante restrita, especialmente nas construções dos sintagmas.

No entanto, nas LDs fica evidente o poder de troca das unidades, numa posição bastante próxima da LN. Cada unidade documentária designa uma realidade dentro do sistema construído, o que torna evidente o valor e a possibilidade de troca, de representação.

De toda forma, as LDs são tributárias da LN, na medida em que são construídas a partir dela. Embora haja um esforço de neutralização de traços que fazem da LN um sistema aberto, heterogêneo e multiforme, as LDs acabam por assimilar algumas particularidades, uma vez que se valem de unidades da LN e são manipuladas, freqüentemente por seres que têm na LN algo, naturalmente, incorporado à sua existência.

A função da LD é tratar o conhecimento dispondo-o como informação. Em outras palavras, compete às LDs transformar estoques de conhecimentos em informações adequa-

das aos diferentes segmentos sociais. É esse partilhamento que está na base do caráter público da informação e que não pode ser obtido na ausência de uma LD. De fato, durante muito tempo acreditou-se que a disponibilização dos estoques seria suficiente para a sua socialização. Mas, atualmente, o fundamental é a existência de uma forma de organização que garanta o partilhamento. Essa organização é a LD.

Coyaud (1972) apontava, na relação entre a Lingüística e a Documentação, aquilo que considerava ser um grande defeito: por um lado, os teóricos da Lingüística acabavam trabalhando sobre questões abstratas, ficando inteiramente ausentes informações sobre as línguas concretas. A seu ver os lingüistas poderiam ser divididos em dois grupos: os teóricos que pendiam para a lógica, para a busca de universais da linguagem, ignorando mesmo as línguas concretas; e os especialistas numa dada língua que acabam por se fechar nela e por força de um terminologia própria chegavam ao limite da incomunicabilidade, até mesmo com o grupo teórico. Entre estes dois extremos havia um vazio.

Evidentemente que esta questão nem é tão simples, nem tão clara. De toda forma, parecem faltar trabalhos de caráter extensional que possam fazer de forma produtiva a ligação de teorias contemporâneas com práticas sociais.

Conhecer primeiro os meandros da linguagem parece ser pré-requisito, razão pela qual foram introduzidos nesse livro conceitos que permitam aprofundar conhecimentos já acumulados.

Do ponto de vista da linguagem três aspectos merecem destaque: a demarcação, a significação e a comunicação. São

esses aspectos que, segundo Kristeva (1969:14), permitem dizer que todas as práticas humanas são tipos de linguagens.

Quanto às linguagens documentárias, é necessário que sejam vistas, simultaneamente, como sistemas e como práticas sociais com todas as suas implicações que vão de seu aspecto material, consubstanciado em cadeias de unidades, à sua natureza comunicativa que pressupõe acordo entre sujeitos que dela se valem. Nesse sentido, uma LD não se apresenta como uma construção universal, segue princípios únicos, mas reflete práticas sociais distintas relacionadas não só às necessidades específicas de informação dos vários segmentos sociais mas também aos vários consensos que os caracterizam.

O livro organiza-se, além da Introdução, em quatro capítulos. O primeiro traz considerações, como o título sinaliza, sobre conhecimento, informação e linguagem verbal ou natural – LN. O segundo enfoca aspectos importantes das linguagens documentárias – LDs. O terceiro discute relações intervenientes nas linguagens documentárias e o quarto retoma conceitos de semântica lingüística que têm papel fundamental na construção de LDs.

#### 1

### Conhecimento, informação e linguagem

#### 1.1 CONHECIMENTO E INFORMAÇÃO

Em estado dicionário *informação* significa "ação ou efeito de informar", "instrução", "indagação", "investigação", "notícia".

O significado de *informação* implica a presença de semas que envolvem apresentação, representação ou criação de idéia, segundo uma forma. Em suma, a informação constitui, ela mesma, um conhecimento potencialmente transmissível.

Sob outro ângulo, pode-se dizer que a informação relaciona-se à identificação de um "sinal" e supõe uma "forma" passível de ser interpretada como mensagem.

De outra ótica, ainda, sabe-se que a *informação* se constitui, na sociedade moderna, em ingrediente indispensável do dia-a-dia das pessoas, graças, de modo especial, aos veículos de comunicação de massa.

Entretanto, é em sentido específico de algum tipo de conhecimento produzido no nível do mundo científico e tecnológico que interessa fazer considerações.

Da mesma forma que a informação acontece nos dois

extremos do circuito da comunicação, o conhecimento acontece no extremo do emissor, responsável pela criação em si e no extremo do receptor, onde se dá a recepção da informação criada.

Neste sentido, parece indiscutível que a *informação* cumpre papel decisivo na mudança dos destinos da humanidade, uma vez que ela está, diretamente, ligada ao conhecimento e ao desenvolvimento de cada uma das áreas do saber, já que todo conhecimento começa por algum tipo de informação e se constitui em informação.

A partir da década de 1970, a noção de informação, bem como os termos que a representam tomam vulto, seja na constituição dos discursos, seja na criação de disciplinas específicas. Acredita-se mesmo que a sua expansão represente, na sociedade ocidental, um dos maiores sucessos de uma palavra no século XX.

A utilização recorrente da palavra gerou, como é natural, uma variação conceitual. Assim, fala-se do conceito de informação em diferentes áreas de conhecimento, podendo a relação informação/conhecimento ser observada a partir de três aspectos que se complementam:

- enquanto o conhecimento é estruturado, coerente e frequentemente universal, a informação é atomizada, fragmentada e particular;
- enquanto o conhecimento é de duração significativa, a informação é temporária, transitória, talvez mesmo efêmera;
- enquanto o conhecimento é um estoque, a informação é um fluxo de mensagens.

Com efeito, o estoque de conhecimentos é alterado com o *input* de novas informações, em virtude de adições, reestruturações ou mudanças.

Mas, para que o *conhecimento* da sociedade não se perca e possa ser compartilhado, ele é registrado num dado suporte: livro, imagem, foto, disco etc., passando a se constituir um documento.

O desenvolvimento científico e tecnológico tem proporcionado à sociedade uma massa enorme de informações geradoras de conhecimentos, portanto de documentos, que precisam ser tratados adequadamente para que haja não só a sua divulgação, como também a criação de novos conhecimentos, cumprindo assim a rotina natural da própria ciência.

Daí o papel fundamental da área de documentação, responsável pela triagem, organização e conservação da informação, bem como pela viabilização a seu acesso.

Há que se considerar que à massa considerável de documentos em papel que constitui volume considerável vêm se juntando, de forma também crescente, documentos em outros suportes como disco, fotografia, fita magnética, vídeo etc.

Segundo Waddington (1975), é praticamente impossível dar uma imagem do mundo moderno, que chegue próxima da exatidão, em termos de conhecimentos acumulados. Entretanto, pode-se chegar a ter uma idéia parcial do problema, quando se consideram os estudos sobre o crescimento da informação científica e técnica nos últimos dois séculos, através só das publicações de revistas especializadas desses dois campos.

As duas primeiras revistas inteiramente dedicadas à ciên-

cia começaram em 1665: The Philosophical Transactions of the Royal Society of London e Journal des Sçavants (França). A partir de 1760, liouve uma implementação de publicações desta natureza que, pouco a pouco, praticamente duplicaram a cada quinze anos.

Sabe-se que até meados da segunda metade do século XX foram fundadas mais de 100.000 revistas científicas. No entanto, não se sabe quantas desapareceram e, hoje, é praticamente impossível dizer o número delas. Para se ter uma idéia, em 1938, calculou-se em 33.000 o número de revistas científicas publicadas, sendo que no final dos anos 1960, ele atingia cerca de 60.000, com um milhão de artigos individuais por ano. Estimava-se, em 1996, que havia 200.000 periódicos em circulação, número que continua crescendo principalmente a partir de sua difusão em formato eletrônico.

Os dados são, sem dúvida, imprecisos, mas suficientes para demonstrar a dimensão do problema cujo desdobramento pode ser observado por meio da criação de revistas secundárias e terciárias, do fenômeno da "redescoberta" científica, da tendência à especialização e da rápida obsolescência da informação.

Com efeito, a primeira revista secundária, cuja função é resumir e sintetizar os artigos publicados nas revistas primárias, surgiu na Alemanha em 1714. De lá para cá, esse tipo de periódico veio aumentando, chegando mesmo a multiplicar-se com, praticamente, a mesma taxa exponencial das revistas primárias. Em 1960, calculou-se em 1.900 o número destas revistas secundárias.

O volume de revistas secundárias levou à criação de revistas terciárias que informavam sobre as revistas de síntese. Sob essa mesma perspectiva, criou-se o Sistema Universal de Informação Científica (UNISIST), com o patrocínio das Nações Unidas, com a tarefa central de armazenar toda a informação científica em um computador central, dotado de um sistema de busca.

O terceiro desdobramento diz respeito a um fenômeno muito comum hoje: é mais fácil redescobrir algo que saber se alguém já o descobriu antes. Acreditam alguns que este fenômeno da "redescoberta" possa tornar-se um dos principais fatores limitadores da taxa de avanço da ciência na sociedade contemporânea. Há, por vezes, um dispêndio enorme de recursos humanos e materiais para descobrir o já descoberto.

A tendência à especialidade constituiu uma característica muito presente nas décadas passadas, chegando mesmo a motivar filósofos e educadores para discutir a questão da interdisciplinaridade. Nos últimos anos, embora ainda persista, com saliência, esta característica, assiste-se a uma forte reação à alta especialidade, de modo particular com os movimentos denominados pós-modernos.

A velocidade de produção de informação tem como conseqüência quase imediata a obsolescência de conhecimentos. De Solla Price discutiu isto em termos do que ele chamou de coeficiente de imediatismo: se a quantidade de informação dobra em quinze anos, ela seria A no início deste período e 2A no fim do mesmo período. O acréscimo de A é A e o coeficiente de imediatismo é A/2A = 1/2. Isto é, ao cabo de quinze anos, 50% das informações disponíveis serão fruto de descobertas realizadas durante o período em questão (Price, 1965).

Não é difícil perceber que em áreas de avanço muito veloz, como a computação, o período de duplicação não é 15 anos, mas muito menor, talvez 4, o que amplia bastante o índice de obsolescência.

Assim, claro está que ninguém pode, nem mesmo numa área de especialidade, aventurar-se a "conhecer" tudo o que se publica. Mas também é claro que uma pessoa pode conseguir informações parciais em níveis satisfatórios, graças aos meios desenvolvidos para guarda e recuperação da informação.

As necessidades, naturalmente, variam de um domínio para outro, de um grupo para outro, segundo o estágio de desenvolvimento da área, a natureza dos usuários, seus objetivos. Apesar dessas variações, é preciso que as informações sejam confiáveis, atuais e imediatamente disponíveis.

Para se chegar a isso é indispensável um trabalho sistemático que se compõe de um conjunto de operações em cadeia, isto é, operações marcadas por íntima relação entre cada uma das etapas: as últimas operações estão ligadas às primeiras e as primeiras vão conduzindo às últimas.

Numa extremidade da cadeia estão os documentos que serão tratados e, na outra, os resultados desse processo expressos em produtos documentários do tipo: referências, descrições de documentos, publicações secundárias e terciárias.

O processo começa pela operação de coleta de dados que se constitui num procedimento de *alimentação*, por meio do conjunto de documentos que passam a integrar uma unidade de informação.

A primeira fase que se decompõe em algumas etapas sucessivas (localização de documentos, triagem e escolha, procedimentos de aquisição propriamente ditos) exige pro-

fissionais atualizados em relação à evolução do conhecimento e à produção no domínio considerado, o que supõe que a unidade de informação esteja bem integrada no circuito científico nacional e internacional, formal e informal.

Quando se trata de publicações disponíveis no mercado, a coleta apóia-se em fontes identificáveis e acessíveis: depósito legal, bibliografias nacionais, catálogos de editores, ou catálogos coletivos, índices, repertórios, bibliografias de toda espécie. Mas, quando se trata de localizar uma literatura dita subterrânea, é fundamental que se possa dispor de uma rede de permuta e de aquisição sistemática, o que implica integração no circuito científico da área.

A segunda fase do processo consiste em operações de controle e registro do material. Nesta fase, é feito o tratamento intelectual dos documentos, por meio de descrição bibliográfica, descrição do conteúdo, estocagem, busca e difusão. Todas essas operações visam encontrar, de imediato, a informação necessária para responder à demanda.

A tarefa inicial consiste em proceder à identificação do documento, o que é feito por intermédio de uma descrição bibliográfica ou de catálogo que explicita suas características formais: autor, título, fonte, formato, língua, data da edição etc.

Em seguida, é feita a descrição do conteúdo, denominada *análise documentária*. Esta etapa recobre operações de descrição das informações que trazem o documento, e a tradução dessas informações numa formulação aceitável pelo sistema adotado.

Daí nasce a relação da ciência da informação com a linguagem natural, relação que precisa ser analisada do ângulo da guarda e da recuperação dos documentos, por meio de sistemas que fazem a representação da informação que veiculam conhecimento.

Não só o volume de documentos constituídos em linguagem natural, como também a natureza da linguagem verbal, justificam uma reflexão específica»

#### 1.2 LINGUAGEM: CARACTERÍSTICAS GERAIS

A linguagem, enquanto objeto de reflexão, perde-se no tempo; entretanto, enquanto objeto de uma ciência, é relativamente recente.

O caráter científico deu à linguagem uma força tal que, hoje, pode-se dizer que ela é tomada como chave de acesso do homem moderno às leis do funcionamento social (Kristeva, 1969).

Embora, desde sempre, ela tenha sido considerada na sua articulação homem/sociedade, hoje, busca-se um isolamento metodológico na tentativa de vê-la como objeto particular em si mesma. O homem como que se distancia, se descola da linguagem que o constitui e obriga-se a "dizer o modo como diz". (Kristeva, 1969, p. 14)

Neste esforço de mais e melhor conhecer a linguagem, os primeiros aspectos que se sobressaem são a demarcação, a significação e a comunicação.

Em primeiro lugar, é preciso dizer que todas as práticas humanas são tipos de linguagens, já que clas têm a função de demarcar, significar e comunicar. Entretanto, como assinala Barthes (1964), qualquer sistema semiológico repassa-se de linguagem verbal.

Ao longo dos tempos, a concepção de linguagem foi se modificando, à mercê do saber constituído e da ideologia reinante. Até o século XVIII, predominou uma concepção teológica que colocava em primeiro plano sua origem e as regras universais da sua lógica. O século XIX foi marcado por uma concepção historicista que via a linguagem como um processo em evolução através dos tempos. Hoje predominam as concepções da linguagem como sistema em funcionamento.

A prática da linguagem é marcada por uma tendência natural do homem: compreender, governar e modificar o mundo. Com efeito, o homem busca, incansavelmente, encontrar uma ordem para as coisas, já que um mundo caótico seria incompreensível, insuportável; por isso ele busca encontrar, em meio à aparência caótica, uma ordem mesmo que subjacente, uma estrutura capaz de explicar as coisas.

Na sua busca reflexiva, o homem trabalha com uma estrutura que é, a um só tempo, estática e dinâmica, isto é, que permite a fixação de cada aparência dentro do esquema geral de referência, ao mesmo tempo em que deixa espaço para que essa mesma aparência surja num outro ponto do quadro, a partir de outras relações, repetindo o mesmo processo.

Assim, situa-se numa ponta a apreensão e, na outra, a compreensão. O primeiro esforço, o de fixação, equivale a uma catalogação do mundo. O segundo, o de coordenação, equivale a uma hierarquização do mundo.

E dentre as coisas a conhecer, provavelmente, seja a linguagem verbal uma das mais intrigantes, já que ela se faz presente no dia-a-dia, de forma inalienável, participando do processo e do produto deste conhecer.

Como é feita de palavras, a grande maioria dos dados

de que o homem dispõe, daquilo que forma seu intelecto, parece importante pensar a palavra, unidade recoberta por inúmeras dificuldades, entre as quais pode-se citar o fato de nem todas as línguas possuírem escrita e, portanto, a identificação da palavra com o espaço em branco ser, em alguns casos, inviável.

Flusser (1963, p. 22) tenta atingir um nível de explicação para a palavra, construindo uma imagem que tenta expressar a passagem das sensações para a linguagem. Diz ele: "Há, aparentemente, uma instância entre sentido e intelecto, que transforma dado em palavra. O intelecto sensu stricto é uma tecelagem que usa palavras como fios. O intelecto sensu lato tem uma ante-sala na qual funciona uma fiação que transforma algodão bruto (dados dos sentidos) em fios (palavras). A maioria da matéria-prima, porém, já vem em forma de fios."

Para ele, ao se definir realidade como conjunto de dados, se está concebendo que a vida do homem se passa numa dupla realidade: por um lado, a realidade das palavras; por outro, a realidade dos dados brutos ou imediatos. Considerando que os dados brutos atinjam o intelecto na forma de palavras, pode-se dizer que a realidade se faz com palavras e palavras in statu nascendi.

Na prática da linguagem natural, sabe-se que as palavras chegam até as pessoas por intermédio dos sentidos de forma organizada, isto é, são agrupadas de acordo com regras preestabelecidas, formando frases.

De um lado, então, a língua pode ser vista como um sistema cujas unidades se articulam no plano da expressão e do conteúdo, planos que se unificam como o único modo de

ser do pensamento, a sua realidade e a sua realização. Assim, a língua integra o universo mais amplo da linguagem e atua como elemento fundamental na comunicação social.

Da mesma forma que não há sociedade sem linguagem, não há sociedade sem comunicação. "Tudo o que se produz como linguagem tem lugar na troca social para ser comunicado" (Flusser, 1963, p. 22).

Na comunicação, observa-se que todo falante assume o duplo papel de destinador e destinatário de mensagens, pois ao mesmo tempo em que é capaz de emiti-las, sabe decifrálas. Ou seja, na situação natural de comunicação, o falante não emite mensagem que ele não seja capaz de decifrar.

Assim se introduz o falante no complexo domínio do sujeito, isto é, no universo da sua constituição e da sua relação com o outro. Na relação consigo mesmo e com o outro falante, opera com o ato de nomear que é feito com a língua, exterior ao indivíduo e submissa a uma espécie de contrato social firmado, naturalmente, para garantir a comunicação.

A língua é, pois, um sistema de signos e regras combinatórias que, de fato, não se realiza completamente na fala de nenhum sujeito. Ela só existe completamente na massa, no conjunto de uma sociedade. Mas também é um sistema de relações virtuais em permanente disponibilidade para o falante.

Enquanto realização, pode-se dizer que, quando as palavras são percebidas, percebe-se uma realidade ordenada, um cosmo, o que permite dizer que a língua é também o conjunto de frases percebidas e perceptíveis.

Por outro lado, as palavras são apreendidas e compreendidas como símbolos, isto é, como tendo significado, por-

que, por meio de um acordo entre vários contratantes, elas substituem algo, apontam para algo, são "procuradoras" de algo.

É, pois, a partir de um acordo entre sujeitos que os sinais são apreendidos e compreendidos, realizando, em sociedade, o caráter simbólico da língua, condição do pensamento.

Tradicionalmente, são distinguidas as palavras a partir de seus significados em substantivos, adjetivos, verbos etc. A mesma tradição ensina que substantivos significam "substâncias", que adjetivos significam "qualidades", que verbos significam "processos modificando substâncias", que preposições e conjunções significam "relações" entre substâncias»

Essa classificação, não obstante ser enfatizada, oferece pontos de conflito muito evidentes. Antes de mais nada, ela pressupõe uma realidade absoluta, um universo uniformemente ordenado, uma estrutura rígida de mundo, espelhada na estrutura da língua. É mais ou menos como na concepção platônica em que o fenomenal espelha a estrutura do mundo das idéias.

Sc a realidade mais ampla mostra línguas como o chinês e, de resto, as línguas aglutinadas e assilábicas onde esta divisão não faz sentido, a presença mais imediata da língua materna mostra realidades que põem em cheque esta divisão.

Enquanto na frase "Isto é uma caixa grande", "caixa" e "grande" são expressões autênticas, respectivamente, das significações substância e qualidade, na frase "Isto é um caixão" a qualidade como que vem engolida pela substância. Já na frase "Viver é lutar" observam-se processos assumidos como substâncias.

Os exemplos poderiam ser multiplicados, para mostrar

que a classificação absoluta não corresponde à realidade. Entretanto, é preciso admitir que a classificação tradicional, mesmo com possíveis defeitos, oferece vantagens, na medida em que ela permite ver a língua como um sistema de símbolos apontando para algo, ou significando algo. Na realidade, a língua não se constitui num conjunto de símbolos equivalentes, mas, antes, num conjunto de símbolos hierarquicamente diferenciados. O significado de cada símbolo só se torna compreensível dentro do conjunto do sistema inteiro.

A língua não é função do sujeito falante nem sucessão de palavras correspondentes a outras equivalentes. É um sistema-estrutura de valores e formas. Os sistemas de valores não são construções particulares de um indivíduo; são, antes, o resultado de todo um contexto sociohistórico que determina as condições de produção do discurso.

#### 2

#### Linguagens documentárias

#### 2.1 NATUREZA, ESPECIFICIDADE E FUNÇÕES

Um rápido retrospecto sobre a área da documentação mostra que, nas décadas de 1950 e 1960, com o crescimento do conhecimento científico e tecnológico, houve dificuldades para armazenar e recuperar informações. A solução foi encontrada com uma mudança do enfoque e da conceituação da recuperação da informação. Com efeito, foi abandonada a perspectiva preferencial de recuperação bibliográfica e normalização classificatória e descritiva, buscando-se a construção de linguagens próprias.

Vem desta época a utilização de Linguagens Documentárias – LDs, para a recuperação da informação. Essas linguagens são, pois, construídas para indexação, armazenamento e recuperação da informação e correspondem a sistemas de símbolos destinados a "traduzir" os conteúdos dos documentos.

Como decorrência desta mudança de conceituação da área, houve grande concentração em estudos de Lingüística e de Estatística, especialmente para viabilizar a automação do tratamento da informação.

Com os estudos de Lingüística esperava-se resolver pro-

blemas de vocabulário, tendo em vista a construção de instrumentos mais adequados. Estes estudos levaram a análises de conteúdos da Linguagem Natural – LN, a buscas de métodos de padronização relativos à passagem da LN para a LD, ao estabelecimento de mecanismos para a estruturação de campos semânticos, de campos associativos e de categorias funcionais.

A Estatística, por sua vez, foi tomada como instrumento de apoio, tendo em vista determinar frequências de descritores, mapeamento de ocorrências e análise de citações, o que levou ao desenvolvimento da Bibliometria.

No amplo universo da linguagem, as LDs possuem um *status* muito particular: por meio delas pode-se representar, de maneira sintética, as informações materializadas nos textos.

Tal como a LN, as LDs são sistemas simbólicos instituídos que visam facilitar a comunicação. Sua função comunicativa, entretanto, é restrita a contextos documentários, ou seja, as LDs devem tornar possível a comunicação usuáriosistema.

Grande parte das discussões teóricas sobre LDs inseremse no âmbito da Análise Documentária que, por sua vez, se define como uma atividade metodológica específica no interior da Documentação, que trata da análise, síntese e representação da informação, com o objetivo de recuperá-la e disseminá-la.

Nesse contexto, as LDs são, pois, instrumentos intermediários, ou instrumentos de comutação, através dos quais se realiza a "tradução" da síntese dos textos e das perguntas dos usuários. Esta "tradução" é feita em unidades infor-

macionais ou conjunto de unidades aptas a integrar sistemas documentários. A formalização das perguntas dos usuários é feita na linguagem do próprio sistema. É por esta razão que as LDs podem ser concebidas como instrumentos de comutação documentária.

Mas, diferentemente da LN, o sistema de relações das LDs não é virtual, bem como seus mecanismos de articulação são extremamente precários, em face daqueles existentes nas línguas em geral. Bem ao contrário, elementos dessa linguagem específica são selecionados de universos determinados e seu sistema de relações é construído, sendo indispensável, para utilizá-la, a existência de regras explícitas. Por esse motivo, as LDs são linguagens construídas.

Cada LD específica representa, por outro lado, um ponto de vista particular sobre a realidade. Como sistema de relações construído, o significado de cada um de seus elementos vai estar diretamente subordinado às definições correspondentes aos elementos colocados nas posições superiores do sistema.

Segundo Gardin, uma LD é um conjunto de termos, providos ou não de regras sintáticas, utilizadas para representar conteúdos de documentos técnico-científicos com fins de classificação ou busca retrospectiva de informações (Gardin et al., 1968).

Para o autor, uma LD deve integrar três elementos básicos:

- um léxico, identificado como uma lista de elementos descritores, devidamente filtrados e depurados;
- uma rede paradigmática para traduzir certas relações essenciais e, geralmente estáveis, entre descritores. Essa

- rede lógico-semântica, corresponde à organização dos descritores numa forma que, *lato sensu*, poder-se-ia chamar classificação; e
- uma rede sintagmática destinada a expressar as relações contingentes entre os descritores, relações que são válidas no contexto particular onde aparecem. A construção de "sintagmas" é feita por meio de regras sintáticas destinadas a coordenar os termos que dão conta do tema.

Embora na LN haja diferença conceitual clara entre léxico, vocabulário, nomenclatura e terminologia, observamse usos sinonímicos de léxico e vocabulário por um lado, e nomenclatura e terminologia por outro.

Nas LDs, por sua vez, é bastante frequente o uso indiscriminado destas palavras, o que pode comprometer o próprio conceito de representação documentária, na medida em que a cada termo deveria corresponder uma função diferente dentro da linguagem.

Entretanto, cada uma dessas palavras remete a conceitos específicos, o que nos permite dizer que cada uma tem características e funções próprias, fator suficiente para impedir sua utilização indiscriminada.

Embora mesmo nos estudos das ciências da linguagem haja, eventualmente, referência a léxico e vocabulário como conjunto de palavras de uma língua ou de um autor, de uma arte ou de um meio social, a rigor, *léxico* designa o conjunto de unidades reais e virtuais que formam a língua de uma comunidade, algo como um depósito de elementos em estado virtual e de regras que permitem a construção de novas unidades, necessárias para a atividade humana da fala.

Já vocabulário refere-se ao conjunto de ocorrências que integram um determinado corpus discursivo, como uma lista de unidades da fala (Dubois et al., 1973). Assim, pode-se falar no vocabulário que encontramos no trabalho de Cunha, relativamente às ocorrências registradas nos discursos sobre política colonial de Adriano Moreira (Cunha, 1990), ou no vocabulário médico, a partir de levantamento em determinadas obras médicas, por exemplo.

Em termos de LDs, não faz sentido falar nem em léxico, nem em vocabulário nas acepções da Lingüística, uma vez que esses elementos são específicos da LN. As LDs, linguagens construídas que são, com finalidades específicas de representação documentária, não são suficientemente articuladas, nem se constituem em unidades geradoras de novos elementos.

Também não integram vocabulários propriamente ditos porque são formadas de palavras preferenciais, combinando palavras de vocabulários de determinados domínios e palavras utilizadas pelos usuários. Desta forma, englobam vários vocabulários, representativos de vários discursos. Assim quando a palavra vocabulário refere-se à LD, deve ser entendida segundo esta última acepção, que privilegia uma constituição a partir de origens diferentes.

Uma nomenclatura, por sua vez, como sugere a própria palavra, diz respeito à ação de chamar algo por seu nome. Assim, se constitui em lista de nomes que supõem biunivocidade da relação significado-significante (Dubois et al., 1973). Talvez se possa melhor caracterizar uma nomenclatura como etiquetas que designam coisas ou conceitos préexistentes, como a nomenclatura da Química, por exemplo,

na qual, independentemente de um sistema nocional particular, algo se chama ouro, nitrogênio ou potássio.

Diferentemente de uma nomenclatura, uma terminologia refere-se ao conjunto de termos de uma área, termos relacionados e definidos rigorosamente para designar as noções que lhe são úteis (idem, ibidem). Assim, por exemplo, a terminologia da educação brasileira pode ser encontrada no Glossário de termos em educação (Brasil, Ministério da Educação e Cultura, 1980). Trata-se de um sistema de termos organizados a partir de noções particulares.

É bom lembrar que todo conhecimento técnico-científico desdobra-se num universo de linguagem. A linguagem condiciona o conhecimento objetivo, determina os limites e sua formulação (Granger, 1974). As linguagens construídas exigem formulações rigorosas de sentido à medida que a própria atividade se encontra subordinada à articulação da linguagem. Desse modo, a atividade terminológica é parte constitutiva da atividade técnico-científica e diz respeito, diretamente, a um conjunto de termos organizados.

Todas as definições analisadas anteriormente levamnos a concluir que as LDs não se confundem com léxicos, vocabulários, nomenclaturas e terminologias, embora incorporem elementos de todos eles. É importante que essa diferenciação seja feita, para melhor delimitar suas características em face da função que devem desempenhar na representação da informação documentária.

A representação documentária é obtida por meio de um processo que se inicia pela análise do texto, com o objetivo de identificar conteúdos pertinentes em função das finalidades do sistema – e da representação desses conteúdos – numa forma sintética, padronizada e unívoca.

A síntese e a representação documentárias advindas do processo de análise podem apresentar-se, geralmente, sob duas formas: o resumo, que é feito sem a intermediação de uma LD e o índice, que, para maior qualidade, deve ser elaborado a partir de uma LD.

A operação de tradução de textos em LN para uma LD denomina-se indexação. Inerente ao processo de indexação estão operações de classificação.

As várias fases do processo analítico apresentam uma complexidade considerável, pois não se trata de adquirir os documentos e armazená-los numa ordem lógica. A documentação é memória, seleção de idéias, reagrupamento de noções e de conceitos, síntese de dados. Trata-se de triar, de avaliar, de analisar, de "traduzir", de encontrar respostas para necessidades específicas.

A utilização da LN neste processo leva, seguramente, à incompreensão e à confusão, devido a fenômenos naturais como a redundância, a ambigüidade, a polissemia e as variações idioletais.

A condição para se obter resultados positivos na busca de informação é que a pergunta e a resposta sejam formuladas no mesmo sistema. Assim, é necessário converter uma pergunta feita em LN para o sistema em que foi traduzido o conteúdo do documento, isto é, para uma LD.

Dito de outro modo, uma LD é utilizada na entrada do sistema, quando o documento é analisado para registro. Seu conteúdo é identificado e "traduzido", de acordo com os termos da LD utilizada e segundo a política de indexação

estabelecida. É da mesma forma utilizado à saída do sistema, quando, a partir da solicitação da informação pelo usuário, é feita a representação para busca. Assim, seu pedido é analisado, seu conteúdo identificado e devidamente "traduzido" nos termos da LD utilizada.

Para realizar tais funções de intermediação, as LDs devem ser construídas de tal forma que seja possível o controle sobre o vocabulário. Tal controle é necessário para que, a cada unidade preferencial integrada numa LD, corresponda um conceito ou noção. Essa correspondência só é assegurada por intermédio das terminologias de especialidade.

Vale lembrar que, isoladas, as palavras não têm significado ou têm todos os significados possíveis. É só no discurso, ou seja, no uso, que as palavras assumem significados particulares. Como, via de regra, os elementos das LDs são desvinculados dos contextos onde aparecem, pode-se correr o risco de que as palavras que as integram assumam todos ou nenhum significado. Por meio das terminologias de especialidade, as palavras passam a ser termos, assumindo significados vinculados a sistemas de conceitos determinados. Confere-se, desse modo, referência às palavras, que passam a significar segundo determinados sistemas nocionais, assegurando interpretações pertinentes.

As LDs mais conhecidas são os tesauros e os sistemas de classificação bibliográfica (Gomes, 1990). As diferenças entre esses dois tipos de LDs residem no maior ou menor grau de reprodução das relações presentes na LN e no universo de conhecimento que pretendem cobrir.

Os primeiros sistemas de classificação bibliográfica conhecidos são de natureza enciclopédica, como a CDD – *Dewey*  Decimal Classification, a CDU – Classificação Decimal Universal e a LC – Library of Congress, e visam cobrir todo o espectro do conhecimento. Sistemas posteriores como as classificações facetadas desenvolvidas a partir do CRG – Classification Research Group, com base na Colon Classification, de Ranganathan, visam a domínios particulares. Os tesauros, por seu lado, originaram-se de classificações facetadas com uma preocupação adicional: a do controle do vocabulário.

Historicamente, verifica-se contínua progressão das LDs a caminho da especialização. Consequentemente, abandonase a pretensão de cobrir todo o universo do conhecimento para voltar-se a domínios cada vez mais específicos.

Todas as LDs, entretanto, são utilizadas para representar o conteúdo dos textos, mas não os textos eles mesmos. A função de representação deve ser entendida, neste contexto, como sendo de natureza eminentemente referencial: as unidades de uma LD devem ser utilizadas como índices relativos a assuntos tratados nos textos, não tendo, portanto, a função de substituí-los.

Os produtos obtidos por meio da intermediação das LDs são, desse modo, generalizantes. Não se representa o texto individual, mas a classe de assunto à qual ele se refere. A maior ou menor especificidade do assunto a ser representado depende da maior ou menor correspondência da LD com o sistema nocional dos domínios de especialidade. Assim, por intermédio de um sistema de classificação enciclopédico, textos muito específicos são classificados em classes de assunto mais gerais; a representação da especificidade dos assuntos de tais textos é mais viável com o uso de uma LD voltada, especificamente, para o domínio correspondente.

Os estudos das LDs têm avançado progressivamente, na direção da definição dos constituintes e de suas interrelações, gerando várias linguagens, de acordo com o domínio de especialidade. Isto, por um lado, permite que a área se libere do monopólio das classificações universais; por outro, tem mostrado inúmeros problemas ligados à falta de rigor na construção de LDs. Tais problemas referem-se à definição do conjunto de termos que comporão a lista de descritores; à organização dos termos numa rede paradigmática (árvores classificatórias ou relações verticais) para reunir descritores; ao estabelecimento da rede sintagmática (relações horizontais entre descritores e mecanismos de sintaxe) para permitir maior possibilidade de representação de novos conceitos e a agilização na recuperação de assuntos; à definição das chaves de acesso ao sistema (compatibilização de linguagem usuário/sistema).

## 2.2 CONFIGURAÇÃO DAS LINGUAGENS DOCUMENTÁRIAS

As LDs mais consistentes para a representação documentária dispõem de um vocabulário que integra elementos, de um lado, da linguagem de especialidade e das terminologias e, de outro, da LN que é a linguagem dos usuários. Essas unidades, acompanhadas ou não de uma notação, constituem o "léxico" das LDs, denominadas, diferentemente, conforme o sistema e a época, como: palavras-chave, descritores, cabeçalhos de assunto etc.

O vocabulário documentário tem por objetivo reunir unidades depuradas de tudo aquilo que possa obscurecer o sentido: ambigüidade de vocábulo ou de construção, sinonímia, pobreza informativa, redundância etc. Além disso, ele é fixado de tal forma que seu uso, bem como suas relações estruturais são codificados e não podem mudar ao sabor dos usuários. Assim, chega-se a um instrumento relativamente estável.

Toda LD tem, também, uma sintaxe. Ela é bastante rudimentar nos sistemas de classificação bibliográfica (Add notes, na CDD; uso de + / , : na CDU, por exemplo) e mais desenvolvida nos tesauros, com a utilização de operadores booleanos. O esquema sintático de uma LD permite a delimitação mais precisa de um assunto, por meio da combinação de seus elementos.

Nos sistemas de classificação convencionais, não há grande preocupação com o controle do vocabulário. É freqüente a utilização de frases, como ocorre, por exemplo, na CDU. Já nos tesauros, a função de controle do vocabulário está mais presente. Para este fim, as LDs incorporam procedimentos de normalização gramatical e semântica. A normalização gramatical refere-se à forma de apresentação dos seus elementos quanto ao gênero (geralmente masculino), ao número (uso de singular ou plural) e ao grau. (Para mais informações, ver Gomes, 1990) A normalização semântica procura garantir a univocidade na representação dos conceitos de áreas de especialidade, por meio das relações lógico-semânticas.

O conjunto nocional básico é apresentado em hierarquias (na vertical), em torno das quais se agregam as unidades informacionais que se relacionam horizontalmente. Nenhuma unidade pode figurar numa LD sem que esteja relacionada a uma outra unidade da mesma linguagem.

Nos tesauros, os diferentes tipos de relações entre as unidades são mais claramente apresentados enquanto sistemas de classificação bibliográfica, não raras vezes amalgamam, numa mesma hierarquia, relações de natureza diferente.

As variações na forma de apresentação das LDs devemse a maior ou menor incorporação dos diferentes tipos de relações existentes entre as palavras na LN e entre os termos de especialidade. Tais variações exprimem, também, o maior ou menor aprimoramento da função de representação documentária.

Algumas LDs foram construídas visando, principalmente, à organização dos documentos nas estantes, sendo que sua função de representação deve ser diferenciada: a representação nesse caso deve ser entendida como a identificação de documentos com classes genéricas de assuntos tradicionalmente reconhecidos.

A estrutura básica de uma LD é dada por relações hierárquicas, que podem ser genéricas, específicas ou partitivas. (relações genéricas e relações partitivas serão tratadas no capítulo 3). O vértice de cada hierarquia é o gênero ou o todo. As subdivisões sucessivas na hierarquia constituem as espécies e/ou as partes, que podem, novamente, se subdividir. As relações hierárquicas provêem as unidades superordenadas e as unidades subordinadas. Unidades subordinadas ao mesmo vértice, quando no mesmo nível da cadeia, denominam-se coordenadas.

Nos sistemas de classificação bibliográfica, a estrutura hierárquica é dada pela notação (decimal, no caso da CDD e da CDU). O vértice das cadeias hierárquicas é constituído por disciplinas convencionais que se subdividem sucessivamen-

te. A indicação dos assuntos é feita por meio da notação numérica ou alfa-numérica, conforme o tipo de sistema.

A organização básica dos tesauros também é hierárquica, existindo tantos vértices, que equivalem a classes, quantos forem os aspectos escolhidos para organizar o domínio de especialidade. Nos tesauros mais modernos, tais vértices são denominados Top Terms e não constituem descritores, mas identificam as classes escolhidas para reunir os descritores. Via de regra, são utilizadas notações numéricas apenas para apresentar as hierarquias básicas e suas principais subdivisões. Tais notações, entretanto, raramente são utilizadas para descrever o conteúdo dos textos. A ligação lógico-hierárquica entre descritores é, no caso dos tesauros, mais clara, uma vez que é identificada pelos códigos TG (Termo Genérico ou Termo Geral), TE (Termo Específico). Alguns tesauros utilizam, também, os códigos TGP (Termo Genérico Partitivo) e TEP (Termo Específico Partitivo) para apresentar as relações hierárquicas do tipo todo/parte.

As LDs apresentam, ainda, unidades que são relacionadas de forma não-hierárquica. As relações não-hierárquicas são, normalmente, denominadas associativas, muito embora não se possa afirmar que as relações hierárquicas também não o sejam. É preciso lembrar, entretanto, que as relações hierárquicas representam associações mais estáveis entre termos, enquanto que as relações não-hierárquicas expressam outro gênero de proximidade entre os termos. Os relacionamentos não-hierárquicos indicam a ligação entre termos que estão em campos semânticos distintos, porém próximos. Cada termo relacionado pode se constituir no ponto de partida para uma família de termos aparentados.

Nos sistemas de classificação bibliográfica, os relacionamentos não-hierárquicos, quando ocorrem, são erroneamente "encaixados" nas hierarquias. É só nos tesauros que estas relações são explicitamente identificadas pelo código TR (Termo Relacionado).

Adicionalmente, as LDs apresentam relações de equivalência. Este tipo de relacionamento entre os termos permite a compatibilidade entre a linguagem do sistema e a do usuário, operando no nível da sinonímia. Desse modo, criam-se as remissivas, indicadas nos tesauros pelas expressões USE (Use) e UP (Usado Para), quase inexistentes nos sistemas de classificação bibliográfica. As relações de equivalência remetem o conjunto dos não-termos ou não-descritores para o conjunto dos termos ou descritores. A finalidade dessas remissivas é encaminhar o usuário para os termos preferidos pelo sistema. Constitui-se, desse modo, uma chave de acesso ao sistema.

O conjunto de relações que constitui a estrutura do tesauro é "um elemento importante para que ele possa cumprir sua função: ela permite ao usuário (indexador ou consulente) encontrar o(s) termo(s) mais adequado(s), mesmo sem saber, de início, o nome específico para representar a idéia ou o conceito que ele procura. A partir de um termo que o usuário conhece, o tesauro, através de sua estrutura, mostra diversos outros que podem ser tão oportunos ou mais do que aquele que lhe veio à mente" (Gomes, 1990, p.16).

Vale ressaltar, ainda, que no uso das LDs, podem ser construídas novas relações entre os termos a partir do conjunto de operadores sintáticos disponíveis, como, por exemplo, as add notes, na CDD; + / : ::, no caso da CDU; operadores booleanos, no caso dos tesauros.

Uma vez elaboradas e postas em uso, as LDs mais desenvolvidas como os tesauros, são permanentemente atualizadas, mediante operações de supressão de termos em desuso, reagrupamento de descritores em função da existência de palavras raramente utilizadas e/ou adição de termos novos. Só assim as LDs se mantêm como instrumentos dinâmicos capazes de incorporar os avanços do conhecimento e as modificações de significado de termos já existentes.



# 3 Sistema nocional

A todo e qualquer campo de conhecimento corresponde um conjunto de noções que lhe é próprio. As áreas especializadas da experiência humana devem ter seu universo nocional devidamente identificado a partir de um dado ponto de vista, para que seja possível organizá-lo de forma sistemática, ou seja, inter-relacionada. Só a organização nocional de uma área permite a utilização de instrumentos eficazes para o tratamento e recuperação da informação.

A ausência de um sistema de noções dévidamente sistematizado, inviabiliza o empreendimento de dar forma a um conjunto de palavras, na medida em que esbarra, necessariamente, em dificuldades advindas da falta de compreensão ou da compreensão incorreta das possibilidades de relacionamento entre termos.

Considerando que as LDs, normalmente, funcionam a partir do controle do "vocabulário" da área, pode-se facilmente depreender que o sistema nocional de uma área constitui-se em um parâmetro básico, ou em uma viga-mestra de sustentação das LDs.

Assim, na prática, a ausência de um sistema nocional compromete não só a indexação, mas também, a economia

da própria atividade documentária, fragmentando-a com questões relativas ao significado e à compreensão dos termos. Além disso, não raro, as respostas às questões formuladas submetem-se a variações, segundo o entendimento que cada indexador tem da área, ou segundo o humor no momento daquele que opera com a informação, o que, fatalmente, introduz deformações, descaracterizando os instrumentos documentários.

Desta maneira, faz-se necessário estabelecer, a priori, que a utilização de qualquer LD supõe a explicitação nocional da área a que se refere e a sua organização na forma de um sistema.

Segundo a norma ISO 1087, um sistema nocional define-se como um "conjunto estruturado de noções que reflete as relações estabelecidas entre as noções que o compõem e no qual cada noção é determinada pela sua posição no sistema". Não basta, portanto, recuperar as noções, enumerando-as. É preciso ir além e estabelecer suas posições relativas, o que se obtém por meio da determinação das relações que as associam.

A noção ou o conceito, por sua vez, define-se como "unidade de pensamento constituído por propriedades comuns a uma classe de objetos" (ISO 1087). Embora não estejam ligadas a línguas específicas, as noções são expressas por termos e símbolos, sendo influenciadas pelo contexto sócio-cultural.

As noções, devidamente relacionadas, constituem, pois, o arcabouço fundamental para a organização de uma área, na medida em que possibilitam um ponto de vista materializado no sistema de noções, para o trabalho documentário.

As relações entre as noções materializam o sistema de noções, que se expressam, documentariamente, em relações hierárquicas e relações não-hierárquicas.

As relações hierárquicas são aquelas que se definem entre noções subordinadas em um ou vários níveis (ISO 1087). Dito de outra forma, as relações hierárquicas são aquelas que acontecem entre termos de um conjunto, onde cada termo é superior ao termo seguinte, por uma característica de natureza normativa.

No conjunto das relações hierárquicas, há que se levar em conta o conceito de ordem e de subordinação. A ordem deve ser observada como uma superordenação que consiste na possibilidade de subdivisão de uma noção hierárquica mais alta em um certo número de noções de nível inferior, chamadas noções subordinadas. É este processo de subdivisão que se denomina subordinação. Inversamente, a noção subordinada é a noção que, num "sistema hierárquico", pode ser agrupada com uma ou mais noções do mesmo nível (noções coordenadas entre si), para formar uma noção de nível superior (ISO 1087), ou seja, uma superordenação.

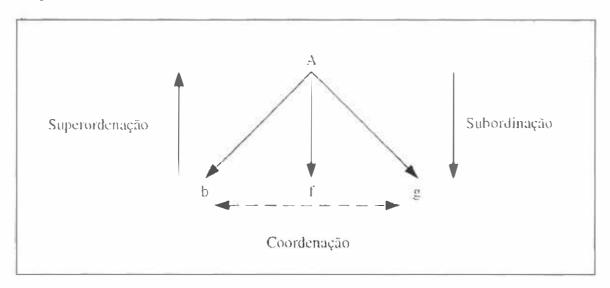


Fig. 1 – Esquema de relações hierárquicas.

A partir das noções de geral/particular e de todo/parte, a análise das relações hierárquicas mostra, pelo menos, três tipos característicos: as relações genéricas, as relações específicas e as relações partitivas que, como os nomes indicam, marcam relações de gênero, portanto globais ou gerais, relações de espécie, logo particulares e relações de parte de nm todo.

As relações genéricas definem-se como relações hierárquicas, baseadas na identidade parcial do conjunto de características das noções superordenadas e subordinadas nelas envolvidas. O gênero, nesse sentido, é entendido como noção superordenada que comporta as mesmas características das noções subordinadas, a partir delaæ

Já as relações específicas definem-se como relações hierárquicas subordinadas que, além de compartilhar das mesmas características da noção que lhes é superordenada, apresenta, pelo menos, uma característica a mais que as diferencia.

A noção genérica impõe-se, portanto, como conjunção de características comuns, enquanto que a noção específica estabelece uma disjunção, a partir da conjunção dada.

A noção específica é, portanto, uma noção subordinada que indica a existência de uma diferença, em face de um conjunto de características comuns. Ao mesmo tempo, apresenta as características comuns e, pelo menos, uma característica que a diferencia da noção genérica.

Assim, por exemplo, ao subdividir o conjunto dos mamíferos em racionais e irracionais, afirma-se, simultaneamente, a existência de uma diferença (racional e irracional) sobre um plano comum ou semelhante (mamíferos).

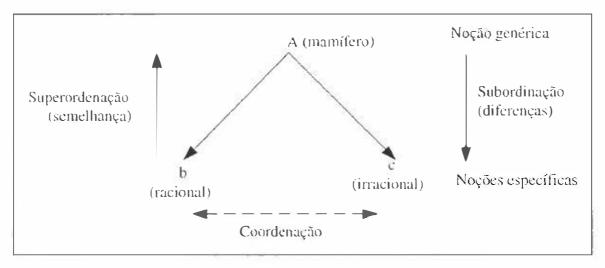


Fig. 2 – Esquema de relação genérica.

Na relação genérica, a superordenação caminha das diferenças para as semelhanças, ou seja, da espécie para o gênero e, inversamente, a subordinação caminha das semelhanças para as diferenças, a partir das primeiras, isto é, do gênero para as espécies.

Exemplificando: a noção de "embarcação" subdivide-se, segundo o "tipo", em noções mais específicas como a de iate, jangada, canoa, navio, chata etc. Em relação a essas últimas, a noção específica "embarcação" é a noção superordenada. É a partir dessas relações que se pode afirmar que iate é uma espécie de "embarcação; que "embarcação" é um gênero; e que iate e canoa são noções coordenadas.

Já a relação partitiva é um tipo de relação hierárquica, na qual a noção superordenada refere-se a um objeto considerado como um todo e as noções subordinadas a objetos considerados como suas partes. Em relação a "navio", a noção de "casco" é uma noção específica partitiva, denotando que navio é uma noção referente ao todo (superordenada) e que "casco" é uma noção referente à parte (subordinada). Do mesmo modo, a noção "convés" denota uma subdivisão

por partes da noção "navio". Relacionadas por coordenação, as noções "convés" e "quilha" são denominadas noções coordenadas.

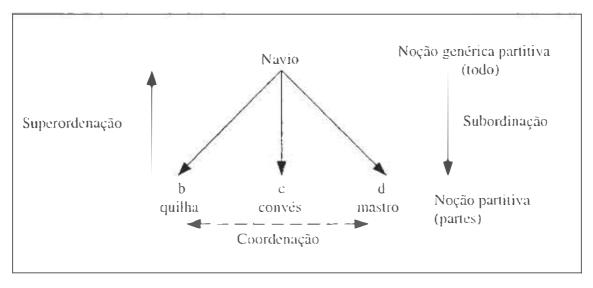


Fig. 3 – Esquema de relação partitiva.

As relações não-hierárquicas, por sua vez, definem-se pela negativa. Elas recobrem o conjunto de relações que não são passíveis de serem descritas como hierárquicas.

É evidente a insuficiência dessa abordagem. No entanto, concretamente, pouco se pode a ela acrescentar. Se as relações hierárquicas supõem ordem e subordinação lógicas, as não-hierárquicas não podem supor, exatamente, essas características.

As relações que não se submetem a uma hierarquia são aquelas que apresentam entre si contigüidade espacial ou temporal. Por esta razão, tais relações também são chamadas de relações seqüenciais.

Consideram-se relações sequenciais as de oposição, as de causa-efeito, as de contradição e outras menos evidentes como aquelas estabelecidas entre as etapas de um processo,

produto e produtor, instrumento e processo. É importante salientar que tais relações não podem ser definidas em toda sua extensão.

De certo modo, a dificuldade de definir as relações nãohierárquicas encontra-se enunciada na sua denominação usual: relações associativas. A impropriedade do termo "associação" deve-se, neste caso, ao fato de que qualquer que seja a natureza da relação, ela é, em certo grau, associativa. O problema então continuaria: haveria relações associativas hierárquicas e relações associativas não-hierárquicas.

Por isso é preciso restabelecer o contexto que a valide, ou seja, indagar a sua natureza. O exemplo transcrito abaixo, da Norma ISO 1087, ilustra este aspecto.

RELAÇÃO ASSOCIÁTIVA	
DOENÇA TRANSMITIDA SEXUALMENTE	ALIENAÇÃO MENTAL INCURÁVEI.
Gonorréia	Demência
Sífilis	Esquizofrenia
Cancro	Psicose maníaco depressiva
Linfogranulomatose inguinal	

Fig. 4 - Motivos médicos juridicamente aceitos para o divórcio.

### 3.1 RELAÇÕES HIERÁRQUICAS

A macro-organização da maior parte das LDs funda-se na organização lógico-hierárquica de suas unidades. A delimitação de classes de assunto é feita a partir de pontos de vista determinados. Tais pontos de vista, por sua vez, estão baseados em postulados de significado ou convenções culturais e ideológicas.

Esse é o caso dos sistemas de classificação bibliográfica como a DEWEY DECIMAL CLASSIFICATION, CDD e a CLASSIFICAÇÃO DECIMAL UNIVERSAL, CDU. Esses sistemas seculares são atualizados por edições periódicas que buscam solucionar problemas da contemporaneidade, já que eles foram organizados, na sua forma inicial em 10 classes principais que cobriam, consensualmente, o conhecimento de então. Tais classes, por sua vez, subdividem-se sucessivamente.

A organização lógico-hierárquica é também a base da organização dos tesauros. Como já mencionado, os tesauros têm sua origem na *Colon Classification* de Ranganathan e nas experiências posteriores desenvolvidas pelo *Classification Research Group*, referentes à estruturação do conhecimento, a partir da noção de "faceta", ou seja, da noção que privilegia determinados pontos de vista no arranjo dos domínios e subdomínios particulares, em função de objetivos específicos do sistema documentário em questão. A fonte de referência para a construção das hierarquias, neste caso, é a estrutura teórico-conceitual de domínios específicos, determinando-se conjuntos de termos do domínio nuclear – a área de especialização propriamente dita –, e domínios periféricos, ou áreas complementares, conforme necessidades objetivas do sistema em questão.

No caso dos sistemas de classificação e dos tesauros, a organização da macro-hierarquia e das hierarquias subsequentes depende, portanto, dos princípios ou características de divisão adotados a cada passo, variando conforme objetivos determinados: a CDD e a CDU pretendem referir-se ao

universo global de conhecimento, tendo-se curvado, para tanto, às referências postuladas por Bacon para a organização do conhecimento; já os tesauros voltam-se para domínios cada vez mais particulares, sendo construídos em função de universos muito determinados. São, por essa razão, mais flexíveis quanto à estruturação do esquema classificatório básico e mais adequados ao atendimento das necessidades informativas de domínios especializados.

A flexibilidade dos tesauros vincula-se a um princípio de utilidade. Desse modo, pode-se construir, para um campo particular do conhecimento, tantos tesauros quantos forem necessários. Cada um deles procurará organizar um dado universo nocional, de acordo com o ponto de vista que se imprime ao domínio, para responder a diferentes necessidades. Para a ISO 704, "um objeto específico pode ser visto de diferentes pontos de vista por disciplinas diferentes".

Assim, por exemplo, "em termodinâmica as características essenciais do conceito 'líquido' são aquelas que indicam que ele é 'uma substância em estado condensado, intermediário entre sólido e gasoso'" (idem, ibidem). "Em hidromecânica, as características essenciais do conceito 'líquido' são que ele é uma substância que é 'incompressível', 'densa e capaz de fluir'" (idem, ibidem).

No exemplo da ISO, as características (propriedades) privilegiadas na definição de "líquido" em termodinâmica ou em hidromecânica, determinam a definição, implicando, portanto, modos específicos de abordagem do assunto e, consequentemente, construção das hierarquias.

Desta forma, é possível construir tantas hierarquias quantas diferentes conjunções realizarmos entre as palavras.

Exemplificando:

ANIMAIS

ANIMAIS

Mamíferos

Carnívoros

Aves

Herbívoros

Répteis

Batráquios

Peixes

Cada uma dessas construções delimita e conforma as noções ou conceitos a serem representados, refletindo escolhas de determinadas propriedades, tal como numa árvore de Porfírio. "O homem é necessariamente mortal somente numa árvore de Porfírio particularmente focalizada no problema da duração da vida" (Eco, 1984, p. 51).

Refletindo tais princípios de organização, a configuração das LDs é fruto da organização empírica das propriedades das palavras (e não das coisas), estando fundamentada em postulados sócio-culturais. As classes, assim obtidas, representam, portanto, pontos de vista determinados sobre os assuntos.

# 3.1.1 Relação genérica

Uma relação genérica supõe uma noção fundamental que inclui noções específicas que, por sua vez, mantém com ela relações hierárquicas (Boutin-Quesnel et al., 1985).

Por exemplo, a noção de árvore agrupa noções mais específicas de folhas e de coníferas; por sua vez, as coníferas são, segundo a persistência das folhas, caducas e não-caducas (idem, ibidem).

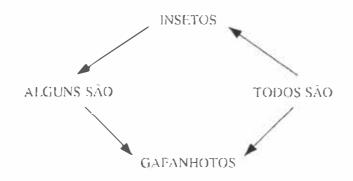
Desse modo, "as relações genéricas indicam que todo

conceito que pertence à categoria do conceito específico (a espécie) é parte da extensão do conceito amplo (o gênero). Um conceito específico possui todas as características do conceito mais amplo, mais, pelo menos, uma característica distintiva adicional que serve para diferenciar conceitos específicos no mesmo nível de abstração" (ISO 704).

A extensão de urna noção corresponde ao "conjunto de indivíduos aos quais uma noção pode ser aplicada (Boutin-Quesnel et al., 1985) e diz respeito à "totalidade de todas as espécies que pertencem ao mesmo nível de abstração ou à totalidade dos objetos que têm todas as características do conceito" (ISO 704).

A noção de extensão vem sempre associada à de intensão ou compreensão. Intensão de uma noção é o conjunto de características que compõem esta noção (Boutin-Quesnel et al., 1985). A intensão de um conceito diz respeito à totalidade das características deste conceito (ISO 704). Portanto, quanto maior a intensão do conceito, menor sua extensão e viceversa. Ou seja, quanto maior o número de características que compõem um conceito, menor é o número de objetos que compartilham destas características (lei da correlação reversa).

A validade de uma relação genérica pode ser constatada por meio de um esquema lógico do tipo "todos/alguns".



O esquema precedente (IBICT, 1984, p. 26; ISO 2788-1986, 1989, p. 605) indica que alguns membros da classe "Insetos" são conhecidos como "gafanhotos", enquanto que todos os "gafanhotos" são "insetos", por definição e independentemente do contexto. Isso porque a classificação tem por base as características que são necessárias e suficientes para distinguir noções. O conjunto de objetos ao qual se atribui características ou propriedades comuns, ou seja, ao qual foi aplicada a mesma característica de divisão, forma a classe.

Pelo teste de classe, garante-se que o termo "gafanhotos" não seja indevidamente subordinado à classe de "pragas", conforme o esquema abaixo:



Podem existir, todavia, casos especiais nos quais o campo "controle de pragas" determina a subordinação de "gafanhotos" a "pragas", atendendo a objetivos muito específicos (*idem*, *ibidem*).

Conforme já mencionado, uma seqüência de conceitos subordinados forma uma seqüência vertical, enquanto que noções diferenciadas no mesmo nível de abstração formam uma seqüência horizontal, denominada coordenação.

A coordenação resulta, pois, da associação entre noções obtidas por intermédio da divisão a partir de uma mesma

característica. São, portanto, coordenadas às noções obtidas a partir de "máquina", resultante da subdivisão por tipo: máquina de moer carne, de costura, de fresar, de macarrão etc.

#### 3.1.2 Relação partitiva

A relação partitiva expressa a relação entre o todo e suas partes. É preciso observar que a relação partitiva não se confunde com a relação genérica, embora geralmente elas sejam representadas do mesmo modo.

Na relação partitiva, o conceito da parte depende do conceito do todo e não pode ser definido previamente à definição do conceito do todo. Não podemos definir "um motor de automóvel", antes de definirmos "um automóvel" (ISO 704).

As diretrizes para a elaboração de tesauros da UNESCO reconhecem quatro tipos principais de classes que representam relacionamento todo/parte: sistemas e órgãos do corpo, localidades geográficas, disciplinas ou áreas de estudo e estruturas sociais hierarquizadas (IBICT, 1984).

Os conceitos que estão em uma relação partitiva podem formar séries horizontais e verticais similares às séries horizontais e verticais formadas por relações genéricas (ISO 704), como no exemplo:

SISTEMA NERVOSO SISTEMA NERVOSO CENTRAL CÉREBRO MEDULA ESPINHAL Noção genérica partitiva Noções partitivas

(Fonte: IBICT, 1984, p. 27)

Tais relações estão presentes nos sistemas de classificação bibliográfica como a CDD e a CDU.

Também os relacionamentos enumerativos podem ser considerados como uma modalidade de relação partitiva, na medida em que indicam "a conexão existente entre uma categoria geral de objetos ou acontecimentos expressos mediante um substantivo comum e um caso individual de tal categoria, que constitui um exemplo ou classe de um só elemento, representado por um nome próprio", como em:

#### REGIÕES MONTANHOSAS

#### Andes

#### Himalaia

Fonte: (ISO 2788)

Neste caso, Andes e Himalaia são subordinados hierarquicamente, porque, mesmo que não sejam tipos nem partes de "regiões montanhosas", representam exemplos ou casos específicos do termo genérico (idem, ibidem).

#### 3.2 RELAÇÕES NÃO-HIERÁRQUICAS OU SEQÜENCIAIS

As relações sequenciais são relações que apresentam, como vimos, uma dependência resultante de uma contigüidade espacial ou temporal (Boutin-Quesnel et al., 1985), do tipo causa/efeito, antes/depois, esquerda/direita, acima/abaixo, produtor/produto, material/produto.

Tais relações podem, também, representar estágios de um processo de desenvolvimento ou de produção, procedimentos legais, procedimentos administrativos. Conceitos deste tipo, com alguma freqüência, representam ações que podem ser subdivididas em ações partitivas, tomando lugar consecutivamente ou simultaneamente (ISO 704).

A grande dificuldade para definir as relações associativas não-hierárquicas provém do fato de que todas as palavras, termos ou conceitos podem se relacionar entre si em algum momento. Isto porque as associações dependem, em larga medida, do universo de referência considerado.

As associações entre termos pertencentes a categorias diferentes são dadas a partir do universo de referência individual. Para o controle de vocabulário, entretanto, é essencial conhecer e explicitar determinados universos de referência.

Tais referências só podem estar assentadas em princípios funcionais, como reconhece Dahlberg, para quem um relacionamento funcional é "aquele em que um termo que denote atividade ou operação se liga, conceitualmente, a uma entidade ou propriedade" (apud IBICT, 1984, p. 31). Assim sendo, a delimitação das associações entre os termos deve se ligar à estrutura conceitual de domínios específicos, operacionalizada pela terminologia, na qual os conceitos deverão estar mapeados e definidos. Escapa-se, desta maneira, da virtualidade associativa passível de ser desencadeada em LN; confere-se, por outro lado, consistência aos procedimentos para a determinação das associações em domínios específicos.

Como ressalta o documento do IBICT, "não existe pesquisa suficiente para determinar as bases teóricas das relações associativas" (IBICT, 1984, p. 31). Em face desse problema, a maior parte das recomendações existentes nos manuais e normas para construção de LDs são resultantes da prática (idem, ibidem).

A experiência na elaboração de LDs permitiu enumerar vários tipos de associação, segundo a sua natureza. Entretanto, a ocorrência e utilidade desta ou daquela associação depende do modo de organização dos domínios de especialidade.

Confrontando-se as recomendações do IBICT e aquelas apresentadas por Lancaster (1987) e por Motta (1987), observa-se grande variedade de relações marcadas por diferentes pontos de vista. Abaixo estão reunidos exemplos destinados a esclarecer as complexas relações entre termos, cuja associação resulta de contigüidade temporal ou espacial:

a) Relação de Atribuição:

ECONOMIA NÍVEL DE ATIVIDADE ECONÔMICA

b) Disciplina ou campo de estudo/objetos ou fenômenos estudados:

ENTOMOLOGIA INSETOS

ESTÉTICA BELEZA

PACIFISMO PAZ

c) Processo ou operação/seu agente ou instrumento:

CONTROLE DA TEMPERATURA TERMOSTATOS

ILUMINAÇÃO LÂMPADAS

AUTOMAÇÃO COMPUTADORES

AQUECIMENTO COMBUSTÍVEIS

POLÍTICA MONETÁRIA TAXAS DE JUROS

d) Relação de Influência:

POLÍTICA MONETÁRIA INFLAÇÃO

e) Matéria-prima/produto:

BAUXITA ALUMÍNIO

f) Coisa/aplicação

ABASTECIMENTO DE ÁGUA IRRIGAÇÃO

g) Ação/resultado da ação TECELAGEM TECIDOS PINTURA (Arte) MURAIS CRESCIMENTO ECONÔMICO DESENVOLVIMENTO ECONÔMICO h) Causalidade ou causa/conseqüência: CRESCIMENTO ECONÔMICO DESENVOLVIMENTO ECONÔMICO i) Efcito/causa: INFECÇÃO FEBRE j) Dependência causal: AGENTES PATOGÊNICOS DOENÇAS PATOGÊNICAS k) Atividade/agente: TABAGISMO **FUMO** l) Atividade/propriedade: CORTE **USINABILIDADE** m) Atividades complementares: COMPRA VENDA n) Opostos: VIDA MORTE DESEMPREGO EMPREGO

o) Ação/seu paciente:

EXTRADIÇÃO **CRIMINOSOS** 

PESCA PESCADO

p) Coisa ou atividade/suas propriedades ou agentes

VENENOS TOXIDADE

CORTE USINABILIDADE

INTELIGÊNCIA CRIANÇA SUPERDOTADA

q) Coisa/seu contra-agente:

**INSETOS** 

INSETICIDAS

r) Atividade/produto:

TEAR

TECIDO

s) Pessoas ou coisas/suas origens:

BRASILEIROS

BRASIL

AUTOMÓVEL.

INDÚSTRIA

AUTOMOBILÍSTICA

t) Associação implícita:

BALANÇO DE PAGAMENTO

COMÉRCIO INTERNACIONAL

Obs.: Esta associação inclui, segundo Motta, todas aquelas que não se conformaram aos exemplos anteriormente referidos (idem, ibidem).

u) Expressões sincategoremáticas/substantivos nelas incluídos:

PEIXES FÓSSEIS

PEIXES

FLORES DE PAPEL

**FLORES** 

v) Interfaceta:

NÍVEL DE ATIVIDADE ECONÔMICA

POLÍTICA MONETÁRIA

Obs: Política Monetária (B) é associado a Nível de Atividade Econômica (A) "porque A já havia sido associado a B, previamente, pelo fato de A ser uma das características de B, e sem que B seja, necessariamente uma das características de A" (Motta, 1987, p. 49).

# 4

# Relações lingüísticas e documentação

Uma vez estabelecido um sistema nocional, existem condições para estabelecer relações entre termos. O rigor com que tais relações se propõem determina o grau de controle de uma linguagem construída. Dito de outro modo, uma linguagem construída é produto de uma operação nas palavras que as transforma em termos. De fato, a linguagem construída neutraliza as diferenças existentes na relação entre a palavra e seus significados em LN. Nela não podem coexistir, por exemplo, duas ou mais palavras que se refiram a um mesmo conceito, ou uma palavra para designar vários conceitos, sem que o fato seja suficientemente registrado, e seja devidamente controlado. Por essa razão as linguagens documentárias integram vocabulários controlados.

Para caracterizar o que vem a ser o controle do vocabulário, é preciso entender como se comporta a significação. Bakhtin (1981) observa que, no plano ideológico, a palavra é uma unidade "neutra", isto é, apta a se adequar a diferentes padrões culturais. E isso ocorre, porque ela é portadora de uma gama de significação que a torna capaz de

assumir sentidos ou valores diferentes, dependendo do contexto.

Assim, a despeito de seus semas básicos que constituem o que se poderia chamar de núcleo "duro" de significação da palavra, ela como que se amolda a cada realidade contextual, permitindo diferentes focalizações.

Desta forma, é impróprio dizer que uma dada palavra tem o significado y, embora seja viável, a partir de um significado básico, afirmar que ela assume vários sentidos ou valores, dependendo de contextos. Não é por acaso, pois, que a LN se propõe como espaço para o exercício da liberdade. O sujeito falante não é apenas um reprodutor de sentido. Ao se apossar da linguagem, ele exercita o ato de significar, que supõe liberdade de escolha.

Esta é uma das razões pelas quais a LN se define, invariavelmente, pela sua dinamicidade, já que, a cada momento, ela se transforma, evolui. É o instrumento de representação da realidade que deve ser caracterizado como múltiplo e plurissignificativo.

As unidades constitutivas das linguagens construídas, ao contrário, significam de maneira precisa. Contrapõem-se às unidades da LN, justamente por imporem significados fixos, de maneira coercitiva. Ao contrário da palavra polissêmica do vocabulário geral da LN, o termo do vocabulário especializado das linguagens construídas tende a se comportar de maneira uniforme, com pequenas variações, isto é, nele as relações entre forma significante e significado tendem a ser unívocas. Diz-se, nesse caso, que o termo, ao contrário da palavra polissêmica, é de natureza monossêmica. Em diagrama:



É preciso entender, portanto, que é intrínseco à palavra significar de maneira própria a cada ocorrência. Esse não e um defeito. É antes uma característica importantíssima para a interpretação do mundo. Não se pode exigir que a LN decline da sua função também interpretativa e criadora para exercitar apenas a função informativa. As LDs, ao contrário, porque são elaboradas para o exercício estrito da função informativa, compreendem unidades capazes de representar informação. Não é suficiente que tais unidades signifiquem. É necessário que elas signifiquem de maneira determinada.

Portanto, quando se afirma que as linguagens documentárias supõem o controle do vocabulário, afirma-se, simultaneamente:

- a) a existência de mecanismos interpretativos próprios, uma vez que não se pode utilizar o mecanismo interpretativo da LN para determinar significados das unidades destinadas à representação da informação;
- b) a possibilidade de se produzir linguagens de natureza monossêmica que participam da elaboração de LDs. Em face da natureza plurissêmica da LN, a elaboração de LDs supõe alterar a fonte de significação, isto é, alterar a possibilidade de significar, orientando-a para a necessidade de fixar significados. Este processo

- permite a transformação da unidade de significação em unidade de informação;
- c) a existência de um vocabulário próprio de uma LD que comporta, preferencialmente, unidades de linguagens de especialidade, isto é, termos, também denominados "vocabulários especializados". O vocabulário geral que se compõe de palavras, se, por um lado, é mais rico que o primeiro, por outro, do ponto de vista do tratamento da informação, é mais limitado.

#### 4.1 POLISSEMIA E AMBIGÜIDADE

Para a Lingüística, a palavra é sempre fonte de significação. Mas há que se distinguir a plurissignificação como fenômeno geral, decorrente da organização sintático-semântica de enunciados, e a polissemia, fenômeno específico da área vocabular.

A ambigüidade, por sua vez, é entendida como a possibilidade de uma comunicação lingüística prestar-se a mais de uma interpretação e ocorre em função, tanto da plurissignificação como da polissemia.

De fato, a ambigüidade pode ser consequência, na área vocabular, da polissemia ou da homonímia e, no plano mais geral, de deficiências na utilização de padrões sintáticosemânticos.

Pela polissemia, como foi mencionado anteriormente, observa-se que uma palavra pode comportar mais de um significado, como em "Hoje trabalhei muito *com* ar-condicionado", onde o enunciador tanto pode estar dizendo que tra-

balhou em aparelhos de ar-condicionado, quanto em ambiente refrigerado ou aquecido por ar-condicionado. Ou na frase "O cachorro do meu vizinho uivou a noite toda", onde ele pode estar dizendo que o cachorro pertence ao vizinho, ou que o vizinho é um cachorro.

Também a homonímia, que consiste em uma mesma forma significante remeter a duas realidades vocabulares diversas, sejam unidades com identidade fônica (homofonia) ou identidade gráfica (homografia), pode gerar ambigüidade. Por exemplo, numa frase como "O mestre entregou a cadeira ao colega", o significante "cadeira" tanto pode remeter à palavra cadeira = objeto para sentar, quanto a cadeira = cátedra de um docente.

Pelas deficiências no uso de padrões sintáticos, evidenciam-se também ambigüidades, geralmente, resolvidas em LN com modificações de colocação, como em "Os juízes encaravam os réus enigmáticos", onde tanto a significação pode ser relativa à atitude dos juízes, quanto ao estado dos réus. A colocação dos sintagmas, ou a seleção de padrões sintáticos pode, entretanto, desambigüizar a frase: "Os juízes enigmáticos encaravam os réus" ou "Os juízes encaravam os réus que eram (estavam) enigmáticos.

Numa linguagem documentária, tanto a polissemia, quanto a ambigüidade devem ser neutralizadas, para que seja garantida a monossemia entre a forma do significante e a do significado.

A ambigüidade evidencia, de maneira inequívoca, a divergência entre a aparência e a realidade do sistema e nos permite dizer que a aparência não é sempre a pista interpretativa mais segura. Levados pela aparência, operamos,

normalmente, com os mecanismos interpretativos habituais e nos acostumamos com significados repetitivos. Isso diz respeito aos nossos hábitos e não ao sistema lingüístico.

Por estarem num sistema relacional, as palavras devem ser observadas em oposição umas às outras. Em si mesma, por exemplo, a palavra "alta" pode ser incorretamente interpretada como ambígua, já que pode estar associada a significações diferentes, como: criança alta e mulher alta. O mesmo pode-se afirmar em relação à palavra "baixo", uma vez que "criança baixa" e "mulher baixa" apresentam igualmente significações diversas. A ambigüidade das palavras inexiste se as observamos como oposição. Fica evidente, desse modo, que "criança baixa"/"criança alta" é uma oposição análoga à oposição "mulher baixa"/"mulher alta". A significação, nesse caso, denomina-se oposicional e possibilita determinar o sentido, propondo limites para a indeterminação original.

Estamos diante, então, de dois fenômenos que devem ser objeto das operações de elaboração de linguagens documentárias: a polissemia e a monossemia. A polissemia é responsável pela passagem de uma significação a outra, de modo que as unidades sejam capazes de representar a informação. A informação, ao contrário da significação geral, deve ser determinada. Para que ela o seja, a significação que a representa não pode ser de natureza polissêmica. A monossemia, por sua vez, desejável nas LDs, é obtida por meio de redes relacionais e definições dos termos. Isto quer dizer que, ao contrário da LN, onde a riqueza vincula-se à polissemia, a fixação de relações e definições precisas é seu princípio organizador elementar e básico.

Assim, ao operar com LDs, devemos analisá-las, tendo em vista desvendar o modo pelo qual nelas as significações são organizadas.

A rigor, não se deseja que um termo se enriqueça. Exige-se que ele expresse conceitos determinados. A definição deve propor uma expressão (sintagma ou palavra) semanticamente equivalente à unidade a ser definida. Não se deve descrever, por exemplo, o objeto concreto ferro ou água, mas o funcionamento lingüístico do termo num sistema nocional em questão, tal como Fe e H<sub>2</sub>O, respectivamente, para o vocabulário da Química.

O termo, também, define-se por suas relações com outros termos. Extraindo o termo do lugar que ocupa, o qual lhe confere seu valor, privamo-nos do único meio possível para definir sua existência lingüística, rigorosa o suficiente para garantir seu funcionamento como unidade de informação.

Sendo assim, Fe e  $\rm H_2O$  passam a ter significados fixados e determinados. Integram um vocabulário especializado (técnico ou científico). Seus correlatos ferro e água integram o vocabulário geral da LN, no qual podem assumir significações diversas. Por exemplo, ferro, em relação ao objeto, a conceito etc.

Por vezes, observa-se confusão entre ambigüidade e polissemia. A ambigüidade lexical impõe-se por meio da polissemia e da homonímia. Na linguagem documentária, a ambigüidade é tratada com o auxílio de modificadores que contextualizam o sentido. Ex.: planta (botânica), planta (arquitetura); companhia (empresa), companhia (pessoa).

Em princípio, em LN, a ambigüidade é facilmente resolvida pelo contexto. O mesmo não ocorre com a polissemia.

A visão ingênua que identifica ambigüidade e polissemia, acaba por acreditar que apenas a ambigüidade leva à indeterminação do sentido. Ela é, de fato, o fenômeno mais aparente e o menos grave. A armadilha é acreditar que a palavra tenha um único significado. Nega-se a polissemia como fenômeno global e estabelecem-se operadores de sentido que pouco têm a ver com o campo nocional, isto é, substitui-se o conceito ou a noção própria dos vocabulários especializados pelas indeterminações do vocabulário geral.

Para neutralizar a polissemia, é preciso lançar mão de dois recursos: elaboração de redes relacionais e estabelecimento de definições e notas de escopo, sempre que as redes se mostrarem insuficientes para a interpretação unívoca da significação. Tais recursos impõem operadores de sentido, isto é, elementos que conduzem o indexador a interpretar adequadamente, em conformidade com o sistema nocional em questão.

#### 4.2 SINONÍMIA

A sinonímia é uma relação de equivalência entre, ao menos, duas palavras. Por meio dela não se afirma a identidade entre os elementos envolvidos na relação. Isto é, x equivale a y indica que x pode, em determinadas circunstâncias, substituir y. A equivalência é um recurso normalizador importante para a compreensão de uma linguagem documentária. De um lado, permite normalizar a polissemia, indicando que várias palavras, uma vez que compartilham significados próximos, expressam-se por um mesmo descritor. De outro,

permite compatibilizar a linguagem dos usuários com a linguagem do sistema, funcionando, assim, como operador de sentido.

É importante entendê-la sempre como consequiência do contexto. Este fator caracteriza a equivalência como uma operação relativamente arbitrária, mas isso é pouco importante, uma vez que a arbitrariedade esteja registrada.

De fato, a transformação da unidade de significação em unidade de informação é a característica fundamental do controle de vocabulário, já que numa linguagem construída, a cada unidade de informação deve corresponder um único sentido referencial.

No entanto, a existência de sinônimos ou quase sinônimos nos leva a considerar relações de equivalência para o trabalho documentário.

A grande importância das relações de equivalência advém do fato que elas intensificam o processo de controle sobre a variação de significado, permitindo maior rigor no tratamento da informação e eficácia na sua recuperação. Como os outros gêneros de relações mencionadas anteriormente, as relações de equivalência introduzem parâmetros para o uso da linguagem determinadas por um grupo.

No sentido estrito, a sinonímia pode ser definida como identidade de significação entre elementos lexicais, porém, a existência de sinonímia absoluta é controversa, sendo causa de debates entre lexicólogos. Alguns autores admitem sua existência para o caso da equivalência entre duas línguas funcionais, como em gaivotas – nome popular/Larus – nome científico; outros, ao contrário, tratam tais equivalências como quase-sinonímia.

Entre lingüistas é mais frequente a aceitação do conceito de quase-sinônimo, ou de para-sinônimo, uma vez que parece muito pouco provável que, em LN, duas palavras portadoras de exatamente o mesmo significado possam sobreviver.

Entre documentalistas, também, são utilizados os conceitos de sinônimo e quase-sinônimo. Enquanto sinônimo indica cada um dos termos de uma língua dada que designam uma mesma noção, mas que se situam em níveis da língua ou de conceptualização diferentes, ou que se empregam em situações de comunicação diferentes; quase-sinônimos designam formas que não são intercambiáveis em todos os enunciados relativos a um mesmo domínio.

A variada gama de quasc-sinônimos, talvez, possa ser resumida em alguns tipos:

- palavras pertencentes a dialetos diferentes (dialetos regionais, sociais, etários etc.), como pesquisa (Brasil)/Investigação (Portugal); avião/aeroplano;
- palavras pertencentes a diferentes estilos ou registros, como dor de cabeça/cefaléia; gaivotas/larídeos; ácido clorídrico (química)/ácido muriático (construção civil);
- palavras que guardam apenas uma diferença emotiva ou valorativa, como países em vias de desenvolvimento/países subdesenvolvidos;
- palavras que têm sua ocorrência limitada, na medida em que só aparecem com outras, como "de barbear" que vem com lâminas: giletes/lâminas de barbear;
- palavras cujos significados são, de fato, muito pró-

ximos e se intersectam, como belo/bonito; casa/residência; falecimento/morte.

Na elaboração de LDs é fundamental um trabalho específico com sinônimos e quase-sinônimos, uma vez que essas linguagens têm por função compatibilizar pelo menos duas outras linguagens: a de especialidade ou da literatura em questão e a do usuário, por meio de termos preferenciais.

Numa acepção mais ampla, como é o caso da sinonímia utilizada na elaboração de tesauros, dois termos são sinônimos quando têm a possibilidade funcional de se substituírem um ao outro, podendo compreender tanto a sinonímia absoluta como a quase-sinonímia. A sinonímia nas LDs é de caráter eminentemente preferencial e visa remeter o usuário de um termo não-preferencial, para um termo selecionado, ou preferencial.

#### 4.3 HIPONÍMIA

Do ponto de vista da Lingüística, a estruturação hierárquica de um vocabulário pode ser dada sob dois modos: por uma relação de hiponímia ou por meio da relação parte/todo.

No nível das relações de sentido o problema da significação pode ser visto sob diversos ângulos, ou seja, a partir de diversas categorias.

A categoria denominada hiponímia opera com a noção de inclusão, a mesma noção que permite reunir unidades numa classe. Assim, rosa e cravo estão incluídas em flor, ou gato e leão estão incluídos em animal, ou escarlate está incluído em vermelho.

A inclusão tem a ver, pois, com a inserção de um dado elemento numa classe. Isso dito de outra forma, indica que a hiponímia expressa "a relação existente entre um lexema mais específico ou subordinado, e um lexema mais geral ou superordenado, tal como é exemplificada por pares como 'vaca': 'animal', 'rosa': 'flor' etc." (Lyons, 1977, p. 235).

Nesta relação há que se considerar dois termos: o superior, denominado por Lyons (1977) Superordenado, e o inferior, Hipônimo.

Os termos constitutivos de uma classe são, pois, cohipônimos. Entretanto, é necessário observar que nem toda classe dispõe de um superordenado. E mais: a existência de um superordenado encabeçando uma classe pode variar de língua para língua.

Lyons menciona a existência, em grego clássico, de uma forma superordenada para abranger todas as profissões e ofícios, desde sapateiro, médico, passando por tocador de flauta e timoneiro. Em inglês e em português não há palavra que possa encabeçar conjunto tão variado. Neste caso, temse uma lacuna lexical.

A hiponímia pode ser definida, também, em termos de implicação unilateral e representa uma relação transitiva, de tal modo que, se 'x' é hipônimo de 'y' e 'y' é hipônimo de 'z', então 'x' é hipônimo de 'z'.

A hiponímia é, ainda, uma propositura analítica, sendo que a leitura e compreensão do significado dos hipônimos

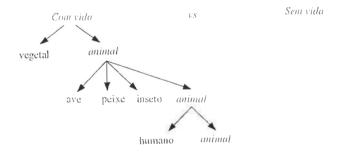
podem ser feitos segundo a fórmula 'x é uma espécie (ou tipo) de y': o gato é uma espécie de animal.

A relação de hiponímia/hiperonímia (ou subordinação/ superordenação) permite verificar que um termo pertence, ou subordina-se a um outro mais geral, o gênero, mas não permite identificar em que os termos subordinados se diferenciam entre si.

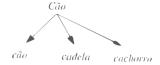
Por outro lado, em virtude da polissemia, uma mesma palavra pode aparecer em vários pontos da hierarquia. Palmer (1976) menciona como exemplo a palavra *animal* que pode ser usada em três pontos da cadeia:

- 1. em contraste com "vegetal", incluindo, neste caso aves, peixes, insetos, mamíferos;
- 2. no sentido de "mamífero", contrapondo-se a aves, peixes e insetos, mas incluindo seres humanos e bichos;
- 3. no sentido de "bicho", opondo-se a seres humanos.

Assim, a palavra *animal* poderá surgir três vezes na classificação hierárquica da natureza, como mostra Palmer (p. 92):



Os exemplos poderiam ser multiplicados. No entanto, parece ser suficiente levar em conta que, em razão da polissemia, um termo como *cão* pode, por ser o genérico da classe, ser tomado como superordenado e como hipônimo, respeitadas, naturalmente, as situações contextuais.



Resta observar que além da noção de inclusão, a hiponímia contém implícita, também, a relação lógica de consequência, já que ao dizer "Isto é uma rosa", tem-se, necessariamente, o pressuposto "Isto é uma flor". Ou seja, a frase que contém o hipônimo pressupõe, necessariamente, o superordenado. O inverso, evidentemente, não é verdadeiro.

Se os membros de uma classe são especificados com "todos", ocorre o inverso: "Todas as flores são belas" inclui "Todas as rosas são belas", mas o inverso não é verdadeiro.

Pode-se dizer que a relação de hiponímia representa uma operação de conjunção em face do termo superordenado, bem como de disjunção, tomando-se a série de termos obtidos a partir da divisão realizada.

Como na LN, nas LDs a superordenação/subordinação representa um caso de implicação unilateral, onde o termo superordenado implica termos subordinados, denominados hipônimos.

Em termos do léxico, o sentido de um hipônimo é produto do sentido de um nome superordenado e de um modificador adjetival real ou potencial, que responde a perguntas do seguinte tipo: 'que espécie de...?'; 'que tipo de...?'.

Por exemplo: - "Que espécie de animal era?"

- "Era um elefante."

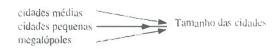
Dito de outro modo, a resposta a perguntas desse gênero – e outras similares, do tipo 'como ...?', 'de que maneira ...?' – se dão a partir da introdução de uma diferença, que produz as subclasses.

Os co-hipônimos – ou os termos coordenados que formam uma mesma série – contrastam em sentido, sendo que a natureza do contraste pode ser explicada em termos de diferentes modificações adjetivais (Lyons, 1977).

Pode-se dizer que as modificações adjetivais no léxico correspondem, nas LDs, a características ou propriedades que realizam a individuação de termos. Do ponto do vista extensional, os termos que se subordinam a um superordenado contêm todas as características que identificam a classe, mais uma que os distingue dos demais.

A relação de hiponímia colocada pela lingüística permite explicar, nas LDs, vários tipos de relacionamentos tomados como hierárquicos que não cabem dentro da classificação gênero/espécie (e tão pouco nas relações todo/parte, parte/parte)\*

Há casos, por exemplo, em que dois ou mais termos encontram-se em contraste e não existem, no léxico, palavras (ou termos, no caso das LDs) que lhes sejam superordenados e, a não ser que se utilizem elementos de natureza diferente, provenientes de outras partes do discurso, não é possível reuni-los.



Este é um caso de relação quase-paradigmática (ou uma quase-hiponímia), uma vez que se utiliza uma expressão mais geral ("tamanho") para reunir os diferentes tipos de cidade.

Aqui, entretanto, não é válida a aplicação da fórmula "x é um gênero de y". A frase obtida da sua aplicação não é natural ou é inaceitável. A estruturação do vocabulário, neste caso, é feita por outras palavras ou sintagmas que desempenham o mesmo papel de "que gênero de ..." (ou que espécie de ...).

"Comparáveis às questões Que gênero de animal era?, e Era uma vaca ou outra espécie de animal?, são Como é que ele obteve isso — comprando-o ou roubando-o? e Ele comprou isso ou arranjou-o de algum outro modo?" (Lyons, 1977, p.237) … ou ainda, no caso de adjetivos "Quando dizes que o teu vestido é carmim, queres dizer que é em tons de vermelho ou de outra cor? Assim como podemos dizer A vaca é um animal de um certo gênero, também podemos dizer ...Comprar alguma coisa é obtê-la de uma determinada maneira e Um objeto carmim é um objeto vermelho de uma certa maneira" (idem, ibidem).

Em resumo, responde-se, nestes casos, a perguntas do tipo "como" e "de que maneira", muito embora elas não possam, também, ser amplamente empregadas com sucesso. A hiponímia, na verdade, pode manifestar-se de muitas maneiras.

Isto explica porque não é possível aplicar, muitas vezes, o esquema lógico "todos/alguns" sugerido pelos manuais de elaboração de vocabulários documentários.



sugere o esquema como meio para a validação de um relacionamento genérico, indicando que alguns membros da classe "insetos" são conhecidos como "gafanhotos", enquanto que todos os "gafanhotos" são "insetos", por definição e independentemente do contexto (IBICT, 1984, p. 26; ISO 2788, 1986, 1989, p. 605).

Entretanto, ele não funcionaria no exemplo anterior, relativo a 'cidades'. A língua, na verdade, não é rigidamente estruturada em termos lógicos.

As diferentes séries formadas a partir de um mesmo termo podem ser vistas como o resultado de diferentes modos de realizar a conjunção, oriunda dos diferentes pontos tomados como origem da subdivisão e/ou das diferentes características tomadas para a construção de cada hierarquia.

Esse aspecto relaciona-se com a adoção das categorias aristotélicas de predicação – substância, modo, quantidade, qualidade etc. e suas atualizações nos seus desenvolvimentos subseqüentes. Em Documentação, por exemplo, Ranganathan utilizou cinco categorias para agrupar os assuntos: personalidade, matéria, energia, espaço e tempo.

Para a área de Ciência e Tecnologia, o *Classification* Research Group – CRG sugeriu que os termos fossem agrupados segundo as seguintes categorias fundamentais: substância (produto), órgão, constituinte, estrutura, forma, propriedade, objeto da ação (materiais brutos, materiais não tratados), ação, operação, processos, agente, espaço e tempo. Barbara Kyle, também integrante do CRG, distinguiu as artes, as atividades, os objetivos, os objetos, as idéias, as abstrações.

Ainda na área de Documentação, Shera & Egan (1969) propuseram as categorias agente, ação, modo, objeto, objeto de ação, tempo, espaço e produto. Grolier (1962), por seu lado, sugeriu categorias constantes de tempo, espaço, ação e categorias variáveis: substância, órgão, analítico, sintético, propriedade, forma e organização.

Na área da Lingüística vale ressaltar os "casos conceptuais" de Pottier (1974): causativo, instrumental, agentivo, nominativo, ergativo, acusativo, associativo, locativo, dativo, beneficiativo, finalidade.

De um modo ou outro, todas essas noções ou facetas remontam às classificações aristotélica e kantiana. A estruturação do vocabulário em áreas distintas definirá as noções funcionais mais generalizantes a serem adotadas. Por outro lado, tal estruturação, dada em função de relações de hiponímia e quase-hiponímia, pode ser realizada por meio de um pequeno número de lexemas (noções generalizantes, categorias, facetas) com sentido muito geral.

Pode-se afirmar, com Lyons, que nem sempre é possível estruturar hierarquicamente os lexemas em termos de hiponímia, dada a ausência de lexicalização, em algumas línguas. Não há, em português, por exemplo, nenhum lexema que seja superordenado a todos os nomes abstratos, ou a todos os nomes concretos. O que se encontra, ao contrário,

são conjuntos de lexemas muito gerais — "'pessoa' (ou 'indivíduo'), 'animal', 'peixe', 'ave', 'inseto', 'coisa', 'lugar', 'substância', 'matéria', 'qualidade', 'estado' etc. — que são superordenados em relação a subconjuntos maiores ou menores destas subclasses de nomes" (Lyons, 1977).

Em resumo, para o autor não existe ordenação hierárquica a partir de um lexema superordenado único, fato que se estende a diversas partes do discurso, pois, além dos nomes, isso se aplica aos verbos e aos adjetivos.

Entretanto, se nos casos de hiponímia pode-se afirmar que existe uma relação paradigmática de sentido entre os lexemas, na ausência de superordenados paradigmáticos para a reunião de lexemas pode ocorrer uma relação quase-paradigmática (idem, ibidem).

Assim, para reunir os adjetivos 'vermelho', 'amarelo', 'azul' etc., pode-se utilizar: 'cor'; para falar de 'redondo', quadrado', 'oblongo': 'forma'.

Mas há casos em que não há, no vocabulário, lexemas para organizar, hierarquicamente, os termos. Trata-se das "lacunas lexicais", devidas, na maior parte das vezes, a fatores culturais.

Uma lacuna lexical pode ser descrita como um "buraco no modelo", ou seja, a ausência de um lexema num dado lugar da estrutura de um campo lexical" (Lyons, 1977)

A relação hiponímia/superordenação corresponde, em lógica, à relação gênero/espécie (ou espécie/gênero). O conjunto desse tipo de relacionamento é denominado, via de regra, relacionamento genérico.

# Bibliografia

AITCHISON, J. & GILCHRIST, A. Manual para construção de tesauros. Rio de Janeiro: BNG/Brasilart, 1979.

BAKHTIN, M. Marxismo e filosofia da linguagem. São Paulo: Hucitec, 1981.

BARANOW, V. G. Aspectos lingüísticos de linguagens de indexação. In: CONFERÊNCIA BRASILEIRA DE CLASSIFICAÇÃO BIBLIOGRÁFICA. Anais. Rio de Janeiro, IBICT, v.1, p.295-310, 1979.

BARTHES, R. Elementos de semiologia. São Paulo: Cultrix, 1971.

BOUTIN-QUESNEL, R<sub>4</sub> et al. *Vocabulaire systematique de terminologie*, Québec: Publications du Québec, 1985.

BRASIL. Ministério da Educação e Cultura. Serviço de Estatística da Educação. *Glossário de termos utilizados na estatística educacional*. *I*. Rio de Janeiro: Fundação MUDES, 1980.

CDU – CLASSIFICAÇÃO DECIMAL UNIVERSAL. Ed. média em língua portuguesa. Brasília: IBICT 2v. (Publicação FID, n.665), 1987

CINTRA, A. M. M.; TÁLAMO, M. F. G. M.; LARA, M. L. G. & KOBASHI, N. Y. Do termo ao descritor: um estudo explora-