BIBLIOTECA
DO
IBICT

Tesauros

Linguagens terminológicas

19106
ex.3
24.10.96
CCI

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA
IDI BIBLIOTECA
FORNECEDOR:
FORMA AQUISIÇÃO: Doação
VALOR: R$ 1,00
TOMBO: 19106
DATA: 24.10.96
PHL: 036566

Emilia Currás

Tesauros

Tradução de Antônio Felipe Corrêa da Costa

Brasília, 1995

025.43
C976t
1995
ex.3
CDU

CNPq - Conselho Nacional de Desenvolvimento Científico e Tecnológico
IBICT - Instituto Brasileiro de Informação em Ciência e Tecnologia

© Emília Currás Puente
© Editorial Paraninfo
ISBN 85-7013-052-X
Magallanes, 25 - 28.015 Madrid

Título original : Thesauros – Lenguajes Terminologicos
Editorial Paraninfo, 1991
Autoria : Emilia Currás Puente. Baseado na obra Thesaurus Construction, de Jean Aitchison e Alan Gilchrist, publicado pela Aslib (Inglaterra).

Direitos desta edição cedidos ao Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT).
É proibida a reprodução de qualquer parte desta obra sem a prévia autorização do IBICT.

© IBICT 1995
Coordenação editorial : Margaret de Palermo
Projeto gráfico : Nair Costa Barreto
Revisão e normalização : Francisco de Paula, Margaret de Palermo e Edith Lima
Composição eletrônica : Cleber da Costa Matos
Editoração eletrônica: Rogério Anderson e Heloisa Neves
Capa: Carlos T. D. Brasil
Arte-final das figuras : Cícero Freitas dos Reis
Tradução : Antônio Felipe Corrêa da Costa

Currás, Emilia
Tesauros, linguagens terminológicas/ Emilia Currás; tradução de Antônio Felipe Corrêa da Costa._Brasília:IBICT, 1995
286 p.
Tradução de Thesauros Lenguajes Terminologicos.
ISBN 84-283-1825-5

1.Tesauro-linguagem de indexação - organização.I.Costa, Antônio Felipe Corrêa da, trad. II. Título.

ISBN 85-7013-052-X CDU 025.43

Primeira edição espanhola : 1991
Tradução para português : 1995

Instituto Brasileiro de Informação em Ciência e Tecnologia – IBICT
SAS, Quadra 5, Lote 6, Bloco H
CEP 70070-914 Brasília, DF
Tel.:(061) 217 6161 - Telex 2481 CICT BR - Fax (061) 226 2677

Impresso no Brasil

No silêncio e no recôndito do pensamento, nasce a claridade da compreensão quando a palavra se transforma em termo.

Emilia Currás

Nota do tradutor

O meu agradecimento a Milton Nocetti pelo prestimoso auxílio prestado na revisão de alguns termos em espanhol.

Sumário

Apresentação — 11
Preliminar — 13

Capítulo 1. A terminologia — 15

Liguagem e linguagens — 15
Importância da terminologia — 18
Definições de terminologia — 21
O conceito e os termos — 23
Terminografia — 30
Os terminólogos — 31
Temas para estudo — 34
Referências bibliográficas — 35

Capítulo 2. Evolução histórica — 37

Períodos evolutivos — 37
Período clássico — 38
Período de iniciação à terminologia moderna — 40
Período de consolidação — 41
Período de teorização — 42
Período de evolução acelerada — 44
A terminologia na Espanha — 47
Temas para estudo — 49
Referências bibliográficas — 50

Capítulo 3. As indústrias da língua — 51

Aplicação das novas tecnologias — 54
Bancos de dados terminológicos — 55
Tipos de bancos de dados terminológicos — 56
Construção de um banco de dados terminológicos — 58
A tradução automática — 61
A normalização nas indústrias da língua — 65
Instituições de normalização — 65
Temas para estudo — 69
Referências bibliográficas — 70

Capítulo 4. Linguagens terminológicas. Tesauros 77

Classes de linguagens terminológicas 78
Os tesauros 83
Definições de tesauros 84
Definições atuais de tesauros 86
Condições a serem cumpridas por um tesauro 88
Evolução histórica 90
A evolução histórica na Espanha 93
Classes de tesauros 95
Temas para estudo 98
Referências bibliográficas 99

Capítulo 5. A estrutura dos tesauros 101

Componentes fundamentais de um tesauro 102
Classes de termos 103
Tipos de relações entre termos 105
Notas de aplicação 106
Notações utilizadas para designar as relações entre termos 108
Exemplos de relações entre termos com sua notação correspondente 110
Formas de termos que compõem um tesauro 111
Termos simples componentes de um tesauro 113
Os termos compostos 117
A utilização da forma singular e plural 119
O tratamento dos números 124
Temas para estudo 127
Referências bibliográficas 128

Capítulo 6. Construção de tesauros 129

Quando faz falta um tesauro novo 130
Considerações que devem ser levadas em conta 131
Condições gerais que um tesauro deve cumprir 135
A apreensão e fixação do termo 137
Termos simples ou termos compostos 140
Estabelecimento das relações entre termos 144
Relações associativas 155
Temas para estudo 162
Referências bibliográficas 163

Capítulo 7. Formas de apresentação de tesauros — 165

Apresentação alfabética de tesauros — 165
Apresentação sistemática de tesauros — 167
Apresentação gráfica de tesauros — 170
Temas para estudo — 182
Referências bibliográficas — 183

Capítulo 8. Os tesauros multilíngües — 185

Definições de tesauro multilíngüe — 186
Peculiaridades dos tesauros multilíngües — 187
A estruturação — 189
Relação de semelhança entre os termos — 190
Técnicas de construção — 194
Apresentação de tesauros multilíngües — 195
Anexo: Confecação de um tesauro estruturado para química — 199
Temas para estudo — 208
Referências bibliográficas — 209

Capítulo 9. Da publicação e outras questões complementares — 211

Integração de tesauros — 211
Desdobramento de tesauros — 216
A revisão e a manutenção — 218
A prática da comprovação — 221
O momento da redação e a publicação — 222
Pacotes informatizados — 227
Temas para estudo — 230
Referências bibliográficas — 231

Capítulo 10. O uso dos tesauros — 233

O tesauro na indexação de documentos — 233
Técnicas de indexação — 235
O truncamento dos termos — 238
Uso de operadores de proximidade e distância — 240
Índice de ponderação — 241
Indicadores de ligação — 241
Indicadores de função — 243
Linguagem de indexação — 246
O tesauro na recuperação da informação — 247
Técnicas de recuperação da informação — 248
Linguagem de recuperação da informação — 249

Diferenças entre uma linguagem de indexação e outra de recuperação da informação — 250
Temas para estudo — 252
Referências bibliográficas — 253

Epílogo — 255
Índice de ilustrações — 265
Índice de autores — 267
Índice de assunto — 277

Apresentação

O processo de indexação compreende duas etapas básicas: análise do documento para indexação dos conceitos-chave nele contidos e tradução destes para a linguagem do sistema de informação/biblioteca. No caso, a linguagem pode ser um vocabulário linear ou estruturado, isto é, um tesauro.

No Brasil, grande número de sistemas especializados de informação/bibliotecas utiliza tesauros no processo de indexação. Por conseguinte, manuais para construção desses valiosos instrumentos são obras de relevante utilidade, tanto prática, quanto teórica. A tradução do original em espanhol do livro *Tesauros - Linguagens Terminológicas* para a língua portuguesa significa, portanto, frutífero avanço em prol da difusão do conhecimento sistematizado no Brasil. Com esse empreendimento, o Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) objetiva contribuir para o desenvolvimento da área de indexação (*lato sensu*) no país e, concomitantemente, estimular a produção bibliográfica de autores nacionais.

Na obra de autoria da doutora Emilia Currás, professora titular da Faculdade de Ciências da Universidade de Madri, o profissional da informação da área científica encontra instrumentos não só para atualizar seus conhecimentos, mas também para tornar mais afiadas suas ferramentas de indexação, quando delas precisar se valer. Nela, o profissional e aquele que se prepara com vistas a tornar-se encontrarão a doutrina justa e a clara exposição das questões básicas da terminologia – reunião e fixação dos termos, tradução automática e indústrias culturais da língua – como etapas distintas, uma vez que a terminologia possui também os seus próprios fins, ao mesmo tempo em que atende à estruturação de "linguagens documentárias".

Este livro apresenta-se muito bem fundamentado e, agora, também auxiliará o estudante brasileiro a fixar conceitos, uniformizar termos e livrar-se de muitas incertezas.

Parece fácil, pois, que o leitor concorde com a pertinência de uma obra como a presente, meditada durante longo tempo e desenvolvida pela autora à luz de sua vasta experiência, para tornar cristalinos aspectos do trabalho do documentalista que ainda se encontram turvos, à sombra da teoria, ou mesmo marginalizados, no momento de tesaurizar conceitos.

Ademais, a obra demonstra a forte inclinação da autora à filosofia e sua tendência a comprovar o exposto com exemplos emprestados da química – ciência que foi o berço da documentação.

Enfim, com a publicação deste precioso livro, o IBICT mostra, de fato, que está mesmo empenhado em promover o desenvolvimento científico e tecnológico do país.

José Rincon Ferreira
Diretor do IBICT

Preliminar

Vivemos, atualmente, com a atenção voltada para o mundo como um todo, ainda que heterogêneo e complexo nem por isso menos interessante e digno de nossa atenção. Na época atual – talvez não muito diferente da chamada Idade Antiga em nossa cultura ocidental, embora com a diferença lógica no âmbito geográfico –, tendemos à universalização e à internacionalização das formas de pensamento e de governo, medidas pela atividade rápida e surpreendente de cada dia. Não podemos viver isolados, os comportamentos estanques desapareceram, os conceitos e as idéias se ampliaram, modificando-se e transformando-se. Novos instrumentos surgiram, cada vez mais complexos em sua construção e funcionamento, mas que, não obstante, apresentam-se com uso cada vez mais simples.

No mundo atual, devemos nos relacionar uns com os outros. É um dever dificilmente iludível, sob pena de ficarmos isolados em determinado lugar e definharmos pouco a pouco, no isolamento que provoca um profundo desgosto em nós mesmos. Por analogia, poderíamos comparar esse fato a um homem jogado no mar, o qual, caso não procure nadar a favor das ondas, sucumbirá irremediavelmente.

Por outro lado, parece que a necessidade de relacionamento com seus semelhantes é tão importante para o desenvolvimento dos seres humanos, que as pessoas que vivem sós escutam diariamente o rádio durante muitas horas e falam ou cantam sozinhas, tudo isso somente para ouvir vozes humanas.

Para estabelecer tais relações, a comunicação entre os seres humanos, contamos com dois esteios importantes. O primeiro é fundamental e o outro, a sua conseqüência: a linguagem e a informação.

Nestes momentos de grande evolução da humanidade, surgem continuamente novas formas de expressão, novas idéias, teorias, novos instrumentos e artifícios. A linguagem, pois, também evolui no mesmo ritmo, e temos que procurar, ou pretender ao menos, controlá-lo, e estabelecer as relações oportunas entre os diferentes idiomas, para não cair novamente em uma Torre de Babel. De onde aparece a grande importância que a terminologia, a tradução e a normalização possuem atualmente. E, também, a necessidade de que se apliquem estas novas e enigmáticas máquinas automáticas, com a intenção de emular o funcionamento da inteligência humana.

Por sua vez, é a informação que nos aproxima do mundo em que vivemos e evoluímos a nosso mundo interior. A informação, entendida como fenômeno independente de nós mesmos

e como processo, é elaborada e solicitada para desenvolver nossas atividades cotidianas.

A peça mais importante nessa elaboração, armazenamento e recuperação da informação é a terminologia. O estudo das linguagens naturais e artificiais. Esta é a causa pela qual nesta obra é abordado o estudo da terminologia e das linguagens terminológicas em conjunto.

Consultando a bibliografia e os manuais especializados, pode-se observar que todo o assunto das linguagens controladas, de tesauros, de indexação, de linguagens de indexação, das relações entre os termos e tudo o mais que se costuma chamar de "Teoria de Tesauros" não foi entendido muito claramente. Talvez isso se deva à sua relativa novidade e contínua evolução. Talvez também a certas traduções de um idioma a outro, de difícil interpretação conceitual. Eis aqui, novamente, a questão da terminologia e da tradução.

Nesta obra, não pretendemos ser exaustivos no que se refere ao tratamento da terminologia. A esta foi feita referência como base e apoio à teoria de tesauros. Procuramos sistematizar e esquematizar determinados conceitos, idéias e teorias, reunindo opiniões de autores autorizados nessas questões. Tudo isso foi feito com o espírito de compendiar os princípios mais fundamentais. Levou-se em consideração obras clássicas, como são as de E. Wüster, Rolling, H. Felber, Galinski, Ingetraut Dahlberg, Gernot Wersig... dentre os mais destacados especialistas de nossa época. Na bibliografia, foram citadas obras que tratam desses assuntos com profundidade, pensando-se naqueles que desejam lhes dar maior atenção.

No que se refere ao assunto das linguagens terminológicas e, de modo mais concreto, aos tesauros, procurou-se, do mesmo modo, reunir opiniões e teorias de outros autores, normas e vários tipos de documentos, com o propósito de esclarecer idéias. Procurou-se igualmente explicar, também de modo coerente e esquematizado, as questões práticas concernentes à construção e utilização de tesauros. Ficamos animados a proceder dessa forma pela freqüência com que temos sido consultados a respeito das formas de trabalhar e dos métodos de trabalho. Nota-se um vazio na bibliografia em espanhol sobre o assunto. Existe alguma obra antiga, bem como capítulos em determinados manuais e várias monografias. Todas essas publicações são baseadas, contudo, nas mesmas obras inglesas e americanas, as quais – por sua vez – falam para um público especializado, ou seja, para profissionais com conhecimentos básicos. Por outro lado, as interpretações, traduções ou transferências para o espanhol não são de compreensão fácil para os que possuem um conhecimento menor sobre o assunto. Esta é uma das razões pelas quais, nesta obra, foi abordado, separadamente, o que é um tesauro, os elementos que o compõem, suas relações, sua estrutura, os tipos de tesauros etc. Em outros capítulos, foram estudadas as formas de construção de tesauros, o seu uso e utilização.

Desejamos que esta obra cumpra a sua missão, tornando-se útil, prática e bem acolhida por todos aqueles que se interessam pelos assuntos aqui tratados.

BIBLIOTECA
DO
IBICT

1

A Terminologia

A terminologia adquiriu tanta importância nas últimas décadas, que se tornou impossível prescindir dela, não obstante em circunstâncias nas quais parece totalmente independente e isolada do contexto dos termos tratados. Nas atividades mais gerais e corriqueiras da vida cotidiana, devemos nos expressar com propriedade e fazer uso das palavras dentro de seu sentido correto, para podermos compreender e nos fazer compreender. E isso mais ocorrerá nas atividades profissionais, pois a compreensão e o entendimento entre os especialistas é a base das boas relações e índice de um progresso em evolução contínua.

Para a finalidade desta obra, interessa fazer referência à terminologia como fundamento e ponto de partida de algumas argumentações lógicas, dedutivas, que desejamos que nos conduzam desde um primeiro processo até uma completa elaboração de sistemas de idéias, por onde o ser humano discorre na execução das atividades, tanto cotidiano-rotineiras como profissionais-especializadas. O ser humano é um indivíduo social. Foi feito para viver em comunidade, relacionando-se com seus semelhantes, com os que tem, infalivelmente, que entender-se. A frase "estamos condenados a nos entender" já é de uso geral e ouvida freqüentemente, o que está rigorosamente certo. Porém, em um sentido menos dramático, poder-se-ia substituir a palavra "condenados" por "destinados". Ou, possivelmente, não se melhora o propósito da expressão. Tudo depende do que podemos entender por "destino" e sua conotação sociomoral.

Com este simples raciocínio, mediante o qual procuramos manifestar que o homem precisa relacionar-se com seus semelhantes, já surgiram disparidades nas palavras utilizadas. E não somente nas próprias palavras com expressão de um sentimento ou idéia. Evidencia-se a complexidade da linguagem (figura 1.1).

LINGUAGEM E LINGUAGENS

Deveríamos dizer linguagem ou linguagens? O debate deve ser centralizado na idéia que se tenha de linguagem ou linguagens. A bibliografia sobre estes assuntos, tanto retrospectiva como recente, é numerosa. Nela se observa uma evolução desde um sentido ontológico, até um pragmatismo técnico prático. Para A. Martinet, a linguagem é um simples meio de comunicação. Ferdinand Saussure já fazia distinção entre linguagem e língua, na qual esta última resulta ser um sistema arbitrário de símbolos fônicos que satisfaz a necessidade de comunicação de uma comunidade humana. Por sua vez, Helmut Felber define a linguagem como um sistema de símbolos que serve para expressar idéias por meio de conceitos. Bruner

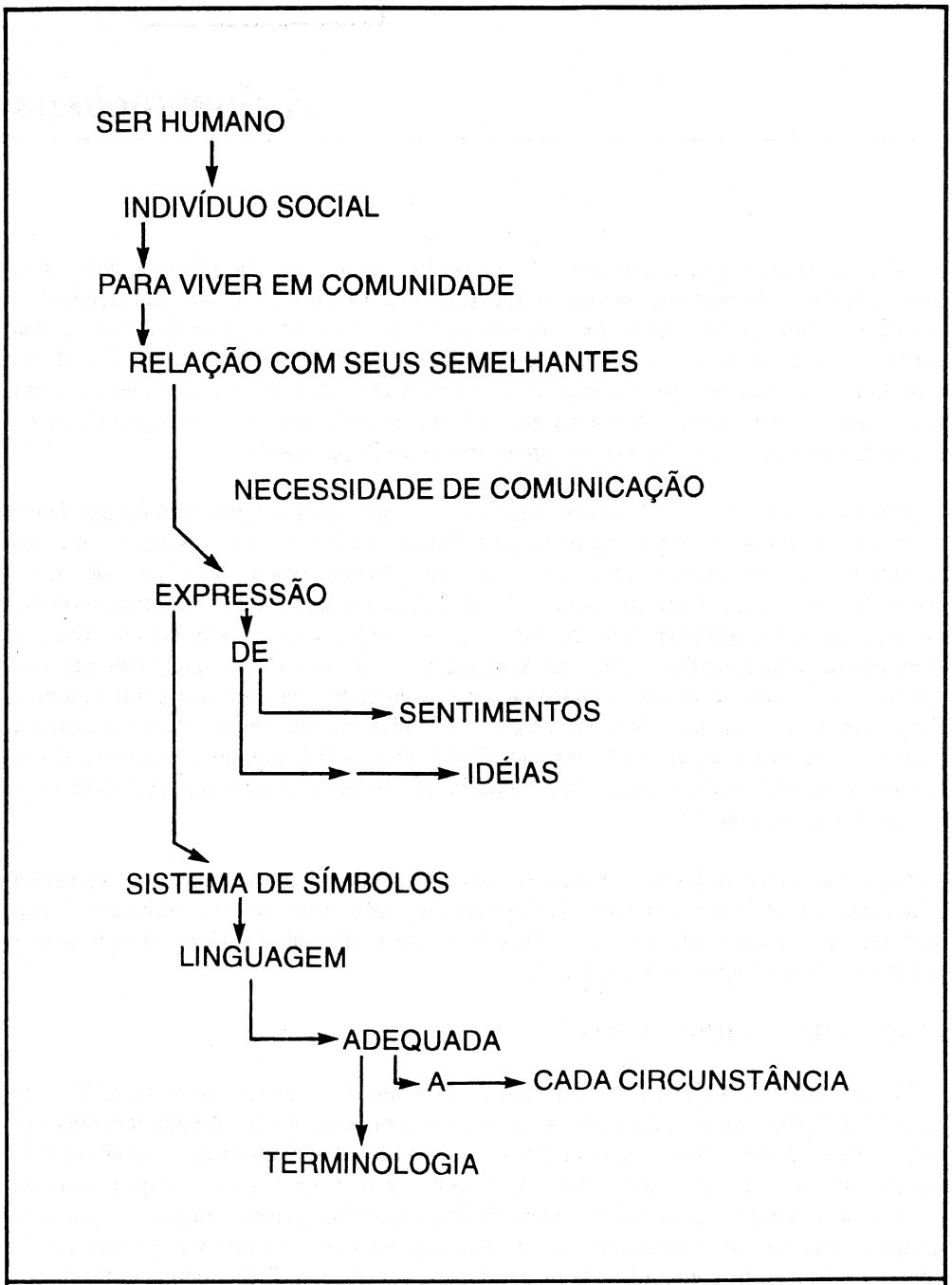


Figura 1.1. - Ser humano → Terminologia

considera a linguagem como um instrumento por intermédio do qual os seres humanos criam, constituem ou convencionam um mundo social do qual possam compartilhar. Tem, pois, a finalidade de construir um mundo social e poder "operar", mover-se, desenvolver suas atividades dentro dele. Para Saussure, a linguagem é uma faculdade humana, propriamente humana, absolutamente necessária para seus fins comunicativos. Também faz uma distinção entre a linguagem e a língua. Esta é um produto social da faculdade da linguagem.

Certamente que a linguagem é uma faculdade humana. Porém, não me atreveria a dizer que "puramente" ou "exclusivamente" humana. Os animais e também as plantas possuem uma forma de comunicação, sua linguagem própria, ininteligível para nós, e, muitas vezes, inclusive inaudível. Digamos, pois que

a ***linguagem***

- é uma manifestação dos seres vivos
  útil para suas relações em sua convivência existencial.

A palavra linguagem, no contexto desta obra, deverá ser considerada sob os aspectos diferentes:

- faculdade humana
  - enquanto que ser vivo;
- conseqüência dessa faculdade humana,
  - enquanto que conjunto de sons articulados em contínua evolução, sugidos como conseqüência da vida em forma associada.

Estas definições poderiam ser qualificadas de clássicas ou convencionais e se referem à linguagem como um todo. Contudo, nem sempre nos comunicamos por "sons". Muitas vezes utilizamos sinais ou símbolos. Assim, dizemos: a linguagem das flores. Também o sistema Morse é um sistema de sinais que serve para a comunicação humana, portanto é uma linguagem. Poderíamos citar muitos outros exemplos, sobretudo relacionados ao campo informático e das telecomunicações. Portanto, o conceito de

linguagem

ampliou-se e poderíamos considerá-lo como

- um conjunto de símbolos convencionais,
  portadores de informação,
  cujo principal fim refere-se à comunicação entre os seres vivos.

Esta definição, formulada de modo tão amplo, enquadra-se, não obstante, perfeitamente no âmbito das ciências da documentação. Também vai nos servir para ligar a lingüística, a terminologia e as linguagens terminológicas, uma vez que a informação está contida nas unidades documentárias, objeto de tratamento para sua análise, armazenamento e posterior recuperação.

Resta-nos ainda outro ponto a considerar. É o que se refere às linguagens: linguagens especializadas, peculiares e particulares de cada uma das distintas atividades humanas e dos diversos ramos do conhecimento humano na sua pluraridade de disciplinas. Qualquer uma das definições de linguagem pode-se aplicar também agora, mas com uma característica adicional, qualificadora – trata-se de "linguagens especializadas", talvez utilizadas somente pelos "especialistas" e talvez, também, somente conhecidas com propriedade por eles. A lingüística, como estudo da linguagem, passa a ser, agora, lingüística especializada: terminologia (figura 1.2).

IMPORTÂNCIA DA TERMINOLOGIA

Voltamos ao ponto de partida, a terminologia. Desejamos colocar o ser humano em sua dimensão humana-terrestre, como centro, agente, ator de um processo pelo qual transforma o macrocosmo e o microcosmo em seu mesocosmo. Trata-se de um processo ontológico, complexo e evolutivo em progressão crescente no tempo. Portanto, o mesocosmo, como produto elaborado, participará dessas mesmas características, quer dizer, será ontológico, intelectivo, complexo e evolutivo. Para conseguir-se entender em seu âmbito de ação, será necessário utilizar uma linguagem ... linguagens apropiadas, especializadas ... uma terminologia própria (figura 1.3).

Atualmente, não é necessário ressaltar a importância da terminologia. Todos compreendem que, sem seu estudo e aplicação, iremos nos perder em uma confusão sem limites.

Da mesma maneira, devemos pensar na defesa de nossos próprios idiomas, os de cada povo, que se vêem ameaçados por estrangeirismos, em uma nova forma de colonialismo científico e econômico. Contudo e não querendo entornar a nota discordante, mas com objetivo de chegar a um ponto de reflexão, atrever-me-ia a dizer que não se deve ter tanto medo dos estrangeirismos, barbarismos e outros ... ismos. É bem verdade que o castelhano origina-se do latim, com influências fortes do grego. A maior parte de nossas palavras assim o demonstram. Todavia, o castelhano enriqueceu-se extraordinariamente com palavras de origem árabe, visigótica e conta, inclusive, com restos celtas e ibéricos. É uma boa mistura dos idiomas que falavam os antigos povoadores. O mesmo latim era a língua do invasor. Tudo isso contribuiu para constituir um idioma rico, vivo e atual.

O castelhano, ao converter-se em espanhol e dar o salto para o outro lado do oceano, enriqueceu-se, deste modo, com a chegada das línguas de outros povos. Portanto, por que temer tanto os estrangeirismos? Talvez não devêssemos estar tão amarrados às raízes gregas ou latinas no momento de criar novos vocábulos.

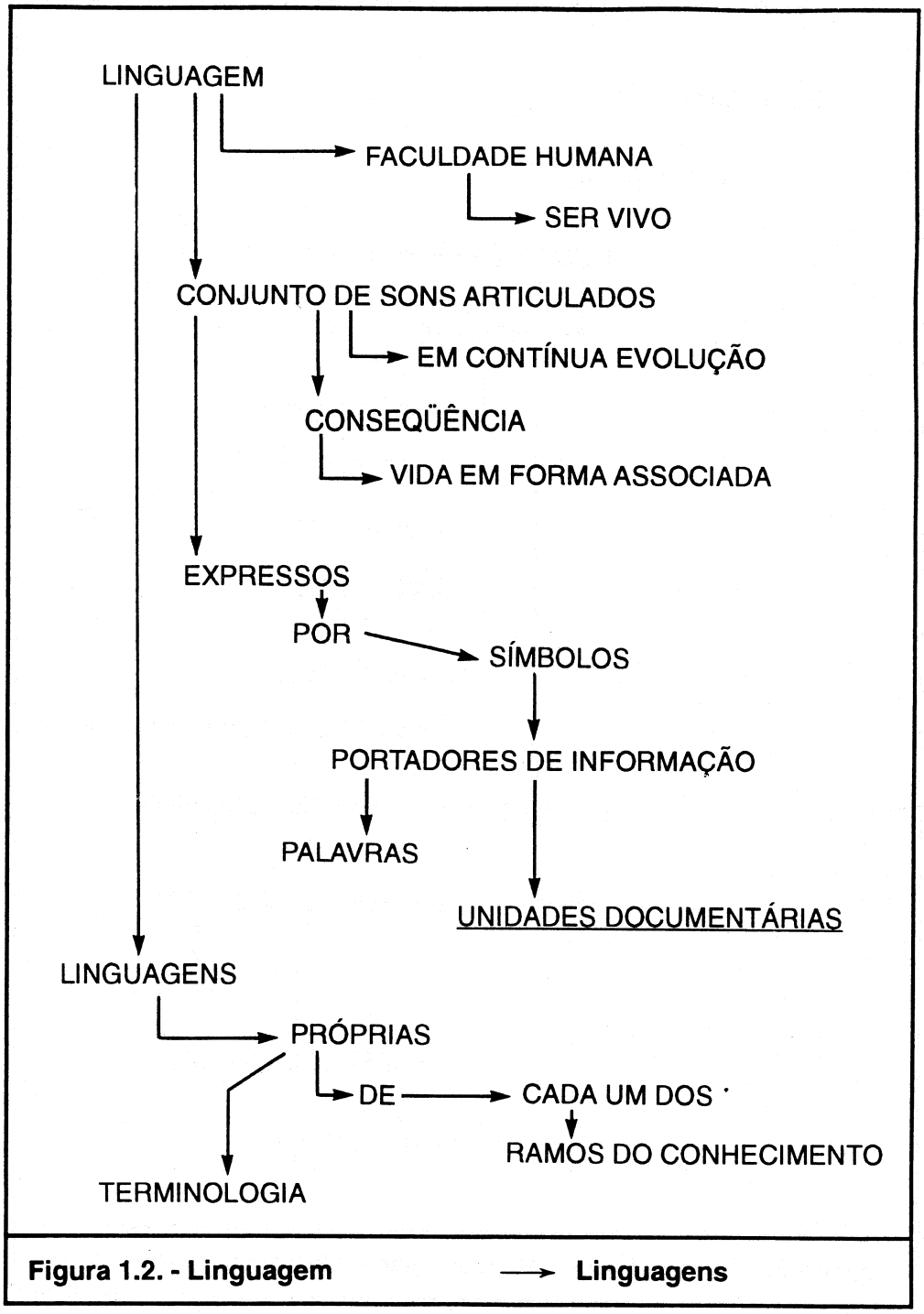


Figura 1.2. - Linguagem → Linguagens

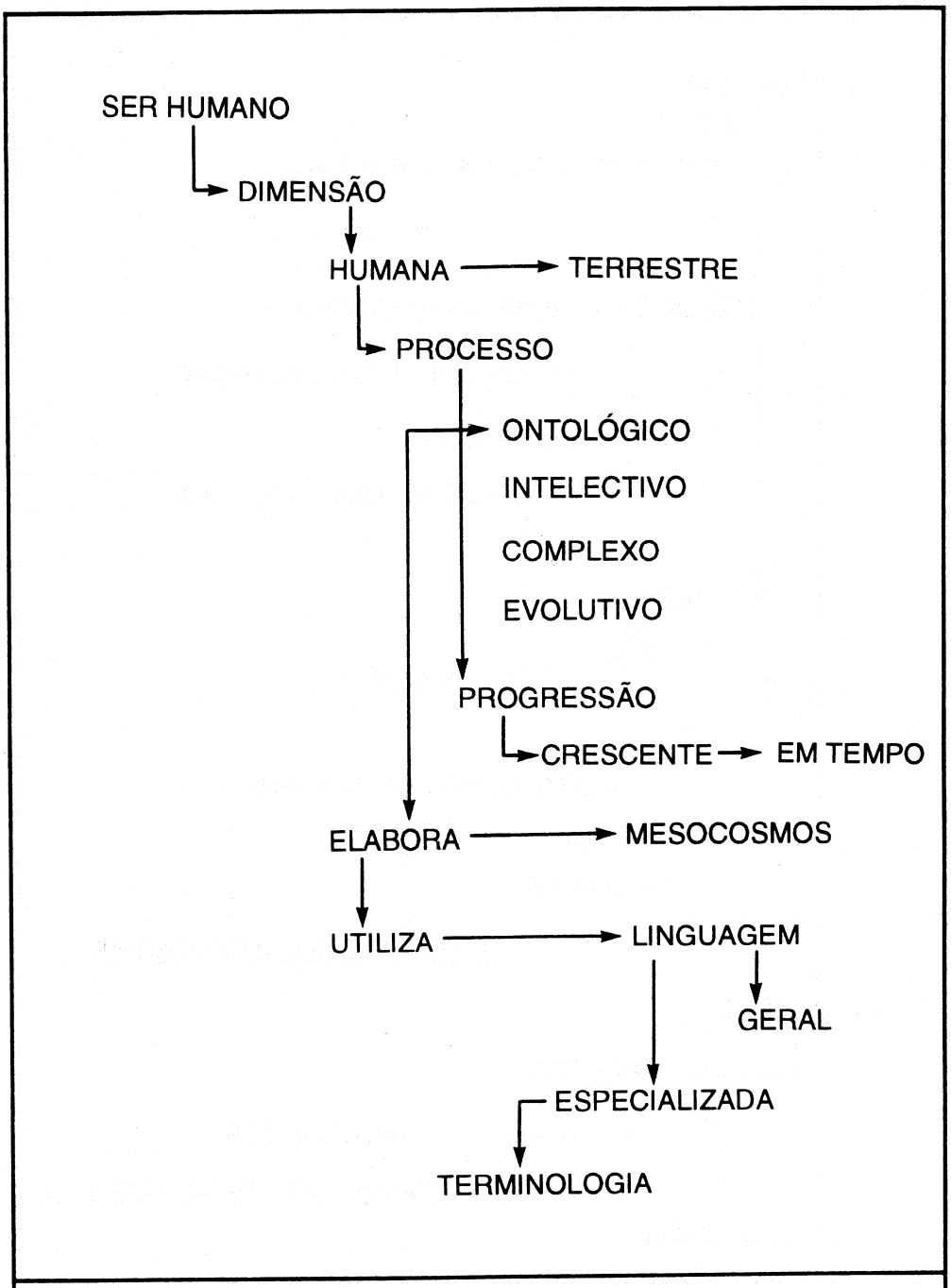


Figura 1.3. - Terminologia Processo de elaboração do mesocosmo

A importância da terminologia manifesta-se no uso correto dos vocábulos, em sua formação apropriada e em sua conservação, protegendo-os da obsolência. Manifesta-se, ainda, na uniformidade da linguagem, na normalização dos vocabulários especializados, na busca de equivalências apropriadas a serem empregadas nas traduções. Com efeito, sem uma terminologia consciente e apropriada talvez não possamos acompanhar o carro do progresso, nem nós, espanhóis, com a nossa diversidade idiomática, nem qualquer outro grupo de povos.

DEFINIÇÕES DE TERMINOLOGIA

Ao revisarmos a extensa bibliografia encontrada sobre o assunto, constatamos que a palavra terminologia, tão intuitiva em seu significado, tem sido objeto de muitos estudos, existindo diversas definições da mesma.

Nos documentos consultados, foram encontradas mais de 20 definições, as quais não resisto em omitir. Por um lado, possuímos uma visão global do seu conteúdo e, por outro, temos os princípios que nos possibilitarão chegar aos argumentos dos próximos capítulos.

Em primeiro lugar, devemos considerar que a terminologia, como toda a ciência, faz referência a duas concepções de idéias diferentes:

- um processo para a consecução de um fim;
- unidades acumuladas, resultado do processo anterior.

Dividiremos, pois, as definições segundo as dos conceitos de idéias e começaremos pela que nos é dada pelo *Diccionario de la Real Academia de la Lengua Española*:

- estudo dos termos de uma linguagem (especializada, ou vocabulário);
- estudo comparativo dos termos de várias linguagens diferentes.

Prosseguiremos, agora, com as definições dadas pelos principais estudiosos e especialistas nestes assuntos:

- disciplina situada na fronteira entre a lingüística e as demais ciências (Eugen Wüster);
- disciplina teórica que enfrenta o problema da comunicação e o entendimento na ciência e na técnica (C. Kuck);
- uma parte da lingüística (L. Guilbert);
- matéria-prima para a lingüística informática (Philippe Rossillon);
- princípio de ordenação de conhecimentos, indispensável para o desenvolvimento da comunicação e entendimento dos povos, sobretudo devido à aparição de tantas invenções, descobertas e novas ferramentas (Christian Galinski e Hans Czap);
- conjunto de termos utilizados em um campo determinado do conhecimento, de modo diferente de como é feito na linguagem comum, definidos pelo uso ou pela normalização (Picht);
- elaboração de um sistema de conceitos, reflexo da sistematização realizada em um campo

do conhecimento, com fins de entendimento entre os especialistas. Um desses fins se concretiza nos processos classificatórios desse campo do conhecimento, para seu melhor uso e aproveitamento (Amelia de Irazazábal);

- campo interdisciplinar do conhecimento que se ocupa do estudo dos conceitos e suas representações (Instituto Nacional de Información y Documentación en Ciencia y Tecnología (ICYT), Curso de Terminologia).

Helmut Felber fornece, em seus vários estudos e tratados, diferentes definições que expomos como continuação:

- sistema de conceitos aos que se lhes atribui um termo;
- padrão para a representação e ordenação do conhecimento;
- conjunto de regras, fatos e estratégias para resolver problemas;
- base da trilogia.

  objeto → conceito → símbolo,

  que liga a realidade com a simbologia;
- ramo do conhecimento humano relacionado com a lingüística, a lógica, a ontologia, a ciência da informação (ciências da documentação), as ciências naturais e as ciências do espírito;
- estuda a atribuição de conceitos, postula definições, pesquisa expressões lingüísticas de coisas, idéias e procedimentos, para a comunicação entre especialistas que trabalham em um determinado campo do conhecimento;
- estuda a correspondência de termo, em um sistema referencial, com outro, em outro sistema referencial.

Por outro lado, a terminologia, enquanto conjunto de palavras especializadas de uso restrito e com significado próprio, pode-se definir, de acordo com a Real Academia de la Lengua Española, como:

- conjunto de termos ou vocábulos de uma determinada profissão, ciência ou assunto.

Nos apontamentos do curso sobre terminologia, organizado pelo Cindoc, encontra-se esta outra definição:

- conjunto de termos que representa o sistema de conceitos de um campo específico (em um ou vários idiomas).

Outras definições relativas a esta acepção significam o mesmo, ainda que com palavras diferentes.

Fruto de minha própria reflexão sobre esta disciplina, aplicando os princípios da teoria de sistemas, acrescento outra definição às existentes:

- processo ontológico pelo qual se chega à normalização dos sistemas formados pelas diferentes linguagens especializadas, de onde o conceito fixado pelo termo está em relação semântica com o resto dos termos desse sistema. A finalidade da terminologia se resume no melhor entendimento dos especialistas entre si e destes com o seu meio (figura 1.4).

Como vemos, todas as definições fazem referência aos processos mentais que acontecem em nosso cérebro quando se deseja nomear um conceito, seja este concreto ou abstrato, aplicado a um determinado ramo do conhecimento. Atualmente, quando as ciências evoluem continuamente e aparecem novas teorias, novos descobrimentos e novos inventos, surgem, deste modo, novos conceitos que devem ser simbolizados por um termo. Portanto, a terminologia deve caminhar à frente do progresso. Não somente deve estudar os termos existentes, mas construir novos termos, para as novas necessidades ... Neste ponto, está a sua importância e força, que não se deve considerar exclusivamente a partir do ponto de vista lingüístico e lexicográfico. Sua repercussão vai muito mais além, até a normalização das linguagens que permitem um domínio, uma soberania sobre o mundo da informação.

O CONCEITO E OS TERMOS

Após havermos feito referência, anteriormente, ao conceito e ao termo repetidas vezes, sem nos aprofundarmos em seus significados e utilização, parece que, agora, chegou o momento de dedicar a esses assuntos um estudo um pouco mais detalhado, que nos sirva de base para argumentações posteriores.

Ficaremos, primeiramente, na acepção de "conceito" a que faze referência a maioria das definições de terminologia. Logo após, consideraremos o "termo" como sua expressão simbólica e utilizável.

Por detrás do conceito, esconde-se toda uma linha de pensamento. Ao pensamento se chega pela elaboração do conhecimento, e este se alcança mediante um processo de assimilação da informação que nos vem de fora, de fora de nosso cérebro. Assim, pois, tanto os estímulos de nosso mundo interior, como os do mundo que nos cerca, são objeto de elaboração da mente, onde os neurônios desempenham um papel primordial. Produz-se o fenômeno da informação, somos conscientes do mesmo, modificamos nosso estado de conhecimento e, em um estado posterior de sua elaboração, formamos nossa linha de pensamento – nascem os conceitos. Desta maneira, une-se a informação com o conceito (figura 1.6). Uma unidade de informação se transforma em uma unidade conceitual, que não guardamos em nosso cérebro para uso exclusivo e interno de nossos pensamentos, mas que desejamos expressar com objetivos de comunicação. Para isto, utilizamos os símbolos. Na maioria dos casos, lançamos mão dos símbolos lingüísticos e formamos palavras escritas ou faladas.

Hoje em dia, temos de considerar que as palavras podem ser expressas de muitas maneiras diferentes, das quais os sinais dos alfabetos distintos não constituem mais que uma parte, embora muito importante e majoritária. Uma definição recente aborda o conjunto de símbolos

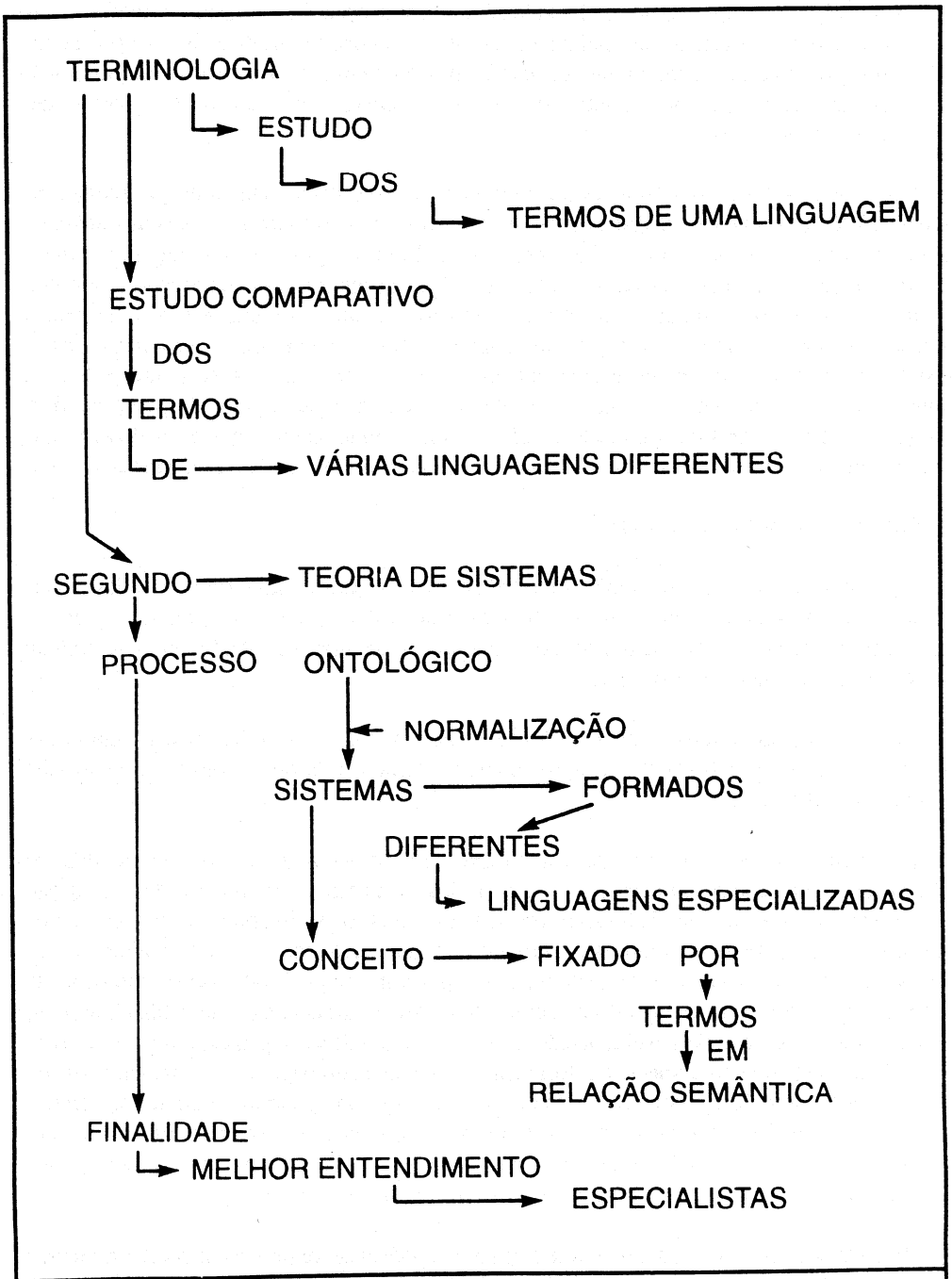


Figura 1.4. - Conceito de terminologia

ou sinais, separados por espaços em branco, que possuem um significado inteligível para o ser humano (nota da autora). Para o propósito deste livro, levaremos em consideração a definição clássica, que supõe utilizar os alfabetos lingüísticos convencionais para representar a palavra.

Então, de

informação
↓
conhecimento
↓
pensamento
↓
conceito
↓
palavra (ou conjunto de palavras).

Esta linha corre paralela com a de Helmut Felber, quando estabelece a relação
objeto → conceito → símbolo.

Para nós, a partir do ponto de vista das ciências da documentação, o processo não termina na palavra, pois essa palavra pode constituir uma unidade informativa, um *quantum* de informação útil, contida em documentos, que deve ser tratada para seu armazenamento e posterior recuperação e utilização. Origina-se, pois, um circuito aberto, irreversível, que, partindo da informação como origem, leva à informação como fim, como produto elaborado (figura 1.7).

O **conceito** é um produto intermediário, primordial. Se não chegarmos ao conceito, não se pode continuar o processo. Por esse motivo, tanto os lingüistas, como os documentaristas e os terminólogos procuram definir o conceito, de que temos encontrado também diversas modalidades. Destacamos as seguintes:

- unidade de pensamento, geralmente expressa por um termo, ou letras como símbolos, ou qualquer outro tipo de símbolos (ISO/R 1087);

- construção mental que serve para classificar objetos individuais ou abstrações (ISO/R 704);

- unidade de conhecimento que dá origem a uma unidade de pensamento, portanto é a representação da verdade, verificável e justificável (Shuji Czehi);

- elemento de conhecimento. Representação mental de objetos individuais que formam uma unidade de pensamento. Expressa-se por lexemas, formados por morfemas ou

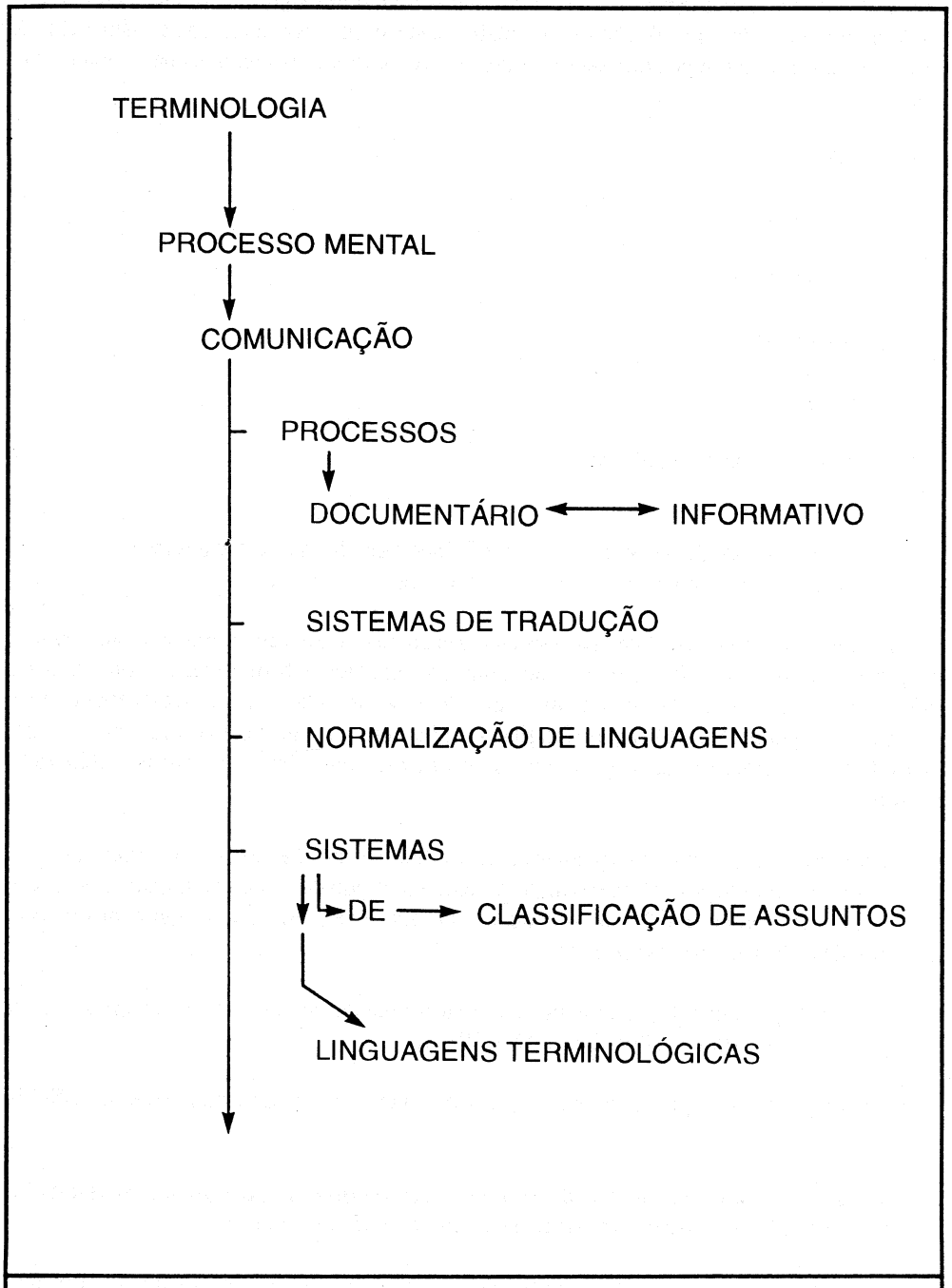


Figura 1.5. - Utilidade da terminologia

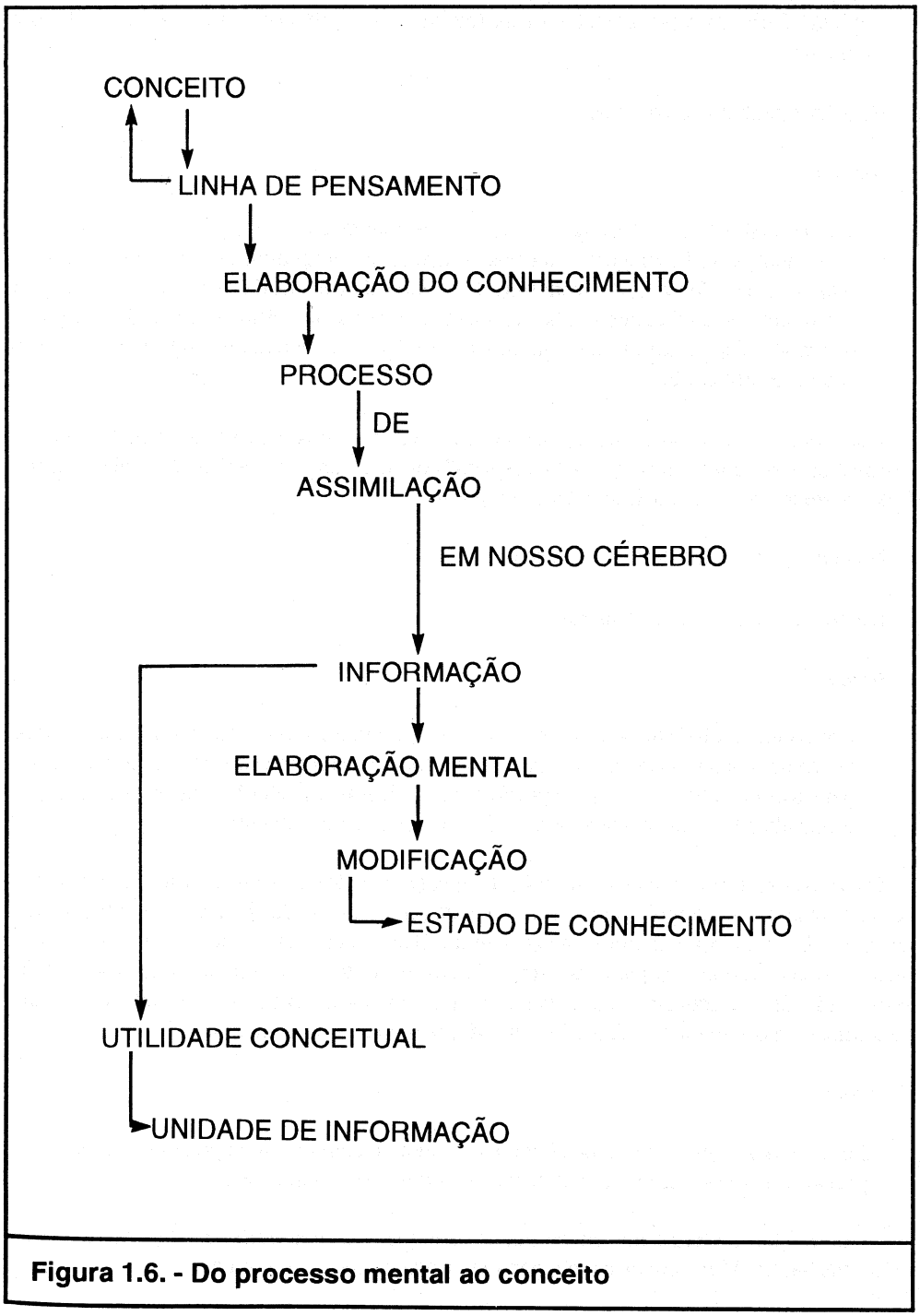


Figura 1.6. - Do processo mental ao conceito

conjuntos de morfemas, e pode apresentar peculiaridades distintas; por exemplo, ser objetos (substantivos), características (adjetivos e advérbios), atos (verbos) (Helmut Felbert).

H. Felber também nos diz que

conceito

é a expressão "real", objetiva de informações, que recebe forma por meio das palavras – elementos da linguagem – termos – dados, que adquirem significado quando se relacionam no contexto da economia, da política, da ciência, da sociologia, ou qualquer outro ramo do conhecimento. Isto é, uma palavra pode ter vários significados, segundo o campo de aplicação. Cada palavra – termo – possui somente um significado em seu campo de aplicação.

Esta é uma das poucas definições de conceito em que se considera a finalidade do conceito e se atribui ao mesmo uma característica especial para um campo específico do conhecimento. Desta maneira, pode-se compreender as etapas de

conceito → palavra → termo.

Deduz-se que se pode considerar o

termo

como uma unidade de pensamento que se configura em nosso cérebro, ao se colocar os neurônios em atividade pelos impactos de informação recebida do exterior, armazenadas para uso posterior e que se expressam por palavras, de acordo com uma linguagem especializada e normalizada para cada parcela do conhecimento.

O termo possui, pois, um único significado. A palavra pode ter vários significados. Costuma-se dar a água como exemplo representativo. Esta palavra varia de significado conforme seja considerada em hidrologia, química, agricultura, física, indústria ... Portanto, um termo será uma palavra especializada para um campo determinado do conhecimento. É a expressão do significado de um conceito, que serve para a comunicação entre os especialistas em um determinado assunto ou atividade. Resumindo, poder-se-ia dizer que um

termo

é um símbolo (ou conjunto de símbolos ou sinais), com que se expressa um conceito, é portador de informação – um elemento *semaforonte* (figura 1.8).

As definições de termo que encontramos na numerosa bibliografia diferem muito pouco umas das outras. Merecem destaque, por sua diferenciação, as seguintes:

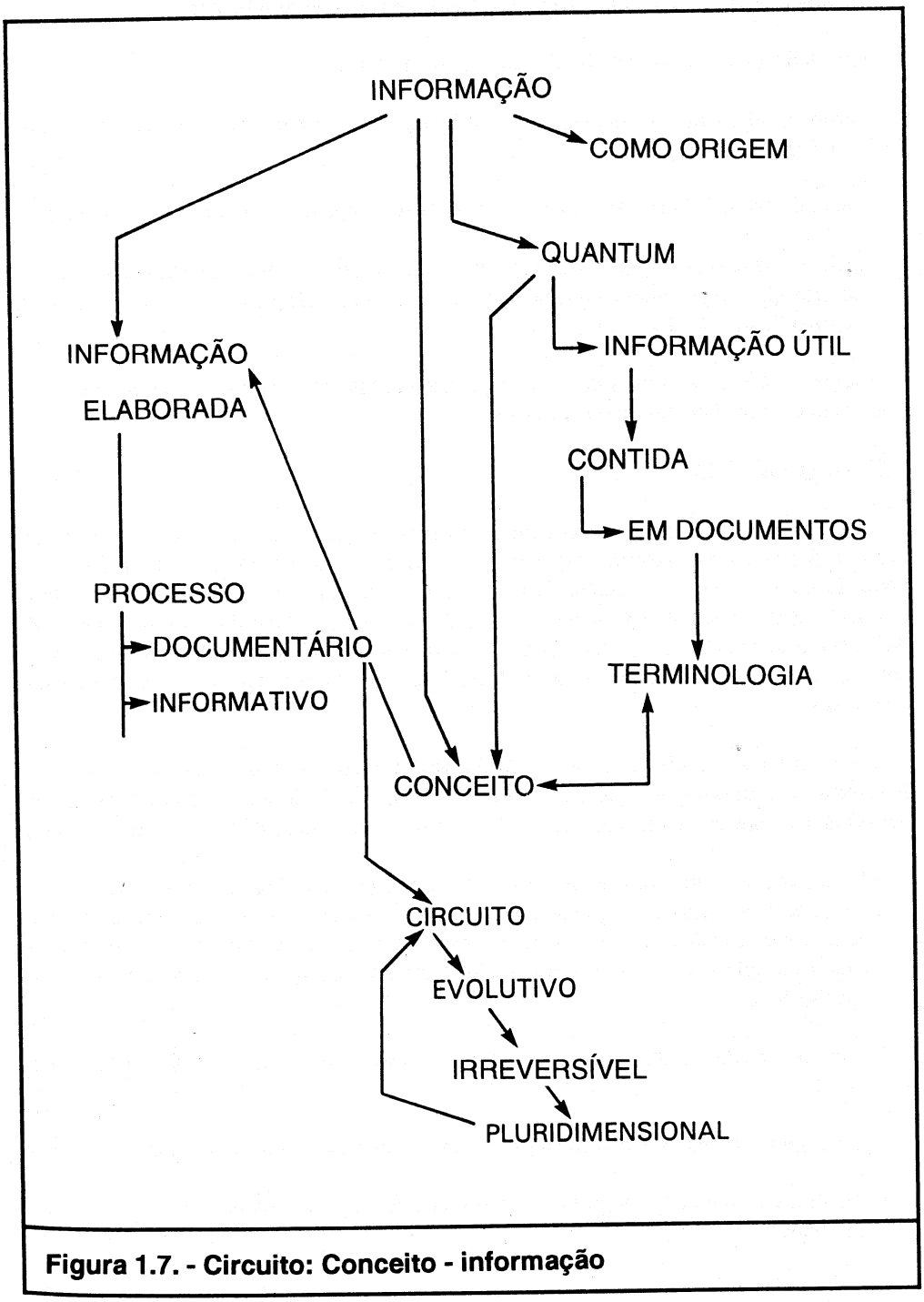


Figura 1.7. - Circuito: Conceito - informação

- unidade lingüística de um vocabulário especializado (Helmut Felber);

- formalização de um conceito (Amelia de Irazazábal);

- primeiro elemento do conhecimento, útil para armazenar os conhecimentos (Ingetraut Dahlberg);

- unidade classificatória (apontamentos e notas do grupo de terminologia do Cindoc);

- palavra ou expressão utilizada para a inclusão temática de um item em um sistema de informação. Representa o conceito em um sistema de indexação (trabalho de classe de Regina Soares de Oliveira).

Citamos estas duas últimas definições, porque já indicam a sua utilização na lingüística documental e nos sistemas classificatórios.

TERMINOGRAFIA

Hoje em dia, quando aparecem mais idiomas autóctones, nos campos científicos, econômicos e técnicos, a forma na qual os termos se apresentam ou podem ser representados adquiriu uma grande importância. É bem verdade que não se ouve falar desse assunto freqüentemente, sendo necessário recorrer às publicações e à bibliografia muito especializada para encontrar referências ou definições. Todavia, não se pode prescindir da terminologia, sobretudo quando se deseja uma informação e normalização sobre opiniões, teorias e métodos de trabalho.

Uma definição concisa e clara de terminografia encontra-se no manual do Curso de Introdução à Terminologia, organizado pelo ICYT. A autoria da definição não é citada, motivo pelo qual não podemos mencioná-la aqui. Diz (embora não tenha sido tomada textualmente):

- Disciplina que tem por objetivo o registro das correspondências permanentes termo-conceito. Determina a posição de um conceito dentro de seu âmbito de aplicação. Trata de registrar os dados terminológicos que fornecem uma descrição precisa de um conceito e indica as relações entre conceitos. Assim, é fixado e registrado o uso de termos em um vocabulário.

Por meio de terminografia, consegue-se (baseado nos apontamento de classe de Amelia de Irazazábal):

- estruturar e ordenar os dados que representam os conceitos (os termos e suas definições);

- registrar os dados terminográficos, compostos pelos dados terminológicos e os dados associados;

- a ordenação dos conjuntos de dados que representam os conceitos, constituídos de unidades de vocabulário, seções monolíngües de unidades multilíngües e formatos;
- estruturar as unidades de um vocabulário, que são unidades orientadas para o conceito e o termo;
- criar símbolos gráficos para as relações entre conceitos, termos etc.;
- tratar os dados terminológicos, tanto manual, como automaticamente;
- formar coleções de dados terminológicos; por exemplo, vocabulário especializados.

BIBLIOTECA DO IBICT

A terminologia adquire sua verdadeira importância na construção dos bancos terminológicos, nos quais cada termo deve estar caracterizado registrado e documentado exatamente.

OS TERMINÓLOGOS

Não se pode terminar este capítulo, onde se aborda a terminologia em sentido geral, sem fazer referência aos que tornam possível o trabalho terminológico e terminográfico – são os terminólogos.

Esta profissão, tão antiga como o aparecimento da linguagem, e seu uso para as diversas atividadades humanas não foram considerados como tal até há muito pouco tempo. Ainda quando se fala dela, há quem sorria maliciosamente, pensando tratar-se de elocubrações e fantasias para justificar atividades e remunerações, e pergunta-se: de que profissão não se pode dizer o mesmo? O que deve ser demonstrado é se os terminólogos são, realmente, necessários ou não.

Depois do que foi descrito aqui e de examinar a extensa bibliografia, concluímos que a terminologia é primordial para o entendimento entre os cientistas e especialistas de qualquer atividade, e também somos levados a admitir, irremediavelmente, que os terminólogos também o são.

Mas quem são esses terminólogos? São os profissionais especialistas em um assunto concreto do conhecimento humano e em lingüística. São os que devem encontrar o termo apropriado ao conceito em estudo. São os que devem definir os novos termos e formá-los, normalizá-los e recomendá-los. São os que precisam defender o idioma contra as deformações. Estas tarefas parecem fáceis, assim enunciadas. Contudo, cada uma delas encerra uma grande complexidade. Por isso, consideramos que devam ser especialistas em um assunto específico e em lingüística. Neste ponto, encontra-se a dificuldade. Normalmente, um físico, por exemplo, conhece sua especialidade dentro da física, todavia não sabe nada ou conhece pouco de lingüística. Com um lingüista ocorre o mesmo, mas em sentido contrário. Incorporar ambos os conhecimentos não é fácil. E não é pela complexidade dos assuntos em si, pois para isso

está a inteligência do ser humano. A dificuldade reside em que nem o físico nem o lingüista possuem interesse pela disciplina contrária. É preciso procurar pessoas que se interessem por ambas as disciplinas, e isso é o que parece não se encontra freqëntemente. Por isso, diz-se que os cientistas e os lingüistas devem trabalhar em conjunto nas tarefas terminológicas.

Esta é uma profissão nova, considerada como tal. Na maioria dos países, não figura nos catátologos de profissões, não se conhece bem quem são, nem se confere a eles a devida importância, tampouco existem muitas escolas onde possam receber formação profissional. Em alguns casos, como a de Viena ou Caracas, por exemplo, constituem exceções à regra. Na maioria das ocasiões, trata-se de autodidatas que nem sequer têm consciência de que são terminólogos.

Não obstante, nestes últimos anos, com a necessidade de se dispor de traduções confiáveis, manuais ou automatizadas, começa-se a pensar na importância de poder recorrer a uma só pessoa que possa resolver os problemas terminológicos e terminográficos. Em nível profissional, está mudando-se a mentalidade, e isso abrirá as portas para esta "nova" profissão.

Um terminólogo deve possuir uma formação muito completa, universal e erudita. Para criar um novo termo, em qualquer que seja o campo, é necessário ter conhecimentos nesse campo determinado e, além disso, dominar a lingüística. Esta disciplina é, por si mesma, multidisciplinar. Está ligada diretamente à história, à sociologia, ao conhecimento de idiomas, à economia ...

Enfim, um terminólogo deve ser um profissional muito completo em relação à sua formação e educação. Devido a isso, até agora, não são em grande número, ainda que a maioria dos países conte com alguns especialistas bem conhecidos no mundo terminológico.

O trabalho nacional e internacional é muito frutífero. Realizam-se congressos e reuniões. Existem escolas de linhas de pesquisa. Estão sendo criados bancos terminológicos nacionais e internacionais. A terminologia é uma ciência em pleno auge, portanto a profissão de terminólogo também deverá ser.

No capítulo seguinte, onde se aborda a evolução histórica da terminologia, serão citadas as escolas de linhas de pesquisa e os especialistas e terminólogos mais importantes e relevantes, tanto os mais antigos, como os atuais.

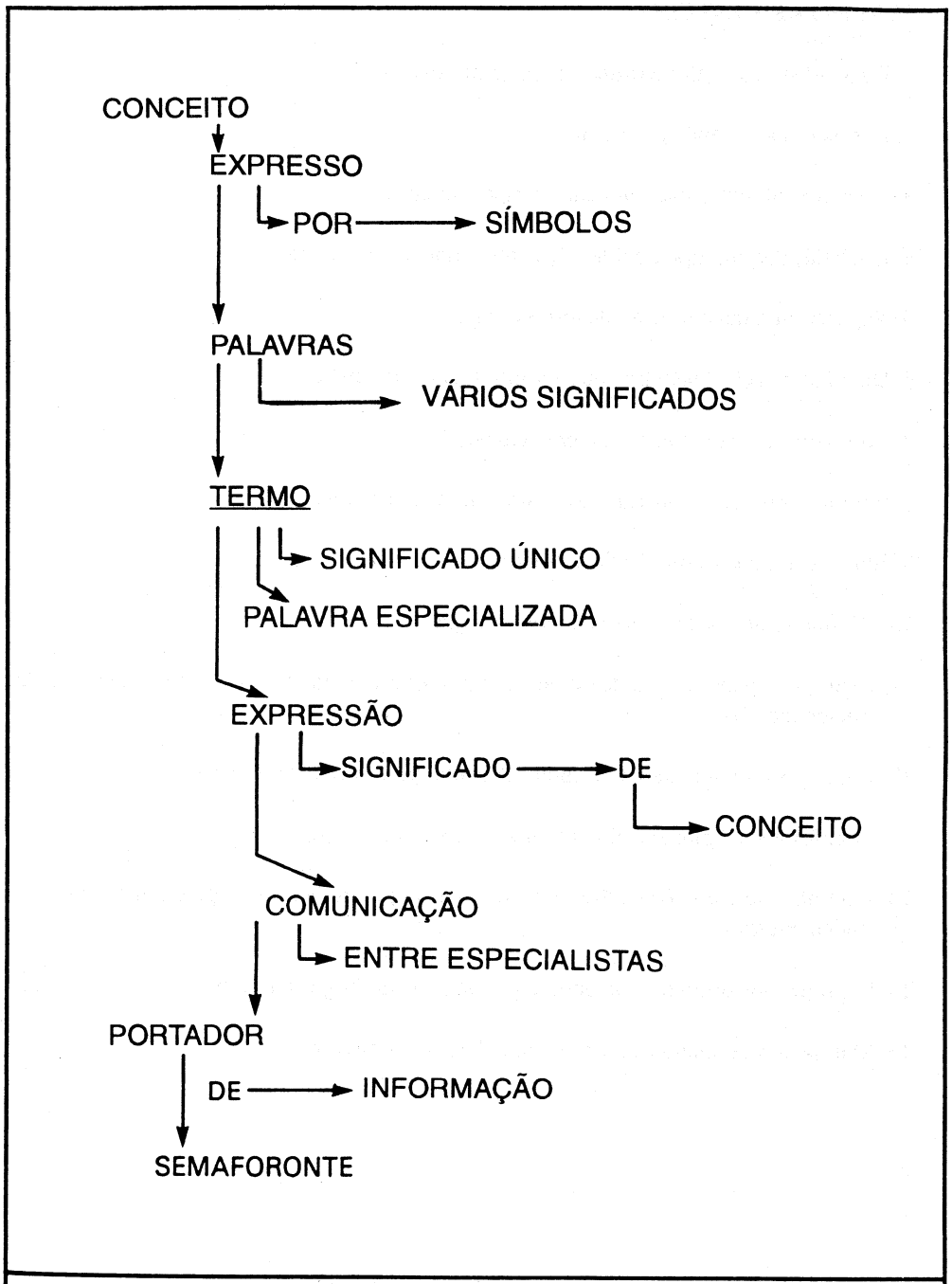


Figura 1.8. - O termo

TEMAS PARA ESTUDO

1. Faça a distinção entre terminologia e terminografia.

2. Defina o que entende por termo.

3. Diga que diferença encontra entre termo e conceito.

4. Estabeleça algum tipo de diferença entre conceito e palavra.

5. Exponha alguma definição de terminologia.

6. Diga que relação encontra entre terminologia e lingüística.

7. Como diferencia um termo de uma palavra?

8. Em que forma relacionaria o conceito com o conhecimento?

9. Enuncie alguma definição de conceito.

10. Dê sua opinião sobre o uso da terminologia.

11. Explique o papel que pode desempenhar a terminologia no contexto das ciências da documentação.

12. Como pode relacionar a terminologia com o processo informativo?

13. Manifeste sua opinião sobre a profissão de terminólogo.

14. Acredita que é possível substituir um lingüista por um terminólogo em um centro de documentação?

15. Explique que diferença encontra entre linguagem, língua e terminologia.

16. Manifeste sua opinião a respeito da defesa da linguagem.

REFERÊNCIAS BIBLIOGRÁFICAS

* Actas de la Exposición de Lingüística Informática y de Terminología Científico-técnica; CSIC y Unión Latina, Madrid, 1987.

* Actas del Primer Seminario Nacional de Terminología; Caracas-Venezuela, Universidad Simón Bolivar, junio de 1984.

ALSORNOZ, O.: "Las lenguas latinas como idioma científico en las ciencias sociales: los problemas políticos y culturales de la lengua sajona en América Latina y el Caribe"; en Actas de la Exposición de Lingüística Informática y de Terminología Científico Técnica, CSIC y Unión Latina, Madrid, 1987, 69-72.

ALVAREZ, S.: "La barrera lingüística. Soluciones", en Apuntes del Curso de Introducción a la Terminología, ICYT, Madrid, 1988, 13-22.

BOSCH DE ROZF, A.: "Lingüística y Ciencias de la Información: un acercamiento integrador"; Rev. Esp. Doc. Cient., 9, 2, 1986, 167-171.

CALVO HERNANDO, M.: "Periodismo científico y respeto a la Ciencia"; Química e Industria, 33, 5, 1987, 400.

CENTENERO GALLEGO, M.A.: Congreso de Terminología Musical Castellana; Campus (Rev. Universidad de Murcia), 17 de febrero de 1989. 10.

CURRÁS, E.: "La necessité d'une cooperation hispano-française en terminologie chimique"; Documentaliste, *18* 1, 1981, 7-9.

* CURRAS, E.: "Taxonomía y Sistemática"; en La información en sus nuevos aspectos: Ciencias de la Documentación, Paraninfo, Madrid, 1988, 170-194.

* Curso de Introducción a la Terminología, ICYT, Madrid, 1988-89, Apuntes.

* Diccionario de Información y Documentación. Versión Espanola del Diccionario Plurilingüe de la FID, a partir de los originales inglés y ruso, Madrid, ICYT, sin fecha.

* Dictionary of Library Science, Information and Documentation. Compiled by W.F. Clason. Amsterdam, Elsevier Scientific Publishing Company, 1973.

ESCORIAL, M.: "Es necesario definir una política hispano-latina en terminología técnica"; en Bol. de Comunicación, Tecnología y Sociedad, *II*, 11 jun., 1988, 8-10.

FEDOR DE DIEGO, A.: "Lenguaje común, tecnolecto, terrninología" en Actas del Primer Seminario Nacional de Terminología, Caracas, 1984.

* FELSER, H.:"Terminology and Knowledge Engineering"; in International Congress on Terminology and Knowledge Engineering; Federal Republic of Germany, University of Trier, Frankfurt, 1987, 3-7.

* FELBER, H.: "Terminology Manual"; UNESCO/Infoterm, Paris, 1985

FELBER, H. y PICHT, H.: "Métodos de terminografía y principios de investigación terminológica"; HISPANOTERM, Instituto Miguel de Cervantes, CSIC, Madrid, 1983.

GOETSCHALCKX, J.: "Rapport de synthese. Deuxième Symposium d'Infoterm. Vienne, 14-17 avril 1985"; TermNet, 15, 1986, 3.

(*) As obras assinaladas com um asterisco (*) também fazem referência aos temas tratados em outros capítulos.

GONZALES SUAREZ, E.: "Conocimiento, Información y Comunicación"; Act. Inf. Cient. Técn., 1 (144), 1989.

HANADA, T. y SASAKI, H.:"Features and problems of japonese technical terms"; In Terminology and Knowledge Engineering, INDEKS Verlag, Frankfurt, 1987, 403-412.

IRAZAZABAL, A.: "La terminología. ¿Por qué?'"; en Noticias de la SEDIC, Madrid, 1988, 7-16. Trabajo mecanografiado.

* KARLGREN, H.: "FID Lingüístic in Documentation", Commitee FID/LD; Int. Forum Inf. and Docum., *11*, 3, july 1986, 51-55.

KOTTER, R. and LUFT, A.L.: "Sprachpragmatische Aspekte der Terminologie und Wissensgewinnung"; in International Congress on Terminology and Knowlegde Engineering; INDEKS Verlag, Frankfurt, 1987, 77-84.

LAPESA, R.: "La necesidad de una política hispánica sobre neologismos técnicos y científicos"; en Rev. Telos, 5, 1986, 84-9.

MARTIN MUNICIO, A.: "Lexicografía de la Ciencia y de la Técnica"; Rev. Telos, 5, 1986, 105-112.

RIEGER, B.: "Definition of terms, wordmeaning, and knowledge structure. On some problems of semantics from a computational view of linguistics"; in Terminology and Knowledge Engineering, INDEKS Verlag, Frankfurt, 1987, 3-7.

RODRIGUEZ JIMENES, M.: "La informática y la degradación de la lengua ; en Bol. de Comunicación, tecnología y sociedad, *21*, 23 ene., 1988, 13-4.

ROSILLON, P.: "Necessités terminologues en la politique d'Etat"; en Actas de la Exposición de Lingüística y de Terminología Científico-técnica, CSIC y Unión Latina, Madrid, 1987,126-131.

SANLLORENTE, A.M.: "La terminología de la educación con América Latina: su compatibilización"; Rev. Esp. Doc. Cient.; *9*, 2, 1986, 151-159.

SOARES, R.M.: "Lingüística y terminología. Traducción y Normalizació""; lección preparada para el Curso de Doctorado sobre "Metodología y Práctica de la Información Científica y la Documentación Automatizada", organizado por el Gabinete de Documentación Científica de la Universidad Autónoma de Madrid, curso 1986-87.

* "Terminology and Knowledge Engineering"; TermNet, 15, 1986, 1-2.

* "Terminology Manual"; PGI y UNISIT, Infoterm, París, 1984.

* "Terminología sobre Información y Documentación y materias afines"; Rev. Esp. Doc. Cient., *10*, 3, 1987, 361-369.

UNGUARY, R.: "Intensional Splittin,: an Empirical Examination of Conceptual Duality"; Int. Classif., *13*, 1, Budapest, 1986, 9-17.

* WERSIG, G. y NEVELING, U.: "Terminology of documentation"; The UNESCO Press, París, 1976.

* YOUNG, H.: "Glosario ALA de Bibliotecología y Ciencias de la Información"; Díaz de Santos, Madrid, 1988.

2

Evolução histórica

Quando desejamos um aprofundamento no estudo da disciplina terminologia, encontramos tantos dados, tantas referências e tantas atividades que tememos não poder abordar o assunto com o suficiente conhecimento de causa. Talvez, se perguntássemos qual o ramo do conhecimento como maior índice de desenvolvimento, não saberíamos como responder. À primeira vista, parece que a informática, junto com a cibernética e a matemática, são as disciplinas que se encontram mais no apogeu atualmente. Contudo, quando alguém se encontra imerso em um mar de documentos avulsos, manuais de estudo, atas de congressos, programas de cursos e publicações periódicas sobre terminologia, não deixará de pensar, ao menos, em sua relevância e importância.

Certamente é assim, já que é a base para o entendimento. E aqui se encerra uma complexa série de questões que repercutem na totalidade das demais ciências. Trata-se do estudo da lingüística, da filosofia, da lógica, da lexicografia, da tradução, da normalização, dos sistemas de classificação, das técnicas documentário-informativas – todas elas ciências em si mesmas de ampla difusão nos meios científicos, técnicos e culturais. Todas elas são de primeira necessidade para se conseguir um desenvolvimento harmônico da humanidade, no qual o progresso marca o modelo de sua evolução. Evolução, hoje em dia, acelerada em progressão crescente, rompendo fronteiras entre os povos, aproximando-os cada vez mais. Não obstante, paradoxalmente, ao mesmo tempo vão surgindo, emergindo, povos minoritários que desejam conservar sua identidade e sua idiossincrasia, precisamente conservando seus idiomas. Quer dizer, por um lado, devemos fazer o possível para encontrar vínculos de união e de entendimento, por meio da linguagem, e, por outro, desejamos isolarmo-nos e identificar-nos, deste modo, pela linguagem. Tão importante é o poder da língua!

Por esse motivo, neste capítulo, procuramos evidenciar a evolução da terminologia no decorrer da história e poder compreender melhor seu desenvolvimento para chegar ao estado atual.

PERÍODOS EVOLUTIVOS

Estudando o desenvolvimento dos problemas terminológicos no decorrer dos tempos, é interessante constatar as linhas de evolução seguidas desde as primeiras épocas até nossos dias. Como conseqüência, será possível vislumbrar o que acontecerá em um futuro próximo.

Todos os ramos do conhecimento marcam seus próprios períodos evolutivos, que nem sempre coincedem com os grandes períodos históricos, geralmente conhecidos como tais. No caso da terminologia, podem ser considerados os seguintes:

1. Período clássico:
   até o final do século XVII.

2. Período de formação:
   (à terminologia moderna), séculos XVIII e XIX.

3. Período de consolidação:
   século XX, até 1930.

4. Período de teorização:
   de 1930 a 1970.

5. Período de evolução acelerada:
   de 1971 até nossos dias.

Estas divisões não supõem mudanças drásticas produzidas nas datas assinaladas. Nenhuma evolução humana acontece drasticamente em um determinado momento histórico. Quando, aparentemente, esse choque brusco se produz, é porque chegou o amadurecimento de um processo que vinha se formando há muito.

Escolhemos essa datas porque marcaram alvos na evolução da terminologia e se ajustam apropriadamente para realizar este estudo histórico.

Deverá ser levado em consideração que o desenvolvimento da disciplina em si, desde seu ponto de vista conceitual, como também a ampliação de seu campo de aplicação e a adaptabilidade da informática como meio auxiliar para seu uso mais prático e racional. Hoje, não se pode prescindir das chamadas novas tecnologias, com a informática e as telecomunicações em suas versões mais modernas.

PERÍODO CLÁSSICO

Fixou-se esse período desde o começo das civilizações até o final do século XVII. Abrange uma ampla extensão de tempo, quando ocorreram acontecimentos muito importantes na história da humanidade. Na realidade, poderia parecer pouco apropriado incluir, pelo menos, alguns milhares de anos, e inclusive séculos, em uma só divisão temporal. Todavia, do ponto de vista da terminologia moderna, sucederam-se apenas acontecimentos que mereceram marcar etapas intermediárias importantes.

Durante o decorrer do século XVII, quando começou a nascer a chamada ciência moderna, tinham-se idéias universalistas das ciências, da filosofia e, em geral, de todos os ramos do

conhecimento. Durante os primeiros séculos de nossa civilização ocidental, predominavam as idéias enciclopedistas e universais. Os mesmos especialistas que surgiam na astronomia, ou na matemáticas, ou nas náutica, por exemplo, tinham consciência de sua especialidade como uma parte de um todo. As linguagens que utilizavam, ainda que próprias e estabelecidas para seu campo de trabalho, constituíam temas de estudo da lingüística. Os novos conceitos que iam aparecendo, como reflexo das sucessivas invenções e descobrimentos, enquadravam-se dentro da lógica, sem uma preocupação por seu estudo em nível terminológico. Tampouco acredito que os enciclopedistas da Idade Antiga, nem Avicena, em seu *Livro da Sabedoria*, nem Raimundo Lulio, em sua *Ars Magna*, nem Juan Huarte, em seu *Examen de Ingenios para las Ciencias*, tiveram consciência de estar ocupando-se de problemas terminológicos, quando idealizaram suas respectivas classificações universais. Os tradutores, imprescindíveis em qualquer época histórica, trabalham por sua conta, consultando os dicionários e buscando equivalências apropriadas para as palavras e expressões que apareciam nos textos.

O que se procura mostrar é que, ainda que se realizassem tarefas consideradas atualmente no campo terminológico, e algumas delas bem importantes, não se possuía consciência do mesmo. Cada grupo de profissionais ou trabalhadores trabalham em seu assunto, estudando lingüística, ou filosofia, idealizando sistemas classificatórios, realizando traduções, elaborando dicionários e enciclopédias, sem se relacionar uns com os outros, mais que para utilizar os resultados de suas atividades.

O aparecimento da linguaguem é inerente e, portanto, contemporâneo ao desenvolvimento do indivíduo como ser humano. Desde tempos imemoriais, é conhecida a utilização de símbolos para representar os conceitos e os sons. As linguagens especiais de cada tipo de atividade também surgiram em uníssono com a evolução da humanidade. Cada profissão, cada associação, utiliza um vocabulário especializado, que vai se modificando naturalmente, conforme mudam as profissões e as atividades. Portanto, ainda com essa idéia universalista e enciclopedista, desenvolveram-se os dicionários e vocabulários especializados. Foram feitas grandes coisas de ressonância universal, como, por exemplo, as traduções da *Bíblia* para o latim e, posteriormente, para os idiomas romances. A Escola de Tradutores de Toledo foi um centro cultural e humanístico importante, que aproximou a cultura clássica à da Idade Média. No decorrer dos séculos, em épocas diferentes, idealizaram-se sistemas de classificação importantes, como os citados nos parágrafos anteriores. Assentaram-se as bases para o estabelecimento da nomenclatura química. Em anatomia, pode-se citar Versalius, o qual, em meados de 1500, idealizou algumas nomenclaturas sistemáticas. Confeccionaram-se enciclopédias e dicionários que compilavam o conhecimento de seu tempo. Durante tantos e tantos séculos, o homem não se mostrou ocioso e, sobretudo, respondeu ao imperativo de seu desenvolvimento como ser humano em seu caminhar pelo planeta Terra.

A característica de toda essa extensa época histórica, do ponto de vista terminológico, consiste em um desenvolvimento das atividades lingüísticas, de tradução e de classificação, independentes umas das outras, ainda que, em muitos casos, amparadas umas nas outras, como disciplinas de apoio.

PERÍODO DE INICIAÇÃO À TERMINOLOGIA MODERNA

Limita-se este período abrangendo os séculos XVIII e XIX, pois são as épocas em que vão se estabelecendo as bases para o desenvolvimento da terminologia tal como hoje a conhecemos. Segundo todos os documentos consultados, assinados pelos terminólogos mais prestigiados, pode-se situar o nascimento da terminologia com Lavoisier e Bertholler, os quais, no século XVIII, idealizaram uma classificação e nomemclatura para a química. A originalidade de Lavoisier, que era químico e humanista, mas não lingüísta, foi relacionar o conceito com o termo e com o que ele chamou de "motivação do termo", isto é, sua expressividade, seu poder de comunicação. Nesse mesmo século XVIII, Linneo idealizou uma classificação para a botânica, que ainda é utilizada atualmente. Foi no Século das Luzes que as ciências se desenvolveram extensamente e começaram a se relacionar entre si de modo que também os estudiosos dos diversos ramos do conhecimento começaram a manter contato uns com os outros. Por esse motivo, começa-se a sentir a necessidade de dispor de boas traduções, embora fosse o latim o idioma internacional dos estudiosos e pesquisadores.

Não apenas se sente a necessidade de um entendimento internacional nos campos científicos. Na vida cotidiana, e também cultural e humanística, deseja-se dispor de um idioma universal comum para todo o mundo. Por esse motivo, os gramáticos de Port Royal decidiram criar um idioma com essas características, surgindo, assim, o esperanto, que conheceu um certo período de esplendor. Todavia, por ser uma língua artificial, complicada, pouco útil e difundida, sua vida foi fraca. Não obteve a aceitação desejada, como tampouco tiveram outras linguagens artificiais idealizadas posteriormente.

Todos os descobrimentos científicos do século XVIII repercutiram no século XIX, em que se desenvolveu, quase vertiginosamente, uma série de invenções técnicas e industriais que provocaram uma grande evolução lingüística e terminológica. Um fato muito significativo foi constituído pelas atividades no campo da normalização do trabalho, as quais implicavam, também, uma normalização da linguagem e poder dispor de vocabulários e léxicos especializados para realizar boas traduções.

Também em nível de política internacional, sente-se a necessidade de contar com algumas boas traduções. Na Europa, predomina a França, e o francês é o idioma culto internacional, que somente é aprendido por minorias. Portanto, a tradução adquire uma grande importância e, conseqüentemente, os dicionários, vocabulários especializados, listas de termos para as diferentes disciplinas etc. Cabe destacar os trabalhos da International Electrotechnical Commission (IEC), que realizou uma ordenação e planificação lingüística da terminologia para o campo eletrotécnico e estabeleceu as bases para o desenvolvimento atual da terminologia.

Durante esse período de tempo, a terminologia une-se à lingüística, à lógica e principalmente à normalização, com a preocupação colocada nos novos vocábulos, termos, que se torna necessário idealizar e traduzir para que os diferentes idiomas comecem a se modernizar, sem se deteriorar, nem empobrecer. A ênfase volta-se para as ciências naturais, a técnica e a indústria.

PERÍODO DE CONSOLIDAÇÃO

Este período situa-se do princípio do século XX até, aproximadamente, 1930. É chamado de "consolidação", pois foi nesta época que começou a criar-se a consciência de que estava surgindo uma disciplina nova, independente em si mesma, porém relacionada intimamente com a lingüística, a lógica, as diferentes ciências naturais, a normalização e as teorias classificatórias.

Na realidade, são os cientistas e os tradutores que necessitam usar termos apropriados para comunicar-se com seus colegas. Os lingüistas parecem viver em sua torre de marfim, sem descer à realidade cotidiana. A economia é o ramo do conhecimento que produz a aproximação entre ambos grupos de pesquisadores, no momento em que Messing, lá pelos anos 20, realizou um estudo da Wirtschaftlinguistik – Lingüística Econômica, no qual relacionava ambos campos do conhecimento como unidos entre si.

Outro fato que merece ser destacado é a elaboração do *International Electrotechnical Vocabulary*, que começou pela Internatinal Electrotechnical Commission, em 1906, e foi publicado em 1938. Foi uma obra enorme e importante, muito útil em seu tempo.

A obra de Frege sobre lógica *Uber Sinn und Bedeutung* (*Sobre Sentido e Significado*), publicada em 1892, foi o ponto de partida para posteriores estudos de lógica aplicada à terminologia. Talvez tenha sido a origem de uma disciplina, ainda incipiente, que pretendia chamar-se "conceitologia". Quer dizer, fixava-se o alvo no estudo do conceito como centro para desenvolver toda uma conceituação e teoria do próprio "conceito". É possível que se tenha chegado a demonstrar que Schloman deixou-se influenciar por aquela obra quando decidiu ordenar os termos sistematicamente em seu *Ilustrierte Technische Worterbuch* (*Vocabulário Técnico Ilustrado*), segundo seu significado, ao invés de fazê-lo alfabeticamente, como era costume até então.

É interessante constatar como a língua serve de união entre os povos. Sempre que se quis conseguir uma unidade nacional, procurou-se fazer com que todos os povos falassem o mesmo idioma. Os exemplos abundam ao correr da história, tanto em tempos passados como na atualidade, e os temos experimentado dentro de nossas próprias fronteiras. Também os alemães, povos com raízes comuns, porém diferentes entre si, desejam conservar sua unidade procurando unificar a língua, mantendo-a viva e fazendo propaganda da mesma. Em épocas passadas, o alemão era o idioma científico, que se deveria aprender caso se quisesse desenvolver um trabalho de pesquisa ou industrial medianamente digno. Por isso, no início do século, em 1907, o *Verein Deutscher Ingenieuren* (*União de Engenheiros Alemães*), publicou um *Technolexikon*, em alemão, inglês e francês.

Embora os alemães saibam que, atualmente, tenham perdido a batalha diante do inglês, continuam trabalhando nas questões terminológicas. O grupo de terminologia do Deutsche Institut für Normung (Instituto Alemão para Normalização) assim o declara.

Os demais países sentiram, do mesmo modo, a preocupação pela terminologia. Neste período histórico que está sendo abordado, talvez mereça ser citada a França, onde existem diferentes entidades oficiais encarregadas de zelar pela pureza, modernização e difusão de seu idioma. A França, todavia, não acredita ter perdido a batalha e desenvolve verdadeiros esforços para conservar e propagar o francês.

Cabe destacar também a Checoslováquia, onde os seguidores de Ferdinand Saussure haviam criado uma escola com suas teorias lingüísticas, que estabeleceram os princípios para o desenvolvimento da conceitologia atual.

Em 1927, realizou-se, em Washington, 2ª Conferência de Padronização Pan-Americana. Uma das conclusões redigidas no documento final postulava que o espanhol, como o idioma oficial dos países hispanoamericanos, deveria ser estudado e levado em consideração nos trabalhos de Lingüística e Tradução. Deve-se unificar as nomenclaturas dos distintos países e criar vocabulários normalizando os termos. Tratava-se de evitar a diversidade de palavras que se utilizam nos diferentes países para as mesmas coisas e os mesmos conceitos. Para realizar esses trabalhos, deveriam ser criadas comissões de trabalho internacionais, que se reunissem periodicamente e avaliassem seus estudos. Esta reunião semeou o gérmen dos centros de traduções e de estudo da linguagem, criados posteriormente. Como exemplo, citamos o Colégio do México.

PERÍODO DE TEORIZAÇÃO

Nos anos transcorridos entre 1930 e 1970, sobretudo nos primeiros tempos, instituíram-se os princípios teóricos da terminologia moderna. O fato mais revelante deve ser situado na publicação da tese de doutoramento de Eugen Wüsterk sobre *internationale sprachnormung in der technik, besonders in der elektrotechnik* (normalização internacional da linguagem na técnica, principalmente em eletrotécnica). Este documento apareceu em 1931 e deve ser cosiderado como um tratado da Teoria Geral da Terminologia. Nele, relacionam-se a lingüística, a lógica, a ontologia, a conceitologia, as ciências da natureza e a ciência da informação. Alguns autores supõem, e talvez tenham razão, que E. Wüster tomou como pontos de partida as teorias de F. Saussure. Sua originalidade consiste em relacioná-las e aplicá-las às linguagens científicas especializadas, colocando a ciência em contato com a lingüística e a filologia.

As teorias deste estudioso e pesquisador austríaco deram lugar à criação da Escola de Viena e foram a origem da maioria dos estudos realizados posteriormente. Sua importância foi manifestada continuamente e, ainda hoje, continua sendo fonte e base de uma grande parte das pesquisas terminológicas.

A tese de doutoramento de Eugen Wüster foi traduzida para o russo e tomada também como fundamento para as pesquisas realizadas nesse país. Deve-se a Drezd o fato de a União Soviética incentivar a criação de uma comissão de terminologia (TC 37), no seio da International Federation of National Standardizing Association (ISA). Sua missão era a elaboração dos fundamentos terminológicos e sua normalização. Inicialmente, pertenciam à

ISA/TC 37 a União Soviética, Grã-Bretanha, França e Alemanha. Como é fácil compreender, durante o período da II Guerra Mundial, interromperam-se os trabalhos. Foi E. Wüster quem conseguiu, até os anos 50, que a ISA, transformada na International Standard Organization (ISO), criasse um comitê para continuar os trabalhos realizados anteriormente. Na atualidade, a ISO/TS#& é um comitê internacional ativo e dinâmico, onde seus membros realizam um trabalho muito eficaz e elogiável.

O período considerado caracteriza-se por uma grande atividade. Os acontecimentos sucedem-se uns após outros. Nessa época, foram criados ou afirmaram-se as escolas de estudo e pesquisa, cada uma com uma tendência determinada. Já mencionamos a Escola de Viena, da URSS e a Praga. Esta última, com I.K. Weisenberg como colaborador e continuador da obra de F. Saussure, é especializada na lingüística funcional. Realiza pesquisas sobre linguagem normalizada com fins comunicativos. A escola russa dedica-se principalmente à normalização, sem esquecer o tratamento da terminologia do ponto de vista da lingüística. Ocupa-se também de coordenar a teoria e a prática. Como conseqüência, elabora vocabulários especializados e dicionários bi e plurilíngües.

Na França, apareceram várias escolas com origens nos distintos organismos que se ocupam de estudar seu idioma. Entre eles, encontram-se o Instituto da Língua Francesa e o Alto Comissariado da Língua Francesa. A escola francesa, tomada em sentido genérico, ocupa-se do estudo dos neologismos e sua normalização. No que se refere ao francês, adquiriu grande importância a Escola Canadense. Após a promulgação da Carta da Língua Francesa, em Quebec, estabeleceu-se o francês como língua oficial. Posteriormente, decretou-se que tanto este idioma, como o inglês fossem considerados oficiais em todo o país. Desse modo, seus estudos centralizam-se em traduções e nos dicionários bilíngües.

Voltando novamente à Europa, citaremos a Escola Húngara, também centralizada nos problemas da tradução e na criação de novos termos. Por sua vez, os países escandinavos, por iniciativa de Heribert Picht, da Escola Superior de Estudos Mercantis, unem-se em suas tarefas terminológicas, dedicando-se, principalmente, a assuntos econômicos, humanísticos e jurídicos.

Na América Espanhola, a Argentina, o Chile, a Venezuela e o México são os países onde se realizam estudos terminológicos dignos de menção. Em 1954, criou-se a União Latina, com o desejo de defender, conservar e difundir os idiomas dos cinco países de raízes latinas: espanhol, francês, português, italiano e romeno. Atualmente, estabeleceu sua sede na República Dominicana, e seu secretário geral, Philippe Rosillon, realiza um magnífico trabalho, editando publicações periódicas e patrocinando encontros de especialistas. No capítulo seguinte, comentar-se-á sobre o mesmo.

É curioso observar que nem a Grã-Bretanha, nem os Estados Unidos da América interessaram-se de maneira especial por esses assuntos. Naturalmente que se realizam trabalhos terminológicos, centralizados, principalmente, na criação de novos termos, porém não se sente preocupação, nem inquietação por tais questões, na forma como se faz em outros países.

PERÍODO DE EVOLUÇÃO ACELERADA

Nos últimos 18 ou 20 anos, desde 1970, aproximadamente, sucederam-se tantos acontecimentos, que se torna difícil relatá-los em tão poucas páginas. E, ainda mais, diante do temor de cometer omissões involuntárias, esboçaremos uma visão de conjunto. Serão ressaltados os acontecimentos que parecem mais significativos, tanto no desenvolvimento dos próprios princípios teóricos da terminolcgia, como na ampliação de suas aplicações. Também será feita referência à introdução da informática e sua técnicas modernas nos trabalhos terminológicos.

Quanto ao desenvolvimento dos fundamentos teóricos, continuaram sendo estudados principalmente os da Escola de Viena. Helmut Felber foi o continuador de E. Wüster. No capítulo anterior, citamos várias definições de terminologia e de termo devidas a esse autor. Atualmente, é Christian Galinski, o sucessor de H. Felber, que continua pesquisando tais questões. Em maio de 1989, expôs sua idéias mais recentes, em um seminário realizado em Madrid, organizado pelo grupo de terminologia do Cindoc. Estabeleceu uma diferença entre terminologia e ciência da terminologia. Segundo ele, a terminologia faz referência ao conjunto de termos sobre um determinado assunto com ou sem definições, com ou sem equivalências em outros idiomas. Aqui, a palavra "terminologia" faria referência ao segundo tipo de acepção e seria sinônimo de vocabulário especializado ou de dicionário. Galinski define a ciência da terminologia como o corpo da doutrina metodicamente formado que estuda a terminologia.

É possível que se me escape alguma argumentação, porém compreendo que, etmologicamente considerada, "terminologia" = termo + logos, supõe o tratado, o estudo do termo. Assim, pois, a ciência da terminologia se referiria à ciência da ciência do termo. Talvez Ch. Galinski quisesse dar-lhe a importância de uma ciência da ciência.

Outro acontecimento importante produzido nos últimos anos é a ampliação do campo da terminologia, abrangendo, sobretudo, até os sistemas classificatórios. Acredito que, há 20 anos, quando uma das primeiras figuras alemãs dedicadas ao estudo da classificação, Ingetraut Dahlberg, pesquisava sobre seus princípios ontológicos ou sobre as grandes classificações universais, não estava conscientizada de estar realizando estudos sobre terminologia. Na minha opinião, o uso de expressões como "linguagens terminológicas", aplicadas aos sistemas de classificação, começou a ser feito como conseqüência do estudo das classificações de assuntos utilizando palavras-chave e os tesauros. Finalmente, um tesauro é um vocabulário especificado, normalizado, no qual as palavras que o compõem adquirem a categoria de termos, com suas relações semânticas associativas, hierárquicas e de equivalência. Deve-se estudar esses termos cuidadosamente, inclusive muitas vezes, defini-los e traduzi-los. Aqui é onde entra a terminologia. Se um tesauro é uma linguagem terminológica e um sistema de classificação, qualquer outro sistema classificatório será, deste modo, uma linguagem terminológica.

A mesma argumentação se pode aplicar aos dicionários, vocabulários e léxicos especializados, utilizados para realizar traduções, trabalho onde é preciso seguir uma normalização das palavras, convertidas agora em termos. Existirão diferentes tipos de

linguagens terminológicas, cada uma delas com suas características específicas e seu campo de aplicação.

Relatemos os acontecimentos mais relevantes ocorridos durante esse período. Uma realização que se tornou um marco importante foi a criação, em 1971, do centro Infoterm, relacionado com o programa Unisist, subordinado à Unesco. Trata-se de um centro internacional de informação sobre terminologia, com sede em Viena e continuador da Escola de Viena. Atualmente, conta com um banco terminológico muito amplo, que originou a rede TermNet. Edita, entre outras publicações, a revista *TermNet News*. Em 1984, foi editado um manual de terminologia, de autoria de Helmut Felber, que se tornou a obra teórico-prática, guia e compêndio para o estudo atual deste campo do conhecimento.

A criação de Infoterm influenciou a criação de centros nacionais de terminologia em outros países. Por exemplo, na França, foi criada a Afterm; na Alemenha Federal, a Sprachteor, hoje, desaparecida; nas Escandinávia, a Nordterm. A Espanha será abordada em parágrafo separado.

No âmbito hispano-americano, a Comissão Econômica para a América Latina (Cepal), criou, em 1972, o Centro Latinoamericano de Documentação Econômica e Social (Clades). O Centro possui um sistema de informação terminológica que publica terminólogos. Também elaborou um macrotesauros, muito difundido e utilizado. O Colégio do México continua seus trabalhos, elaborando vocabulários e realizando pesquisas em normalização e na formação de terminólogos e tradutores. Na Argentina, a Unesco patrocinou, em 1983, a criação do Serviço Latino-Americano de Pesquisa e Tradução, que, em 1986, foi reorganizado e enquadrado dentro de seu programa geral de informação. Desde 1981, na Venezuela, funciona o Grupo de Informação Terminológica da Universidad Simón Bolívar, de Caracas. Trabalha conjuntamente com o Departamento de Línguas. Suas pesquisas são orientadas para os princípios teóricos, mas elabora, também, terminologias. Em abril de 1983, organizou o 1° Seminário Nacional de Terminologia, em cujas atas constam exposições sobre os fundamentos teóricos, a história e a prática dessas matérias. No Brasil, a Embratel, em Brasília, é uma das organizações que mais se interessam pelos assuntos terminológicos. Outros países ibero-americanos também possuem organizações dedicadas à terminologia e estão em contato com as instituições internacionais.

Agora, voltaremos a abordar a União Latina, que criou delegações em países distintos. Uma delas situa-se em Paris. Seu secretário geral chama-se Daniel Prado, e os programas de trabalho compreendem várias ações. Entre elas, podem ser citadas as exposições de lingüística informática e a elaboração de mapas lingüísticos do mundo e de vocabulários científicos e técnicos. Promove também a criação de conselhos internacionais de coordenação das atividades de elaboração e normalização de terminologia. Entre outras publicações, figuram as revistas *Terminómetro*, com informações e notícias, e *Cortocircuito* com informações sobre a atualidade do mundo audiovisual latino na América e Europa. Ambas as revistas são publicadas nos cinco idiomas dos países neolatinos.

Na América do Norte, devemos ressaltar as atividades do Canadá, que organizou, entre outras atividades, o Bureau of Translations, com a criação de bancos terminológicos, em francês-inglês, de uso internacional.

No plano internacional, merecem ser citadas as atividades do Comitê ISO/TEC 37, que continua em atividade. Por sua parte, a Federação Internacional de Informação e Documentação (FID) criou, em 1972, um comitê, FID/DT, dedicado a esses assuntos. No decorrer desses anos, tem realizado trabalhos muito importantes, entre as quais citaremos: o *Dicionário de Termos de Informação* (em 14 idiomas); *Bibliografia sobre Terminologia em Documentação*; *Análise da Pesquisa sobre Teoria e Prática da Terminologia em Documentação*; *Dicionário de Abreviaturas*; *Recomendações e Diretrizes para a Recompilação e Criação de Termos*; *Estudo de neologismos*; *Dicionário de Termos sobre Informação* (inglês e japonês); *Vocabulário Sistemático em Cinco Idiomas* (inglês, francês, alemão, russo e espanhol) *para melhor uso da Classificação Decimal Universal (CDU)*; Publicação da série de pesquisa intitulada *Essential Problems in Terminology for Informatics and Documentation*. Posso falar com detalhes sobre esses trabalhos que conheço bem, por ser membro deste comitê da FID.

Desde 1970 até o momento, as atividades terminológicas cartacterizaram-se, do mesmo modo, pela aplicação da informática e das telecomunicações. Criaram-se bancos de dados e redes nacionais e internacionais para assuntos especiais e alguns abrangem vários campos do conhecimento. Citemos, como exemplo, o sistema EURODICAUTOM, criado pelos países da Comunidade Européia, que consiste em um banco de dados terminológicos e de um método de tradução automática, para os 12 idiomas da Comunidade.

Também nestes anos, aplicam-se as técnicas da inteligência artificial e planejam-se sistemas especialistas para serem usados em trabalhos de tradução. A tecnologia foi tão aplicada e tanta repercussão está tendo na indústria e no mundo dos negócios, que hoje se fala da engenharia e da indústria da língua.

Os acontecimentos no campo terminológico sucedem-se continuamente. Mencionaremos a Exposição de Lingüística Informática e de Terminologia Científico-Técnica, organizada pelo Conselho Superior de Investigações Científicas (CSIC) e pela União Latina, realizada em Madrid, em fevereiro de 1987. No final das sessões de estudo, foi redigido o Manifesto de Madrid. No documento, defendia-se a sustentação do espanhol (castelhano) e insistia-se com os governos para conscientizarem-se da importância da terminologia. Como conseqüência, para "formularem" políticas racionais em uma margem maior que suas políticas lingüísticas, solicitando-se que "apóiem e financiem ações encaminhadas para estudar, criar, unificar, normalizar e difundir terminologias em nossa língua comum e a formação de profissionais na especialidade".

Em setembro de 1987, realizou-se, em Trier, Alemanha Federal, um congresso dedicado à *Termnology and Knowledge Engineering* (Terminologia e Engenharia do Conhecimento), no qual foram apresentados trabalhos bem interessantes sobre terminologia informática e

engenharia das línguas. Outro acontecimento importante foi a celebração do 1º Simpósio Latino-Americano de Terminologia, em abril de 1988, em Caracas, organizado pelo Grupo de Investigação em Terminologia, da Universidad Simón Bolivar, de Caracas, e pela União Latina. Desse encontro resultou a criação de uma Rede Ibero-Americana de Terminologia (RITERM), para realizar trabalhos conjuntos de pesquisa e formação de terminólogos. Em Madrid, em dezembro de 1988, realizou-se um colóquio sob o tema "Madrid, Capital da Terminologia Européia", organizado pelas sociedades Logomotiv e C + L e pelo grupo de pesquisa terminológica TermEsp, do Cindoc, que reuniu, entre outros, terminólogos portugueses e espanhóis para tratar de trabalhos e questões comuns. Um dos acontecimentos mais relevantes teve lugar na primavera de 1989, em Túnez, e significou um primeiro passo para a criação de uma rede euro-árabe de terminologia. As reuniões, simpósios, colóquios e congressos sucedem-se continuamente, o que torna complexo relatar todos eles. Aconselha-se consultar as publicações especiais, como *Terminómetro*, ou *TermNet News*, por exemplo, para obter uma boa informação.

Como final deste título, citam-se as sociedades profissionais de maior renome internacional, entre as quais estão as de tradutores. Estes trabalham em íntima relação com os terminólogos, pelo menos para fazer uso de vocabulários, léxicos e bancos de terminologia: Associação Internacional de Lingüística Aplicada; Federação Internacional de Tradutores (FIT); Associação de Universidades Parcial ou Inteiramente de Língua Francesa (AUPELP); Associação Internacional de Termninologia de Quebec (TERMIA); Liga Internacional de Cientistas utilizadores da Língua Francesa; Associação de Terminologia e Transferência do Conhecimento, criada em Trier, em 1986.

A TERMINOLOGIA NA ESPANHA

Há muito tempo, realizam-se trabalhos terminológicos na Espanha, principalmente na Real Academina de la Lengua. Contudo, no que se refere à terminologia moderna, o seu começo pode ser dado como 1962, data em que a Real Academina de Ciencias Exactas, Físicas y Naturales, por iniciativa de Julio Garrido, criou um grupo de trabalho sobre terminologia, com a missão expressa de elaborar um vocabulário científico e técnico. Esse vocabulário foi editado em 1983. Atualmente, publica uma segunda edição. Como colaboradora de Julio Garrido, tenho a honra de participar da elaboração dessa obra.

Outra data importante pode ser 1970, quando se realizou uma reunião franco-espanhola em Barcelona, que deu lugar à criação de Fonds International des Terminologies Romanes (FITRO), ao qual se uniram outros países latinos. Em 1972, voltaram a se reunir em Florença, com a participação do Instituto Miguel de Cervantes, do Consejo Superior de Investigaciones Cientificas, cujo Diretor era Manuel Criado del Val. Em 1977, realizaram-se em Madri as Jornadas Internacionais de Investigação Humanística, que incluíam uma mesa-redonda sobre terminologia. Nasceu, então, a idéia de criar um centro de terminologia científica e técnica de espanhol. A idéia cristalizou-se, e criou-se o Hispanoterm, com Manuel Criado del Val como presidente ou diretor. Sua missão mais importante era elaborar um banco de dados terminológicos, que vinha se realizando manualmente durante anos. Posteriormente, passou

a constituir-se em sociedade sem fins lucrativos e de utilidade pública. Tenho a honra de participar de sua organização e de pertencer à sua Junta Diretiva desde o seu início. Em 1983, realizou-se, em Madri, o 1° Seminário de Terminologia, organizado por Hispanoterm e Infoterm. Helmut Felber e Heribert Picht foram convidados a participar.

Hispanoterm teve de enfrentar numerosas dificuldades e incompreensões por parte das autoridades do Conselho Superior de Investigações Científicas. Atualmente, não desapareceu, porém a aposentadoria de Criado del Val implicou uma nítida diminuição de suas atividades. É compreensível que, atualmente, não se pode realizar um banco de dados da importância pretendida por Hispanoterm sem contar com os recursos da informática.

O Consejo Superior de Investigaciones Cientificas (CSIC), em 1981, reorganizou seus programas de pesquisa e incluiu o 1° Programa de Pesquisa, Estudo e Coordenação da Terminologia Científico-Técnica Espanhola, com o qual deviam colaborar tanto as instituições humanísticas como as científicas. Em 1985, o CSIC aprovou dois programas para a normalização da língua espanhola e encarregou o Cindoc e o Isoc de sua realização, com um amplo plano de atividades nacionais e internacionais. Citamos, como exemplo, a introdução do espanhol no *Dicionário de Informação e Documentação*, elaborado pelo comitê FID/DT, versão espanhola da base de dados terminológicos do Eurodicautom, e o *Catálogo de Recursos Terminológicos em Língua Espanhola*. A realização desses trabalhos ocasionou a criação de um grupo de terminologia, no Cindoc, dirigido por Amélia de Irazazábal. Atualmente, denomina-se TermEsp. Para obterem-se detalhes de todas as suas variadas atividades, programas de cursos, reuniões etc., aconselha-se dirigir-se diretamente a TermEsp.

A mudança constitucional da Espanha para converter-se no Estado das Autonomias deu como um dos resultados a exaltação das línguas de âmbito regional. Assim, a Academia de la Lengua Gallega está modernizando e atualizando a gramática e o léxico do gallego. Na Cataluña e em Euskadi, respectivamente, foram criadas as entidades Tercat e Euskalterm, onde se realizam pesquisas, procuram-se novos termos e se elaboram bancos de dados terminológicos.

No que se refere à aplicação da informática e das novas tecnologias, a Espanha realiza pesquisas sobre o tratamento das linguagens naturais, sobre a tradução automática e sobre a elaboração de tesauros especializados. Elaborou-se um dicionário do castelhano em computador para a composição e verificação de textos, pelo grupo dirigido por Luis Sopeña, do Centro de Investigación IBM-Universidad Autónoma de Madri-UAM. Este mesmo pesquisador idealizou um sistema especialista para o reconhecimento de formas flexionadas dos morfemas. No Instituto de Estudos Avançados de Blanes, estuda-se a realização de um sistema especialista para a elaboração de tesauros especializados.

Existem, na Espanha, outros vários organismos que realizam trabalhos terminológicos, entre os quais deve-se mencionar a Associação Espanhola de Tradutores e Intérpretes (Apeti) e a Escola de Tradutores de Granada, por parte de tradutores. Durante as jornadas do Simo, de 1988, realizou-se uma mesa-redonda sobre terminologia, da qual participaram todos os grupos ocupados com o assunto.

TEMAS PARA ESTUDO

1. Diga o que é TermEsp.

2. Cite algum centro de informação terminológica.

3. Diga que instituições trabalham em terminologia na Venezuela.

4. Quem é considerado o pai da terminologia moderna?

5. Diga algo sobre o Colégio do México.

6. Por que pensa que o esperanto não se difundiu?

7. Nomeie alguma escola de tradutores.

8. Nomeie alguma escola de estudos terminológicos.

9. Nomeie algumas instituições internacionais dedicadas à terminologia.

10. Diga o que é o Manifesto de Madrid.

11. Cite algum centro dedicado principalmente a realizar traduções.

12. Que significa a bonita palavra Eurodicautom?

13. Cite algum membro da Escola de Praga.

14. Diga o que é Infoterm.

15. Diga o que é União Latina.

16. Cite alguma publicação dedicada aos assuntos terminológicos.

17. Cite alguma instituição brasileira e espanhola que trabalhe em terminologia.

REFERÊNCIAS BIBLIOGRÁFICAS

ALONSO BAQUER. M.: "Terminología de la historia militar"; en Introducción a las Fuentes y Métodos de la Historia Militar; Seminario organizado por el Ilustre Colegio Oficial de Doctores y Licenciados en Filosofía y Letras y en Ciencias de Madrid, 11-14 de abril, 1989.

CABRE I CASTELLVI, T.: "Realidad de TERMCAT", en Actas de la Exposición de Lingüística Informática y de Terminología Científico-técnica, CSIC y Unión Latina, Madrid, 1987, 85-86.

Coloquio sobre la historia de la terminología; Terminómetro, *VI*, 1, 1988, 16.

FEDOR DE DIEGO, A.: "Lingüística informática, lexicografía general y terminología científico-técnica en castellano", en Actas de la Exposición de Lingüística y de Terminología Científico-técnica, CSIC y Unión Latina, Madrid, 1987, 166-169.

FONDAZIONE UGO BORDONI: "Comunicazione concernente la lingüística computazionale in Italia"; en Actas de la Exposición de Lingüística y de Terminología Científico-técnica, CSIC y Unión Latina, Madrid, 1987, 157`158.

GORKOVA, V.I.: "FID - Terminology of Information and Documentation Committee (FID/DT); Int. Forum Inf. and Docum. *11*, 3, July 1986, 39-42.

"Infoterm, TermNet y TermNet News"; Terminómetro, *IV*, 2, 1984, 3.

IRAZAZABAL, A.:"Presentación del Grupo de Terminología TERMESP del ICYT (CSIC)"; en Actas de la Exposición de Lingüística y de Terminología Científico-Técnica, CSIC y Unión Latina, Madrid, 1987, 107-111.

IRAZAZABAL, A.: "La Terminología en Espana y en el mundo", en Apuntes del Curso de Introducción a la Terminología, ICYT, Madrid, 1988, 132-139.

MENCIONI, G.: "Recherches lexicographiques en Italie", en Actas de la Exposición de Lingüística y de Terminología Científico-Técnica, CSIC y Unión Latina, Madrid, 1987, 137-139.

PICHT, H.: "Breve historia y situación actual de la teoría, la investigación y la práctica terminológica"; Actas del Primer Seminario Nacional de Terminología; Caracas-Venezuela, Universidad Simón Bolívar, junio 1984, 24-32.

"La terminología en Canadá"; Terminómetro, *VI*, 1, 1988, 12.

UTNE, I.: "Technical Terminology at the Norwegian termbank – a basis for knowledge representation and information retrieval"; in Terminology and Knowledge Engineering, INDEKS Verlag, Frankfurt.

VIDAL CASTRO MELO, M.:"A Linguistica informática no Brasi"; en Actas de la Exposición de Lingüística y de Terminología Científico-Técnica, CSIC y Unión Latina, Madrid, 1987, 148-150.

3

As Indústrias da Língua

A expressão "as indústrias da língua", com a qual começamos o capítulo, merece algumas considerações. Já estamos acostumados com o seu uso e talvez não nos tenhamos dado conta de todo o conteúdo conceitual que encerra. Em primeiro lugar, nos fala da "indústria", que está relacionada com o comércio, a produção, o mercado, a técnica... Depois, aparece a "língua", expressão genuína de nosso pensamento, de nossa maneira de sentir, de atuar... É como se a técnica e as humanidades estivessem enfrentando-se.

À primeira vista, a indústria da língua parece um tremendo contra-senso: como é possível industrializar a língua? Todavia, devemos admitir que se pode realizar e é um fato real com o qual esbarramos diariamente. Não somente se industrializa a língua; fala-se também da engenharia da língua e da linguagem, da informática lingüística informatizada e da lexicografia computacional. Todas estas expressões refletem uma característica de nosso tempo, que faz referência à tecnologia aplicada às atividades cotidianas, sejam elas quais forem. Vivemos em um mundo de máquinas e engenhocas, com as quais pretendemos tornar a existência algo mais cômoda, mais fácil de ser vivida (figura 3.1).

Consideremos, por outro lado, "a indústria da língua" sem esssas conotações de engenharia, informáticas e de aplicações das novas técnicas eletrônicas. Então, perceberemos que a língua sempre esteve relacionada com a indústria; Enquanto a produção de dicionários, vocabulários, enciclopédias impressas, por exemplo, está sujeita aos imperativos das demandas comerciais, necessidades concretas de grupos de especialistas, situações econômicas – sem esquecer toda a complexa maquinaria necessária para a impressão, incluídas as indústrias químicas de tintas e tinturas, e todo o material necessário para que um livro possa chegar às livrarias para ser comprado por um leitor. Não apenas se imprimem e editam obras técnicas, como também a bela poesia, as novelas e tantas e tantas outras manifestações literárias estão sujeitas aos imperativos da indústria.

Assim, pois, entendemos a indústria da língua em toda a sua plenitude, embora, neste caso concreto, iremos nos referir à aplicação das novas tecnologias, incluída a informática, ao tratamento da língua para a sua melhor utilização.

Indústria da língua ou da linguagem? Para o propósito desta obra, devemos considerar que se trata de expressões sinônimas. Sabemos que, estudando o assunto com rigor lingüístico, existem determinadas diferenças derivadas das definições de língua e linguagem, conforme os diferentes critérios e teorias existentes.

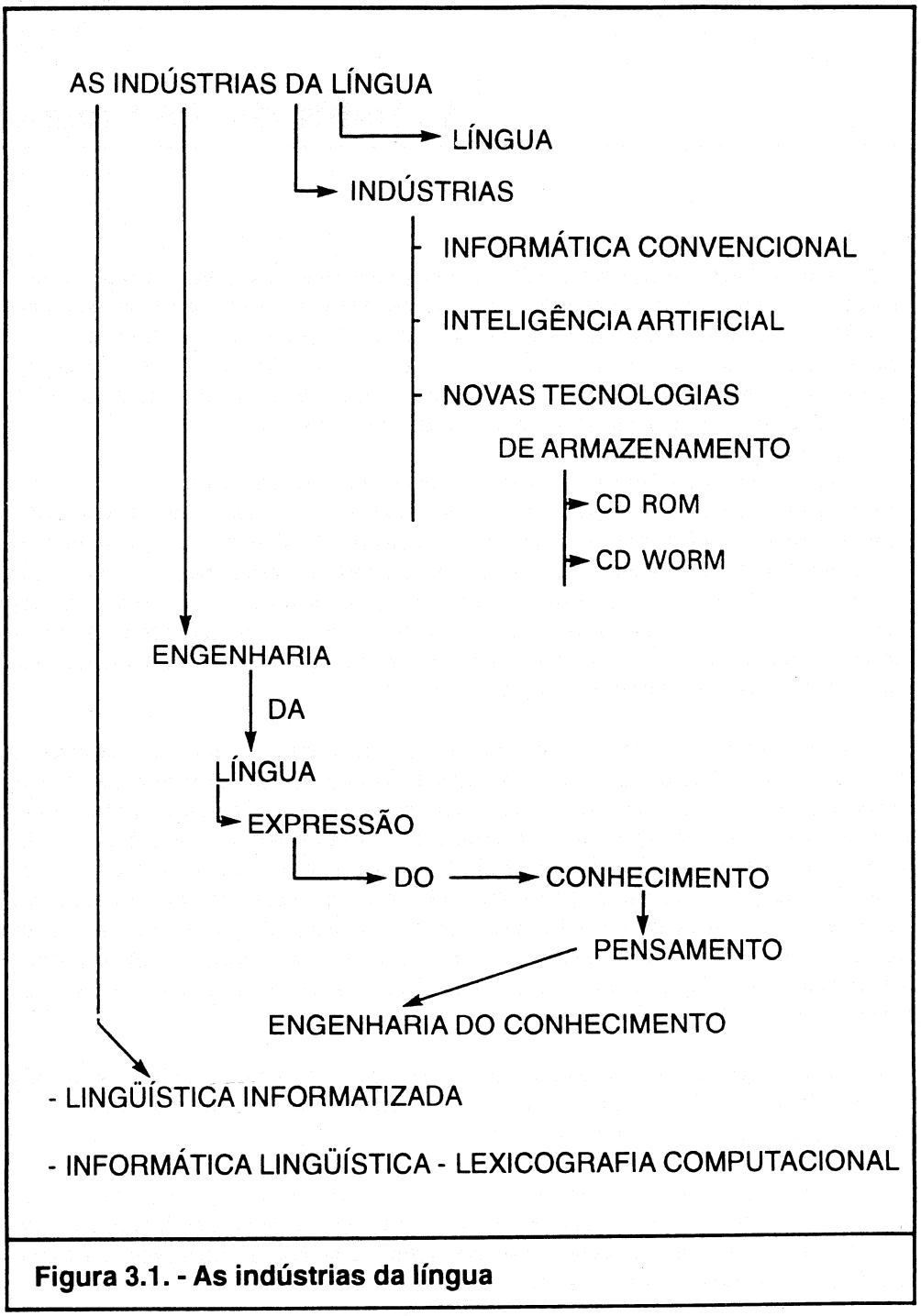


Figura 3.1. - As indústrias da língua

No título do capítulo, figura a expressão no plural: "indústrias da língua", porque faremos referência a mais de uma indústria. Abordar-se-á a informática convencional, a inteligência artificial, com seus sistemas especialistas, e as técnicas de armazenamento em discos compactos.

De alguns anos para cá, pesquisas estão sendo realizadas sobre todos esses assuntos. Como exemplo significativo, citaremos o Observatório Francês da Indústria da Linguagem, criado pelo Secretário de Estado encarregado da Francofonia e o Comissário Geral da Língua Francesa. Compila, ainda, todos os métodos e técnicas modernos que aparecem em qualquer parte do mundo. Em prosseguimento, seus assuntos de pesquisa são citados:

- reconhecimento de linguagem natural;
- tradução automática;
- interfaces aplicáveis à linguagem natural;
- tratamento de escrita impressa e manuscrita;
- análise e criação de textos;
- novos meios de comunicação;
- informática documental;
- editoração eletrônica.

Poderíamos ter citado, igualmente, qualquer outra entidade dedicada ao estudo e à pesquisa, já que as questões principais não variam. Tomemos como exemplo o Centro de Investigações para as Indústrias da Linguagem (Cipil), da Argentina, que inclui também:

- teoria da lingüística computacional;
- construção de bancos de dados terminológicos;
- estudo de implantação de instalações terminológicas.

O Programa da Língua da Comunidade Européia organizou uma sessão aberta de opiniões (HEARING), na qual se decidiu acrescentar, aos assuntos básicos, os seguintes:

- automação de escritórios;
- publicidade;

- ensino de idiomas;
- formação de terminólogos.

APLICAÇÃO DAS NOVAS TECNOLOGIAS

Neste segmento, trataremos da aplicação das novas tecnologias de uma maneira geral, quanto à utilização que se faz de cada uma delas nos diferentes usos da terminologia.

A informática convencional, assim chamada porque serve como base para posteriores estudos mais avançados, possibilitou a construção, entre outras aplicações, de bancos de dados terminológicos. Deste modo, criaram-se bases de dados bibliográficas ou referenciais sobre terminologia. Ambos, bases e bancos de dados, podem ser consultados mediante o acesso em linha ou à distância, por telefone ou correio. Trouxeram uma grande ajuda para o trabalho terminológico. Os lingüistas, os pesquisadores e os tradutores beneficiaram-se com o seu uso que lhes facilitou as tarefas.

No seu momento, o aparecimento das técnicas de inteligência artificial, com os programas de sistemas especialistas, possibilitou, igualmente, a realização de trabalhos terminológicos de modo mais cômodo e com maior confiabilidade.

A construção de sistemas especialistas é uma técnica que ainda se encontra em período de desenvolvimento. A cada dia, aparecem novos sistemas especialistas, com um número maior de aplicações e de mais fácil realização. No que concerne à terminologia, os estudos de treinamento da linguagem natural e de treinamento de texto são os mais apropriados e relevantes. Na bibliografia do capítulo, citam-se as obras onde são dadas informações sobre os sistemas especialistas, os quais aplicados a esta disciplina, já somam um número elevado. (*)

Neste trabalho, são mencionados alguns sistemas especialistas, como, por exemplo, o sistema PRIMUS, que se trata de um sistema especialista japonês destinado à automatização do Sistema de Classificação Decimal Universal (CDU) e à classificação de documentos com o auxílio do computador; o sistema MEDES, o qual é um sistema especialista destinado ao processamento da informação médica, realizado em Groningen (Países Baixos); o sistema LOTOS, que é um sistema especialista idealizado para a produção de um método de tradução automática na área de transportes, desenvolvido pelo grupo internacional de especialistas da International Standard Organization (ISO).

A nova técnica, que revolucionou os métodos de armazenamento e de recuperação da informação de documentos e dados, permite o uso de discos compactos. Estes discos são gravados e lidos por intermédio de procedimentos com base nos raios laser. Até o momento, foram comercializados os CD-ROM (Compact Disk-Read Only Memory), que são discos nos quais não se pode apagar o que está gravado. Admitem uma grande quantidade de dados. É conhecido por todos que a *Enciclopédia Britânica* não ocupa nem mesmo um disco. Por

(*)Em meu trabalho ***Estado das pesquisas sobre os sistemas especialistas na Espanha*** (1988), encontram-se detalhes sobre estas questões.

esse motivo, são sumamente úteis em terminologia. É possível gravar dicionários, bancos de dados, enciclopédias e outros e, logo, obter informação confiável e rápida. Todavia, são caros, motivo pelo qual seu uso se restringe às grandes instituições. Existe um disco CD-ROM que contém um dicionário multilíngüe para os sistemas vídeo em inglês, francês, alemão, espanhol, italiano, holandês, japonês e chinês, elaborado por Zanichelli. É um exemplo de um caso especializado dedicado a um assunto bem concreto.

Para os sistemas de informação convencionais, os CD-ROM apresentam o inconveniente de que os documentos ou dados gravados ficam ali, ocupando espaço para sempre, o que pode não ser desejável. As técnicas modernas já estão oferecendo os CD-ROM, ou, simplesmente, Wright-Read Memory - WORM, que podem ser apagados. Ainda custam muito caro e são algo imperfeitos. As pesquisas futuras possibilitarão que o seu uso venha a ser do domínio de todos, beneficiando muito os trabalhos terminológicos.

BANCOS DE DADOS TERMINOLÓGICOS

Uma das aplicações mais relevantes dos bancos de dados é seu uso em terminologia para construir bancos de dados terminológicos. Aqui se registra o termo, ou o conceito, e se acrescentam as referências e notas auxiliares necessárias e convenientes para cumprir sua finalidade.

Portanto, um

Banco de dados terminológicos

é um registro de conceitos ou termos, realizado com a finalidade de ser utilizado por pesquisadores ou tradutores, onde os dados são acompanhados de referências, definições e outras informações auxiliares para sua melhor compreensão e utilização.

Para completar o conceito de banco de dados terminológicos, se faz referência a alguma outra definição, entre as várias que podem ser lidas nas obras de consulta. Cita-se como exemplo a do grupo de trabalho de terminologia TermEsp, do Conselho Superior de Pesquisas Científicas da Espanha, dada a seguir:

- Está composto por registros e entradas, onde cada registro é uma unidade lexicográfica, de um campo específico do conhecimento, acompanhada dos termos a ela atribuídos e outros dados complementares. Pode tratar-se de unidades isoladas como tais, ou de frases (apontamentos do Curso de Terminologia, 1989).

Outra definição que também merece ser citada é a do Grupo de Trabalho sobre Linguagens de Indexação, do Caicyt, da Argentina, que lhe confere um sentido utilitário, prático, quando afirma que,

- é um "arquivo automatizado de termos científicos e técnicos, utilizado como ferramenta para conhecer os verdadeiros significados dos termos em sua aplicação profissional".

Note-se que somente faz referência aos termos científicos e técnicos ... Certamente, existem também bancos de dados terminológicos para outras muitas disciplinas e em todo o campo do conhecimento.

Para não nos concentrarmos unicamente nos pesquisadores de fala espanhola, cita-se, agora, a definição de Elisabeth Maier e Renate Mayer, do Instituto para la Organización y Economía del Trabajo, de Stuttgart, República Federal da Alemanha, que afirma tratar-se de

- um conjunto de termos armazenados na memória do computador, recuperáveis por meios informáticos, onde devem figurar as definições, origem "e outra série de dados auxiliares".

É curioso que cite "meios informáticos" para a recuperação, e não para a composição – armazenamento – dos dados.

Todas essas definições coincidem em seus pontos mais importantes. Trata-se de um registro de unidades – conceitos ou termos – e alguns dados auxiliares. É interessante, pois, saber quais são esses dados e como se constrói um banco de dados terminológicos, assim como quais são as classes desses bancos que existem.

TIPOS DE BANCOS DE DADOS TERMINOLÓGICOS

No mundo terminológico, existem vários tipos de bancos de dados, segundo sua construção, sua finalidade e sua temática. Em primeiro lugar, deve-se salientar que estamos falando de procedimentos informáticos para sua elaboração. Efetivamente, quase não se concebe, atualmente, que possam ser realizados de outra maneira. Contudo, segundo as circustâncias, poderá ser aconselhável, inclusive, construí-los manualmente, portanto uma primeira divisão seria:

bancos de dados terminológicos

- manuais,
- informatizados.

Nas definições, vimos que se fala de "conceitos" e "termos". Quer dizer que, em alguns casos, o registro será formado por um conceito e, em outros, por um termo. A finalidade de seu uso será, pois, diferente. Por conseguinte, devemos considerar que existem vocabulários e dicionários. Os pesquisadores e tradutores utilizarão uns e outros respectiva ou indistintamente:

- vocabulários
  registro = conceito

- dicionários
  registro = termo.

No que concerne ao âmbito do assunto tratado, poderão ser:

- universais
  - multidisciplinares;

- especializados
  - unidisciplinares.
  - pluridisciplinares.

Não é fácil encontrar bancos de dados do tipo universal. Na realidade, não possuem muito sentido, porque se trata é de precisar o uso do termo para uma atividade concreta, em um léxico especializado. Contudo, existem alguns que cobrem duas ou três – ou muito poucas – disciplinas que estão estreitamente relacionadas entre si.

Deve-se levar em consideração, deste modo, se tais bancos de dados contêm equivalências dos termos, ou conceitos, em outros idiomas, e, neste caso, se falará de:

bancos de dados terminológicos

- monolíngües,
- bilíngües,
- multilíngües.

Os bancos de dados terminológicos do tipo vocabulário serão, na sua maioria, monolíngües, embora muitas vezes contenham também equivalências em outros idiomas. Por sua vez, os bancos de dados do tipo dicionário serão, ao menos, bilíngües. De fato, a maioria deles é multilíngüe, se a sua construção o permitir, pois são mais úteis na realização de traduções.

Todavia, falta-nos considerar a finalidade de um banco de dados terminológicos, quanto ao seu uso, do ponto de vista relativo ao pesquisador e usuário. Se os mesmos se destinarem, por exemplo, à pesquisa e normalização de linguagens especializadas, dentro de um idioma determinado, deverão ser estudados os desenvolvimentos da linguagem, os neolismos, as definições de novos termos etc. Do mesmo modo, dever-se-á levar em consideração o uso dos termos na tradução, onde o que interessa, sobretudo, é a equivalência dos termos, ou conceitos, em outros idiomas. Assim, pois, do ponto de vista de sua finalidade, dedicar-se-ão à:

- pesquisa,
- normalização,
- tradução.

Os bancos de dados terminológicos do tipo vocabulário serão utilizados, principalmente, na pesquisa e normalização; e os do tipo dicionário serão utilizados na tradução.

Talvez se pudesse mencionar algum outro tipo de bancos de dados terminológicos, para completar o panorama de sua tipologia. Contudo, o exposto parece suficientemente ilustrativo para atingir as idéias de conjunto (figura 3.2).

CONSTRUÇÃO DE UM BANCO DE DADOS TERMINOLÓGICOS

Depois de conhecer os tipos de bancos de dados terminológicos que podem ser necessários para desenvolver uma eficiente tarefa terminológica, trata-se, agora, de sua realização. Para isso, deve-se considerar alguns processos fundamentais, tais como:

- coleta das unidades léxicas,
  - conceitos,
  - termos,
  - frases terminológicas;
- formulação das definições;
- estabelecimento da origem das unidades léxicas;
- fixação das equivalências em outros idiomas;
- indicação da data em que o termo foi encontrado;
- indicação da data de elaboração do banco de dados;
- atualização;
- colocação do nome do autor do banco de dados.

Supondo-se que se trata de confeccionar um novo banco de dados terminológicos, o primeiro passo será averiguar se já existe algum em uso em algum local. Neste caso, tratar-se-ia somente, talvez, de uma adaptação, modificação ou atualização.

Contudo, se for necessário iniciar o trabalho desde o princípio, será necessário consultar os dicionários e vocabulários existentes na bibliografia para recolher os conceitos, termos ou frases. Será necessário, também, ler os textos de trabalhos, livros, teses de doutorado e outros documentos sobre a especialidade abordada. É muito difícil precisar quando foi conseguido o número suficiente de termos para concluir que se dispõe de uma terminologia completa sobre esse assunto. Se este for um assunto muito novo, não se encontrarão muitos termos, e, caso seja um assunto já conhecido, sentir-nos-emos sobrecarregados por sua quantidade. Esta não é uma questão que deva nos preocupar demasiadamente. O importante é levar em consideração que, muitas vezes, ao mesmo tempo em que encontramos o conceito, dispomos dos demais dados necessários, como sua definição, o documento onde foi localizado e sua data. Não se deve, pois, ir marcando etapas separadamente. São processos simultâneos e paralelos. Portanto, deve-se ir trabalhando continuamente e anotanto os dados que vão surgindo.

Tornar-se muito difícil elaborar fichas terminológicas para a coleta de dados. É aconselhável utilizar uma folha no tamanho UNE(DIN)A4, onde devem constar os dados de que se necessita e que vão sendo escritos conforme aparecem. Um modelo dos dados que deve figurar na ficha terminológica poderia ser o seguinte:

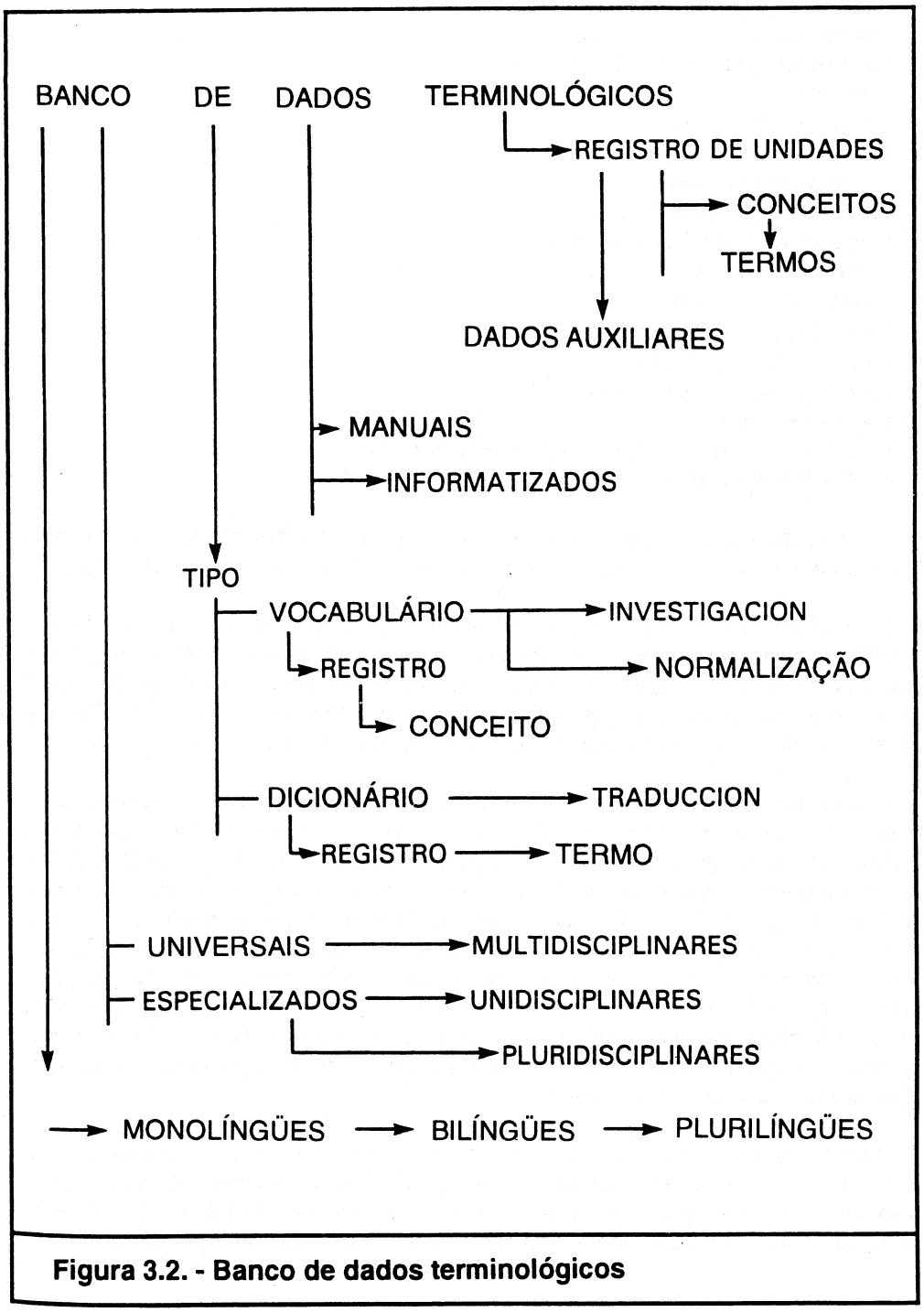


Figura 3.2. - Banco de dados terminológicos

- número de identificação do registro;
- unidade léxica;
- documento (s) (*) onde foi encontrado;
- definição;
- origem da definição;
- campo temático;
- designação do conceito;
- forma léxica de onde deriva;
- raiz e origem da palavra utilizada;
- designação do termo;
- forma léxica do termo;
- documento(s) onde o termo foi lido;
- data deste(s) documento(s);
- equivalente(s) em outro(s) idioma(s);
- data de anotação;
- .....(espaço para modificações ou atualizações);
- nome da pessoa que realizou as anotações.

No caso em que se irá informatizar o banco de dados terminológicos, será necessário acrescentar os dados e as anotações apropriadas para obter logo as informações desejadas.

Nem sempre será preciso fazer constar todos os dados aqui citados. No que se refere ao registro de um ou outros, dependerá do uso que se irá fazer deste banco de dados, do orçamento e do pessoal disponível. Estes são assuntos que a pessoa responsável por sua elaboração deverá levar em consideração. São decisões próprias da política de trabalho e da direção e organização da instituição encarregada do trabalho terminológico.

Pode chamar a atenção o fato em que figura o autor ou pessoa que realiza as anotações na ficha terminológica para a coleta de dados. Os dados são neutros em si mesmos e existem ou não existem. Parece que não cabem interpretações subjetivas com respeito a eles. O problema deve ser encarado sob outro ponto de vista, sob o aspecto da lingüística e dos diferentes costumes dos povos. Todos nós sabemos que, na Espanha, utilizam-se palavras que, nos diversos países hispano-americanos, possuem significados distintos. Também sabemos que, num ou noutro país, usam-se palavras diferentes para conceitos idênticos. E o mesmo acontece, por exemplo, com o idioma inglês empregado no Reino Unido e o dos Estados Unidos da América do Norte. Portanto, saber quem fez as anotações pode ser muito útil. Não somente por este motivo. Talvez surjam, também, certas dúvidas e, por tal, será necessário esclarecê-las com a pessoa responsável pelo trabalho.

O objetivo de registrar a data consiste em poder considerar o desenvolvimento da linguagem. Muitas vezes, certos vocábulos mudam com o tempo. Na Espanha, podemos citar, entre muitos outros, como exemplo típico, o de "locutor". No início da utilização do rádio, dizia-se *speaker*, termo que depois caiu em desuso. A data nos indicaria se o termo é atual ou obsoleto.

(*) Estes documentos podem ser livros, dicionários, vocabulários, artigos de revistas etc.

A grande novidade surgida na elaboração dos bancos de dados terminológicos é o seu armazenamento em discos CD-ROM. Na verdade, são muito apropriados, pois, geralmente, o que está gravado neles não precisa ser apagado. Além do mais, o seu manuseio e transporte torna-se cômodo e fácil. As grandes instituições e empresas dispõem de discos CD-ROM. Foram mostrados nas exposições terminológicas realizadas nos últimos anos. Observa-se que seu número vai aumentando de exposição para exposição e que as técnicas melhoram e os preço baixam, causas que estimulam o seu uso.

Como é natural, em todas as questões nas quais intervém uma pessoa para realizar um trabalho, sua formação profissional e suas qualidades pessoais são muito importantes. Retorno, assim, à questão da necessidade de dispor de bons terminólogos, bem formados, questão que menciono mais uma vez para insistir em sua importância.

Para terminar este item, alguns bancos de dados terminológicos são citados como exemplo dos muitos que já existem no mundo. Praticamente, a maioria dos países industrializados e os que se encontram em vias de industrialização dispõem, pelo menos, de um banco de dados terminológicos central, onde se compilam suas terminologias e onde se pode buscar informação sobre estes assuntos. Já citamos a Eurodicautom, da Unidade Européia. No número 1, volume 12, da *Revista Española de Documentación Científica*, foi publicado um artigo de autoria de A. Irazazábal, S. Alvarez e J. Zarco, no qual se explica detalhadamente a composição desse banco de dados.

Outros bancos de dados terminológicos conhecidos e úteis são o de INFOTERM, AFTERM, NORDTERM, TERMIUM e vários outros. Vale a pena acrescentar, talvez, como igualmente relevantes, o GOSSTANDART, da antiga (N.T.) URSS; INFOTERRA, do Canadá, para assuntos de meio ambiente; INSPEC, para física; TEAM-SIEMENS, multidisciplinar; EWF, de Dresden, Alemanha. É aconselhável consultar os catálogos, repertórios e a bibliografia em geral, pois a cada dia aparecem novos bancos de dados terminológicos.

A TRADUÇÃO AUTOMÁTICA

No momento em que foi possível dispor da informática para várias aplicações, pensou-se em utilizá-la no campo da tradução. Em princípio, pensou-se que isso significaria uma grande ajuda, embora não isenta de dificuldades. Os problemas apresentam-se, porque não se trata, somente, de encontrar os equivalentes dos termos de alguns idiomas em outros. Os termos se apresentam formando parte de uma frase, na qual, além disso, deve-se considerar a intencionalidade que sofre. O importante é o significado do texto, o que este quer dizer, e aqui é onde a máquina não alcança. Ou melhor dizendo, não chegou até agora. Talvez os sistemas especialistas desenvolvidos para este fim nos tragam a solução para a tradução e interpretação de textos.

A história da tradução automática é bastante longa, quando comparada com outras aplicações da informática. Já em 1933, realizaram-se alguns projetos na França e na antiga (N.T.) URSS. Tratava-se, na realidade, de dicionários sobre suporte eletrônico, onde se podiam encontrar os termos e suas equivalências. Esses sistemas foram aperfeiçoando-se no decorrer

dos anos. Mas, como digo, o que se precisava era entender e traduzir o sentido da frase. Deve-se considerar que uma mesma palavra pode exercer funções diferentes em uma frase. Tomemos, por exemplo, a palavra "casa". Ela pode ser um substantivo, mas também, uma forma do verbo "casar". Portanto, era precioso determinar essas funções. Além disso, podia-se prescindir de artigos e de algumas preposições. Assim, os primeiros sistemas do que já poderíamos chamar tradução automática consistiam em preparar a frase do idioma de origem, convertendo-a em uma série de palavras e sinais convencionais, que eram os que substituíam as formas gramaticais e sintáticas. Depois, eram introduzidas no computador e este devolvia uma nova combinação de palavras e sinais, que deveriam ser interpretadas para se obter a frase correta no idioma alvo. Este foi o sistema utilizado pelo primeiro programa TITUS, para a indústria têxtil, e SYSTRAM, da Unidade Européia.

Com o aperfeiçoamento continuado das técnicas, foi sendo possível elaborar programas lógicos de computador mais complicados, que repercutiram nas técnicas de tradução automática. Existe, inclusive, um programa para computador pessoal denominado MICROCAT, realizado para PC/XT, da IBM.

Hoje em dia, tem-se colocado muita esperança na inteligência artificial, com os sistemas especialistas para realizar o tratamento, tradução e interpretação de textos. Os próprios sistemas TITUS e SYSTRAM já os utilizam, embora ainda estejam sujeitos a estudo e aperfeiçoamento posterior.

Dispõe-se de programas de tradução automática muito variados. Os últimos estudos indicam que são o Japão e o Canadá os países onde se realiza maior número de pesquisas. A China começou, há pouco tempo, a sua caminhada por estes assuntos, do mesmo modo que a maioria dos restantes países do mundo. Contudo, até agora, torna-se muito dispendioso realizar um programa de pesquisa nestes assuntos; por esse motivo, as líderes são as grandes instituições e os países aderem às mesmas, colaborando em conjunto nas pesquisas que se realizam. Um exemplo ilustrativo do que foi dito é dado pela Comunidade Européia, com seus 12 países membros e suas necessidades de tradução de um idioma para outro. Atualmente, estudam-se os sistemas especialistas para serem aplicados nos seus programas Eurotra, em conjunto com o Eurodicautom. A Espanha participa, por meio do Cindoc e de seu grupo TermEsp, na medida do possível, possuindo orçamento e pessoal necessário a essa tarefa.

A informática é utilizada, também, para confeccionar bases de dados de referências bibliográficas sobre as traduções existentes. Recentemente, está sendo iniciado o armazenamento de textos de traduções completas em discos CD-ROM. O motivo é devido ao esforço para conservar as traduções e saber quais são os fundos disponíveis. Tudo isto, com vistas a uma possível utilização futura.

As traduções são caras, tanto as feitas de modo manual, como as que são realizadas automaticamente. Deve-se, portanto, procurar obter um bom rendimento das mesmas. Com este objetivo, foram criados os bancos de traduções e as bases com os dados de referência para sua localização e uso posterior. Um dos bancos de traduções mais importante está situado

em um centro internacinal, criado para este fim pela OCDE em Delft, nos Países Baixos. Chama-se Centro Internacional de Traduções. Ali são compiladas também as traduções realizadas em outros países. A Unidade Européia, junto com o Centro Nacional de Pesquisas Científicas da França, criou, em 1978, uma base de dados sobre traduções. Utiliza o sistema PASCAL e publica um periódico chamado *World Transindex*. Sobre esse assunto, deve-se assinalar a conveniência de se consultar o periódico *Journals in Translations*, que fornece informações sobre os periódicos que se traduzem, quer seja pelo sistema página a página, que seja diretamente artigo por artigo. Entre estes últimos, merece ser citada a publicação periódica *International Forum for Information Science*, editada pela Federação Internacional para a Informação e a Documentação (FID), que o faz em russo, inglês e espanhol (figura 3.3).

Existem, também, centros de traduções especializados, de âmbito nacional e internacional. Entre os últimos, ocupam um lugar destacado o Instituto Francês do Petróleo, o BRGM, para mineração e ciências da terra; o Imperial Chemical Industries (ICI), para a indústria química, para citar alguns entre muitos outros.

Na Espanha, cabe mencionar o Cindoc, ligado ao Centro de Delft e com a base de dados da UE. Lá mesmo são realizadas traduções científicas e técnicas que conservam em seus arquivos e que podem ser recuperadas por meio de suas bases de dados. Publicam um periódico trimestral com as referências dessas traduções. A vantagem de se consultar tais publicações repercute em benefício econômico. Caso a tradução de que se necessita já tenha sido realizada, deve-se pagar somente o preço de uma fotocópia.

A tradução automática é assunto de grande atualidade pela importância que adquire, a qual vai aumentando dia a dia. Enquanto se pretende dispor de um idioma internacional com o qual possamos nos fazer compreender, estão surgindo, por outro lado, idiomas e mais idiomas de novas nações ou novas regiões autônomas, que reclamam um lugar no concerto internacional, publicando seus documentos em seus idiomas autóctones ou restringem-se ao publicado fora de seu âmbito. Ao mesmo tempo, aparecem disciplinas, no campo terminológico, que antes não se interessavam por estas questões. Um exemplo recente chegou-me às mãos, com um programa de um seminário sobre terminologia de interesse na história militar.

Tamanha é a importância da terminologia e das indústrias da língua, que já se editam publicações periódicas dedicadas a estes assuntos, como o periódico *Language Terminology*. Foi criado o Prêmio Internacional da Inovação em Matéria de Indústria da Linguagem.

Para um aprofundamento maior nestes assuntos, é aconselhável consultar a bibliografia existente, para obter-se informação ampla e detalhada.

BIBLIOTECA
DO
IBICT

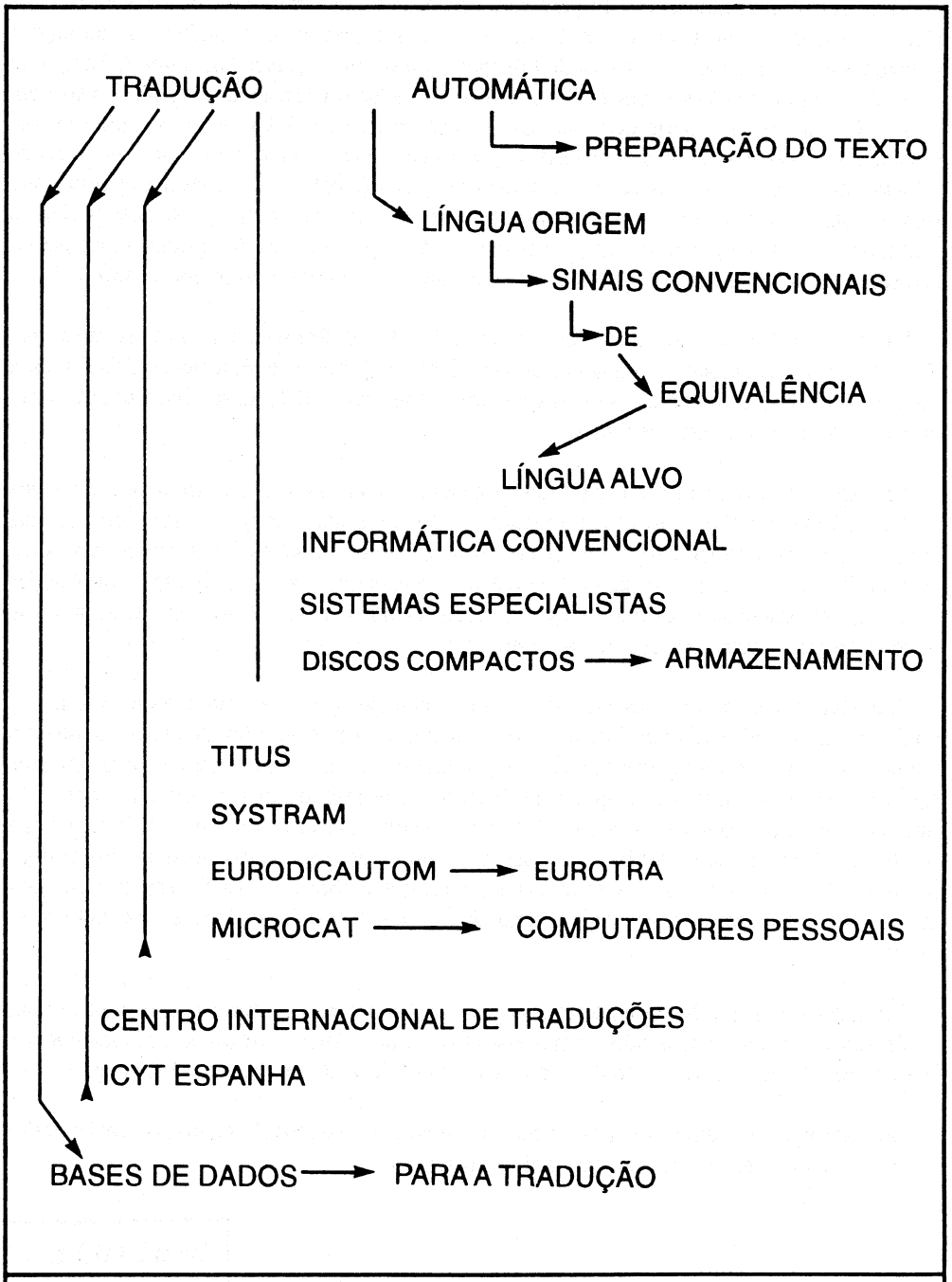


Figura 3.3. - A tradução automática

A NORMALIZAÇÃO NAS INDÚSTRIAS DA LÍNGUA

Antes de terminar este capítulo, parece oportuno falar da normalização em relação às indústrias da língua e da terminologia em geral. Dois aspectos da questão devem ser considerados:

A normalização
↓
como →

— origem
— causa ↔ efeito

da terminologia

Um dos motivos principais pelo qual se iniciaram os estudos terminológicos, que adquiriram tanta importância nestes últimos tempos, foi a necessidade de normalizar as linguagens especializadas, próprias de cada uma das diferentes disciplinas e subdisciplinas dos distintos campos do conhecimento. Todos sabem que uma mesma ferramenta, por exemplo, recebe numerosos nomes nas diferentes regiões dentro de um mesmo país.

Com a explosão científica, técnica e industrial, apareceram novos objetos que reclamam uma denominação apropriada e a destinação de um termo correto. Todos conhecem, também, a anarquia que existe nestas questões e a introdução de termos provenientes do idioma anglo-saxão, por influência dos Estados Unidos da América do Norte, principalmente. Neste país, foi observada a maior evolução do progresso em nosso século. Do mesmo modo, de lá veio a maior influência até o mundo ocidental. Não negamos o desenvolvimento das repúblicas do Leste Europeu, nem do Japão e China. É que, simplesmente, tiveram menos repercussão nos países europeus ocidentais, e de lá a influência lingüística foi menor.

Por todas essas considerações, torna-se necessário normalizar as línguas e sua indústria. Deve-se firmar denominações, definir conceitos, unificar os termos. É aqui onde entra a normalização.

Pensemos, também, que o próprio trabalho terminológico necessita ser normalizado, para que haja uma unidade de critério no momento, por exemplo, de elaborar um vocabulário, um programa de computador, ou algumas definições de conceitos teóricos terminológicos. Igualmente as linguagens terminológicas, em toda a sua extensão, precisam de regras, geralmente teóricas. Assim, pois, a normalização, como causa e efeito da terminologia e da indústria da linguagem, volta a ser imprescindível nestas tarefas e atividades (figura 3.4).

INSTITUIÇÕES DE NORMALIZAÇÃO

As instituições de normalização, tanto nacionais como internacionais, tornaram-se conscientes de que deviam dedicar parte de seu trabalho à terminologia. A International Federation of National Standarizing Association (ISA) já havia organizado um Comitê Técnico

TC 37. Desde então, vem elaborando as normas dedicadas à terminologia, à indústria da língua e às linguagens terminológicas em geral.

As normas elaboradas pelo ISO/TC 37 são de tipos diferentes: umas dedicam-se aos métodos de trabalho e outras aos utensílios ou aparelhos. Assim, por exemplo, regulam as características das fitas magnéticas ou dos formatos para elaborar os bancos de dados terminológicos.

Para cumprir melhor sua missão, a ISO/TC 37 dividiu-se em subcomitês que estudam, cada um, um aspecto diferente da questão. No que se refere à documentação (ciências da documentação), a ISO conta com um comitê técnico dedicado exclusivamente a normalizar os assuntos documentários. É o ISO/TC 46. Dentro dele, existe o subcomitê ISO TC 46/SC3, que trabalha na terminologia da documentação. O Comitê ISO/TC 4 – Computer Applications in Information and Documentation –, como seu nome indica, é o que se dedica a normalizar as aplicações dos computadores na informação e documentação.

Como exemplos, citemos algumas normas ISO mais usadas:

- ISO/R 1087 - Vocabulary of Terminology;
- ISO/DIS 704 - Principles and methods of terminology, que, depois, passou a ser ISO/704 - Naming Principles;
- ISO/860 - International unification of concepts and terms;
- ISO/6156:1987 - Magnetic tape exchange format for terminological/lixicographical redord (MATER);
- ISO/2788 - Documentation, Guidelines for the establishment and development of monolingual thesauri;
- ISO/5964 - Documentation, Guidelines for the establishment and development of multilingual thesauri;
- ISO/5127 - Information and documentation. Vocabulary Section 3a: Adquisition, identification and analysis of documents and data.

Para se obterem melhores informações sobre as normas existentes, é aconselhável consultar os catálogos de normas ISO. Os mesmos encontram-se à disposição do público na AENOR.

Outra instituição internacional dedicada à elaboração de normas sobre terminologia é a Infoterm, na Áustria. Não podemos deixar de citar a Federação Internacional de Informação e Documentação (FID), que, em seu comitê FID/DT, estuda os problemas de terminologia e possui um extenso programa de trabalho, detalhado no capítulo 2.

A maioria das instituições nacionais de normalização possui comitês dedicados à terminologia. Nos catálogos e publicações nacionais de normalização, encontra-se extensa informação sobre as normas publicadas e em elaboração, assim como outras notícias referentes a este assunto. Aqui se faz referência somente à Espanha, por ser o país que nos concerne mais diretamente. Em nosso país, é a Asociación Española para la Normalización (Aenor) sucessora de Iranor, a instituição que se ocupa da normalização do trabalho. Nela existe o

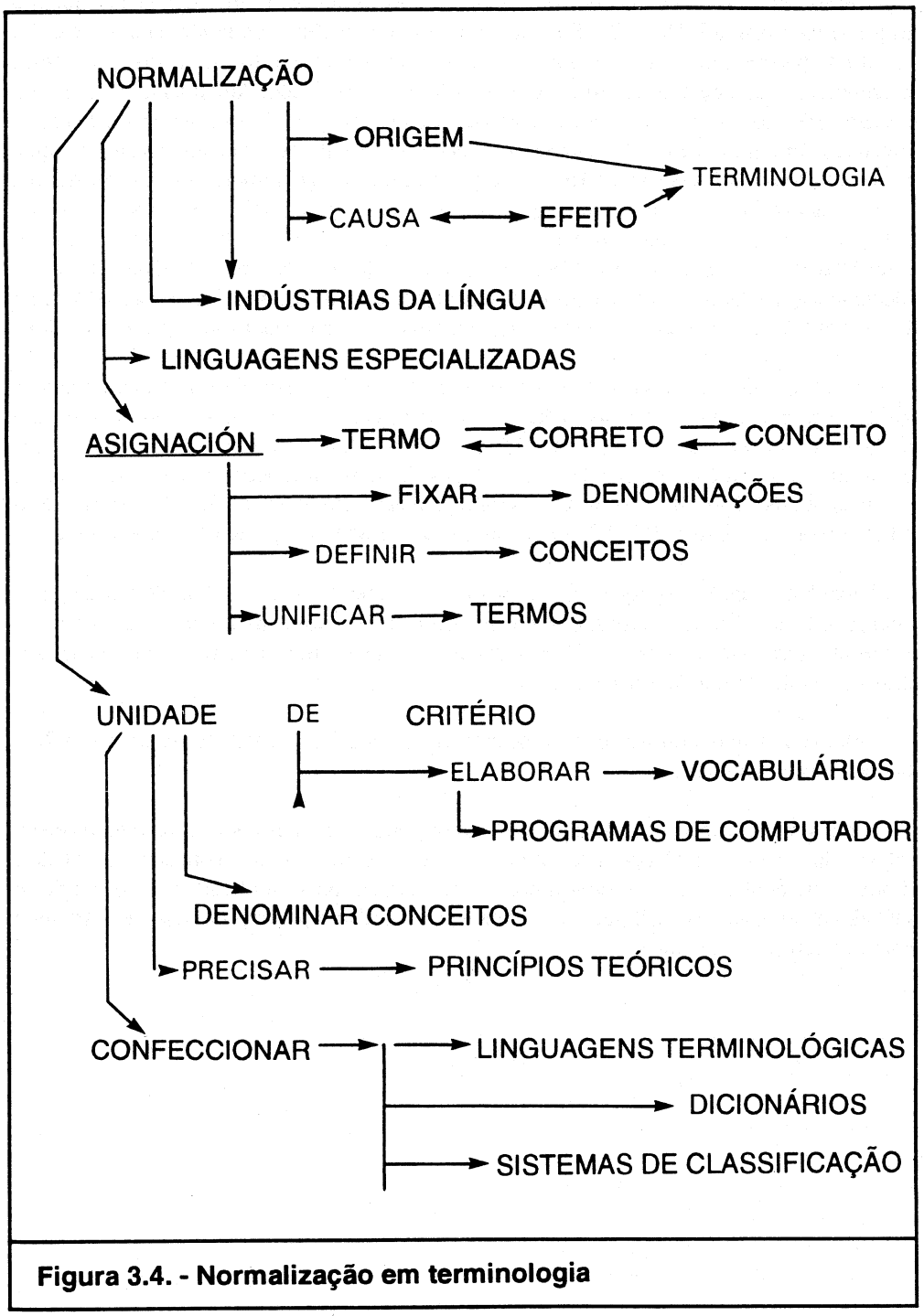


Figura 3.4. - Normalização em terminologia

Comitê Técnico TC 1, dedicado, precisamente, aos assuntos terminológicos. Colabora estreitamente com a ISO/TC 37. Eu tenho a honra de participar, em nome da Espanha, das reuniões internacionais de trabalho, levando-se uma vida muito ativa. Por problemas econômicos e orçamentários, minha colaboração foi-se diluindo. Atualmente, houve uma revitalização no interesse dos documentalistas pelos trabalhos do Comitê TC-1, tendo-se nomeado Amélia de Irazazábal sua representante. O Comitê 50 - Documentação, no qual venho colaborando desde 1974, é o que se ocupa diretamente da terminolgia em documentação. Recentemente, iniciou-se a elaboração e publicação das normas UNE. Seu trabalho começa pela tradução e adaptação das normas ISO/. Uma delas é a norma ISO/5127 (3a), que passou a ser UNE 50-113/3a, ou a norma ISO/2788, convertido em UNE 50-106. Este comitê 50-Documentação é também o que se ocupa da elaboração e atualização da Classificação Decimal Universal (CDU), pois esta é considerada como um sistema normalizado de classificação.

O Comitê UNE/TC 1 está encarregado de elaborar, ainda, as normas referentes a assuntos gerais. Neste conjunto, estão incluídas as que se referem aos trabalhos terminológicos e bibliotecários, como, por exemplo, a UNE 1 - 085-79 - Repertórios de Bibliotecas, Centros de Informação e Documentação; UNE 1 - 84-84 - Códigos para a Representação dos Nomes dos Países; UNE 1 - 088-76 - Folhas de Título para um livro; UNE 1 - 070-79 - Vocabulário da Terminologia; UNE 1 - 072-82 - Apresentação de Vocabulários Sistemáticos Multilíngües.

Parece um contra-senso que seja o Comitê CT 1 que prepare este tipo de normas. Por conseguinte, nos últimos tempos, realizou-se uma reestruturação das atividades dos comitês, tendo sido designada ao Comitê CT-50 - Documentação a confecção das normas que tratam dessa disciplina, em toda a sua extensão.

Para obter maiores informações, é aconselhável consultar o catálogo de normas UNE, também disponível na Aenor.

Como vemos, a normalização é um assunto importante em terminolgia, na qual recebe a relevância que merece. Coopera-se em níveis nacional e internacional com estas atividades, motivo pelo qual se tem a esperança nos resultados para poder unificar critérios, teorias e métodos de trabalho. Isto, naturalmente, com as peculiaridades próprias de cada país e levando em consideração seu caráter.

TEMAS PARA ESTUDO

1. Diga que expressão parece mais apropriada: "indústria da língua" ou "engenharia da linguagem".

2. Defina um banco de dados terminológicos.

3. Considera útil ocupar-se da normalização da linguagem?

4. Diga em que caso construiria um banco de dados terminológicos.

5. Pensa que, se dispuser de um banco de dados terminológicos, poderá necessitar, também, de uma base de dados terminológicos?

6. Diga o que entende por informática lingüística.

7. Cite alguma instituição internacional de normalização que se ocupe de terminologia.

8. Dê sua opinião sobre a tradução automática.

9. Caso conheça, descreva algum sistema de tradução automática.

10. Descreva como elaboraria um banco de dados terminológicos.

11. Diga quem se ocupa das atividades terminológicas na Espanha.

12. Cite alguma norma UNE relacionada com a terminologia.

13. Construiria um banco terminológico manualmente?

14. Cite algumas diretrizes na pesquisa sobre a indústria da linguagem.

15. Quando utilizaria um disco CD-ROM em terminologia automatizada?

16. Dê sua opinião sobre a utilização dos sistemas especialistas na tradução.

17. Diga quando justificaria automatizar uma instituição dedicada à tradução técnica.

18. Cite algum banco de dados terminológicos.

REFERÊNCIAS BIBLIOGRÁFICAS

BEAUMONT, J.F.: "La moderna industria de las lenguas"; Madrid, El País, 18 de marzo de 1987.

"El CD-ROM y el lenguaje"; Terminómetro, *IV*, 2, 1984, 12.

Coloquio "Terminología e industrias del lenguaje"; Terminómetro, *IV*, 2, 1984, 6.

FILLON, L.: "TERMIUN"; en Actas de la Exposición de Lingüística y de Terminología Científico-técnica, CSIC y Unión Latina, Madrid, 1987, 188-192.

GARCIA CAMARERO, E.: "La informática y la creatividad"; en Cursos sobre la informática y el vídeo en la ensenanza. Instituto Juna de Valdés, Cuenca, 1985, 101-114.

GONZALEZ ABASCAL, J., ARES DE BLAS, F., GURRUCHAGA VAZQUEZ, J.: "Un método de comprensión de textos escritos en lenguaje natural. Aplicación al euskera"; en II Congreso Mundial Vasco, *II*, 22, 1987.

GÜNTZER, W., JÜTTNER, G., SEEGMÜLLER, G.: "TEGEN- ein lernfahiges Information Retrieval System"; in International Congress on Terminology and Knowledge Engineering; INDEKS Verlag, Frankfurt, 1987, 323, 337.

"Industrias de la lengua: un desafío para las lenguas latinas"; Terminómetro, *0*, diciembre, 1987.

"Informatique et Langage Natural: Presentation Générale des Méthodes d'Interpretation des Textes Ecrits"; Technique et Science Informatiques, 2, 9, 1986, 103-126.

KUCK, C.: "Knowledge engineering by quantum logic"; in International Congress on Terminology and Knowledge Engineering, Trier, Federal Republic of Germany, University of Trier, 1987, 57-68.

LELOUP, C.: "CD-ROM: Nouvelles Technologies"; en Actas de la Exposición de Lingüística Informática y de Terminología Científico-técnica, CSIC y Unión Latina, Madrid, 1987, 87-91.

LEMAITRE, C.: ''Un sistema semántico para interfaces en lenguaje natural''; en Comunicaciones y Posters Iberamia 88, 1988, 285-97.

"Lengua, tecnología e industria"; Bol. FUNDESCO, *81*, mayo 1988, 2.

"Lexicography in the electronic age"; Proceed. of Simp. Luxembourg, by J. Goetschalkx y Z. Rolling, july, 1981.

MALACA CASTELEIRO, J.: ''El portugués y las industrias del lenguaje"; Terminómetro, *VI*, 1, 1988, 20-21 .

"Manifiesto de Madrid", en Actas de la Exposición de Lingüística Informática y de Terminología Científico-técnica, CSIC y Unión Latina, Madrid, 1987, 197.

MANU, A.: "La diffusion de l'Information Terminologique au sein du Réseau TermNet"; en Actas de la Exposición de Lingüística Informática y de Terminología Científico-técnica, Madrid, 1987, 99-103.

MARCOS MARIN, F.: "Lingüística informática"; Bol. FUNDESCO, 73, 1987, 4.

MAYER, R. and MAIER, E.: "Spezification eines konzeptionellen Schemas für Terminologiedatenbanken"; in International Congress on Terminology and Knowledge Engineering, INDEKS Verlag, Frankfurt, 1987, 151-16.

NAKAI, H. et alii.: "Representation rules for scientific words in a Knowledge base system"; in International Congress on Terminology and Knowledge Engineering; INDEKS Verlag, Frankfurt, 1987, 205-216.

NISSAN, E.: "Onomaturge: An expert Systems for Wordformation and morphosemantic Clarity evaluation"; in International Congress on Terminology and Knowledge Engineering, INDEKS Verlag, Frankfurt, 1987, 167-189.

PAUL, J.: '"Développements récents des supports logiques linguistiques"; en Actas de la Exposición de Lingüística y de Terminología Científico-técnica, CSIC y Unión Latina, Madrid, 1987, 158-161.

RODRIGUEZ JIMENEZ, M.: "Tecnología: las industrias de la lengua, una oportunidad para el espanol"; en Bol. de Comunicación, Tecnología y Sociedad, *II*, 28, 12 marzo, 1988, 35.

SADLER, V.: "AI-directed Interlingual Terminography in tomorrow s MT systems"; in International Congress on Terminology and Knowledge Engineering, INDEKS Verlag, Frankfut, 1987, 369-376.

SCHWARTZ, C.: "Un proyecto experimental (COPSY) en Alemania"- en Actas de la Exposición de Lingüística Informática y de la Terminología Científico-técnica, CSIC y UNION Latina, Madrid, 1987, 104-107.

"El segundo programa de Unión Latina"; Terminómetro, *VI*, 1, 1988, 33-34.

TERMINOLOGIE, Bull. n.38, Com. CEE, Doc. Nr. 4283/81.

VELASCO, G.: "Language Industry Hearing calls for Community Action", in Information Market (I'M), Commission of the European Communities Directorate General XIII: Telecommunications, Information Industries and Innovation, *56*, feb. 1989-april 1989.

VERDEJO, M.F.: "Interfaces en lenguaje natural"; Primeras jornadas de Inteligencia Artificial para Bases de Datos, Barcelona, 17-18 octubre, 1985.

VIDAL-BENEYTO, J.: "La industria de las lenguas"; en EL PAIS, 24 de febrero, 1987, 11-2.

WEINSERG, H.B.: "Observations on the Field of Computational Linguistics... by and Information Scientist"; in Bull. ASIS, December/January 1988, 33-34.

WRIGHT, L.D. Jr.: "World Writer and Interpreted «Term Software»"; TermNet, 15, 1986, 13.

ZAMPOLLI, A.: "Linguistique computationelle: application des ordinateurs aux domaines de la lexicographie et de la terminologie", en Actas de la Exposición de Lingüística Informática y de Terminología Científico-técnica, CSIC y Unión Latina, Madrid, 1987, 81-84.

Bancos de Dados Terminológicos

BRADLEY, D.: "Procedimientos que se emplean en la elaboración de una base de datos de neologismos espanoles"; en Actas de la Exposición Lingüística y de Terminología Científico-técnica, CSIC y Unión Latina, Madrid, 1987, 175-179.

"Las Cámaras de Comercio Europeas desarrollarán una Base de Datos conjunta de términos aduaneros"; en Rev. Actualidad Electrónica, *464*, 12 jun., 1987, 17.

CASAJUANA CASANOVAS, R. y RODRIGUEZ MAGRO, C.: "Clasificación de los verbos castellanos para un diccionario en ordenador"; Actas del I Congreso de Lenguajes Naturales y Lenguajes Formales, Universidad de Barcelona, Barcelona, 1984.

"Catálogo de recursos terminológicos en lengua espanola"; CSIC, Espana, 1987.

"Creación de la Red Iberoamericana de Terminología"; Terminómetro, *VI*, 1, 1988, 3.

DE MEER, J.: "Introduction to the formal description technique LOTOS"; in InternationalCongress on Terminology and Knowledge Engineering, INDEKS Verlag, Frankfurt, 1987, 415-430.

EURODICAUTOM: "La banque de données et terminologiques de la Commision des Communautés Europèennes. Reichling A."; TermNet News, 4/5, 1982.

"EURODICAUTOM" – Manual, Luxembourg, 1984 – European Commission Host Organization. 177, route d'Ésch, Luxembourg.

FIGUEIREDO CASTRO, R.C., LAERTE PACKER, A., DE CASTRO, E.: "Proyecto LILACS/CD-ROM. Literatura Latinoamericana y del Caribe en Ciencias de la Salud en disco compacto"; en Rev. Esp. Doc. Cient., *12*, 1, 1989, 23-31.

FUIIKAWA, M. and ISHIKAWA, T.: Topics for "Terminology and Machine Dictionaries in Japan"; in International Congress on Terminology and Knowledge Engineering; INDEKS Verlag, Frankfurt, 1987, 85-122.

GRUPO DE TRABAJO SOBRE LENGUAJES DE INDIZACION CAICYT: "Bancos terminológicos"; Rev. Esp. Doc. Cient., *9*, 1, 1986, 57-62.

IRAZAZABAL, A., ALVAREZ, S., ZARCO, J.: "EURODICAUTOM. El Banco de Datos Terminológico de la CEE"; en Rev. Esp. Doc. Cient., *12*, 1, 1989, 11-23.

MANU, A.: "La banque de données terminologiques des normes autoschiennes"; en Actas de la Exposición de Lingüística y de Terminología Científico-técnica, CSIC y Unión Latina, Madrid, 1987, 150-154.

MATARO, F.: "Le système SYSTRAN"; en Actas de la Exposición de Lingüística y de Terminología Científico-técnica, CSIC y Unión Latina, Madrid, 1987, 172-175.

RODRIGUEZ JIMENEZ, M.: "Comunicación, nuevas tecnologías e idioma (La base de datos FUNDESCO como observatorio del lenguaje) en Bol. FUNDESCO, 57, 1986, 7-10.

SKOROKHOD'KO, E.F. y STOGNNII, A.A.: Creación de un banco terminológico: el aspecto lingüístico"; Rev. Esp. Doc. Cient., *10*, 3, 1987, 345-347.

SOLANELLAS TORRENTS, J.M. y CABRE CASTELLVI, T.: BTERM, El banc terminologic del catalá, una solució per el seo desenvolupament (BTERM, El banco terminológico catalán, una solución para su desarrollo)"; en CIL 87, 1987, 118-35.

SOPEÑA, L.: "Diccionarios del castellano en ordenador para la composición y verificación de textos"; Madrid, Centro de Investigación U.A.M.-I.B.M., 1987.

Yu, Y: "Aufbau einer on line – mehrsprachingen Terminologie – Datenbank mit chinesichen Schriftzeichen" ; in International Congress on Terminology and Knowledge Engineering; INDEKS Verlang Frankfurt, 1987, 393-402.

Normalização

BORDONABA, V.: "Qué es la normalización?"; en Quím. e Indus., *35*, 4 Abril 1989, 406 407.

DE VRIES ROBBE, P.F.: "Knowledge representation in relation to ISO principles and methods of terminology"; in International Congress of Terminology and Knowledge Engineering; INDEKS Verlag, Frankfurt, 1987, 269-271.

DIN 2330: "Begriffe und Benenungen. Allgemeine Grundsatze. Normennausschuss Terminologie (NAT)", im Deutchen Institut für Normung e.V., Berling, 1979.

EMA, E.: "Normalización terminológica en el campo de la tecnificación"; en Actas de la Exposición de Lingüística Informática y de Terminología Científico-técnica, CSIC y Unión Latina, Madrid, 1987, 97-99.

GALINSKI, Chr. y NEBODITY, W.: "Especial Languages, terminology planning and standardization"; in Standardization of technical Terminology, by R.A. Strehlow, ASTM, 2, Philadelphia, 1988, 6-13.

Iso: ''Structure of ISO/TC46", Separata de ISO, 1987.

MARTINHO, A.M.: "Normalização de terminologie informática portuguesa", en Actas de la Exposición de Lingüística Informática y de Terminología Científico-técnica, CSIC y Unión Latina, Madrid, 1987, 76-77.

Norma UNE 1-070-79. Vocabulario de la terminología. Julio 1979.

Norma 50-113/3a. Información y Documentación – Vocabulário – Capítulo 3a): Aquisición, Identificación y Análisis de Documentos y Datos.

PICHT, H.: "La normalización de la terminología"; en Actas del Primer Seminario Nacional de Terminología; Caracas-Venezuela, Universidad Simón Bolívar, junio 1984, 213-224.

STREHLOW, R.A.: "Standarization of Technical Terminology: Principles and Practices (Second Volume)"; ASTM, Philadelphia, 1988, 29-66.

Tradução Automática

"Algunas empresas de traducción y documentación"; Terminómetro, *VI*, 1, 1988, 28-29.

ALVAREZ, S.: "La traducción en el ICYT (CSIC)", en Actas de la Exposición de Lingüística y de Terminología Científico-técnica, CSIC y Unión Latina, Madrid, 1987, 11-13.

ANXOLABEHERE, V.: "Sistemas de traducción asistida por ordenador", en Actas de la Exposición de Lingüística Informática y de Terminología Científico-técnica, CSIC y Unión Latina, Madrid, 1987, 37-38.

CHAPPUY, S.: "Travaux et systèmes de traduction automatique", en Actas de la Exposición de Lingüística Informática y de Terminología Científico-técnica, CSIC y Unión Latina, Madrid, 1987, 154-157.

GONELLA, D.: "II Progetto DIMA di traduzione automatica"; en Actas de la Exposición de Lingüística y de Terminología Científico-técnica, CSIC y Unión Latina, Madrid, 1987,145-148.

HUTCHINS, W.J.: "Progress in Documentation Machine Translation and Machine-aided, Translation"; J. of Documentation, *34*, 2, 1978.

"Index translationum, Répertoire international des traductions"; Vol. 36, 1983, UNESCO.

LAMB, J.: "Sistemas traductores"; en Rev. Datamation, *14*, 1986, 60-3.

MANU, A.: "Bibliographie pour le traducteur", en Actas de la Exposición de Lingüística Informática y de Terminología Científico-técnica, CSIC y Unión Latina, Madrid, 1987, 9-10.

MARCHUK, Yu N.: "Machine Translation in the USSR"; Int. Forum Inf. and Doc., *9*, 2, 1984, 3-8.

REICHLING, A.: "Originalités de la tecnique au service de la traduction multilingue"; en Actas de la Exposición de la Lingüística Informática y de Terminología Científico técnica, CSIC y Unión Latina, 1987, 115-119.

SOLER, J.: "El programa EUROTRA de traducción automática"; en Actas de la Exposición de Lingüística Informática y de Terminología Científico-técnica, CSIC y Unión Latina, Madrid, 1987, 184-188.

"Traducción"; Terminómetro, *VI*, 1, 1988, 27.

4

As Linguagens Terminólogicas: Tesauros

Há aproximadamente 10 anos, utiliza-se a expressão "linguagens terminológicas" aos sistemas de classificação de assuntos. Em muitos ambientes científicos ainda parecerá estranha, e pensarão que se trata de modernismos empregados com a intenção de afastar-se das normas preestabelecidas. Contudo, esta denominação é correta, pois uma linguagem terminológica presume que se trata de um sistema lingüístico no qual os componentes principais são os termos. E sabemos que estes podem ser definidos como "elementos primeiros do conhecimento para armazenar saberes" (I. Dahlberg), ou "unidades lingüísticas de um vocabulário especializado" (H. Felbert).

Não obstante, pode ser interessante refletir um pouco sobre como se chegou a este ponto na evolução da terminologia e dos sistemas de classificação. Na minha opinião, estamos, novamente, diante da conseqüência do processo evolutivo de nosso tempo e da marcha acelerada dos acontecimentos, quer científicos, técnicos ou humanísticos. Deve-se fazer referência, mais uma vez, à já famosa "explosão da informação".

Em tempos não muito remotos – situemo-nos no período entre as duas guerras, neste século XX –, contudo o mais importante eram os documentos que continham a informação. Por isso, sua classificação, isto é, sua colocação em classes, de acordo com o seu conteúdo, foi uma das principais tarefas dos profissionais que trabalhavam com tais documentos. Já naquele tempo, começou-se a pensar que o documento não presumia mais que um veículo, um suporte para o conteúdo temático que encerrava e que se convertia em informação útil e necessária por causa de seu próprio significado e demanda para seu uso posterior. Do mesmo modo, difundia-se a idéia de que não se queriam livros, somente seu conteúdo. Este devia dissociar-se de seu aspecto e considerar-se como um material diferente. Era necessário, pois, dividir a bibliografia em partes e ordená-las sistematicamente.

Esta mudança de mentalidade e forma de pensar acelerou-se, principalmente, após o término da II Guerra Mundial, quando se voltou a experimentar um grande progresso tecnológico. Verdadeiramente, as classificações por assunto existentes tornaram-se demasiadamente rígidas e excessivamente estáticas para enquadrar nelas as novas invenções e descobertas, os novos conceitos, os novos termos que apareciam diariamente. Assim surgiu a idéia de tirar do documento o seu conteúdo e representá-lo por palavras (palavras-chave), termos, que seriam submetidos a uma ordenação posterior.

Todo o conjunto de palavras – termos – relativos a um assunto, que formavam parte de um

vocabulário especializado, compunha uma linguagem, por sua vez, especializada. Rapidamente, observou-se que essas linguagens deveriam estruturar-se de alguma maneira, relacionando os termos de uma forma lógica, para a sua melhor localização dentro da própria linguagem e seu uso posterior. Passava-se de uma linguagem natural, a dos documentos, para uma linguagem estruturada, a da informação neles contida (figura 4.1).

CLASSES DE LINGUAGENS TERMINOLÓGICAS

Assim apareceram expressões como "linguagem controlada", "linguagem normalizada", "linguagem documentária", que se aplicavam principalmente a estes sistemas de classificação. Todavia, os sistemas de classificação convencionais, tais como a Classificação Decimal Universal, reclamavam um lugar dentro desse mundo documentário-terminológico. Finalmente, são métodos utilizados também para classificar os assuntos contidos nos documentos.

A extensa literatura existente sobre estes assuntos não fornece idéias eslcarecedoras sobre tais questões. De modo geral, dá a impressão de que falam para especialistas com alguns conhecimentos, e a única coisa que se pretende é salientar um ou outro ponto, para limitar determinados extremos.

Todavia, nem sempre são especialistas os que se interessam pelos sistemas de classificação ou as linguagens documentárias. Penso nos futuros profissionais, alunos de escolas ou universidades e outros profissionais, que se aproximem das ciências da documentação para seu uso particular. Atualmente, os computadores individuais tornaram-se uma ferramenta eficiente para ordenar livros ou documentos, inclusive no próprio lar. Por isso é que se deve procurar sistematizar o assunto e tratá-lo a partir dos princípios mais fundamentais.

Começaremos por estabelecer algumas definições que servirão para argumentações posteriores:

Linguagem controlada

A linguagem controlada pode ser definida como:

- uma linguagem na qual os elementos (palavras) que a compõem são controlados de acordo com regras preestabelecidas.

Pode-se definir, também como:

- um sistema lingüístico no qual as unidades que o compõem estruturam-se de acordo com normas prefixadas.

No momento em que se utilizam algumas "normas", sempre as mesmas, prefixadas, seguindo um critério lógico e coerente, dever-se-á falar de linguagem normalizada. As unidades que a compõem imediatamente adquirem a categoria de termos. Por conseguinte:

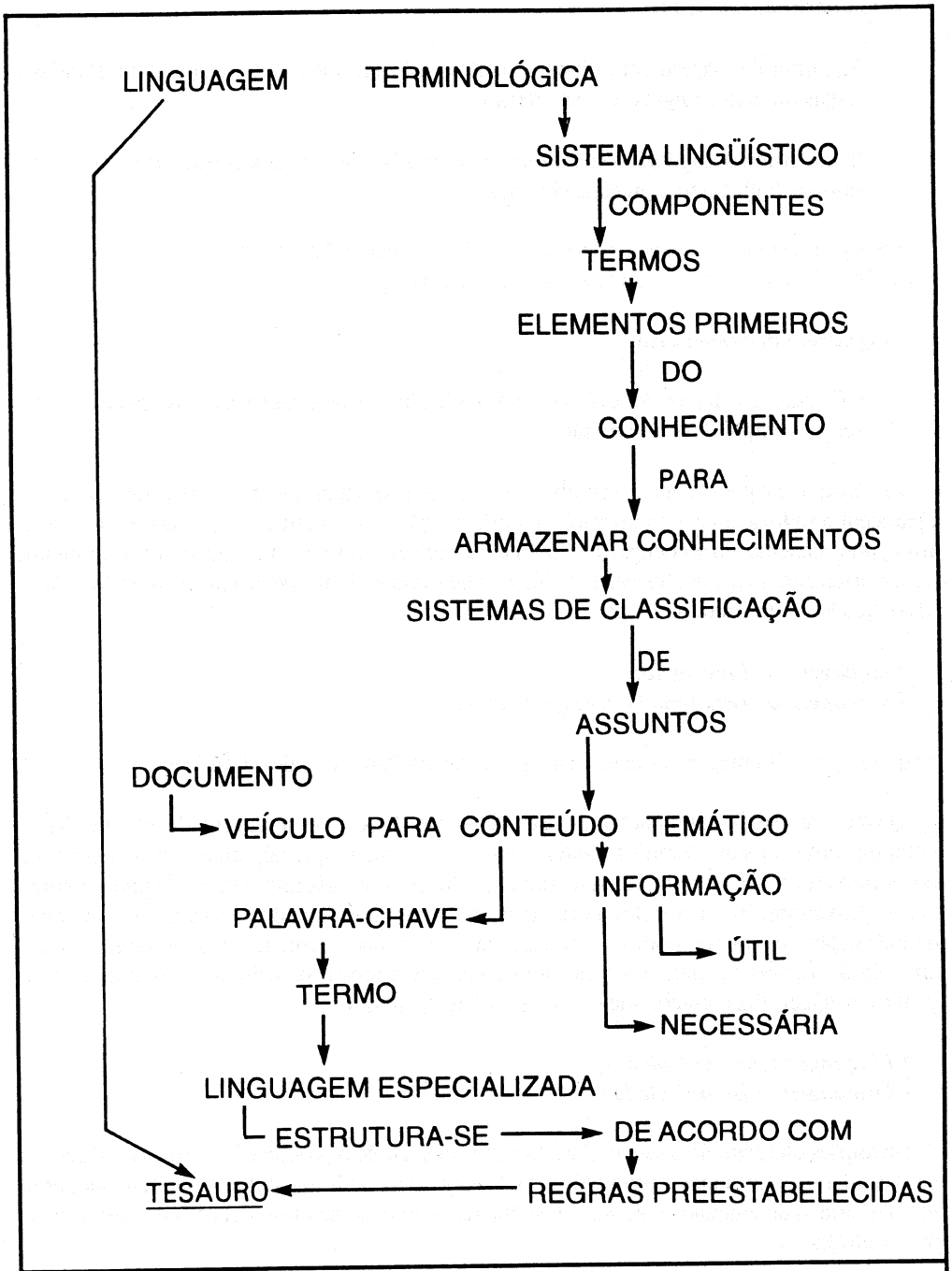


Figura 4.1. - Linguagem terminológica → Tesauro

Linguagem normalizada

1. é uma linguagem controlada, a qual se aplicam normas logicamente prefixadas e cujas unidades lingüísticas são termos.

2. um sistema lingüístico constituído por termos relacionados entre si, de acordo com normas logicamente preestabelecidas.

Eis aqui que estes sistemas possuem sua razão de ser atendendo a seus fins classificatórios no âmbito documentário; de onde se deduz que uma

Linguagem documentária

• É uma linguagem controlada – normalizada – usada com fins classificatórios, no sentido amplo desta expressão.

Ao darmos um passo a mais, introduzimo-nos nas operações documentárias, entre as quais aparecem a indexação e a recuperação da informação. Finalmente, todas estas idéias foram fruto, precisamente, da preocupação e do interesse em encontrar a informação e os documentos necessários para desenvolver uma atividade determinada. Estes deveriam estar indexados e classificados. Surgiram, pois, as

Linguagens de indexação;
Linguagens de recuperação da informação;

as quais, como linguagens documentárias, não necessitam de maior definição.

Quando se estuda detidamente os sistemas de classificação – linguagens documentárias –, tanto os convencionais como os modelos novos, observa-se que, algumas vezes, o controle dos termos foi feito *a priori*, considerando-se a matéria, ou assunto, como conjunto. Outras vezes, pelo contrário, o controle – normalização – foi feito estudando-se os termos – unidades de indexação ou classificatórias – em si mesmas, independentemente umas das outras, mas, em relação recíproca umas com as outras, considerando-se os princípios semânticos ou genéricos. Quer dizer, então, que devem ser consideradas as

• ***Linguagens pré-controladas;***
• ***Linguagens pós-controladas.***

Exemplos das primeiras são as classificações hierárquicas, como a Classificação Decimal Universal, ou as facetadas, como a Colon Classification, de Ranganathan. Por seu lado, uma lista de termos de indexação ou um cabeçalho de assunto serão classificações documentárias pós-controladas.

Seguindo esta ordem de raciocínio, deve-se considerar, agora, de que forma ou maneira

serão controlados os termos de uma linguagem documentária. Novamente se irá pensar em um processo prévio – *a priori* – e em um processo posterior – *a posteriori*. O controle pressupõe o estabalecimento de relações hierárquicas, associativas e de equivalência. Contudo, deverá ser levado em consideração que os termos podem ser simples ou compostos, e que estes últimos poderão se formar coordenando seus elementos, precisamente a priori ou a posteriori. Isto explica que se fale de:

Linguagens pré-coordenadas

- Linguagens documentárias nas quais os termos que as compõem se coordenam em um processo prévio à sua utilização.

Um exemplo típico é dado por um sistema de cabeçalhos de assuntos.

Linguagens pós-coordenadas

- Linguagens documentárias nas quais os termos que as compõem se coordenam em processo posterior à sua determinação, por exemplo, no momento de seu estabelecimento ou de seu uso.

Um exemplo clássico é constituído pelos tesauros (*).

As linguagens de indexação e de recuperação da informação poderão ser pré ou pós-coordenadas, de acordo com as necessidades do sistema de informação ou do centro de documentação onde são usadas.

De tudo o que foi dito, deduz-se que uma linguagem pré-controlada poderá ser pós-coordenada, pois este processo se realiza na fases de indexação – classificação – e de recuperação da informação.

Tomemos o seguinte exemplo: deseja-se classificar um documento sobre o uso dos computadores no ensino, pelo sistema de Classificação Decimal Universal (CDU), sistema pré-controlado. Pesquisaram-se os assuntos relativos à informática e ensino, e combinaram-se convenientemente os termos encontrados para dar ao documento a classificação apropriada (processo de pós-coordenação).

No caso concreto da CDU, a pré-coordenação pressupõe um pré-controle, ao estabelecer as hierarquias de sua estrutura preestabelecida.

Por outro lado, uma linguagem documentária pós-controlada poderá ser pré-coordenada ou pós-coordenada, ou ambos os casos ao mesmo tempo, conforme se considere sua fase de construção ou sua fase de utilização tanto no processo de indexação, como no de recuperação

(*) Até agora, não se desejou utilizar a palavra "tesauro", para proceder logo à sua definição.

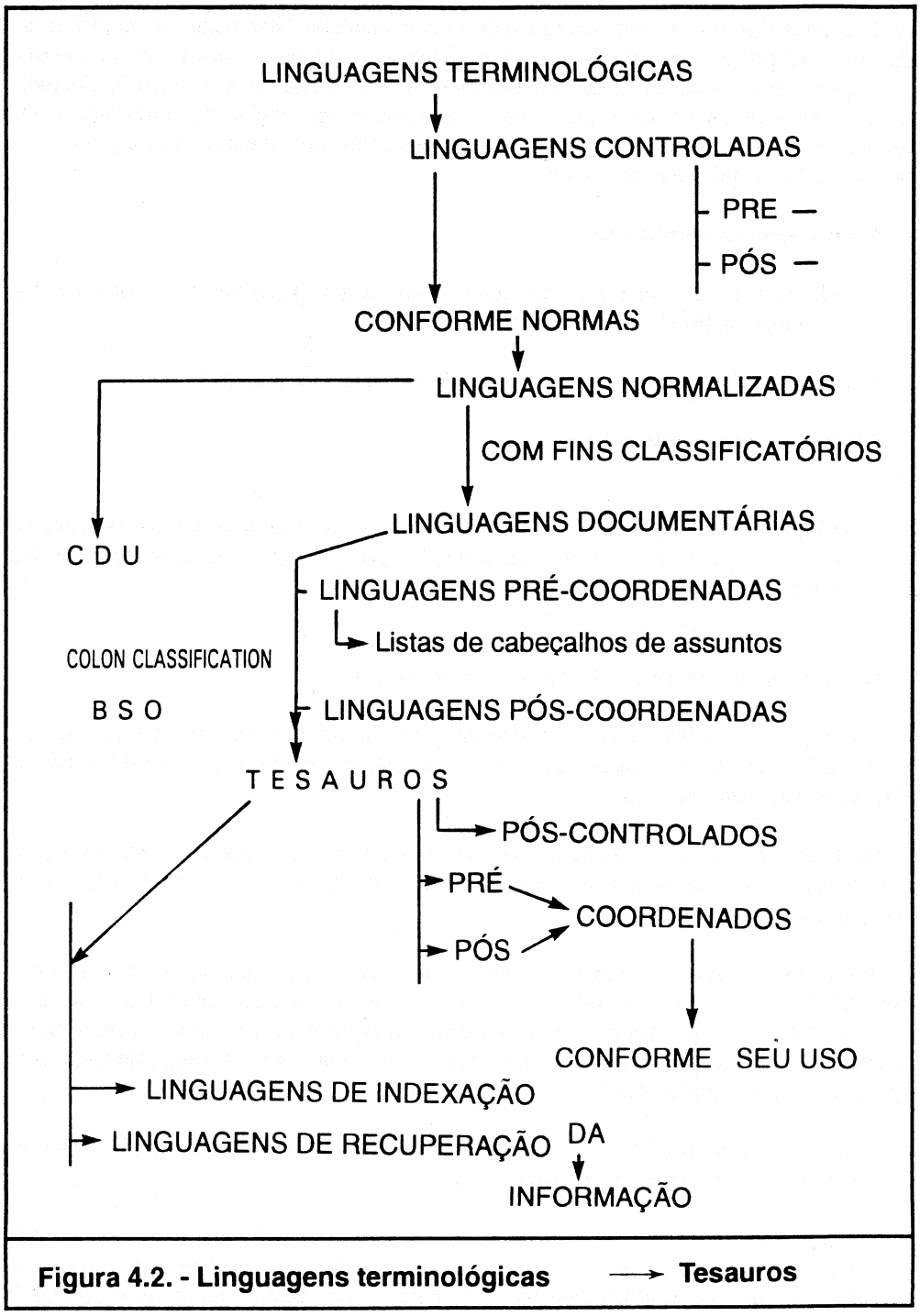


Figura 4.2. - Linguagens terminológicas → Tesauros

da informação, Uma lista de termos (de indexação), tomemos como caso, será uma linguagem documentária pós-controlada, na qual se coordenam os elementos que a compõem na fase prévia à indexação (pré-coordenação), ou em uma fase posterior (pós-coordenação). Na figura 4.2, está resumido o que foi exposto aqui.

OS TESAUROS

Intencionalmente, desejou-se não dar mais de um exemplo no parágrafo anterior, referente aos tesauros, para poder dedicar-lhes mais atenção no item seguinte, dada sua importância nos processos documentários. Há quem afirme, contudo, que a época de esplendor está chegando ao fim(Os novos métodos informatizados, com os sistemas especialistas nos processos de inteligência artificial, possibilitam o uso da linguagem natural, o que faz com que não precisemos quebrar a cabeça com todo esse assunto complicado das relações entre os termos de uma linguagem documentária. Certamente, pode chegar esse momento. Porém, pensamos que somente diminui o trabalho do usuário, ou talvez o do documentalista intermediário, que realiza buscas de informação. Àquele que desenvolve o programa do sistema especialista, àquele que possibilita a recuperação de documentos, pergunte-se como se deve perguntar, a um sistema informatizado de documentação. Esse documentalista-analista deve saber muito bem como estabelecer as relações entre os termos e como coordenar os conceitos. O trabalho intelectual não é feito pelas máquinas, e sim pelos homens.)

Assim, pois, foi dada a este tema a importância que merece, pelo menos, nestes últimos anos do século XX. Passemos a definir o que se entende por tesauro.

É precisamente na extensa bibliografia que se nota um marcado interesse pelos tesauros, de uns anos para cá. É curioso que todos os profissionais documentalistas, com maior ou menor experiência, queiram dizer algo sobre tesauros. Esta tendência não aparece somente na Espanha. Manifesta-se também, em outros países, o interesse por estes atraentes sistemas de classificação. Procurando uma explicação para o caso, ocorre-me pensar que tudo situa-se na proliferação dos sistemas informatizados de documentação e na possibilidade de usos domésticos, com o auxílio dos computadores de uso individual (N.T. Microcomputadores da linha PC - Personal Computers). Penso também como são pouco esclarecedoras algumas definições e os métodos de construção. Todos nós desejamos clarear mais um assunto que se apresenta tão complicado, apesar de ser, por outro lado, muito simples. O mal vem, precisamente, da simplicidade do assunto, que se complica logo com normas e explicações complexas. O mal vem, igualmente, das origens, no que se refere ao fato de tratar-se de um assunto desenvolvido, originalmente, para o idioma inglês, com suas peculiaridades. Ao desejar-se aplicar aos outros idiomas, esbarra-se nas diferentes peculiaridades de cada um deles. Em nosso próprio país, desejamos elaborar uma norma UNE, e o que fazemos é traduzir a norma internacional ISO, a qual, embora "internacional", refere-se ao idioma inglês. Devo confessar, para minha maior culpa, que participei da comissão encarregada da tradução. E devo reconhecer, também, que a norma ISO facilita em grande modo o trabalho, pois há muito de aproveitável nela.

DEFINIÇÕES DE TESAUROS

Passemos, agora, às definições de tesauros, para logo começar a estudar sua evolução histórica e os tipos de tesauros existentes. Sua estrutura e construção, bem como o seu uso, serão tratados nos capítulos seguintes.

Na bibliografia consultada, foram encontradas até mais de 30 definições de *thesaurus*.

Esta palavra, em espanhol, converteu-se em tesauro, generalizando-se o seu uso em todos os países de língua espanhola.

De todas aquelas definições, faremos constar somente as que parecem mais relevantes.

Comecemos pelas próprias definições de *thesaurus* encontradas nos dicionários ingleses e americanos, as quais são a origem de seu uso posterior, no âmbito das ciências da documentação.

Assim, o Shorter Oxford Dictionary, de 1736, diz (*):

> • Um tesouro, ou depósito de conhecimentos: por exemplo, um dicionário, uma enciclopédia e outras obras semelhantes.

Por sua vez, no Webster's American Dictionary, em edições sucessivas escreve-se:

> • Um livro de palavras, ou informação sobre um determinado assunto, ou conjunto de conceitos, especialmente um dicionário de sinônimos.

Em 1852, Peter Mark Roget publicou sua obra "Thesaurus of English Words and Phrases", na qual se lê:

> • Uma coleção de palavras e frases ordenadas, não em ordem alfabética, como se encontram em um dicionário, mas de acordo com as idéias que representam.

Isto significa que se tem a idéia, devendo-se, porém, procurar a palavra, ou palavras, que se ajustem mais exatamente a essa idéia.

Esta definição de Roget foi a que serviu como base para as sucessivas interpretações e uso do termo "thesaurus".

Passemos, agora, ao final da década de 50, quando começaram o desenvolvimento e a utilização dos sistemas de indexação e classificação à base de palavras-chave.

(*) Todas as traduções e transcrições são livres, adaptando-se o mais possível ao texto em inglês ou alemão e original em castelhano, conforme o caso.

Uma das primeiras definições feitas naqueles anos é a de Howerton, que diz:

- Uma lista autorizada, que pode conduzir o usuário de um conceito a outro, por meio de relações heurísticas ou intuitivas. Pode-se usar a lista manual ou mecanicamente, para indicar cabeçalhos de indexação.

Vários autores, como Brian Vickery e Alan Gilchrist, comentam que Helen Brown, em 1957, foi a primeira pessoa que usou a palavra *thesaurus*, que apareceu impressa durante a Dorking Conference on Classification, quando foi escrito:

- O problema da recuperação da informação é transformar conceitos e suas relações, da forma como se expressam na linguagem dos documentos, em uma linguagem mais regularizada, com os sinônimos controlados e sua estruturas sintáticas simplificadas.

Deixemos passar alguns anos e nos situemos no começo da década de 70. Precisamente em 1971, apareceram duas obras importantes, uma no Reino Unido, da autoria de Alan Gilchrist, outra na República Federal da Alemanha, da autoria de Gernot Wersing. Ambas se concentram na construção de tesauros. A obra de Gernot Wersing, do mesmo modo, é um estudo detalhado dos princípios teóricos, dando lugar a uma "teoria de tesauros". As duas publicações serviram de base para a maioria dos estudos posteriores.

Alan Gilchrist escreveu:

- Uma lista autorizada de léxicos, sem notação, que difere de uma lista de cabeçalhos de assuntos, na qual as unidades léxicas, sendo menores, são mais maneáveis e se utilizam na indexação coordenada.

Por sua vez, Gernot Wersing fala de:

Listas de termos, prefixados com antecedência, porém tirados do texto dos documentos, que desdobram os conceitos em unidades simples. Estas se coordenam posteriormente, para evitar ambigüidades. Entre elas, estabelecem-se relações hierárquicas, associativas e de equivalência.

Naquela época, iniciei umas aulas na Universidad Autónoma de Madri e não posso omitir uma definição de tesauro, muito intuitiva e apropriada para introduzir os alunos universitários neste assunto:

- Vocabulário especializado, no qual as palavras que o compõem estão relacionadas umas com as outras semanticamente.

Logicamente, deve-se explicar logo como se estabelecem essas relações e sua forma de uso.

Em 1976, foram publicados os manuais da Unesco, dentro do programa Unisist, onde são definidos os tesauros, segundo sua função e estrutura:

- Segundo sua função,

  ◦ tesauros constituem um instrumento de controle terminológico usado para transferir os descritores retirados da linguagem natural dos documentos para um sistema lingüístico.

  ◦ segundo sua estrutura:

  ◦ tesauros são vocabulários controlados e dinâmicos de termos relacionados, semântica e genericamente, que cobrem um domínio específico do conhecimento.

Nos tesauros multilíngües, os conceitos estão representados em mais de uma língua.

É meu desejo lembrar os profissionais, colegas, muitos deles amigos, como R. Fugmann, I. Dahlberg, N. Henrichs, F.W. Lancastaer, I. Aitchson, Maria Oyarzún, André Deweze, entre vários outros dedicados a estes assuntos. Desejo lembrar, igualmente, meus compatriotas, alguns muito insígnes, como D. Javier Lasso de la Vega, Nuría Amat, García Guitiérrez, Mercedes Caridad, María Luz Terradas, entre outros conhecidos de todos. Com o propósito de tornar esta epígrafe mais compreensível, foram manisfestadas somente as tendências evolutivas em alguns determinados momentos. Este é o motivo pelo qual passamos, agora, aos anos desta década de 80-89, citando algumas definições surgidas recentemente.

DEFINIÇÕES ATUAIS DE TESAUROS

Para compreender melhor as transformações experimentadas com a chegada das ciências da documentação à terminologia, menciona-se, em seguida, a definição de B. Long, surgida em 1980, que estabelece um marco importante nos estudos da teoria dos tesauros:

- Tesauro é um conjunto semiológico, utilizado como sistema de classificação, cujas as unidades classificatórias formam conjunto de morfemas ligados por relações do tipo paradigmático, a partir das quais se constroem diferentes classes. Estes morfemas utilizam relações do tipo sintático, para conseguir que os termos tenham uma representação documentária. O importante não são os documentos, são as relações dos assuntos. Esta tarefa é realizada pelo indexador.
  (Nota: apanhada do artigo de Angelina Boch de Doze, publicado na R. Esp. Doc. Ci., v. 9, n. 2, p. 70. 1970).

Prosseguindo, cita-se outra definição, que representou, também, um marco importante, pelo fato de ser da autoria de um terminólogo, o austríaco Helmut Felber. Considera os tesauros do ponto de vista da terminologia, no seu *Terminology Manual*, de 1984. Ali, diferencia os vocabulários, os dicionários e os *documentation thesauri*. Destes últimos, diz-se que consistem em:

- Elementos – dados terminológicos – que, por um lado, ajudam a controlar uma linguagem específica de indexação e, por outro, indicam as relações entre conceitos.
  (Nota: conceitos que foram estudados separadamente em outra parte desta obra).

Nestes últimos anos, apareceram várias publicações, livros, apontamentos de aula e manuais, nos quais figuram definições de tesauros interessantes de destacar, precisamente, por sua atualidade.

Entre os autores além-fronteiras, parece oportuno citar, novamente, Alan Gilchrist, para poder comprovar sua evolução ao longo dos anos. Em 1987, publicou o livro *Thesaurus construction*, onde define tesauro como:

- um vocabulário de uma linguagem controlada de indexação, organizado formalmente, de forma que as relações entre conceitos se estabeleçam *a priori*, para ser usado em sistemas de recuperação da informação, sejam estes bases de dados, índices impressos ou catálogos.

Deste lado das linhas fronteiriças, cabe citar, por ordem cronológica, as definições surgidas bem recentemente.

Em seu livro *Manual de Bibliotecas*, de 1987, Manuel Carrión dedica uma extensa referência ao assunto. Ali, lê-se, entre outras descrições que

- é um conjunto de termos, ou palavras, com indicação dos sinônimos e das relações mútuas entre eles por motivo de significado;

- é um vocabulário curto, modular, que permite sua própria ampliação constante.

O grupo TermEsp organizou, em 1988, repetido em 1989, um curso de introdução à terminologia, em cujos apontamentos Consuelo Ruiz oferece várias definições do termo tesauro, das quais destacamos:

- Tesauro é uma linguagem documental combinatória. Surge como um anel de cadeia essencial na cadeia que assegura a união documento-usuário, de tal modo que, se partimos de um documento para indexar, como de um pedido de informação, aparece a figura do tesauro, que nos permite passar da linguagem natural para uma linguagem documental.

- Tesauro é uma relação organizada de descritores e não descritores, que obedecem a regras terminológicas próprias, unidos entre si por relações hierárquicas ou semânticas.

Não poderíamos deixar de citar Amelia de Irazazábal, da qual destacamos seu artigo publicado na *Revista Espanhola de Documentação Científica*, em junho de 1989, realizado em parceria com E. Laguna e A. Valle, onde se define um tesauro como:

- uma linguagem documental, uma linguagem artificial;

- expoente mais elaborado de uma linguagem artificial (usada com fins documentais – indexação ou recuperação da informação –), considerada como linguagem controlada, na qual se encontram as relações semânticas, além das de significado.

Por último, menciona-se, aqui, a norma UNE 50-106, de 1989, tradução da norma ISO 2788, em sua segunda edição de 1986, onde se define, entre outras acepções, um tesauro como:

- um vocabulário de uma linguagem de indexação controlada (onde esta é uma série de termos extraídos da linguagem natural e utilizados para representar, de forma breve, os assuntos dos documentos), organizado formalmente, com objetivo de tornar explícitas as relações *a priori* entre conceitos (por exemplo: é mais genérico que, ou é mais específico que).

Não gostaria de terminar esta lista de definições sem mencionar uma das que tenho utilizado em meus cursos nos últimos tempos e que segue a linha geral de evolução (figura 4.3).

- Tesauro é uma linguagem especializada, normalizada, pós-coordenada, usada com fins documentários, onde os elementos lingüísticos que o compõem – termos, simples ou compostos – encontram-se relacionados entre si sintática e semanticamente.

CONDIÇÕES A SEREM CUMPRIDAS POR UM TESAURO

De toda esta série de definições, clássicas ou recentes, de tesauros, conclui-se que, para que se considere como tal, o tesauro deve cumprir as seguintes condições:

- ser uma linguagem especializada;

- estar normalizado
  - em um processo
    - pós-controlado;

- as unidades lingüísticas que o compõem, por tratar-se de léxicos dedicados a um assunto concreto, adquirem a categoria de termos
  - convertidos em palavra-chave,
    pois determinam o assunto de que trata o documento;

- estas palavras-chave relacionam-se entre si,
  hierarquicamente, de forma associativa ou por semelhanças de equivalência;

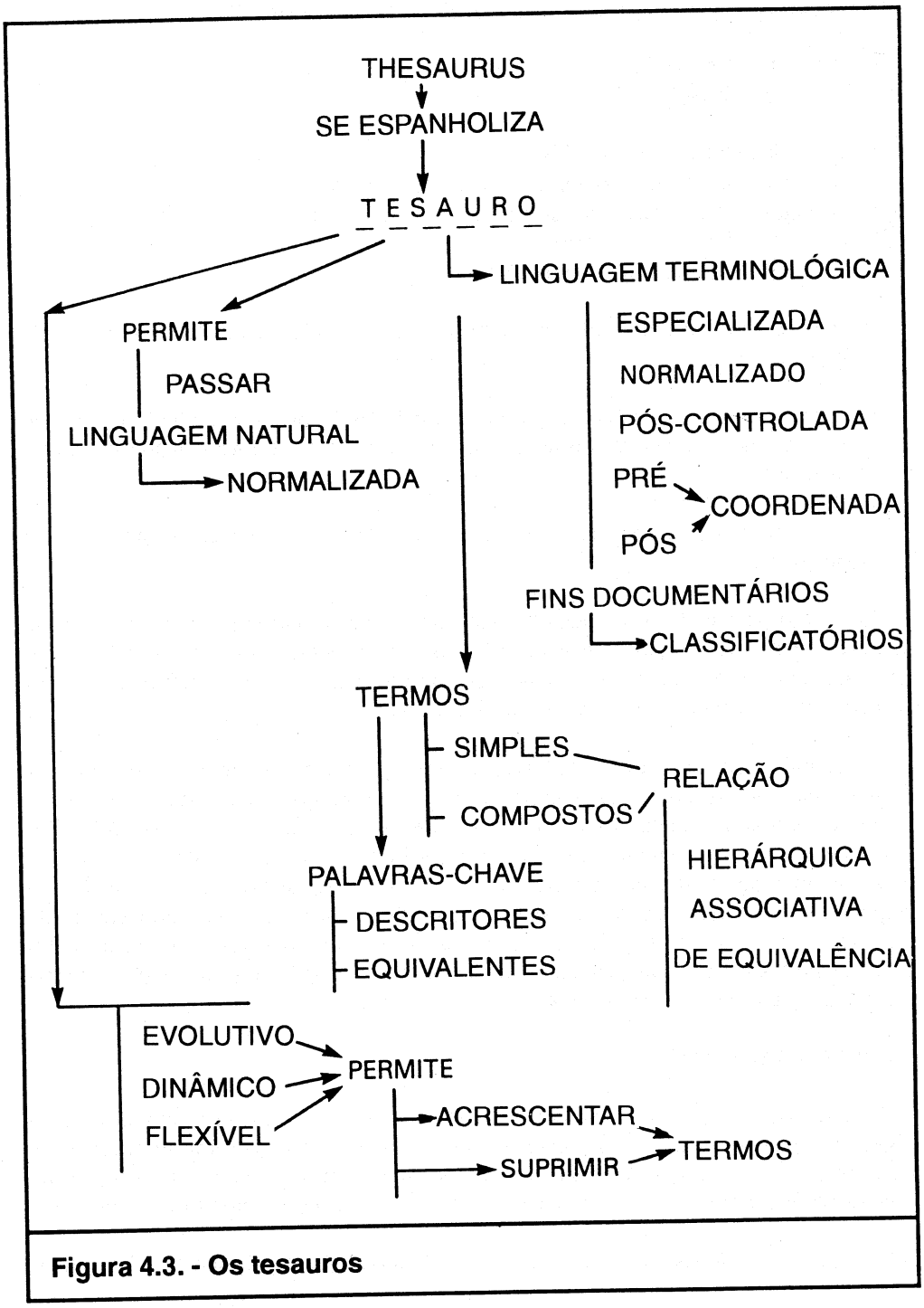


Figura 4.3. - Os tesauros

- estes processos de relação poderiam se realizar com métodos de:
  - pré-coordenação
  - pós-coordenação;

- trata-se de linguagens terminológicas, usadas com fins documentários, portanto, convertem-se em:
  - linguagens documentárias utilizadas nos processos de
    - indexação ou classificação
    - recuperação da informação;

- devem permitir a introdução ou supressão de termos, para manter sua atualidade constantemente;

- devem servir para converter a linguagem natural, dos documentos, ambígüa e livre, em uma linguagem concreta, normalizada, apta a controlar a informação contida no documento;

- devem servir de ligação entre o documento e o usuário, onde o documentalista é o elo fundamental.

Talvez possam parecer demasiadas as condições que deva cumprir um bom tesauro e que se trate de uma questão por demais complicada para aprofundar-se na mesma. Se forem analisados os diferentes capítulos, poderá ser comprovado que todos eles se reduzem a simples regras, explicadas com maior ou menor detalhe (figura 4.3).

A dificuldade consiste, talvez, na forma em que se deve estabelecer as relações entre os termos e sua composição. Dá a impressão de que os idiomas naturais tornaram-se menores. Anteriormente, comentei que possuímos mais idéias do que palavras para expressá-las. Por outro lado, parece que sentimos uma espécie de inércia em utilizar as palavras adequadas para expressar cada pensamento ou aparelho e, com isso, começamos a perder expressões e a empobrecer o idioma. Talvez a gramática e a lingüística não estejam sendo suficientemente estudadas. De todo modo, no que concerce à construção e ao uso de tesauros, ainda não foi formado um consenso. Como poderemos ver, nos capítulos correspondentes, algumas regras contribuem para a solução desses problemas.

EVOLUÇÃO HISTÓRICA

Sempre parece útil estudar a evolução, ao longo da história, dos assuntos abordados, para poder compreender melhor como foram se desenvolvendo e sofrendo transformações até chegar ao estado atual, que, naturalmente, não pressupõe um ponto final, mas somente um estágio intermediário, como os anteriores, suscetível de mudanças e modificações futuras.

No que se refere à chamada "teoria de tesauros", tudo começou pelo aumento de assuntos

– conhecidos e de criação nova – que apareciam na literatura cada vez mais extensa e de todo tipo: os sistemas hierárquicos ou facetados não respondiam adequadamente às demandas de informação. As bibliotecas ou os centros de documentação tornavam-se obsoletos e eram subutilizados, com o conseqüente prejuízo econômico que isso ocasionava. Tornava-se necessário procurar uma solução.

A solução veio das mãos dos que pensavam que o problema consistia em encontrar, planejar novos sistemas de classificação que permitissem mais flexibilidade no tratamento dos assuntos contidos nos documentos. Pensou-se, todavia, que uma ordenação alfabética daqueles assuntos deixaria lacunas no momento de intercambiar suas relações. Retornou-se ao passado, fixando a atenção em documentos que apresentavam estruturas semânticas na ordenação dos conceitos que continham. M. Carrión menciona o *Libro de los Epítomes* e o *Libro de Materias o Proposiciones*, ambos de autoria de Hernando Colón, e o *Dictionarium Históricum*, de Charles Estienne, publicado em 1561, onde figuram classificações sistemáticas. Em outros segmentos desta mesma obra, já mencionamos documentos ordenados dessa forma. A idéia não era nova. Já havia sido utilizada. Era necessário, porém, saber se seria aplicável nos processos documentários. Desta forma, surgiram os primeiros métodos de classificação utilizando conceitos tirados dos próprios documentos, sem ligação prévia. Receberam o nome de *thesaurus*, atendendo a definições e documentos também existentes aos quais foram feitas referências neste mesmo capítulo.

Parece que tudo começou pelos meados dos anos 50, quando Howerton e Helen Brown, cada um separadamente, utilizaram pela primeira vez a palavra *thesaurus*. Este fato coincidiu com o desenvolvimento do Sistema Uniterm, elaborado por Mortimer Taube, e surgido em 1951.

Aqui, os termos que determinam os conceitos eram unidades lingüísticas simples. Segundo a bibliografia, ele foi o primeiro que utilizou a denominação de palavras-chave para designar as palavras que determinavam o conteúdo dos documentos. Criaram-se categorias de palavras-chave. Umas eram as palavras-chave principais, as que se utilizavam, e as outras eram seus sinônimos. Às palavras-chave principais chamou de "descritores", cunhando, dessa forma, a expressão para a posteridade.

Como, naquele tempo, também os computadores eletrônicos começaram a ser utilizados, a ordenação de unidades isoladas de classificação não importava grande inconveniente. Era no processo de busca de informação que se podia combinar os conceitos, usando o sistema Booleano, e se obtinha a informação desejada. É atribuída a esse autor a expressão *information retrieval*, recuperação da informação, que também passou a ser de uso generalizado. O Sistema Uniterm continua sendo utilizado na atualidade, em grandes instituições, com um volume considerável de documentos muito variados e pluralidade de assuntos. Uma dessa instituições é a Unesco.

Aquele sistema, contudo, não satisfazia a maioria, sobretudo aos que tratavam de assuntos muito concretos e de volume menor de documentos. Este foi o motivo que fez com que Calvin

N. Moores, segundo os dados disponíveis, já em 1949, criasse seu sistema Zator, onde as palavras-chave podiam ser termos compostos ou simples. Desta maneira, tornava-se possível a sua combinação na fase de construção do sistema classificatório. Por isso, Moores chamou seu sistema de fichas, já confeccionadas e devidamente dispostas, de Zatocoding.

Com os descritores construíam-se índices, o processo de obter aquelas palavras dos documentos com as quais se confeccionavam os índices chamou-se indexação. Não se tem informação exata da data em que se começou a usar essa expressão. Deve ter sido ainda antes de 1945, pois, nas referências às classificações facetadas, já se fala da indexação.

Os anos seguintes, de 1947 a 1960, aproximadamente, foram tempos de grande atividade no que se refere ao estabelecimento dos princípios teóricos e das definições dos diferentes processos.

A questão da evolução histórica da teoria de tesauros não foi, todavia, abordada com seriedade e rigor científicos. Por isso, os dados que se encontram são contraditórios, tornando-se difícil precisar momentos ou situações. Dissemos que Howerton e Helen Brown foram os primeiros a oferecer uma definição de tesauro. Lemos, em outras publicações, que essa honra é devida a Joyce e Neddham, em 1956. Lemos, também, que foi Farradane, em 1952, quem falou de tal vocabulário pela primeira vez.

Em outro local, atribuiu-se ao próprio M. Taube a expressão indexação coordenada, ainda que em data muito próxima ao final da década de 50. Encontramos outra referência, no livro de Alan Gilchrist, publicado em 1971, que informa ter sido Jolley, em 1963, quem definiu a indexação pré-coordenada. Na realidade, não é muito importante conhecer datas com detalhe. O relevante é observar a evolução que se foi produzindo e que é mostrada nos documentos publicados ao longo de todos esses anos.

Os primeiros tempos, naturalmente, são aqueles em que se procura estabelecer as bases teóricas e definir os conceitos. Assim, na época anteriormente mencionada, foram fixados os conceitos de descritor, indexação, indexação coordenada, relevância e pertinência. Nota-se a grande preocupação em construir um sistema que determinasse o conteúdo dos documentos.

Os primeiros tesauros formalmente construídos apareceram a partir de 1960. Por exemplo, o *Thesaurus da Armed Service Technical Information Agency* (ASTIA), convertido, posteriormente no *Defense Documentation Center* (DDC), foi publicado, precisamente, em 1960. Começa, assim, uma etapa em que a preocupação está centralizada na construção de tesauros. Contudo, esta época não se destaca por sua atividade excessiva. A década seguinte, a dos anos 70, apresenta um florescimento de acontecimentos, continuando a ênfase na construção de tesauros. Observou-se que o interesse pelos tesauros foi aumentando, com uma utilidade preponderante na recuperação da informação. Esta tendência não diminuiu, todavia.

Entre 1969 e 1970, publicou-se *The Precis Index System - Preserved Context Indexing*

System, que foi o documento guia para a confecção da bibliografia britânica. Em 1971, apareceram as obras de A. Gilchrist e G. Wersing.

Anteriormente, em 1967, havia sido criado um grupo de trabalho, na República Federal da Alemanha, para construir um tesauro para a indústria química, do qual tive a honra de participar.

O *TEST - Thesaurus - Thesaurus of Engineering and Scientific Terms* foi publicado em 1967 e constituiu a base na qual o Reino Unido fundamentou-se para construir seu British Standard 5723. Do mesmo modo, a ISO tomou esse documento como modelo para elaborar a norma ISO 2788, em sua primeira edição de 1974.

O desenvolvimento do campo da informática e o aparecimento no mercado de computadores cada vez mais manuseáveis e econômicos potencializou a automatização das ciências da documentação, e este foi o motivo do apogeu verificado no que se refere ao assunto dos tesauros, não somente em países mais ou menos industrializados, mas em todo o mundo. Merecem destaque, pela nossa proximidade cultural, os países ibero-americanos, cujos trabalhos realizados no Chile, na Colômbia, Venezuela, Argentina e no Brasil ressaltamos.

São numerosas as obras publicadas na década de 70, como também o foram as obras editadas na década seguinte. Nesses anos e até nossos dias, continua-se sentindo a necessidade de dispor de tesauros eficazes, tanto para a indexação, como para a recuperação da informação. Atualmente os computadores são, todavia, mais acessíveis e mais manuseáveis, o que faz com que, conseqüentemente, a atenção se centre no seu uso, inclusive para elaborar os tesauros. Os trabalhos mais recentes baseiam-se nas obras de R. Fugmann, Lancaster, Bonnie Snow, M.E. Sievert e B.R. Boyce, A. Gilchrist, D.J. Foskett, entre vários outros que desejaríamos aqui enumerar, mas que são citados na bibliografia.

No que se refere aos estudos teóricos, registra-se a tendência de se realizar uma união com a terminologia, com a inclusão da lingüística. Estas pesquisas servirão de base para a elaboração de sistemas especialistas, utilizáveis tanto em procedimentos documentários, como na própria elaboração de tesauros.

EVOLUÇÃO HISTÓRICA NA ESPANHA

Nosso país seguiu as linhas gerais de evolução, embora com a conseqüente defasagem provocada pelo grau de desenvolvimento industrial em que nos encontrávamos em cada momento histórico. De toda maneira, no que se refere à automatização de bibliotecas e centros de documentação, notou-se um grande impulso de dez anos para cá. A conseqüência apareceu nas atividades documentárias, principalmente no empenho de realizá-las com os instrumentos necessários para o tratamento da documentação. Essas atividades originaram a tradução de numerosos tesauros, que vinham sendo utilizados em inglês, e a construção de outros especializados, para os mais variados assuntos. As instituições que se automatizam, que já são numerosas, elaboram ou adotam os tesauros úteis para seus objetivos. O maior dinamismo foi verificado no Icyt e no ISOC, os quais, como institutos de informação e documentação, cada um especializado em seus assuntos, de caráter nacional, têm a responsabilidade de atender

às procuras provenientes de qualquer lugar do país. É aconselhável dirigir-se a eles para resolver qualquer problema a este respeito.

No âmbito relativo aos estudos teóricos sobre tesauros, devemos reconhecer que até a década de 80 não se verificou uma grande atividade. Os casos isolados, como o de Dom Javier Lasso de la Vega, em seu *Manual de Documentación*, ou a realização da Reuniber 78, que repercutiu seu trabalho nos anos seguintes, ou minhas contribuições a congressos ou simpósios, são precisamente as exceções à regra.

Os fatos que podemos destacar – entre outros e procurando não esquecer ninguém – são a criação da disciplina de lingüística documental e de teledocumentação nas Faculdades de Ciências da Informação e das Escolas Universitárias de Biblioteconomia e Documentação (Fuinca), com a promoção e o estudo dos tesauros. Pela mesma razão, a Fundesco também colaborou.

Em outra ordem de estudos, continua a linha internacional de aproximação com a terminologia, para construir tesauros com maior fundamento lingüístico e terminológico, mais aptos para realizar a recuperação da informação.

Trabalhos sobre a construção de tesauros com a utilização de sistemas especialistas estão sendo realizados no Instituto de Estudos Avançados de Blanes, sob a direção de López de Mántaras.

A atividade mais recente neste campo é a elaboração da norma UNE 50-106, traduzida da ISO 2788, com as modificações correspondentes. Sabemos que as normas não obrigam o seu uso e que são pouco consultadas. Porém, encontra-se à disposição de quem desejar fazer uso dela.

Jornadas de Documentação Automatizada foram realizadas em 1984, 1986 e 1990, com o comparecimento de grande número de profissionais interessados nesses assuntos. É aconselhável consultar as *Actas*, para obter maiores informações sobre o assunto.

CLASSES DE TESAUROS

Após a abordagem do assunto relativo à definição dos tesauros e de expor as características gerais que eles devem cumprir, passaremos, agora, ao estudo de suas classes e tipos (figura 4.4).

Sabemos que uma das condições principais que devem ser cumpridas por um tesauro é a de restringir-se a um assunto especializado. Contudo, construíram-se alguns de âmbito geral, de modo que uma primeira classificação será:

- gerais,
- especializados.

Apesar da especialização de um tesauro, este pode tratar de mais de uma disciplina, com níveis de importância distintos. Certamente que os tesauros são, por definição, multidisciplinares. Portanto, outra classificação feita é:

- multidisciplinares,
- monodisciplinares.

Observa-se que é muito difícil utilizar um tesauro estritamente monodisciplinar. Sempre será incluído um assunto marginal, complementar ao assunto principal. Por isso, muitas vezes, junto a um tesauro principal, deve-se utilizar outro auxiliar. Assim, aparecerão tesauros:

- principais,
- auxiliares,

Isto em relação ao âmbito do assunto que abrangem.

Quando se considera a ordenação do assunto, ou assuntos tratados, as classificações se deversificam. Sabemos que outra das características principais de um tesauro é sua simplicidade no que se refere às relações de seus termos. Parece, pois, que uma ordenação alfabética das palavras-chave será a melhor maneira de ordenar seus assuntos. Não obstante, a prática tem demonstrado que, na maioria dos casos, uma ordenação sistemática é necessária e útil. Por conseguinte, será necessário considerar tesauros:

- alfabéticos
- sistemáticos, chamados, igualmente, de temáticos.

Nos tesauros sistemáticos, os assuntos ali compreendidos podem ser ordenados hierarquicamente, distribuídos em facetas, ou disporem-se em gráficos apropriados ao caso. De onde concluímos que os tesauros poderão ser:

- hierárquicos,
- facetados,
- gráficos.

A maior parte dos tesauros apresenta uma estrutura composta por uma parte:

- alfabética,
- sistemática,
- gráfica.

A partir do momento em que começou a generalizar-se o uso dos computadores para armazernar os tesauros e a sua utilização tornou-se fácil e rápida, começaram a desaparecer as ordenações gráficas dos termos nos tesauros. Deve-se considerar que são muito trabalhosas em sua confecção, consomem muito tempo e sua utilidade é pequena, pois, em qualquer

caso, é necessário dispor de uma lista alfabética para a localização das palavras-chave.

Agora, cabe mencionar outro tipo de tesauros:

- macrotesauros,
- microtesauros.

Ao contrário do que se poderia pensar, não se trata de avaliar o número de termos com que se conta, mas, sim, a sua distribuição no conjunto de seu contexto. Assim, chama-se macrotesauro o que contém diferentes assuntos relacionados entre si, mas dispostos em lugares diferentes, como se tratasse de um conglomerado de vários tesauros menores. Realmente, isto é o que sucedeu. Um macrotesauro é composto de vários microtesauros, relacionados entre si por referências cruzadas, e abrange uma ampla gama de assuntos, cada um deles constituindo objeto de um microtesauro especializado. Estes podem separar-se do macrotesauro e constituir um tesauro por si mesmo. Citemos, como exemplos clássicos de macrotesauros, o da OCDE e o do Clades.

Todavia, cabe mencionar os idiomas em que se pode escrever os termos de um tesauro, conforme os quais se pode considerar os tesauros:

- monolíngües,
- bilíngües,
- plurilíngües.

A utilidade de cada um deles dependerá do centro onde sejam empregados e da finalidade que se queira atingir. Podem desempenhar um bom papel como dicionários auxiliares em tarefas de tradução.

No que se refere à entidade que confecciona um tesauro, deve-se considerar que os tesauros podem ser:

- públicos,
- privados.

Melhor deve ser considerado, se os tesauros foram elaborados com meios públicos ou privados e se são exeqüíveis ou não.

Para todo o assunto relacionado com tesauros, é aconsellhável consultar a numerosa bibliografia que existe em bibliotecas e centros de documentação.

TEMAS PARA ESTUDO

1. Como definiria um tesauro?
2. Dê sua opinião sobre a utilidade de um tesauro.
3. Considera apropriado enquadrar este sistema de classificação dentro da terminologia?
4. Como diferenciaria o processo de indexação do de classificação?
5. Diga o que entende por descritor.
6. Mencione alguma característica própria de um tesauro.
7. Diga o que entende por linguagem terminológica.
8. Diga em que se diferencia uma linguagem controlada de outra coordenada.
9. Como definiria uma linguagem de indexação?
10. Cite algum autor estrangeiro que tenha escrito sobre estes assuntos.
11. Diga o que entende por macrotesauro.
12. Como definiria um tesauro gráfico?
13. Diga que diferença existe entre um tesauro hierárquico e outro temático.
14. Qualificaria a CDU como uma linguagem terminológica?
15. Recorde que norma espanhola trata da confecção de tesauros.
16. Diga quando utilizaria a expressão "termo" ou quando utilizaria a expressão "descritor" para designar a unidade lingüística de um tesauro.
17. Diga a diferença que existe entre um tesauro multidisciplinar e outro multilíngüe.
18. Dê sua opinião sobre a utilização dos computadores para armazenar um tesauro.
19. Até onde limitaria o uso de tesauros em um centro de documentação?

REFERÊNCIAS BIBLIOGRÁFICAS

* AITCHINSON, J. G. A.: ''Thesaurus construction. A practical manual''; London, ASLIB, 1987.

* ALVARO BERMEJO, C., VILLAGRA RUBIO, A., SORLI ROJO A.: ''Desarrollo de lenguajes documentales formalizados en lengua espanola: una evaluación. I. Vigencia teórica y práctica de lenguajes controlados''; Rev. Esp. Doc. Cient., *12*, 2, 1989, 154-159.

* ASLIS: ''Information (monthly)''; London, ASLIB, 1973 to date.

CARRION GUTIEZ, M.: "El catálogo alfabético de materias", en Manual de Bibliotecas, Madrid, Fundación Germán Sánchez Ruipérez-Pirámide, 1987.

CURRÁS, E.: Curso sobre Teoría de Thesauros e Indización. Madrid, ICE Univ. Autónoma de Madrid, Feb.-Mayo 1983, Apuntes de clase.

DAHLSERG, I.: "Ontical structures and universal classification"; Bangalore, India, Sarada Ranganathan Endowment for Library Science, 1978, 42.

DEWEZE, A.: "Informática documental"; Lyon (Francia), Universidad Claude Bernard, Massen ed., 1988.

EVANGELISTA, C.: "El CLADES y los lenguajes controlados"; Actas de la Exposición de Lingüística Informática y de Terminología Científico-técnica, CSIC y Unión Latina, Madrid, 1987, 94-6.

FUGMANN, R.: "The complementarity of natural and indexing languages"; in Perreault, J.M. and Dahlberg, 1. (eds). Universal Classification II: Subject analysis and ordering systems. Proceedings of the 4th. International Study Conference on Classification Research, FID/CR, Augsburg, 28 June-2 July 1982, Frankfurt, INDEKS Verlag, 1982-83, 86-89.

GARCIA GUTIERREZ, A.L.: "Documento, Thesaurus y Ordenador"; Rev. Fac. Doc. Univ. Complutensen, *12*, Septiembre, 1986, 289-304.

* LANCASTER, F.W.: "Vocabulary control for information retrieval"; Arlington, Virginia, Information Resources Press, 1986.

LASSO DE LA VEGA, J.: "Los Thesaurus. Definición y Exposición del Sistema", en Técnicas de investigación y documentación, Madrid, Paraninfo, 1980.

NIEHOFF, R.T. and MACK, G.: "The Vocabulary Switching System: description of evaluation studies"; International Classification, *12*, 1, 1985, 2-6.

* Norma Colombiana 1476 la. Revisión. Documentación para el Establecimiento y Desarrollo de Tesaurus Monolingües; Bogotá, FID/CLA - ICFES, 1980.

ROS GARCIA, J.: "Lingüística documental"; Proyecto docente para la plaza de Catedrático del Area de Biblioteconomía y Documentación de la Escuela de Biblioteconomía y Documentación de la Universidad de Murcia; Murcia, 1989, Ejemplar mecanografíado.

NOTA: As obras citadas nos capítulos 4, 5 e 6 contêm partes que fazem referência a mais de um capítulo. Também, daqui pra frente, as obras assinaladas com um asterísco (*) possuem carácter geral e correspondem a outros capítulos.

5

A Estrutura nos Tesauros

Após haver estudado detalhadamente o que é um tesauro e as condições que um vocabulário especializado e especial deve cumprir para que seja considerado como tal, passemos, agora, a tratar da estrutura interna que deve ter e quais são seus componentes básicos. Veremos também, como se formam esses componentes básicos e de que maneira podem se relacionar.

Para isto, tomou-se como ponto de partida, principalmente, a norma UNE 50-160, que é a tradução-adaptação da norma ISO 2788 em sua edição de 1986.

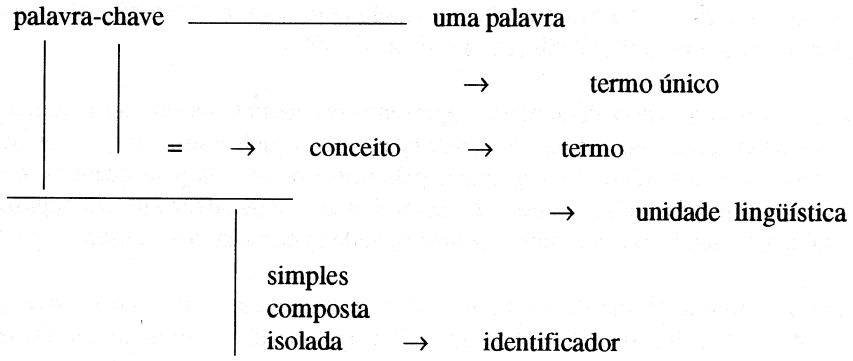
Tanto em uns como em outros documentos, apresenta-se o assunto da estrutura interna de um tesauro, mesclada com suas formas de construção e uso, o que resulta, talvez, em algo complicado para os que abordam estas questões pela primeira vez, ou para quem as tenha estudado sem se aprofundar nelas. Nesta obra, serão tratados separadamente, em capítulos diferentes, embora em um se faça referência a outro, quando as circunstâncias assim exigirem.

Ao estudar a estrutura interna de um tesauro, parece importante expor a diferença que existe entre este e uma lista de termos, ou uma lista de cabeçalhos de assuntos. Os três documentos são vocabulários, linguagens terminológicas, especiais ou especializadas, cujas unidades básicas estão constituídas por palavras normalizadas – termos –, geralmente referidas a um assunto concreto do conhecimento. A diferença consiste no fato de que a lista de termos, em linhas gerais, não possui outra relação entre eles, a não ser a própria ordenação alfabética de seus mesmos termos. Por sua vez, em uma lista de cabeçalhos de assuntos, os termos que a compõem acham-se relacionados entre si, *a priori*, em um processo de pré-coordenação, que confere ao mesmo uma determinada rigidez. Em um tesauro, os termos – simples ou compostos – encontram-se relacionados entre si de forma que permitem a sua combinação em um processo de pós-coordenação. Portanto, são mais flexíveis em si mesmos. Sua atualização torna-se mais dinâmica e rápida. Certas teorias pressupõem que, na construção dos termos, ao combinar umas palavras com outras para construir a unidade lingüística, utilizam-se técnicas de pré-coordenação. Certamente isso acontece. Contudo, deve-se considerar que as palavras-base – origem dos termos – tomam-se isoladamente dos documentos, dicionários, enciclopédias etc. Estas são questões importantes que devem ser tornadas bem claras, para evitar confusões no momento de se desejar construir ou utilizar um tesauro.

COMPONENTES FUNDAMENTAIS DE UM TESAURO

No início, há algum tempo, lá pelos últimos anos da década de 40, supunha-se que o elemento fundamental constituinte de um tesauro era a palavra, a palavra simples. O conceito, a idéia que se queria reter em um tesauro, devia ser expressa por uma palavra: a palavra que determinava o conteúdo do texto de um documento. Surgiu, assim, a expressão palavra-chave. O Uniterm System (sistema unitermo) baseia-se neste princípio.

Simultaneamente, desenvolveram-se outros sistemas que partiam da base de que a palavra-chave não tinha por que representar uma só palavra. Na realidade, o que representava era um conceito e este podia expressar-se por uma ou por várias palavras. Por outro lado, os nomes próprios de autores, entidades e locais geográficos também deviam ser considerados. Portanto, podia-se escrever:



Alguns autores denominam as palavras-chave isoladas de *identificadores*, pois auxiliam a identificar um documento. Não consideram que possam ser determinantes entre os componentes de um tesauro. Dizem, inclusive, que devem ser suprimidas para eliminar termos e conseguir que o tesauro torne-se mais compreensível. Estas informações devem ser realçadas, pois tudo dependerá do assunto coberto pelo tesauro e do uso que se faça do mesmo, embora não pareça ser aconselhável prescindir dos autores dos documentos, por exemplo.

Em outros documentos, sobretudo os anteriores a 1980, denominam-se identificadores os que classificam os termos a que fazem referência. Na norma UNE 50-106, foram chamados de indicadores classificatórios.

Nos últimos anos, parece que se está aceitando empregar a palavra identificadores para denominar essas palavras-chave isoladas e indicadores de função para denominar os *indicadores classificatórios*.

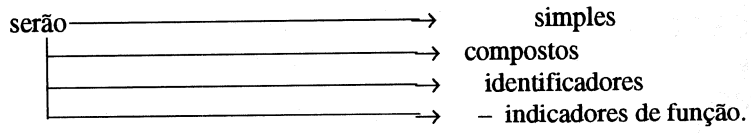
Com a aproximação à terminologia, as palavras-chave, embora expressando conceitos concretos, referidos a um vocabulário especializado, converteram-se em termos. Portanto, os componentes fundamentais dos tesauros são os termos. Termos que constituem unidades

lingüísticas factíveis de serem examinadas lingüística e gramaticamente. É importante expor esta característica para poder compreender mais facilmente a construção ou separação de determinados termos, tanto no momento de sua composição, como no de seu uso.

CLASSES DE TERMOS

Do que foi exposto no parágrafo anterior, deduz-se que, em uma primeira classificação (figura 5.1), os

termos



Isto em relação a sua composição ou sua função no tesauro.

Vejamos, agora, que classes de termos podem ser consideradas, tendo em vista sua importância ou sua relevância. Quando se estudam detidamente os termos de uma linguagem especializada, nota-se que alguns determinam o assunto com precisão, que são imprescindíveis. Outros, pelo contrário, ou são sinônimos reais ou podem ser substituídos pelos primeiros, sem que o conteúdo de tesauro perca o sentido ou fique incompleto. Isto significa que haverá

termos

- principais,
- secundários.

Pode-se dizer que os termos principais são os que descrevem o assunto, motivo pelo qual foram chamados de descritores. Tanto na obra de A. Gilchrist como na norma UNE 50-106, fala-se de termos preferidos, pois ambas se baseiam na norma ISO 2788.

Por contraposição, os termos sinônimos e quase sinônimos, que não são considerados importantes e que podem ser substituídos pelos termos preferidos, foram denominados "não-descritores" ou "não-preferidos", segundo os casos. Esta forma de designação por contraposição, usando a negação, nunca me agradou, pois, além disso, pressupõe utilizar um termo composto com um hífen intermediário, e isso sempre é embaraçoso, principalmente para os programas de computador. Na minha opinião, torna-se mais apropriado, inclusive lingüisticamente, procurar um termo na forma afirmativa. Em outros trabalhos, empreguei a expressão termo secundário, como oposto a "principal". Recentemente, resolvi chamá-los "termos equivalentes", o que está na linha de sua relação semântica com os descritores ou preferidos. Alguns especialistas nestes assuntos optaram também por esta denominação.

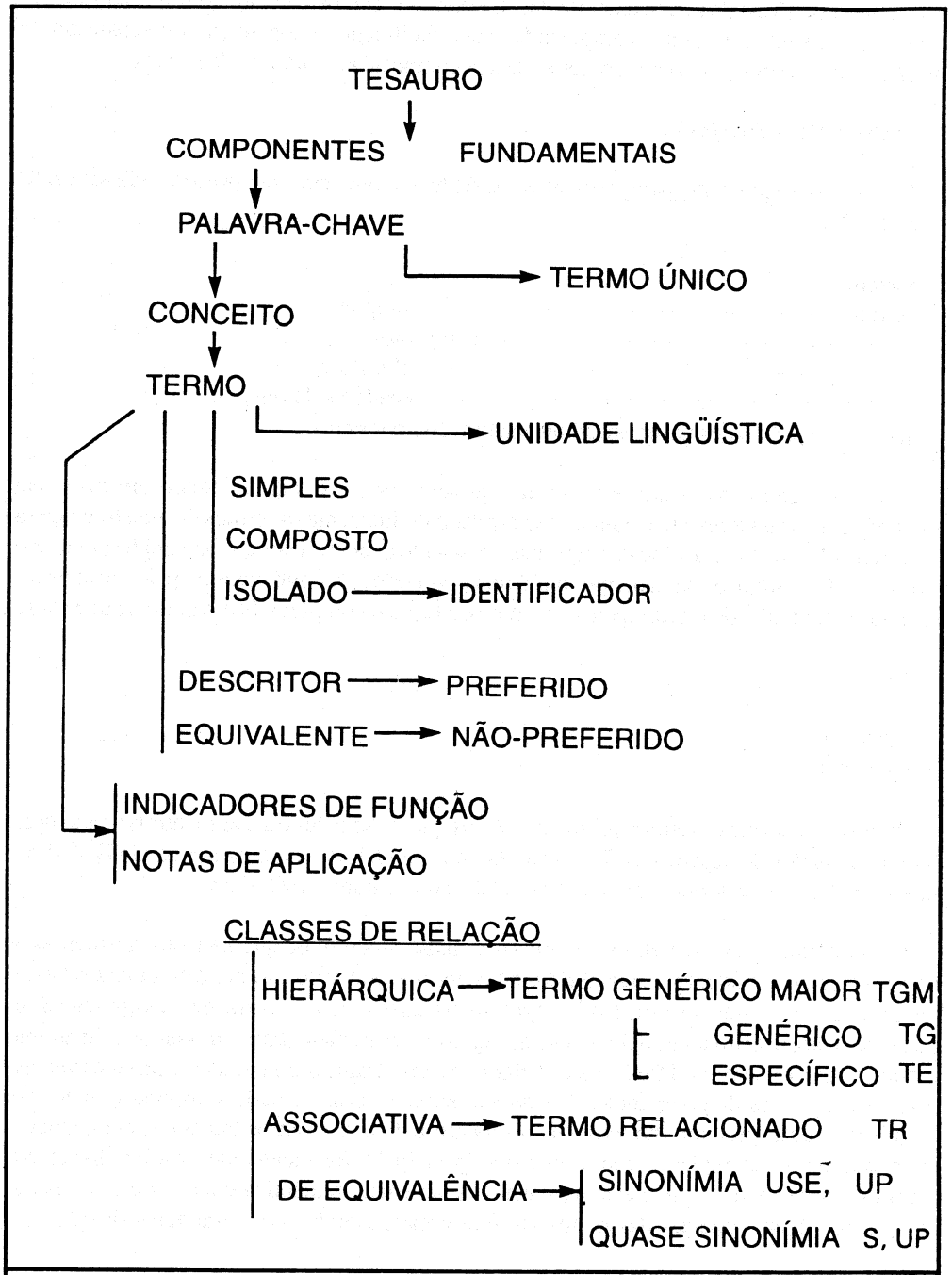


Figura 5.1. - Componentes fundamentais de um tesauro

Portanto, segundo sua relevância, os

termos

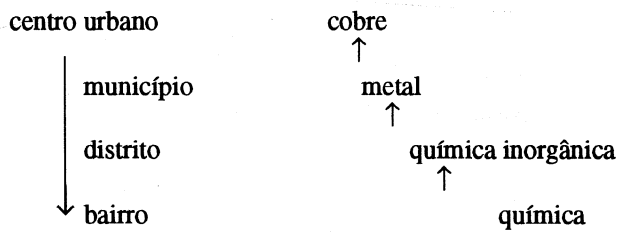
serão

- descritores → preferidos
- equivalentes → não-preferidos

TIPOS DE RELAÇÃO ENTRE TERMOS

Acabamos de abordar as relações semânticas entre os termos e devemos continuar dizendo que essas relações não se limitam somente àquele caso, mas que se deve considerar outras situações que se apresentam entre os termos de um tesauro.

Precisamente, uma das características de um tesauro é que os termos se liguem uns aos outros, segundo indicações estabelecidas pelo seu próprio significado, de modo que uns nos conduzam a outros. Por isso, o mais prático é organizá-los em famílias ou grupos, onde se consideram as afinidades semânticas (figura 5.1). Se as famílias se organizam em grupos onde figure um termo superior (genérico) e termos de menor conteúdo de significado, termos subordinados (específicos), teremos estabelecido uma relação hierárquica de maior a menor, ou de menor a maior. Tomemos um exemplo:



Se as famílias se organizam por associação de idéias em nível horizontal, onde não é possível estabelecer uma hierarquia, mas sim relações semânticas, se falará, precisamente, de relação associativa e os termos serão relacionados entre si.

Tomemos os seguintes exemplos:

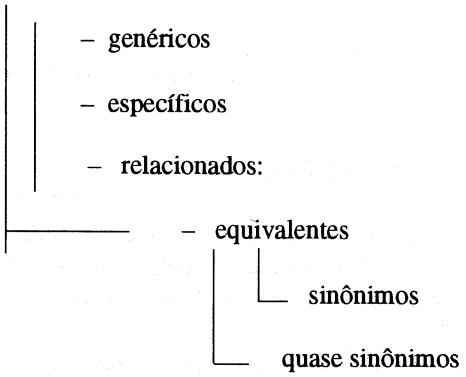
Estarão relacionados no mesmo nível:

- mobiliário
→ carpintaria

.....

- química inorgânica
→ química orgânica

Ainda deverão ser consideradas as relações de equivalência abordadas anteriormente. Conseqüentemente, segundo as relações semânticas, os termos serão:



Na norma ISO 2788 e, por conseguinte, na norma UNE 50-106, mencionam-se, ainda, outras categorias nas relações hierárquicas, designadas como

- termos
- termos principais
- termo genérico - genérico,
- termo genérico - específico → partitivo
- termo específico - genérico,
- termo específico - específico → partitivo.

Todos eles indicam distintos níveis de hierarquia dentro de uma mesma família e podem ser substituídos utilizando-se formas mais simples. Em muitos casos, empregam-se os subíndices com numeração correlativa.

OBSERVAÇÃO: Todas estas relações e dependências entre os termos são recíprocas.

NOTAS DE APLICAÇÃO

Em certas situações, é necessário precisar o uso de um termo em um contexto determinado. Neste caso, acrescenta-se uma nota, que pode ter duas funções diferentes: definir o termo neste determinado assunto ou tornar claro o seu uso também em assunto prefixado.

Por exemplo, o termo "administração" necessitará de uma definição concisa que indique se o termo é usado para referir-se aos funcionários de um Estado ou para referir-se à ação de administrar.

Um exemplo clássico, no qual se necessita de uma nota esclarecedora, é o dos homônimos. Estas são palavras escritas de uma mesma maneira, mas que possuem significado diferente, ou que podem ser aplicadas a diferentes campos do conhecimento. Um caso muito ilustrativo é dado pela palavra mercúrio, que pode ser um metal, um deus da mitologia grega, ou uma marca comercial. Outro exemplo igualmente significativo é dado pelo termo "plantas", que pode referir-se a vegetais, animais ou indústrias. As notas de aplicação que determinam o assunto a que se referem os termos homônimos formam parte do termo.

Na norma ISO, abordam-se os termos homógrafos. Em castelhano, são escassíssimos estes casos, já que é a acentuação o que faz distinguir os termos. Exemplo: cartel, cártel, já não são homógrafos. Contudo, temos que fazer referência aos termos homônimos, que, por sua vez, são homófonos e polígrafos. Em espanhol, contamos com alguns casos.

Exemplo:

erva zinco (N.T. Em espanhol, hierba cinc)

erva zinco (N.T. Em espanhol, yerba zinc)

Estes termos serão tratados empregando-se a relação de sinonímia. Escolhe-se a forma mais usual como descritor e a outra como termo equivalente.

Exemplo:

HIERBA ZINC

UP Yerba UP Cinc

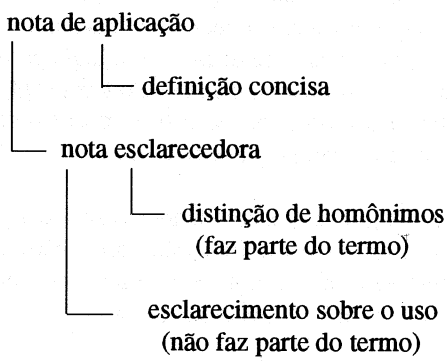
Citemos outro exemplo deste tipo, no qual a nota de aplicação não faz parte do termo. Em um tesauro sobre construção, pode ser necessário esclarecer que o termo

MADEIRA

Usa-se como elemento empregado em decoração.

Para designar estas notas, generalizou-se a expressão "nota de aplicação", que abrange os casos aqui descritos.

Assim, pois,



NOTAÇÕES UTILIZADAS PARA DESIGNAR AS RELAÇÕES ENTRE TERMOS

Seguindo os acordos internacionais, adotaram-se notações alfabéticas – com subíndices no caso – para designar as relações entre termos. São citadas em continuação, na mesma ordem em que se usam nos tesauros. São colocadas sempre precedendo o termo ao qual fazem referência.

NA	Nota de aplicação.
USE	Indica o descritor – termo preferido –, que se escolhe entre vários termos sinônimos ou quase sinônimos.
UP	Indica o termo equivalente - não-preferido
TGM	Termo Genérico Maior. Aplica-se aos termos que identificam o nome da classe mais ampla a qual pertence o conceito específico; emprega-se, algumas vezes, na seção alfabética dos tesauros.
TG	Termo genérico.
TG	Termo genérico (genérico).
TGP	Termo genérico (partitivo).
TE	Termo específico.
TE	Termo específico (genérico).
TEP	Termo específico (partitivo).
TR	Termo relacionado.

Exemplo:

BAÍA
TC OCEANOGRAFIA
TG MAR

Se for necessário dispor de termos genéricos ou genéricos partitivos etc., será necessário somente aplicá-los nos níveis de hierarquização correspondentes. Isto costuma ser utilizado em assuntos muito complexos, para obter a maior informação possível sobre um termo e que possa ser localizado e utilizado com propriedade. Nestes casos, o que se pretende é evitar ambigüidade, naturalmente, complicando ou tornando mais explícito o tesauro.

Relação associativa

Exemplo:

BAÍA
TR PORTO DESPORTIVO
PORTO DESPORTIVO
TR BAÍA

No exemplo seguinte, desenvolveu-se um termo com todas as suas relações possíveis:

BAHIA

NA	Geograficamente considerada
UP	ENSEADA
TC	OCEONOGRAFIA
TG	MAR
TGG	OCEANO
TGP	GOLFO
TE	PORTO DE MAR
TR	PORTO DESPORTIVO

A figura 5.2 reproduz uma página do "inspec-thesaurus - Institute of Electrical Engineers", de 1987, na qual aparecem os descritores com seus termos relacionados, mostrando os distintos tipos de relações.

TT significa Top Term = Termo Genérico Maior, CC faz referência a uma classificação interna, DI indica a data de inclusão do termo no tesauro.

FORMAS DE TERMOS QUE COMPÕEM UM TESAURO

Foram estudados os termos e suas relações do ponto de vista semântico. Agora, serão estudados do ponto de vista sintático, considerando o termo em si mesmo e suas conotações internas.

Inspec - thesaurus, 1987
The Institute of Electrical Engineers

inert gas compounds cont.
PT *none*

inert gases
UF noble gases
rare gases
NT argon
helium
krypton
neon
radon
xenon
DI January 1973

inertial navigation
BT navigation
TT navigation
RT inertial systems
CC B7630 B7650 C3360L
DI January 1973

inertial systems
RT inertial navigation
CC C3200 C3300
DI January 1973

inference processes
USE artificial intelligence

infinite impulse response filters
USE digital filters

infinite series
USE series (mathematics)

inflammability
USE combustion

information analysis
US citation analysis
NT abstracting
cataloguing
classification
indexing
vocabulary
BT information science
TT computer applications
RT information centres
information retrieval
information services
information storage
information use
CC C7240
DI July 1973
PT information science

information analysis centers
USE information centres

information analysis centres
USE information centres

information centers
USE information centres

information centres
UF information analysis centers
information analysis centres
information centers
BT information science
TT computer applications
RT information analysis
information services
libraries
CC C7210
DI January 1973

information dissemination
UF dissemination of information
SDI
selective dissemination of information
BT information science
TT computer applications
CC C7220
DI January 1973

information needs
BT information science
TT computer applications
RT information services
information use

information needs cont.
libraries
CC C7220
DI January 1985
PT information use

information Network System
USE ISDN

information retrieval
UF document retrieval
online literature searching
retrieval, information
BT information science
TT computer applications
RT bibliographic systems
information analysis
information storage
query languages
records management
CC C7250
DI January 1973

information retrieval languages
USE query languages

information retrieval system evaluation
BT information retrieval systems
TT computer applications
CC C7250
DI January 1973

information retrieval systems
UF information storage systems
NT bibliographic systems
information retrieval system evaluation
BT information science
TT computer applications
CC C7250
DI January 1973

information science
UF documentation
librarianship
library science
NT information analysis
information centres
information dissemination
information needs
information retrieval
information retrieval systems
information services
information storage
information use
vocabulary
BT computer applications
TT computer applications
RT language translation
libraries
microforms
publishing
text editing
CC C7200
DI January 1973

information services
UF databases
services, information
technical information centres
BT information science
TT computer applications
RT information analysis
information centres
information needs
libraries
teletext
viewdata
CC C7210

information storage
BT information science
TT computer applications
RT bibliographic systems
information analysis
information retrieval
records management
CC C7250
DI July 1973

Figura 5.2.

A questão sendo encarada desse modo, deverão ser considerados dois grandes grupos de

termos
- → simples → uma só palavra
- → compostos → várias palavras
- → frases

Como seu nome indica, os termos simples são os formados por uma só palavra e os termos compostos são os formados por mais de uma palavra. A este segundo caso pertencem os substantivos nominais e os adjetivados, as frases preposicionais, as expressões adverbiais e as frases mistas:

termos compostos
- substantivos nominais
- substantivos adjetivados
- frases preposicionais
- expressões adverbiais
- frases mistas

Em continuação, serão estudados esses tipos de termos separadamente.

TERMOS SIMPLES COMPONENTES DE UM TESAURO

Aceitou-se internacionalmente que os termos que integram um tesauro devem ser substantivos, salvo em alguns casos muito particulares, como quando se trata de certas expressões adverbiais, que sejam absolutamente necessárias, no campo do conhecimento abordado pelo tesauro. Por exemplo, "altamente corrosivo" ou "muito perigoso".

Por conseguinte, os

termos simples

serão
- → substantivos
- → adjetivos substantivados
- → verbos substantivados

Quando esses nomes são conhecidos por suas siglas ou de uma maneira abreviada, deverão ser tratados como sinônimos, escolhendo-se a forma oficial como descritor, salvo no caso de aquela ser muito conhecida como tal.

Exemplo:

PAÍSES BAIXOS

UP HOLANDA
CONSELHO SUPERIOR DE PESQUISAS CIENTÍFICAS
UP CSPC

Pelo contrário

UNESCO
UP ORGANIZAÇÃO DAS NAÇÕES UNIDAS PARA A EDUCAÇÃO, A CIÊNCIA E A CULTURA:
BRASIL
UP Estados Unidos do Brasil.

OS TERMOS COMPOSTOS

Entre os termos compostos, devem considerar-se as expressões

- substantivadas ou nominais
- adjetivais
- preposicionais
- adverbiais
- mistas

As expressões nominais ou substantivadas são compostas por dois substantivos.

Exemplo:

GERONTOPSIQUIATRIA

ANCIÃO + PSIQUIATRIA

TERMÔMETRO

TEMPERATURA + MEDIDA

Pode ocorrer que as necessidades de uso do tesauro obriguem a decompor essa espécie de termos. Pelo contrário, na maioria dos casos, usam-se como um só termo, com caráter de

descritor ou preferencial.

As **expressões adjetivais** são as que constam de um substantivo e de um adjetivo modificador.

Exemplo:

APARELHOS PORTÁTEIS
FORMA RETANGULAR
FICHA TOPOGRÁFICA
FORMIGÃO REFORÇADO
MULHERES TRABALHADORAS

NOTA: Nos capítulos seguintes, estudar-se-á o seu tratamento na construção e no uso dos tesauros. O mesmo se diz para as frases proposicionais ou adverbiais, os sinônimos, os quase sinônimos etc.

Também devem-se considerar, neste capítulo, as frases adjetivais que possuam um sentido figurado:

Exemplo:

MESA-REDONDA

OURO CINZA

LITERATURA CINZA

As **expressões** ou **frases preposicionais** são aquelas compostas de um substantivo modificado de forma que se torne necessário utilizar uma preposição.

Exemplo:

FONTES DE INFORMAÇÃO

FILOSOFIA DA EDUCAÇÃO

SISTEMA EM LINHA

PROGRAMA PARA COMPUTADOR

Também neste caso será necessário considerar as expressões com sentido figurado.

Exemplo:

SOM À MEIA VOZ

OVOS DE OURO

Entre expressões adverbiais, deve-se considerar como

Exemplo:

SUBSTÂNCIA ALTAMENTE CORROSIVA
PRODUTO MUITO VENENOSO

Finalmente, consideremos alguns exemplos de termos compostos por ***expressões mistas***, adjetivais e proposicionais ao mesmo tempo.

Exemplo:

CADEIAS COM ANÉIS SOLDADOS

RECUPERAÇÃO DA INFORMAÇÃO CORRENTE

Consideremos estes outros

Exemplos:

CADEIAS SOLDADAS COM ANÉIS

RECUPERAÇÃO CORRENTE DA INFORMAÇÃO

que possuem significados diferentes, o que deverá ser levado em consideração no momento de se construir um tesauro.

A UTILIZAÇÃO DA FORMA SINGULAR E PLURAL

Prosseguindo, estudaremos o uso da forma singular ou plural dos substantivos que figuram nos tesauros, tanto como descritores ou como termos equivalentes. Quatro situações diferentes podem se apresentar:

1. As formas singular e plural possuem o mesmo significado.

2. As formas singular e plural possuem significados diferentes.
3. As formas singular e plural possuem o mesmo significado e podem ser usadas indistintamente.

4. As formas singular e plural são escritas com ortografia diferente.

Como norma geral, sempre que for possível, deve-se preferir a forma singular à plural. Contudo, em determinados casos, é mais prático escolher a forma plural, para dar um sentido maior aos termos e evitar ambigüidades e confusões. Vejamos, pois, quando se pode usar uma ou outra forma.

1. As formas singular e plural possuem o mesmo significado.

Distinguem-se entre

1.1 Substantivos

- quantificáveis;

onde se deve considerar as palavras que podem responder à pergunta:

- quantos?

e outras palavras que podem responder à pergunta:

- quanto?

Exemplo:

quantos ANIMAIS	há
quantos LIVROS	há
quantos PARTIDOS POLÍTICOS	há

Estes termos são usados no plural.

Se nos referimos, agora, a

Exemplo:

quanta FARINHA	há
quanta MICA (mineral)	há
quanta NEVE	há

Estes termos serão usados no singular.

1.2. No caso de tratar-se de termos que fazem referência a um todo e suas partes, as que se encontram em número diferente dentro do todo se empregam no singular ou plural, conforme esse número.

Exemplo:

As partes do corpo humano:

CABEÇA

BOCA

MÃOS

OLHOS

SISTEMA RESPIRATÓRIO

SISTEMA NERVOSO

NERVOS

Uma árvore:

TRONCO

GALHOS

RAÍZES

1.3 Os substantivos usados em sentido geral, como um conjunto de elementos, incluem-se no plural.

Exemplo:

AÇOS
↓
CIMENTOS
SISTEMAS DE CLASSIFICAÇÃO
↓
FONTES DE INFORMAÇÃO

1.4 Quando um substantivo faz referência a uma entidade abstrata, emprega-se o singular.

Exemplo:

AUTORIDADE

SENSIBILIDADE

GOVERNO

1.5 Também quando se trata de um substantivo que faz referência aos fenômenos naturais, utiliza-se o singular.

Exemplo:

ECLIPSE

TERREMOTO

CHUVA

DISTÂNCIA SOCIAL

1.6 Emprega-se, igualmente, o singular nos seguintes casos:

1.6.1 Propriedades

Exemplo:

SOLUBILIDADE

EVAPORAÇÃO

TOXIDADE

1.6.2 Atividades

Exemplos:

INDEXAÇÃO

CATALOGAÇÃO

IMIGRAÇÃO

1.6.3 Disciplinas ou ciências

Exemplo:

QUÍMICA

HISTÓRIA

BIBLIOTECONOMIA

1.6.4 Crenças

Exemplo:

CATOLICISMO

BUDISMO

SOCIALISMO

1.6.5 No caso de um termo que represente um conceito onde se considerem vári
será utilizado o plural.

Exemplo:

FICHAS PERFURADAS

CIÊNCIAS DOCUMENTÁRIAS

REAÇÕES QUÍMICAS

CONJUNTOS

2. As formas singular e plural possuem significados diferentes.

Exemplo:

HUMANIDADE
HUMANIDADES

MEMÓRIA
MEMÓRIAS

ATA
ATAS

Aqui se utilizarão ambas as formas, acrescentando-se uma nota esclarecedora que determine o seu significado e o seu uso.

3. As formas singular e plural possuem o mesmo significado e podem ser utilizadas indistintamente.

Exemplo:

TESOURA # CALÇA

TESOURAS # CALÇAS

Nestas circustâncias, deve-se incluir a forma singular, com uma nota de relação de equivalência para o plural.

Exemplo:

TESOURA

UP TESOURAS

TESOURAS

USE TESOURAS

4. As formas singular e plural são escritas com ortografia diferente.

Em espanhol, é muito difícil encontrar-se termos deste tipo. Ao contrário, em inglês, há vários exemplos: man, men; child, children. Deve-se empregar o singular, com uma nota de equivalência para o plural.

Exemplo:

MAN

UF MEN

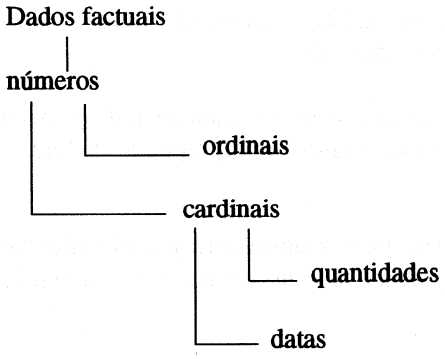
O TRATAMENTO DOS NÚMEROS

Na maioria das obras que tratam da estrutura e construção de um tesauro, não se abordam os números, assim considerados no sentido geral, dando-se mais ênfase ao assunto. Considera-se que são dados auxiliares e não se lhes atribui um posto, nem entre os descritores, nem entre os termos equivalentes. Nas publicações a que nos temos referido, *Thesaurus Construction* e a norma UNE 50-106, não se faz menção a números.

Contudo, deve-se considerar que estes são componentes dos chamados, também em sentido geral, dados factuais, e se compõem de cifras que indicam, por um lado, datas, e, por outro, quantidades ou números de ordem.

Pode-se compreender facilmente que, em uma base de dados sobre história, ou sobre economia, ou sobre propriedades físicas, as cifras compõem um elemento importante que deverá ser considerado no tesauro. Assim, pois, a elas serão dedicados alguns parágrafos, para explicar em que grau e de que forma entram na formação de parte da estrutura de um tesauro.

No esquema seguinte, separam-se os dados factuais em suas classes, considerados do ponto de vista que aqui interessa:



Vejamos que tratamento pode receber cada um deles.

Os números ordinais: primeiro, segundo etc. consideram-se descritores, embora sendo adjetivos, pois não são suscetíveis de substantivação. No caso dos termos, primeiro, PRIMER, este último substitui-se pela outra forma de expressão. O mesmo se faz com os restantes números que apresentam duas formas. Se for considerado muito necessário, pode-se utilizar a fórmula da relação equivalente.

Exemplo:

PRIMEIRO

UP PRIMER

Entre os ***números cardinais***, consideremos em primeiro lugar as ***quantidades***, que, como regra geral, serão escritas com letras e serão tomadas como termos preferidos _ descritores. Se tiverem de ser escritas com números, os algarismos relativos ao milhar são separados dos restantes por um ponto.

Exemplo:

1.492

Desse modo, diferenciam-se das datas, nas quais se omite o ponto.

Exemplo:

1492

Esta é uma norma aceita internacionalmente, que deve ser adotada para evitar confusões e ambigüidades.

As ***datas***, que serão também descritores, serão escritas com algarismos, embora, no caso de expressões tais como (os) anos 50, o artigo seja suprimido.

Para indicar datas exatas, com dias e meses, ou para períodos relativos a determinados séculos, podem-se utilizar algumas das formas usuais, como, por exemplo, as incluídas na Classificação Decimal Universal.

Em qualquer circunstância, quando for preciso contar com um tesauro que contenha dados factuais, cada um conhece suas necessidades e, portanto, a forma em que deve estruturá-los para lhes dar entrada no tesauro.

Com estes dados e situações estudados, foi feita uma revisão nos elementos constituintes de um tesauro e na forma na qual devem ser incluídos para formar o vocabulário principal. A partir daqui, pode-se falar das instruções para a sua confecção e uso.

TEMAS PARA ESTUDO

1. Diga quando considera que um termo é sinônimo de outro em um tesauro.

2. Descreva uma relação associativa entre termos e dê um exemplo.

3. Acredita ser importante relacionar os termos hierarquicamente?

4. Qual é o elemento estrutural, principal, de um tesauro?

5. Cite um caso no qual incluiria o termo no plural como descritor.

6. Dê sua opinião sobre a utilidade dos termos equivalentes.

7. De que outra maneira poderia chamar o "termo não-preferido"?

8. Em sua opinião, quantas classes de palavras-chave citaria?

9. Que objetivo se pretende alcançar, ao tomar as frases preposicionais como descritores?

10. Diga o que considera do uso da relação "termo específico".

11. Acredita que é mais útil usar somente termos isolados, simples, ou termos compostos como descritores em um tesauro?

12. O que faria com uma data?

13. Quando incluiria o nome de um país em um tesauro?

14. Como tratará um adjetivo?

15. Em que caso incluiria como termo equivalente "Alemanha Federal"?

16. Onde incluiria os verbos?

17. Diga o que entende por identificador.

18. Quando usaria um indicador de função?

19. Cite uma norma estrangeira que se ocupe destas questões.

REFERÊNCIAS BIBLIOGRÁFICAS

AITCHISON, J.: "Indexing languages: classification schemes and thesauri"; in Handbook of special librarianship and information work, Anthony, L.J. (ed.). London, ASLIB, 1982, Chapter 10, 207-261.

AITCHISON, J.: "Bliss and the thesaurus: the Bibliographic Classification of H.E. Blis as a source of thesaurus terms and structure". Paper Submitted to the International Conference on Ranganathan's Philosophy, New Delhi, India, November, 1985. Subsequently published as: "A classification as a source for a thesaurus: the Bibliographic classification of H.E. Bliss as a source of thesaurus terms and structure", J. of Document., *42*, 3, Sept.. 1986, 160-181.

GESELLSCHAFT FÜR INFORMATION AND DOKUMENTATION (GID): "Thesaurus guide. Analytical directory of selected vocabularies for information retrieval"; Prepared for the Commission of the European Communities by GID, Amsterdam, North Holland, 1985.

* GILCHRIST, A.: "The Thesaurus in retrieval", London, ASLIB, 1971.

* "Inspec Thesaurus"; London, Institute of Electrical Engineers, 1987.

LOEBEN, M., RUNGE, W.: "Sprachbarrieren bei der naturwissenschaftlich-technischen Fachkommunikation. Ergenbnisse einer Befragung"; Nach, f. Dokum., *36* 3, 1985, 127-136.

LOPE DE SOSA-AGA, G. : "Sobre la automatización del Thesaurus", en Actas del I Congreso Iberoamericano de Informática y Documentacion"; CREI, Medellín (Colombia), 1985, 1, 429-435.

PEREZ ALVAREZ-OSSORIO, J.R.: "El análisis documental"; en Introducción a la Información y Documentación Científica, Madrid, Alhambra, 1988, 43-56.

RANGANATHAN, S.R.: "The Colon Classification"; Rutger's, The State University. Graduate School of Library Science, 1965.

SIEVERT, M.E., BOYCE, B.R.: "Hedge trimming and the resurrection of the controlled vocabulary on line"; Online Review, 7, 6, 1983, 489-494.

6

Construção de Tesauros

A questão da construção de um tesauro preocupou e tem preocupado a todos os profissionais das ciências documentárias. Em geral, vai nela o prestígio pessoal, por se tratar da elaboração de um instrumento, uma ferramenta, como se diz comumente, com a qual é possível ordenar os dados contidos em qualquer tipo de documento, de modo que, posteriormente, torne-se possível encontrá-los. Há algum tempo, colocava-se ênfase na localização correta e na classificação do próprio documento: por isso, os tesauros eram muito confusos, trazendo variadas instruções, remetendo de um tema a outro, com suas diversas formas de representação. Atualmente, trabalha-se colocando o interesse na recuperação da informação. Agora, o que interessa é o assunto, o conceito, e o termo com que se expressa. Os tesauros devem, igualmente, ser elaborados minuciosamente, com instruções e envios e reenvios de um assunto a outro. Contudo, sua apresentação ou sua edição costuma ser mais simples. Muitas vezes, uma simples lista alfabética dos termos, com todas as suas relações de equivalência, hierárquicas e associativas, costuma ser suficiente. O uso dos computadores também simplifica o trabalho, pois permite misturar os termos com os quais se opera e ralacioná-los, ordená-los ou substituí-los facilmente. A difusão dos computadores pessoais (N. T. do inglês *Personal Computers* - PC) quase obrigou a excluir para sempre o trabalho manual.

De todas as formas, quer se utilize um método ou outro, não se pode confiar a responsabilidade intelectual às máquinas. Do mesmo modo, para mandar fazer algo, quer seja por pessoa, quer seja automaticamente, deve-se saber fazê-lo. Iremos tratar, aqui, da construção dos tesauros.

Em honra à verdade, devemos dizer que existem outras publicações onde se explica também como se deve construir um tesauro e que são também recomendáveis. Estas são mencionadas nas *Obras de Referência*. Além do mais, em muitos tesauros, em suas instruções preliminares, cita-se o procedimento empregado pelo grupo que o elaborou, o que costuma ser altamente instrutivo. Recentemente, chegou às minhas mãos o *Tesauro sobre Urbanismo*, elaborado pelo Instituto de Información y Documentación en Ciencias Sociales y Humanidades (ISOC), em 1988, onde figura uma introdução desse tipo.

Uma vez que estudamos quais são os elementos fundamentais de um tesauro e como se relacionam, iremos tratar, agora, dos assuntos relativos a todo o processo que se deve seguir para elaborar um tesauro. Começarmos por considerar as situações nas quais pode ser necessário confeccionar um tesauro novo e a forma de elaborarem-se, com a base de conhecimento, os termos que se deve empregar. Após, será abordada a maneira em que estes

devem se relacionar, como devem se apresentar, e, por último, abordaremos a possível publicação do tesauro já incluído. Em capítulo separado, considerar-se-á a forma de uso dos tesauros.

QUANDO FAZ FALTA UM TESAURO NOVO

É a primeira pergunta que um documentalista especializado deve se formular no momento de realizar a indexação de uma série de documentos relativos a seu assunto de trabalho.

Nesta fase, apresenta-se outra questão importante e que é tratada superficialmente na maioria dos trabalhos. Referimo-nos ao dilema entre

- tesauro de indexação
    - linguagem de indexação;
- tesauro de recuperação da informação
    - linguagem de recuperação da informação.

Em princípio e quase por definição, um tesauro é necessário para indexar os documentos. Conseqüentemente, é necessário o mesmo tesauro para recuperar a informação. Pois bem, em muitos centros ou serviços de documentação, usam-se bases de dados e redes de informação externas, que possuem seus tesauros oficiais. Contudo, não precisam manejar tesauros complicados e volumosos porque seu campo de atuação abrange somente uma parte reduzida do assunto geral, ou porque compreende, também, outros assuntos marginais. Neste caso, constrói-se, adapta-se um tesauro à medida das próprias necessidades. Esse será o tesauro de recuperação da informação, que será elaborado conforme as mesmas normas seguidas quando se trata de um tesauro dedicado à indexação. As diferenças são estabelecidas no seu uso, que não coincidirá em um caso ou outro.

Portanto, neste capítulo, abordar-se-á a construção de tesauros em geral, e, no capítulo dedicado a seu uso, serão assinaladas as divergências de ambas as classes de tesauros.

Sempre que for necessário usar um tesauro em um serviço de informação, seja de que tipo for, a primeira coisa que se deverá fazer é saber se já existe algum para o campo de aplicação em questão. Para isso, deve-se consultar a extensa bibliografia sobre o assunto. Normalmente, os centros de documentação nacionais ou especializados e as bibliotecas especializadas dispõem deste tipo de publicações. Também se podem encontrar catálogos dedicados a este assunto. Na Espanha, recomenda-se dirigir-se ao Cindoc ou ao ISOC, conforme os campos do conhecimento com que se trabalhe, e à Reuniber ou Fuinca. Também pode ser aconselhável consultar as publicações da Aslib, no Reino Unido, principalmente o *Current Awareness Bulletin* e *On-line Notes*.

Se já existe um tesauro sobre nosso campo de trabalho, devemos provar até onde poderá ser utilizado por nós. É muito difícil que um tesauro elaborado por uma entidade ou grupo de profissionais especializados em determinado assunto possa ser utilizado por outro grupo diferente de especialistas nesse mesmo assunto. Várias razões podem ser apresentadas. O âmbito do assunto, as necessidades dos usuários, o número de consultas etc. Assim, pois, não se poderá evitar que se tenha de adaptá-lo às necessidades concretas de cada um.

Nessa circunstância, o tesauro existente servirá de guia e consulta para ajudar a estabelecer as relações entre os termos e sua classificação ou ordenação, se for necessário contar com uma seção sistemática ou uma representação gráfica.

Deve-se advertir que é muito difícil que não exista algum tesauro sobre o assunto em que se trabalha. Atualmente, todos os campos do conhecimento estão cobertos. Porém, sua adaptação a nosso caso particular implica partir do zero. Devemos nos conhecer para poder tomar emprestado o que se nos oferece.

Como conseqüência, deve-se concluir que se deverá agir sempre como se tratasse de construir um tesauro novo.

CONSIDERAÇÕES QUE DEVEM SER LEVADAS CONTA

Várias e variadas são as considerações que se devem levar em conta no momento de confeccionar um tesauro. Umas são do tipo externo: externo ao próprio sistema documentário, quer dizer, ao conjunto de documentos, e outras são do tipo interno, inerente aos próprios documentos e suas características.

Entre as

considerações do tipo interno,

deverão ser examinados:

- assunto principal;
- assuntos auxiliares;
- assuntos marginais;
- âmbito dos assuntos:
- grau de precisão;

- número de documentos;
- classes de documentos
- tipo de dados contidos nos documentos;
- tipo de documentos
  - primários
  - secundários;
- fontes de coleta
  - dos documentos;
- aumento progressivo no tempo;
- necessidades de atualização.

Pode ser muito útil construir um formulário como o da figura 6.1, onde se façam constar todos os dados para que resulte mais intuitivo tê-los presentes.

É conveniente pensar sempre nos assuntos auxiliares e marginais ao nosso assunto principal. Na realidade, é muito raro trabalhar com um único assunto como foco de toda atenção. Sempre há ramificações e conexões com outros assuntos, tratados de forma menos explícita. Por isso, deve-se levar em consideração o âmbito que os assuntos devem abranger. Compreende-se que o assunto principal será tratado em toda extensão, enquanto os assuntos complementares serão abordados em função de sua conexão com o primeiro. Como conseqüência, será importante considerar o grau de precisão com que se deve trabalhar. Será preciso pensar na exatidão com que se necessita indexar um assunto.

Por outro lado, o número de documentos a utilizar será um dado indicativo do tipo e o volume de termos que se deve empregar. Também deve-se pensar nas classes de documentos que se usam. Indexar livros não é o mesmo que indexar artigos de periódicos, atas de congressos e patentes, ou uma miscelânea de todos eles, que costuma ser o corrente. É fundamental considerar os tipos de dados de que se dispõe. Estes podem ser dados documentários ou factuais, quer dizer, palavras, números ou fórmulas. O tipo de tesauro também será diferente em um caso ou em outro. Disto resulta que também será determinante considerar se os documentos a serem tratados são primários ou secundários. Variará o grau de precisão dos termos em um caso ou em outro. Em uma base de dados bibliográficos, onde já existe um resumo, é mais fácil decidir quais são os termos utilizáveis, do que em um documento completo, onde se deve começar por fazer esse resumo, com a conseqüente diversidade de termos que podem parecer determinantes do assunto.

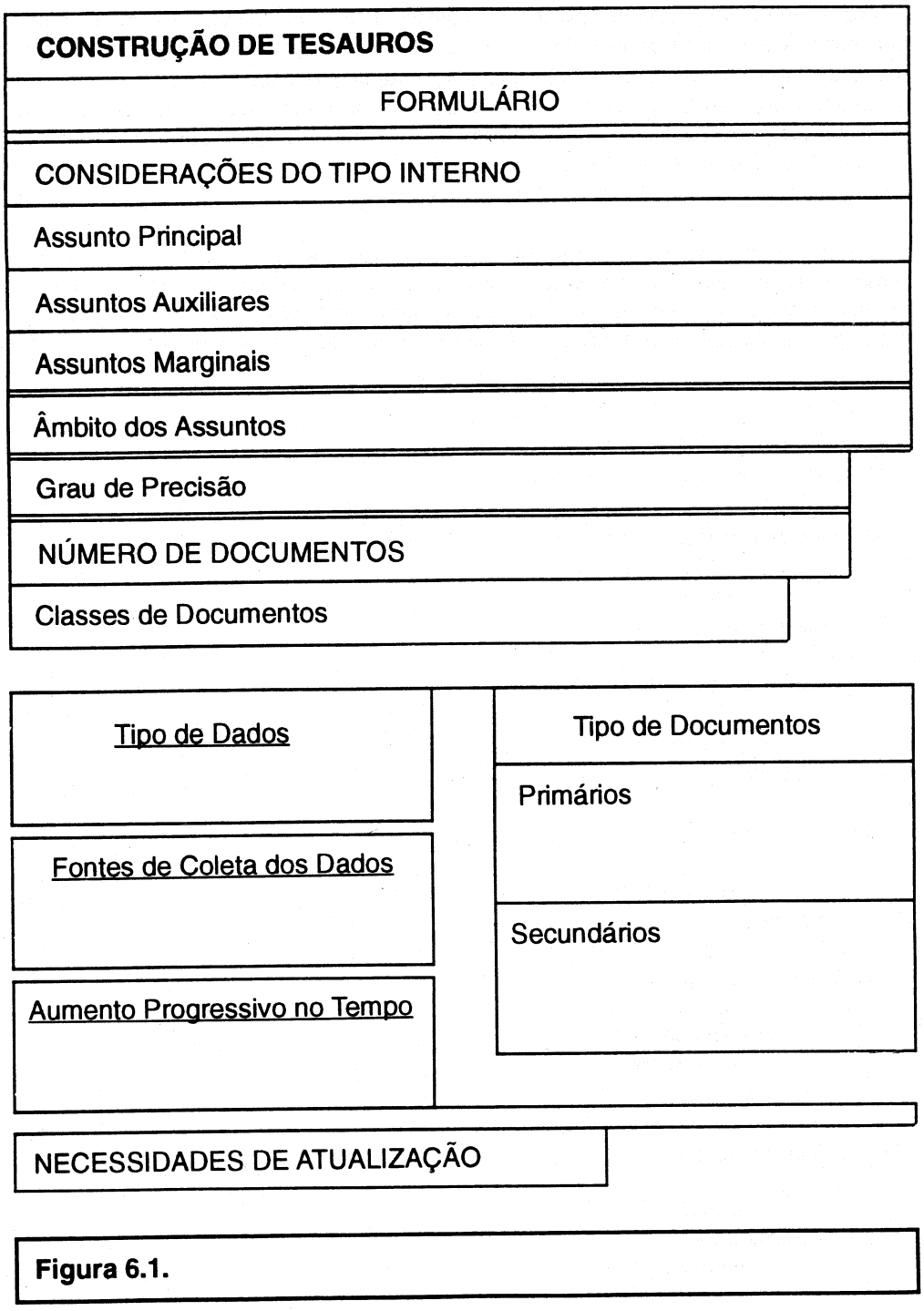

CONSTRUÇÃO DE TESAUROS
FORMULÁRIO
CONSIDERAÇÕES DO TIPO INTERNO
Assunto Principal
Assuntos Auxiliares
Assuntos Marginais
Âmbito dos Assuntos
Grau de Precisão
NÚMERO DE DOCUMENTOS
Classes de Documentos
Tipo de Dados
Fontes de Coleta dos Dados
Aumento Progressivo no Tempo
Tipo de Documentos
Primários
Secundários
NECESSIDADES DE ATUALIZAÇÃO

Figura 6.1.

Por último, é importante, igualmente, considerar a proveniência dos documentos. Suponhamos que se trate de dados bibliográficos, resumidos no centro de origem. Conforme seja este, dar-nos-á confiança ou não, em sua forma de trabalho. Também a origem dos documentos avulsos pode ser motivo de confiança, na razão da exatidão das palavras especializadas empregadas. Indicará, também, o país de onde procede e, portanto, a forma de usar um determinado idioma. Pensamos nos países hispano-americanos e outros de língua espanhola, por exemplo, o Norte de Marrocos.

A avaliação do aumento de documentos em períodos de tempo prefixados faz com que pensemos nas previsões que devem ser tomadas e nas decisões relativas a reter ou excluir os termos atuais ou desatualizados. Deve-se ser muito precavido com esta questão. O desenvolvimento das ciências e da tecnologia pode nos oferecer muitas surpresas.

Entre as

considerações do tipo externo

Deve-se ter em conta:

- As características do centro

  → interno
  → externo

- A classe do centro

  → privado
  → público

- O tipo do centro

  → especializado
  → geral

- A forma de trabalho

  → com o usuário
  → direta
  → indireta

- O sistema de trabalho

  → manual
  → automatizado
  → em linha

→ fora de linha

• O pessoal disponível
→ especializado
→ auxiliar

• As características dos usuários

• O tipo de consulta

→ sistema de perfis
→ disseminação seletiva da informação

→ informação retrospectiva

• O volume de consultas

• O grau de crescimento das consultas

• A previsão de crescimento do próprio centro

• O sistema de financiamento.

Também neste caso, deve-se confeccionar um formulário. Pode-se tomar como exemplo o da figura 6.2.

Todos estes dados não precisam de maior comentário. Cada um conhece sua situação e pode deduzir se necessita de um tesauro muito exaustivo ou mais reduzido; se os termos devem ser muito precisos ou mais gerais; como deverão ser as relações entre eles. Em cada caso, o responsável pela elaboração do tesauro saberá que conclusões deve tirar diante das respostas a todas essas perguntas e talvez a alguma outra não pensada neste capítulo.

Deve-se reconhecer que a situação ideal, a melhor, é aquela na qual um documentalista especializado neste campo do conhecimento deve construir um tesauro especializado neste campo do conhecimento, para um serviço de informação interno, de uma empresa privada; sobre um volume de documentos manuseáveis, para usuários conhecidos, que trabalhem com assuntos bem determinados, e onde se dispõe de sistemas automatizados. Poucas são as vezes em que se apresenta esta oportunidade! · ·

CONDIÇÕES GERAIS QUE UM TESAURO DEVE CUMPRIR

Diante de todos os dados obtidos das considerações expostas anteriormente, deve-se determinar as condições gerais que um tesauro deverá cumprir para as específicas necessidades de cada caso.

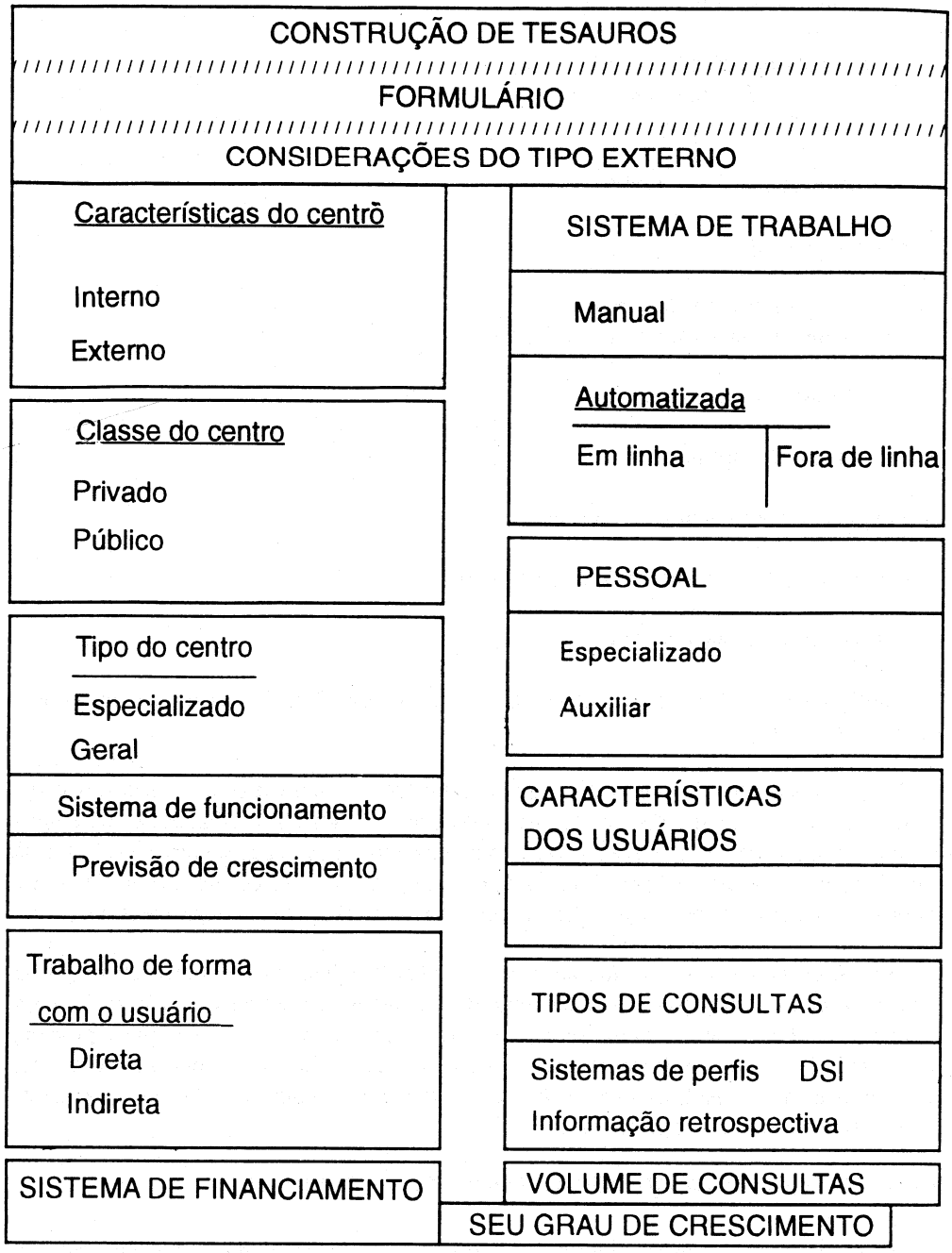

CONSTRUÇÃO DE TESAUROS
FORMULÁRIO
CONSIDERAÇÕES DO TIPO EXTERNO
Características do centrõ
Interno
Externo
Classe do centro
Privado
Público
Tipo do centro
Especializado
Geral
Sistema de funcionamento
Previsão de crescimento
Trabalho de forma
com o usuário
Direta
Indireta
SISTEMA DE FINANCIAMENTO
SISTEMA DE TRABALHO
Manual
Automatizada
Em linha
Fora de linha
PESSOAL
Especializado
Auxiliar
CARACTERÍSTICAS
DOS USUÁRIOS
TIPOS DE CONSULTAS
Sistemas de perfis DSI
Informação retrospectiva
VOLUME DE CONSULTAS
SEU GRAU DE CRESCIMENTO

Figura 6.2.

Assim, saberemos se faz falta um

tesauro

- geral e pluridisciplinar
  - ◦ inclusive, multilíngüe

- geral monodisciplinar
  - ◦ mono ou multilíngüe

- especializado
  - → bem determinado
  - → com áreas marginais

- exaustivo
  - → considerando a maior quantidade possível de termos;
  - → todas as relações possíveis;

- específico
  - → com um número reduzido de termos
  - → todas as relações possíveis;

- genérico
  - → com um número reduzido de termos
  - → poucas relações entre eles.

Não é muito comum a necessidade de se confeccionar tesauros gerais pluridisciplinares e mono ou multilíngües. Existem alguns, muito conhecidos, como o TEST-Thesaurus, Precis ou o Spines da Unesco. Este último foi construído na forma de macrotesauro. São tesauros construídos aproximadamente há uns 10 anos, planejados com um fim preconcebido, embora de uso constante. A tendência atual está voltada para os tesauros especiais e específicos para assuntos bem determinados e atuais. Aí temos o exemplo, entre muitos outros, do tesauro sobre urbanismo, confeccionado pelo ISOC, recentemente.

A APREENSÃO E FIXAÇÃO DO TERMO

Em um caso ou em outro, o ponto de partida deverá ser o conhecimento dos termos com que teremos de nos enfrentar e como devemos relacioná-los e apresentá-los por intermédio de algo que seja coerente, manuseável e útil. Devemos voltar a pensar nas condições que os tipos de termos devem cumprir, de onde devemos tirá-los etc. A lista seguinte pode nos servir de guia.

Questões que se deve considerar sobre o próprio

termo

- assunto principal
- assuntos auxiliares
- assuntos marginais
- classe do termo
  → descritor
  → equivalente
- tipo de relações
  → de equivalência
  → hierárquicas
  → associativas
- a escolha de termo simples ou composto
- compilação do termo
- freqüência com que aparece
- origem lingüística do termo
- idioma de origem
- ortografia
- formas de apresentação
- atualização.

Será muito útil confeccionar uma ficha, talvez em tamanho UNE A4, onde seja possível ter todos os dados à vista. Estes dados não variam no caso em que se trabalhe com computador. Sempre que for possível, é preferível trabalhar de forma automática. Atualmente, os computadores individuais (N. T. - do inglês, PC = Personal Computers) oferecem programas simples e econômicos, bastante aptos para trabalhar com os termos.

O primeiro passo a dar é

delimitar os assuntos de trabalho,

tanto o principal como os auxiliares ou marginais.

Assim, poderemos saber se um termo deverá ser considerado como principal (descritor), preferido, ou como termo secundário (equivalente), não-preferido. Um termo não será descritor somente por pertencer ao assunto principal de trabalho. As considerações sobre os sinônimos e quase-sinônimos pressupõem um passo posterior na confecção de um tesauro.

Aqui o que podemos argumentar é a relevância de um termo e a importância que merece. Por exemplo, um termo (palavras-chave) do assunto principal deverá ser desenvolvido em toda a sua extensão, enquanto outro, de um assunto marginal, não necessitará mais referência que a um assunto genérico ou específico, ou a uma indicação de relação associativa.

o passo seguinte consiste em

reunir os termos.

Para isto, todas as obras aconselham a consulta à maior quantidade possível de documentos de todo o tipo:

- dicionários,
- voacabulários especializados,
- enciclopédias,
- léxicos,
- sistemas de classificação,
- tesauros,
- livros,
- documentos avulsos,
- bases de dados,
- sistemas de perfis,
- consultas de usuários,
- partes de estratégias de busca de informação,
- consultas com especialistas.

Deve-se conseguir abranger o mais possível o âmbito de documentação onde possam aparecer os termos que nos interessam. Conforme vão aparecendo, deve-se ir anotando na ficha planejada para tal, tanto em papel como na memória do computador ou em disquetes.

Já se pode ir anotando ***a fonte, ou fontes, onde o termo foi encontrado, sua origem lingüística; suas formas de ortografia, a zona cultural onde é empregado*** etc. Quer dizer que, ao compilar os termos, realizamos uma série de etapas simultaneamente.

É muito aconselhável anotar a freqüência com que um termo se apresenta. Isto dará idéia de sua importância para decidir sobre a sua inclusão no tesauro e que categoria lhe deve ser atribuído, ainda que se deva advertir sobre a conveniência de não desvalorizar nenhuma palavra. O progresso da humanidade, assim considerada em sentido geral, pode mudar a importância e o uso de qualquer termo.

Outra questão sobre a qual sempre surgem perguntas refere-se à forma de poder determinar quando foram conseguidas todas as palavras necessárias para a confecção do tesauro. O momento em que, ao ler um novo documento, não apareça nenhuma palavra que não tenhamos anotado, é quando já dispomos de todas elas, relativas ao assunto tratado. Contudo, chegar a essa situação pode pressupor o emprego de tempo demasiado e dificultar o ritmo de trabalho. O normal é reunir um número razoável de termos e começar a trabalhar com eles, independentemente de que se continue recolhendo os termos que sigam aparecendo. Porém, qual é esse "número razoável"? Li, em algum livro, estar ele "entre quinhentos e mil". Na realidade, não se pode citar cifras. Tudo dependerá da quantidade total de termos com que se deve trabalhar, da abrangência do tesauro, do tamanho e da finalidade do serviço de informação onde será utilizado e de uma infinidade de outras condições particulares de cada caso.

O aconselhável é realizar as diferentes etapas simultaneamente. O trabalho torna-se mais natural, mais eficaz e menos monótono.

TERMOS SIMPLES OU TERMOS COMPOSTOS

Entre as operações que se deve realizar para incluir os termos em um tesauro, figura a tarefa de esclarecer se devem ser tomados termos simples ou termos compostos. O primeiro caso pressupõe a necessidade de se trabalhar com uma grande quantidade de termos, que, muitas vezes, possuem diversos significados. Isto implica ter de acrescentar notas descritivas e complicar o tesauro. Quando se usam termos compostos, reduz-se o tamanho do tesauro e se evitam confusões, porém é preciso saber como ordená-los para não perder informação.

Consideremos o exemplo

fontes de informação:

se separarmos essa frase em

fontes + informação,

deveremos considerar que

fontes → apresenta vários significados em diferentes assuntos;
informação → também pode apresentar várias interpretações.

pelo contrário, se dissermos
minério de carvão:

será melhor tomar
minério + carvão;

pois, assim, poderemos combinar
minério → com outros elementos ou compostos químicos.

Como conseqüência, deve-se dizer que o mais prático e útil será dispor de termos simples ou compostos, conforme a necessidade de uso do tesauro.

A escolha dos termos simples, como tais, não apresenta nenhuma dificuldade, pois serão tomados na forma como se apresentam nos documentos. Porém, no que se refere aos termos compostos, deve-se estudar um por um cuidadosamente. Em certos casos, um termo que se apresenta em forma composta deverá ser decomposto ou desdobrado nas partes que o formam, para obter-se maior clareza na elaboração do tesauro, ou maior confiabilidade em seu sucesso. Em outros casos, pelo contrário, será mais útil conservar o termo composto da forma como se emprega nos documentos.

Nos termos compostos, consideram-se duas partes:

- foco ou núcleo,
- modificador.

Exemplo:
Centro de documentação
centro ... núcleo,
documentação ... modificador.

Vejamos, pois, em que casos é possível
fatorar os termos

1. A norma UNE diz que se deve decompor termos como
termômetro em

TEMPERATURA + MEDIÇÃO + INSTRUMENTO,

"embora não seja recomendável". Realmente, pode ser que seja melhor usar o termo em conjunto: TERMÔMETRO, já que temos um significado muito marcante, inconfundível e cunhado pelo uso.

Este será um caso de ***fatoração semântica.***

2. Casos de fatoração sintática

2.1. Quando permitem uma análise morfológica, de forma que cada componente possa ser aceito como termo independente.

Exemplo:

Fiação de algodão
FIAÇÃO + ALGODÃO

NOTA I: Em todos esses casos, subentende-se que pode ser muito útil incluir as duas formas. Deverão ser consideradas como sinônimos e deverão levar a notação correspondente: USE ou UP, conforme os casos.

2.2. Quando o modificador se refere a uma parte do núcleo.

Exemplo:

Rodas de bicicletas
RODAS + BICICLETAS

Motores de aviação
MOTORES + AVIAÇÃO

2.3. Quando uma ação transitiva do núcleo recai no modificador.

Exemplo:

Automação de Escritórios
AUTOMAÇÃO + ESCRITÓRIOS

Classificação de Documentos
CLASSIFICAÇÃO + DOCUMENTOS

NOTA II: Como observação complementar, pode-se considerar que, nestes casos, parece que mudam as funções das partes do termo.

Assim:

OFICINAS
DOCUMENTOS — núcleos
modificados por
AUTOMAÇÃO
CLASSIFICAÇÃO

2.4. Quando uma ação intransitiva do núcleo recai no modificador.

Exemplo:

Migração de aves
MIGRAÇÃO + AVES

Restauração de pergaminhos
RESTAURAÇÃO + PERGAMINHOS

Aqui também se poderia aplicar a NOTA II acrescida ao caso anterior.

NOTA III: Neste caso, como em todos os relativos à construção de tesauros, aconselha-se consultar a norma UNE, caso se deseje informação complementar.

Prosseguindo, expõem-se os casos em que se ***usarão termos compostos.***

1. Os nomes próprios.

Exemplo:

LASSO DE LA VEGA

2. Os termos que significam nomes próprios.

Exemplo:

EXÉRCITO DA SALVAÇÃO
COMPLEXO DE ÉDIPO
FICHÁRIO KARDEX

3. Os termos nos quais o modificador perdeu seu significado.

Exemplo:

COLÉGIO INVISÍVEL

4. Os termos com sentido figurado.

   Exemplo:

   ECONOMIA SUBTERRÂNEA
   LITERATURA CINZENTA

5. Os termos que não se pode voltar a utilizar separadamente, sem acrescentar uma explicação.

   Exemplo:

   LABORATÓRIO DE IDIOMAS
   ENERGIAS DE COMPENSAÇÃO

6. Os termos nos quais o modificador não corresponde a uma espécie ou tipo do núcleo.

   Exemplo:

   PANTERA ROSA
   FLORES DE PAPEL
   OURO NEGRO

Quando os termos em sua forma simples ou composta já foram adotados e foram feitas as separações oportunas, com suas notações correspondentes, será o momento de avançar um passo a mais e começar a estabelecer as relações entre os termos. Na prática, estas duas tarefas podem ser simultâneas.

ESTABELECIMENTO DAS RELAÇÕES ENTRE TERMOS

Aqui se estudará a maneira como se deve estabelecer as relações de equivalência, hierárquicas e associativas entre os termos, para obter a correlação entre eles e poder configurar um tesauro apropriado à medida das necessidades de cada local onde deva ser utilizado.

Começaremos pelas

Relações de equivalência.

No capítulo anterior, falamos deste tipo de relações, fazendo referência à estrutura do tesauro. Agora, iremos estudá-las, com o objetivo de conhecer seus tipos e que termos deverão figurar como descritores e quais deverão figurar como equivalentes.

Sabemos que se deve considerar dois tipos de

termos

- sinônimos,
- quase sinônimos.

Vejamos que formas de sinonímia se apresentam normalmente. Em sua maioria, foram retiradas da obra *Thesaurus Construction* e da norma UNE 50-106.

1. Nomes populares e científicos.

Exemplo:
aranhas = aracnídeos

2. Nomes comuns e científicos e marcas comerciais.

Exemplo:
Citrovit = vitamina C
PVC = policloreto de vinil

3. Nomes instituídos cientificamente e os devidos à linguagem popular.

Exemplo:

equipamentos de alta fidelidade = equipamentos HI-FI
aparelhos de televisão = TV

4. Palavras com ortografia diferente.

Exemplo:

PSICOLOGIA = SICOLOGIA
RUMÂNIA = ROMÂNIA

5. Termos de origem lingüística diferente.

Exemplos:

trabalho doméstico = trabalho caseiro
geomagnetismo = magnetismo terrestre

6. Termos originados de grupos étnicos diferentes, mas que usam uma língua comum.

N.T.: É o caso típico ocorrido entre Portugal e outros países de língua portuguesa. Os exemplos são numerosos, dentre os quais, são citados os que se seguem.

Exemplos:

rapariga (em Portugal) = moça (no Brasil)
carruagem (em Portugal) = carro (no Brasil)

7. Termos atuais e outros de menor uso.

Exemplo:

ensino à distância = ensino por correspondência
países em vias de desenvolvimento = países subdesenvolvidos

8. Abreviaturas e nomes complexos.

Exemplo:

FID = Federação Internacional para a Informação e a Documentação
CDU = Classificação Decimal Universal

NOTA: Alguns estão tão introduzidos na linguagem comum que são usados como tal.

Exemplo:

laser
raios x

Como norma geral, deve-se tomar como descritores – termos preferidos – os termos científicos, os mais usados no país onde o tesauro será utilizado, e os mais atuais. De toda a maneira, as necessidades do centro de documentação ou serviço de informação e a utilização que o tesauro tenha serão os determinantes para a escolha dos descritores. Os termos tomados como sinônimos deverão ser incluídos com sua notação correspondente.

O mesmo se aplica para o caso dos termos quase sinônimos.

Em continuação, citam-se alguns ***casos de quase sinonímia***.

1. Termos de significado muito semelhante, porém onde se pode estabelecer algumas diferenças.

Exemplo:

áreas urbanas = cidades
gênios = superdotados
congresso = simpósio

2. Termos que podem ser agregados sob outro com maior abrangência de significado.

Exemplo:

PROPRIEDADES TERMODINÂMICAS
UP Entalpia
Entropia
Calor de absorção

MICROFORMAS
UP Microfichas
Microfilme

Geralmente se recorre ao uso de termos quase sinônimos quando se deseja reduzir o tamanho de um tesauro e o número de termos que se deve utilizar. Deve-se ser muito cauteloso no momento de escolher estes termos, levando em consideração seu significado. O exemplo seguinte pode ser muito ilustrativo:

clarabóia
= cobertura de cristal
casa de vegetação

Trata-se, contudo, de coberturas diferentes, com usos distintos; portanto, não poderão ser consideradas quase sinônimos.

Nos parágrafos seguintes, serão abordadas as

relações hierárquicas,

Com o objetivo de poder discernir quais são os termos que se consideram genéricos e quais são os que se consideram específicos, outros detalhes sobre este assunto foram expostos no Capítulo 5.

A melhor forma de trabalhar no sentido de estabelecer as relações hierárquicas é ir agrupando os termos por famílias conforme seu significado. É recomendável começar por grupos de famílias pequenos, com poucos termos, e logo aumentando a abrangência, até constituir grandes famílias. Estas podem chegar a formar classes próprias de uma seção sistemática do tesauro.

Como exemplo:

1. Podemos começar por ÁRVORE e reunir todos os termos que fazem referência às suas partes : FOLHAS, RAMOS, RAÍZES, FLORES ...

2. Após, ordenaremos as fichas segundo o seu significado.

ÁRVORE
- RAMOS
- FOLHAS
- FLORES

RAÍZES

Com isto, já conseguimos uma primeira ordenação em hierarquias.

3. Depois, passaremos a ordenar os termos que se referem a ARBUSTOS, PLANTAS ... Conseguiremos outra família de maior abrangência.

BOSQUE
- ÁRVORE
- ARBUSTO
- PLANTA
- .
- .
- .
- .

que logo poderá integrar uma outra, ainda mais ampla, e assim sucessivamente. Depois, estabelecer-se-ão as relações recíprocas entre os termos, como é indicado no capítulo anterior.

As figuras 6.3 e figura 6.4, retiradas da obra *Catering Thesaurus*, mostram duas etapas das sucessivas distribuições dos termos em categorias, uma mais detalhada que a outra.

Se necessário, pode-se estabelecer diferentes níveis de hierarquia até o termo genérico ou até o especificado, que se pode assinalar com subíndices TG1, TG2, TE1, TE2, ou com as notações TGG, TGE, TEG, TEE.

Na figura 6.5, mostra-se um exemplo retirado da obra *CAB Thesaurus*, onde aparecem níveis distintos de hierarquização.

Broad subject categories
Catering and related fields

General concepts

Geographical areas
Research, design and testing
Standardization
Management
Organizations

Peripheral subjects

Communication and information
  Computers
Philosophy
Mathematical sciences
Sciences
  Physical sciences
  Earth sciences
  Biology
Agriculture. Food production
Technology and Engineering
  including Construction
Biomedical sciences. Medicine
  Nutrition. Dietetics
  Environmental health
    Food hygiene
  Health services
    Hospitals
Psychology
Education
Social sciences
  Sociology
  Social problems and Welfare
  Politics and Government
    Armed forces
  Law
  Humanities. History. Religion
Arts, literature, leisure

Closely related subjects

Food technology
Household technologies
  Cleaning technology
  Corporate housekeeping
    Hotel management
  Home economics

Catering

Figura 6.3.

Allocation of terms to broad subject categories
Catering and related fields

General concepts

...
Management
Catering management
Catering managers
Financial management
Office management
Personnel management
Planning
Purchasing
Recruitment
Work measurement

Peripheral subjects

...
Psychology
Attitude
Conflict
Interpersonal relations
Social interaction
Stress

Education

Colleges of further education
Colleges of higher education
Curricula
Educational courses
In-service training
Polytechnics
Practical training
Schools
Teachers
Teaching
Training
Universities

etc.

Figura 6.4.

Multi-level alphabetical thesaurus
CAB thesaurus

DOMESTIC PRODUCTION (C)
rt self sufficiency

domestic science
USE home economics

DOMESTIC TRADE
uf trade domestic
BT1 trade

DOMESTICATED BIRDS
BT1 birds
NT1 pigeons
NT1 poultry
NT2 capons
NT2 chickens
NT3 bantams
NT3 broilers
NT3 chicks
NT4 day old chicks
NT3 cockerels
NT3 pullets
NT2 drakes
NT2 ducks
NT3 ducklings
NT3 wild ducks
NT2 geese
NT3 ganders
NT3 goslings
NT2 guinea fowl
NT2 hens
NT2 turkeys
NT3 poults
NT3 toms
rt domestic animals
rt poultry

DOMESTICATION
uf speciation
rt crops
rt domestic animals
rt evolution

DOMIATI CHEESE
BT1 cheeses
BT2 milk products
BT3 dairy products
BT4 products
BT3 livestock products

DOMINANCE
NT1 apical dominance
NT1 codominance
NT1 epistatic deviation
NT1 overdominance
NT1 semidominance
rt alleles
rt dominant genes
rt dominant lethals
rt dominant species
rt ecological succession
rt genetics
rt overdominant genes

DOMINANT GENES
BT1 genes
rt dominance

DOMINANT LETHALS
uf lethals dominant
BT1 lethals
BT2 genes
BT2 genetic defects
BT3 abnormalities
BT3 defects
rt dominance

DOMINANT SPECIES
BT1 species
rt climax communities
rt competitive ability
rt dominance
rt mutispecies fisheries
rt plant interaction
rt plant succession

DOMINICA
BT1 caribbean
BT2 america
rt acp
rt caribbean community
rt commonwealth of nations
rt windward islands

DOMINICAN REPUBLIC
BT1 caribbean
BT2 america
BT1 central america
rt threshold countries

DON
BT1 goat breeds
BT2 breeds
BT1 horse breeds

DONAX
BT1 mollusca

DONGOLA
BT1 horse breeds
BT2 breeds

doob
USE cynodon dactyion

DOOR TO DOOR SALES
BT1 marketing techniques
BT2 marketing
BT2 techniques
BT1 retail marketing
BT2 marketing channels
BT3 marketing

DOORS
BT1 buildings
rt gates

DOPA
uf 3 4-dihydroxyphenylalanine
BT1 catecholamines
BT2 ammonium compounds
BT1 neurotransmitters
BT1 phenylalanine
BT2 essential amino acids
BT3 amino acids
BT1 sympatholytics
BT2 neurotropic drugs
NT1 aidomet
rt levodopa

DOPAMINE
BT1 catecholamines
BT2 ammonium compounds
BT1 neurotransmitters
NT1 bromocriptine
rt cholinergic mechanisms

DOPAMINE BETA-MONOOXYGENASE
(ec 1 14 17 1)
BT1 oxidoreductases
BT2 enzymes

DOPATRIUM
BT1 scrophulariaceae
NT1 dopatrium junceum

DOPATRIUM JUNCEUM
BT1 dopatrium
BT2 scrophulariaceae

DOPING
BT1 drug therapy
BT2 therapy
rt depression
rt stimulation
rt varnishes

doralis
USE aphis

doralis fabae
USE aphis fabae

DORITIS
BT1 orchidaceae
NT1 doritis pulcherrima

DORITIS PULCHERRIMA
BT1 doritis
BT2 orchidaceae

DORITIS PURCHERRIMA
BT1 ornamental plants

DORMANCY
BT1 developmental stages
BT2 development
NT1 aestivation

DORMANCY (C)
NT1 diapause
NT1 sedd dormancy
NT1 torpor
rt biological rhythms
rt bud break
rt buds
rt dormancy breakers
rt germination
rt hard seeds
rt hibernation
rt rest
rt sleep
rt stratification
rt vernalization

DORMANCY BREAKERS
BT1 growth regulators
rt biostimulators
rt dormancy
rt gibberellins

DORMANCY BREAKING
rt germination

DORONICUM
BT1 compositae
NT1 doronicum cordifolium
NT1 doronicum macrophyllum
NT1 doronicum oblongifolium

DORONICUM CORDIFOLIUM
BT1 doronicum
BT2 compositae
BT1 ornamental plants

DORONICUM MACROPHYLLUM
BT1 doronicum
BT2 compositae
BT1 medicinal plants

DORONICUM OBLONGIFOLIUM
BT1 doronicum
BT2 compositae
BT1 medicinal plants

DOROSOMA
BT1 clupeoidei
BT1 fishes
rt shad

DOROTHEANTHUS BELLIDIFLORUS
uf mesemoryanthemum crinitiorum

DORPER
BT1 sheep breeds
BT2 breeds

DORSET DOWN
BT1 sheep breeds
BT2 breeds

DORSET HORN
BT1 sheep breeds
BT2 breeds

DORYALIS
uf dovyalis
BT1 flacourtiaceae
NT1 doryalis caffra

DORYALIS CAFFRA
uf dovyalis caffra
BT1 doryalis
BT2 flacourtiaceae
BT1 tropical fruits

DORYANTHES
BT1 amaryllidaceae

DORYLAIMIDAE
NT1 enchodeilus
NT1 eudorylaimus

DORYLUS
BT1 hymenoptera
NT1 dorylus labiatus

DORYLUS LABIATUS
BT1 dorylus
BT2 hymenoptera

DORYPHORA
BT1 atherospermataceae

Figura 6.5.

Na bibliografia, assinalam-se os seguintes tipos de relações hierárquicas:

1. hierárquica genérica,
2. hierárquica partitiva = todo → com a parte.
3. hierárquica enumerativa,
4. poliirárquica.

Estudemos cada uma delas.

1. ***Relação hierárquica genérica.***

Refere-se a relacionar o todo – cabeça de família – com seus membros.
Aconselha-se a utilizar a comparação

todos
em sentido ascendente e descendente
alguns

para poder estabelecer este tipo de hierarquia.

todos	→	anfíbios	
↓		↓	são
alguns		vertebrados	

todos		anfíbios	
↑		↑	são
alguns	→	vertebrados	

1.1 Espécies de animais ou plantas.

Exemplo:

VERTEBRADOS
TE ANFÍBIOS
AVES
MAMÍFEROS
RÉPTEIS

1.2 Objetos e suas classes

Exemplo:

DOCUMENTOS
TE DOCUMENTOS PRIMÁRIOS
DOCUMENTOS SECUNDÁRIOS

1.3 Ações e propriedades

Exemplo:

TRATAMENTOS TÉRMICOS
TE TEMPERADO
ENDURECIMENTO
CLASSIFICAÇÃO DE DOCUMENTOS
TE CLASSIFICAÇÃO ALFABÉTICA
CLASSIFICAÇÃO FACETADA

1.4 Conceitos abstratos e seus efeitos

Exemplo:

VALORES
TE VALOR ECONÔMICO
VALOR MORAL
VALOR SOCIAL

1.5 Profissões e suas classes

Exemplo:

BIBLIOTECÁRIOS
TE BIBLIOTECÁRIOS GERAIS
BIBLIOTECÁRIOS DE HOSPITAIS

2. ***Relação hierárquica partitiva***

o todo → a parte

2.1 Sistemas e órgãos do corpo

Exemplo:

OUVIDO
TE OUVIDO INTERNO
TROMPA DE EUSTÁQUIO
TÍMPANO

2.2. Localidades geográficas

Exemplo:

ESPANHA
TE1 CASTELA
TE2 BURGOS

2.3. Ciências e disciplinas

Exemplo:

CIÊNCIAS DA DOCUMENTAÇÃO
TE BIBLIOTECONOMIA
ARQUIVOLOGIA
DOCUMENTAÇÃO

GEOLOGIA
TE GEOLOGIA ECONÔMICA
GEOLOGIA FÍSICA
PETROLOGIA

2.4. Estruturas hierárquicas sociais

Exemplo:

ARQUIDIOCESE
TE1 DIOCESE
TE2 PARÓQUIA

3. ***Relação hierárquica enumerativa***

Aplica-se em casos em que, não sendo a parte um componente do todo, pertence a ele como uma classe.

Exemplo:

MARES
TE MAR BÁLTICO
MAR CÁSPIO
MAR MEDITERRÂNEO

Observe-se que não são tipos de mares, porém mares com seu nome próprio, que se diferenciam por isto. Enumeram-se os diferentes mares.

4. *Relações poli-hierárquicas*

Por último, passemos a considerar os casos em que um termo pertence a duas ou mais famílias, quer dizer, pode depender de dois ou mais termos genéricos.

Exemplo:

Petroquímica
↓ ↓
Geologia Química
↓ ↓
Bibliotecas Universitárias
↓ ↓
Bibliotecas Universidade

O termo subordinado – específico – deverá figurar nas duas famílias.

Exemplo:

BIBLIOTECAS UNIVERSITÁRIAS
TG UNIVERSIDADE
TG BIBLIOTECONOMIA

UNIVERSIDADE
TE BIBLIOTECAS UNIVERSITÁRIAS

BIBLIOTECONOMIA
TE BIBLIOTECAS UNIVERSITÁRIAS

RELAÇÕES ASSOCIATIVAS

Sabemos que se estabelecem relações associativas entre os termos quando as famílias, ou grupos de termos afins, estudam-se a partir de um plano horizontal, levando-se em consideração diferentes pontos de vista.

Exemplo:

ENSINO À DISTÂNCIA
TR MÉTODOS DE ENSINO

ALIMENTAÇÃO
TR NUTRIÇÃO

É opinião generalizada que se deve ser cauteloso no momento de escolher relações associativas, porque se corre o perigo de exceder-se, com o pretexto de construir o mais exato tesauro. Contudo, o que se conseguirá é torná-lo mais confuso e difícil de utilizar, ao contar com um grande número de termos e relações.

De todas as formas, para os casos em que seja necessário utilizar estas relações, especificaremos, em continuação, que tipos podem se apresentar.

Tipos de relações associativas

1. Termos que pertencem à mesma categoria.

Exemplo:

BIBLIOGRAFIAS NACIONAIS
TR BIBLIOGRAFIAS ESPECIAIS

2. Termos que pertencem a categorias diferentes.

Aqui devem se considerar diversos casos:

2.1 Relação do todo com a parte.

Exemplo:

EDIFÍCIO
TR PORTAS

2.2 Disciplinas, ou campos de estudo, e os objetos ou fenômenos estudados.

Exemplo:

SISMOLOGIA
TR TERREMOTOS
ETNOGRAFIA
TR SOCIEDADES PRIMITIVAS

2.3. Uma operação ou processo e seu agente ou instrumento.

Exemplo:

SALÃO DE CABELEIREIRO
TR SECADOR DE CABELO
VELOCIDADE
TR VELOCÍMETRO

2.4. Uma ocupação e a pessoa que a exerce.

Exemplo:

CONTABILIDADE
TR CONTABILISTA

PROCESSAMENTO DE DADOS
TR ANALISTA DE SISTEMAS

2.5. Uma ação e um produto dessa ação.

Exemplo:

CONSTRUÇÃO DE ESTRADAS
TR ESTRADAS

2.6. Uma ação e um sujeito paciente.

Exemplo:

ENSINO
TR ESTUDANTE

ANÁLISE DE DADOS
TR DADOS

2.7. Conceitos relacionados com suas propriedades

Exemplo:

LIGAS DE AÇO
TR RESISTÊNCIA À CORROSÃO
MULHERES
TR FEMINILIDADE

2.8. Conceitos relacionados com sua origem.

Exemplo:

ÁGUA
TR POÇOS

INFORMAÇÃO
TR FONTES DE INFORMAÇÃO

2.9. Conceitos unidos por dependência de causa.

Exemplo:

EROSÃO
TR USO

AVARIA
TR DESGASTE

2.10. Uma matéria-prima e seu produto.

Exemplo:

COURO
TR PELES

FARINHA
TR PÃO

2.11. Uma ação e a propriedade associada a ela.

Exemplo:

APARELHOS DE PRECISÃO
TR PRECISÃO

MEIOS DE COMUNICAÇÃO
TR VERACIDADE

2. 12 Um conceito e seu oposto.

BIBLIOTECA
DO
IBICT

Exemplo:

SECA
TR UMIDADE

Como é óbvio, sempre que aparecer um termo com a notação de relação associativa, deverá aparecer outro, em seu lugar correspondente, com a mesma notação.

Exemplo:

UMIDADE
TR SECURA

SECA
TR UMIDADE

Na figura 6.6, é reproduzida uma página do *TEST-Thesaurus – thesaurus of engineering and scientific terms*, na qual se pode apreciar os descritores com uma boa variedade de termos relacionados.

Outro exemplo muito ilustrativo é apresentado pela figura 7.7, onde se expõe uma página do *Tesauro da Unesco*.

Com isto, revisamos todas as formas de relação que podem se apresentar entre os termos de um tesauro e que devem ser levadas em consideração no momento de escolher aquelas que integrarão sua base de conhecimento.

Convencional alphabetical display
Thesaurus of engineering and scientific terms (TEST)

—Food services
Cooking devices 0608
UF Electrical cooking devices
French fryers
Pressure cookers
Toasters
Waffle irons
RT Kitchen equipment and supplies
Kitchens
Cooking liquors (pulping) 1112
NT Spent sulfite liquors
Sulfite cooking liquors
White liquors
RT—Chemical pulping
Semichemical pulping
—Spent liquors (pulping)
Cook off 2102
RT Propellant storage
—Solid rocket propellants
Coolant pumps 1311
BT Pumps
RT—Air conditioning equipment
Cooling systems
—Refrigerating machinery
Coolants 1301 1107
RT Air conditioning
Air cooling
—Brines
—Coolers
—Cooling
Cooling systems
Cutting fluids
Dry ice®
Gas cooling
—Liquid cooling
Liquid metal coolants
Nuclear reactor coolants
—Nuclear reactors
Refrigerants
Water cooling
Coolers 1301
UF Aftercoolers
Intercoolers
Precoolers
NT Air coolers
Beverage coolers
Milk coolers
Oil coolers
Unit coolers
Water coolers
RT—Air conditioners
—Air conditioning equipment
—Compressors
Coolants
—Cooling
Cooling systems
Dehumidifiers
Freezers
Ice refrigeration
—Refrigerating
—Refrigerating machinery
—Refrigerators
Cool flames 2102
BT Flames
RT—Combustion
Cooling 1301 2013
UF Chilling
Heat dissipation
NT Adiabatic demagnetization
Air cooling
—Evaporative cooling
Expansion cooling
Film cooling
—Liquid cooling
Radiant cooling
Sublimation cooling
Sweat cooling
Thermoelectric cooling
Water cooling
RT Ablation
Air conditioning
Cold treatment
—Condensing
Contraction
Coolants
—Coolers
—Cooling coils
Cooling curves
Cooling load
Cooling rate
Cooling systems
Cooling towers
Desuperheating
Environmental engineering
—Fans
Flooding
—Freezing
—Heating
Heat loss
Heat radiators
—Heat transfer
Hilsch tubes
Jackets
—Melting
—Quenching (cooling)
Recalescence
—Refrigerating
Supercooling
—Temperature
Temperature control
Temperature distribution
Thermal cycling tests
Thermal shock
Thermal stresses
—Ventilation
Venting
Wetting
Cooling coils 1301
NT Direct expansion cooling coils
RT—Air conditioning equipment
—Condensers (liquefiers)
Condenser tubes
—Cooling
Cooling systems
Evaporators
Expansion valves
Gas expanders
—Refrigerating machinery
Cooling curves 1301
BT Charts
Graphs (charts)
RT—Cooling
Cooling rate
Phase diagrams
—Thermal analysis
Cooling fans 1301
BT Fans
Cooling fins 1301
UFt Finned tubes
Fins (coolers)
BT Fins
RT Cooling systems
Engine blocks
Engine cylinders
Heat exchangers
Heat radiators
Cooling load 1301
RT Air conditioning
—Cooling
Enthalpy
Fluid infiltration
Heating load
Heat storage
—Heat transfer
Heat transmission
—Loads (forces)
—Thermal insulation
Cooling rate 1301
BT Rates (per time)
RT Air conditioning
—Cooling
Cooling curves
Cooling systems
—Freezing
—Precipitation heat treatment
—Refrigerating
Thermal shock
Thermal stresses
Cooling systems 1301
UF Water cooling systems
RT Absorbers (equipment)
Absorption refrigeration
Air circulation
—Air conditioners
Air conditioning
—Air conditioning equipment
Air cooling
Beverage coolers
Blowers
—Compressors
—Condensers (liquefiers)
Coolant pumps
Coolants
—Coolers
—Cooling
—Cooling coils
Cooling fins
Cooling rate
Cooling towers
Dehumidification
Ducts
Engine blocks
—Evaporative cooling
Evaporators
—Exhaust systems
—Fans
Gas expanders
Heat exchangers
Heat pumps
Heat radiators
Heat sinks
Humidity control
—Intake systems
—Liquid cooling
Lubrication systems
Mechanical refrigeration
Piping systems
—Porous metals
Refrigerants
—Refrigerating
—Refrigerating machinery
Registers (air circulation)
Steam jet apparatus
Steam jet refrigeration
Temperature control
Thermoelectric refrigeration
—Transpiration
Transport refrigeration
Unit coolers
—Ventilation
Vents
Water coolers
Water cooling
Cooling towers 1301 0701 1309
UF Water cooling towers
RT—Air conditioning equipment
Columns (process engineering)
—Cooling
Cooling systems
—Evaporative cooling
—Refrigerating machinery
Water conservation
Water coolers

Subject Category Index numbers follow main terms; (—) = See main entry for narrower terms; t = Consult main entry;

Figura 6.6.

B/H Ciencia y tecnología

B •
UF Financiacion de la tecnología
B24 Administración de la ciencia
UF Administración de la ciencia
B24.30 Organizaciones científicas
B24.30.10 Instituciones científicas
UF Instituciones tecnológicas
B25/29 Investigación científica
UF Investigación tecnológica
Investigación oceanográfica D86
Instalaciones científicas B36/39
Investigación agrícola G06
Investigación biológica F05
Investigación química C32
Investigación ecológica E11
Investigación geográfica D35.10
Investigación hidrológica D67
Investigación médica F51
B26 Expediciones científicas
B28 Investigación y desarrollo
UF Investigación industrial
R&D
B287.10 Desarrollo experimental
UF Investigación sobre el desarrollo
Proyectos piloto A81.40
Desarrollo de los procesos N38.70.20
Desarrollo de los productos N38.70.30
Proyectos experimentales A81.10
Experimentación A88
B29 Resultados de la investigación científica
SN Incluidas las aplicaciones de dichos resultados
Concesión de patentes A65/69
Descubrimientos científicos B16.10
Innovaciones científicas B16.15
Transferencia de tecnología B30
B30 Transferencia de tecnología
Resultados de la investigación científica B29
Industrias extranjeras N40.15
Desarrollo industrial N37
Industrias autóctonas N40.10
B30.10 Transferencia horizontal de tecnología
UF Adaptación de tecnología
B30.20 Transferencia vertical de tecnología
B30.30 Difusión de tecnología
B30.40 Implantación de tecnología
B30.60 Transferencia internacional de tecnología
Países beneficiarios L84.65
Asistencia técnica L84.50.60
Países donantes L84.60
Instrumentos internacionales M75
Ayuda internacional L84
B30.80 Servicios de transferencia de tecnología
Oficinas de patentes A68.20
Sistemas de información científica Z68.50
B34 Personal científico
UF Personal técnico
Personal tecnológico
Personal médico F98.20
B34.10 Comunidades científicas
B34.25 Personal técnico superior
B34.30 Científicos
B34.35 Tecnólogos
B34.38 Ingenieros
B34.40 Ambientalistas
B34.42 Geofísicos

Figura 6.7.

TEMAS PARA ESTUDO

1. Diga quando construiria um tesauro novo.
2. Usaria termos simples ou compostos?
3. Cite alguma classe de termos simples.
4. Usaria um verbo como descritor?
5. Cite algum tipo de relação hierárquica entre termos.
6. Diga como construiria o índice de um tesauro.
7. Cite alguma classe de termos compostos que deva conservar-se como tal.
8. Cite algum caso em que dividiria um termo.
9. Utilizaria um substantivo como descritor?
10. Cite algum caso de relação de quase sinonímia.
11. O que quer dizer "termo relacionado"?
12. Diga de onde tiraria os termos para elaborar um tesauro.
13. Levaria em consideração assuntos marginais ao tema principal de trabalho?
14. Quando consideraria que dois termos são sinônimos?
15. Caso existisse um tesauro sobre sua especialidade, conStruiria um novo?

REFERÊNCIAS BIBLIOGRÁFICAS

ALVARO BERMEJO, C., VILLAGRA RUBIO, A., SORLI ROJO A.: "Desarrollo de Lenguajes Documentales Formalizados en Lengua Espanola"; II Evaluación de los Tesauros disponibles en Lengua Espanola"; Rev. Esp. Doc. Cient., *12*, 3, 1989.

BSO: Broad System of Ordering; schedule and index, third revision: by FID/BSO, FID Nr. 564, The Hague-París, FID, 1978.

CURRÁS, E.: "Elaboración de lenguajes gráficos y científicos para su automatización '; en I Seminario Hispano-Luso sobre sistemas generales de Información Tecnológica, Soc. Esp. de Sist. Generales, Madrid, 17-21 Noviembre, 1980.

* CHAUMIER, J.: "Análisis y lenguajes documentales '; Barcelona, Mitre, 1986.

* FOSKETT, D.I.; ''The construction of a faceted classification for a special subject"; London, Butterworth, 1959.

GALINDO, F.: "Desarrollo experimental de un modelo de thesaurus jurídico inteligente"; en jornadas sobre Informática y Derecho, Madrid, 1985, 1-19.

MALDONADO MARTINEZ, A. y GAVIRA GULPE, C.: ''El control del lenguaje en la producción de bases de datos; el tesauro de urbanismo ISOC"; en II Jornadas Españolas de Documentación Automatizada, Málaga, 1986, 125-136.

* Norma ISO 2788. Documentation-Guidelines for the establishment and development of monolingual thesauri, 1986.

Norma UNE 50-106. Documentación-Directrices para el establecimiento y desarrollo de tesauros monolingües.

* ORNA, E.: "Build yourself a thesaurus: a step by step guide"; Norwich, Running Angel, 1983.

PEREZ, E.: ''Text enhacement. Controlled vocabulary vs free text''; Special Libraries, *73*, 3, July 1982, 183-192.

Thesaurus of Engineering and Scientific Terms – Test – , a list of scientific and engineering terms and their relationship for as a vocabulary in indexing and retrieving technical information, New York, Engineers Joint Council and US Departament of Defense, 1967.

VICKERY, B.C.: ''Faceted classification: a guide to construction and use of special schemes''; London, ASLIB, 1960.

7

Formas de Apresentação de Tesauros

Uma vez que já se tenham os termos escolhidos e se tenham estabelecido as relações entre eles, o passo seguinte é pensar a forma de apresentá-los, para seu melhor uso.

À primeira vista, parece que uma ordenação alfabética dos termos principais, com seus termos relacionados colocados debaixo, indicando, com a notação correspondente, a forma de relação, pode ser suficiente para localizá-los e para conhecer deles tudo o que for necessário. Realmente, na maioria dos casos, basta essa forma de apresentação. Todavia, a experiência tem demonstrado que, muitas vezes, torna-se necessário ordenar os termos conforme as famílias que havíamos formado inicialmente. Isto significa que se deve distribuir por classes que proporcionem uma visão de conjunto do assunto.

Por outro lado, em outra quantidade de casos, uma representação gráfica dos termos e suas relações ajuda a compreender sua ordenação e facilita o uso do tesauro.

Assim, pois, deve-se considerar três tipos de representações de tesauros:

1. representação alfabética;
2. representação em grupos ou classes, quer dizer, sistemática;
3. representação gráfica.

APRESENTAÇÃO ALFABÉTICA DE TESAUROS

Além da apresentação alfabética descrita anteriormente, pode-se realizá-la de modo mais simples. Assim, poder-se-á colocar um ou dois termos com suas relações correspondentes sob o descritor e acrescentar a notação referente à sua posição na representação sistemática. Um exemplo é mostrado na figura 7.1. Este tipo de representação alfabética é a que se costuma usar como índice e figura no final do tesauro para poder procurar os termos mais facilmente.

Cabe acrescentar, além disso, que alguns tesauros aparecem acompanhados de um índice tipo KWIC para completar a localização dos termos. Veja-se o exemplo tomado da obra *Relación Permutada Descriptores Índice Español*, na figura 7.2.

	DESCRIPTORES	FREC
	BETULACEAE	23
CARPE (CARPINUS	BETULUS)	0
	BETUN ASFALTICO	28
ANILLOS DE	BEZOUT	1
	BIBIONIDAE	0
	BIBLIOGRAFIA	161
REVISION	BIBLIOGRAFICA	610
	BIBLIOMETRIA	5
	BIBLIOTECA	4
	BICARBONATO SODICO	15
	BICARBONATOS	10
COMPUESTOS	BICARBONILICOS	0
CETONAS	BICICLICAS	2
COMPUESTOS	BICICLICOS	4
	BICICLOOCTANO	1
	BICICLOOCTENO	0
	BICORNIFERA	3
RIO	BIDASOA	0
CUENCA DEL	BIDASOA	3
	BIDENTADO	1
	BIDIMENSIONAL	64
	BIDUAL	0
	BIELAS	0
	BIENES DE EQUIPO	19
INDUSTRIA DE	BIENES DE EQUIPO	16
	BIENESTAR	4
EL	BIERZO	6
	BIESTABLE	15
MULTIVIBRADOR	BIESTABLE	4
	BIFASICO	2
	BIFENILO (VER: DIFENILO)	0
	BIFIDOBACTERIUM	0
	BIFLUORURO AMONICO	0
VICIA	BIFOLIATA	1
	BIFONAZOL	0
	BIFOSFOGLICERATO	0
	BIFURCACION	1
	BIFURCACION DE HOPF	0
	BIGUANIDA	1
	BIGUANIDAS	11
	BILBAO	26
RIA DE	BILBAO	1
VESICULA	BILIAR	2
SECRECION	BILIAR	9
SALES	BILIARES	3
ACIDOS	BILIARES	2
CALCULOS	BILIARES	3
PIGMENTOS	BILIARES	1
MODELO	BILINEAL	4
MODULO	BILINEAL	0
OPERADOR	BILINEAL	0
INTEGRACION	BILINEAL	2
FORMAS	BILINEALES CONTINUAS	1
	BILIRRUBINA	6
	BILIS	35
	BILIVERDINA	0
MEZCLA	BINARIA	82
ESTRELLA	BINARIA	18
MEZCLAS	BINARIAS	145
ESTRELLAS	BINARIAS ECLIPSANTES	0
	BINARIO	7
SISTEMA	BINARIO	13
DISTRIBUCION	BINOMIAL	7
	BIOCANINA	1
	BIOCENOSIS	49
	BIOCIDAS	3
	BIOCLASTOS	0
SISTEMA	BIOCLIMATICO DE EMBERGER	0
	BIOCLIMATOLOGIA	47
	BIOCOMPATIBILIDAD	8
	BIOCONCENTRACION	1
	BIOCRISTALES	0
	BIOCRONOLOGIA	10
	BIODEGRADABILIDAD	12
POLIMEROS	BIODEGRADABLES	4
RESIDUOS NO	BIODEGRADABLES	0
	BIODEGRADACION	55
	BIODEPOSITOS	9
	BIODISPONIBILIDAD	113
SEÑALES	BIOELECTRICAS	0
	BIOELEMENTOS	16
	BIOESTADISTICA	21
	BIOESTRATIGRAFIA	324
	BIOESTROMA	0
	BIOFACIES	18

Figura 7.2 - Relação permutada descritores. Índice em espanhol.

- lugar,
- tempo.

A norma UNE 50-106 propõe como exemplo o termo AVIÕES, que aqui continuou expandindo-se da seguinte forma:

AVIÕES

Materiais – Elementos para a sua construção
ESTRUTURAS DE ALUMÍNIO
PLÁSTICOS

Propriedades e medidas
AVIÕES A REAÇÃO
HIDROAVIÕES
ALTÍMETROS

Agentes – Usuários
AVIÕES MILITARES
AVIÕES CIVIS
AVIADORES
CONTROLADORES DE VÔO

Operações e processos

ATERRISSAGEM
HANGARES
PISTAS DE PROVAS

Entidades

As diferentes marcas de fabricação
As companhias aéreas

Lugar
Não necessita de maiores comentários
Tempo

Para realizar estas tarefas, poder-se-á desenhar um ficha onde se irá anotando os termos nos lugares correspondentes, conforme venham surgindo.

Dentro de cada um dos grupos, poder-se-á estabelecer outras divisões, também facetadas, até o momento de elaborar um sistema de classificação completo, que terá os lugares apropriados a cada assunto.

A figura 7.3 mostra outras formas de apresentar as divisões das facetas e foi reproduzida da obra *Catering Thesaurus*. Observam-se as diferenças entre as diversas seções A, B,C, D, que representam etapas diferentes na confecção do tesauro.

Para designar as diferentes classes que se criam em uma classificação destas características, pode-se utilizar uma notação numérica, alfabética ou alfa-numérica. A figura 7.4, reproduzida da seção Clothing Technology do tesauro Precis, mostra um exemplo de notação alfabética.

Os epígrafes ou indicadores das diferentes classes em que se vão enquadrando os termos para elaborar a parte sistemática do tesauro formam os chamados indicadores de função. * Na norma UNE 50-106, recebem o nome de "indicadores classificatórios".

É nosso desejo que as recomendações e exemplos aqui expostos possam servir de ajuda para elaborar a apresentação sistemática de um tesauro. Não é tarefa fácil. Além de um profundo conhecimento do assunto, torna-se necessário um bom critério para colocar cada termo em seu lugar e manter a ficha aberta para receber possíveis modificações.

APRESENTAÇÃO GRÁFICA DE TESAUROS

Em determinadas circunstâncias ou situações especiais, torna-se necessário ordenar os termos graficamente. Com isto, consegue-se maior compreensão das relações dos termos dentro de uma mesma família, além de uma visão de conjunto de tudo aquilo que compõe essa família.

Contudo, apresenta o inconveniente de reduzir o número de termos dentro de cada família ou classe, para poder elaborar um gráfico compreensível. Por outro lado, torna-se difícil estabelecer relações entre termos de famílias diferentes. Além do que, planejar um gráfico apropriado é muito trabalhoso e consome muito tempo. Todos esses motivos desaconselham as representações gráficas dos tesauros. Não se deve realizar a não ser que seja absolutamente necessário. Muitos documentalistas pensam que não possuem um tesauro finalizado enquanto não construam uma apresentação gráfica. Deve-se argumentar, contudo, que isto é, geralmente, uma tarefa supérflua e que consome um número considerável de recursos humanos, econômicos e de tempo. Na maioria dos casos, a utilidade obtida não compensa o esforço despendido.

De toda a maneira, para aqueles que necessitam apresentar os tesauros graficamente, abordaremos o assunto. Pode-se escolher tantas formas de representação gráfica, como variada pode ser a imaginação humana, portanto qualquer diagrama será aceito. Não obstante, algumas normas gerais devem ser cumpridas:

1. O termo genérico maior, que abrange todo o âmbito semântico da família de termos, deverá ser colocado em um lugar central, remarcando sua situação, e ser escrito com outro tipo de letra para que ressalte à simples visão.

* N.T.: Em português, chamamos também indicadores de faceta ou rótulo nodal.

Subject field analysis and thesaurus derivation
Catering and related fields

Step A

Find facets/groups
(Arrangement of terms previously listed alphabetically under Catering)

CATERING
(Personnel - general)
Catering personnel

(Equipment)
Catering quipment
Cookers
Microwave ovens

(Operations and personnel associated)
Catering management
Catering managers
Cookery
Chefs
Cooks
Food purchasing
Food preparation
Food service
Waiters
Waiting

(Areas and associated operations)
Bars (licensed)
Canteens
Cafes
Kitchens
Restaurant management
Restaurants

(Food and meals)
Bakery products
Beverages
Convenience food
Dairy products
Fast food
Fish
Food
Food dishes
Fruit
Meals
Menus
Vegetables

(Applications)
(Not further analysed)
Educational catering
Hospital catering
Hotel catering
Institutional catering
School catering

Step B

Make hierarchies/arrays
Code terms NT/RT to term one step above

CATERING
(Personnel)
(RT) Catering personnel

(Equipment)
(RT) Catering equipment
(NT) Cooking appliances
(NT) Cookers
(NT) Microwave ovens

(Operations)
(NT) Catering operations
(NT) Catering management
(Personnel)
(RT) Catering managers

(NT) Food purchasing
(NT) Food preparation
(Personnel)
(RT) Chefs
(Operations)
(NT) Cookery
(NT) Food service
(NT) Waiting
(Personnel)
(RT) Waiters

(Areas)
(RT) Catering areas
(NT) Bars (licensed)
(NT) Canteens
(NT) Restaurants
(Management)
(RT) Restaurant management
(NT) Kitchens

(Food and meals)
(RT) Food
(Dishes)
(NT) Food dishes
(By special properties)
(NT) Convenience food

(Individual foods)
(RT) Beverages
(NT) Meat
(NT) Fish
(NT) Dairy products
(NT) Fruit
(NT) Vegetables
(NT) Bakery products
(RT) Meals
(RT) Menus

Step C

Add SNs, UFs and BT/NT/RTs from other subject fields, and links within the catering field

CATERING
(Personnel)
(RT) Catering personnel
NT Catering managers
Chefs
Waiters

(Equipment)
(RT) Catering equipment
SN Includes food mixers, cooking ware, tableware, etc.
UF Kitchen equipment
RT Dishwashers
(NT) Cooking appliances
BT Domestic appliances
(NT) Cookers
(NT) Microwave ovens

(Operations)
(NT) Catering operations
(NT) Catering management
BT Management
NT Restaurant management
RT Food purchasing

(Personnel)
(RT) Catering managers
BT Catering personnel
BT Managers

(NT) Food purchasing
BT Purchasing
RT Catering management

(NT) Food preparation
(Personnel)
(RT) Chefs
UF Cooks
BT Catering personnel
(Operations)
(NT) Cookery

(NT) Food service
(NT) Waiting
(Personnel)
(RT) Waiters
BT Catering personnel

(Areas)
(RT) Catering areas
(NT) Bars (licensed)
RT Licensing laws
(NT) Canteens
(NT) Restaurants
(Management)
(RT) Restaurant management
BT Catering management
Kitchens

(Food and meals)
(RT) Food
(Dishes)
(NT) Food dihes
RT Meals

(By special properties)
(NT) Convenience food
UF Fast food

(Individual foods)
(RT) Beverages
UF Drinks
(NT) Meat
(NT) Fish
(NT) Dairy products
(NT) Fruit
(NT) Vegetables
(NT) Bakery products
(RT) Meals
RT Food dishes
(RT) Menus
RT Diet

Step D

Make conventional alphabetical display.
(A few selected entries. See Figure 36 for complete thesaurus)

CATERING PERSONNEL
NT Catering mangers
Chefs
Waiters
RT Catering

COOKING APPLIANCES
BT Catering equipment
Domestic appliances
NT Cookers
Microwave ovens

CATERING MANAGEMENT
BT Catering operations
Management
NT Restaurant management
RT Catering managers
Food purchasing

CHEFS
UF Cooks
BT Catering personnel
RT Food preparation

FOOD SERVICE
BT Catering operations
NT Waiting

BARS (licensed)
BAT Catering areas
RT Licensing laws

RESTAURANT MANAGEMENT
BT Catering management
RT Restaurants

FOOD DISHES
RT Food
Meals

BEVERAGES
UF Drnks
RT Food

MEALS
RT Catering
Food dishes
Menus

Figura 7.3.

Faceted classification

M	CLOTHING TECHNOLOGY
	(Personnel)
MC	Clothing technology personnel
MCE	Clothing technology managers
MCP	Clothing workers
MCP.D	Sewing machinists
	(Equipment)
ME	Clothing technology equipment
MEH	Sewing machines
	(Operations)
MG	Clothing manufacturing processes
MGH	Patternmaking
MGJ	Cutting
MGL	Sewing
MGT	Tailoring
MGV	Dressmaking
	(Materials)
MI	Clothing materials
MIK	Textile fabrics
MIR	Leather
MIV	Fur
	(Parts)
MKL	Collars
MKP	Sleeves
MKQ	Cuffs
MKV	Linings
	(End-products)
MM	Clothing
	(By process used)
MNP	Knitwear
	(By property)
MOR	Lightweight clothing
	(By material)
MPI.T	Woollen clothing
MPI.V	Fur clothing
	(By purpose)
MQQ	Outerwear
MQQ.C	Coats
MQQ.H	Dresses
MQQ.N	Suits
MQR	Hosiery
MQS	Headwear
MQT	Footwear
MQU	Underwear
MSN	Nightwear
MST	Sportswear
MTU	Uniforms
	(By user)
MVC	Baby clothing
MVE	Childrens clothing
MVM	Womens clothing
MVM.P	Maternity wear
MVR	Mens clothing

Figura 7.4.

2. O restante dos termos deverá ser distribuído pelo gráfico de forma que fiquem claros a sua dependência ao termo genérico maior e os níveis de hierarquização ou sinonímia.

3. As relações associativas deverão ser mostradas explicitamente, para sua melhor compreensão.

4. Também deverão ser explicitadas as relações com famílias irmãs.

5. O diagrama deverá estar suficientemente assinalado de modo a encontrar com facilidade o termo que se busca.

6. Tipos e tamanhos diferentes de letras deverão ser usados, assim como linhas e outras pontuações para poder distinguir categorias e classes de termos.

7. Cada diagrama deverá vir acompanhado de uma lista alfabética de seus termos, com a referência ao lugar onde se encontra.

8. Uma lista dos diferentes diagramas deverá ser confeccionada, com a indicação do assunto ou termos genérico maior a que faz referência.

Quando se pensa no papel como suporte, fala-se de gráficos de duas dimensões. Atualmente, os computadores permitem alguma outra dimesão e outras características. Naturalmente, quando se faz uma cópia na impressora, voltamos a nos encontrar em um plano de duas dimensões. Isto também deve ser considerado no momento de planejar uma representação gráfica de um tesauro.

Nos tesauros mais conhecidos e nos manuais com regras para sua construção, foram observadas as seguintes formas de ***apresentação gráfica***, que são, portanto, as mais usuais.

1. Forma em árvore.

2. Forma quadriculada
   – em um espaço fechado, dividido em quadrículos.

3. Em famílias de termos interligadas.

4. Formando uma espiral em sentido ascendente.

5. Formas várias.

A forma em árvore, como seu nome indica, pressupõe que se coloque o termo principal na parte superior do diagrama e, logo, o restante dos termos, em sucessivas ramificações, segundo seu nível de dependência daquele. As figuras 7.5 e 7.6 mostram dois exemplos de representação deste tipo, tiradas do ISDD tesauros e PRECIS, respectivamente. Na figura 7.5, incluem-se

uma lista com um termo descritor e os termos que estão relacionados com ele, indicando seus tipos de relação. Na figura 7.6, os termos levam sua notação classificatória referida à seção semântica do tesauro. So,, Sn,, Sm indicam o tipo de relação de uns termos com outros.

Para representar um tesauro em forma quadriculada, quer dizer, por meio de um diagrama em quadrículos, costuma-se escolher o retângulo como padrão mais adequado, ainda que exista algum caso no qual se usem os quadrados, os triângulos ou os polígonos. Cada retângulo, tomemos por exemplo, divide-se em quadrículos eqüiláteros e constitui um diagrama por si mesmo. Cada diagrama é destinado a um termo genérico maior, com seus termos subordinados. Normalmente, coloca-se o termo genérico maior em algum quadrículo central, bem assinalado, de modo a ser distinguido facilmente. O resto dos termos é distribuído pelo quadrículos, em maior ou menor distância do termo genérico maior e de uns com outros, conforme seja seu nível de hierarquização. As dependências entre termos, que indicam, por sua vez, os tipos de relações, costumam ser assinaladas com flechas. Cada diagrama situa-se em um eixo de coordenadas para poder determinar a posição de cada quadrículo. Um bom exemplo é dado na figura 7.7.

A representação em famílias interligadas é a que proporciona maior informação em forma mais compreensível. Apresenta o inconveminente de ser muito trabalhosa em sua confecção e de requerer um tempo maior para esta tarefa. Em cada família, coloca-se um termo genérico maior com todos os seus subordinados e os sinais de relação. As famílias ligam-se umas às outras conforme o seu grau de dependência ou de relação. Na realidade, é como se cada família contivesse um microtesauro. Dependerá do detalhe de cada grupo de termos, o que se pode considerar verdadeiramente como um microtesauro. Muitas vezes, não é necessário tratar os termos de modo muito minucioso nas representações gráficas. A figura 7.8 nos mostra um bom exemplo, tirado do tesauro de Eudised.

Agora, aborda-se a ***apresentação formando um espiral***. Na realidade, costuma-se desenhar círculos concêntricos, porém, como se coloca o termo genérico maior no centro e os restantes de seu grupo nos diferentes círculos, cada vez, em maior distância, conforme seja sua dependência semântica, adotou-se a denominação "em espiral". O exemplo da figura 7.9 foi tirado da publicação do Conicyt, do Chile, chamada *Construcción de Tesaurus*.

Como curiosidade, são reproduzidas, nas figuras 7.10 e 7.11, uma apresentação em forma **poligonal** e outra **flechada**, respectivamente, tiradas do *Euroatom Thesaurus*.

Como se observa, não existe nada preestabelecido sobre estas questões.

É importante considerar que as apresentações gráficas dos tesauros poderiam substituir as seções sistemáticas, pois costumam pressupor a distribuição dos termos em classes e subclasses formando facetas. Contudo, deve-se ser muito precavido com estas situações. Os diagramas não devem ser muito complicados, nem muito numerosos, para que se tornem inteligíveis, embora, na maioria das vezes, uma classificação facetada seja extensa e complicada. Por conseguinte, nem sempre é aconselhável substituir uma apresentação sistemática por um simples gráfico. Mais aconselhável seria fazer o contrário.

Family tree structure
ISDD thesaurus

CHOICE OF DRUG USE

UF Drug use of choice
Preference of drug use

NT * MODIFICATION OF DRUG USE
* POLYDRUG USE
* SEQUENCE OF DRUG USE
* SINGLE DRUG USE
* UNUSUAL DRUG USE

SN Considerations of drug use which focus primarily on which drug or combination of drugs is being used. Use an NT wherever possible.

RT DOSAGE
Drug use status
EFFECTS OF DRUGS
PATTERNS OF USE
Studies

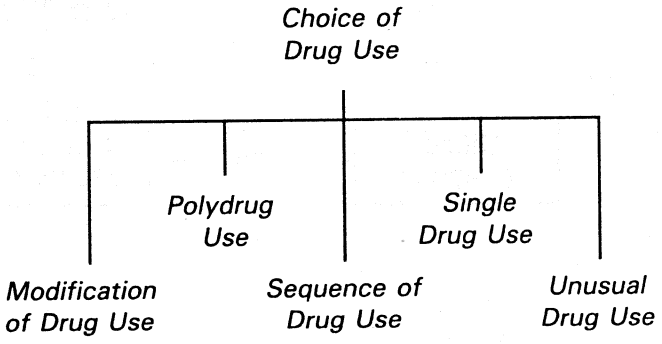


Figura 7.5.

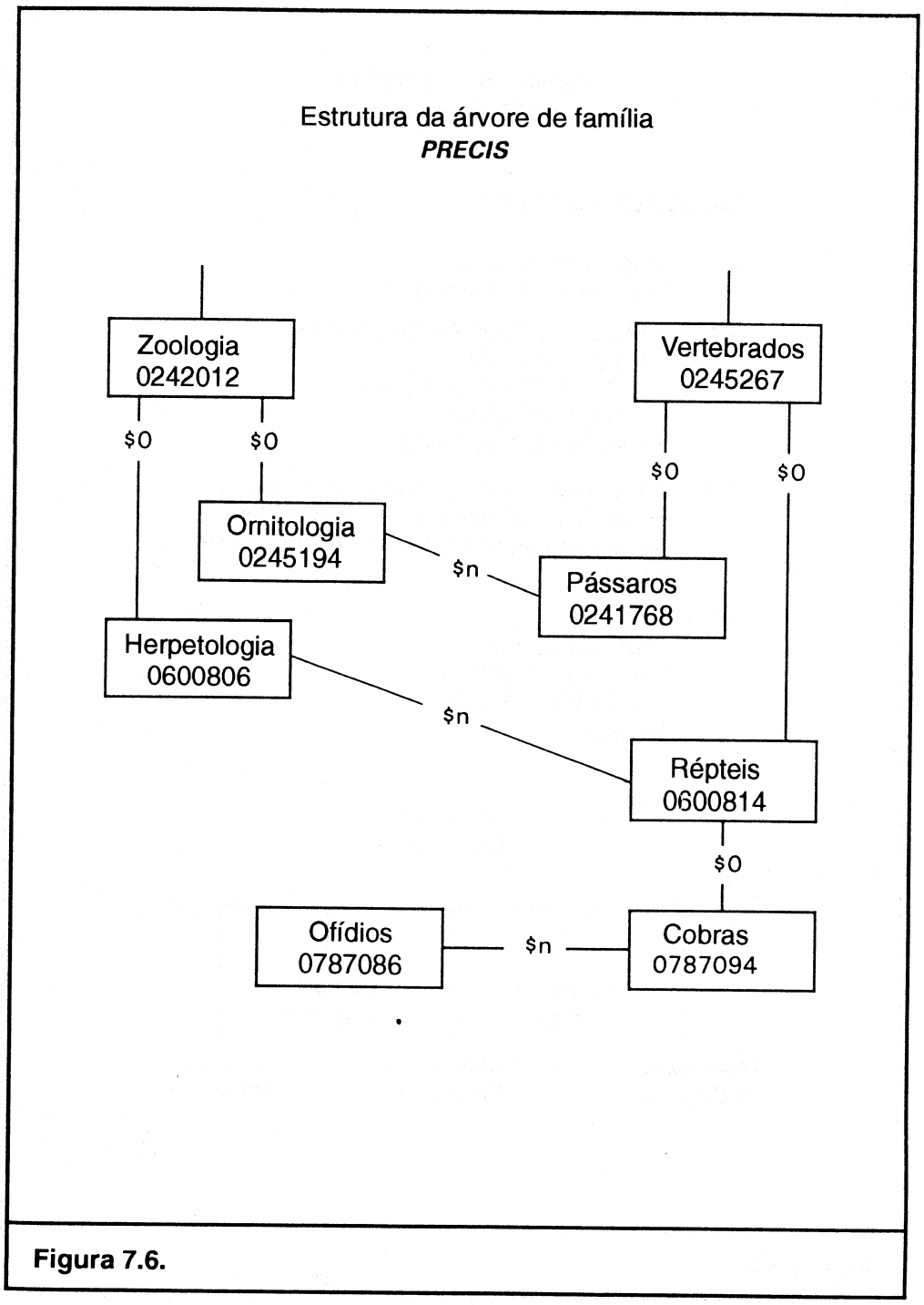

Estrutura da árvore de família
PRECIS
Zoologia
0242012
Vertebrados
0245267
$0
$0
$0
$0
Ornitologia
0245194
$n
Pássaros
0241768
Herpetologia
0600806
$n
Répteis
0600814
$0
Ofídios
0787086
$n
Cobras
0787094

Figura 7.6.

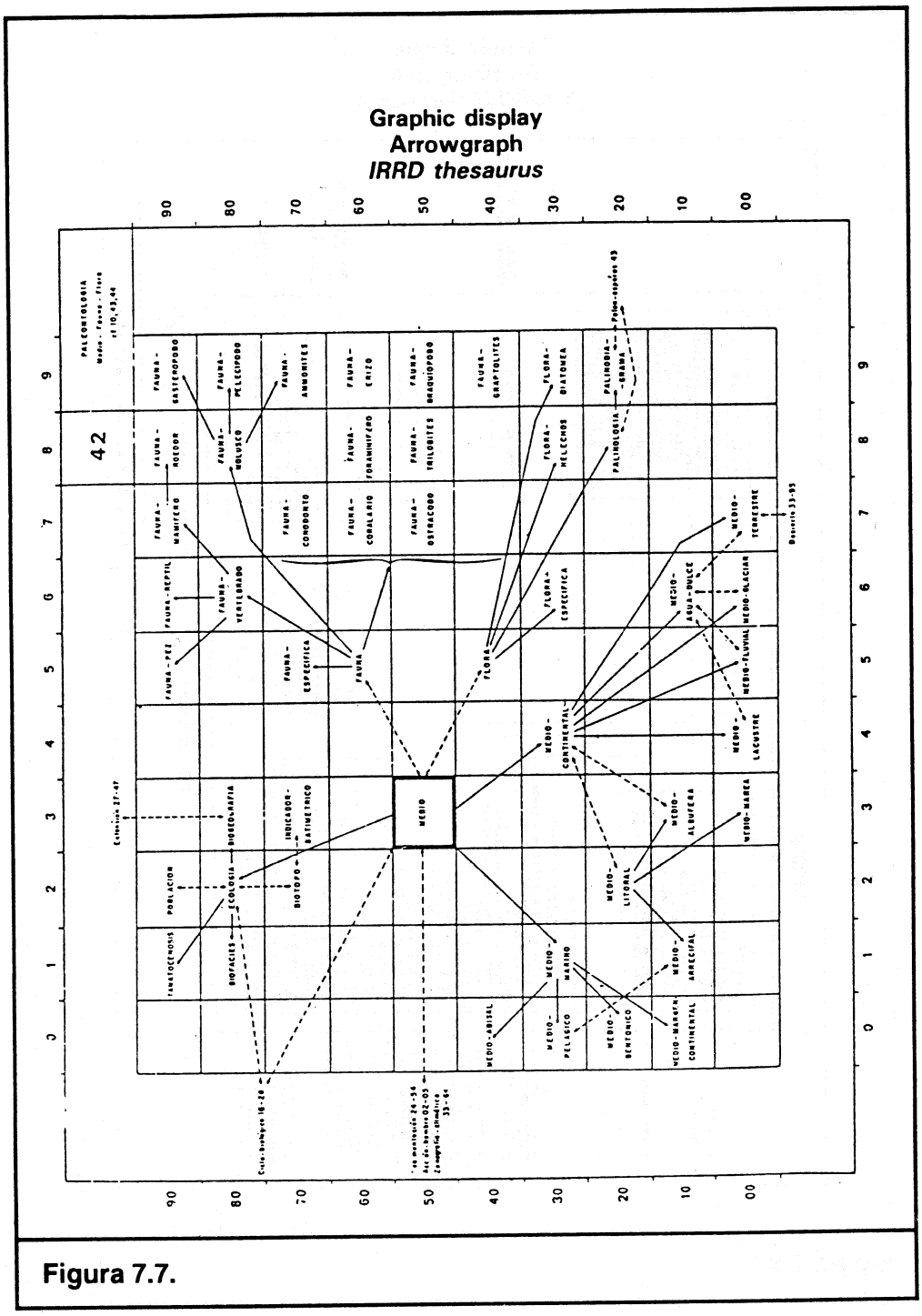

Graphic display
Arrowgraph
IRRD thesaurus
42
MEDIO
FAUNA
FLORA
PALEOECOLOGIA
FAUNA-ESPECIFICA
FAUNA-REPTIL
FAUNA-MAMIFERO
FAUNA-PEZ
FAUNA-MOLUSCO
FAUNA-ARTROPODO
FAUNA-GASTEROPODO
FAUNA-PELECIPODO
FAUNA-AMMONITES
FAUNA-ERIZO
FAUNA-FORAMINIFERO
FAUNA-CORALARIO
FAUNA-BRAQUIOPODO
FAUNA-TRILOBITES
FAUNA-OSTRACODO
FAUNA-GRAPTOLITES
FAUNA-CONODONTO
FLORA-ESPECIFICA
FLORA-DIATOMEA
FLORA-MELECHOS
PALINOLOGIA
PALINOLOGIA-ESPORA
MEDIO-CONTINENTAL
MEDIO-AGUA-DULCE
MEDIO-GLACIAR
MEDIO-FLUVIAL
MEDIO-LACUSTRE
MEDIO-TERRESTRE
MEDIO-MARINO
MEDIO-LITORAL
MEDIO-ALBUFERA
MEDIO-MARISMA
MEDIO-ARRECIFAL
MEDIO-MARGEN CONTINENTAL
MEDIO-ABISAL
MEDIO-PELAGICO
MEDIO-BENTONICO
INDICADOR-ACTIVIDAD TECTONICA
BIOTOPO
POBLACION
TANATOCENOSIS
BIOFACIES
BIOESTRATIGRAFIA
ECOLOGIA

Figura 7.7.

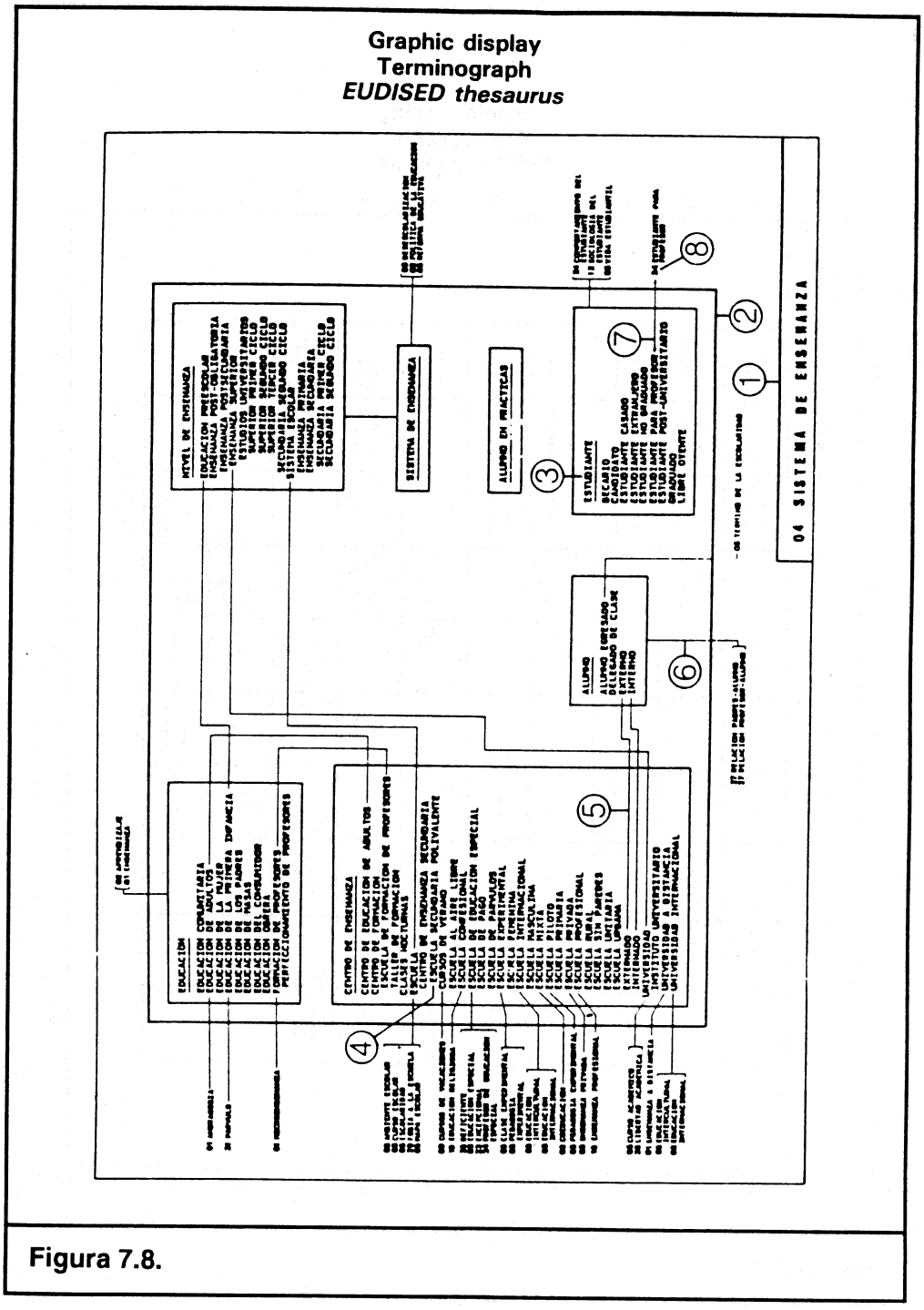

Graphic display
Terminograph
EUDISED thesaurus
04 SISTEMA DE ENSEÑANZA
NIVEL DE ENSEÑANZA
SISTEMA DE ENSEÑANZA
ALUMNO EN PRACTICAS
ESTUDIANTE
ALUMNO
CENTRO DE ENSEÑANZA
EDUCACION
1
2
3
4
5
6
7
8

Figura 7.8.

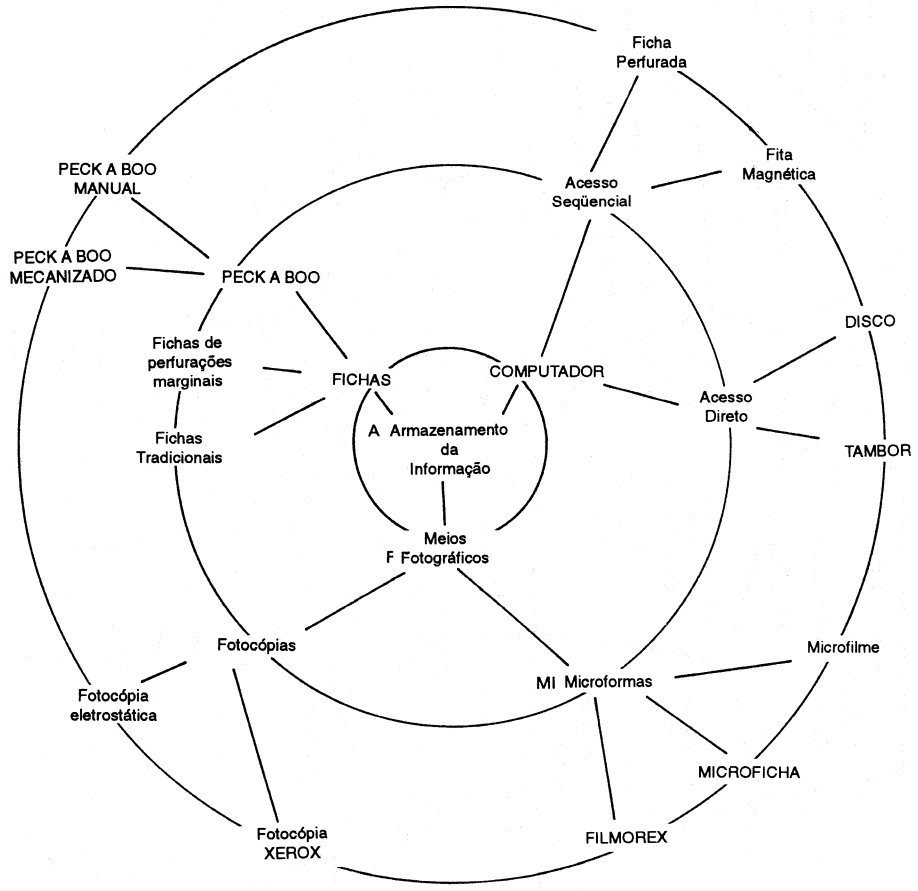


Figura 7.9. - Representação circular. Apresenta os termos distribuídos em círculos, cada um dos quais apresenta um nível de hierarquia, com o termo mais específico no centro

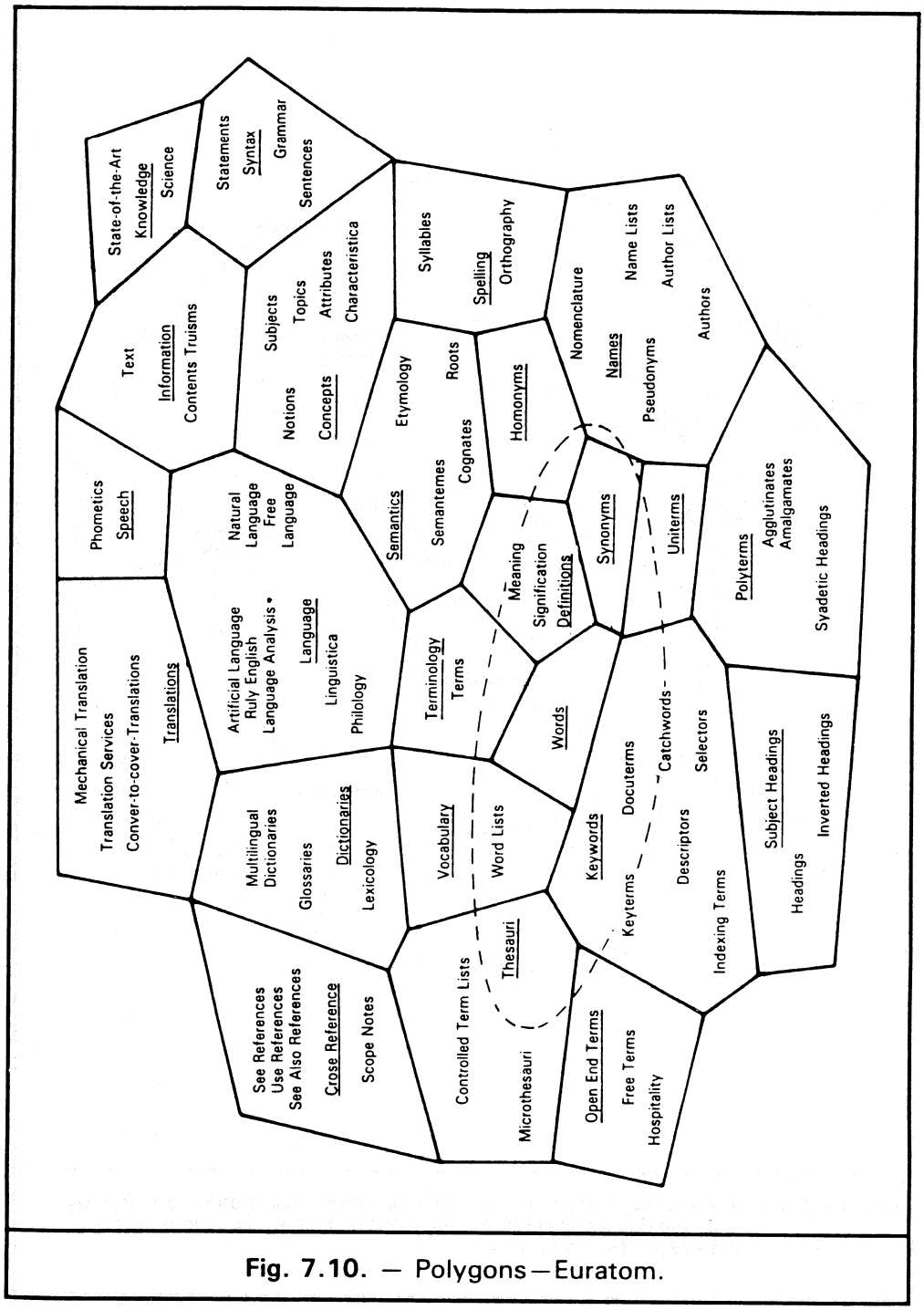


Fig. 7.10. — Polygons—Euratom.

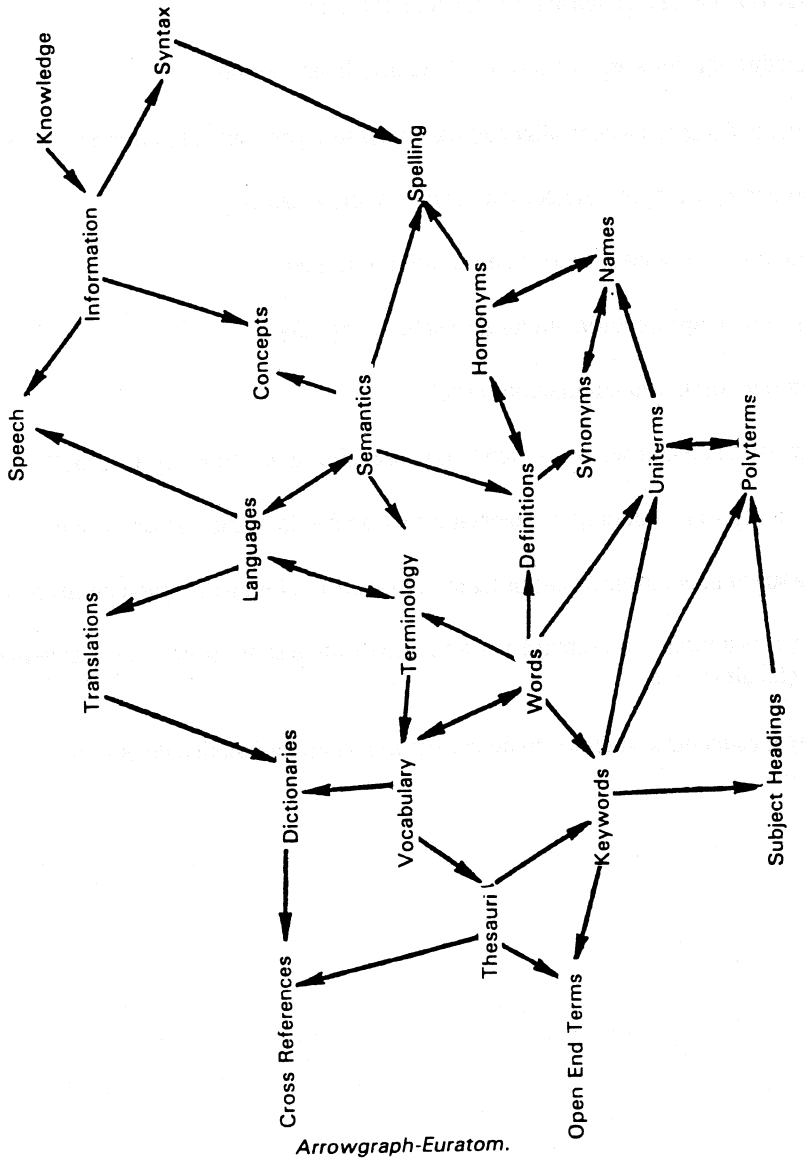


Figura 7.11. - Diagrama Flechado - Euratom

TEMAS PARA ESTUDO

1. Diga quando apresentaria graficamente um tesauro.

2. Acredita ser útil a apresentação sistemática de um tesauro?

3. Pensa que a apresentação alfabética de um tesauro pode substituir a apresentação gráfica?

4. Esquematize a apresentação sistemática de um tesauro.

5. Quantos índices alfabéticos pode conter um tesauro?

6. Diga como apresentaria um tesauro sobre geografia.

7. Como apresentaria um macrotesauro?

8. Que forma de apresentação escolheria para um tesauro sobre agricultura?

9. Na sua opinião, quem deve planejar a representação gráfica de um tesauro?

10. Na sua opinião, o que é mais útil: a apresentação em forma de espiral ou em quadrículos?

11. Acompanharia a apresentação sistemática de um tesauro sobre meio ambiente de uma seção alfabética?

12. Diga, esquematicamente, como construiria a parte alfabética de um tesauro.

REFERÊNCIAS BIBLIOGRÁFICAS

AITCHISON, J., BREWIN, P., COTTON, J.: ''DHSS-DATA Thesaurus''; London, Department of Health and Social Security, 1985.

AITCHSON, J., GEMERSALL, A., IRELAND, R.: "Thesaurofacet: a thesaurus and faceted classification for engineering and related subjects"; Whetstone, Leicester. England: English Electric Company Ltd., 1969.

AUSTIN, D.: ''PRECIS: a manual of concept analysis and subject indexing", 2nd. ed., London, British Library, 1984.

"CAB Thesaurus", by G. Eric Tidbury, Slough, England, Commomwealh Agricultural Bureaux, 1982.

CONICYT-Comisión Nacional de Investigación Científica y Tecnológica: ''Construcción de Tesauros''; Serie: Información y documentación, nr. 12, Santiago de Chile, 1975, 34.

DEFRIEZ, P.: ''The ISDD thesaurus. Keywords relating to the non-medical use of drug dependence"; London, Institute for the Study of Drug Dependence, 1980.

International Road Research Documentation (IRRD) Thesaurus; 2nd. edition, París, OECD, 1985.

NOTA: Recomenda-se a consulta, além dessas, às obras citadas do capítulo 4 ao 10.

8

Os Tesauros Multilíngües

Nesta época em que não é mais possível adotar o isolamento em nenhuma atividade humana, torna-se necessário dispor de instrumentos com os quais se possa fazer contato com os nossos congêneres. Já não é mais possível querer viver em uma torre de marfim. Isto nos levaria ao desaparecimento por extenuação. É claro que é possível pensar em uma escolha de vida em um nível puramente individual e pessoal, porém, no plano coletivo, impõe-se a comunicação.

As fronteiras, tanto espirituais como materiais, ampliam-se de forma que devemos nos relacionar com nossos semelhantes para sobreviver. Por outro lado, o que se apresenta assim como um destino fatal de caráter obrigatório afigura-se como uma necessidade inerente ao ser humano enquanto um indivíduo vivo. Em outros trabalhos, fizemos referência à informação como base e meio fundamental de apreensão e conhecimento de nosso mundo exterior. A informação que nos chega também sai de nós, como instrumento de comunicação.

O ser humano foi feito para viver em sociedade. Isolado não pode bastar-se a si mesmo, ainda que as novelas ou casos muito excepcionais desejem nos demonstrar o contrário. E, mesmo assim, o protagonista comunica-se com os animais ou com outros elementos da natureza.

Sim, o homem precisa comunicar-se. Sempre foi assim, quanto mais hoje, quando o mundo tornou-se menor por causa da velocidade. Velocidade no transporte, nas telecomunicações – nas formas de vida. Assim, pois, e dado que nós todos nos expressamos na mesma língua, necessitamos das traduções e dos instrumentos que nos ajudem a realizá-las para melhor nos entendermos uns com os outros.

Entre esses instrumentos figuram os tesauros multilíngües, dos quais os bilíngües são um caso particular. Iremos dedicar-lhes, deste modo, o espaço que lhes é correspondente na obra dedicada aos tesauros. Toma-se como base as recomendações da Unesco e a norma ISO 5964, de 1985, em sua versão inglesa. Também recomenda-se, entre outras publicações, ler as instruções preliminares do tesauro Eudised em sua versão espanhola, de 1984.

DEFINIÇÕES DE TESAURO MULTILÍNGÜE

Todas as publicações citadas tomam como ponto de partida as normas e obras referentes aos tesauros monolíngües. Estabelecem que um

tesauro multilíngüe
é um caso particular de

> um tesauro monolíngüe ao qual foram acrescentadas as equivalências dos termos em um ou mais idiomas.

As diretrizes da Unesco explicitam a definição, considerando-a dos pontos de vista de sua função e de sua estrutura.

Segundo sua função:

1. É um instrumento usado em *translating* = transladar a linguagem natural dos documentos para outro "sistema lingüístico" mais restrito, usado com fins documentários.

   Diferencia-se de um tesauro monolíngüe, por possuir a habilidade de proporcionar a *translation* = tradução = translado em mais de uma língua.

   NOTA: Já vimos aqui como a palavra *translation* pode ter duas interpretações diferentes em espanhol. É uma prova das dificuldades encontradas na elaboração deste tipo de tesauros.

Segundo sua estrutura:

2. É um vocabulário controlado e dinâmico de termos relacionados semântica e genericamente, que cobre um campo determinado do conhecimento. Os conceitos são representados por termos em mais de uma língua.

   Observe-se que estas definições são as mesmas dadas para os tesauros monolíngües, aos quais se acrescenta que os termos são representados em mais de uma língua. Na bibliografia consultada, não foi encontrada uma definição específica conveniente.

   Em minhas notas de aula, costumo apresentar uma definição que faz referência à Teoria de Sistemas:

3. É um sistema aberto, evolutivo e dinâmico no tempo, no qual os nódulos – termos – que o compõem encontram-se em relação dialética por meio de vetores de equivalência, hierárquica e de associação, onde os idiomas em que se representam os termos encontram-se também em inter-relação dialética, de forma que se facilita a comunicação até fora do sistema.

Nesta definição, as inter-relações dialéticas referem-se a relações entre os termos e com os conceitos que representam, retirados dos documentos em sua indexação, ou usados para localizá-los no processo de recuperação da informação (figura 8.1).

PECULIARIDADES DOS TESAUROS MULTILÍNGÜES

Estudando as definições anteriores, deduz-se que a questão primordial neste tipo de tesauros é encontrar as equivalências dos termos em um idioma com seus correspondentes em outro ou outros idiomas.

Pode-se estabelecer uma divisão entre

- tesauros bilíngües,
- tesauros multilíngües,

que não pressupõe mais que uma diferença do número de línguas presentes e que pode causar efeito no momento de seu uso, conforme as necessidades do caso.

Os tesauros multilíngües podem ser ordenados nas mesmas classes que os monolíngües e devem cumprir as mesmas condições. Possuem uma vantagem adicional, já que podem ser usados como simples dicionários, onde acrescentam informações muito úteis no momento de escolher um termo de um idioma equivalente em outro ou outros. Não obstante, é conveniente estabelecer as diferenças entre um tesauro multilíngüe e um dicionário.

Em um *tesauro multilíngüe*

- os vocábulos que compõem encontram-se relacionados entre si, genérica, semântica e sintaticamente.

Podem apresentar-se em forma alfabética, facetada ou gráfica e em duas ou três dessas formas ao mesmo tempo.

Em um *dicionário,*

- os vocábulos que compõem não apresentam tais relações, embora possam trazer definições ou notas de aplicação.

  Normalmente, apresentam-se em ordem alfabética. A representação sistemática é pouco freqüente.

Parece ser tão fácil encontrar as equivalências de uns termos com outros, porém este é precisamente o ponto onde se localiza a dificuldade. São necessários sérios conhecimentos lingüísticos, terminológicos, domínio de idiomas e do assunto abordado. Estas condições se apresentam raras vezes concentradas em uma só pessoa. Motivo pelo qual, geralmente, recorre-

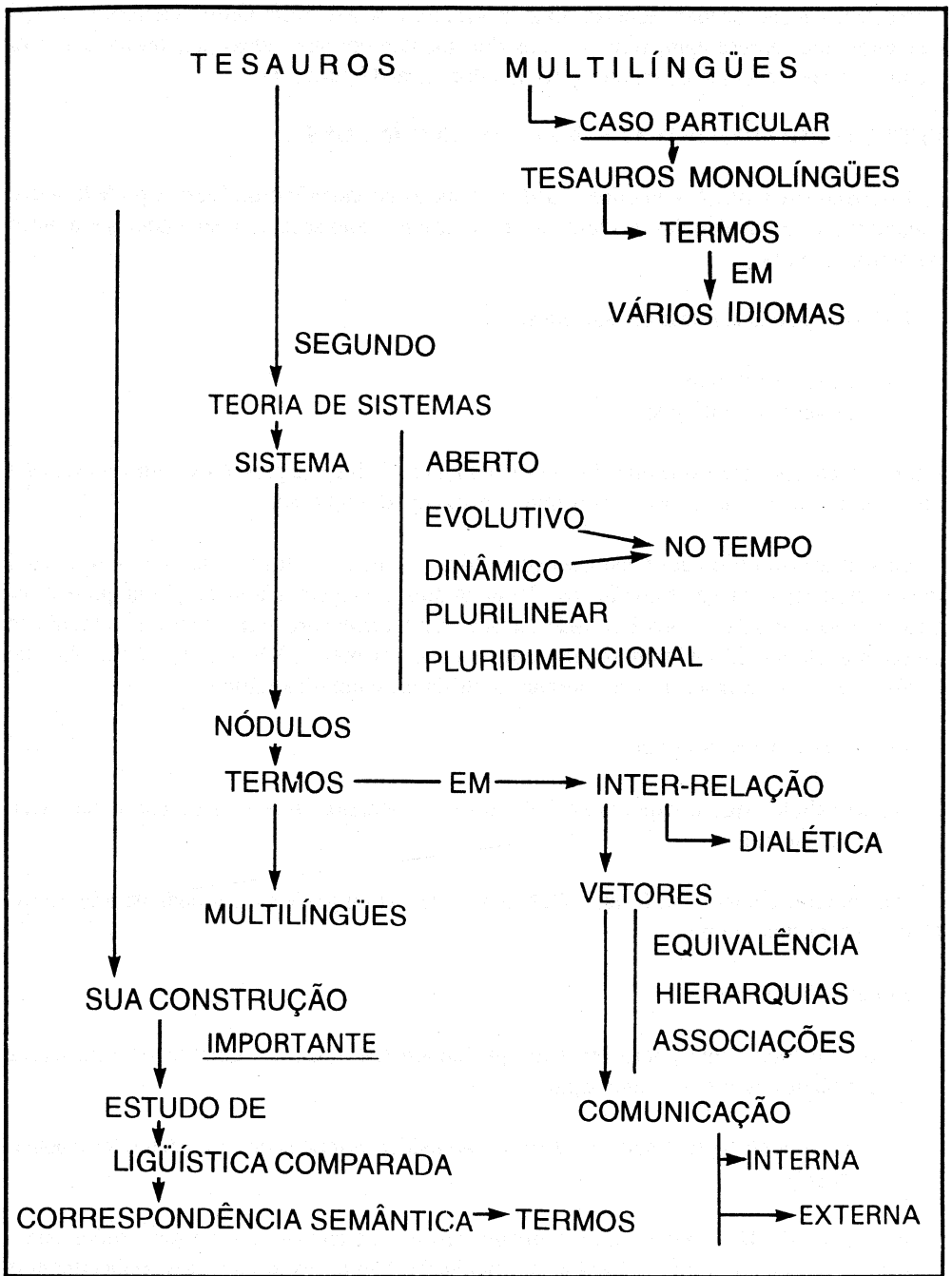
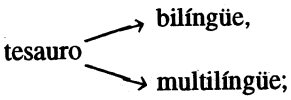

TESAUROS
MULTILÍNGÜES
CASO PARTICULAR
TESAUROS MONOLÍNGÜES
TERMOS
EM
VÁRIOS IDIOMAS
SEGUNDO
TEORIA DE SISTEMAS
SISTEMA
ABERTO
EVOLUTIVO
DINÂMICO
NO TEMPO
PLURILINEAR
PLURIDIMENCIONAL
NÓDULOS
TERMOS
EM
INTER-RELAÇÃO
DIALÉTICA
MULTILÍNGÜES
VETORES
EQUIVALÊNCIA
HIERARQUIAS
ASSOCIAÇÕES
SUA CONSTRUÇÃO
IMPORTANTE
ESTUDO DE
LIGÜÍSTICA COMPARADA
CORRESPONDÊNCIA SEMÂNTICA
TERMOS
COMUNICAÇÃO
INTERNA
EXTERNA

Figura 8.1.

se à formação de grupos de especialistas para que trabalhem em equipe. É muito desejável o grupo ser formado por pessoas provenientes de diferentes países, pois, desse modo, são trazidas opiniões mais adequadas ao uso das línguas em seus respectivos países, do que quando o grupo é composto por especialistas de diferentes procedências, mas que habitam no mesmo país.

Resumindo, as primeiras condições que devem ser consideradas no momento de confeccionar um tesauro deste tipo podem ser formuladas da seguinte maneira:

1. deve cumprir as mesmas condições que um tesauro monolíngüe;
2. pode-se construir as mesmas classes;
3. estabelecem-se os mesmos tipos de relações;
4. devem ser consideradas as diferenças que existem com um dicionário;
5. não obstante, podem ser utilizados como um dicionário;
6. deve se considerar se é necessário um



7. para a sua elaboração, será constituído um grupo de especialistas em nível internacional, oriundos de diferentes países;
8. do grupo de trabalho deverão participar lingüistas, terminólogos, tradutores e especialistas no assunto sobre o qual será construído o tesauro;
9. o grupo de trabalho será constituído em caráter permanente, pois é preciso acompanhar as modificações surgidas e fazer a manutenção do tesauro.

A ESTRUTURAÇÃO

Quando se desejar estruturar um tesauro multilíngüe e se procurar os termos correspondentes de um idioma em vários outros, é preciso considerar os seguintes pontos:

- o conhecimento profundo das línguas com as quais se deve trabalhar;
- o nível de preferência de um idioma, ou idiomas, em relação aos demais;
- as relações de semelhança entre os termos que constituirão o tesauro.

Em relação ao primeiro ponto, é imprescindível conhecer com profundidade as línguas com as quais se trabalha. Para isso, é conveniente saber que estas podem se apresentar dos seguintes tipos:

- flexíveis,
  como o latim.

  As línguas flexivas contam com declinações e conjugações completas. As raízes das palavras e as terminações são as que estabelecem as formas dos substantivos e dos verbos.

- aglutinantes,

  como o turco, o finlandês e o húngaro, onde as raízes das palavras e as terminações são palavras isoladas que logo se unem.

- monossilábicas,

  como o chinês ou o japonês.

- analíticas,

  como o inglês, que conserva vestígios de declinações e conjugações e permite formar palavras unindo-se duas ou mais delas.

Esses conhecimentos serão importantes no momento de saber em que casos é possível separar as palavras em suas partes e como se deve fazer para que os significados dos conceitos que representam não variem. Também são importantes no momento em que se deve unir os termos para obter a correspondência idiomática que se pretende.

A outra questão, relativa ao segundo ponto e que deve ser considerada, é o nível de preferência de um idioma em relação aos demais. Se for dada a um idioma a categoria principal, sendo os idiomas restantes que devem figurar no tesauro considerados somente elementos auxiliares, todos os termos pertencentes ao primeiro idioma são os descritores – termos preferidos – e figuram com essa categoria nas publicações impressas ou nas bases de dados de um computador. Os termos dos outros idiomas serão tratados com a categoria de sinônimos. No caso em que todos os idiomas devam receber o mesmo tratamento, cada termo, em seu idioma, conservará o nível de relação correspondente, quer dizer, será desenvolvido completamente, como descritor ou como termo equivalente.

RELAÇÃO DE SEMELHANÇA ENTRE OS TERMOS

A terceira questão, exposta no título anterior, refere-se à relação de semelhança entre os termos dos diferentes idiomas que constituem o tesauro objeto de estudo.

Sabemos que nem sempre uma palavra de um idioma corresponde exatamente à outra de outro idioma. Mesmo que haja correspondência, pode acontecer que se escrevam de outra maneira, ou que uma seja composta por palavras com significado próprio.

Por tais motivos, pode ser útil registrar aqui os tipos de relação de semelhança que podem se apresentar.

1. Semelhança exata.

Trata-se de palavras que possuem o mesmo significado e a mesma ortografia. Podem apresentar diferença fonética.

Exemplo:

Espanhol	Inglês
hospital	hospital

2. Semelhança semântica/não ortográfica,

Considera-se esta semelhança quando as palavras possuem o mesmo significado, porém se escrevem de modo diferente.

Exemplo:

Espanhol	Inglês
patrocinador	sponsor

3. Semelhança ortográfica/não semântica,

A fonia pode ser diferente.

Trata-se de palavras escritas da mesma maneira – ainda que apresentem diferenças, mas que não possuem o mesmo significado. A fonia também pode diferir de uma ou outra palavra.

Exemplo:

Espanhol	Inglês
librería	library

4. Semelhança parcial,

Aqui, as palavras não correspondem exatamente umas às outras. Uma delas abrange um conteúdo significativo maior que a outra.

Exemplo:

Espanhol	Inglês
biblioteca	library

5. Semelhança simples ou múltipla,

Nestes casos, as palavras não possuem correspondência exata, mas se pode conseguir isso unindo-se dois ou mais vocábulos. É requerida uma análise profunda dos idiomas com os quais se trabalha, o que pressupõe possuir bons conhecimentos científicos e lingüísticos.

Norma geral: admitem-se três situações diferentes neste tipo de semelhanças.

5.1. A palavra de um idioma é expressa no outro idioma por várias delas.

Exemplo:

Espanhol	Inglês
librería	bookshop (book + shop)

5.2. Há casos em que, para expressar um conceito, deve-se utilizar termos que, separados, possuem outro significado.

Exemplo:

Espanhol	Inglês
Energia solar	solar heating

. *energía* não significa o mesmo que *heating*, porém, quando se une à expressão solar, passa a possuir o mesmo significado.

5.3. A correspondência de um termo é obtida com dois ou mais termos do outro idioma que não se podem unir.

Devem ser expressos separadamente.

Exemplo:

Inglês	Espanhol
congressasistance	congresso + asistência

6. Semelhança explicativa.

Aparece quando não existe uma correspondência exata e, para compreender uma delas, deve-se acrescentar uma nota explicativa.

Exemplo:

Espanhol	Inglês
filing	ordenar ou sortear (ao mesmo tempo, classificando segundo um método)

7. Não- semelhança

Trata-se dos casos em que uma palavra não existe no outro idioma e deve-se compor várias palavras para conseguir a correspondência, pois, do contrário, ela deve ser utilizada em seu próprio idioma.

Exemplo:

Espanhol	Inglês
software	logical? programa lógico? programa de computador? *software*?

Deve-se ter muito cuidado com este tipo de palavras que não apresentam correspondência ou tradução. Pode tratar-se de uma palavra nova, referente a um conceito de aparecimento recente e que ainda não foi estudada suficientemente no outro idioma no qual se deseja encontrar uma equivalência, para fixar-lhe um termo. Com o tempo, os terminólogos ou os lingüistas acabam por encontrar o vocábulo apropriado. Pode ocorrer, também, que se imponha o termo "estrangeiro" como um neologismo e se generalize seu uso. Também pode ocorrer o caso em que o termo "estrangeiro" se nacionalize, mudando, talvez, a ortografia de modo a adaptá-la a sua fonética. Finalmente, aparece o caso de não existir correspondência entre os termos, porque não existe ou não se concebe tal conceito em um idioma ou em outro. Cada povo possui sua mentalidade e seus costumes.

Como conseqüência de tudo isso, as palavras de um idioma que não apresentam semelhanças com outra ou outras de um idioma diferente devem ser submetidas a estudo e observação, para que se possa ir detectando sua evolução.

Como é fácil compreender, o uso dos computadores diminui consideravelmente a tarefa de estruturar um tesauro multilíngüe. A tarefa ficará ainda mais facilitada quando começarem a ser operacionalizados os sistemas especialistas para técnicas de inteligência artificial, que se encontram em fase de estudos. Até agora, está sendo dada pouca informação sobre as pesquisas realizadas neste campo. Talvez isso se deva à existência de interesses comerciais no assunto. Talvez sim, talvez não.

TÉCNICAS DE CONSTRUÇÃO

Neste segmento, faz-se referência às técnicas de construção quanto à sua diversidade idiomática, já que a confecção do tesauro como tal será feita seguindo os mesmos procedimentos descritos para os tesauros monolíngües.

Assim, pois, a questão está em seguir qualquer procedimento dos que são apropriados para realizar traduções. Nestes casos, considerar-se-á uma língua "origem" e as outras como "resultado" ou "produto final".

Quando se trata de realizar traduções múltiplas,costuma ser aconselhável dispor de um veículo intermediário, comum para todas elas, que sirva como referência no momento em que se necessita procurar as equivalências de alguns termos com outros. Como instrumento intermediário, pode-se escolher uma linguagem artificial, planejada expressamente para esse fim, formada por palavras da linguagem natural, por símbolos preestabelecidos ou por números, estruturados de acordo com normas prefixadas. Na maioria dos casos, utilizam-se números, por estes apresentarem várias vantagens, tais como as seguintes:

1. é possível escolher a quantidade de posições – cifras – que seja conveniente;
2. apresentam caráter internacinal;
3. podem ser compostos do modo que for necessário, variando as posições das cifras;
4. as cifras admitem todas as posições e sua repetição é ilimitada;
5. são legíveis por computador e armazenáveis em qualquer suporte de memória;
6. seu significado não está ligado a conceitos referentes a seres ou objetos abstratos ou concretos.

Realmente, uma quantidade considerável de sistemas de tradução automática utiliza notações numéricas para conectar os idiomas com os quais trabalha.

O procedimento é bem simples. Consiste em atribuir um número a cada palavra e sempre o mesmo número, para as mesmas palavras nos diferentes idiomas. Tomemos um exemplo arbitrário:

biblioteca =

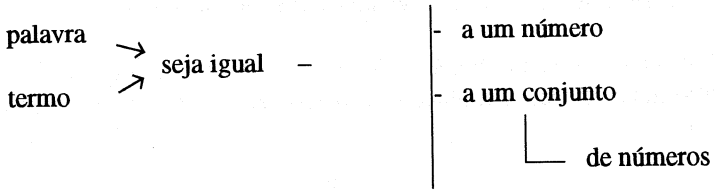
library = 46

bibliotek =

O número 46 é o meio de ligação dessa palavra em cada um dos três idiomas.

Sabemos que nem sempre é tão fácil encontrar equivalências de algumas palavras com

outras. Se for preciso unir dois conceitos, por exemplo, para formar o equivalente em outra língua, será necessário poder unir, também, os números. Do que se conclui que



O planejamento de uma maneira de estruturar os números de forma que os mesmos façam referência a características lingüísticas, propriedades dos conceitos etc. é bastante aconselhável e facilitará o trabalho. Definitivamente, do que se trata é que o número – meio – deve conter informação sobre o conceito ou os termos que representa. Isto pode requerer um processo muito trabalhoso, no qual o importante é estabelecer os princípios para a configuração dos diferentes números e atribuir a cada palavra – ou conceito unidade – aquele que lhe corresponde. O computador fará o resto, combinando-se conforme seja necessário. No anexo deste capítulo, inclui-se um artigo escrito há alguns anos entitulado *Confección de un Thesauro Estructurado para Química*, onde se explica este método detalhadamente e citam-se exemplos.

APRESENTAÇÃO DE TESAUROS MULTILÍNGÜES

Os tesauros deste tipo podem se apresentar, como acontece habitualmente entre os monolíngües, em forma alfabética ou sistemática, com seu índice alfabético correspondente. Não costuma ser comum encontrar apresentações gráficas, por resultarem extremamente complicadas, sem trazer informação adicional.

Também não é comum que um tesauro multilíngüe apareça em seções, com todos e cada um de seus termos ao mesmo tempo, registrados em todos os idiomas para os quais tenha sido confeccionado. Geralmente, dividem-se em partes por idiomas, e cada um delas é apresentada separadamente. Isto quer dizer que se adota um idioma, o espanhol, por exemplo, como principal, que se desenvolve em sua totalidade, como se fosse um tesauro monolíngüe. Debaixo de cada termo, colocam-se seus correspondentes nos outros idiomas nos quais foi confeccionado o tesauro multilíngüe em questão. Em outra seção independente, será desenvolvido o francês, por exemplo, anotando-se as equivalências nos outros idiomas, e assim sucessivamente. Nas figuras 8.2 e 8.3, reproduzem-se, respectivamente, uma página da seção alfabética da versão em espanhol do tesauro Eudised - Tesauro Multilíngüe para el Tratamiento de la Información en Materia de Educación, de 1984, e do Tesauro Spines, elaborado pela Unesco, em 1985.

Na realidade, um tesauro multilíngüe completo se comporta como um grande macrotesauro, no qual cada seção, para cada idioma, pressupõe ser um microtesauro que pode ser utilizado independentemente. Essa classe de tesauro existe somente como suporte de memória de computador. Isto significa que existe a possibilidade de serem publicados impressos em papel, porém esta é dificultosa e pouco prática, razões pelas quais não costuma ser feita. Pelo contrário,

publicam-se os microtesauros, com seções independentes, para cada idioma. É possível reunir o cojunto de seções dedicadas aos diferentes idiomas, obtendo-se, desse modo, o macrotesauro multilíngüe completo, em partes, ou fascículos, ou volumes separadamente.

Quando se desejar editar um tesauro multilíngüe, serão seguidas as mesmas recomendações propostas para um tesauro monolíngüe. A publicação constará das mesmas partes e estará sujeita às mesmas normas.

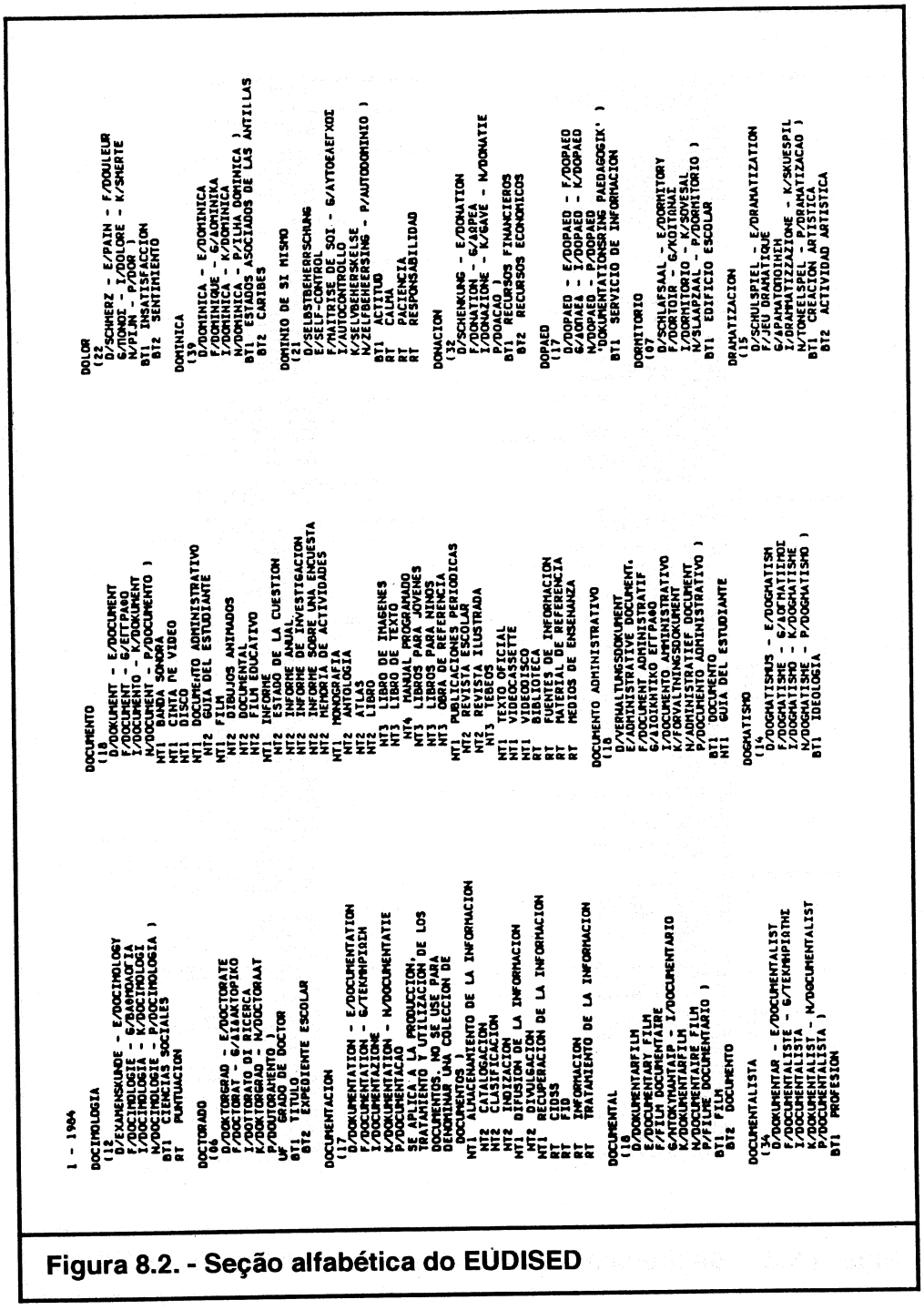
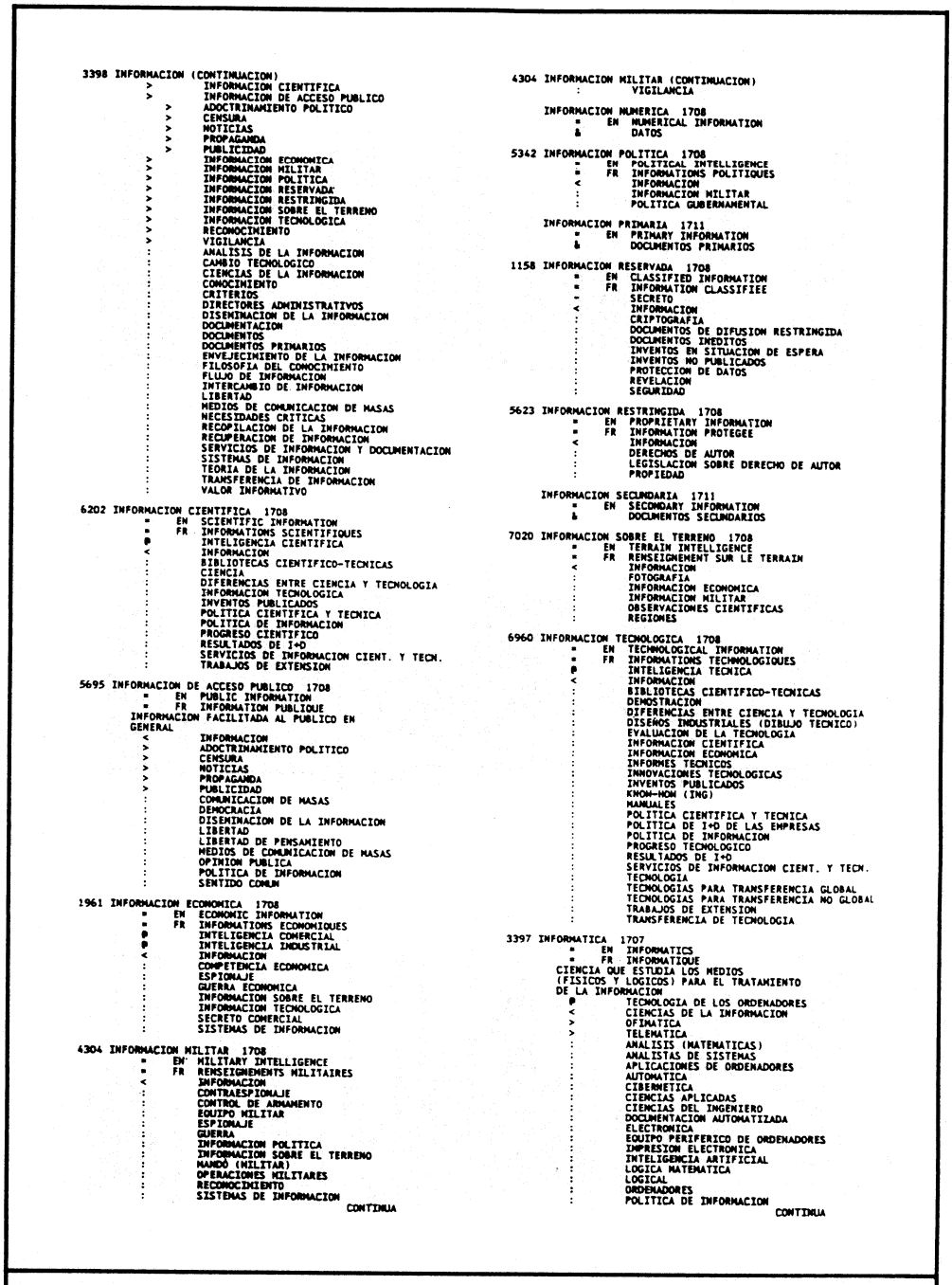


Figura 8.2. - Seção alfabética do EUDISED



```
3398 INFORMACION (CONTINUACION)
          >       INFORMACION CIENTIFICA
          >       INFORMACION DE ACCESO PUBLICO
            >     ADOCTRINAMIENTO POLITICO
            >     CENSURA
            >     NOTICIAS
            >     PROPAGANDA
            >     PUBLICIDAD
          >       INFORMACION ECONOMICA
          >       INFORMACION MILITAR
          >       INFORMACION POLITICA
          >       INFORMACION RESERVADA
          >       INFORMACION RESTRINGIDA
          >       INFORMACION SOBRE EL TERRENO
          >       INFORMACION TECNOLOGICA
          >       RECONOCIMIENTO
          >       VIGILANCIA
                  ANALISIS DE LA INFORMACION
                  CAMBIO TECNOLOGICO
                  CIENCIAS DE LA INFORMACION
                  CONOCIMIENTO
                  CRITERIOS
                  DIRECTORES ADMINISTRATIVOS
                  DISEMINACION DE LA INFORMACION
                  DOCUMENTACION
                  DOCUMENTOS
                  DOCUMENTOS PRIMARIOS
                  ENVEJECIMIENTO DE LA INFORMACION
                  FILOSOFIA DEL CONOCIMIENTO
                  FLUJO DE INFORMACION
                  INTERCAMBIO DE INFORMACION
                  LIBERTAD
                  MEDIOS DE COMUNICACION DE MASAS
                  NECESIDADES CRITICAS
                  RECOPILACION DE LA INFORMACION
                  RECUPERACION DE INFORMACION
                  SERVICIOS DE INFORMACION Y DOCUMENTACION
                  SISTEMAS DE INFORMACION
                  TEORIA DE LA INFORMACION
                  TRANSFERENCIA DE INFORMACION
                  VALOR INFORMATIVO

6202 INFORMACION CIENTIFICA  1708
          =     EN  SCIENTIFIC INFORMATION
          =     FR  INFORMATIONS SCIENTIFIQUES
          @         INTELIGENCIA CIENTIFICA
          <         INFORMACION
                    BIBLIOTECAS CIENTIFICO-TECNICAS
                    CIENCIA
                    DIFERENCIAS ENTRE CIENCIA Y TECNOLOGIA
                    INFORMACION TECNOLOGICA
                    INVENTOS PUBLICADOS
                    POLITICA CIENTIFICA Y TECNICA
                    POLITICA DE INFORMACION
                    PROGRESO CIENTIFICO
                    RESULTADOS DE I+D
                    SERVICIOS DE INFORMACION CIENT. Y TECN.
                    TRABAJOS DE EXTENSION

5695 INFORMACION DE ACCESO PUBLICO  1708
          =     EN  PUBLIC INFORMATION
          =     FR  INFORMATION PUBLIQUE
        INFORMACION FACILITADA AL PUBLICO EN
        GENERAL
          <         INFORMACION
          >         ADOCTRINAMIENTO POLITICO
          >         CENSURA
          >         NOTICIAS
          >         PROPAGANDA
          >         PUBLICIDAD
                    COMUNICACION DE MASAS
                    DEMOCRACIA
                    DISEMINACION DE LA INFORMACION
                    LIBERTAD
                    LIBERTAD DE PENSAMIENTO
                    MEDIOS DE COMUNICACION DE MASAS
                    OPINION PUBLICA
                    POLITICA DE INFORMACION
                    SENTIDO COMUN

1961 INFORMACION ECONOMICA  1708
          =     EN  ECONOMIC INFORMATION
          =     FR  INFORMATIONS ECONOMIQUES
          @         INTELIGENCIA COMERCIAL
          @         INTELIGENCIA INDUSTRIAL
          <         INFORMACION
                    COMPETENCIA ECONOMICA
                    ESPIONAJE
                    GUERRA ECONOMICA
                    INFORMACION SOBRE EL TERRENO
                    INFORMACION TECNOLOGICA
                    SECRETO COMERCIAL
                    SISTEMAS DE INFORMACION

4304 INFORMACION MILITAR  1708
          =     EN  MILITARY INTELLIGENCE
          =     FR  RENSEIGNEMENTS MILITAIRES
          <         INFORMACION
                    CONTRAESPIONAJE
                    CONTROL DE ARMAMENTO
                    EQUIPO MILITAR
                    ESPIONAJE
                    GUERRA
                    INFORMACION POLITICA
                    INFORMACION SOBRE EL TERRENO
                    MANDO (MILITAR)
                    OPERACIONES MILITARES
                    RECONOCIMIENTO
                    SISTEMAS DE INFORMACION
                                          CONTINUA

4304 INFORMACION MILITAR (CONTINUACION)
                    VIGILANCIA

     INFORMACION NUMERICA  1708
          =     EN  NUMERICAL INFORMATION
          &         DATOS

5342 INFORMACION POLITICA  1708
          =     EN  POLITICAL INTELLIGENCE
          =     FR  INFORMATIONS POLITIQUES
          <         INFORMACION
                    INFORMACION MILITAR
                    POLITICA GUBERNAMENTAL

     INFORMACION PRIMARIA  1711
          =     EN  PRIMARY INFORMATION
          &         DOCUMENTOS PRIMARIOS

1158 INFORMACION RESERVADA  1708
          =     EN  CLASSIFIED INFORMATION
          =     FR  INFORMATION CLASSIFIEE
          -         SECRETO
          <         INFORMACION
                    CRIPTOGRAFIA
                    DOCUMENTOS DE DIFUSION RESTRINGIDA
                    DOCUMENTOS INEDITOS
                    INVENTOS EN SITUACION DE ESPERA
                    INVENTOS NO PUBLICADOS
                    PROTECCION DE DATOS
                    REVELACION
                    SEGURIDAD

5623 INFORMACION RESTRINGIDA  1708
          =     EN  PROPRIETARY INFORMATION
          =     FR  INFORMATION PROTEGEE
          <         INFORMACION
                    DERECHOS DE AUTOR
                    LEGISLACION SOBRE DERECHO DE AUTOR
                    PROPIEDAD

     INFORMACION SECUNDARIA  1711
          =     EN  SECONDARY INFORMATION
          &         DOCUMENTOS SECUNDARIOS

7020 INFORMACION SOBRE EL TERRENO  1708
          =     EN  TERRAIN INTELLIGENCE
          =     FR  RENSEIGNEMENT SUR LE TERRAIN
          <         INFORMACION
                    FOTOGRAFIA
                    INFORMACION ECONOMICA
                    INFORMACION MILITAR
                    OBSERVACIONES CIENTIFICAS
                    REGIONES

6960 INFORMACION TECNOLOGICA  1708
          =     EN  TECHNOLOGICAL INFORMATION
          =     FR  INFORMATIONS TECHNOLOGIQUES
          @         INTELIGENCIA TECNICA
          <         INFORMACION
                    BIBLIOTECAS CIENTIFICO-TECNICAS
                    DEMOSTRACION
                    DIFERENCIAS ENTRE CIENCIA Y TECNOLOGIA
                    DISEÑOS INDUSTRIALES (DIBUJO TECNICO)
                    EVALUACION DE LA TECNOLOGIA
                    INFORMACION CIENTIFICA
                    INFORMACION ECONOMICA
                    INFORMES TECNICOS
                    INNOVACIONES TECNOLOGICAS
                    INVENTOS PUBLICADOS
                    KNOW-HOW (ING)
                    MANUALES
                    POLITICA CIENTIFICA Y TECNICA
                    POLITICA DE I+D DE LAS EMPRESAS
                    POLITICA DE INFORMACION
                    PROGRESO TECNOLOGICO
                    RESULTADOS DE I+D
                    SERVICIOS DE INFORMACION CIENT. Y TECN.
                    TECNOLOGIA
                    TECNOLOGIAS PARA TRANSFERENCIA GLOBAL
                    TECNOLOGIAS PARA TRANSFERENCIA NO GLOBAL
                    TRABAJOS DE EXTENSION
                    TRANSFERENCIA DE TECNOLOGIA

3397 INFORMATICA  1707
          =     EN  INFORMATICS
          =     FR  INFORMATIQUE
        CIENCIA QUE ESTUDIA LOS MEDIOS
        (FISICOS Y LOGICOS) PARA EL TRATAMIENTO
        DE LA INFORMACION
          @         TECNOLOGIA DE LOS ORDENADORES
          <         CIENCIAS DE LA INFORMACION
          >         OFIMATICA
          >         TELEMATICA
                    ANALISIS (MATEMATICAS)
                    ANALISTAS DE SISTEMAS
                    APLICACIONES DE ORDENADORES
                    AUTOMATICA
                    CIBERNETICA
                    CIENCIAS APLICADAS
                    CIENCIAS DEL INGENIERO
                    DOCUMENTACION AUTOMATIZADA
                    ELECTRONICA
                    EQUIPO PERIFERICO DE ORDENADORES
                    IMPRESION ELECTRONICA
                    INTELIGENCIA ARTIFICIAL
                    LOGICA MATEMATICA
                    LOGICAL
                    ORDENADORES
                    POLITICA DE INFORMACION
                                          CONTINUA
```

Figura 8.3. - Seção alfabética do tesauro SPINES da UNESCO

ANEXO
CONFECÇÃO DE UM TESAURO ESTRUTURADO PARA QUÍMICA*

INTRODUÇÃO

Mediante o estudo e comparação dos diferentes tesauros existentes, que estão se tornando numerosos, pensou-se que talvez faltasse um sistema no qual as palavras componentes desses tesauros estivessem estruturadas de modo que se poderia encontrar sem dificuldade relações de umas com as outras. Também, notou-se a falta de tesauros facilmente utilizáveis para traduções automáticas e para o armazenamento das palavras-chave em fitas magnéticas, com o emprego de determinadas linguagens próprias dos computadores eletrônicos. Esta falta mostra-se mais acentuada no caso da química, considerada em sentido geral.

Por outro lado, existem sistemas de classificação e vocabulários especiais que dão aos conceitos hábeis sinais de identificação, tanto para poder ordenar tais conceitos, como para lacalizá-los ao longo de todos os processos de análise, armazenamento e recuperação da informação. De todos eles, é conhecida a Classificação Decimal Universal, por exemplo. No campo da química, também é conhecido o sistema utilizado pelo *Chemical Abstracts Service*, que acompanha cada composto químico de um número de identificação. Existe o *Euroatom Thesaurus*, no qual as palavras-chave levam um número, também, de identificação. Podeiríamos citar algum outro caso, aplicável ao mesmo tempo a assuntos tais como economia, história etc. Como exemplo, aparece o elaborado pelo então chamado Instituto de Cultura Hispânica, atualmente Centro de Cooperação Ibero-americana, que estrutura os números de identificação de tal forma que, em uma simples leitura, pode-se saber a que tipo de conceitos se referem.

Como dissemos, refletindo sobre todos esses métodos, principalmente sobre a Classificação Decimal Universal, surgiu a idéia de tentar pesquisar um método lógico e instrutivo para identificar os termos de um tesauro, para o conjunto da química, de modo que fosse possível reconhecer e também traduzir esses termos facilmente, sobretudo com o objetivo da utilização de computadores.

A Classificação Decimal Universal utiliza números e símbolos estruturados de maneira lógica e intuitiva, o que faz com que nosso trabalho possa parecer supérfluo. Contudo, devemos considerar que tais números apresentam certos inconvenientes, como, por exemplo, de tornarem-se demasiadamente extensos, para assuntos muito específicos; outras vezes, são pouco extensas, e, em outras, apresentam demasiadas cifras, pois a lógica utilizada não é aplicável a todos os campos do conhecimento. Finalmente, todos conhecem os inconvenientes deste método de classificação, o qual, por outro lado, apresenta numerosas vantagens, sendo uma delas a fácil tradução dos conceitos e vocábulos de um idioma para outro.

(*) Actas do Congresso REUNIBER. REUNIBER 78. CENIDOC. Madrid, setembro 1978. p. 297-318.

Os outros tesauros e métodos aqui mencionados para a química apresentam, também, alguns inconvenientes ou limitações. Por exemplo, *Euroatom-Thesaurus* refere-se somente a este assunto determinado.

O sitema utilizado pelo *Chemical Abstracts Service* não corresponde a nenhum método, nem lógica, por dar aos compostos químicos e somente a estes um número de ordem correlativo, de acordo com o surgimento destes.

Assim, pois, como ponderamos aqui e pensando no tesauro confeccionado pelo Centro de Cooperação Ibero-Americano, empreendemos a tarefa de elaborar um sistema que pudesse ser aplicado à química em geral e que fosse capaz de atenuar alguns dos inconvenientes antes citados.

EXPOSIÇÃO DO MÉTODO

Este método consiste em acompanhar cada palavra de um número, o que, como vimos, não pressupõe novidade alguma. A novidade fundamenta-se na estrutura dada a esse número (figura A.1).

Este número é criado na base dos 10 algarismos simples e se admitem todas as posições possíveis, o que pressupõe poder dispor de uma grande quantidade de combinações.

Para a maior modalidade das cifras, admite-se que possam ser separadas por pontos (.), também colocado em qualquer posição.

Além disso, a posição de determinados algarismos no número pode indicar tipos de conceitos diferentes. Também a posição do ponto (.) pode-se utilizar com este fim. Por exemplo, os números que começam com 00, os números que possuem um 0 no meio; os números que possuam os dois primeiros algarismos separados por um ponto (.). Todos eles farão referência, como dissemos, a um tipo de palavra-chave e sempre ao mesmo.

Para o caso da química, pensou-se que será necessário dispor de uma número de seis algarismos, quando se fizer referência a um assunto de química inorgânica, ou química geral, e de um número de sete algarismos, quando se fizer referência aos compostos da química orgânica. Desta maneira, teremos uma forma de identificação para esse tipo de compostos.

As palavras-chave foram divididas em dois tipos (figura A.1), a saber:

- nomes de elementos e compostos químicos;
- assuntos gerais e específicos = palavras temáticas.

Neste último grupo, incluímos conceitos como destilação, catálise, ligações, amálgamas, reação, temperaturas etc. , quer dizer, todas aquelas palavras que não façam referência a nome de elementos, nem de compostos químicos. Essas palavras são chamadas, em alemão, *sachverhalten*; neste trabalho, iremos denominá-las "palavras temáticas".

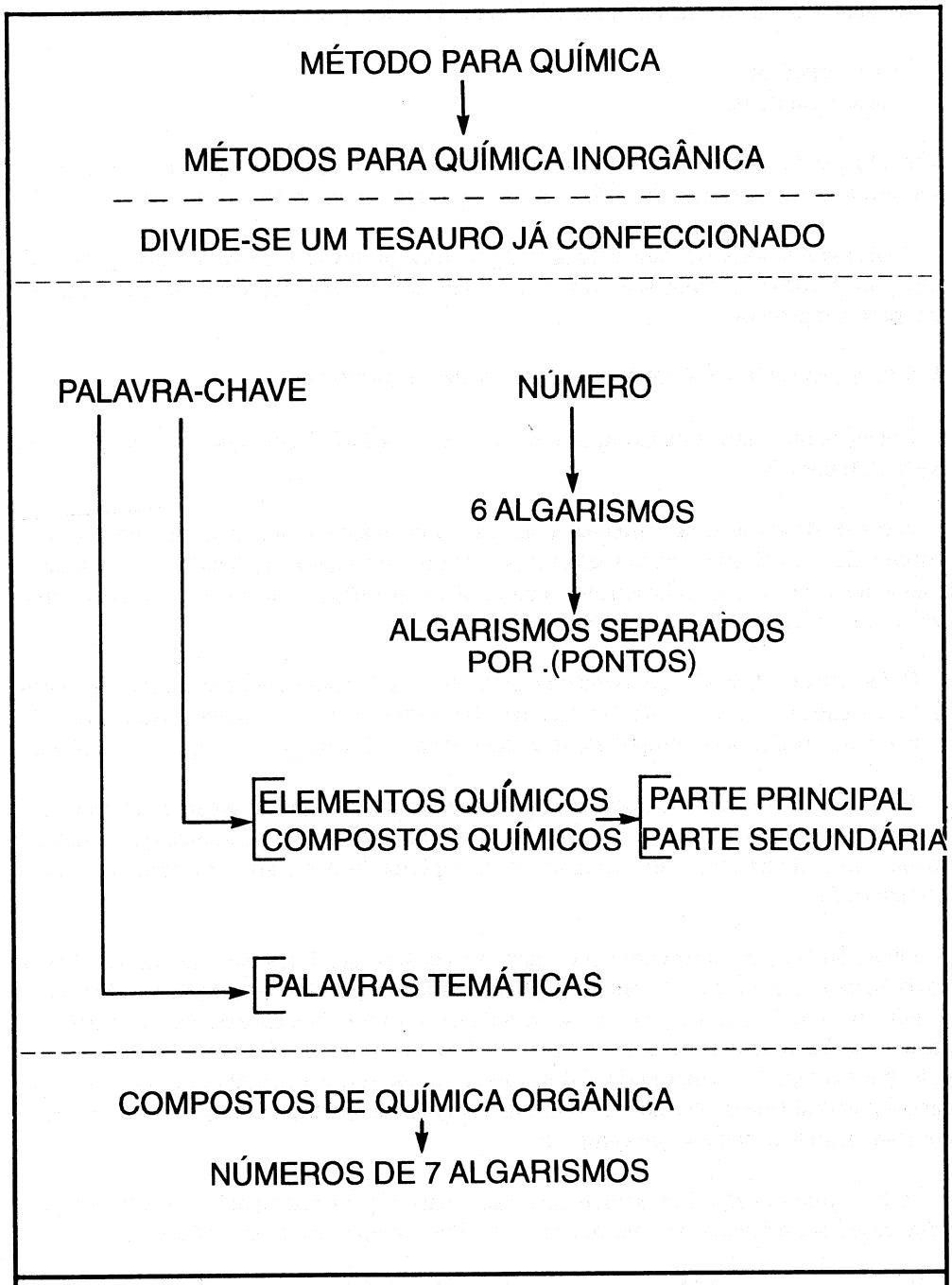


Figura A.1. - Confecção de um tesauro estruturado para química

Os nomes dos componentes químicos foram divididos, por sua vez, em

- parte principal,
- parte secundária,

de modo que no número haverá alguns algarismos comuns a uma série de conceitos e outros que variarão; mas que, por seu lado, poderão ser comuns a outra série de termos.

As palavras temáticas foram incluídas no grupo das que levam números de seis algarismos. Desejou-se, como dissemos anteriormente, os números de sete algarismos somente para os compostos orgânicos.

APLICAÇÃO DO MÉTODO À QUÍMICA INORGÂNICA

Começamos nosso trabalho aplicando-o à química inorgânica, que é o assunto mais conhecido por nós.

Partimos de tesauros já existentes, pois neles a linguagem química já está organizada e estruturada. As palavras estão ordenadas por famílias e foram considerados os sinônimos, homônomos e os termos relacionados. Parece óbvio que não se tenha desejado repetir uma tarefa já feita e que pressupõe um bom ponto de partida.

O que estamos fazendo, por o trabalho estar em vias de realizaçao, é atribuir um número de seis algarismos a cada uma das palavras-chave do tesauro e estruturar esse número de forma, como foi dito, a nos proporcionar um meio de reconhecer a palavra a que faz referência.

Dissemos que tínhamos dividido em duas partes as palavras que designam os elementos e os compostos químicos. Dado que os elementos ultrapassam 100, pensamos que podemos atribuir três algarismos para designá-los, completando o restante do número com 0 (figura A. 2).

Em nosso caso, consideramos como "parte secundária" precisamente a que faz referência aos elementos, e, como estes levam três algarismos, foram separados por um ponto (.). Portanto, os números que levam um ponto (.) separando os três primeiros algarismos designarão os elementos. Além do mais, convencionamos colocar 000 (três zeros) antes do ponto. Assim, pois, por exemplo, o número 000.124 nos indicará que se trata de um elemento. Também cabe a possibilidade de considerar os diferentes graus de oxidação; por exemplo: cúprico, cuproso, quando se trata de um composto.

Os três primeiros algarismos foram reservados para a "parte principal" do vocábulo e que, neste caso, fará referência ao tipo de composto. Por exemplo: cloratos, sulfatos etc.

Se designarmos por 182 os sulfatos, o número 182.000 fará referência ao íon sulfato.

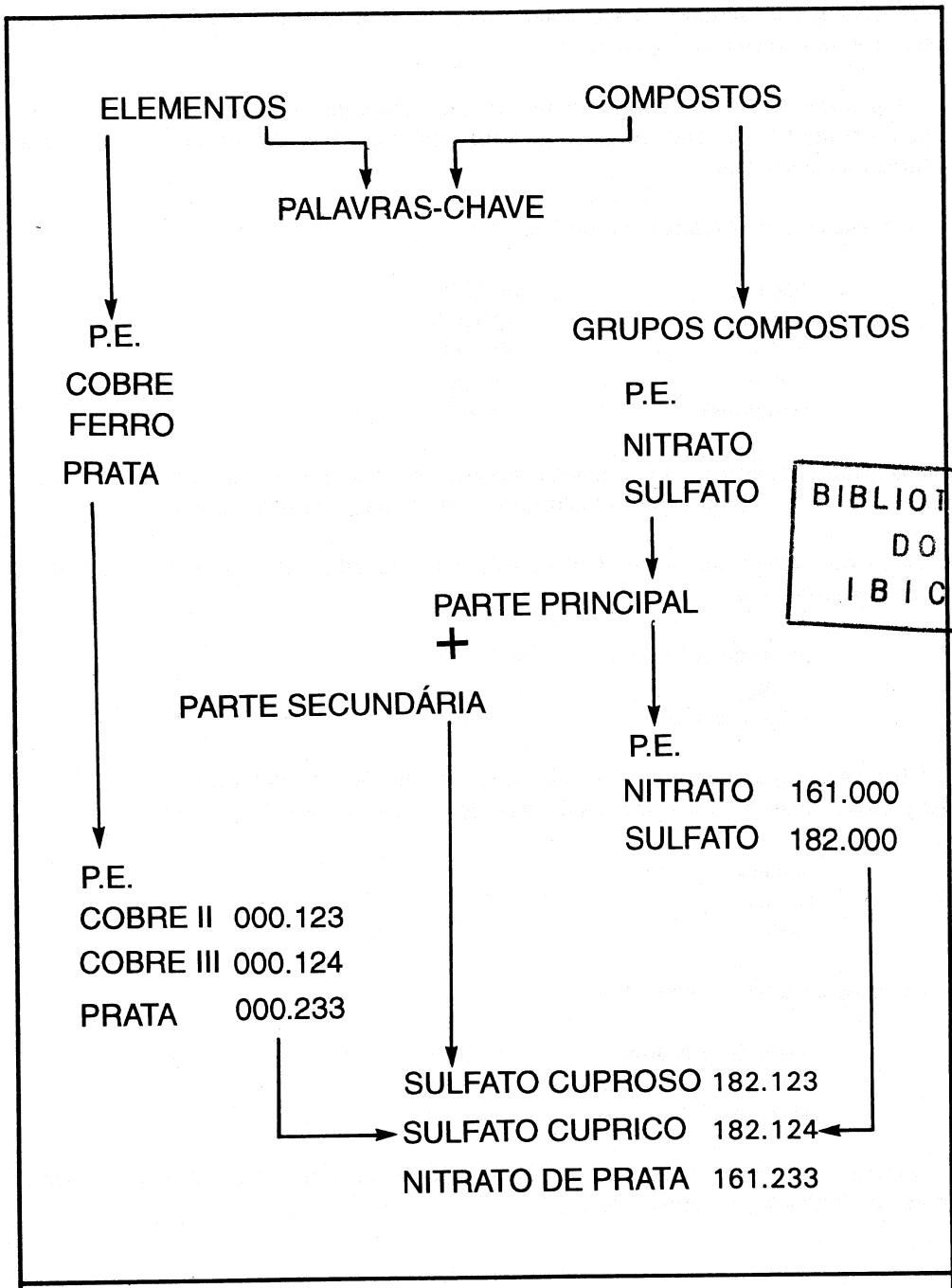


Figura A.2. -Confecção de um tesauro estruturada para química.

Se dermos o número 000.124 ao cobre com grau III de oxidação, o número 182.124 se referirá ao sulfato cúprico (figura A.2).

As palavras temáticas poderão ser reconhecidas em seguida pela falta desta estrutura. Terão o ponto (.) colocado em qualquer outra posição, menos depois aos três primeiros algarismos (figura A.3).

Suponhamos, por exemplo, que designamos:

obtenção	00.37.00
ligação	06.0.000
fusão	90.0.000
vácuo	4283.82
cristal(ais)	00.00.12

Ao se fazer a análise de um documento, deverão ser considerados estes números, os quais irão nos servir, em seguida, para a tradução, ou para a recuperação da informação.

Se tivermos um trabalho sobre "Obtenção de cristais de sulfato cúprico no vácuo, esta fase se terá convertido em

obtenção de cristais	00.37.12
vácuo	4283.82
sulfato cúprico	182.124

Observamos que com nosso método se resolve também o problema dos descritores compostos de mais de uma palavra, sobretudo quando uma delas é tão geral como:

obtenção,
destilação,
análise.

No nosso exemplo, combinamos:

obtenção cristal(ais)	00.37.00 - 00.00.12
para escrever "obtenção de cristais"	00.37.12

O mesmo poderíamos fazer com "análise gravimétria, análise por via úmida, destilação no vácuo, destilação em refluxo etc. etc.".

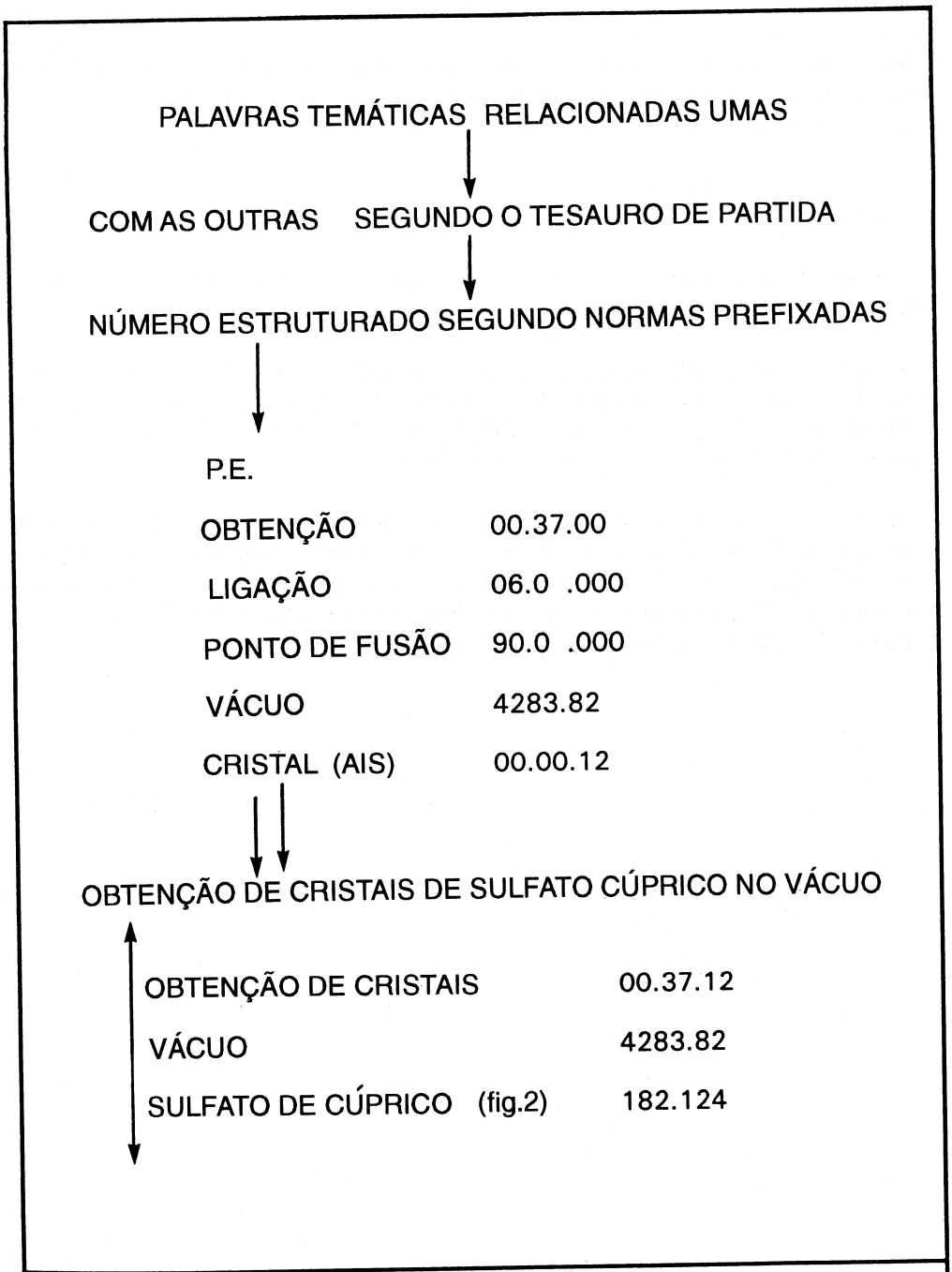


Figura A.3. - Confecção de um tesauro estruturado para química.

CONCLUSÃO

Este método para confeccinar um tesauro para a química, com numeração estruturada para seus descritores, é um método lógico e dedutivo, simples de aplicar e prático, sobretudo nos casos, cada vez mais freqüentes, de ser preciso realizar traduções automáticas.

Neste campo, presume-se, além do mais, uma simplificação, ao se tornar necessário considerar os sinais de transliteração, quando se trata de idiomas com grafias diferentes.

Este método pode, também, ser aplicado às linguagens de computador mais modernas e complicadas.

É preciso, naturalmente, colocar muita atenção à designação dos números nos diferentes conceitos e considerar todos os termos de um tesauro, sobretudo no que se refere às "palavras temáticas". É aqui que se deve estudar detidamente a estrutura do número para suas possíveis combinações e economia, conseqüentemente, de números.

Apresentou-se uma mostra deste trabalho no 2° Congresso de Euronet, que, sob o tema geral *"Como libertar as barreiras da linguagem"*, realizou-se em Luxemburgo, em maio de 1977. Além do que, é para nós uma satisfação comprovar que uma parte de nossas idéias foi utilizada no tesauro para ciências sociais e economia, que está sendo elaborado no Peru, com o objetivo de integrá-lo na rede de documentação hispano-americana.

BIBLIOGRAFIA CONSULTADA

CDU - Introdução. Qualquer edição.

FONDIN, H. La structure et le vocabulaire de l'analyse documentaire; contribution pour une mise point. Documentalist, v. 14, n. 2, p. 11-16, mars-avr. 1971.

CHOSE, A. Problems of thesaurus construction. J. Chem. Inf. Comput. Sci., v. 28, n. 4, p. 211-218, jul. 1977.

ILJON, A. Le logiciel Astute, un outil informatique pour la création, la mise à jour, l'édition et l'impression de thesauri mono et multilingues. Documentalist, v. 14, n. 4, p. 25-34, jul.août 1977.

LISTADO de descritores econômicos, sociales y políticos, 1977. Madri, Instituto de Cultura Hispânica, 1976. p. 50.

SIDORCHENKO, V.D. Thesaurus semantic structure: current state and propects of development. Auto. Doc. Math. Ling., v. 10, n. 3, p. 52-65, 1976.

SPINES - Thesaurus. 4v. Paris, UNESCO, 1977.

EUROATOM - Thesaurus. Bruxelas, European Atomic Energy Community Center for Information and Documentation, 1966. p. 91.

YANEZ, A. Colloque international de terminologie. Paris, juin 1976. Problèmes de lexicographie dans un système plurilingue de stockage et de recherche d'information. Documentalist, v. 14, n. 2, p. 25-28, mars-avr. 1977.

ZOLOTAREV, A.E. & PEREGUDOV, A.N. Struture and software for a learning IRS bases on an M-222 computer. Auto. Doc. Math. Ling., v. 11, n. 1, p.22-34, 1977.

NOTA: Devem ser consultadas, além dessas, as obras citadas dos capítulos 4 ao 10.

TEMAS PARA ESTUDO

1. Cite alguma condição fundamental que deva cumprir um tesauro multilíngüe.

2. Diga como apresentaria um tesauro multilíngüe.

3. Justifique por que acredita ser importante, ou não, a interverção de um terminólogo na confecção deste tipo de tesauros.

4. Quando construiria um tesauro bilíngüe?

5. Diga que diferença pode haver entre um tesauro trilíngüe e um dicionário.

6. Diga qual é a utilidade de um tesauro multilíngüe.

7. Possuindo um bom dicionário multilíngüe, é possível confeccionar um tesauro?

8. Cite algum tipo de dificuldade que pode se apresentar ao elaborar um tesauro multilíngüe.

9. Acredita ser necessário possuir conhecimentos lingüísticos ou é suficiente um bom dicionário especializado para a construção de um tesauro bilíngüe?

10. Diga que classes de tesauros multilíngües podem ser feitos.

11. Justifique por que se esforçaria, ou não, em procurar a equivalência de um termo novo em outro idioma.

12. Quando confeccionaria um tesauro trilíngüe manualmente?

13. Esquematize a forma como elaboraria um tesauro multilíngüe.

14. Diga quando daria a primazia a um idioma sobre os restantes.

REFERÊNCIAS BIBLIOGRÁFICAS

CURRÁS, E.: "Confección de un Thesauro Estructurado para Química"; Actas de REUNIBER 70, Madrid, CENIDOC, 1978, 119-126.

LAGUNA SERRANO, E., IRAZAZABAL NERPELL, A., VALLE BRACERO, A.: "Confección Automática de Tesauros"; Rev. Esp. Doc. Cient., *12*, 2, 1989, 129-139.

ORTIZ GONZALEZ, C., FERNANDEZ GARCIA, J.: "Normalización Automática de los Descriptores en la toma de Datos y Elaboración de Indices de Materias Bilingües"; en II Encuentro Hispano-Luso de Información Científica y Técnica, Salamanca, 1988, 332-337.

SPINES Thesaurus, a controlled and structural vocabulary of science and technology for policy making, management and development. Paris, UNESCO, 1976.

Tesauro Spines. Versión provisional en espanol. ICYT (Instituto de Información y Documentación en Ciencia y Tecnología) y Centro de Cálculo Electrónico, CSIC, Madrid, 1985.

UNISIT, Guidelines for the Establishment and Development of Multilingual Thesauri; Paris, UNESCO, 1976.

VIET, J., VAN SLIPE, G.: "EUDISED. Thesauro multilingüe para el tratamiento de la información en materia de Educación. Versión Espanola, Berlín / New York / Amsterdam, Consejo de Europa, Comisión de las Comunidades Europeas y Mouton Publishers, 1984.

NOTA: Devem ser consultadas também as obras citadas dos capítulos de 4 a 10.

9

Da Publicação e outras Questões Complementares

Neste capítulo, inclui-se uma série de questões que não podem compor, por si mesmas, um capítulo independente, como são os assuntos de ligação e desdobramento de tesauros, sua revisão e manutenção, a forma como devem ser redigidos, sua publicação etc.

Ao assunto de automatização não foi dado um capítulo especial, pois, em todos os processos de construção, manipulação ou uso ela intervém de uma maneira direta. Atualmente, quase não se concebe mais o trabalho manual em muitos casos específicos, com um volume mínimo de documentos e outro não muito maior de termos. Existem numerosos programas e "pacotes" com instruções de uso e tesauros completos aplicáveis a qualquer tipo de computador. Os preços também são acessíveis para pequenos orçamentos. A ningém escapam as vantagens advindas do uso dos computadores em qualquer tratamento dos tesauros ou seus termos.

INTEGRAÇÃO DE TESAUROS

Pode acontecer o caso em que seja necessário ampliar o campo de trabalho de um determinado serviço de informação, por motivo de haver aumentado o número de assuntos como que se opera. Pode acontecer que seja preciso confeccionar mais de um tesauro. Nessas circunstâncias, talvez resulte mais prático integrar dois ou vários dos tesauros existentes em um.

É oferecida tanta variedade de tesauros que é imprescindível estudá-los com detalhe para decidir que procedimento será o mais adequado para unir aqueles com os quais se irá trabalhar.

Para estudar os tesauros em questão, o mais operacional é utilizar as técnicas de comparação, atendendo às seguintes características:

1. Especificação.

   Um tesauro pode ser muito específico e outro conter nada mais que termos genéricos.

2. Exaustividade.

   Um tesauro pode conter todas as áreas do assunto tratado e outro omitir alguma.

3. Estruturação dos termos.

Um tesauro pode conter termos compostos e outro somente unitermos.

4. Uso de termos proferidos.

Pode ocorrer o caso de um tesauro não conter mais termos descritores e o outro incluir termos eqüivalentes.

5. Relações de sinonímia.

Em um tesauro, podem aparecer relações de sinonímia que não aparecem no outro.

6. Relações de quase sinonímia.

Também pode ocorrer que variem as relações de quase sinonímia de um tesauro para outro.

7. Relações de hierarquização.

Outro caso que pode se apresentar é o de um tesauro onde se explicitem mais as relações hierárquicas que em outro.

8. Relações de associação.

Também neste tipo de relações podem apresentar-se variações ao considerarem-se termos relacionados em um tesauro que não aparecem no outro.

9. Aplicações e definições.

Pode ocorrer que os mesmos termos sejam usados ou sejam definidos de maneira diferente nos tesauros que se esteja comparando.

10. Origem dos termos.

Os termos usados em um tesauro e em outro podem ter origem diferente. Por exemplo:

10.1. lingüística.
10.2. zona cultural.
10.3. nível científico.
10.4. costume de uso.

11. Idioma em que estão escritos.

Em algumas circunstâncias, não coincidirão os idiomas em que estão escritos um ou outro tesauro.

Como se vê, os casos que podem se apresentar são vários e variados. Aqui foi considerado o caso de integração de dois tesauros. Pensemos nas diversas situações que devem ser consideradas para resolver as possíveis situações surgidas no momento de se integrarem vários tesauros. Além disso, algumas das características citadas podem ocorrer ao mesmo tempo. Na prática, o problema não é tão crítico e a integração pode se realizar com relativa facilidade.

Em qualquer caso, uma vez já tendo sido comparados e estabelecidas as diferenças e analogias, o passo seguinte consistirá em decompô-los em seus termos e estudar termo por termo. Pode-se trabalhar escrevendo um termo e anotando suas peculiaridades; por exemplo, se é descritor ou sinônimo etc. Logo após, será escrito sua equivalência no outro tesauro, também com suas peculiaridades.

Se os tesauros forem escritos no mesmo idioma, o termo naturalmente será idêntico, a menos que não exista tal termo, quando se tornará necessário pesquisar-se um sinônimo ou quase sinônimo com o qual se possa comparar. Quando os tesauros pertencem a idiomas diferentes, a união torna-se mais delicada, porque se deve começar por traduzir o termo para encontrar seu correspondente no outro tesauro. Estes casos são muito pouco freqüentes e complicados. Poderão ser utilizadas as técnicas da tradução automática ou as indicadas para a confecção de tesauros multilíngües.

A figura 9.1, reproduzida da obra *Thesaurus Construction*, mostra uma maneira de comparar termos de vários tesauros, usando o computador. A figura 9.2, reproduzida da mesma obra, faz referência ao caso de um tesauro multilíngüe e inclui a tradução do termo para o francês e para o espanhol.

Uma vez que foram estudados os termos, será procedida a sua integração em um ou outro tesauro. Normalmente, tratar-se-á de:

- acrescentar,
- substituir,
- modificar,
- suprimir

└──── termos

no lugar correspondente da seção apropriada, seja alfabética, sistemática ou gráfica. Neste úlitmo caso, será necessário recompor os diagramas e listas alfabéticas a que fazem referência. Considere-se que, quando se deve modificar a seção sistemática, pode-se contar com uma modificação no sistema de classificação usado. Certamente, será necessário adequar-se o índice situado no final do tesauro.

Vocabulary Switching System

```
WELCOME TO *VSS* - VOCABULARY SWITCHING SYSTEM
USING BATTELLE'S DATA MANAGEMENT SYSTEM: BASIS

VSS  CONTAINS FOUR VOCABULARY SETS
  1- BUSINESS
     A. ABI                    B. MANAGEMENT CONTENTS
  2- BEHAVIOUR SCIENCE
     A. ERIC                   B. PSYCH ABSTRACTS
  3- LIFE SCIENCE
     A. BIOSIS                 B. CHEM ABSTRACTS
     C. MESH
  4- PHYSICAL SCIENCE
     A. DOE                    B. CHEM ABSTRACTS
     C. EI                     D. INSPEC
     E. IRON                   F. NASA

PLEASE SELECT 1 OF THE 4 VOCABULARY SETS
  BY ENTERING EITHER 1, 2, 3, OR 4

   ?  4

  VSS PROVIDES FOR 6 SWITCHING OPTIONS:
   1- SYNONYMS
   2- BROWSE
   3- NARROWER TERMS
   4- BROADER TERMS
   5- NARROWER/BROADER TERMS
   6- OTHER (USER-DEFINED)

PLEASE SELECT 1 OF THE 6 OPTIONS
  BY ENTERING EITHER 1, 2, 3, 4, 5, OR 6

   ?  2

SPECIFY THE MAXIMUM NUMBER OF TERMS
  TO BE DISPLAYED PER VOCABULARY.
  ENTER A NUMBER.

   ?  10

PLEASE ENTER A SINGLE SEARCH TERM OR COMMAND

   ?  HEAVY WATER

SWITCH SUCCESSFUL
```

TERM TYPE	VOCAB	TERM
YOUR TERM	DOE	HEAVY WATER
YOUR TERM	EI	HEAVY WATER
YOUR TERM	INSPEC	HEAVY WATER
YOUR TERM	NASA	HEAVY WATER
RELATED	DOE	COOLANTS
RELATED	NASA	COOLANTS
RELATED	DOE	DEUTERIUM COMPOUNDS
RELATED	EI	DEUTERIUM COMPOUNDS
RELATED	INSPEC	DEUTERIUM COMPOUNDS
RELATED	NASA	DEUTERIUM COMPOUNDS
RELATED	DOE	MODERATORS
RELATED	CHEM A	MODERATORS
RELATED	INSPEC	MODERATORS
RELATED	NASA	MODERATORS
RELATED	DOE	TRITIUM COMPOUNDS
RELATED	INSPEC	TRITIUM COMPOUNDS
RELATED	DOE	DUAL TEMPERATURE PROCESS
RELATED	DOE	DEUTERIUM
RELATED	EI	DEUTERIUM
RELATED	INSPEC	DEUTERIUM
RELATED	NASA	DEUTERIUM
RELATED	DOE	TRITIUM
RELATED	EI	TRITIUM
RELATED	INSPEC	TRITIUM
RELATED	NASA	TRITIUM
RELATED	INSPEC	FISSION REACTOR MATERIALS
RELATED	DOE	HEAVY WATER PLANTS
RELATED	DOE	SGHWR REACTOR

Figura 9.1.

Descriptor bank entry
'Integrated thesaurus of the social sciences'

	*IC	IL	IS	MA	PM	SP	UT
NUPTIALITY							
IC : 22230							
IL : 14.01							
IS : 15421							
MA : 14.02.05							
PM : 13.01.00							
UT : R10.71							
F = Nuptialite		IL	IS	MA	PM		
S = Nupcialidad		IL		MA	PM		
= Marriage rate						SP	
< Population dynamics						SP	
< Population events							UT
— Divorce						SP	
— Families						SP	
— Family system						SP	
— Genetic counselling						SP	
— Marriage	IC	IL	IS	MA	PM		UT
— Nuptiality rate					PM		
— Nuptiality table					PM		
— Sexual union					PM		
— Statistical data						SP	

* Codes indicating thesauri included in the descriptor bank, e.g. MA = *Macrothesaurus,* UT = *UNESCO thesaurus.*

Figura 9.2. - Entrada do banco de descritores. "Tesauro integrado de Ciências Sociais. Construção de tesauro.

Em algumas circustâncias, convém ressaltar os termos que foram acrescentados ou modificados. Para isso, acrescenta-se uma notação ou se escrevem os mesmos com outro tipo de letra, se forem impressos em papel.

Sempre será conveniente indicar os critérios que foram seguidos na integração, para facilitar seu uso posterior (figura 9.3).

DESDOBRAMENTO DE TESAUROS

Quando se deve desdobrar um tesauro, deve-se pensar que se trata de confeccionar um de menor âmbito de conteúdo e outro mais amplo. O caso mais simples é aquele em que se deve separar um microtesauro de um macrotesauro, pois é necessário apenas separá-lo no computador ou imprimi-lo separadamente.

As outras situações que podem se apresentar também não oferecem dificuldades. Deve-se considerar:

1. Âmbito do assunto que se deseja abranger.

2. Especificidade do tesauro

   que se deseja separar, quer dizer, determinar se é necessário tomar todos os termos possíveis ou será suficiente um número reduzido.

3. Especificidade dos termos:

   deverá ser estabelecido, caso se deseje contar com todo o termo em sua extensão, incluindo o conjunto de relações, ou será suficiente com alguma referência aos termos relacionados com este. Por exemplo, dever-se-ão constar todos os termos superiores, ou somente o primeiro grau de hierarquia.

4. Adequação dos termos:

   será preciso determinar que termos são os mais adequados para expressar os conceitos do assunto de trabalho.

5. Composição do termo:

   do caso anterior, será deduzido de que maneira deve-se tomar o termo e qual será sua forma e sua composição.

6. Idiomas escolhidos:

   pode ocorrer que se trate de separar um tesauro em um determinado idioma, retirado de outro, multilíngüe.

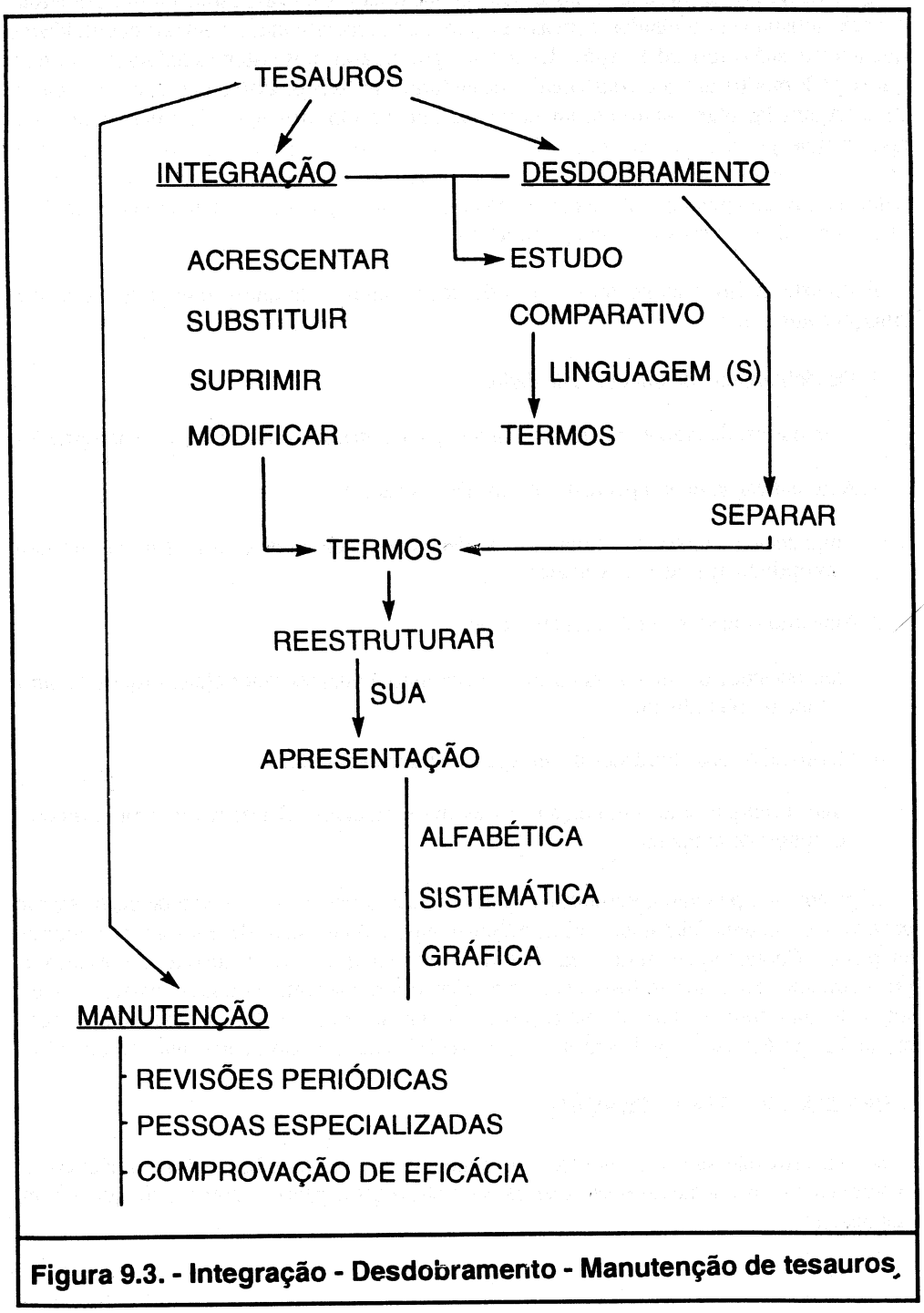


Figura 9.3. - Integração - Desdobramento - Manutenção de tesauros

Quando for preciso desdobrar um tesauro, também será necessário estudar termo por termo e ir selecionando os escolhidos, com as características, especificidade e outras peculiaridades que tenham sido prefixadas. Após, dever-se-á apresentá-los nos sistemas habituais. Note-se que se pode modificar a seção sistemática de tal forma que seja preciso procurar outra maneira de estruturá-la, planejando-se um novo método de classificação. Também mudará a apresentação gráfica. Menciona-se este tipo de apresentação, por completar os procedimentos utilizados, embora, atualmente, esteja caindo em desuso. Um tesauro prático e operacional, realizado por computador, não necessita de uma apresentação gráfica, totalmente supérflua. Os tesauros de publicação recente prescindem dela.

Feito este inciso, continuemos com o desdobramento de tesauros, que pode dar-se em situações tais como:

1. Desdobramento de uma base de dados

   em outras, de menor conteúdo temático, para a qual faz falta um tesauro apropriado.

2. Automatização de um pequeno serviço de informação

   que decida adaptar-se a uma base de dados já criada e que necessita de um tesauro apropriado às suas necessidades.

3. Aumento considerável do assunto tratado

   que obrigue a dividir o assunto em parcelas mais reduzidas, determinando que o tesauro também seja reduzido.

4. Modificação das atividades de um centro de documentação,

   que implique uma diminuição dos assuntos tratados. O tesauro deverá, também, diminuir de tamanho.

Logicamente, poderão apresentar-se outras situações nas quais se tenha de desmembrar um tesauro, e alguma delas implicará, ao mesmo tempo, a necessidade de realizar uma posterior adaptação. Comentou-se, nesta obra, que, nos casos em que se deva adaptar um tesauro, o procedimento será como se fosse necessário construir novamente. O tesauro existente pode servir de guia, fonte de termos, indicação de formas de relação entre eles etc. Sempre será muito útil, porém não se pode suprir as peculiaridades de cada caso particular (figura 9.3).

A REVISÃO E A MANUTENÇÃO

Os tesauros não são seres mortos nem estáticos. São vivos e bem vivos, dinâmicos e evolutivos. Precisamente essas são suas características principais e o motivo pelo qual foram planejados.

Por conseguinte,devem ser submetidos à revisão contínua para conseguir uma manutenção atualizada (figura 9.3). A situação ideal seria aquela na qual o grupo internacional encarregado da construção do tesauro se reunisse periodicamente e estudasse sua situação e as necessidades de modificação:

- alterando os termos
  - seu significado
  - suas relações com outros termos;
- introduzindo termos;
- suprimindo termos.

Se os especialistas na área se reunissem, seria possível que não variassem os critérios estabelecidos e que fossem seguidas as mesmas diretrizes de trabalho. Porém, nem sempre é possível escolher o melhor. Contudo, o que é factível é dar a uma pessoa a função de manutenção e atualização do tesauro, que, seguramente, não precisará ocupar-se do mesmo em tempo integral, nem mesmo diariamente. Naturalmente, as horas destinadas a essas tarefas dependerão do tamanho do centro ou base de dados, do assunto ou assuntos tratados, do uso que se faça do tesauro e de outras várias considerações. Pode ocorrer, inclusive, que seja necessário empregar mais de uma pessoa, em tempo integral, nestas atividades.

Menciona-se este assunto de pessoal dedicado aos trabalhos de revisão e atualização de tesauros, porque costuma ser normal, sobretudo por parte de diretores ou do pessoal administrativo, pensar que estas são tarefas secundárias às quais se deve dedicar pouca atenção e somente de maneira esporádica. Quantas falhas podem ocorrer em um serviço de informação ou em uma base de dados, por não se ter o tesauro? Quantos exemplos se poderia citar!

Para manter um tesauro atualizado, deve-se fazer revisões periódicas com uma freqüência que varia conforme os casos. Não se pode dar conselhos aqui. Cada um deve conhecer suas necessidades e situações. Porém, é importante que sejam periódicas. Do contrário, corre-se o risco de deixar para mais adiante e, logo, a tarefa torna-se maior e mais complicada. Por outro lado, trabalhando-se irregularmente, pode-se perder a visão de conjunto e não perceber a própria utilidade do tesauro. Um tesauro defasado é "computador" molhado.

Para revisar um tesauro, deve-se dar toda a atenção ao estudo dos termos que o compõem e a sua utilidade. Deve-se considerar se:

1. um termo é muito usado;
2. outro termo caiu em desuso;
3. aparecem novos conceitos que necessitem de termos novos;
4. deve-se introduzir termos novos;

5. é preciso alterar as relações entre os termos;
6. há mudança no conteúdo conceitual dos termos.

É de opinião geral que, quando um termo aparece com bastante freqüência, deve ser suprimido, pois torna-se supérfluo. Deve-se ter muita cautela nesses casos e estudar detidamente o termo em questão e o conceito que representa. Sobretudo é preciso fixar-se em sua categoria. O assunto principal de trabalho faz referência a um assunto secundário? Trata-se de um termo de conteúdo amplo, muito geral, como, por exemplo: obtenção, análise? Talvez seja conveniente combiná-lo com outro? São várias as considerações que devem ser feitas. É aconselhável que não se suprima o termo, colocando-o de lado, e observando-se os resultados que ocorrem quando se prescinde desse termo.

No caso de um termo já antigo no tesauro, que seja muito pouco utilizado, é necessário também agir com muita cautela antes de se pensar que o termo caiu em desuso e deve ser eliminado. Considere-se que o desuso pode ser temporal, por muitas razões variadas, entre as quais, a pouca utilidade do assunto, a modificação dos planos de trabalho do centro em questão, a redução de orçamentos e, conseqüentemente, a diminuição de atividades, bem como outros motivos variados. É muito aconselhável, também agora, não se desfazer desse termo e deixá-lo de lado, até comprovar-se a eficácia do trabalho documentário, seja ele de indexação ou de recuperação da informação, quando, então, prescinde-se dele. Por outro lado, aquelas condições podem mudar e repercutir na freqüência do uso do termo citado.

Quando aparecem novos conceitos, pela própria evolução do futuro da humanidade e por ampliação do campo de trabalho do serviço de informação, será preciso atribuir-lhes um termo. Este deve ser introduzido no tesauro.

Costuma-se tomar como norma a comprovação da utilidade e eficácia de um novo termo, antes de introduzi-lo no tesauro, o que evitará o emprego de más considerações. É necessário situá-lo em seu contexto, em sua família de termos afins, com todas as relações apropriadas ao caso, ou seja, é preciso construir um termo completo, à medida que seja possível, e, então, comprovar sua eficácia. Como resultado, será aceito ou desprezado. Talvez, deva-se escolher outro termo diferente para designar o novo conceito. Neste caso, deve-se começar novamente, até encontrar o termo apropriado.

Conseqüentemente, ocorrerá o caso em que se deva introduzir novos termos no tesauro. Será muito conveniente seguir os seguintes passos:

1. Comprovar que o termo corresponde ao conceito.
2. Consultar especialistas na área e terminólogos.
3. Consultar obras sobre o assunto tratado.
4. Adotar o termo escolhido.
5. Construir o termo completo, à medida que for necessário.

6. Determinar sua posição no tesauro.

7. Estabelecer suas relações com os termos afins.

8. Incluir o termo fisicamente no tesauro.

9. Alterar as seções alfabéticas, sistemáticas e gráficas, conforme o caso.

10. Comprovar sua eficiência.

Todas essas operações requerem um trabalho consciente e metódico, que deverá ser realizado por especialistas com boa formação, formação essa que deverá ser tanta como quando se trata de confeccionar um tesauro de nova fábrica.

Todavia, verificou-se outro momento no qual deverá ser feita a atualização de um tesauro; quando será preciso alterar as relações entre seus termos. Pode ocorrer que se introduza uma nova área do trabalho em um serviço de informação, ou mude o sistema de perfis de um usuário, ou se deseje modificar algum aspecto nos campos da pesquisa documentária. Em uma série de situações diversas, será imposta a mudança das relações entre os descritores entre si, ou entre os descritores e seus termos equivalentes. Considere-se um tesauro sobre informática, no qual se amplia sua abrangência de assunto mediante técnicas de inteligência artificial. A introdução dos novos termos implica uma ampliação das relações hierárquicas, associativas e equivalentes. Por exemplo, a palavra *software* terá aumentado o seu conteúdo conceitual e os seus níveis de hierarquia com relação a outros termos, assim como as suas relações associativas, com o acréscimo de novos termos. Como se observa, aqui não somente se modifica um termo com relação a outro, como também se muda o seu conteúdo conceitual. E esta é outra questão importante que deve ser considerada quando se deseja atualizar um tesauro. Os ramos do conhecimento evoluem e os idiomas também. Portanto, deve-se ficar atento para detectar essa evoluções.

Muitos são os pontos que devem ser cumpridos na atualização de um tesauro. A tarefa se tornará leve, caso se trabalhe periodicamente e com um método estabelecido, de acordo com as necessidades de cada um, e com o uso do computador.

A PRÁTICA DA COMPROVAÇÃO

É aconselhável também comprovar a eficácia de um tesauro e observar até que ponto o mesmo é útil, ou se é preciso realizar alguma modificação.

Há quem pense que essa comprovação deve realizar-se no momento em que um tesauro encontra-se completamente finalizado. A opinião geral, contudo, inclina-se para a idéia de não se deixar para mais tarde. É mais operacional ir realizando-se ensaios, conforme se vá cumprindo etapas, ou quando se tenha concluída uma família de termos, ou ainda quando se tenha confeccionado uma parte da seção sistemática ...

A melhor maneira de realizar uma comprovação de um tesauro é empregar algumas das seguintes práticas, ou todas elas, simultaneamente:

1. utilizar situações reais de uso

   → na indexação de documentos,
   → na recuperação da informação;

2. planejar estratégias fictícias de busca de informação;

3. planejar situações fictícias de indexação;

4. solicitar aos usuários que sugiram buscas de documentação ou assuntos de trabalho;

5. remeter o tesauro para colegas que trabalhem na mesma área de assunto, para que eles comprovem o seu uso.

Realmente, o uso do tesauro é que nos dirá se o mesmo é correto e onde se encontram as falhas.

O MOMENTO DA REDAÇÃO E DA PUBLICAÇÃO

Quantas páginas foram escritas até chegarmos a esta epígrafe! Visto que não se deve começar a redação do tesauro antes de se ter todos os elementos necessários, antes de terminar a sua construção. Em alguns casos, pode ser necessário redigir alguma parte, com fins determinados. Considere-se que cada uma dessas seções, como também a parte gráfica, compõem tesauros por si mesmas. Portanto, não estamos deixando de cumprir a norma indicada anteriormente (figura 9.4).

Sempre que for necessário redigir um documento, deve-se considerar separadamente:

- o conteúdo,
- a forma.

No que se refere ao conteúdo, trata-se de redigir o próprio tesauro para o que é necessário levar em consideração:

1. o tipo de tesauro;

2. a maneira como se redigirá.

Figura 9.4. - Publicação de tesauros.

Ou seja, devemos considerar se estamos diante de um tesauro alfabético, sistemático, gráfico ou composto por alguma dessas partes. Deve-se lembrar também se foi confeccionado um macrotesauro.

Em todos esses casos e satisfazendo a forma na qual deve ser redigido um tesauro, elaborar-se-á um roteiro no qual figurem os capítulos em que se dividirá o conjunto de obra. Pode ser feito da seguinte maneira:

Roteiro

Justificativa

Introdução

Palavras de agradecimento

Índice

Normas seguidas para a construção

Instruções de uso

Abreviaturas utilizadas

Seção alfabética desenvolvida[1]

Seção sistemática desenvolvida[2]

→ com índice alfabético

→ um tesauro completo

→ simples lista de termos

Seção gráfica[3]

→ com índice alfabético

→ um tesauro completo

→ simples lista de termos

→ índice de diagramas

Índice final alfabético de termos

Nota: Pode suprir ao registro temático.

Bibliografia

Obras de consulta

Registro temático

Relação de autores citados[4]

De todos os capítulos mencionados, somente os assinalados necessitam de comentário.

Comecemos pelo número 1, referente à Seção Alfabética, que pode constituir, por si só, todo o tesauro. Trata-se da parte onde figuram todos os termos que constituem o tesauro, ordenados alfabeticamente, acompanhados, cada um, dos termos restantes com esse relacionados. Este tipo de tesauro é muito fácil e flexível.

O número 2 corresponde à Seção Sistemática, que também pode constituir por si só, o tesauro, ou fazer parte de outro mais completo. Em qualquer caso, não se deve prescindir de um índice alfabético próprio, onde apareçam os termos com a referência à classe ou faceta a que pertencem. Em algumas circunstâncias, esse índice complica-se, com o acréscimo a cada termo de outros com ele relacionados em níveis de hierarquia, associação ou equivalência. Converte-se, assim, este índice em um tesauro alfabético, que pode formar, por si mesmo, uma seção alfabética do tesauro principal. Deve-se contrabalançar cuidadosamente quando se necessita de tanto detalhe. Uma boa norma de trabalho consiste em não complicar o trabalho. Quanto mais simples for um tesauro, mais eficaz e "simpático" será o seu uso.

A Seção Gráfica está assinalada com a nota 3. É imprescindível que esta seção seja acompanhada de uma relação alfabética de termos onde se faça referência ao diagrama, devidamente assinalado, em que se encontra. Este índice pode acompanhar cada diagrama separadamente, ou pode ser escrito após o último diagrama, com o conjunto completo de termos que contém. Querendo ou devendo complicar as coisas, também pode-se confeccionar o índice como o tesauro completo, ou retirá-lo diretamente da Seção Alfabética. Não é muito comum que um tesauro conste somente da Seção Gráfica, ainda que ocorram alguns casos. Na maioria das vezes, confeccionam-se como apoio e complemento explicativo para qualquer das seções restantes.

A nota 4 acompanha a relação de autores citados, para explicar que esta relação somente será incluída no tesauro quando forem citados autores nas primeiras partes, tais como na justificativa, na introdução e nas normas seguidas para a construção. Os autores arrolados na bibliografia ou nas obras de consulta não costumam ser relacionados novamente em outra lista independente. Isto implicaria uma duplicação desnecessária de trabalho.

Pode ser alvo de atenção o fato de haver dedicado uma epígrafe às obras de consulta independentemente da bibliografia. Isto se justifica pelo fato de se tratar de publicações com função diferente. Naquelas, aparecerão as publicações, livros, separatas, documentos avulsos, outros tesauros etc. que façam referência ao assunto "tesauros" e que mereçam ser citados por sua importância. Poderiam ser considerados como "leituras recomendadas". Pelo contrário, na bibliografia, aparecerão os documentos usados na confecção do tesauro. O objetivo dessa separação é fazer a bibliografia completa mais leve e consultável, com uma divisão clara entre o necessário e o complementar. É sempre bom e útil dispor de um grande número de obras de referência.

Abordemos, agora, a forma física na qual se apresentará um tesauro. Na realidade, trata-se de sua publicação, que poderá ser feita

- impressa em papel
- legível por computador
  - → em fita ou cinta magnética
  - → em disquetes
  - → visível em pantalha
  - → reproduzida por impressora
- em microformas.

Atualmente, quando se dispõe de excelentes impressoras a raio laser e os computadores tornam-se acessíveis financeiramente a um número maior de pessoas, parece que não se deveria falar mais do que uma forma de publicar um tesauro, pois se tem em memória de computador para seu uso e manutenção, em disquetes, por exemplo. Ao mesmo tempo, pode-se fazer uma cópia impressa de qualidade magnética, com entradas de capítulos e diferentes tipos de letras. Contudo, haverá casos – e, de fato, existem – em que um tesauro deverá ser publicado por impressão gráfica. Em um caso ou no outro, sempre que for oportuno, as normas e leis concernentes a cada caso devem ser cumpridas.

Para terminar esta epígrafe, lembramos que é muito recomendável e desejável remeter uma nota aos centros nacionais de informação e documentação, anunciando que foi iniciada a construção de um tesauro e indicando o assunto abordado e outras características. Da mesma maneira, quando a construção do tesauro for concluída, também deve ser comunicada.

Para os tesauros em espanhol, deve-se recorrer à Reuniber, ao Cindoc, ao Isoc e à Fuinca-Fundesco. O Consórcio de Catalunha compila e confecciona tesauros em Catalão. Quando se tratar de tesauros em outros idiomas, é aconselhável dirigir-se a:

• Aslib
Information House
20-24 Old Street
London EC1V 9AP
United Kingdom

• D.G.D. Komitee für Klassifikation und Thesaurusforchung
Deutsche Gesellschaft für Dokumentation
Westendstr. 19
D. 6000 Frankfurt / Main 1
Alemania

• Thesauros Clearinghouse
The Library
Faculty of Library Science
University of Toronto
140 St. George Street
TORONTO
Ontario M55 1A1
Canadá

Para tesauros em outras línguas:

• Instituto Informacj i Naoukowej, Technicznej i Ekonomiecznej (IINTE)
Zurawia 3/5
00-926 WARSAW
Poland

PACOTES INFORMATIZADOS

Em continuação, citam-se alguns pacotes de instruções de construção e uso, com tesauros incluídos, e as sociedades onde se pode dispor dos mesmos. É uma lista orientadora, uma vez que a situação é muito versátil e muda freqüentemente. Aparecem novos pacotes, novos programas de computador, novos tesauros automatizados, ou novas sociedades que os comercializam e onde podem ser consultados. Com a mesma facilidade desaparecem. Por este motivo, aconselha-se dirigir-se aos centros nacionais de documentação, ou aos que possuem essas funções, para obter informação recente e atualizada.

Aqui está a lista mencionada anteriormente.

ASTUTE
Sistema Automatizado para la Actualización, Control e Impresión de Tesauros (para tesauros multilingües)
Comission of the European Comunities
Directorate General for Scientific and Technical
Information and Information Management
DG XIII-B
Batiment Jean Monnet Plateau Kirchberg
Luxembourg L-2920
Tel. (352) 2920

BATTELLE INSTITUTE LTD.
BASIS Coordinator
15 Hanover Square
London W I R 9AJ
United Kingdom
Tel. (01) 493 0184

BIBLIOTECH LIBRARY SYSTEM
Biblio-techniques Library and Information System (BLIS)
Advanced Data Management Inc.
c/o Comstow Information Services
302 Boxboro Road
Stow, MA 01775
USA
Tel. (617) 897 7163

CALM
Pyramid Software Products
9 Church Street
Reading
Berks. RGI 2SB
United Kingdom
Tel. (0734) 595633

MINISIS
International Development Research Centre
Information Science Division
PO Box 8500
60 Queen Street
Otawa, ON
Canada KIG 3H9
Tel. (613) 236 6163

POLYDOC
Norwegian Centre for Informatics
Postboks 350
Blindern
0314 Oslo 3
Norway
Tel (2) 452010

TEMAS PARA ESTUDO

1. Diga em que forma publicaria um tesauro.

2. Faça um roteiro para a publicação de um tesauro.

3. Que diferença estabeleceria entre a publicação de um tesauro monolíngüe e outro trilíngüe?

4. Dê sua opinião sobre o desdobramento de tesauros.

5. Quando se torna necessário integrar dois tesauros?

6. Faça um esquema do procedimento a seguir no desdobramento de um macrotesauro.

7. Diga quando usaria as técnicas de informática para publicar um tesauro.

8. Na sua opinião, quem deve realizar a atualização de um tesauro?

9. Qual é o papel desempenhado por um terminólogo em todas estas questões?

10. Na sua opinião, quando é oportuno atualizar um tesauro?

11. Que faria para introduzir um novo termo em um tesauro bilíngüe?

12. Publicaria um tesauro em forma gráfica?

13. Diga o que faria com um termo muito freqüente, no momento de desejar incluí-lo na parte sistemática de um tesauro.

14. Suprimiria um termo pouco usado?

15. Na sua opinião, quem deve fazer uma comprovação sobre a utilidade de um tesauro?

16. Onde pesquisaria bibliografias sobre tesauros?

REFERÊNCIAS BIBLIOGRÁFICAS

AITCHISON, J.: "Integrated Thesaurus of the social science; design study"; París, UNESCO, 1981.

NOTA: Devem ser consultadas também as obras citadas dos capítulos 4 ao 10.

conteúdo e a **forma**. O conteúdo refere-se ao assunto, e a forma, à maneira como é usado o idioma em que se apresentam os documentos analisados.

Este último ponto será importante no momento de planejar os termos compostos ou as relações de sinonímia, quando se estudam as palavras que o autor empregou para expressar suas idéias. Também pode ser determinante, neste caso, a zona cultural de origem do autor e o uso da língua materna ou outra estrangeira.

TÉCNICAS DE INDEXAÇÃO

Do ponto de vista da utilização do tesauro, para indexar um documento, deve-se considerar:

o conteúdo

• quer dizer,

o assunto tratado;

onde se determinará:

1. o assunto principal;
2. os assuntos auxiliares;
3. os assuntos marginais;
4. a extensão do assunto;
5. a reunião de algumas de suas partes.

a forma

• fará referência à maneira como se expressa o assunto tratado;
onde se considerará:

1. o idioma;
2. as suas características lingüísticas;
3. a sua redação;
4. o tipo de linguagem usada:

→ científica,
→ vulgar;

5. a especificidade da linguagem;
6. o tratamento dos conceitos;
7. o conhecimento da terminologia própria do assunto tratado;
8. o espaço cultural de procedência do autor;
9. o uso do idioma materno ou de um estrangeiro;
10. o conhecimento do assunto por parte do autor;
11. a instrução ou profissionalização do autor.

Naturalmente, é importante levar em consideração o idioma com o qual se deve trabalhar. O tratamento não será o mesmo para o inglês, o alemão ou o italiano, por exemplo. Cada um deles apresenta suas próprias características lingüísticas, que devem ser conhecidas para determinar as peculiaridades gramaticais ou sintáticas. Deve-se dar atenção também à maneira como está redigido o documento, verificando-se se são usados mais substantivos, adjetivados, ou maior número de frases preposicionais, ou outras formas de expressão, de onde se deduzirá como foram apresentados os conceitos. E ainda se deve levar em consideração se o autor utilizou uma linguagem científica, ou se tratou o assunto com palavras de linguagem vulgar. Em um caso ou outro, o indexador deverá concluir se deverá tomar os termos de indexação como estão expressos no documento, ou se terá de substituí-los por outros mais adequados aos que figuram no tesauro. Pelo mesmo motivo, será preciso considerar se o assunto é tratado muito especificamente ou de maneira geral.

Uma questão de suma importância relacionada com os casos anteriores refere-se ao conhecimento da terminologia do assunto, ou seja, ao uso adequado dos conceitos por parte do autor. Se este conhece a terminologia, o documentalista pode confiar nela e usá-la sem receio. Deverá conhecer igualmente a zona cultural de onde procede o autor. Por exemplo, um peruano usará expressões castelhanas que talvez não possuam o mesmo significado na Espanha, ou que não sejam utilizadas. Será necessário, então, consultar os vocabulários ou dicionários e o tesauro, para encontrar termos relacionados ou sinônimos, conforme o caso. Se um autor não escreve em seu idioma materno, também pode confundir os termos ou não usá-los no sentido correto.

Para determinar se o autor se expressa de modo adequado, utilizando a terminologia apropriada corretamente, é fundamental saber se o mesmo conhece o assunto de que trata, seu nível de instrução e sua formação profissional. Por exemplo, um sismólogo não se expressará da mesma forma que um jornalista quando escreve sobre terremotos. Como também não se expressarão da mesma forma um lavrador e um engenheiro agrônomo ao tratarem de colheitas, embora o lavrador saiba muito mais sobre o assunto que o engenheiro agrônomo.

Todas essas questões são realmente importantes quando se deve escolher os termos de indexação de acordo com o tesauro que deve ser usado. Assim, será possível determinar:

• *as características dos termos com as quais se deve trabalhar; por exemplo:*

1. sua freqüência,
2. sua composição,
3. sua especificidade,
4. sua adequação ao tema,
5. sua adequação ao conceito que representam,
6. sua ortografia correta.

Com todos esses dados, poderer-se-á esquematizar uma **estratégia de indexação** que indique quais os termos que devem ser usados e de que forma devem ser tratados. Por exemplo, um termo que aparece com freqüência no documento indicará que não se pode prescindir dele na indexação. Também é importante saber de que maneira foram compostas as palavras para formar um termo, ou seja, se foi escrito "terapia educacional" ou "terapia para a educação".

Igualmente deve-se levar em consideração a especificidade com que se trata a linguagem, ou se são usados os termos adequados, ou se estão escritos corretamente... Quando o autor não escreve em sua própria língua, deve-se ter mais cuidado, porque talvez aumentem as incidências.

Uma vez que se tenham avaliado todas essas características, poder-se-á seguir as seguintes etapas:

1. escrever os termos da mesma forma feita pelo autor, exceto nos casos em que se apresentem erros ortográficos (pense-se nos estrangeiros);
2. consultar o tesauro, para ver se são encontrados os termos;
3. comprovar de que forma aparecem os termos no tesauro;
4. comprovar, em seu caso, que não existe o termo no tesauro;
5. adotar termos diferentes dos termos usados pelo autor, caso seja necessário.

Como primeira medida, deve-se escrever os termos tal como foram usados pelo autor. Logo se comprovará se eles existem no tesauro e de que forma. Pode ocorrer que o termo não apareça, porém alguma variante de um sinônimo. Por exemplo, no lugar de "ciências da documentação", aparece "ciências documentárias". Supondo-se que o termo em questão não apareça no tesauro, deve-se verificar se há alguma maneira de adequá-lo a outro existente. Em caso contrário, pensar-se-á que se está diante de um novo termo.

Após realizar todos esses estudos e comprovações, chega-se à formulação do termo de indexação, que será o que ficará designado para determinar o assunto contido no documento.

Isso tudo não é tão fácil, como é possível deduzir do que foi escrito até aqui. Cada um dos passos citados nos parágrafos precedentes necessita de estudos cuidadosos e delicadas deliberações e do uso de determinadas estratégias para conseguir um termo de indexação adequado e eficaz (figura 10.2).

O TRUNCAMENTO DOS TERMOS

Uma dessas estratégias auxiliares consiste em cortar os termos, separando a raiz de seus prefixos e sufixos. Esta é uma técnica muito útil quando se trabalha com documentos que tratam de um assunto geral em seu conjunto, porém muito específico em cada uma de suas partes.

Nesses casos, diversas formas do mesmo termo geral aparecem, repetidamente, expressas de maneiras diferentes. Um procedimento prático é limitá-las a fazer referência a somente um, ou a uma parte dele, de modo que fiquem interligados. Por exemplo, pensemos em um assunto marinho, onde aparecem termos como: oceano, oceanografia, nível oceânico, furnas oceânicas e oceanos glaciais. Todas essas palavras possuem uma mesma raiz *ocean*, portanto poderá ser tomada como vínculo para ligar todas as demais.

Os programas de computador já foram planejados para tratar este tipo de termos cortados. Realizar esta tarefa manualmente é quase impossível, pela quantidade de etapas que requer.

Para designar o corte, costuma-se utilizar o asterisco (*) como notação. Na prática, apresentam-se os seguintes casos:

1. corte à direita;
2. corte à esquerda;
3. corte bilateral.

Tomemos como exemplos:

1. docum*:
   se indexação ao mesmo tempo:

   docum-entação;
   docum-ento;
   docum-entarista;
   docum-entalismo;
   docum-ento primário;
   docum-entos legais.

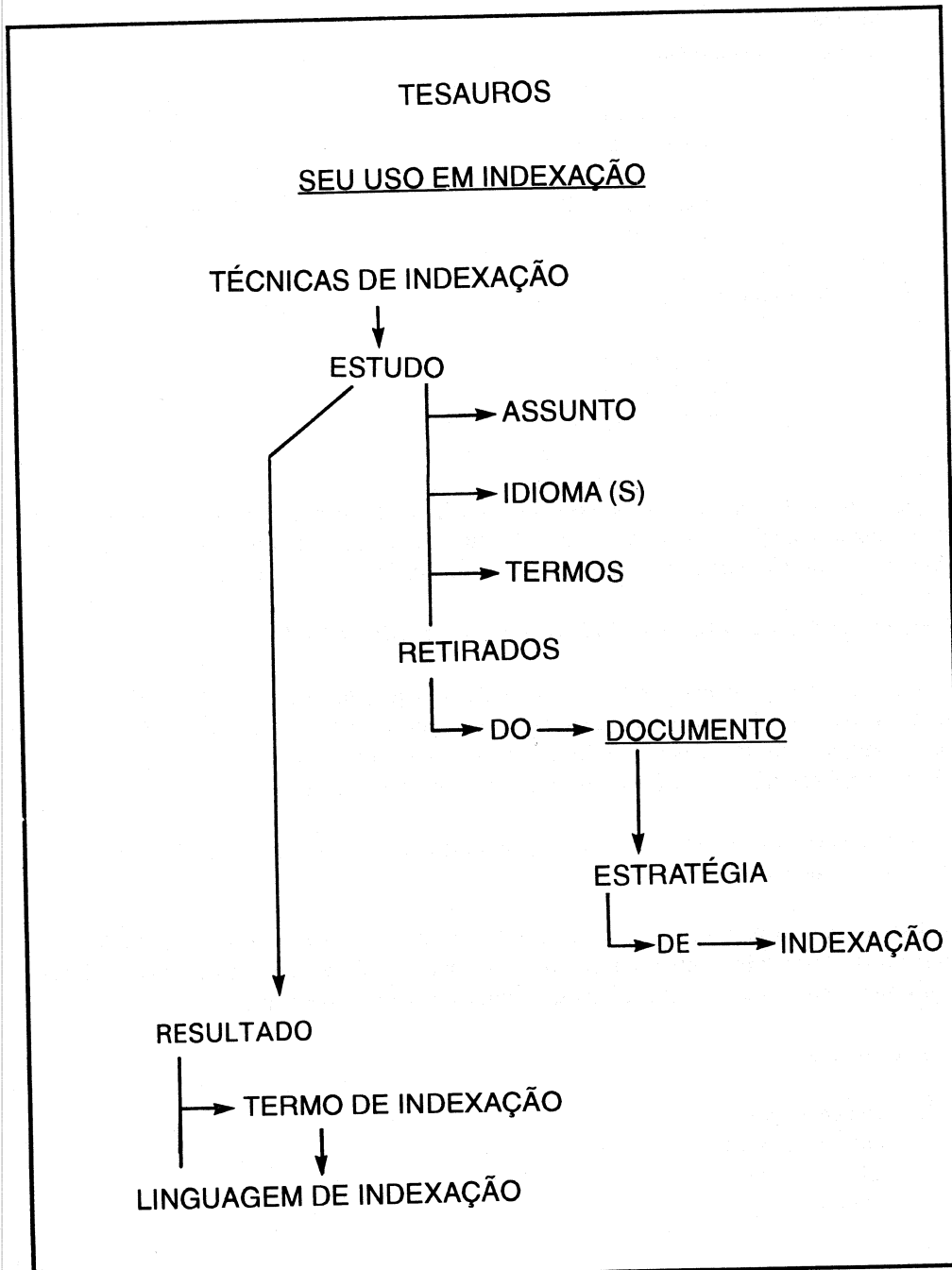


Figura 10.2 - Uso do tesauro na indexação.

2. *análise:
se indexarão também:

psico- análise;
eletro- análise.

técnicas de análise.

3. *-genétic-*
se indexarão também:

cito-genétic-a
bio-genétic-a
bio-genétic-ista

As técnicas de truncagem são úteis para serem aplicadas particularmente em idiomas com o espanhol, que apresenta terminações diferentes para os gêneros masculino e feminino e para os plurais. Com uma só forma, abrangem-se todos os casos possíveis.

USO DE OPERADORES DE PROXIMIDADE E DISTÂNCIA

Utilizou-se esta técnica, geralmente, nos casos em que um termo composto se repete com freqüência em um documento, como também, em uma série deles. Para poder localizar onde se encontram cada uma das partes do termo, utiliza-se uma notação que o computador sabe identificar. Costuma-se usar a letra w. Por exemplo: sala w leitura.

Na memória do computador, ficam registrados todos os termos "leitura" e "sala", onde quer que se encontrem.

Esta é uma técnica muito prática quando se trabalha com tesauros construídos com unitermos. Do contrário, é mais útil combinar os conceitos e formar termos adequados. Talvez seja conveniente, também relacionar estes com a notação de associação TR.

Por exemplo:

sala de acordos
sala de reuniões
sala de espera
telefone de sala
.
.
.
sala de leitura

.
.
.
hábito de leitura
leitura infantil
.
.
SALA DE ACORDOS
TR SALA DE REUNIÕES
.
.
HÁBITO DE LEITURA
TR LEITURA INFANTIL

ÍNDICE DE PONDERAÇÃO

Outra técnica que ajuda a fixar a importância de um termo de indexação é aquela pela qual se pode determinar sua relevância, indicando seu "peso específico". Normalmente, estabelece-se uma escala de valores arbitrariamente. Pode ser do 1 ao 9, ou pode-se utilizar letras. Isto não é importante. O importante é que cada termo venha acompanhado da notação correspondente, segundo sua relação com o assunto central do documento. Os termos que fazem referência ao assunto principal receberão a qualificação máxima e, conforme se distanciem dele, esta irá diminuindo.

Não é muito habitual usar índices de ponderação em indexação. Costuma não ser necessário. Somente circunstâncias especiais e muito particulares aconselham o seu uso. Um caso do qual se tem notícia, onde foram utilizados, foi na indexação dos documentos relativos à reforma do ensino do Ministério de Educação e Ciência, o qual deu muito bons resultados.

INDICADORES DE LIGAÇÃO

Os indicadores de ligação são os que, em inglês, chamam-se *links*. Originam-se nos tempos em que os computadores eram dispendiosos e as fitas ou discos magnéticos também. Era preciso economizar espaço e utilizar o menor número de termos. Para isso, deveriam ser criadas maneiras de uni-los e evitar o uso de termos compostos, onde alguma de sua partes se repetisse. Também era e continua sendo útil nos tesauros unitermos onde é obrigatório combinar termos para expressar determinados conceitos.

Vejamos o seguinte exemplo: deve-se indexar um documento sobre "importação de ferro e cobre e o preço do cobre no mercado nacional".

Uma vez consultado o tesauro, foram encontrados os termos:

importados,
cobre,
ferro,
preços,
mercado nacional,

que coincidem com as palavras-chave do documento.

Para unir os conceitos, conforme convém, acrescentam-se a essas palavras umas notações que sirvam de ligação.

Por exemplo:

importação	A
cobre	A., B
ferro	A
preços	B
mercado nacional	B

É dessa foram que ficam formados os termos de indexação que devem figurar junto ao documento ou ao seu resumo.

Em alguns cados, foram construídos tesauros usando-se estes indicadores de ligação, como no *API-Thesaurus* para a indústria do petróleo.

Atualmente, as coisas podem ser simplificadas, pois os dispositivos de memória dos computadores admitem uma grande concentração de dados: por exemplo, os discos compactos e as previsões futuras, portanto será melhor formar tantos termos compostos quanto as necessidades o exijam, e tão extensos quanto seja preciso, para determinar os assuntos tratados.

No exemplo anterior, poder-se-ia escolher os seguintes termos:

importação de ferro,
importação de cobre,
preços do cobre,
mercado nacional;

e, inclusive

preços do mercado nacional,

se fosse um termo suficientemente comum.

INDICADORES DE FUNÇÃO

Chamam-se indicadores de função aqueles que, em inglês, recebem o nome de *roles*. Em outras publicações, usa-se a expressão "indicadores de função". Em castelhano, não é possível, pois estes são os que indicam o ponto de vista do que se considera um termo, como é detalhado no Capítulo 5. Da mesma maneira que os indicadores de união, têm sua origem nos primeiros tempos, quando se começou a utilizar as técnicas de indexação e os tesauros. Sua função consiste em relacionar as palavras-chave e, muitas vezes, também os termos, de forma que se saiba qual é sua posição em uma frase e, por conseguinte, possa-se determinar como se deve formar os termos de indexação e qual é seu lugar no tesauro.

Consideremos os seguintes exemplos:

1. Deseja-se indexar frases como as seguintes:
   - importação da Inglaterra de ferro de importação.

   Neste caso, deve-se fixar a posição da palavra "ferro".

   Será produto de partida:
   - ferro de importação.

   Será produto final:
   - importação de ferro.

Para distinguir ambas as posições, acrescenta-se uma notação, um indicador de posição, que as distinga:

Por exemplo:

- ferro	X Y
- importação	X
- importação	Y
- Inglaterra	Y

2. Outros exemplos significativos são aqueles nos quais se deve fazer a diferença entre expressões tais como:

2.1. Cadeias com anéis soldados

Cadeias soldadas com anéis

2.2. Recuperação da informação retrospectiva
Recuperação retrospectiva da informação

NOTA: Os exemplos são muito numerosos.

Estes tipos de expressões devem ser analisados cuidadosamente para que se torne possível captar o sentido do seu significado. No exemplo 2.2, talvez se diga o mesmo, porém expresso de uma das duas maneiras. O exemplo 2.1 não apresenta dúvidas.

No exemplo 2.1 e em cada um dos que o compõem, para diferenciar se são as cadeias que estão soldadas ou os anéis, acrescenta-se um indicador de função que o determina.

Consideremos como caso:

primeira hipótese:

- cadeias X
- anéis X Y
- soldagem Y

Assim ficam ligadas as cadeias aos anéis, e estes, por sua vez, ao processo de soldagem.

segunda hipótese:

- cadeias X Y
- anéis Y
- soldagem X

Desse modo, ficam ligadas também as cadeias aos anéis, porém o processo de soldagem faz referência às cadeias.

Todavia, deve-se citar outro uso dos indicadores de função, que é o empregado para deferenciar homônimos.

Por exemplo:

o termo minerais pode fazer referência a:

minerais	-	geologia	(A)
	-	mineração	(B)
	-	geografia	(C)
	-	economia	(D)
	-	joalheria	(E)

Para determinar a que campo pertence, acrescenta-se uma notação diferenciadora, que pode ser alfabética, numérica ou alfanumérica, conforme o estabelecido. No Capítulo 5, também se abordam os termos homônimos.

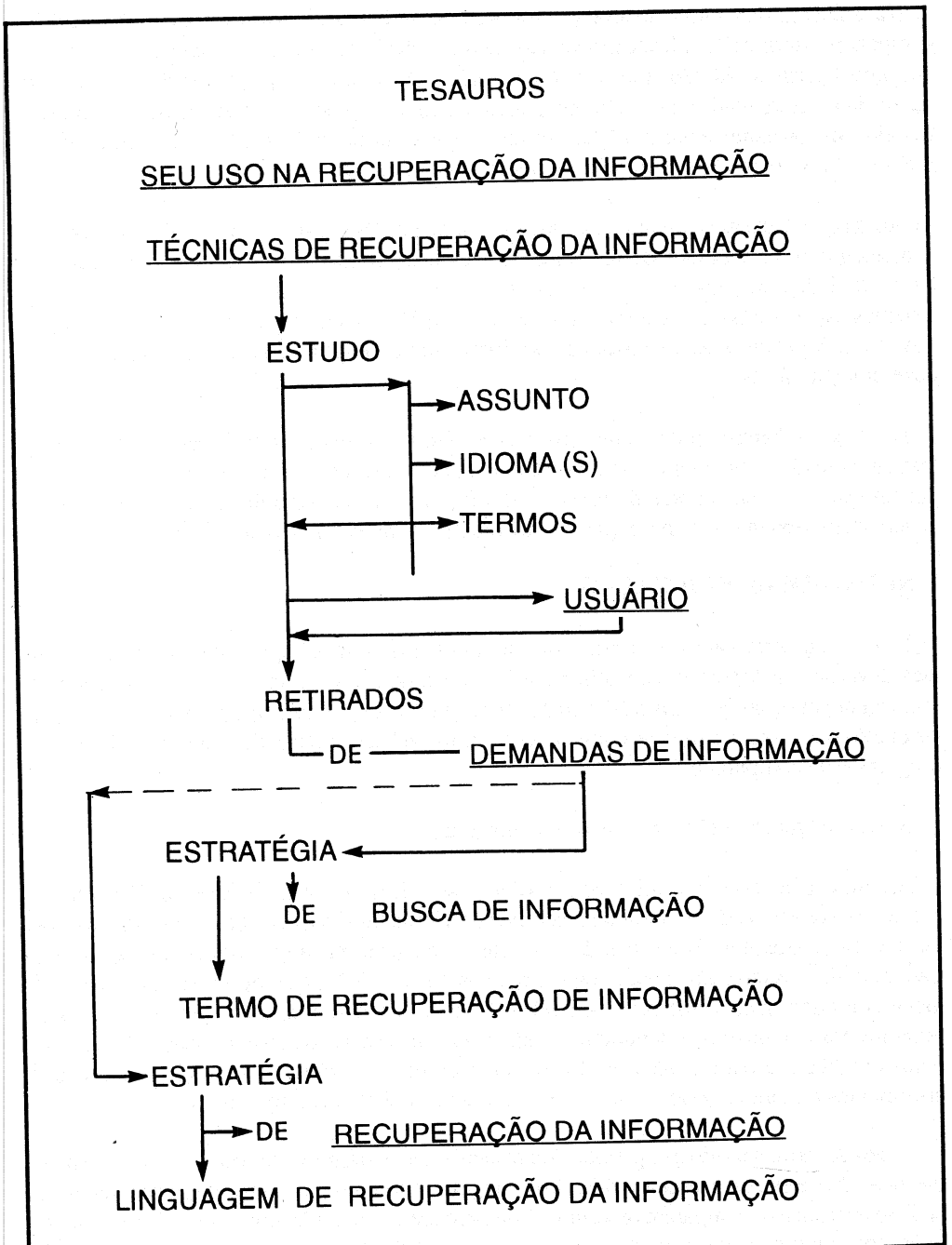


Figura 10.3. - Uso do tesauro na recuperação da informação.

Na prática, é feito uso moderado deste tipo de indicadores, pois se pode evitar fazê-lo, compondo os termos de indexação adequadamente, e não importa que se repitam certas palavras em vários termos. Atualmente, não é necessário economizar espaço nos dispositivos de memória de computador, mas, em compensação, deve-se precisar o mais possível o assunto tratado. Não obstante, existem alguns tesauros construídos com indicadores de função, como o *E.J.C. Thesaurus*, elaborado pelo Engineers Joint Council.

Ainda que pensemos, atualmente, que estes indicadores possuem pouca utilidade, porque contamos com os métodos da informática, seus programas de computador, estruturas das bases de dados etc., não nos aventuramos a dizer o mesmo relativamente ao seu uso nos sistemas especialistas. Aí, tanto os indicadores de união, como os de função e como outros novos que venham a ser idealizados adquirirão grande importância. Por conseguinte, não devemos ignorá-los.

Devemos salientar ainda, como observação geral, que tanto as notações que indicam o truncamento de termos como as de distanciamento entre os termos – como os índices de ponderação e os indicadores de união e de função – fazem parte do termo de indexação, devendo acompanhá-lo sempre que for utilizado para marcar a notação em questão.

LINGUAGEM DE INDEXAÇÃO

Uma vez determinados os termos que serão usados na indexação e qual a forma que se lhes deve dar, será possível construir com eles a linguagem de indexação. Esta será uma linguagem pela qual se "translada" a linguagem natural dos documentos para uma linguagem terminológica, com finalidade documentária utilizável em indexação e inteligível para os programas de computador.

NOTA: Algumas serão somente de uso manual.

Assim, estabece-se uma diferença entre tesauro e linguagem de indexação. O tesauro é construído considerando-se o conjunto possível de documentos concercentes a um determinado assunto, isto é, o assunto é considerado como caso geral, de forma impessoal, anônima, objetiva e atemporal. É usado como padrão ou matriz, como sistema de classificação de onde se deve tomar o modelo para construir a segunda. Já esta é elaborada com o conjunto de termos retirados dos documentos e devidamente adequados ao tesauro, ou tesauros disponíveis, para casos concretos e para grupos de documentos usados na prática. O assunto contido nos documentos é tratado, neste caso, de forma pessoal, subjetiva e temporal.

O fato de termos usado a expressão "ou tesauros", no parágrafo anterior, não deve chamar a atenção. No nosso caso, estamos falando no singular, referindo-nos a um tesauro de referência, onde se buscam e se comparam os termos. Compreende-se, contudo, que, caso seja necessário, serão consultados os tesauros disponíveis no serviço de documentação ou biblioteca onde esteja sendo feito o trabalho de indexação.

Tampouco deve chamar a atenção o fato de se dizer que se trata o assunto de forma subjetiva. Não estamos nos referindo à subjetividade do indexador, deixando-se influenciar por suas opiniões pessoais, mas ao assunto e a cada um de seus termos, que, neste caso, tornam-se protagonistas individuais e independentes.

Precisamente por causa deste caráter temporal, este tipo de linguagem necessita de revisões freqüentes para manter-se atualizada, de modo mais freqüente do que no caso do tesauro padrão ou matriz.

Em algumas obras de outros autores, denominam-se as linguagens de indexação de tesauros de indexação. Certamente podem ser consideradas como tais, considerando-se que se trata de um caso particular, de mais uma classe de tesauros. É muito importante advertir que um "tesauro de indexação" deve ser utilizado somente no centro de documentação onde foi construído, porque foi realizado de acordo com as necessidades particulares e temporais daquele local. O seu uso em um local diferente pode conduzir a erros. É natural que possam ocorrer situações onde o seu emprego seja, inclusive, recomendável; porém serão casos isolados e excepcionais, com características muito peculiares.

O TESAURO NA RECUPERAÇÃO DA INFORMAÇÃO

Como no caso da indexação, abordaremos a recuperação da informação, do ponto de vista da utilização do tesauro, e não em sentido geral (figura 10.3).

Situemo-nos no caso em que devemos atender a uma busca de informação formulada por um cliente. A situação, agora, muda de aspecto. Não é o documentalista que se defronta diretamente com os documentos, seus assuntos, suas formas de apresentação e de expressar os conceitos. Entre ambos, encontra-se o cliente, o chamado "usuário", que traz o seu próprio assunto e sua forma particular de expressar-se. Por conseguinte, a primeira coisa que se deve fazer é entender e interpretar a pergunta do usuário, levando-o a realizar previamente um estudo deste.

Para fazer um estudo do usuário deve-se considerar:

1. o seu conhecimento do assunto;
2. a sua forma de expressá-lo;
3. o tipo de linguagem que emprega;
4. o seu domínio do idioma – pensemos nos usuários estrangeiros;
5. o seu grau de instrução ou de profissionalismo.

Como se observa, todas essas características são muito semelhantes às que devem ser consideradas quando se estuda um autor para interpretar suas idéias. Também aqui, das respostas a essas questões, pode-se deduzir como o usuário emprega a linguagem.

Será possível determinar:

1. o domínio do assunto:
   – qual é o assunto de seu interesse;
2. o conhecimento da terminologia;
3. o uso concreto dos termos;
4. a especificidade com que se expressa;
5. a interpretação da pergunta;
6. a sua necessidade de informação.

Os pontos 5 e 6 merecem um comentário adicional, pois é freqüente o caso em que um usuário se apresenta com uma pergunta ambígüa, sem fornecer algum dado real ou alguma referência por onde seja possível começar a busca de informação, que também não se sabe em que consiste. É preciso, então, interpretar a demanda da informação, que pode requerer a formulação de uma série de perguntas adicionais, além de consumir um número considerável de horas.

TÉCNICAS DE RECUPERAÇÃO DA INFORMAÇÃO

Tanto a partir do estudo do usuário em si mesmo como de seu contato e relação com o assunto pelo qual ele se interessa, deverá ser elaborada uma **estratégia de busca de informação**. Do mesmo modo que no caso da estratégia de indexação, deve-se considerar o **conteúdo** e a forma como se deve delinear a busca. Também será aplicável aqui tudo o que foi dito para aquela e deverão ser consideradas as mesmas incidências. Geralmente, os centros de documentação têm confeccionado formulários, nos quais fazem constar os dados necessários para poder determinar as reais necessidades de informação de seus usuários.

Para realizar uma recuperação da informação do ponto de vista da utilização do tesauro, deverão ser seguidas as seguintes etapas:

1. realizar um estudo do usuário;
2. cumprir o formulário preparado para tal;
3. detectar as necessidades de informação;
4. preparar uma estratégia de busca, que consistirá em:
   4.1. formular os termos de busca de informação;
   4.2. compará-los com os que figuram nos tesauros;
   4.3. adaptá-los aos termos do tesauro;

5. formular o termo de recuperação da informação;
6. realizar a busca da informação no sistema documentário, base de dados, arquivos ...
7. interpretar a resposta obtida;
8. comprovar a eficácia da resposta.

Quando a resposta não for satisfatória, deve-se revisar todo o processo, para detectar onde se poderá encontrar o erro. Na maioria dos casos, trata-se do uso incorreto de algum termo. Nessa situação, voltar-se-á a traçar a estratégia de busca, modificando os termos que se consideram não adequados a essa busca de informação. Os sistemas conversacionais em linha facilitam estas tarefas.

Como se observa, é feita uma distinção entre

- termo de busca de informação e
- termo de recuperação.

O **termo de busca de informação** será aquele que se formula depois de considerar as características do usuário e de detectar as necessidades de informação. Quando esse termo tiver sido trabalhado de modo que possa ser usado na base de dados, por exemplo, ele, então, será convertido no termo de recuperação da informação.

Para realizar essa adequação, deverão ser seguidos os mesmos passos que foram estipulados para o caso da indexação. Aqui, deve-se insistir na comprovação da correta combinação dos termos para construir tesauros compostos. Um equívoco pode levar o documentalista a um grande fracasso.

Agora, também deverá considerar se é necessário acompanhar os termos de alguma das notações e estratégias descritas nas páginas anteriores. Por exemplo: pode ser muito útil empregar as técnicas de truncamento, ou as de ponderação, ou algum indicador de função, para ajudar a localização dos dados ou dos documentos buscados.

LINGUAGEM DE RECUPERAÇÃO DA INFORMAÇÃO

Também neste caso, com os termos de recuperação da informação, pode-se construir uma **linguagem** própria, que servirá exclusivamente para **recuperar a informação** buscada.

Será uma linguagem mediante a qual se "translada" a linguagem natural do usuário de um determinado serviço de informação para uma linguagem terminológica com finalidade documentária, que pode ser legível por procedimentos de informática e que se utiliza exclusivamente para a recuperação da informação.

Conclui-se que suas características principais aparecem no seu emprego para determinados usuários, em determinados lugares e com finalidade própria. Trata-se, igualmente, de uma

linguagem construída de forma pessoal, subjetiva e temporal. Portanto, será estabelecida a mesma diferença entre um **tesauro** e uma **linguagem de recuperação da informação**, como no caso das linguagens de indexação. Cada um deles possui uma finalidade distinta e, portanto, o seu uso também difere.

Na bibliografia sobre estes assuntos, encontram-se referências aos tesauros de recuperação da informação. Referimo-nos ao que foi explicitado no caso dos tesauros de indexação, que pode ser aplicado sem nenhuma modificação.

Observa-se que ambas linguagens apresentam muitas características e peculiaridades comuns. Aplicam-se as mesmas técnicas e seguem-se as mesmas etapas para a sua construção e uso. Surge, pois, a pergunta que faz com que se abordem as suas diferenças.

DIFERENÇAS ENTRE UMA LINGUAGEM DE INDEXAÇÃO E OUTRA DE RECUPERAÇÃO DA INFORMAÇÃO

Para estabelecer as diferenças entre estas duas classes de linguagens documentárias, deve-se fazer referência obrigatoriamente a tudo o que foi dito nas páginas anteriores. Com isso, pode-se formular as seguintes divergências:

1. a origem das palavras-chave;
2. as características dos impulsores de tais linguagens;
3. a adequação das palavras-chave aos termos que formarão as respectivas linguagens;
4. o conhecimento de sua eficácia;
5. o seu uso

   → primeira e principal diferença.

Examinemos os pontos enunciados. Quando se fala da origem das palavras-chave, faz-se referência à maneira como se apresentam diante do documentalista. No caso da indexação, estão contidas em uma série de documentos, de forma impessoal. Pelo contrário, quando se deve recuperar informação, elas são trazidas pelo usuário, de forma pessoal e, talvez, ambígua.

Isto está relacionado com as características dos "propulsores" de tais linguagens, que serão os autores dos documentos e os usuários, respectivamente. É mais fácil esclarecer essas características e as peculiaridades de um autor que as de um usuário. Este se mostra mais mutável. Hoje, formula sua pergunta de uma maneira e, amanhã, de outra. Também pode mudar de assunto. Sua situação é hesitante. E o cliente também nem sempre é o mesmo.

Para a adequação das palavras-chave à indexação, o indexador se defronta com um texto coerente e suficientemente extenso para poder compreender o seu significado com clareza. Portanto, será "relativamente" fácil precisar os termos que deve utilizar, como deve combiná-los, estabelecer suas relações semânticas e sintáticas etc. Em compensação, quem deve preparar

uma busca de informação depara-se com frases curtas, com palavras separadas, com expressões sem mais coerência que um assunto geral comum. Conseqüentemente, preparar os termos adequados para a recuperação pressupõe mais dificuldade e mais risco.

O indexador realiza o seu trabalho, e, logo após, remetem-se os termos escolhidos para o sistema documentário, por exemplo, para uma base de dados. Não alcança o resultado de seu trabalho, a eficácia da linguagem de indexação, até o momento em que alguém deva recuperar aqueles termos para responder uma busca de informação. A indexação é mais impessoal que a busca de informação, pois, neste caso, logo se conhece a utilidade da linguagem utilizada. Em grande parte, o êxito ou o fracasso depende da forma como o indexador tenha introduzido a informação no sistema, ou seja, que não somente se deve preparar a estratégia de busca de acordo com a demanda do usuário, como também com as técnicas de indexação usadas. Será, pois, mais difícil construir uma linguagem para recuperar que para indexar a informação.

No que concerne aos diversos usos dos tesauros, poder-se-ia dizer que a sua utilização correta repercute exclusivamente nos profissionais que devem utilizá-los. Independentemente das peculiaridades de cada linguagem, efetivamente as últimas conseqüências de um uso adequado, com suas divergências e suas semelhanças, repercutem em todos os que os utilizam. Normas podem ser estabelecidas, conselhos podem ser dados, etapas de trabalho podem ser criadas – que devem ser realizadas por quem? Pelas pessoas, neste caso, pelos profissionais das ciências da documentação. Portanto, em última instância, sobre estes recaem as responsabilidades do trabalho realizado. Desta maneira, deseja-se enfatizar a atenção que se deve dar à realização destas tarefas documentárias, importantes e fundamentais para determinar o êxito de qualquer sistema ou serviço de informação.

TEMAS PARA ESTUDO

1. Defina uma linguagem de indexação.
2. Acredita ser muito difícil o trabalho de um indexador?
3. Pensa que pode ser muito importante conhecer a nacionalidade do autor de um documento que se deve indexar?
4. Diga o que entende por termo de recuperação da informação.
5. Diga quando usaria um macrotesauro.
6. Na sua opinião, é importante conhecer a pessoa que demanda informação?
7. Diga o que é um indicador de função.
8. Explique para que usaria o truncamento de termos para a direita.
9. Diga o que entende por tesauro de recuperação da informação.
10. Diga a diferença que encontra entre um tesauro e uma linguagem de indexação.
11. Quando um tesauro sistemático terá utilidade?
12. Cite alguma particularidade de uma busca de informação.
13. Esquematize uma técnica de indexação.
14. Na sua opinião, é útil um estudo detalhado dos usuários?
15. O que faria com uma busca de informação ambígua?
16. Dê sua opinião sobre os indicadores de união.
17. Utilizaria alguma vez os índices de ponderação?
18. Para que serve um tesauro matriz?
19. Na sua opinião, qual é a influência que pode ter na indexação quando o autor é estrangeiro?
20. Esquematize uma estratégia de recuperação da informação.

REFERÊNCIAS BIBLIOGRÁFICAS

CARIDAD SEBASTIAN, M.: "La teledocumentación; Madrid, Forja, 1984, 446 pp.

CURRÁS, E.: "Thesaurus", en Tipos de documentos científicos. Métodos de Clasificación, Rev. Química e Industria, *21* 3, Marzo 1975, 269.

DUBOIS, C.P.R.: "The use of thesauri in online retrieval"; J. Inf. Scien., 8, 2, March 1984, 63-66.

JOHNSTON, S.M.: "Effect of thesaurus indexing on retrieval from machine-readable destabases"; Quarterly Bull. of IAALD, 27, 3, 1982, 90-96.

RODRIGUEZ RICARD, T: ''Epigrafiario: Lenguaje de Indización de la Biblioteca Nacional de Ciencia y Técnica"; Act. Inf. Cient. Técn.,*1* (144), 1987, 38-53.

ROZO RIOS, D.S., y GARZON HERRAN, H.E.: "Tesauro en Documentación e Información"; Bogotá, FID-CLA, ICFES, FID 587, 1980.

SALRA, X., LOPEZ DE MANTARAS, R.: ''Desarrollo de técnicas de inteligencia artifical de soporte a gestión de fondos documentales (catalogación y recuperación)"; Proyecto de investigación. Departamento de Matemáticas de la Facultad de Informática. Centro de Cálculo de la UPC, Universidad Politécnica de Cataluna, 1985.

SNOW, B.: ''Why use a database thesaurus?"; Online, 9, 6, November 1985, 92-96.

TURUGUET MAYOL, D.: "CDU frente a Tesauro en la indización temática para la automatización de una biblioteca científico-técnica", en Actas de las II Jornadas Espanõlas de Documentación Automatizada, Málaga, 1986, 275-285.

NOTA: Devem ser consultadas também obras de tipo geral resenhadas nos capítulos dedicados ao assunto tesauros, do capítulo 4 ao 10.

Epílogo

EM QUE TERMOS IREMOS NOS ENTENDER?

Que a luz do entendimento
não se torne obscurecida
pela ineficácia da língua.

E. C.

Seguramente, quando os homens começaram a descer da Torre de Babel, confusos, desorientados e sem conseguirem fazer-se entender com seus semelhantes, teriam feito a seguinte pergunta uns aos outros: Que termos, que palavras deveriam utilizar para compreender e fazer-se compreender, bem como comunicarem-se com aqueles com os quais, infalivelmente, deviam relacionar-se?

Os pobres apóstolos, pessoas simples, diante da sua missão de percorrer o mundo, também se fariam esta pergunta, e não somente com relação aos próprios termos – palavras – que deviam utilizar, mas também ao seu significado e sua intencionalidade, à mensagem que tinham sido chamados a transmitir.

Os gramáticos de Port Royal, no século XVIII, fizeram-se igualmente esta pergunta, no propósito de buscar soluções para a profusão de línguas que começavam a ser empregadas na divulgação das pesquisas e avanços científicos do momento.

Os políticos de todos os tempos, quando desejam se dirigir às pessoas para pedir o seu voto ou o seu beneplácito, procuram escrupulosamente os termos pelos quais devem dirigir-se a elas.

Os profissionais, estudiosos ou pesquisadores procuram, também, buscar os termos que devem usar para ser compreendidos, ainda que nem sempre o consigam; e os pensadores, filósofos e escritores encontram-se na mesma situação.

Não podemos deixar de mencionar os documentalistas, quando se encontram diante da pergunta de um usuário ansioso, a qual deve ser respondida com rapidez, eficácia e relevância. Este documentalista enfrenta o dilema de decidir qual o termo que deve ser utilizado, para que o sistema responda adequadamente.

A dúvida se acentua no caso em que o sistema documentário é mecanizado e deve-se recorrer a determinados artifícios para que a máquina consiga compreender a pergunta formulada e responda de modo inteligente.

O tradutor e o profissional da informática, da mesma maneira, perguntam-se como devem utilizar um termo, ou que interpretação devem dar a um outro.

Uma pergunta como a do título deste trabalho, à primeira vista, muito simples, provoca situações e respostas diferentes, a que, todavia, devemos, indesculpavelmente, poder responder e solucionar, caso queiramos, realmente, fazer-nos compreender por nossos semelhantes – compreensão essa imprescindível para suportar a vida associativa. Em outros trabalhos, declaramos que, sem comunicação, sem intercâmbio de informação, nem transferência de experiências, a vida do ser humano é impossível. É bem verdade que um indivíduo isolado em uma ilha deserta, onde haja luz, ar e alimentos, pode desenvolver sua existência física de modo mais ou menos precário; porém seu cérebro não poderá desenvolver-se na mesma proporção que o faria, se o indivíduo vivesse em comunidade e expressasse suas experiências por meio da linguagem, veículo de comunicação com o mundo exterior.

A linguagem é essencial para o ser humano. Supõe-se que, na evolução do Homonóide ao Homo Sociologicus – *Homo sapiens*, produziu-se, igualmente, uma evolução com seus semelhantes. Acreditava-se que, em princípio, emitia sons disformes, acrônicos e discordantes e que, pouco a pouco, esses sons foram se aperfeiçoando, até chegar o momento atual. Na minha concepção, a linguagem é uma característica a mais da evolução do ser humano no Planeta Terra, produzida segundo leis preestabelecidas em determinados momentos desse processo evolutivo. Segundo esta teoria, a linguagem progride ao mesmo tempo em que o homem o faz. Se este é cada vez mais perfeito e complicado, aquela também deverá sê-lo. E, se o indivíduo caminha até uma degeneração em seu processo existencial, a linguagem também declinará.

Por outro lado, se podemos demonstrar que a linguagem se encontra em período de esplendor, teremos de admitir que a espécie humana encontra-se, também, em pleno auge. Que não poderia ser capaz de assegurá-lo, muitos caminhos ainda devem ser percorridos, e muitos estágios de desenvolvimento que não podemos nem mesmo prever deverão ser cumpridos. Os 10, 20, 25 mil anos – tais diferenças não vêm ao caso– de evolução recente não trazem consigo mais que um pequeno elo em todo o conjunto do processo vital de nosso Planeta.

Não obstante, tão diminutos, tão insignificantes diante do Universo e tão importantes e fundamentais diante de nós mesmos, devemos viver nossos momentos de cada dia e devemos nos relacionar continuamente com as pessoas com as quais – gostemos ou não – estamos obrigados a nos entender, usando a linguagem.

A CONFUSÃO DAS LÍNGUAS

A linguagem não é única nem comum para todos. O seu estudo torna-se necessário para procurar uma compreensão melhor de nós mesmos e do mundo que nos cerca. Surgem, assim, a lingüística, por um lado, e a terminologia, por outro.

Até aqui, demos um longo passeio para chegarmos a citar a terminologia. Não nos proporemos a defini-la de maneira formal, pois foi definida profusamente por aqueles que possuem mais autoridade que eu para fazê-lo. Basta-nos ressaltar a sua importância como disciplina com características ontogênicas, imprescindível não somente para a compreensão entre profissionais dedicados ao trabalho ou estudo de um determinado assunto, como também, em geral, para o desenvolvimento do indivíduo como ser humano dentro de sua espécie. Também, igualmente como indivíduo na comunidade da qual participa.

Hoje, talvez mais do que nunca, ou talvez hoje novamente, repita-se a situação da confusão das línguas. As interpretações míticas falam de um castigo divino por se desejar ir além do limite permitido. Outras interpretações que podemos dar estão centradas precisamente nesse processo evolutivo do ser humano em seu caminho em direção à paz do planeta. Chega um momento em que a quantidade de informações que recebe em seu cérebro – estamos nos referindo ao ser humano como espécie, não como indivíduo isolado – é tão grande, que supera a sua capacidade de compreensão, novas formas de linguagem. Surgem novas idéias, novos conceitos, novas situações, novas palavras que cada subgrupo humano, segundo as características físicas de formação de seu organismo, expressa de maneira diferente. Se esta for, realmente, uma explicação convincente do aparecimento das diferentes linguagens, ou línguas, ou idiomas, de cada raça ou povo, é algo por demonstrar, como tantas outras teorias para as quais faltam bases experimentais para apoiar-se, mas que, ao menos, não carecem de argumentação lógica e coerente.

A questão seguinte seria analisar se a situação de confusão das línguas em que nos encontramos atualmente pode ser imputada a um processo semelhante. Na minha opinião, pelo menos, é parecida. O aumento de informação tornou-se maior nos últimos tempos. As necessidades de comunicação, conseqüentemente, também. As pessoas – seja em sua vida profissional ou cotidiana – buscam formas de expressão que estejam em concordância com sua própria idiossincrasia e utilizam suas linguagens, seus idiomas próprios. Desse modo, aparecem cada vez mais escritos em línguas vernáculas de diferentes povos e etnias, e, desta maneira, inclusive criam-se gramáticas e estudos lingüísticos para linguagens que, antes, eram consideradas simples dialetos ou modismos idiomáticos.

Talvez nos encontremos em momentos como aqueles que deram lugar à criação de gramáticas para as línguas romances que iam se formando ou, em épocas posteriores, com a elaboração de regras gramaticais-formais para fixar um forma única de entendimento entre todos os povos germânicos.

Seja como for, as línguas, os idiomas, estão aí, e os escritos nesses idiomas também. E há de se encontrar a forma de entendê-los. Este é o motivo pelo qual fazemos referência à confusão de línguas e não à profusão de línguas. E surge-nos também a dúvida se o mito do castigo voltará a se repetir, pelo fato de querermos ir, com as máquinas inteligentes, mais além dos limites permitidos.

O DOM DO DOMÍNIO DAS LÍNGUAS

Conforme venho me expressando, é como se fosse contra o direito natural de cada indivíduo expressar-se em sua própria língua. Contudo, esta não é minha intenção. Cada um sabe o que quer dizer e o que diz utilizando as palavras que lhe parecem essenciais, que aprendeu na sua infância e que conserva impressas em sua informação genética. É curioso que os alemães chamem o seu próprio idioma de *muttersprache* – idioma da mãe – e que, em outros países, receba o nome de linguagem do país, do país de nascimento, pressupõe-se.

Precisamente essa **confusão** ou **profusão** de línguas nos obriga a falar e poder nos expressar em mais de um idioma, sobretudo nas zonas geográficas onde também se utiliza outro idioma com caráter oficial. Falar – e, por pressuposto, ler e escrever – vários idiomas foi considerado desde sempre como um dom, um privilégio. Nem todos os indivíduos são capazes disto, ao menos não de fazê-lo com certa facilidade e desenvoltura. Uma pessoa que pode expressar sua idéias, pensamentos e sentimentos em vários idiomas revela que o seu cérebro está em um grau de desenvolvimento um pouco mais elevado que seus semelhantes. Pelo menos, a sua capacidade de memorização é maior, porque lembra os diferentes vocabulários e acepções próprias de cada idioma. Pode ser que seu ouvido o traia no momento de procurar adquirir uma pronúncia correta, porém esta é outra questão.

De acordo com o que foi dito, o conhecido fato de que os cidadãos dos países do Leste Europeu possuem esta facilidade para o aprendizado de idiomas – há alguns que falam e escrevem até 14 idiomas – revelaria a sua maior capacidade de inteligência por ter o cérebro mais desenvolvido. Talvez seja assim, talvez não.

A necessidade crescente de aprender ao menos um segundo idioma também implicaria, segundo o que foi dito anteriormente, que os seres humanos contemporâneos tendessem a uma maior capacidade de inteligência. Talvez seja assim, talvez não. Não devemos esquecer ainda o que foi exposto em outros trabalhos meus, onde se atribui o maior grau de desenvolvimento intelectual dos indivíduos de nosso tempo e sua evolução para o aumento progressivo, ao impacto causado pelos sucessivos e inumeráveis quanta de informação útil que bombardeiam continuamente seu cérebro. Efetivamente, recebemos essa informação, contudo ela não fica aprisionada nos neurônios. É elaborada e emitida, logicamente, por meio da linguagem, que, atualmente, devemos apostrofar como "natural".

UNIFICAÇÃO DAS LÍNGUAS

Como se sabe, uma ação produz uma reação, e o desejo de cada um de expressar-se em seu próprio idioma originou a idéia de poder dispor de uma língua comum na qual, ao menos para assuntos profissionais, sejamos capazes de nos entender em nível nacional e internacional. Em nosso mundo ocidental, em tempos passados, foi o grego e, posteriormente, o latim a língua comum. Atualmente, ambos caíram em desuso, por maiores que tenham sido os esforços feitos pelos professores. Trata-se de línguas mortas, as que, além desse uso, não apresentam utilidade. Por outro lado, suas gramáticas, sintaxe e outras regras lingüísticas resultam demasiadamente complicadas. Hoje em dia, as pessoas não estão dispostas a dispender esforço mental, tempo e dinheiro em algo que se considera retrógrado e obsoleto.

Ao longo dos últimos tempos, foram feitas tentativas para criar idiomas comuns internacionais, dos quais o mais destacado foi o esperanto, por ser o mais difundido e contar com maior número de adpetos. Tais esforços também não tiveram êxito, talvez pelos mesmos motivos anteriores. Neste caso, tratava-se de linguagens elaboradas artificialmente, com o propósito de contentar e unificar critérios das pessoas dos países considerados importantes naquele momento. Ou seja, línguas mortas, complicadas, de uso restrito, sem utilidade prática.

O INGLÊS, PASSAPORTE PARA A CELEBRIDADE

Outro motivo que fez com que as línguas internacionais comuns não prosperassem foi a difusão da tecnocultura anglo-saxônica, principalmente nos Estados Unidos da América do Norte. Os estadunidenses irromperam na vida profissional com toda a força de sua jovem cultura e trouxeram com eles o inglês, além de seu desinteresse em aprender outras línguas.

O Inglês é uma língua que oferece muitas vantagens de adaptação aos tempos modernos e para converter-se em idioma universal. Pensemos que sua gramática é muito simples, que novas palavras podem ser construídas com o acréscimo de prefixos, utilizando os gerúndios e unindo dois ou mais vocábulos. Sua sintaxe também é simples e facilmente exeqüível a uma análise para processamento de textos por computador. Além do que, é falado por um número considerável de milhões de habitantes que possuem a sagacidade de colonizar os continentes sem que estes mesmos tomem consciência do fato.

A nós, espanhóis de fala castelhana, causa-nos um certo desconforto constatar que o inglês é o idioma que se impôs no mundo de maneira superior ao espanhol. O espanhol é uma língua viva, sonora, com um vocabulário rico e uma gramática estruturada. É falado também por milhões de pessoas, em vários continentes. Por tais razões, a União Latina desenvolve esforços para manter o espanhol em destaque, promovendo a comunicação entre os países de fala espanhola, atitude elogiável e digna de todo o apoio.

Todavia, não podemos fugir da realidade. Se alguém quiser ser internacionalmente conhecido e compreendido, deverá expressar-se por meio do inglês. Em vários trabalhos recentes, onde se defende a utilidade da tradução automática, esta é justificada manisfestando-

se a necessidade de traduzir os trabalhos para o inglês e deste para outros idiomas. Inclusive, chegou-se a dizer que as pesquisas ou os resultados científicos que não forem armazenados em bases de dados elaboradas por ingleses ou norte-americanos – naturalmente em inglês – perdem-se para o acervo cultural e científico da humanidade.

Avaliaram-se, estruturaram-se e planejaram-se numerosos sistemas de tradução de alguns idiomas a outros. Em alguns casos, diretamente, isto é, sem passar pelo inglês com intermediário – por exemplo, na Alemanha, na Espanha e na França – estes sistemas de tradução permanecem em seu âmbito territorial, com uma clientela muito restrita, limitada e especializada. Na maioria dos casos, busca-se a informação científica, técnica ou humanística nas grandes bases de dados internacionais elaboradas em inglês. Se o trabalho de determinado autor não figurar ali, será conhecido apenas por um número reduzido de pessoas.

No auge e no impulso da tecnocultura estadunidense e, em menor escala, canadense, no momento em que o inglês se impõe como língua internacional – como para assuntos de turismo, por exemplo –, aparece, em situação não menos considerável de auge e impulso, a tecnocultura japonesa. Para o Japão, o resto do mundo mostra-se mudo e surdo. Sobretudo, estes parecem mudos e surdos, pois, idiomaticamente, situam-se em outras esferas de compreensão. Como melhor veículo para se mostrarem e introduzirem nos mercados internacionais, escolheram o idioma que, junto à vantagens mencionadas anteriormente, já desfrutava de preferência e importância: o inglês.

Pouco a pouco, foram somando-se a esse processo os chineses, indianos, coreanos... Desse modo, pois, não podemos ir contra a corrente, ainda que mantenhamos o espanhol como base para o entendimento entre os povos de língua espanhola, onde quer que se encontrem, em qualquer lugar do planeta.

O ASSUNTO DA TRADUÇÃO AUTOMÁTICA

Encontramo-nos, portanto, diante da questão da tradução. Seja como for, a única saída é recorrer a ela, hoje e sempre. Não se trata de coisa atual. Tenho a impressão de que o homem sempre precisou valer-se das diferentes formas de expressão de outros indivíduos para poder ser compreendido por estes. Contudo, hoje muito mais do que antes, é dada uma atenção substancial à tradução. Por exemplo, somente os escritórios da Comunidade Européia necessitam traduzir todos os documentos de 12 idiomas. E é lógico que se pense em utilizar as máquinas disponíveis e que se planejem programas, códigos e métodos para realizar as traduções.

À primeira vista, a questão parece simples, porque se deve somente buscar a equivalência em outro idioma diferente. Se tudo fosse tão simples! Em primeiro lugar, nem tudo se chama da mesma maneira nos diversos idiomas. Em segundo lugar, não se trata de palavras colocadas uma ao lado das outras, mas de compreender o significado desse conjunto e o que é mais importante: compreender a sua intencionalidade.

Os sistemas para realizar traduções automáticas com os recursos oferecidos pela informática convencional não chegaram a dar resultados satisfatórios. Por esse motivo, recentemente, começou-se a estudar uma forma de adotar os procedimentos empregados pela inteligência artificial, planejando-se sistemas especialistas para o tratamento da linguagem natural e o processamento de textos.

Trata-se de métodos que estudam a sintaxe e a semântica das palavras e das frases, para decompô-las em sintagmas simples, para os quais se estabelece correlações com outras estruturas semelhantes, em outros idiomas. É uma técnica delicada, minuciosa e meticulosa, para a qual é necessário possuir não somente qualidades específicas, como também conhecimentos de lingüística, matemáticas, terminologia, psicologia, informática, sociologia e, talvez, de algum outro campo do conhecimento humano. Muitos são os estudos recentes sobre o tratamento da linguagem natural e o tratamento de textos. Para citar alguns, mencionaremos os que estão sendo feitos na Espanha e que deram resultados concretos, como, por exemplo, o *Diccionário del Castellano em Ordenador para la Composición y Verificación de Textos*, realizado pela equipe dirigida por Luis Sopeña, do Centro de Investigación da Universidad Autónoma de Madrid – UAM e a empresa IBM; e o *sistema interativo de apoio à construção de tesauros*, realizado por Eric Plaza, no Instituto de Estudios Avanzados de Blanes, do Consejo Superior de Investigaciones Científicas de Cataluña.

No campo próprio da terminologia, temos o prazer de mencionar o *Catálogo de Recursos Terminológicos en Lengua Española*, de 1987, que contém uma relação exaustiva dos recursos terminológicos existentes na Espanha. É uma ferramenta indispensável para os terminólogos, bem como para os tradutores, e foi auspiciado pela Unión Latina, publicado por esta mesma entidade e realizado pelo grupo de terminologia do Cindoc, em Madri.

Também realizaram-se estudos sobre tradução automática, para os quais se encontram em fase de elaboração e experimentação alguns sistemas especialistas. Trabalha-se em colaboração com instituições estrangeiras e foram depositadas grandes esperanças em seus resultados.

REFERÊNCIAS BIBLIOGRÁFICAS

Bancos terminológicos, estudio realizado por el Grupo de Trabajo sobre Lenguajes de Indización Centro Argentino de Informaci6n Cientifica y Tecnológica - CAICYT; Rev. Esp. Doc. Cient. *2*, 1, 1986, 57-62.

BOSCH DE ROZE, A.: "Lingüistica y Ciencias de la Información: un acercamiento integrador"; Rev. Esp. Doc. Cient. *2*, 2, 1986, 167-171.

CALVO HERNANDO, M.: "Periodismo cientifico y respecto a la Ciencia"; Quimica e Industria, *33*, 5, 1987, 400.

CASAJUANA CASANOVAS, R., y RODRIGUEZ MAGRO, C.: "Clasificación de los verbos castellanos para un diccionario en ordenador"; Actas del I Congreso de Lenguajes Naturales y Lenguajes formales, Universidad de Barcelona, Barcelona, 1984.

Catálogo de recursos terminológicos en lengua espanola, Madrid, CSIC, 1987.

Coloquios con Academias de Ciencias Hispanoamericanas sobre Terminologia Científica e Historia de la Ciencia; Real Academia de Ciencias Exactas, Físicas y Naturales, Madrid, 20-30 de abril de 1976.

CURRÁS, E.: "La nécessité d'une coopération hispano-française en terminologie chimique"; Documentaliste, *18*, 1, 1981, 7-9.

CURRÁS, E.: "La información como cuarto elemento vital y su influencia en la cultura de los pueblos"; Discurso de Ingreso en la Real Academia de Bellas Artes y Ciencias Históricas de Toledo; Toletum, marzo 1986.

CURRÁS, E.: Neuester Stand der Forschung zu Experten Systemen bei Anwendung im Gebiet der Informadionswissenchaft in Spanien; in 15. Kolloquium über Information und Dokumentation, November 1987, Ilmenau.

FERNANDEZ RIOS, L., y BUELA CASAL, G.: "Educación de la inteligencia: posibilidades y límites"; Madrid, Organizacion de Estados Iberoamericanos para la Educación, la Ciencia y la Cultura OEI, 1986.

GOETSCHALCKX, J.: "Rapport de synthèse". Deuxième Symposium d'Infoterm. Vienne 14-17 avril 1985; TermNet, *15*, 1986, 3.

GORKOVA, V.I.: FID Terminology of Information and Documentation Committee (FID/DT); Int. Forum Inf. and Docum. *11*, 3, 1986, 39-42.

INTERTEXT Fremdsprachendienst der DDR; TermNet, *15*, 1986, 22.

JAÑEZ ESCALADA, L.: "Inteligencia artificial: algunas aplicaciones prácticas"; FUNDESCO, 67, 1987, 4.

JARVELIN, K. and REPO, A.J.: "A Taxonomy of Knowledge Work Suppor Tools"; in Chalenges to an Information Society. Proceedings of the 47th. AIS Annual Meeting, vol. 21, Philadelphia, PA, Oct. 21-25, 1984, Knowledge Industry Publications, Inc., White Plains, New York, 59-62.

Jornada Corporativa sobre el tema Sistemas Expertos para la Información y Documentación; Madrid, C.I.T.E.M.A., Memoria 1986, 86.

14. Kolloquium über Information und Dokumentation; Vom 26. bis 28. November 1985, Ilmanau, 1985.

KOTOV, R. G.: "Language in Real Communication as an Object of Applied Linguistics", Int. Forum inf. and. Doc. *2*, 4, 1984, 17-21.

LOEBEN, M. and RUNGE, W.: Spranchbarrieren bei der natuwissenschaftlichtechnischen Fachkommunikation. Ergebnisse einer Befragung; Nach. f. Dokum., *36*, 3, 1985, 127-136.

MARCHUK, Y. N.: Machine Translation in the USSR: Int. Forum Inf. and Doc., *2*, 2, 1984, 3-8.

SOARES, R. M.: "Lingüística y Terminología. Traducción y Normalización"; lección prepada para el curso de Doctorado sobre Metodología y Práctica de la Información Científicas y la Documentación Informatizada, organizado por el Gabinete de Documentación Científica de la Universidad Autónoma de Madrid. Curso 1986/87.

SOPEÑA, L.: "Dccionarios del castellano en ordenador para la composición y verificación de textos"; Madrid, Centro de Investigación U.A.M.-IBM, 1987.

Terminology and Knowledge Engineering; TermNet, *15*, 1986, 1-2.

VERDEJO, M.F.: "Interfaces en lenguaje natural"; Primeiras Jornadas de Inteligência Artifi cial para Bases de Dados, Barcelona, 17-18 octubre, 1985.

WRIGH T, L.D. Jr.: "World Write and Integrated Term Software"; TermNet, *15*, 1986, 13.

ÍNDICE DE ILUSTRAÇÕES

Figura 1.1 Ser humano — Terminologia ——— 16

Figura 1.2 Linguagem — Linguagens ——— 19

Figura 1.3 Terminologia — Processo de elaboração do microcosmo ——— 20

Figura 1.4 Conceito de terminologia ——— 24

Figura 1.5 Utilidade da terminologia ——— 26

Figura 1.6 Do processo mental ao conceito ——— 27

Figura 1.7 Circuito: Conceito — Informação ——— 29

Figura 1.8 O termo ——— 33

Figura 3.1 As indústrias da língua ——— 52

Figura 3.2 Banco de dados terminológicos ——— 59

Figura 3.3 A tradução automática ——— 64

Figura 3.4 Normalização em terminologia ——— 67

Figura 4.1 Linguagem terminológica — Tesauro ——— 79

Figura 4.2 Linguagens terminológicas — Tesauro ——— 82

Figura 4.3 Os tesauros ——— 89

Figura 4.4 Classes de tesauros ——— 94

Figura 5.1 Componentes fundamentais de um tesauro ——— 104

Figura 5.2 Inspec - thesaurus, 1987. The Institute of Electrical Engineers ——— 112

Figura 6.1 Modelo de formulário para a construção de tesauros I ——— 133

Figura 6.2 Modelo de formulário para a construção de tesauros II ——— 136

Figura 6.3 Broad subject categories. Catering and related fields ——— 149

Figura 6.4 Allocation of terms to broad subject categories. Catering and related fields ——— 150

Figura 6.5 Multi-level alphabetical thesaurus. CAB thesaurus ——— 151

Figura 6.6 Conventional alphabetical display. Thesaurus of engineering and scientific terms (TEST) ——— 160

Figura 6.7 B/H Ciencia y Tecnología. Tesauro de la Unesco ——— 161

Figura 7.1 Exemplo de índice alfabético que faz referência às classes a que o termo pertence, sem mais especificação ——— 166

Figura 7.2 Relação permutada descritores. Índice espanhol ——— 168

Figura 7.3 Subject field analysis and thesaurus derivation. Catering and related fiels ——— 171

Figura 7.4 Faceted classification ——— 172

Figura 7.5 Family tree structure. ISDD thesaurus ——— 175

Figura 7.6 Family tree structure. PRECIS ——— 176

Figura 7.7 Graphic display. Arrowgraph. IRRD thesaurus ——— 177

Figura 7.8 Graphic display. Terminograph. EUDISED thesaurus ——— 178

Figura 7.9 Representação circular. Apresenta os termos distribuídos em círculos, cada um dos quais apresenta um nível de hierarquia, com o termo mais específico no centro ——— 179

Figura 7.10 Polygons - Euratom ——— 180

Figura 7.11 Diagrama flechado - Euratom ——— 181

Figura 8.1 Tesauros Multilíngües ——— 188

Figura 8.2 Seção alfabética do EUDISED - Tesauro multilíngüe ——— 197

Figura 8.3 Seção alfabética do tesauro SPINES da Unesco ——— 198

Figura A.1 Confecção de um tesauro estruturado para química ——— 201

Figura A.2 Confecção de um tesauro estruturado para química ——— 203

Figura A.3 Confecção de um tesauro estruturado para química ——— 205

Figura 9.1 Vocabulary Switching System. Thesaurus Construction ——— 214

Figura 9.2 Descriptor bank entry. "Integrated thesaurus of the social sciences". Thesaurus construction ——— 215

Figura 9.3 Integração - Desdobramento - Manutenção de tesauros ——— 217

Figura 9.4 Publicação de tesauros ——— 223

Figura 10.1 Técnicas auxiliares no uso de tesauros ——— 234

Figura 10.2 Uso do tesauro na indexação ——— 239

Figura 10.3 Uso do tesauro na recuperação da informação ——— 245

ÍNDICE DE AUTORES

AENOR. 66, 68.

AENOR. CTT-50 Documentação. 66,67.

AFTERM. 45, 61.

ASLIB. 227.

Aitchison. 99, 128, 183, 231.

Albornoz, O. 35.

Alonso Baquer, M. 50.

Alto Comissariado da Língua Francesa. 43.

Álvarez, S. 35.

Álvaro Bermejo, C. 99, 163.

Anxolabehere, V. 75.

Austin, D. 183.

B.S.O. - Broad System of Ordering. 163.

Beaumont, J. F. 70.

Bordonaba, V. 74

Bosch de Rozf, A. 34, 262.

Boyce, B.R. 127.

Brewin, P. 183.

Brown, H. 99.

Buela Casal, G. 262.

CDU - Classificação Decimal Universal. 68.

CEPAL. 45.

CLADES. 45.

CONICYT - Comissão Nacional de Investigação Científica e Tecnológica. 183.

Cabré I Castellvi, T. 50.

Calvo Hernando, M. 34, 262.

Casajuana Casanovas, R. 262.

Caridad Sebastián, M. 253.

Carrión Gútiez, M. 99.

Castro, E. de. 72.

Centenero Gallego, M. A. 34.

Colégio do México. 45.

Comité UNE/TC 1. 66.

Cotton, J. 183.

Criado del Val, M. 48.

Currás, E. 34, 98, 163, 209, 253, 262.

Czap, H. 21.

Chappuy, S. 75.

Chaumier, J. 163.

Chose, A. 207.

Dahlberg, I. 14, 30, 44.

Defriez, P. 183.

Deutsche Institüt-für Normung. 41.

Deweze, A. 99.

Drezd. 42.

Dubois, C. P. R. 253.

EURODICAUTOM. 73.

EUSKALTERM. 48.

Ema, E. 74.

Escorial, M. 35.

Escola da U.R.S.S. 43.

Escola de Praga.43.

Escola de Viena. 42.

Escola Canadense. 43.

Escola Francesa. 43.

Escola Húngara. 43.

Escola Nórdica. 43.

Estienne, Ch. 91.

Evangelista, C. 99.

FID - Federação Internacional de Informação e Documentação. 66.

FID/DT. 66.

FUINCA. 95.

FUNDESCO. 95.

Fedor de Diego, A. 34, 50.

Felber, H. 25, 35.

Fernández García, J. 207.

Fernández Rios, L. 262.

Figueiredo Castro, R. C. 73.

Fillón, L. 70.

Fondazione Ugo Bordoni. 50.

Fondin, H. 207.

Foskett, D. J. 93, 163.

Fugmann, R. 99.

Fujikawa, M. 73.

Galindo, F. 163.

Galinski, Ch. 14, 21, 28.

García Camarero, E. 70.

García Gutiérrez, A. L. 99.

Garrido, J. 47.

Garzón Herrán, H. E. 253.

Gavira Gulpe, C. 163.

Gemersal, A. 163.

Gesellschaft für Information and Dokumentation (GID). 128.

Gilchrist, A. 85, 92, 128.

Goetschalckx, J. 35, 262.

Gonella, D. 75.

Gonzáles Suárez, E. 36.

Gonzáles Abascal, J. 35, 70.

Gorkova, V. I. 50, 262.

Grupo de Terminologia do ICYT (TermEsp). 30.

Guilbert, L. 21.

Güntzer, W. 70

HISPANOTERM. 48.

Hanada, T. 35.

Hernando Colón. 95.

Howerton. 95.

Hutchins, W. J. 75.

ICYT - Instituto Nacional de Informação y Documentação. 22, 93, 226.

INFOTERM. 45.

IRANOR. 66.

ISO - Internacional Standard Organization. 66.

ISO/TC 37. 66.

ISO/TC 4. 66.

ISO/TC 46. 66.

ISO/TC 46/SC 3. 66.

ISOC - Instituto de Informação e Documentação em Ciências Sociais e Humanidades. 95.

Ilión, A. 207.

Instituto da Língua Francesa. 40.

International Electrotechnical Commission. 41.

Ishikawa, t. 74.

Irazazabal, A. de. 35, 50, 73.

Ireland, R. 183.

Jañez Escalada, L. 263.

Jarvelin, K. 262.

Johnston, S. M. 253.

Jüttner, G. 70.

Karlgren, H. 35.

Kotov, R. G. 264.

Kotter, R. 35.

Kuck, C. 21, 70.

Laerte Packer, A. 70.

Laguna Serrano , E. 209.

Lamb, J. 75.

Lancaster, F. W. 93, 99.

Lapesa, R. 35.

Lasso de la Vega, J. 99.

Leloup, C. 70.

Lemaitre, C. 70.

Loeben, M. 127, 264.

Long, B. 86.

Lope de Sosa-Aga, G. 127.

López de Mantaras, R. 127.

Marchuk, Yu N. 75, 262.

Mack, G. 99.

Maier, E. 73.

Málaga Casteleiro, J. 71.

Maldonado Martínez, A. 163.

Manifesto de Madri. 71.

Manu, A. 71, 75.

Marcos Marín, F. 71.

Martín Municio, A. 35.

Martinet, A. 15.

Mataro, F. 73.

Martinho, A. M. 74.

Mayer, R. 71.

Meer, J. de. 73.

Mencioni, G. 50.

Moores, C. N. 103.

NORDTERM. 45, 61.

Nakai, H. 71.

Nebodity, W. 75.

Neveling, U. 36.

Niehoff, R. T. 99.

Nissan, E. 71.

Orna, E. 163.

Ortiz González, C. 209.

Pardo, D. 45.

Paul, J. 71.

Pérez, E. 163.

Pérez Alvarez Ossorio, J. R. 128.

Picht. 21, 50.

Plaz, E. 261.

REUNIBER. 226.

RITERM. 47.

Ranganathan, S. R. 128.

Reichling, A. 75.

Repo, A. J. 263.

Rieger, B. 35.

Rodríguez Jimenes, M. 35, 71, 73.

Rodríguez Ricard, T. 253.

Rolling, Z. 14.

Ros García, J. 99.

Rossilon, P. 21, 35.

Rozo Rios, D. S. 253.

Ruiz, C. 85.

Runge, W. 263.

Sadler, V. 71.

Salra, X. 253.

Sanllorente, A.M. 35.

Sasaki, H. 36.

Saussure, F. 17.

Schwartz, C. 71.

Seegmüller, G. 70.

Shuji Czehi. 25.

Sidorchenko, V. D. 207.

Sievert, M. E. 127.

Skorokhod'ko, E. F. 73.

Snow, B. 253.

Soares de Oliveira, R. M. 35, 73, 263.

Solanellas Torrents, J. M. 72.

Soler, J. 75.

Sopeña, L. 73, 263.

Sorli Rojo, A. 99.

Stognnü, A. A. 73.

Strehlow, R. A. 74.

TERMCAT. 47, 48.

TermEsp. 47, 48.

TermNet. 45.

Taube, M. 91.

Turuguet Mayol, D. 253.

UNE CT-50 Documentação. 68.

UNISIST. 45, 206, 209.

Unguary, R. 36.

Unión Latina. 35.

Utne, I. 50.

Valle Bracero, A. 209.

Van Slipe, G. 209.

Velasco, G. 71.

Verdejo, M. F. 71, 72, 263.

Vian Ortuño, A.

Vickery, B. C. 164.

Vidal Beneyto, J. 71, 72.

Vidal Castro Melo, M. 50.

Vidal de Canelles.

Viet, J. 209.

Villagra Rubio, A. 99.

Vries Robbe, P. F. de. 72.

Weinberg, H. B. 72.

Weisenberg, I. K. 43.

Wersig, G. 36, 93.

Wright, L. D. 72, 263.

Wüster, E. 21, 43, 70.

Yañez, A. 207.

Young, H. 36.

Yu, Y. 73.

Zampolli, A. 72.

Zarco, J. 61.

Zolotarev, A. E. 207.

ÍNDICE DE ASSUNTO

Adjetivos substantivados, 63.

Apresentação em forma flechada, 173.

Apresentação em forma poligonal, 173.

Apresentação formando uma espiral, 173.

Apresentação gráfica arborescente, 173.

Apresentação gráfica quadriculada, 173.

Apresentação gráfica em famílias interligadas, 173.

Bancos de dados de traduções, 55, 56.

Bancos de dados terminológicos, 55, 56.

Bancos de dados terminológicos bilíngües, 56.

Bancos de dados terminológicos informatizados, 56.

Bancos de dados terminológicos manuais, 56.

Bancos de dados terminológicos monolíngües, 57.

Bancos de dados terminológicos multidisciplinares, 56, 57.

Bancos de dados terminológicos multilíngües, 56, 57.

Bancos de dados terminológicos unidisciplinares, 56, 57.

Bancos de dados terminológicos universais, 56, 57.

Bancos terminológicos, 56, 57.

Bibliografia, 227.

Borader Term, 109.

Broader term (generic), 109.

Broader term (patititve), 109.

CD-ROM, 69, 70.

CD-WORM, 70.

CDU - Classificação Decimal Universal, 81.

Catálogo de Recursos Terminológicos em Língua Espanhola, 262.

Centro de Investigações para as Indústrias da Linguagem (CIPIL), 53.

Centro Internacional de Traduções, 63.

Ciência da Terminologia, 54.

Ciências da Documentação, 25.

Classes de tesauros, 93, 95, 96, 97.

Classificação de documentos.

Comissariado Geral da Língua Francesa, 47.

Comparação entre tesauros, 221.

Comunicação, 23.

Conceito, 25, 26.

Conceitologia, 44, 45.

Construção de bancos de dados terminológicos, 58.

Construção de tesauros, 85.

Corte à direita, 238.

Corte à esquerda, 238.

Corte bilateral, 238.

Corte de termos, 238.

Dados factuais, 125.

Datas, 127.

Definição de conceito, 25.

Definição de linguagem, 18.

Definições de terminologia, 18, 22.

Definições de tesauros, 84, 85, 86.

Descritor(es), 92.

Desdobramento de tesauro, 71.

Dicionário, 56, 70.

Discos Compactos (CD - Compact Discs), 54.

Documentação, 11.

EURODICAUTOM, 61.

EUROTRA, 62.

Engenharia da língua, 235.

Escolha do termo de indexação, 235.

Estratégia de indexação, 235.

Estudo do usuário, 247.

Expressões adjetivas, 118.

Expressões adverbiais, 118, 119.

Expressões preposicionais, 119.

Expressões mistas, 119.

Expressões nominais, 119.

Formas de quase sinonímia, 146.

Formas de sinonímia, 144.

Divisão de termo(s), 141.

Divisão semântica, 141.

Divisão sintática, 141.

Frases preposionais, 113.

Frases mistas, 113.

GOSSTANDART, 61.

Homófono(s), 107.

Homógrafo(s), 107.

Homônimo(s), 107.

INFOTERRA, 61.

INSPEC, 61.

Identificador(es), 102.

Indicadores classificatórios, 102.

Indicadores de união, 102.

Indicadores de função, 102, 243.

Indicadores de posição, 243.

Indicadores de ponderação, 241.

Indexação, 92.

Indexação pré-coordenada, 92.

Indexação coordenada, 92.

Indexar, 92.

Indústria(s) da Língua, 51.

Indústria(s) da Linguagem, 51.

Informação, 227.

Informação sobre tesauros, 227.

Informática lingüística, 51.

Information retrieval, 91.

Integração de tesauros, 211.

Jornals in Translation, 63.

LOTOS, 48, 49, 54.

Língua, 17,

Linguagem(s), 17,

Linguagem(s) documentária(s), 80, 81.

Linguagem(s) controlada(s), 80, 81.

Linguagem(s) pré-coordenadas(s), 80, 81.

Linguagem(s) pós-controladas(s), 80, 81.

Linguagem(s) classificatória(s), 81.

Linguagem(s) de indexação, 81.

Linguagem(s) de recuperação da informação, 80.

Linguagem(s) documentária(s).

Linguagem normalizada, 80, 81.

Linguagem(s) terminológica(s), 80.

Lexicografia computacional, 51.

Livro, 17.

Lingüística, 17.

Links, 241.

Lista de cabeçalhos de assuntos, 101.

Lista de termos, 101.

MEDES, 68, 69.

MICROCAT, 62.

Macrotesauro(s), 97.

Manutenção de tesauros, 218.

Manual de Terminologia, 45.

Microtesauro(s), 97.

NORDTERM, 61.

Narrower term, 109.

Narrower term (generic), 109.

Narrower term (partitive), 109.

Norma UNE 50-106, 88.

Normalização, 65.

Normalização na indústria da língua, 65.

Normalização em terminologia, 65.

Nota de esclarecimento, 108, 109.

Notas de aplicação, 108, 109.

Números cardinais, 125.

Números ordinais, 125.

Observatório Francês da Indústria da Linguagem, 53.

Obras de referência, 225.

PRIMUS, 48, 54.

Programa da Língua da Comunidade Européia, 48, 53.

Palavra, 28.

Palavra(s) chave, 91.

Pacotes de informática, 227.

Pertinência, 91, 92.

Polígrafo(s), 107.

Publicação de tesauros, 227.

Quantum de informação útil, 28.

Redação de tesauro(s), 221.

Relação(ões) associativa(s), 155.

Relação(ões) de equivalência, 155.

Relação(ões) hierárquica(s), 155.

Relação(ões) hierárquica(s) enumerativa(s), 154.

Relação(ões) hierárquica(s) genérica(s), 156.

Relação(ões) hierárquica(s) partitiva(s), 155.

Relação(ões) poli-hierárquica(s), 155.

Relação(ões) semântica(s), 157.

Related term, 109, 157.

Relevância, 91, 92.

Recuperação da informação, 92, 93.

Revisão de tesauro(s), 218.

Roles, 243.

SYSTRAM, 62.

Scope note, 156, 157.

Semaforonte, 26.

Sistema Uniterm, 91, 92.

Sistema Zator, 91,92.

Sistema(s) clasificatório(s), 78.

Sistemas de classificação, 77.

Sistemas especialistas, 54.

Sistemas especialistas para tratamento de textos, 261.

Substantivos, 108, 114.

Substantivos adjetivados, 113.

Substantivos nominais, 113.

TEAM-SIEMENS, 61.

TermEsp, 63.

Terminómetro, (.T. - Nome de periódico publicado pela Unión Latina), 46.

TERMIUM, 61.

TITUS, 62.

Técnicas de distanciamento, 240.

Técnicas de indexação, 240.

Teoria de tesauros, 14.

Teoria de Sistemas, 15.

Teoria Geral da Terminologia, 42.

Termo, 140.

Termo(s) de busca de informação, 248.

Termo(s) principal(is), 106.

Termo(s) composto(s), 140.

Termo(s) quase sinônimo(s), 106.

Termo(s) de indexação, 237.

Termo(s) de recuperação da informação, 249.

Termo(s) equivalente(s), 108.

Termos(s) genérico(s)-específico(s), 108.

Termo(s) específico(s)-específico(s), 108.

Termo(s) específico(s)-genérico(s), 108.

Termo(s) genérico(s), 108.

Termo(s) genérico(s)-genérico(s), 108.

Termo(s) não-descritor(es), 113.

Termo(s) não-preferenciais, 108, 109.

Termo(s) partitivo(s), 108, 109.

Termo(s) preferenciais, 113.

Termo(s) relacionado(s), 113.

Termo(s) secundário(s), 113.

Termo(s) simples, 113.

Termo(s) sinônimo(s), 113.

Terminografia, 30.

Terminologia, 15, 18, 21.

Terminologia moderna, 38.

Terminólogo, 30, 31.

Terminology Manual, 86.

Thesaurus, 93, 94.

Thesaurus Construction, 114.

Tesaurisar.

Tesauro(s), 78.

Tesauro(s) alfabético(s), 96, 97.

Tesauro(s) auxiliar(es), 96, 97.

Tesauro(s) bilíngüe(s), 96, 189.

Tesauro(s) de recuperação da informação, 132.

Tesauro(s) de referência, 245.

Tesauro(s) gráfico(s), 96, 97.

Tesauro(s) de indexação, 132.

Tesauro(s) especializado(s), 96, 97.

Tesauro(s) facetado(s), 96, 97.

Tesauro(s) geral(is), 96, 97.

Tesauro(s) hierárquico(s), 96, 97.

Tesauro(s) marginal(is), 96, 97.

Tesauro(s) monodisciplinar(s), 96, 97.

Tesauro(s) monolíngüe(s), 96, 187.

Tesauro(s) multidisciplinar(es), 96, 186.

Tesauro(s) multilingüe(s), 96, 187, 195.

Tesauro(s) plurilíngüe(s), 96, 186, 189.

Tesauro(s) temático(s), 96, 187.

Tesauro(s) temático(s), 96, 187.

Tesauro(s) sistemático(s), 96, 167, 186.

Top Term, 111.

Tradução automática, 61, 62.

Unidade de informação, 21.

Unidade(s) lingüística(s), 105.

Use, 109.

Use for, 109.

Uso de tesauros, 187.

Verbo(s) substantivado(s), 118.

Vocabulário, 56.

World Transindex, 63.

Zatocoding, 92.