

Grupo Temma

BIBLIOTECA
DO
IBICT



Análise documentária:

a análise da síntese

4.05

BRASILEIRO
MACÃO
IA E TECNOLOGIA

Presidente da República

José Sarney

Ministro da Ciência e Tecnologia

Renato Archer

Secretário-geral

Luciano Coutinho

Presidente do Conselho Nacional de Desenvolvimento Científico e
Tecnológico (CNPq)

Crodowaldo Pavan

Vice-presidente

José Duarte de Araújo

Diretor do Instituto Brasileiro de Informação em Ciência e
Tecnologia (IBICT)

Antônio Agenor Briquet de Lemos

Vice-diretora

Vania Maria Rodrigues Hermes de Araújo

Grupo Temma:

Johanna W. Smit (coordenação)
Anna Maria Marques Cintra
Eunides A. do Vale
Isabel M.R. Ferin Cunha
Maria de Fátima G.M. Tálamo
Nair Yumiko Kobashi
Regina Keiko Obata F. Amaro



INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA	
DO DEPARTAMENTO DE DOCUMENTAÇÃO	
IBICT	
DATA:	1987
VALOR:	1.000,00
TIPO DE AQUISIÇÃO:	77

Análise documentária:

a análise da síntese

Brasília
1987

MCT
Ministério da Ciência e Tecnologia

 **CNPq**
CONSELHO NACIONAL DE DESENVOLVIMENTO
CIENTÍFICO E TECNOLÓGICO

IBICT
INSTITUTO BRASILEIRO
DE INFORMAÇÃO
EM CIÊNCIA E TECNOLOGIA

025.4.05
A 532
EX. 4
EDU

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA	
DDI BIBLIOTECA	
F. P. ELOR:	
FORMA AQUISIÇÃO:	VALOR:
<i>Adquirido</i>	<i>1.00</i>
TOMBO:	DATA:
<i>21067</i>	<i>15-09-97</i>

PHK 003949

Análise documentária : a análise da síntese / coordenação de Johanna W. Smit. Brasília : IBICT, 1987.

P.

I. Cunha, Anna Maria Marques. II. Vale, Eunides A. do III. Cunha, Isabel M.R. Ferin. IV. Tálamo, Maria de Fátima G.M. V. Kobashi, Nair Yumiko. IV. Amaro, Regina Keiko Obata F. 1. Análise de conteúdo - coletânea.

CDU 002

Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT)

SCN - Quadra 2 - Bloco K
70710 Brasília, DF

Tel. (061) 225 9677 Telex 612481 CICT BR

S U M Á R I O

INTRODUÇÃO	5
<i>Johanna W. Smit</i>	
1. LINGUAGENS DE INDEXAÇÃO	11
<i>Eunides A. do Vale</i>	
2. ESTRATÉGIAS DE LEITURA EM DOCUMENTAÇÃO	27
<i>Anna Maria Marques Cintra</i>	
3. ANÁLISE DOCUMENTÁRIA	37
<i>Isabel M. R. Ferin Cunha</i>	
4. ESTRUTURAÇÃO DE VOCABULÁRIO	61
<i>Isabel M, R. Ferin Cunha</i>	
5. A DEFINIÇÃO SEMÂNTICA PARA A ELABORAÇÃO DE GLOSSÁRIOS	87
<i>Maria de Fátima G. M. Tálamo</i>	
6. A ANÁLISE DA IMAGEM: UM PRIMEIRO PLANO	99
<i>Johanna W. Smit</i>	
7. REVISÃO BIBLIOGRÁFICA	113
<i>Isabel M. R. Ferin Cunha</i>	
<i>Nair Yumiko Kobashi</i>	
<i>Regina Keiko Obata F. Amaro</i>	

INTRODUÇÃO

O presente volume reúne as discussões que, com o passar do tempo, foram dando origem ao grupo TEMMA. No início, eram discussões para melhor integrar disciplinas a serem oferecidas no curso de graduação em biblioteconomia e documentação da Escola de Comunicações e Artes da USP. Depois, as preocupações foram se tornando mais específicas, uma vez que, por um lado, havia uma forte demanda por cursos extra-curriculares na área da "análise e representação temática", as quais se tentou responder organizando cursos de extensão; e por outro lado a linha de pesquisa "análise documentária", a nível de pós-graduação, na medida em que se estruturava, reunia pesquisadores com preocupações convergentes.

A discussão, aliada às respostas por ela recebidas, quer seja nos cursos oferecidos (graduação, pós-graduação e extensão), ou na avaliação da pesquisa acadêmica de mestrado ou doutorado, foi responsável por uma dinâmica que se instaurou e que obrigou cada um a buscar cada vez mais longe seu horizonte. Os problemas que ontem nos preocupavam estão, em parte, superados, o que somente significa que estamos em melhores condições de avaliar o caminho que resta a percorrer.

No início da década de 70 uma vaga intuição direcionou as pesquisas para um campo que, pelo menos na época, não integrava o currículo habitual da biblioteconomia no Brasil: a interseção entre a documentação e a lingüística. Lembro-me de uma conferência que James Perry fez em 1970: seria incapaz de relembrar o assunto geral da conferência; a única imagem que ficou gravada na minha memória dizia respeito aos anéis de benzeno que o homem desenhou na lousa, discorrendo sobre os diferentes radicais e as relações que se estabeleciam entre o anel e os radicais. Naquele momento, um pararelo se estabeleceu entre o anel de benzeno e o "miolo" das palavras, modificáveis pelos diferentes radicais. Na época, tive a impressão de que, se conseguisse

isolar melhor os "miolos", a documentação trabalharia com maior economia e eficácia, preocupando-se unicamente com o "cerne" da questão.

A busca pelo "cerne" foi longa, mas, hoje, podemos dizer que o grupo TEMMA está se aproximando do objetivo. Esta busca levou o grupo, quase que imediatamente, à lingüística e, conseqüentemente, a um titubear entre as diferentes escolas que se apresentavam. A busca pela escola lingüística que pudesse responder às nossas preocupações de documentalistas fez-nos, por momentos, trilhar becos sem saída; perdemos tempo em buscas infrutíferas, mas agora, pouco a pouco, a poeira está assentando e a inter-relação documentação/lingüística passou a constituir um questionamento razoavelmente controlado.

O bibliotecário leigo ou semi-leigo em lingüística tem, ao se debruçar sobre o assunto, uma sensação inicial de estar presenciando um tiroteio no qual ele não consegue distinguir o mocinho do bandido. O mocinho, forçosamente, seria aquela tendência lingüística que pudesse ser aplicada à documentação. A identificação do mocinho tornou-se possível a partir do momento em que ficou claro que não havia um único mocinho, mas vários. O tiroteio transformou-se numa sessão de pescaria, em que operações e conceitos relevantes para a documentação eram selecionados de várias escolas, tendo-se o cuidado de não justapor conceitos contraditórios. A somatória de conceitos e operações compatíveis reverte numa opção pragmática que, se por um lado resolve uma série de problemas, cria, sem dúvida alguma, inúmeros outros. Hoje, estamos em condições de dizer que, no caso, não houve a eleição de uma escola e o descarte das outras: houve a eleição de operações relevantes para a documentação, guiando-se a pesquisa pela aplicabilidade a questões de documentação e não pelo maior ou menor acerto atribuído a esta ou aquela escola.

Assim, pouco a pouco, chegou-se à gramática dos casos e à

lingüística e semântica dos textos, ultrapassando-se a barreira da frase (limitante essencial em boa parte das escolas lingüísticas).

Resumidamente, pode-se dizer que as preocupações acerca da explicitação do processo de análise documentária (e as questões lingüísticas inevitavelmente envolvidas pela discussão) datam da década de 70. A automação desnudou a relativa ingenuidade das operações de análise documentária: para responder e estar à altura dos acontecimentos, era necessário estabelecer regras claras, dar opções unívocas à máquina, explicitar em suma. A necessidade oficializou um casamento que já existia, mas que não se assumia como tal: análise documentária e lingüística passaram a ser discutidas em conjunto, uns abordando a lingüística como "fornecedora" de insumos básicos, outros, mais ousados mas igualmente corretos, oferecendo a experimentação de procedimentos de análise documentária em larga escala (e automatizada) como possibilidade de validação de teorias lingüísticas.

Desde 1957, o entrelaçamento entre a documentação e a lingüística existe. O idealizador dos índices KWIC e KWOC, Luhn, fundamentou-se na lingüística distribucionalista de Harris e desde esta época a discussão acerca das correntes lingüísticas permeia os "progressos" da análise documentária (o inverno não é verdadeiro, cumpre acrescentar).

As discussões suscitadas pela automação da indexação, análise e classificação, com o passar dos anos, revelaram que a problemática da análise documentária não se resumia a questões de sintaxe e semântica, uma vez que, ao ultrapassar o limite da frase e passando a se preocupar com o texto, um outro componente intervinha igualmente: a lógica.

A tradução da informação contida nos documentos em informação documentária supõe a criação de instrumentos para normalizar e

homogeneizar esta tradução: ambos estão hoje irremediavelmente associados a operações lógico-semânticas. Esta ampliação do universo frásico inicial levou, pouco a pouco, a uma modificação na abordagem da problemática, na medida em que a lógica interna ao texto é considerada igualmente relevante para sua análise. Neste movimento, ultrapassa-se várias barreiras que podem ser enumeradas como segue:

a. Para fins de análise documentária, a frase não é forçosamente uma unidade relevante de informação, uma vez que a informação relevante pode estar numa palavra, frase, parágrafo ou texto. A unidade de análise pré-determinada deixa de ser um conceito pertinente;

b. quando a unidade de análise é o próprio texto, para gerenciá-lo de forma eficaz é necessário dominar, entre outras, sua estruturação interna (ver Capítulo 2), o que leva a uma valorização das relações lógico-semânticas presentes no mesmo;

c. na hipótese precedente, o texto (inclusive o texto científico) deixa de ser um espaço neutro, composto de uma única camada unívoca, passando a ser entendido como uma sobreposição de várias camadas, algumas internas ao texto, outras externas. Surge a discussão da pragmática: as condições de produção do texto, bem como suas condições de consumo. O texto deixa de ser um objeto neutro, isolado, fechado, e passa a ser um espaço de circulação de informações, ou seja: as informações contidas no próprio texto acrescidas das "leituras" que delas são feitas no momento da produção, análise documentária ou consumo.

A ideologia irrompe (ver Capítulo 3), desfazendo alguns tabus perseverantes no meio documentário e que se referem ao "fazer" da análise: o "texto neutro", com sua "leitura única e absoluta", objetivando a seleção de "conceitos neutros" etc.

Inevitavelmente, o espaço atribuído ao texto devendo ser cada vez maior, como vimos, aumentou consequentemente a complexidade e diversidade de variáveis em presença no texto a ser analisado. A busca de uma solução para este novo desafio está levando a documentação a se interessar pela inteligência artificial, uma vez que esta se propõe a entender os procedimentos lógicos e lingüísticos envolvidos pela produção e o consumo da informação. Assim, a análise documentária passa a ser uma modalidade de leitura de texto que leve em conta, além do próprio texto, as condições de produção e consumo.

As propostas da inteligência artificial permeiam este volume, com uma certa timidez ainda, mas o grupo TEMMA pretende continuar pesquisando nesta direção, certo de que soluções econômicas e procedimentos inteligentes para a documentação brasileira poderão ser alcançados.

O grupo TEMMA é composto por sete pesquisadores (bibliotecários e lingüistas), na maior parte pertencentes ao corpo docente do Departamento de Biblioteconomia e Documentação da Escola de Comunicações e Artes da USP.

Eunides Aparecida do Vale (bibliotecária da Faculdade de Filosofia, Letras e Ciências Humanas da USP e ex-docente da ECA), no Capítulo 1, sistematizou uma panorâmica geral dos instrumentos clássicos de análise documentária, sua evolução e suas limitações. Este primeiro capítulo, resumindo o estado atual das discussões mais frequentes sobre o assunto, situa a problemática geral da análise documentária hoje e anuncia os temas das discussões futuras (e dos capítulos subsequentes).

A lingüista Anna Maria Marques Cintra (PUC/SP) vem pesquisando há anos a inter-relação entre documentação e lingüística, e está ministrando um curso de pós-graduação na ECA, como professora convidada, sobre a mesma questão. O Capítulo 2, que discute as "estratégias de leitura", contextualiza a análise documentária

enquanto estratégia e discute as implicações destas estratégias em relação à eficiência da própria análise.

Um aspecto da tese de doutorado de Isabel M. R. Ferin Cunha (ECA) foi resumido no Capítulo 3, no qual ela discute profundamente a questão da ideologia na análise documentária. Este trabalho, extremamente inovador, vem acompanhado de colocações sobre os recortes pragmáticos dos conceitos lingüísticos operacionais para a documentação: a operacionalização do conjunto está descrita no Capítulo 4.

A mesma preocupação com a seleção de conceitos lingüísticos relevantes para a documentação está presente no Capítulo 5, desta vez sob a ótica da lingüista Maria de Fátima G. M. Tálamo (ECA), que complementa a problemática iniciada no Capítulo 4.

Numa tentativa de ampliar a abrangência da discussão da análise documentária para o documento em geral (e não somente o documento escrito), Johanna Smit (ECA) aborda alguns problemas específicos da análise da imagem no Capítulo 6. A análise dos documentos audiovisuais, com objetivos documentários, carece totalmente de bibliografia, razão pela qual o capítulo restringe-se à discussão da problemática em si, abordando as interseções desta problemática com os conceitos lógicos e lingüísticos explanados nos capítulos anteriores de forma altamente elíptica.

Finalmente, a revisão bibliográfica elaborada por Nair Yumiko Kobashi e Regina K. Obata F. Amaro (ambas da ECA) detalha as origens dos conceitos mais recorrentes na atual discussão, propiciando uma visão histórica do encadeamento de linhas de pensamento e contextualizando conceituações retomadas nos diferentes capítulos de forma mais isolada.

Johanna W. Smit

São Paulo, junho de 1987

1. LINGUAGENS DE INDEXAÇÃO

Eunides A. do Vale

1.1 LINGUAGENS DE INDEXAÇÃO PRÉ-COORDENADAS

1.1.1 Sistemas tradicionais de classificação

1.1.1.1 CDD/CDU

1.1.1.2 LC

1.1.1.3 Hierarquização

1.1.2 Listas de cabeçalhos de assunto

1.1.2.1 Remissivas (LCSH)

1.1.3 Classificações facetadas

1.1.3.1 Classificações facetadas pós-Ranganathan

1.1.4 Ordem de citação em sistemas pré-coordenados

1.2 LINGUAGENS DE INDEXAÇÃO PÓS-COORDENADAS

1.2.1 Unitermo

1.2.2 Tesouros

1.3 COORDENAÇÃO: LINGUAGENS PRÉ E PÓS COORDENADAS

1.3.1 Pré-coordenação

1.3.1.1 Ordem de citação

1.3.2 Pós-coordenação

1.3.2.1 Artíficos de precisão utilizados para evitar associações falsas

BIBLIOGRAFIA

1. LINGUAGENS DE INDEXAÇÃO

Eunides A. do Vale

A escolha de uma linguagem de indexação é fator essencial para a eficácia de um sistema de recuperação da informação. Deve-se levar em consideração os objetivos do sistema, o tipo de usuário e a abrangência ou especificidade do assunto a ser tratado.

As linguagens se dividem em dois grupos:

- Pré-coordenadas, que combinam ou coordenam os termos no momento da indexação. São utilizadas principalmente em sistemas manuais: catálogos de bibliotecas, bibliografias e índices impressos etc.
- Pós-coordenadas, que combinam ou coordenam os termos no momento da busca. São utilizadas principalmente em sistemas automatizados: bases/bancos de dados, bibliotecas ou centros de documentação que operam com computador.

1.1. LINGUAGENS DE INDEXAÇÃO PRÉ-COORDENADAS

1.1.1. Sistemas tradicionais de classificação

São as mais antigas linguagens documentárias e abrangem todas as áreas do conhecimento. Atualmente, têm como objetivo principal a arrumação dos livros nas estantes por assunto. Representam a indexação através de uma notação, basicamente composta por números, letras ou a mistura de ambos.

Ex.: 324.63 = Sistemas eleitorais

Os sistemas mais utilizados no Brasil são:

- Classificação Decimal de Dewey (CDD);
- Classificação Decimal Universal (CDU);
- Classificação da Library of Congress (LC).

Esses sistemas são estruturados hierarquicamente, dividindo o conhecimento em dez classes principais, subdivididas sucessivamente.

É através dessas subdivisões que muitas vezes se forja a inserção de novos assuntos, os quais em outras estruturas teriam um tipo de relação não-hierárquica.

No exemplo abaixo, retirado da CDD, percebe-se a relação hierárquica existente entre a Sociologia e a Antropologia Cultural e Social.

Ex.: CDD:	18ª ed.	301	Sociology
		301.2	Culture and cultural processes
			Class here primitive races (formerly 57217), cultural (formerly 390) and social anthropology, comprehensive works on anthropology.
	19ª ed.	301/307	Sociology
		306	Culture and institutions
			Class here cultural and social anthropology.

Outro aspecto a ser considerado refere-se à diferença de conceituação de determinados assuntos, como pode ser observado no exemplo a seguir:

Ex.: CDD:	19ª ed.	572	Human races
			Class here Ethnology
	19ª ed.	306	Culture and institutions

Class here cultural and social anthropology.

Tesauro de Sociologia (VIET, 1971):

Social and cultural Anthropology/Ethnology

UF Ethnology

Enquanto a CDD atribui conceitos diferentes e os classifica como assuntos diversos, em classes completamente diferentes, o Tesauro de Sociologia (VIET, 1971) considera-os como termos equivalentes em língua inglesa e francesa. O *Dicionário Aurélio* define Etnologia como: 1. Ramo de Antropologia que estuda a cultura dos chamados povos naturais. 2. Estudo e conhecimento, sob o aspecto cultural, das populações primitivas. 3. Estudo comparativo de todos os povos. (Sinônimo geral: Antropologia cultural).

Referindo-se especialmente à CDU, lembramos que ela apresenta dispositivos de síntese, que dão maior flexibilidade ao sistema, embora, em muitos casos, isso leve à formação de notações muito grandes, dificultando sua leitura e compreensão, principalmente por parte dos usuários. A CDU é também utilizada para a elaboração do catálogo sistemático.

1.1.1.2 Library of Congress (LC)

Foi elaborada para ser aplicada especialmente ao acervo da Library of Congress dos EUA. Também está estruturada hierarquicamente e suas classes são criadas de acordo com o perfil e desenvolvimento do acervo. É suplementada por lista de cabeçalho de assunto.

1.1.1.3 Hierarquização

A interdisciplinaridade das ciências mostra que a hierarquização do conhecimento não é suficiente para evidenciar os vários tipos

de relações possíveis entre assuntos, num sistema de recuperação de informação; daí a necessidade de maior reflexão sobre as linguagens documentárias que utilizam esse tipo de estruturação. O exemplo abaixo mostra o assunto **Eleição** estruturado por um sistema de classificação tradicional e por um tesouro.

Ex.: CDD	- 19ª ed.	Tesouro de Sociologia (VIET,
32	Political Science	1971)
324	The Political process	
324.6	Elections	Elections/Elections
324.63	Electoral systems	RT Electoral system/Système électoral
	Relação hierárquica	Relação Associativa

1.1.2 Listas de Cabeçalhos de Assunto

Representam os assuntos sob a forma de cabeçalhos já estruturados. Geralmente são listas gerais que arrolam termos de todas as áreas do conhecimento. No Brasil, as mais conhecidas e usadas são:

Library of Congress of Subject Headings (LCSH) e Sears List of Subject Headings (Sears), sendo a primeira de maior porte que a segunda. Essas listas apresentam-se alfabeticamente, utilizando alguns sinais como o traço, a vírgula e o parêntese para estruturar os cabeçalhos indiretos.

No exemplo a seguir, retirado da LCSH e citado por Foskett (1973), observa-se o uso do singular e plural e o uso do cabeçalho direto e indireto, o que interfere na recuperação de assuntos similares.

Ex.:	Libraries, naval
porém	Libraries, military see military libraries
	Library administration (não Libraries-Administration)
porém	Acquisitions (Libraries)
	Cataloging of moving-pictures
porém	Classification - Moving-pictures

Contudo, o singular e o plural tornam-se importantes quando diferenciam o sentido do cabeçalho.

Ex.: Alemão - língua
 Alemães - povo

As listas tradicionais também apresentam sinais de relação em função dos cabeçalhos, como por exemplo:

X (see) cabeçalhos sinônimos

XX cabeçalhos mais amplos ou genéricos e cabeçalhos relacionados

see also - cabeçalhos específicos e relacionados

em negrito - cabeçalhos eleitos para entrada

Com a utilização apenas desses sinais, as listas não fazem uma distinção clara entre os vários tipos de relações, como pode-se observar no exemplo a seguir, em que o mesmo assunto "Plantas" aparece estruturado por uma lista de cabeçalho de assunto e um tesouro. Na lista de cabeçalho de assunto os termos específicos e relacionados aparecem num mesmo grupo, precedido pelo termo see also (relação de associação "livre"), enquanto que no tesouro cada termo aparece especificado na sua relação com os demais (relação de associação estruturada).

Ex.: (Sears, 11ª ed.)	Transformação para uma estrutura de tesouro
Plants	Plants
see also	TE alpine plants
Fertilization of Plants	TR Fertilization of plants
Alpine Plants, Flowers,	TR Flowers
Gardening	
X Flora	TR Gardening
XX Botany	UP Flora
	TG Botany

1.1.2.1 Remissivas (LCSH)

Foskett (1973) cita um estudo de J. Daily sobre a LCSH, em que foi detectado que, de 36.468 cabeçalhos principais, cerca de 7.033 não possuem remissivas, o que demonstra que só poderiam ser localizados pelas entradas previstas na lista, exigindo, portanto, um profundo conhecimento do instrumento pelo indexador.

1.1.3 Classificações Facetadas

Ranganathan é considerado o criador das classificações facetadas. Denominou sua classificação de Colon Classification, a qual ficou conhecida como Classificação dos Dois Pontos.

Revolucionou a estrutura dos sistemas tradicionais de classificação, introduzindo o princípio da divisão de assuntos em categorias ou facetas.

Suas categorias ou facetas principais são conhecidas pela sigla PMEST (FOSKETT, 1973):

ST - Espaço/Tempo

E - Energia (como força compulsora), ex.: Exportações em economia, currículo em educação.

M - Matéria, ex.: Ouro como material de dinheiro dentro da Economia. Instrumento em música. Marfim na Escultura.

P - Personalidade - corresponde às coisas, tipos de coisas, tipos de ação. Ex.: Pessoas em Sociologia/Psicologia.

A classificação de Ranganathan em si não tem aplicação no mundo ocidental. É uma classificação geral, com 42 classes principais, sendo aplicada, a cada uma dessas classes, a análise em facetas.

1.1.3.1 Classificação facetada pós-Ranganathan

Desde 1952, na Inglaterra, o Classification Research Group (CRG)

vem se dedicando aos estudos das classificações facetadas (Foskett, Shera, Vickery, etc.). No Brasil, os primeiros estudos datam de 1970 - curso de pós-graduação do IBICT (BARBOSA, 1972).

A classificação facetada define-se como a aplicação com rigor da análise em facetas à sua estruturação. Através da síntese de mais de uma faceta representa-se um assunto composto.

Ex.: Vegetação própria de terra roxa (VICKERY, 1980):

9ir/7n

tipo de solo 9

terra roxa ir

relação entre as facetas

constituintes 7

vegetação n

O quadro apresentado mostra parcialmente as facetas e suas subdivisões aplicadas à Ciência do Solo, que deram origem ao exemplo (VICKERY, 1980):

Facetas

9 tipos de solo

8 estrutura

7 constituintes

6 propriedades

5 processos no solo

4 operações

3 técnicas de laboratório

1 geral

Desdobramento das facetas

9 tipos de solo

b orto-elúvio

:

:

i subtropical e tropical

ir terra roxa

:

:

8 estrutura

b simples-grão

:

:

7 constituintes

b químicos

c inorgânicos

:

:

n vegetação

:

6 propriedades

As classificações facetadas normalmente se aplicam a uma área do conhecimento, tendo, portanto, um caráter de especialização. Essa especialização ou assunto núcleo é considerada o Universo ou campo que se quer classificar, por ex.: Sociologia, Petróleo, Transporte etc.

Faceta é a diferença ou característica de uma parte do conhecimento analisado (CURSO..., 1978).

Subfaceta é a divisão de uma faceta pelas diferenças apresentadas (CURSO..., 1978).

Além de Ranganathan, outros estudiosos como, por exemplo, Vickery, estabeleceram novas categorias para divisão do Universo a ser classificado. Acredita-se que, ao elaborar uma classificação facetada, primeiro deve-se verificar se as categorias já existentes podem ser aplicadas ou se é necessária uma adaptação ou até mesmo a elaboração de novas categorias.

Segue abaixo um exemplo de uma categorização "livre" aplicada ao campo: Alimentação de grupos indígenas, baseado no texto de Barbosa (1972).

Ex.:	Facetas	Subfacetas
	obtenção	caça pesca coleta roça
	preparo	cru cozido assado fermentado
	consistência	sólido líquido

Facetas

Subfacetas

ocasião de consumo

cotidiano

festa

funeral

propriedade nutriente

vitaminas

A

B

C

proteínas

sais minerais

As classificações facetadas serviram de base para elaboração de tesouros.

1.1.4 Ordem de citação em sistemas pré-coordenados

Normalmente, cada sistema sugere uma ordem de citação. Essas ordens não devem ser encaradas como inflexíveis, uma vez que podem ser modificadas de acordo com as necessidades e objetivos do sistema de informação.

Ex.: PMEST - ordem de citação de Ranganathan;
lugar, tempo, forma, língua - ordem de citação da CDU.

1.2 LINGUAGENS DE INDEXAÇÃO PÓS-COORDENADAS

1.2.1 Unitermo

Criado e empregado por Mortimer Taube em 1953. Pode ser definido como a indexação por palavras únicas retiradas do contexto, na maioria das vezes, sem controle de vocabulário e destinado aos sistemas pós-coordenados.

Utiliza-se das chamadas Fichas Unitermo, como suporte para sua aplicação. Essas fichas são divididas em dez colunas numeradas de 0 a 9, deixando no alto um espaço para a palavra ou termo de

entrada.

Os documentos são numerados seqüencialmente e o número individual de cada um é anotado em todas as Fichas Unitermo que correspondem aos conceitos escolhidos pelo indexador como entrada. Por exemplo, o nº 298 de um documento sobre ensino de Biblioteconomia, é anotado nas fichas: Ensino coluna 8 e Biblioteconomia também na coluna 8. Nota-se que há uma correspondência entre o dígito final do número do documento com o número da coluna da Ficha Unitermo.

No momento da busca, comparam-se as fichas referentes aos conceitos que se está procurando; por exemplo: para encontrar documentos que tratam do Ensino de Biblioteconomia, conferem-se as fichas sobre Ensino e sobre Biblioteconomia, as de números comuns certamente corresponderão à pesquisa em questão.

É um método manual, sendo posteriormente também empregado o uso de fichas de coincidência óptica.

Para um sistema pequeno, de fácil manipulação, pode-se considerar um método satisfatório, mas à medida que o acervo cresce e conseqüentemente cresce também o número de fichas de assunto, torna-se difícil o seu manuseio. Muitos sistemas acabam se transformando em sistemas pré-coordenados ou então passam a se utilizar de computador.

1.2.2 Tesouros

Aplicados preferencialmente aos sistemas automatizados, são usados, por vezes, como base para indexação pré-coordenada em sistemas manuais, desvirtuando-se de seu objetivo principal.

Estrutura:

Sinais relacionados com a função do Descritor:



TG (BT) - termo geral
TE (NT) - termo específico
TR (RT) - termo relacionado
NA (SN) - nota de alcance
TGM (TT) - termo genérico maior
UP (UF) - usado para
USE (USE) - use

Tipos de relações (AITCHISON & GILCHRIST, 1979):

Relações de equivalência, incluindo sinônimos verdadeiros e quase sinônimos.

Ex.: Cosmonauta
USE Astronauta

Relações hierárquicas incluindo gênero/espécie e coisa/tipo:

Ex.: Árvore frutífera
TE macieira
TG Árvore
Máquina
TE Máquina de escrever

A relação parte/todo é, em alguns casos, considerada como hierárquica.

Ex.: Brasil
TE São Paulo
TG América do Sul

Poli-hierarquias - termo que pertence a mais de uma hierarquia:

Ex.: Repressão sexual Repressão política
TG Repressão TG Repressão

Relações Associativas ou não-hierárquicas, incluindo: coisa/parte; coisa/propriedade; coisa/processo etc.

Ex.: Fichário
TE Gavetas

A apresentação dos tesouros pode ser alfabética ou sistemática, suplementada por um índice alfabético. Alguns incluem também a classificação facetada.

1.3 COORDENAÇÃO: LINGUAGENS PRÉ E PÓS-COORDENADAS

A coordenação dos termos é o que permite a precisão na recuperação da informação.

1.3.1 Pré-coordenação

O primeiro termo é o que determina a recuperação, o que significa que, na pré-coordenação, é necessário o uso de muitas remissivas ou entradas múltiplas para explicitar todos os conceitos significativos, o que causa um aumento de custo para o sistema na fase de entrada. Podemos observar o número de entradas no exemplo abaixo:

Ex.: Entradas múltiplas

301.32:33 - Sociologia: C.Política: Economia

32:33:301 - C.Política: Economia: Sociologia

33:301:32 - Economia: Sociologia: C.Política

Remissivas

Demarcação de terras - Índios - Goiás

X Terras - Demarcação - Índios - Goiás

X Índios - Demarcação de terras - Goiás

X Goiás - Demarcação de terras - Índios

1.3.1.1 Ordem de citação

A ordem de citação pré-fixada, na maioria das vezes, favorece a precisão na recuperação e dispensa o uso de artifícios de ligação

para evitar coordenações falsas. Mesmo assim, é necessário alertar que algumas vezes a prefixação dos termos acaba dispersando elementos de conceitos relacionados.

Ex.: Código penal na Nicarágua, classificado pela CDU, ficaria assim representado:

342.2 (728.5) (094.4)

342.2 - Direito Penal

(728.5) - Nicarágua

(094.4) - Código

O conceito código penal (342.2(094.4)) fica disperso, porque a ordem de citação da CDU recomenda o local antes da forma.

1.3.2 Pós-coordenação

Necessita de um suporte computacional adequado, visto que a saída manual é praticamente inviável em sistemas de médio ou grande porte. Dispensa a ordem de citação e possibilita múltiplas combinações no momento da busca. Opera preferencialmente com conceitos simples, sendo que esses conceitos podem ter uma ou mais palavras.

Ex.: uma só palavra: Tuberculose - para expressar doença no pulmão.

mais de uma palavra: Cientista político (diferente de Político).

Os manuais de construção de tesouros sugerem o uso da pré-coordenação nos casos em que a pós-coordenação pode alterar o sentido do conceito, por exemplo: Vida/padrão pode significar tanto padrão de Vida como Vida Padrão.

O uso de conceitos compostos, próprios de determinadas áreas, colaboram para a especificidade da linguagem, possibilitando

melhor precisão na recuperação.

1.3.2.1 Artificios de precisão utilizados para evitar associações falsas (AITCHISON & GILCHRIST, 1979)

Elos (links): Usados para mostrar quais os termos que devem ser relacionados num mesmo documento.

Ex.: Febre Amarela ataca turistas em uma comunidade de pescadores no Rio Amazonas.

Febre Amerela - 1000 A

Turistas - 1000 A

Comunidade - 1000 B

Pescadores - 1000 B

Rio Amazonas - 1000 A

Os elos de ligação evitam associações falsas, como por exemplo Febre amarela/Pescadores e Pescadores/Rio Amazonas, que não figuram no texto. No entanto, nem sempre os elos conseguem resolver todos os problemas de coordenação, como pode ser observado no exemplo acima em que a combinação correta Comunidade de pescadores (B)/Rio Amazonas (A), fica excluída.

Aitchison & Gilchrist (1979) afirmam que os "elos são valiosos em grandes coleções, mas devem ser usados com cautela".

Indicadores de função: Determinam a função que o termo apresenta no contexto: matéria-prima, produto final, componente etc.

Ex.: Utilização de madeira na construção de habitações populares na Amazônia Habitações populares/
Madeira/Amazônia (1)

Habitações populares em madeira na Amazônia Habitações populares/
Madeira/Amazônia (2)

Indicadores de função:

- (1) matéria-prima
- (2) produto

As linguagens de indexação envolvem problemas gerais e específicos. Os primeiros enunciam-se como: conhecimento da área, interdisciplinaridade e ideologia; os segundos como: elaboração e adaptação - questões que serão abordadas nos capítulos posteriores.

BIBLIOGRAFIA

- AITCHISON, J. & GILCHRIST, A. (1979). *Manual para construção de tesouros*. Rio de Janeiro, BNG/Brasilart.
- BARBOSA, Alice Príncipe (1972). Classificação facetada. *Ciência da Informação*. Rio de Janeiro, 1 (2):73-81.
- CURSO sobre lenguajes de indización de tesouros (1981), por el Grupo de Trabajo sobre Lenguajes de Indización. Paris, UNESCO. (Basado em el Seminario Regional sobre Lenguajes de Indización, Buenos Aires, UNESCO/CAICYT, 1978).
- FOSKETT, A.C. (1973). *A Abordagem temática da informação*. São Paulo, Ed. UnB/Polígono.
- VICKERY, B.C. (1980). *Classificação e indexação nas ciências*. Rio de Janeiro, BNG/Brasilart.
- VIET, Jean (1971). *Thesaurus pour le traitement de l'information en Sociologie*. Paris, Mouton.

2. ESTRATÉGIAS DE LEITURA EM DOCUMENTAÇÃO

Anna Maria Marques Cintra

2. ESTRATÉGIAS DE LEITURA EM DOCUMENTAÇÃO

Anna Maria Marques Cintra

Tomando a leitura como processo interativo leitor/texto, Kato (1985) aponta três fatores básicos que atuam como suporte para a legibilidade: a qualidade do texto, o conhecimento prévio do leitor e o tipo de estratégias que o texto exige.

Para a qualidade do texto concorrem vários fatores tais como: a manutenção do tema, a correção gramatical, a adequação lexical, a estruturação do texto.

Na sua prática profissional o bibliotecário depara-se com textos de diversos níveis. Em se tratando de bibliotecas especializadas, é provável que, na medida em que a especialidade evolua, haja um aprimoramento natural na construção dos textos, por exigência dos próprios cientistas. Se o texto, por exemplo, segue padrões canônicos em sua estrutura e apresenta-se bem redigido, a leitura pode não só ser facilitada, como também aumentar a probabilidade do trabalho documental ganhar em precisão. Entretanto, a qualidade do texto não depende do bibliotecário.

Ao contrário, o conhecimento prévio e as estratégias de leitura estão, de alguma forma, vinculados a habilidades do profissional da documentação e é nesses aspectos que delimitamos nosso trabalho.

Numa visão mais global do processo, dizemos que todo texto é produzido para determinados receptores e que a sua eficácia depende, em boa parte, da capacidade do autor em estabelecer com seus leitores potenciais uma relação cooperativa (GRICE, 1967).

Na leitura para fins documentários, entretanto, o princípio de cooperação autor/leitor é rompido, já que, normalmente, o autor

não previu o documentalista como leitor. Neste caso, à complexidade natural do processo tem de ser acrescido esse fator.

Ainda no âmbito do mais global, descartamos a possibilidade de uma codificação "neutra", tendo em vista qualquer objetivo de leitura, mesmo que seja uma análise documentária. Sendo a linguagem intrinsecamente comprometida com o cultural e o ideológico, tanto o processo de produção do texto, quanto o de recepção não se isentam destes componentes. E é esse comprometimento que garante o caráter interativo da leitura.

Ao considerar conhecimento prévio e tipo de estratégias que o texto exige como fatores de legibilidade num processo interativo, estamos reconhecendo que toda leitura envolve mais que o conjunto de sinais visuais que compõem o texto. Na verdade estamos admitindo também, como básicos, aspectos cognitivos relacionados com conhecimentos armazenados na memória do leitor e comportamentos específicos durante a leitura.

A noção de conhecimento prévio vincula-se à teoria dos "esquemas" que vem se desenvolvendo com maior ênfase a partir de fins da década de 70, graças a estudos desenvolvidos na área da cognição. Esses estudos têm assumido diferentes feições (MARCUSCHI, 1985 e MEURER, 1985), dentre as quais coloca-se a de Van Dijk (1980), para quem "esquemas" são unidades que organizam seqüências de eventos. Para Kato (1983), que adota posição semelhante à de Rumelhart (1980, p. 26), "um Esquema é uma teoria prototípica do significado, porque corresponde ao significado de um conceito codificado em termos de situações ou eventos típicos, ou normais, que instanciam esse conceito".

Neste trabalho entendemos por "esquema" algo mais próximo a essa última concepção, uma vez que a prática profissional do bibliotecário não exige que ele estoque na memória apenas seqüências de eventos como ocorre, freqüentemente, com os relatos da vida diária. O bibliotecário deve lançar mão de conhecimentos armaze-

nados em sua memória os quais constituem uma espécie de quadro de referência, formado por uma rede multidimensional de unidades conceituais, a partir da qual o *input* visual é avaliado.

Essa concepção de "esquema" permite dizer que o leitor trabalha tanto com um quadro de referência composto de entidades lingüísticas, quanto conceituais e dele depende para a compreensão do texto, para a construção de novos conhecimentos e para o trabalho operacional com o texto. Até mesmo as situações do dia-a-dia recebem respostas analógicas ou automáticas em função de "esquemas" armazenados em nossa memória, através de uma organização ativa das reações do nosso passado (MEURER, 1985).

No trabalho de leitura, entendemos ser indispensável considerar como entidades lingüísticas estocadas, não apenas o nível da palavra, de blocos significativos e de frases, como também de "esquemas" formais de organização textual. Sabemos que também na leitura para fins documentários, o leitor não lê letra por letra, palavra por palavra e até mesmo não lê, freqüentemente, todas as palavras do texto, quando as seqüências são previsíveis (CINTRA, 1983). Sobrepondo ao texto conhecimentos prévios de vários níveis, o leitor avança na compreensão e imprime maior velocidade à leitura.

Por outro lado, todo texto se constrói segundo alguns princípios tipológicos. Os constituintes básicos de um determinado tipo de texto definem a sua superestrutura.

Pesquisas têm demonstrado que leitores com conhecimento prévio específico sobre superestruturas textuais executam a tarefa de ler de forma mais fácil que leitores que não "vêm" essa superestrutura e por isso são obrigados a um maior apoio na leitura palavra por palavra, dificultando a integração das informações no texto como um todo (MEURER, 1985).

pondência entre a ordem linear do texto e a ordem temporal dos eventos, de interpretação de vários sintagmas do texto como sendo correferenciais. Kato (1984) salienta que essa última tendência é explicada pelo princípio da parcimônia de Fillmore e Kay (1981, pp. 104-105), "segundo o qual o leitor tende a reduzir os personagens do cenário mental que se forma a partir do texto".

Assim o leitor, automaticamente busca as informações principais procedendo a uma espécie de descarte do que é secundário.

Kato ainda considera outros dois princípios como explicativos das estratégias cognitivas: o princípio da canonicidade e o princípio da coerência. O princípio da canonicidade atua a nível sintático e semântico, na medida em que o leitor decodifica automaticamente, por exemplo, a ordem natural sujeito/verbo/predicado, a oração principal antes da subordinada (plano sintático); o animado antes do inanimado, o agente antes do paciente, a causa antes do efeito, a tese antes da antítese (plano semântico).

Ao nível textual também atua cognitivamente o princípio da canonicidade, uma vez que o leitor espera encontrar sentenças tópico nos parágrafos e constituintes básicos de superestrutura textuais.

O outro princípio é o da coerência do texto. O leitor espera um texto coerente em todos os níveis. Ao iniciar a leitura, até mesmo pela percepção primeira do texto, cria expectativas de coerência global e temática. A coerência local é algo esperado passo a passo na leitura do texto.

As estratégias cognitivas ficam, então, como expectativas inconscientes, como atividades realizadas automaticamente em função de "esquemas" prévios armazenados na memória de longo termo do leitor. No caso de qualquer insucesso o leitor dispõe de mecanismos ou estratégias de superação que aplicadas, conscientemente, podem levar à compreensão.



Com base nas estratégias elencadas por Brown (1980), dizemos que o leitor experiente se caracteriza por uma prontidão para:

- monitorar, enquanto lê, a finalidade e a compreensão da leitura;
- identificar as partes mais importantes do texto;
- concentrar mais atenção sobre conteúdos principais, basicamente sobre o tema;
- fazer a segmentação do texto, identificando as macro-proposições semânticas, isto é, as seqüências que contêm as informações principais;
- proceder a ações corretivas, quando são detectadas falhas no processo.

Concluindo dizemos que, na leitura para fins documentários, os textos são desautenticados, na medida em que são deslocados de seus contextos naturais e mesmo assim os dois fatores ativos que determinam a legibilidade têm, nessa tarefa, papel preponderante.

Enquanto o conhecimento prévio viabiliza, por força de "esquemas", uma leitura mais rápida, as estratégias, especialmente as metacognitivas, conduzem à eficácia da tarefa.

Embora as estratégias cognitivas tenham nesse tipo de leitura papel tão fundamental quanto em outro qualquer, pois comandam a compreensão ortográfica, sintática e semântica do texto, entendemos que deva haver um trabalho específico para a passagem de estratégias cognitivas para metacognitivas, especialmente a nível de texto. Julgamos que o domínio consciente de algumas etapas permite identificar "fatias" ou segmentos significativos do texto, estabelecer relações de sentido e referência entre certos segmentos, manter a coerência entre as macro-proposições semânticas do texto, o que permite avaliar a consistência das informações, inferir até mesmo o significado e o efeito pretendido pelo autor.

O leitor que domina as superestruturas textuais, capta com mais facilidade as idéias centrais do texto, pois tem como parâmetro a identificação dos constituintes básicos. Assim, num texto dissertativo argumentativo, por exemplo, o leitor faz predições que vão sendo testadas sobre a introdução, a tese, os argumentos, a conclusão, ordenando as idéias centrais e descartando as idéias secundárias.

A necessidade eventual de leitura palavra por palavra permite aos especialistas dizer que os "esquemas" do leitor são ativados por dois movimentos complementares: o movimento *botton-up*, segundo o qual o leitor vai lendo na dependência estrita do contexto, ou seja, vai extraíndo, linearmente, dos símbolos impressos o significado, caminhando assim das partes para o todo o movimento *top-down*, no qual há maior dependência de conhecimento prévios do leitor, pois ele vai fazendo generalizações e predições a partir de "esquemas" que tem armazenados em sua memória. Ao caminhar do todo para as partes, vai confirmando ou não suas hipóteses sobre o texto.

Em especial o segundo movimento dá conta de um aspecto de importância fundamental: a capacidade de predição do leitor que corresponde à formulação de hipóteses decisivas para a compreensão, além de fortalecer a própria organização do conhecimento individual.

Estudos atuais sobre leitura consideram a predição como componente essencial do processo de ler e tomam os vários comportamentos hipotetizados pelo leitor, enquanto lê, como estratégias de leitura, ou seja, ações mentais que o levam a atingir seus objetivos.

Na leitura, então, o leitor toma o texto como suporte para predições, extraíndo dele mais do que está expresso linguisticamente, graças a ações mentais estruturadas, ou estratégias de

leitura.

O fenômeno da inferência, por exemplo, que corresponde à apreensão de informações não explícitas no texto, depende, basicamente, de "esquemas" estocados na mente do leitor, a ponto de ser, praticamente, impossível reproduzir literalmente as informações impressas, sem a interferência dos "esquemas" mentais nos conteúdos inferenciais (MEURER, 1985).

A seleção de estratégias durante a leitura, além dos "esquemas" prévios do leitor, depende de outros fatores: da finalidade da leitura, da experiência do leitor, ou da sua maturidade frente à tarefa de ler, do tipo de texto lido, da atenção mais concentrada em partes do texto, do grau de novidade do texto e até mesmo da motivação para ler.

No processo de leitura pode-se distinguir dois grupos de estratégias: as cognitivas que compreendem comportamentos automáticos e inconscientes e as metacognitivas que supõem comportamentos desautomatizados, na medida em que o leitor tem consciência de como está lendo.

Ainda que toda leitura envolva esses dois tipos de estratégias, é provável que quanto menos atividades metacognitivas exigir, mais legível será o texto. Entretanto, é também provável que a leitura apenas automática conduz à incompreensão. Desta forma, admite Kato (1985, p. 64) que "a legibilidade pode ser igualmente uma função do equilíbrio entre a natureza parcialmente prototípica e não prototípica do *input*: a primeira que favorece a assimilação e a segunda que motiva o interesse. Em outras palavras, o texto legível é aquele que exige uma aplicação equilibrada de estratégias cognitivas e metacognitivas".

As estratégias cognitivas compreendem processamentos automáticos e inconscientes de interpretação de palavras e frases, de corres-

- BROWN, A.L. (1980). Metacognitive development and reading. In: SPIRO et al. (org.). *Theoretical issues in reading comprehension*. Hillsdale, N.J., Lawrence Erlbaum Associated Publishers, pp. 453-482.
- CINTRA, A.M.M. (1983). Elementos de lingüística para estudos de indexação. *Ciência da Informação*, Brasília, 12(1):15-22.
- FILLMORE, C. & KAY, B. (1981). Three properties of the ideal reader (xerox), apud KATO, M.A. Estratégias cognitivas e metacognitivas na aquisição da leitura. In: ENCONTRO INTERDISCIPLINAR DE LEITURA, Londrina, 1984. *Anais*, pp. 102-115.
- GRICE, H.P. (1967). Lógica e conversação. In: DASCAL, M. *Fundamentos metodológicos da lingüística*. Campinas, 1982, pp. 81-103 (Pragmática, 4).
- KATO, M.A. (1983). Estratégias em interpretação de sentenças e compreensão de textos. *Lingüística*. São Paulo, EDUC, pp. 9-33 (Cadernos PUC, 16).
- _____. (1984). Estratégias cognitivas e metacognitivas na aquisição da leitura. In: ENCONTRO INTERDISCIPLINAR DE LEITURA, Londrina, 1984. *Anais*, pp. 102-115.
- _____. (1985). Uma visão interativa da legibilidade. *Ilha do Desterro*, (13):57-66.
- MARCUSCHI, L.A. (1985). Leitura como processo inferencial num universo cultural cognitivo. *Leitura: Teoria & Prática*, Porto Alegre, 4(5):3-16 jun.
- MEURER, L. (1985). Schemata in reading comprehension. *Ilha do Desterro*, (13):31-46.
- RUMELHART, D.E. (1980). Schemata: the building blocks of cognition. In: SPIRO et al. (ed.). *Theoretical issues in reading comprehension*. Hillsdale, Lawrence Erlbaum Associates Publisher, pp. 35-58.
- VAN DIJK, T.A. Textwissenschaft, apud MARCUSCHI, L.A. (1985). Leitura como processo inferencial num universo cultural cognitivo. *Leitura: Teoria & Prática*, Porto Alegre, 4(5):3-16, jun.

3. ANÁLISE DOCUMENTÁRIA

Isabel M. R. Ferin Cunha

3.1 PARÂMETROS TEÓRICOS

3.1.1 As análises possíveis e a análise documentária

3.1.2 Lingüística da análise documentária

3.1.3 Lógica e "lógicas"

3.1.4 Ideologia. Linguagem/ideologia

3.2 PARÂMETROS EXPERIMENTAIS

3.2.1 Metodologia de análise

3.2.2 Metodologia de síntese

BIBLIOGRAFIA

3. ANÁLISE DOCUMENTÁRIA

Isabel M. R. Ferin Cunha*

Comumente a Análise Documentária (AD) é definida como um conjunto de procedimentos efetuados com o fim de expressar o conteúdo de documentos, sob formas destinadas a facilitar a recuperação da informação.

Até recentemente, os métodos desta análise foram considerados como operações empíricas de "bom senso", pelas quais se atribua a um texto/documento qualquer, uma ou várias palavras-chave destinadas a facilitar a recuperação do documento, no momento da pesquisa, sobre um dado tema. À leitura dos textos/documentos com fins de Análise Documentária, a Biblioteconomia/Documentação atribui, implícita ou explicitamente, uma leitura única, fechada e universal, independentemente do leitor/analista da documentação. Um outro preconceito surge na atribuição de palavras-chave aos textos/documentos, na medida em que se consideram as palavras "neutras", isto é, desprovidas de sentidos.

Contudo, esta passagem de um texto original para um tipo de representação é uma operação semântica, mesmo que não obedeça a nenhuma regra precisa e varie em função de cada organismo e do analista que seleciona as palavras-chave, normalmente de forma intuitiva, em função da sua ocorrência e do seu interesse para a instituição. No entanto, mesmo trabalhando de uma forma empírico-pragmática o analista da documentação tem de ter em conta algumas variáveis, notadamente, as do discurso (com pretensões científicas ou não) realizadas através de uma língua dada e que compreende a sua organização interna e a linguagem/ideologia. Estas

* O presente capítulo constitui um extrato da tese de doutoramento "Do mito à análise documentária: a luso-assimilação", defendida na ECA no 1º semestre de 1987.

variáveis surgem não só na fase de análise, como na de obtenção do produto documentário. Nesta última, é necessário ter em conta o sistema em que se encontra inserida a disseminação desse produto-informação e a variável produtor-tradutor-consumidor da mesma, sendo que a estas é inerente a existência de uma linguagem/ideologia refletora do pensamento e da situação em que se encontram inseridos.

Com a automação, o produto da análise deixou de se limitar a três ou quatro palavras-chave, passando-se a obter representações muito mais ricas, na medida em que paráfrases completas do enunciado foram registradas em memória, numa linguagem simbólica adequada. Contudo, a automação, que trouxe enormes vantagens no isolamento de conceitos por ocorrência e frequências, mostra-se ainda ineficiente quando aplicada a textos de Ciências Humanas, com fins documentários, mesmo quando esses textos se dizem científicos, o que pressuporia uma preocupação lógica/metodológica maior e uma maior precisão nos conceitos e termos utilizados.

Nesse sentido, a Análise Documentária, tanto realizada de forma manual como automática, tem procurado desenvolver léxicos de base ou léxicos intermediários destinados a preencher o papel de sistemas de conversão entre conceitos apresentados de forma independente nas diversas linguagens, e conceitos de leitura "universal" definidos pela própria Análise Documentária.

Esta formulação leva-nos a considerar que no momento da Análise Documentária existem três linguagens (para fins documentários, não lingüísticos) interligadas: a linguagem natural (LN), as linguagens especializadas (LE) e as linguagens documentárias (LD). Sendo que, quer por processos empíricos ou de "bom senso", quer usando um procedimento metodológico ou automático, o que se pretende é conseguir que um mesmo texto submetido a analistas diferentes dê lugar a representações semelhantes, condição necessária não só à transmissão da informação, como à validação

de instrumentos como tesouros, vocabulários controlados etc.

Contudo, esta interligação entre linguagens naturais (LN), linguagens especializadas (LE) e linguagens documentárias (LD) realiza-se de forma diferente nas Ciências Exatas e nas Ciências Humanas, devido à maior ou menor imprecisão dos respectivos sistemas simbólicos. Isto é, enquanto nas Ciências Exatas o vocabulário especializado (LE, em documentação) tem grandes chances de se transformar automaticamente na LD, independentemente dos problemas de controle, hierarquização e tradução que possa acarretar, nas Ciências Humanas existe uma enorme margem de confusão entre LN e LD, na medida em que o vocabulário especializado se encontra ligado a escolas de pensamento e ideologias veiculadas através da LN. Esta questão leva-nos a afirmar que as dificuldades da AD crescem em proporção ao uso em textos/documentos científicos da linguagem natural, por conseguinte, ao ter-se como objeto os discursos em Ciências Humanas.

Na realidade, as Ciências Exatas caracterizam-se pelo emprego de sistemas simbólicos próprios e exigências epistemológicas testadas por validação, enquanto nas Ciências Humanas dominam "dialetos" estritamente aparentados a esta ou àquela língua natural, cujos sistemas simbólicos restringem-se a locais e campos específicos, sem exigências epistemológicas concretas de validação.

Para J. C. Gardin (GARDIN, LAGRANGE e outros, 1981) esta oposição não é senão uma oposição derivada do "estado primitivo" em que se encontram a maior parte das Ciências Humanas, a meio caminho de uma simbolização dos fatos que estudam, incapazes de organizar através de uma língua natural um discurso de caráter estritamente científico. Segundo o mesmo autor, é na capacidade de formular um sistema próprio à captação dos fenômenos em Ciências Humanas que se encontra a viabilidade de um discurso científico nas mesmas. Sem ele, as CH continuarão quer a

socorrer-se de sistemas simbólicos já estabelecidos, quer a manter como critério único o estilo e a erudição do autor, distanciando-se das exigências epistemológicas.

Esta problemática resulta, para nós, numa nova atribuição à AD em Ciências Humanas, para além da descrição do sentido/conteúdo dos textos com vista à recuperação da informação. Esta atribuição é, ao mesmo tempo, hipótese metodológica de procedimento, isto é, a AD teria como função não só analisar como o autor/produzidor passa de um conjunto de observações dadas à sua interpretação, mas também, como são montados os discursos em CH e de que forma esses discursos contribuem para a ciência em estudo. Só através destas duas funções de análise chegar-se-ia a uma real recuperação da informação.

Estas e as problemáticas levantadas anteriormente, levam-nos a propor alguns parâmetros teóricos para a AD.

3.1 PARÂMETROS TEÓRICOS

3.1.1 As análises possíveis e a análise documentária

Como foi anteriormente dito, a Análise Documentária visa expressar o conteúdo dos textos/documentos com o fim de recuperar e disseminar a informação. Na medida em que procura expressar o conteúdo dos textos/documentos ela se aproxima de outras realizadas com fins idênticos. É o caso das análises de conteúdo, literárias, semióticas e lingüísticas, realizadas quer manualmente quer de forma automática (GARDIN, 1974).

Em função desta coincidência de objetivos, é interessante observar até que ponto estes diferentes tipos de análises aproximam-se e diferenciam-se da Análise Documentária.

A análise de conteúdo é um conjunto de técnicas destinadas a

ressaltar o conteúdo de diversas categorias de documentos escritos, certos elementos singulares de onde se inferem características psico-sociológicas do objeto analisado. Ela pretende enfocar tanto a expressão como a coisa, independentemente da espécie de textos considerados (isolados ou por conjuntos), desde que o analista suspeite que haja no conteúdo inferências úteis para o conhecimento dos homens que as produzem ou do homem em geral.

A hipótese de trabalho que se encontra subjacente a esta análise é a da existência no texto, ou textos, de elementos que tenham valor de indicadores, individualizando a distribuição das ocorrências do texto ou "corpus". No processo de evidenciar os indicadores que singularizam o texto, os analistas recorrem normalmente à interpretação, o que torna "análise de conteúdo" um nome convencional dado a toda espécie de estudos de textos destinados a esclarecer questões relativas aos produtores dos textos e seus consumidores.

A necessidade de extração de elementos com valores de indicadores, na análise de conteúdo, explica o fato das primeiras tentativas de formalização desta análise terem sido, na maior parte dos casos, orientadas para a elaboração de metalinguagens de tradução — indicadores semânticos — que objetivassem os mecanismos de interpretação semântica. Os problemas de tradução foram resolvidos quer matematicamente por estudos de ocorrência, quer pela passagem de uma linguagem natural a uma metalinguagem, sem contudo chegar-se à formulação de uma teoria que a justificasse.

Estes procedimentos transferem para os instrumentos a validação da análise, caso da automação ou da linguagem de representação como meio e não produto de análise.

Existem, no entanto, aproximações entre a análise documentária e a análise de conteúdo, na medida em que o que está presente é um

texto/documento a ser analisado e a necessidade de se chegar a um produto através de um sistema simbólico baseado na língua do texto/documento considerado, sendo que as duas análises têm como exigência de procedimento a obtenção de uma metalinguagem tradutora, ou seja, palavras representativas do texto/documento.

Em relação à análise literária, a análise documentária encontra algumas diferenças fundamentais quanto aos seus objetivos, procedimentos e finalidades. Dois tipos de análises literárias são normalmente praticadas: uma privilegiando o valor de uma escola analítica e definindo-se em função dessa gramática específica (o formalismo, o estruturalismo, o transformacionalismo) ou de uma semântica própria (psicanálise, marxismo, fenomenologia), sendo que a gramática das interpretações é validada em função da credibilidade atribuída ao aparelho utilizado. O outro tipo de análise literária é realizada tendo como ponto de partida uma hipótese a demonstrar e a ser confirmada pelo texto, utilizando para tal mecanismos e categorias de um sistema interpretativo singular.

O primeiro tipo de análise oscila em função das escolas e dos modelos de pensamento e, dado este fato, a sua contribuição para a análise documentária apenas interessa como experiências adquiridas e resultados obtidos, visto que a discussão enfoca, antes, a excelência dos modelos possíveis do que os produtos obtidos. Quanto ao segundo tipo de análise, interessa à análise documentária na medida em que propõe um tipo de raciocínio hipotético-dedutivo muito usado na atividade documentária enquanto "bom-senso".

Dados estes fatores, não existe, na análise literária, uma preocupação a nível da metalinguagem ou da tradução, assim como da explicitação e formalização dos procedimentos na organização de um texto, objetivo da análise documentária.

Na análise semiótica a preocupação pela formulação de modelos e

pela simbolização da linguagem é semelhante à Análise Documentária. Neste sentido, interessou-nos, principalmente, a proposta de Greimas quanto ao estudo das modalidades e organização lógica do discurso científico. Na sua proposta de identificação do discurso do saber e da pesquisa, no estudo das modalidades, encontram-se possibilidades não só para a Organização Lógico-Semântica do texto, mas também para a determinação das funções/categorias dentro de um vocabulário controlado. Na realidade, a análise contextual das diferentes hipóteses de tradução dos verbos modalizadores ser/fazer, leva-nos a acreditar na possibilidade de desenvolver um modelo de análise documentária capaz de melhor determinar as funções e categorias das palavras-chave, construtoras de instrumentos documentários.

Contudo, a Análise Documentária diverge da semiótica, na medida em que à primeira interessa um fazer pragmático, tendo como objetivo tanto o sentido do texto como a sua tradução em linguagens documentárias e não somente a apreensão de linguagens em sistemas de modelos universais.

3.1.2 Linguística da Análise Documentária

A importância da linguística para a Análise Documentária não se pode restringir apenas às suas propostas de análise. É necessário, dado que a Análise Documentária trabalha com a língua e elabora linguagens de documentação, fazer um recorte teórico da linguística no que ela apresenta de instrumental para a Análise Documentária, não só na fase de procedimento, como na elaboração de instrumentos e produtos a serem obtidos.

Podemos definir a linguística como a ciência que estuda as línguas naturais tendo como objetivo preferencial o estudo da sua modalidade oral. Ela visa, também, ao estudo da teoria das línguas, na medida em que pretende sistematizar as suas invariantes.

A língua funciona segundo dois eixos (estudados pela gramática): o eixo sintagmático, que rege os agrupamentos possíveis dos diferentes signos lingüísticos e o eixo paradigmático, que reflete as relações existentes entre os signos capazes de assumir a mesma função. Esses dois eixos são virtuais como a própria língua e determinam a organização do discurso. No caso da Análise Documentária, que tem como objeto textos, isto é, enunciados finitos e espaços discursivos limitados, o reconhecimento das relações paradigmáticas e sintagmáticas leva não só à formulação de campos semânticos dos textos analisados como à estruturação e elaboração das linguagens tradutoras, permitindo a passagem de uma LN — LD. É através das relações paradigmáticas e sintagmáticas que se torna possível a reunião de unidades lexicais de uma língua em grupos estruturados, de tal modo que cada unidade se defina pelo lugar que ocupa relativamente às demais. Para tal, a formação do Campo Semântico possibilita à Análise Documentária uma melhor organização dos seus procedimentos de análise e a construção de instrumentos e produtos.

Para que seja possível a identificação das relações paradigmáticas e sintagmáticas e a formação de Campos Semânticos, com o fim de passar de uma LN — LD, a Documentação tem necessidade de elementos de filologia, no sentido do estudo diacrônico e sincrônico dos vocábulos, assim como da Lexicografia e da Lexicologia, principalmente no que diz respeito à organização e estruturação de dicionários, léxicos e glossários. Neste sentido, a lingüística colabora com a Análise Documentária, tanto no reconhecimento dos textos como na estruturação das LD's.

Sendo que não existem estudos interdisciplinares com objetivos definidos que envolvam a documentação e lingüística, dizemos que a primeira recorta da segunda as propostas que lhe interessam no seu fazer pragmático. Damos alguns exemplos:

A diferenciação realizada por Saussure (SAUSSURE, 1969) entre

Linguagem (abstração humana, capacidade do homem de comunicar-se com os seus semelhantes através de signos verbais), língua (conjunto de todas as regras fonológicas, morfológicas, sintáticas e semânticas que determinam o emprego de sons, das formas e relações sintáticas necessárias para a produção dos significados) e fala (combinatória individual que utiliza elementos discriminados dentro do código) permite à Análise Documentária em embasamento na distinção entre linguagem natural e linguagem documentária. Na realidade, viabiliza a definição de linguagem documentária enquanto gramática, sintaxe construída a partir de um campo semântico previamente determinado, assim como remete a construção da LD para o conhecimento de um todo língua efetivado através de enunciados singulares - o discurso.

O desdobramento das propostas Saussurianas pelo estruturalismo, através das Escolas de Genebra, Círculo Linguístico de Praga, de Copenhague e Escola Mecanicista de Bloomfield, tiveram como preocupação dominante o estudo de enunciados efetivamente realizados, o estudo das unidades mínimas da língua, preocupando-se em trabalhar esses enunciados fora da situação ou da enunciação em "corpus" delimitado. Resultado destas preocupações foram descrições e taxionomias de conceitos operatórios de base e estruturas elementares.

O estruturalismo interessa à Análise Documentária na medida em que permite estudos distribucionais e de ocorrência, importantes quando da elaboração de instrumentos de análise, como é o caso dos vocabulários controlados, tesouros, sistemas de classificação. No entanto, ao ter como preocupação unidades mínimas da língua e o estudo de enunciados deslocados da situação de enunciação, afasta-se da Análise Documentária, que tem como objeto o discurso e como objetivos a identificação do seu conteúdo.

A proposta transformacionalista de Chomsky (CHOMSKY, 1970) dá à gramática o papel de gerador de um número infinito de frases, sendo que ela é necessariamente finita. Neste sentido, a gramática

deverá consistir num conjunto finito de regras capazes de gerar um conjunto infinito de frases gramaticais de uma língua dada. Dá-se, portanto, uma inversão em relação à proposta da gramática estrutural. Em lugar de uma análise, propõe-se uma síntese, manipulam-se sistemas de regras e não um sistema de elementos. Dado isto, o que mais interessa à Análise Documentária é a tentativa de explicar, através das gramáticas, as estruturas lógicas do pensamento, capazes de levar ao sentido dos discursos, além do aprimoramento de sistemas automáticos de indexação como o KWIC e o KWOC, iniciados dentro de uma perspectiva distribucionalista de ocorrências, frequências, posteriormente, incorporando programas de função ao nível de enunciado.

O desdobramento da proposta transformacionalista, centrando o objeto de estudo na possível correspondência entre estrutura gramatical e estruturas lógicas do pensamento, levou a novas propostas: a gramática de casos de Fillmore (FILLMORE, 1977) e os casos conceptuais de Pottier (POTTIER, 1974). Para estes linguistas o importante é isolar os universais linguísticos, através de conceitos universais, capazes de receber em uma função operacional, fundamentada em perguntas básicas: quem fez isso, a quem aconteceu isso, o que vai mudar, onde aconteceu. Nestas perguntas, casos universais, a Análise Documentária encontra as categorias necessárias para a ordenação dos campos semânticos, já identificados, possibilitando a estruturação de uma LD (o que foi automatizado no PRECIS). Ao ter como objeto de análise, tanto as unidades mínimas do discurso como a análise da estrutura lógico-gramatical da frase e do sentido gerado pelo seu conjunto no interior do discurso, a gramática de casos converge para os procedimentos e objetivos da Análise Documentária. Apesar desta convergência, as diferenciações surgem na medida em que à Análise Documentária interessa um produto a ser traduzido de uma linguagem natural para uma linguagem documentária e, nesse sentido, já não são as unidades mínimas que estão em jogo, mas sim o todo discurso/texto/documento.

Para terminar o recorte lingüístico, com vista à Análise Documentária, é necessário ter em conta que esta utiliza a lingüística enquanto universo teórico capaz de subsidiar a sua atividade pragmática. Neste sentido, não lhe interessam as propostas das diferentes correntes lingüísticas enquanto escolas de pensamento, com as suas terminologias, modelos e sistemas fechados, mas sim, os elementos e conceitos possíveis de serem utilizados no seu fazer pragmático. Esta questão leva-nos, por um lado, a uma dificuldade de precisão terminológica e conceitual dentro da Análise Documentária quando utiliza o jargão das diferentes correntes lingüísticas utilizadas, por outro, confere à Análise Documentária uma mobilidade de procedimentos, permitindo-lhe uma abordagem teórica mais rica e globalizante do objeto em análise.

Convém ainda referenciar o que se entende por análise automática e o que ela tem em comum com a análise documentária. Na realidade, a análise automática é um procedimento aplicável a qualquer análise anteriormente descrita. Ela aproxima-se da documentária na medida em que ambas pretendem ser um exercício de explicitação das organizações lógico-semânticas que sustentam certos textos ou classes de textos. Para tal, ambas propõem procedimentos interpretativos explícitos e avaliações de resultados, isto é, têm como preocupação a validação dos seus programas e procedimentos. Contudo, a Análise Documentária difere da análise automática, na medida em que tem de ter em consideração os seus objetivos específicos de recuperação da informação. Para tal, necessita colocar em equação as variáveis produtor/produção/contexto de produção/consumo/consumidor, ou seja, equacionar os vários sistemas em que se encontra inserida a informação contida no texto, a ser traduzida automaticamente.

3.1.3 Lógica e "lógicas"

A utilização do "bom-senso" por parte dos bibliotecários analistas da documentação implica numa série de procedimentos "lógicos". Esses procedimentos resumem-se, normalmente, na identi-

ficação e estruturação, através de instrumentos específicos, das chamadas "informações significativas ou relevantes". Nesta operação, estão implícitas relações de inclusão/exclusão, todo/parte, gênero/espécie, embasadas em juízos e raciocínios, além da capacidade de programar a tradução do conteúdo do texto em etapas sequenciais lógicas. No mesmo procedimento de identificação e tradução de informações significativas encontram-se, implícitas, as questões relativas à construção do discurso do autor/produtor, que quando científico envolve a identificação e avaliação (na medida em que o bibliotecário/analista da documentação decide o que é significativo ou não) dos métodos, hipóteses, leis, teorias e resultados, utilizados pelo autor/produtor com o fim de chegar à determinada "informação nova".

Apesar de todas estas implicações a nível do conhecimento, a Lógica enquanto ciência encontra-se bem distante das preocupações dos bibliotecários/analistas da documentação, mesmo quando se utiliza automação. Desta forma, consideramos que, do mesmo modo que a Análise Documentária utiliza a lingüística, enquanto instrumento teórico para a sua atividade pragmática, a lógica interessa à Análise Documentária enquanto parâmetros de procedimentos, entendidos mais como aparatos formais que como metodologia ou epistemologia da ciência do discurso.

Dado isto, a Análise Documentária deve ter em consideração a lógica enquanto lógica geral e lógica formal. Na lógica geral, e principalmente no que se refere ao discurso científico em CH, propomos examinar no texto os processos de aquisição dos conhecimentos científicos, assim como a construção dos princípios, das hipóteses, leis gerais e teorias. Na lógica formal interessa-nos determinar como se elaboram os raciocínios dum ponto de vista formal.

Em ambos os sentidos de lógica que teremos de considerar, encontram-se inerentes as questões tematizadas na Teoria do Conhe-

cimento, isto é, as questões relativas ao ponto de partida do conhecimento, às fontes primeiras; o processo com que os dados se transformam em juízos ou afirmações acerca de algo; a maneira como o sujeito se posiciona frente ao objeto a ser conhecido e o âmbito do que pode ser conhecido.

Contudo, ao levantar estas questões, não pretendemos entrar no debate filosófico estabelecido ao longo dos tempos, nem trabalhar as diferentes formas como foram encaminhadas e solucionadas as diferentes abordagens. A nossa orientação dá-se no sentido de resolver, pragmaticamente, a nossa problemática, o que nos leva a recortar o CONHECIMENTO, tendo em mente tanto a coerência do recorte teórico realizado, como o nosso objetivo específico, resumível em dois itens fundamentais: identificação das macroproposições semânticas*, Traços Descritivos apresentados pelo autor/produtor no seu texto e dos quais decorre a Organização Lógico-Semântica (OLS) do discurso; demonstrar como se passa dessas macroproposições semânticas/traçoes descritivos para palavras-chave/conceitos que os representem.

Dados os objetivos expostos, o nosso posicionamento teórico realiza-se no sentido de levantar instrumentais capazes de assegurar uma maior objetividade, tanto nos procedimentos de análise do texto como na fixação de palavras-chave/conceitos, de forma a explicitar-se o "bom senso" usual através de parâmetros teóricos estabelecidos, mesmo que extrapolados. É o caso da utilização que propomos de definições de juízo, conceitos, argumentos e modalidades lógicas, independentemente das escolas e princípios gerais que a originaram.

* Proposições que resumem seqüências de proposições sem entretanto pretenderem dar uma descrição completa das mesmas, mas apenas uma amostragem (VAN DIJK & KINTSCH, 1983).

Voltando às duas conceituações lógicas, anteriormente expostas e à sua aplicação pragmática na Análise Documentária, surge-nos ao nível formal a identificação de como o objeto do discurso dito científico, é isolado pelo autor/produtor, como o autor/produtor passa de um conjunto de observações dadas à sua interpretação, reconstruindo um novo discurso dentro do discurso. Enquanto procedimento, a Análise Documentária visa, nesta fase, montar um arcabouço simbólico-abstrato capaz de sintetizar, esquematicamente, os procedimentos do autor/produtor. Ao nível da lógica geral, torna-se necessário ter em presença o método, os princípios, leis gerais e teorias formuladas em relação ao objeto do discurso.

Estas questões levam-nos a afirmar que a cientificidade de um discurso dependerá da articulação entre as exigências da lógica formal e as exigências da lógica geral, sendo que a esta última a análise documentária apenas colocará a questão de saber até que ponto o discurso em análise contribui, ou não, para a ciência em discussão.

Ao utilizar a Análise Documentária como instrumento de análise do discurso dito científico, a nossa proposta é identificar a cientificidade do discurso através dos procedimentos desta análise, ou seja: como procede o autor/produtor para isolar o seu objeto de análise; como passou de um conjunto de observações dadas à sua interpretação; se contribuiu ou não para a ciência em discussão.

Estamos, então, perante dois níveis de análise - contrariando a tradicional visão bibliotecária do texto "absoluto" - a ser realizada pelo bibliotecário/analista da documentação; a análise do texto, tentando desmontar a construção do autor/produtor e a Análise Documentária, propriamente dita, visando à identificação das macro-proposições semânticas/traços descritivos/informações "significativas", capazes de traduzir o conteúdo sob a forma de conceitos/palavras-chave. Na realidade, estamos também perante dois objetivos: o do autor/produtor, na medida em que constrói o

texto com fins específicos, e o do analista da documentação, interessado em decodificar esse texto em "informação nova" a ser traduzida de uma linguagem natural para uma linguagem documentária.

3.1.4 Ideologia. Linguagem/ideologia

Nestes dois níveis de análise com fins documentários, coloca-se uma nova problemática: a questão da linguagem/ideologia, visto que os textos/discursos estão escritos numa língua natural e pretendem "convencer" o leitor da tese que defendem; e a ideologia do analista da documentação (enquanto ser humano integrado num sistema social), tradutor dos conteúdos dos textos/discursos para todos, conceitos/palavras-chave.

Ao colocarmos esta problemática, pretendemos pôr em discussão o preconceito bibliotecário de "leitura única e absoluta", assim como a existência de conceitos/palavras-chave e bibliotecários/analistas da documentação "neutros". Isto é, consideramos que o leitor/bibliotecário/analista da documentação tem sempre uma visão ideológica, sobrepondo-a à linguagem/ideologia do texto/discurso a analisar. Essa sobreposição se manifesta pela opção "ideológica" que faz em relação ao uso ou descarte de determinados conceitos/palavras-chave, mesmo quando se reporta às regras de objetividade e neutralidade aconselhadas pelos manuais e pela ética profissional vigente.

Neste sentido, consideramos necessário colocar a questão da ideologia e linguagem/ideologia. No entanto, da mesma forma que não se pretende trabalhar a lingüística e a lógica enquanto ciências, a nossa abordagem em relação à ideologia e linguagem/ideologia, far-se-á recortando, das várias teorias, as concepções que nos parecem mais aplicáveis ao nosso fazer pragmático e ao texto a analisar como amostra.

Dado isto, e tomando como pressuposto a afirmação de Veron (VERON, 1978, p.7) "L'ordre de l'idéologique et l'ordre du pouvoir traversent de part en part une société", pensamos ser o preconceito da "neutralidade bibliotecária" uma escamoteação dos problemas de análise, na medida em que qualquer discurso surge como um complexo de idéias e representações articuladas entre si, sendo que elas são detectáveis pela própria análise e apresentam uma natureza essencialmente política, tomando-se o termo político na sua acepção sócio-econômica mais ampla.

Na realidade, negar o componente ideologia em análise de texto para fins documentais é não colocar em questão o papel do indivíduo no processo de produção/consumo e tradução do todo/conteúdo para o todo/conceito/palavra-chave. Ao mesmo tempo, é negar que a linguagem, quer do autor/produtor do texto/discurso, quer da linguagem documentária, são aquisições, construções e instrumentos elaborados respectivamente por uma sociedade e para um grupo social, que pode ser o mesmo do produtor ou não, refletindo, em ambos os casos, o pensamento e os interesses dominantes.

Apresentados estes pressupostos teóricos, eles terão de ser levados em consideração em todo o processo de AD, e implicarão nas seguintes interrogações: até que ponto a ideologia que permeia o discurso é reflexo de uma linguagem de classe ou grupo? Será a língua um meio de dominação e a linguagem documentária um instrumento escamoteador dessa realidade? Como identificar a ideologia do Estado e de seus aparelhos? Que processo de interação social se realiza quando a classe dominante impõe a incorporação de sistemas de signos "viciados" e como isso se reflete na AD? Em que grau a função prescritiva se impõe em detrimento da função informativa e quais as suas conseqüências para a documentação?

Na realidade, ao considerarmos estas questões, colocamos a hipótese de existir uma correspondência entre linguagem documentária

e estado social, entre linguagem documentária e objetivos de dominação e controle da informação por parte das classes dominantes. Isto é, a nossa hipótese coloca a AD e o seu produto como resultantes da ideologia do autor/produtor, do analista e do sistema em que se encontram inseridos.

3.2 PARÂMETROS EXPERIMENTAIS

As considerações que foram feitas anteriormente nos permitem analisar o discurso científico. Perante este objeto, nossa proposta é, num primeiro momento, observar como o autor/produtor passa das macro-proposições semânticas/TDs para a construção de um novo discurso e, num segundo momento, chegar a conceitos/palavras-chave capazes de traduzir esses Traços Descritivos em informações documentárias.

No primeiro momento estamos perante a construção e organização metodológica do discurso realizada pelo autor/produtor e a tentativa de identificação dessa organização metodológica por parte de quem realiza a AD. Estamos, portanto, com dois níveis de preocupações: o do autor/produtor e o do analista, o que nos leva a propor o seguinte esquema de trabalho:

- delimitação do "corpus" (do autor/produtor, do analista);
- definição dos objetivos da análise (do autor/produtor, do analista);
- identificação do método de análise (do autor/produtor, através da proposta da AD);
- desmontagem da construção/interpretação (do autor/produtor, realizada com base na proposta do analista);
- validação do resultado obtido (tanto pelo autor/produtor como pelo analista).

No segundo momento, a Análise Documentária visa chegar a conceitos/palavras-chave capazes de traduzir o conteúdo desse discurso em informações, exigindo para tal a segmentação do texto e a

selação dos conceitos/palavras-chave tradutoras desse conteúdo. Esta questão remete para a necessidade de estudo diacrônico/sincrônico dos conceitos/palavras-chave passíveis de serem fixados como tradutores do conteúdo e integrantes dum vocabulário controlado.

3.2.1 Metodologia de análise

Esta fase visa, primeiramente, identificar a organização metodológica do discurso do autor/produtor através da segmentação do texto e em seguida isolar conceitos/palavras-chave tradutoras do conteúdo desses segmentos.

Para isto, convém ter em consideração que o autor/produtor trabalha em função de um universo de referência. Isto é, ao trabalhar em função de um universo de referência e ao ter como objetivo a "ciência", formula as suas "teses", "teorias", "hipóteses" e "leis", apresentando-as e defendendo-as nas suas obras, por meio da construção "lógica" do discurso, das citações e argumentações.

Em AD efetuam-se, normalmente, cortes na obra do autor. Isto é, não se analisa a obra como um todo, mas partes consideradas significativas, sendo que o critério de significação é determinado em função dos sistemas (instituições, consumidores, custos/benefícios etc.) em que é realizada a análise.

Isto nos leva a afirmar que em documentação haveria uma correspondência entre "corpus" e material a analisar, sem contudo poder-se deixar de ter em consideração o "corpus maior" e o universo de referência/conhecimento da área.

Posto isso, interessa-nos identificar os objetivos da análise feita pelo autor/produtor com base nos dados/informações manipulados e nos resultados/conclusões por ele obtidos. Neste sentido,

teremos de considerar as propostas explícitas e implícitas (onde surgem as variáveis linguagem/ideologia/estado da ciência) identificadas pela análise do texto que, traduzidas em informação, implicarão um maior ou menor grau de acerto em relação ao conteúdo do texto/discurso a ser transmitido aos consumidores potenciais.

Passamos, então, para a identificação do método de análise do autor/produzidor através da proposta da AD. É o momento da segmentação do texto com o objetivo de "desmontar" a construção do discurso do autor/produzidor, cabendo à AD explicitar a passagem dos Traços Descritivos à sua interpretação e esclarecer as transformações dos dados de entrada ao longo do processo analítico, usando um procedimento complementar de descrição e ordenação dos traços. Este processo de descrição faz-se tanto na fase de segmentação do texto, ao isolar os Traços Descritivos portadores de informação "significativa", como na atribuição de palavras-chave identificadoras dessas unidades portadoras de informação. Pretende-se, então, que o conjunto dos Traços Descritivos preencham a função de léxico ou metalinguagem e sejam trabalhados no momento da seleção e fixação desses conceitos/palavras-chave tradutoras.

Levantados estes Traços Descritivos, torna-se necessário ordená-los. É a fase de ordenação, etapa em que o analista começa a desviar-se de uma visão descritiva externa para um visão interpretativa interna, baseada na observação de relações entre os termos da descrição, de recorrência, equivalência, oposições, paralelismos, simetrias, inversões etc. Relações que só poderão ser precisadas na medida em que se tiver visualizada a totalidade do "corpus" e selecionado conceitos/palavras-chave correspondentes aos Traços Descritivos.

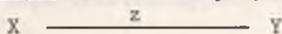
No momento da segmentação do texto e identificação dos Traços Descritivos, é necessário ter em conta o apoio de outras unidades de informação introduzidas pelo autor/produzidor, quer como apoio à

informação principal, quer como justificativa teórico/exemplificativa dessa informação primeira. Os argumentos são implícitos, quando refletem o saber geral comumente aceito, uma espécie de semântica universal, ou explícitos, quando têm origem no saber estabelecido cientificamente, uma semântica local, que se manifesta pela menção ou citação de autores e obras ou pelo saber adquirido pelo autor através da sua formação. Quer nos argumentos implícitos, quer nos explícitos, surgem Traços Descritivos que se apóiam numa espécie de intuição, hipóteses levantadas pelo autor e a serem verificadas pelo analista, que seriam na perspectiva de Greimas (GREIMAS, 1976), os antecedentes metodológicos do "discurso da descoberta".

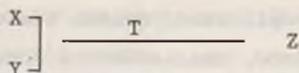
Na realidade, esta fase de identificação do método de análise, se bem que teoricamente linear e metodológica, em exposição, envolve problemas graves quanto à margem de subjetividade, principalmente no que se refere aos critérios de segmentação do texto, na medida em que é entregue ao analista a tarefa de encontrar a estrutura lógica/metodológica da construção do discurso, a partir da identificação das macro-proposições semânticas/TDs.

Identificados e ordenados esses traços, detectada a quantidade ou não de argumentos, o analista passa a observar como o autor/ produtor realizou a interpretação/construção de um novo discurso com base nos traços descritos apresentados. Na maior parte dos casos, toda a construção é gerada a partir de um jogo de seleções e transformações dos TDs previamente ordenados. A seleção dá-se no momento da descrição e da ordenação (na altura em que o autor/ produtor abandona ou desenvolve determinada informação) enquanto a transformação se realiza na construção através de dois mecanismos fundamentais: a reformulação e a assimilação (NATALI, 1976, pp. 32-35).

Esquematizando: a transformação por reformulação dá-se quando a um grupo de Traços Descritivos incorpora-se um dado externo ou novo, tendo como resultado uma construção, uma interpretação



A transformação por assimilação ocorre quando Traços Descritivos distintos se unem através de uma operação de síntese



Na base destas transformações encontram-se operações lógico-semânticas, que no momento da análise do discurso precisam ser identificadas, assim como os dados exteriores e novos que foram utilizados. Estes, também denominados argumentos, podem ser implícitos ou explícitos.

Surge, então, a fase de testar o resultado obtido, tanto pelo autor/produtor, como pelo analista da documentação. No que se refere ao autor/produtor, a validação do seu discurso remete tanto para a eficiência, lógica e solidez da teoria exposta (resultado possível de ser identificado no interior do texto através da análise documentária) como para a comprovação externa ao texto face às suas teorias.

Neste último caso, já não estamos apenas perante a análise de texto mas sim, frente ao estado da ciência o qual exige para sua compreensão o mínimo de conhecimento da área.

Para o analista da documentação, o teste de validação consiste numa simulação capaz de refazer o texto do autor/produtor, montando o "esqueleto" do texto/discurso analisado e construindo, em cima dele, um outro, portador do mesmo arcabouço simbólico.

3.2.2 Metodologia de síntese

Nesta fase, a análise documentária visa chegar a conceitos/palavras-chave capazes de traduzir o conteúdo do discurso analisado. Procedese, então, primeiramente, a uma seleção e depois, a uma fixação desses conceitos/palavras-chave.

Na seleção precisa-se de ter em conta dois níveis: o formal e o de conteúdo, sendo que no primeiro há que se considerar o sistema em que se encontra inserido o trabalho de análise e disseminação da informação, isto é, qual o tipo de instituição, qual o perfil (perfis) do(s) consumidor(es) dessa informação, qual a sua linguagem/ideologia de leitura/consumo. Destas variáveis formais resultarão condicionamentos específicos à seleção e fixação de conceitos/palavras-chave.

Com estas considerações, poderemos fixar os conceitos/palavras-chave capazes de traduzir o conteúdo do texto/discurso. Contudo, ao proceder a essa fixação é necessário definir e justificar a opção por determinados termos, explicitando a problemática envolvida na passagem de uma linguagem natural para um linguagem documentária, ao mesmo tempo que se torna necessário determinar o Campo Semântico e as categorias funcionais. Na realidade, são eles que vão permitir a estruturação e construção de instrumentos documentários, como tesouros e vocabulários controlados, possibilitando a organização e disseminação da informação por meio de uma leitura "única/universal".

BIBLIOGRAFIA

- CHOMSKY, N.A. (1970). *La gramatica generativa transformazionale*. Turin, Boringhieri.
- FILLMORE, C.J. (1971). Em favor do caso. In: *A semântica na lingüística moderna: O Léxico*. Rio de Janeiro, Livraria Francisco Alves, pp. 277-365.
- GARDIN, J.C. (1974). *Les analyses de discours*. Neuchâtel, Delachaux et Niestlé.
- GARDIN, J.C., LAGRANGE, M.S., MARTIN, J.M., MOLINO, J., NATALI, J. (1981). *La logique du plausible: essais d'epistemologie pratique*. Paris, Ed. Maison des Sciences de l'Homme.
- GREIMAS, A.J. (1976). *Semiótica do discurso científico: das modalidades*. São Paulo, DIFEL.

- NATALI, J.W. (1976). *De l'analyse documentaire à l'analyse poétique: a propos des critiques des "Chats" de Baudelaire.* Paris, École des Hautes Études en Sciences Sociales. Thèse de 3e. Cycle.
- POTTIER, B. (1974). *Linguistique générale: théorie et description.* Paris Klincksieck.
- SAUSSURE, F. de (1969). *Curso de lingüística geral.* São Paulo, Cultrix/EDUSP.
- VERON, E. (1978). *Sémiosis de l'idéologique et du pouvoir.* Communications, (28):7-20.

4. ESTRUTURAÇÃO DE VOCABULÁRIO

Isabel M. R. Ferin Cunha

4.1 FORMULAÇÃO DAS CATEGORIAS OPERACIONAIS

4.2 FORMAÇÃO DO CAMPO SEMÂNTICO

4.3 DISCUSSÃO DOS CONCEITOS

4.4 CONTROLE DE VOCABULÁRIOS: OS DIAGRAMAS

4.5 CONCLUSÃO

BIBLIOGRAFIA



4. ESTRUTURAÇÃO DE VOCABULÁRIO

*Isabel M. R. Ferin Cunha**

Em documentação a necessidade de um vocabulário controlado em determinado assunto, condição de acesso à informação, leva primeiramente ao levantamento dos "termos significativos" e à sua estruturação. Quando esta operação se realiza de forma manual, o levantamento destes termos faz-se com base no "bom-senso" bibliotecário, o qual tem implícito tanto o conhecimento da área específica da documentação como o perfil do consumidor. Quanto à estruturação do vocabulário, o "bom-senso" exige que se tenham em consideração algumas regras básicas articuladas em torno de duas vertentes teóricas: a Lógica, contemplando relações de todo e parte, gênero e espécie, inclusão/exclusão, associação/oposição e a Lingüística, ocupando-se da formação do Campo Semântico, relações paradigmáticas e sintagmáticas, polissemia, sinonímia, homonímia etc.

Quando se procede de forma automática, nomeadamente com base em programas de frequência/ocorrência (tipo KWIC e KWOC), o levantamento surge-nos como uma listagem a ser triada e estruturada em função dos termos significativos, sendo que a identificação destes depende igualmente do conhecimento da área e do perfil do consumidor. A tentativa de indexação automática realizada pelo PRECIS (*PREserves Context Index System*), fundamentada na análise sintática e semântica da língua e prevendo uma série de operadores de função para caracterizar a oposição e o significado dos

* O presente capítulo constitui um extrato da tese de doutoramento "Do mito à análise documentária: a luso-assimilação", defendida na ECA no 1º semestre de 1987.



termos no contexto de um cabeçalho de assunto, anula o texto, o qual nem sempre se reflete no cabeçalho, principalmente quando se trabalha com textos de Ciências Humanas, ao mesmo tempo que descarta a discussão acerca dos "sentidos" dos termos, atendendo-se preferencialmente ao plano denotativo (LYONS, 1977).

Na realidade, existe uma dificuldade básica em relação aos procedimentos automáticos em documentação, principalmente quando aplicados a textos de Ciências Humanas, na medida em que os programas apenas consideram o plano denotativo (e mesmo a este nível se restringem à sintaxe), uma vez que o "sentido profundo do texto" ou do enunciado se encontra ao nível conotativo/conceptual (LYONS, 1977). Esta dificuldade ocorre no momento em que o receptor da mensagem realiza uma redução seletiva das informações emitidas, sendo que estas se encontram ligadas tanto a fatores sócio-culturais como a situações etnolinguísticas, determinantes, por sua vez, da linguagem/ideologia.

Ao trabalharmos um texto/discurso/documento, temos no plano denotativo, no que diz respeito ao leitor/receptor, um processo que se inicia pela identificação da organização do significante, através das estruturas sintáticas e dos sistemas semânticos, sendo que a conceptualização/conotação se daria na fase final de identificação do texto/discurso/documento, diretamente dependente do conhecimento da área, da representatividade atribuída ao autor/produtor do texto/discurso/documento em relação a esse mesmo momento, e ao posicionamento do leitor/receptor quanto ao conteúdo transmitido.

Quanto ao autor/produtor o processo é inverso na medida em que parte de uma conceptualização/conotação individual (conhecimento do assunto, seu ponto de vista, posicionamento ideológico) da organização do discurso, tendo em vista a sintaxe-semântica e a organização dos significantes.

No processo de análise documentária a proposta é reduzir a clivagem entre estes dois procedimentos, tentando reconstruir a conotação/conceptualização feita pelo autor/ produtor e atribuir-lhe conceitos/palavras-chave tradutoras da informação conceptualização/conotada.

Neste sentido é necessário ter em conta as várias leituras possíveis do texto/discurso/documento, captando as hipóteses conceptuais/conotativas no levantamento dos termos "significativos" possíveis de integrar o vocabulário controlado.

Essas várias leituras possíveis remetem-nos para a necessidade de se ter ou criar esquemas de análise de texto/discurso/documento embasados na identificação das macro-proposições semânticas/ Traços Descritivos, os quais permitirão localizar no texto/discurso/documento os segmentos portadores de "maior informação" e subseqüentemente os termos "significativos" do vocabulário de base, capazes de integrar o vocabulário controlado.

Posto isto, a exigência de selecionar e controlar o vocabulário de base e determinar as relações lógico-lingüísticas e semânticas existentes torna necessário logo à partida a construção de blocos aglutinadores que funcionem como categorias representativas dos diferentes aspectos conceptuais/conotativos da formulação do texto.

Esta questão pressupõe por um lado, a consideração de paradigmas (SAUSSURE, 1969), dado que determinam relações de elementos que podem ser colocados no mesmo ponto de uma cadeia, sendo substituíveis ou comutáveis entre si; pressupõe, por outro lado, o conjunto de relações sintagmáticas (SAUSSURE, 1969) possíveis, na medida em que conjugam unidades consecutivas onde o valor de cada uma se define por relação ao valor da outra.

Exemplo: "O usuário da biblioteca pública lê jornais diários"

Eixo paradigmático ou das
escolhas possíveis

Eixo sintagmático ou das
escolhas realizadas

associações

do autor/
produtor

{ USUÁRIO
INFORMAÇÃO

usuário da biblioteca
pública...

Na medida em que as relações paradigmáticas determinam associações abstratas possíveis, elas tornam-se capazes de definir categorias aglutinadoras de atributos, enquanto que as relações sintagmáticas tendem a pulverizar-se através das escolhas realizadas, particularizando as associações e contribuindo para a divisão das categorias em novos blocos de menor capacidade aglutinadora a que poderemos chamar de classes ou subclasses.

Neste sentido, tornou-se necessária a definição formal de conceitos determinantes de categorias que pudessem ter a função operacional de aglutinar palavras portadoras do mesmo sentido/função no texto, atribuídas pelo autor quando da sua elaboração conceptual.

4.1 FORMULAÇÃO DAS CATEGORIAS OPERACIONAIS

Com este pressuposto tomou-se como hipótese de trabalho os vários Estudos de Casos realizados pela lingüística, nomeadamente, os de Fillmore e Pottier.

O embasamento nos Estudos de Casos para o tratamento da informação tem sido proposto na estruturação de vocabulários realizados de forma manual, caso das sugestões apresentadas por Anna Maria Cintra em "Elementos de Lingüística para estudos de indexação" (1983, pp. 5-22) como na proposta automática do PRECIS (AUSTIN, 1974). No entanto, enquanto Anna Maria Cintra apresenta a contribuição lingüística e aponta caminhos para adequação a ser reali-

zada, no manual do PRECIS e no trabalho de adaptação ao português realizado na dissertação de Mariângela Spotti L. Fujita (1985), os casos surgem como dados estabelecidos sem se fazer referência à passagem do teórico-lingüístico ao pragmático-documentação.

Dada esta questão, trabalhamos os textos de Fillmore e Pottier, comparando o desenvolvimento teórico realizado pelos dois autores de forma a identificar as propostas que mais interessariam ao tratamento da informação. Com este objetivo recuperamos de Fillmore *Em favor do caso* (FILLMORE, 1977) a especulação acerca da existência e possibilidade de determinação de casos numa língua e de Pottier *Linguistique Générale: théorie et description* (POTTIER, 1974) os exercícios de identificação dos casos conceptuais. Isto tendo em conta o fato de que para nós, não interessa a determinação dos casos numa língua, mas sim a possibilidade de organizar os termos "significativos", o vocabulário base levantado em categorias formais que possibilitem a sua estruturação sem referência a uma língua particular. Assim, ao considerarmos os Estudos de Casos atendemos, por um lado, à estrutura formal por eles apresentada e por outro, aos nossos objetivos específicos de organização e estruturação de um vocabulário previamente definido.

Ao atendermos à estrutura formal dos Estudos de Casos lingüísticos, afirmamos com Fillmore (1977, p. 128) que existem "muitos relacionamentos sintáticos semânticos relevantes envolvendo substantivos e as estruturas que os contêm, que esses relacionamentos são em grande parte, não manifestos mas, assim mesmo possíveis de serem descobertos empiricamente, que formam um conjunto finito específico...".

Esta afirmação salvaguarda-nos, na medida em que ao pretendermos organizar e estruturar o nosso vocabulário, estamos tanto trabalhando o texto ao nível conceptual/conotativo, como tentando construir uma sintaxe/semântica própria ao nosso vocabulário, condições necessárias para atingir o nosso objetivo de recuperação

da informação.

Neste sentido, a afirmação de Fillmore de que "As noções de casos incluem um conjunto de conceitos universais, presumivelmente inatos que identificam certos tipos de julgamentos que os seres humanos são capazes de fazer acerca dos assuntos tais como quem fez, com quem aconteceu, e o que foi mudado? (FILLMORE, 1977, p. 299) leva-nos a propor uma leitura do texto em simultâneo com os termos "significativos" levantados pelo vocabulário de base, na tentativa de determinar os grandes blocos capazes de operacionalizar as nossas perguntas em relação ao texto, tendo como objetivo a recuperação da informação. Essas perguntas fundamentariam com grande margem de segurança os blocos/categorias organizativas da informação.

Esta questão leva-nos a repensar a proposta de Pottier de 11 casos conceptuais capazes de determinar o encadeamento lógico do enunciado. Tomando como exemplo o enunciado CONTRIBUIÇÃO DOS BIBLIOTECÁRIOS PARA A VALORIZAÇÃO DO HOMEM NO BRASIL aplicamos os casos conceptuais propostos por Pottier, resultando no seguinte:

Agentivo - Bibliotecários

Locativo - Brasil

Causativo - Valorização do Homem

O Causativo, "valorização do Homem", pressupõe um Instrumento de valorização, incluindo na valorização, o Benefício para alguém (os brasileiros, presumivelmente!?) e o alguém, Dativo (o Homem enquanto abstração). Ao mesmo tempo, a "contribuição" pressupõe uma relação de poder entre quem contribui e quem recebe a contribuição, determina a necessidade de uma Finalidade que pode ser identificada como Causativo, "valorização do Homem".

Teríamos, então, como hipótese de trabalho alguns casos formulados neste exemplo: Agentivo, Locativo, Causativo e os que tomamos

como implícitos, o **Instrumental**, **Beneficiativo** (Benefactif), **Dativo** e **Finalidade**.

Esta hipótese da organização de vocabulário decalcada nos casos conceptuais de Pottier traz-nos problemas relativamente à organização da informação.

Por exemplo, em relação ao agente que pratica ou sofre a ação, Pottier prevê a sua distribuição pelos casos Ergativo, Nominativo, Agentivo, Acusativo, Associativo, Dativo e Beneficiativo. Esta pulverização do Humano-agente dificulta a formação de blocos/ categorias aglutinadoras de palavras, na medida em que a pergunta que nos convém fazer ao vocabulário levantado é: quem pratica ou sofre a ação? A pulverização da resposta à pergunta em vários casos impedirá a organização do vocabulário e a posterior recuperação da informação. Na verdade, não nos interessa num primeiro momento distinguir quem pratica ou sofre a ação (casos contemplados no Agentivo, Dativo e Beneficiativo), mas sim a formação de uma categoria geral organizada em torno do atributo "Humano", possível de se dividir em classes diferenciadas pela intensidade de participação na ação: ativo, receptivo, passivo etc.

Por outro lado o Causativo tem tendência a confundir-se com o Instrumental, na medida em que se organiza em torno do "Não-Humano", sendo que a possibilidade de diferenciação se dá no grau de aproximação (Causativo) ou distanciamento (Instrumental) com o Humano.

Por exemplo: **CONTRIBUIÇÃO DOS BIBLIOTECÁRIOS PARA A VALORIZAÇÃO DO HOMEM NO BRASIL** pela informação

"A Valorização do Homem", Causativo, encontra-se ligada ao Agentivo, "bibliotecários", contendo mais elementos de Humano do que "informação" (Instrumental), apesar de ser "Não-Humano" mas produzido e concretizado pelo "Humano".

Na organização do nosso vocabulário, os dois casos, Causativo e Instrumental poderiam apresentar-se reunidos em torno da pergunta "quais os meios concretos que permitiram a ação consumir-se?", formando um bloco/categoria formal de aglutinação de palavras rotulado de Instrumento(s), caracterizado pelo Não-Humano, Não-Animado, absorvendo a margem fluida do que é criação do Homem e existe per si.

O Locativo que nos aparece no exemplo anterior em "Brasil" é aplicado por Pottier a uma das três zonas de experiência: espaço, tempo, situação (notion). O uso deste caso como categoria aglutinadora do vocabulário levantado interessa-nos, tendo em vista o conhecimento da área e a capacidade de resposta à pergunta básica: "aonde e quando se passa a ação?".

O Acusativo que não surgiu no exemplo dado apresenta-se problemático por derivar de uma relação ativa entre uma ação e uma não ação determinada pelo verbo. Na realidade, pode-se pressupor como sendo um modo de relação entre a ação e a não ação e neste sentido adequar-se-á à organização do nosso vocabulário através de duas perguntas básicas determinantes de dois blocos aglutinadores de palavras. A primeira pergunta seria: "qual o objeto que sofre e/ou motiva a ação?"; a segunda pergunta: "como se realiza a ação?".

Teríamos, portanto, decalcados no Acusativo dois blocos/categorias possíveis de serem rotulados, respectivamente como: Objeto e Modo.

Em relação aos casos conceptuais Dativo, Beneficiativo e Finalidade, eliminado o fator Humano absorvido pelo Agente, teríamos a anulação do Dativo, na medida em que prevê "o caso do ser animado afetado pela ação identificados pelo verbo" (FILLMORE, 1977, p. 300) e a consideração em relação ao Beneficiativo do resultado desse benefício.

O Beneficiativo não estaria (implicitamente) nos "brasileiros" mas sim no produto desse benefício: "Valorização do Homem". Em virtude de termos proposto a absorção do caso Causativo pelo Instrumental, tendo como finalidade a formulação das nossas categorias organizativas formais, o Beneficiativo distingue-se do Causativo ao realizarmos a seguinte pergunta: "qual o benefício/produto/resultado da ação?". Isto nos dá margem para um novo bloco aglutinador do nosso vocabulário em torno do rótulo Produto.

Da mesma forma, o caso conceptual Finalidade adequar-se-á ao nosso vocabulário ao pressupor a pergunta: "qual a intenção/fim da ação?" Distingue-se do bloco aglutinador Produto por exigir projeto/intenção/virtualização e não somente realização/competência/"performance".

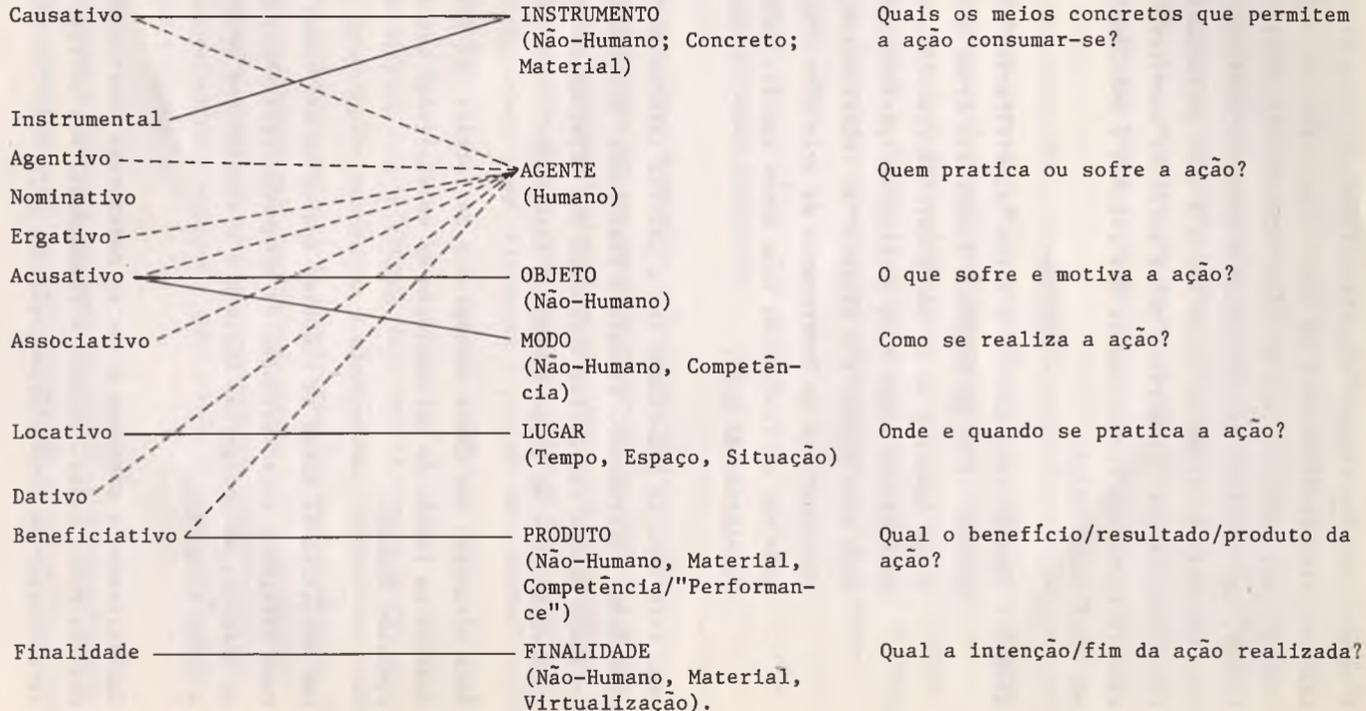
Quanto ao caso Associativo, ao aglutinar o fator Humano no Agente, restar-nos-ia o conceito de associação, o qual se apresenta na documentação diretamente dependente das relações paradigmáticas (no sentido das escolhas possíveis) e das relações sintagmáticas (no sentido das escolhas realizadas) sendo, portanto, inadequado formular um bloco aglutinador com base neste caso conceptual.

Esquemmatizando os casos conceptuais de Pottier e a adequação realizada teríamos:

CASOS CONCEPTUAIS DE POTTIER

ADEQUAÇÃO REALIZADA

PERGUNTAS A FORMULAR



A organização dos termos "significativos" em categorias/blocos aglutinadores realizar-se-á em função das perguntas sugeridas. Trabalha-se, portanto, não ao nível sintático/semântico dos textos, mas ao nível dos conceitos e suas funções/ações dentro da área em análise, tendo em conta tanto o eixo paradigmático (determinante das categorias, pelas escolhas possíveis), como o eixo sintagmático (determinante das classes e subclasses, pelas escolhas realizadas).

Exemplo: Tomando isoladamente o termo "significativo" INFORMAÇÃO, podemos lê-lo da seguinte forma: ele exige a existência de um Agente e de um receptor/ativo/passivo, dá-se em determinado Lugar e em determinado período, pode ser um Modo de o agente e o receptor se relacionarem, pode funcionar com um Instrumento de ação para o agente/ativo, pode ser um Produto da ação desse agente, pode ser a Finalidade da ação.

Dado este exemplo poderíamos ter a palavra INFORMAÇÃO figurando como Modo, Instrumento, Produto e Finalidade. A inclusão ou exclusão dessas categorias dependerá das relações paradigmáticas estabelecidas e do Campo Semântico integrador.

Esta adequação dos Casos Conceptuais de Pottier foi feita e testada em função da estruturação de vocabulários nas áreas de Ciências Humanas.

Uma adequação às áreas de Ciências Exatas ou Aplicadas resultará numa diferente caracterização da adequação realizada, mantendo-se, no entanto, as categorias básicas e o sentido das perguntas a serem formuladas.

Explicitando: a proposta é que as categorias Instrumento, Agente, Objeto, Modo, Lugar, Produto e Finalidade, são categorias de estruturação de vocabulários, válidas para todas as áreas do

Conhecimento. Portanto, Categorias Universais, sendo que as diferenças das áreas refletir-se-iam nas caracterizações das categorias e na formulação das perguntas que, no entanto, mantêm o mesmo sentido. Isto é, o Agente em Química, um reagente, não se caracterizaria pelo atributo Humano como na área de Ciências Humanas, mas sim pelo fato de provocar uma ação - o que muda a caracterização mas não a pergunta.

Esta proposta pressupõe o levantamento exaustivo e atualizado dos vocabulários de base das diferentes áreas do Conhecimento e posteriores estruturações com base nas categorias propostas, de forma a ser possível determinar tanto o grau de universalização dessas categorias e eventuais desvios que caracterizarão necessariamente as diferentes áreas, como identificar as tipologias dessas mesmas áreas.

4.2 FORMAÇÃO DO CAMPO SEMÂNTICO

Entendemos por Campo Semântico o "conjunto de classificações correlatas que associam um grupo de palavras ou ainda a possibilidade que as unidades lexicais de uma língua têm de se reunir em grupos estruturados de tal modo que cada unidade fica definida pelo lugar que ocupa respectivamente à posição das demais" (LOPES, 1976, p. 242).

O ponto de partida para a estruturação destes campos faz-se a partir de associações mnemônicas e aleatórias. Independentemente da escolha do léxico inicial de construção do Campo Semântico, as associações realizadas em cima dele dependem tanto de fatores sócio-culturais, como dos elementos envolvidos - destinatário/receptor da mensagem/enunciado - quer ainda da finalidade das associações realizadas.

Em documentação a formação do Campo Semântico determina a possibilidade de organizar a informação, na medida em que é em relação

a ele que se podem estruturar as categorias, agrupar conceitos e estabelecer relações lógico-lingüísticas entre eles. Contudo ao recorrer à lingüística para a determinação do Campo Semântico, a documentação depara-se com alguns problemas de adequação.

Na verdade, a noção de palavra-ponto-de-partida para a associação é mais intuitiva e etnocêntrica do que científica. Por outro lado, ao trabalharmos com documentação fazemo-lo com textos, os quais sendo compostos por palavras não têm o sentido/valor da soma linear do conjunto das palavras, o que quer dizer que o sentido do todo é diferente da soma linear dos sentidos das suas partes. Além disso, numa língua dada as palavras não possuem apenas um sentido/leitura; elas têm mais de um sentido, determinado tanto pelas relações sintagmáticas de um contexto dado, como pela competência/"performance" do autor-receptor. Esta questão leva-nos a concordar num primeiro momento com a afirmação de Edward Lopes (LOPES, 1976, p. 282): "Quanto aos campos semânticos, eles não passam de ser, na maioria das vezes, um conjunto de termos reunidos numa área aprioristicamente delimitada, não a partir da rede de relações inter-sígnicas de um sistema lingüístico, mas a partir da concepção sociológica do 'campo'". Afirmação que parece eliminar 'a priori' a possibilidade de um procedimento "universal" para a formação do Campo Semântico, requisito fundamental à Biblioteconomia/Documentação para a estruturação da informação de qualquer área do conhecimento.

Como resolver, então, o problema da construção do Campo Semântico em documentação, sendo que ainda a maioria dos lingüistas trabalha a construção e estruturação do Campo Semântico com base em termos concretos (objetos, coisas, animais) como por exemplo embarcações, animais domésticos, assento, relações de parentesco etc., e não em conceitos abstratos?

A nossa resposta a estas dificuldades centra-se em dois momentos de trabalho a ser realizado pela Biblioteconomia/Documentação.

Primeiramente na necessidade de se levantar vocabulários nas áreas e subáreas de Conhecimento existentes, determinando as definições, usos e funções do vocabulário de base dessas áreas, o que implica não só uma formação especializada dos bibliotecários como trabalhos interdisciplinares de equipe. Simultaneamente a este trabalho impõe-se a identificação das tipologias dos discursos das áreas e subáreas de Conhecimento com que se trabalha, de forma a ser possível rastrear os percursos de produção e consumo de informações e subseqüentemente montar os conjuntos de relações paradigmáticas e sintagmáticas de uso, tanto por parte dos produtores como dos consumidores da informação.

Só com este longo trabalho de equipe interdisciplinar nos parece possível trabalhar a questão do Campo Semântico de uma forma profunda e "universalizante". Ao contrário, resta-nos trabalhos paliativos, se bem que necessários no conjunto da proposta anteriormente exposta, de tentar organizar os Campos Semânticos através da atribuição de traços específicos e traços genéricos, agrupando os vocábulos com base em denominadores comuns e em relação aos casos/funções já referidos (questão a ser colocada no Capítulo 5)

Dadas estas questões teremos que admitir que para além da especialização do bibliotecário, do conhecimento da área e da concepção sociológica do Campo Semântico - que implica autor/receptor, condições de produção e consumo do texto - são os procedimentos lógico-matemático-linguísticos sistematizados que permitirão determinar o Campo Semântico operacional em Biblioteca/Documentação.

4.3 DISCUSSÃO DOS CONCEITOS

Na fixação dos conceitos em documentação com vista à recuperação da informação, o "bom-senso" bibliotecário determina que se tenha

em consideração a especificidade dos descritores, isto é, que eles possam traduzir adequadamente a informação contida no documento analisado. No caso de ambigüidade em relação ao conceito-descriptor está previsto o uso de notas de alcance com a função de precisar, limitando e determinando, o uso desse conceito-descriptor no instrumento que se insere, seja ele um tesauro ou um vocabulário controlado.

As notas de alcance pretendem conferir uma leitura "única" ao conceito-descriptor, eliminando as sinônimas e polissemias possíveis, enquadrando o conceito na hierarquia do tesauro/vocabulário controlado, tornando os conceitos e os instrumentos operacionais.

Contudo, ao definir e limitar os conceitos/descriptores por notas de alcance, o "bom-senso" bibliotecário procede a um raciocínio que envolve a necessidade de determinar os diferentes usos, ou os usos sucessivos do conceito, as suas formas de coexistência, além de proceder à identificação dos procedimentos de intervenção realizados em diferentes níveis e situações com o fim de os sistematizar (FOUCAULT, 1971).

Exemplificando. No caso do conceito INFORMAÇÃO estamos perante a necessidade de determinar os diferentes usos e acepções deste conceito nas várias áreas que o utilizam, o que resultaria no seguinte esquema:

	INFORMAÇÃO	-	área 1	-	conceito 1
	"	"	"	-	conceito 1.1
	"	"	"	-	conceito 1.2
	"	"	"	-	-----

* Diacronia, eixo da sucessividade. Os conceitos surgem como imutáveis. Contudo, eles alteram-se em função de fatores culturais, científicos, políticos e econômicos internos e externos, determinados pelo eixo da simultaneidade.

(Na realidade, as áreas não têm consciência do processo de sobreposição. A Biblioteconomia/ Documentação, dado que trabalha com todas as áreas e todos os conceitos de INFORMAÇÃO, precisa conscientizar-se deste processo para melhor resolver os problemas dele derivado).

INFORMAÇÃO

—————→ Sincronia. Eixo da simultaneidade.

Conceito 1 = definição de INFORMAÇÃO +
variáveis de área 1

Conceito 2 = Conceito 1 + variáveis de área
2

Conceito 3 = Conceito 1 + Conceito 2 +
variáveis de área 3

O que nos permite deduzir que o conceito INFORMAÇÃO é utilizado numa determinada área (área 1) em diferentes acepções resultantes de fatores culturais, científicos, políticos e econômicos (conceito 1.1, 1.2, e assim sucessivamente). Essas interferências podem resultar na apropriação desse Conceito 1, da área 1, por outra área (área 2, por exemplo), dando origem a um novo conceito (conceito 2) de INFORMAÇÃO, que sendo "novo" é simultâneo em uso ao Conceito 1 da área 1 e assim sucessivamente.

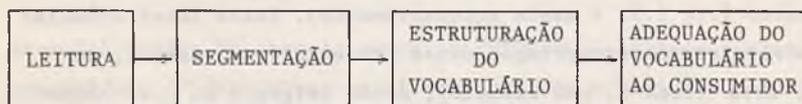
Esta questão leva-nos a admitir a necessidade de um procedimento circular de estruturação de vocabulários, na medida em que, partindo de um vocabulário de base, se passaria pela sua organização em categorias e Campo Semântico, obtendo-se um vocabulário estruturado, sendo que este necessitaria, para se adequar aos diferentes consumidores, sofrer novo refinamento, o qual exigiria uma volta às instituições de produção, autor/produtor (portanto, ao texto/discurso/documento inicial) e conseqüentemente às definições apresentadas no vocabulário de base.

4.4 CONTROLE DE VOCABULÁRIOS: OS DIAGRAMAS

Colocamos até este momento o problema da estruturação do vocabulário com vista à disseminação da Informação. No entanto, dado o enfoque proposto de ter-se em consideração os esquemas de produção e consumo dessa Informação, já não estamos, apenas, ao nível da estruturação, mas sim do controle de vocabulários. Este, implica a formação de tipologias dos discursos/documentos e dos vocabulários capazes de identificar tanto as áreas de produção como os seus diversos consumidores.

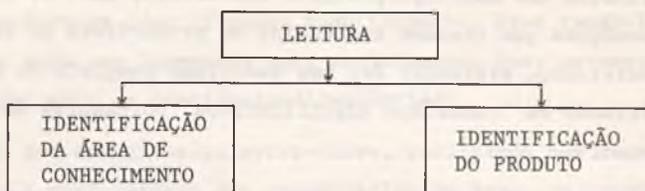
Para melhor visualizar-se estas propostas, apresentamo-las em diagramas que pretendem representar, desde o procedimento geral efetuado no controle do vocabulário (Diagrama nº 1), até a adequação aos vários consumidores possíveis (Diagrama nº 5), fechando o percurso tanto da estruturação como do controle da Informação através de vocabulários.

Diagrama nº 1 - Elaboração de vocabulários controlados -
Diagrama de Precedências



A elaboração de vocabulários controlados passa em primeiro lugar, pela Leitura do texto; seguidamente, pela Segmentação que tem como procedimento relevar as "informações significativas". Com base nestas, procede-se ao levantamento do vocabulário de base e sua transformação em conceitos/palavras-chave/descriptores e à sua Estruturação. Para finalizar, realiza-se a Adequação desse vocabulário-informação ao consumidor potencial.

Diagrama nº 2 - Elaboração de vocabulários controlados: Leitura -
Diagrama Hierárquico Funcional



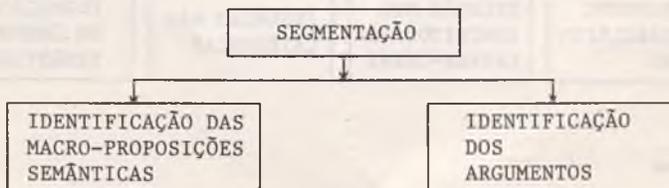
A Leitura, primeiro procedimento para elaboração de vocabulários controlados, compreende duas etapas:

- a identificação da área de Conhecimento em análise e suas características;
- identificação do autor/produtor e suas características.

A caracterização das áreas e seus autores/produtores apresenta-se como condição necessária à formulação de tipologias de discursos e de vocabulários, específicos às diversas áreas e subáreas. Essas caracterizações constituem, também, etapas que permitem reconstruir o fluxo de produção-consumo da Informação.

Diagrama nº 3 - Elaboração de vocabulários controlados:
Segmentação

Diagrama Hierárquico Funcional



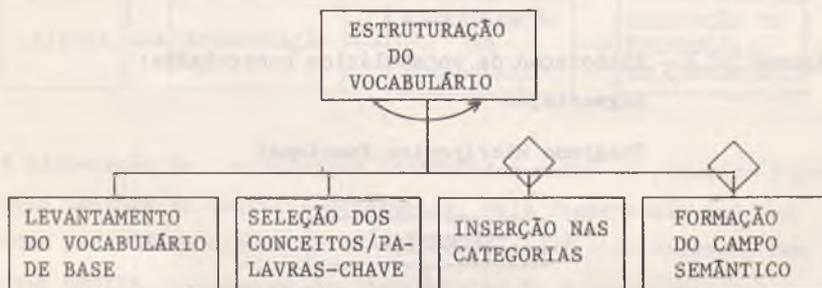
A Segmentação, segundo procedimento para elaboração de vocabulário controlados, compreende duas etapas:

- identificação das macro-proposições semânticas, isto é, seleção de proposições que resumem seqüências de proposições do texto, sem, entretanto, pretender dar uma descrição completa do mesmo, mas relevando os "conteúdos significativos" portadores de informação;
- identificação dos argumentos e códigos utilizados em função do "Estado da Ciência".

Estas duas etapas envolvem procedimentos lógico-linguísticos, os quais visam tanto a identificação do processo de produção do autor/produtor (pressupostos iniciais relativos ao "Estado da Ciência", Organizações Lógico-Semânticas e epistemológicas) como o processo de construção do analista da documentação - o seu discurso/leitura em cima do texto/documento.

Diagrama nº 4 - Elaboração de vocabulários controlados: Estruturação do vocabulário

Diagrama Hierárquico Funcional



Símbolos

Processo circular

Condição

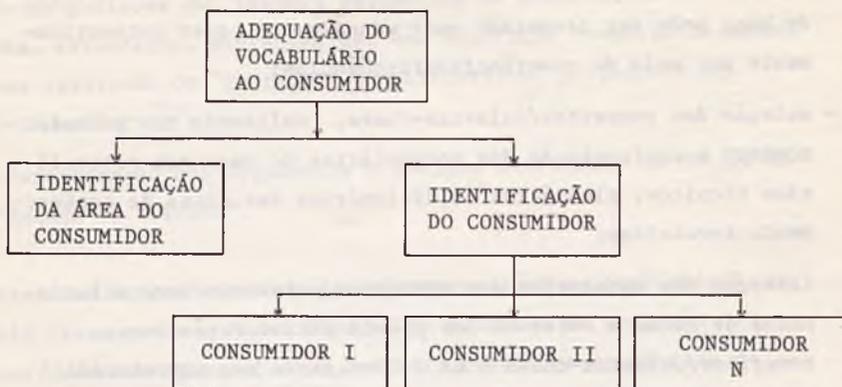
A Estruturação do vocabulário, terceiro procedimento para a elaboração de vocabulários controlados, compreende quatro etapas:

- levantamento do vocabulário de base contido nas macro-proposições semânticas identificadas anteriormente. Esse vocabulário de base pode ser levantado quer manualmente, quer automaticamente por meio de ocorrências/freqüências;
- seleção dos conceitos/palavras-chave, realizando num primeiro momento a confrontação dos vocabulários de base com vocabulários técnicos, glossários ou dicionários das áreas de conhecimento envolvidas;
- inserção nas categorias dos conceitos/palavras-chave selecionados de forma a obter-se uma prévia estruturação desses conceitos/palavras-chave e da informação a ser estruturada;
- formação do Campo Semântico com base nas categorias e nos conceitos previamente estruturados.

Esta última etapa de Estruturação do Vocabulário é, na realidade, o início de um processo circular que volta de novo ao vocabulário de base, à seleção dos conceitos/palavras-chave e à inserção nas categorias, tendo em vista um refinamento do produto-informação. Por outro lado, a inserção em categorias dos conceitos/palavras-chave selecionados é condição para a formação do Campo Semântico e este torna-se, por sua vez, condições para o refinamento da estruturação do vocabulário através da inserção em categorias.

Diagrama nº 5 - Elaboração de vocabulários controlados: Adequação do vocabulário ao Consumidor

Diagrama Hierárquico Funcional



A Adequação do Vocabulário ao Consumidor, quarto procedimento para elaboração de vocabulários controlados, compreende duas etapas:

- identificação da área do consumidor e suas características;
- identificação do consumidor e suas características.

A caracterização das áreas dos consumidores e dos próprios consumidores apresenta-se como condição necessária à formulação de tipologias de discursos e de vocabulários específicos às diversas áreas e subáreas, assim como à identificação de tipos-padrões de consumidores potenciais. Tanto as caracterizações das áreas e subáreas dos consumidores como a identificação de tipos-padrões de consumidores contribuem para a construção do circuito de produção-consumo da Informação.

4.5 CONCLUSÃO

A apresentação destas propostas para a estruturação e controle de vocabulário e portanto, da Informação, pressupõe algumas conside-

rações de âmbito geral e específico.

No que concerne às primeiras - as de âmbito geral - a questão pode-se equacionar da seguinte forma: a atividade da Biblioteconomia/Documentação insere-se numa Sociedade Industrializada, que se articula em torno de esquemas de produção e consumo; qualquer atividade que nela seja exercida tem de obedecer a esse condicionamento, caso contrário tenderá à extinção.

Ao inserirmos a Biblioteconomia/Documentação na Sociedade Industrializada, a Informação torna-se insumo e a tecnologia instrumento imprescindível, o que não anula a "função-social" da Biblioteconomia/Documentação, mas retira-a da perspectiva "assistencialista", colocando-a na da cidadania e obrigando-a a rever seus parâmetros teóricos, filosóficos e metodológicos.

É nesta revisão de parâmetros teóricos - que terá de visar fundamentalmente uma tecnologia de produção - que se encontram os maiores problemas da Biblioteconomia/Documentação, na medida em que, podendo-se utilizar o já existente (a Informática, por exemplo), tornam-se necessárias adaptações e construções próprias à sua atividade.

Na realidade, estamos perante a exigência de criar Sistemas Inteligentes, através da Inteligência Artificial, que possam desenvolver sistemas de representação de conhecimentos e de raciocínios, assim como, de sistemas de compreensão da linguagem natural. Estes sistemas não fazem o trabalho humano, nem lêem textos e dão a Informação por si, mas permitem transpor nossos conhecimentos em dados que a máquina registra, estoca e estabelece relações frente a um problema específico colocado (CATTENAT, 1984).

Neste sentido, apresentamos nossa proposta de métodos de análise de texto e de construção de vocabulários, de forma a sistematizar

procedimentos lógico-lingüísticos realizados no processo de obtenção da Informação. Neste sentido, também, nossa insistência na necessidade de montar os esquemas de produção/consumo de Informação, de forma a ser possível estabelecer o(s) fluxo(s) da mesma e, portanto, saber quem produz, como o faz e para quem, identificando as construções lógicas do produtor e do consumidor, tornando possível tipologias de produção e consumo (GARDIN, 1985).

Restam-nos ainda algumas considerações de âmbito específico acerca das propostas apresentadas para a estruturação e controle de vocabulários e que se encontram implícitas no que foi, anteriormente, dito. Trabalhos deste tipo são de equipe e exigem uma racionalização de recursos humanos e materiais a nível regional ou nacional; a formação do profissional bibliotecário/documentalista deve ser revista ou tenderá à extinção.

BIBLIOGRAFIA

- CATTENAT, A. e PAUL, G. (1984). *Intelligence Artificielle*. In: *Les nouvelles technologies dans l'Information Scientifique et Technique*. Paris, IRNA/SOPHIS-Antipolis.
- CINTRA, A.M.M. (1983). Elementos de lingüística para estudos de indexação. *Ci.Inf.*, Brasília, 12 (1):5-22.
- FILLMORE, C.J. (1977). Em favor do caso. In: *A semântica na lingüística moderna: O Léxico*. Rio de Janeiro, Livraria Francisco Alves, pp. 277-365.
- FOUCAULT, M. (1971). *A Arqueologia do saber*. Petrópolis, Ed. Vozes.
- FUJITA, M.A.S.L. (1985). *PRECIS na língua portuguesa: uma experiência com a indexação de artigos de periódicos brasileiros*. São Paulo, ECA/USP. Dissertação de Mestrado.
- GARDIN, J.C. (1985). *Informática e Arqueologia*. Lisboa, INIC.
- LOPES, E. (1976). *Fundamentos da lingüística contemporânea*. São Paulo, Editora Cultrix.
- LYONS, J. (1977). *Semântica I*. Lisboa, Ed. Presença/Martins Fontes.

POTTIER, B. (1974). *Linguistique générale: théorie et description*.

Paris, Klincksieck.

SAUSSURE, F. (1969). *Curso de lingüística geral*. São Paulo,

Cultrix/EDUSP.

IV. A DEFINIÇÃO SEMIÓTICA PARA A ELABORAÇÃO DE
SISTEMAS

1.1. O CONCEITO DE SEMIÓTIKA

1.2. O CONCEITO DE SEMIÓTIKA

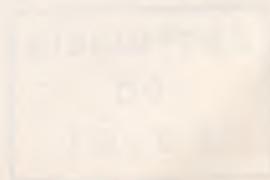
1.3. O CONCEITO DE SEMIÓTIKA

1.4. O CONCEITO DE SEMIÓTIKA

1.5. O CONCEITO DE SEMIÓTIKA

1.6. O CONCEITO DE SEMIÓTIKA

1.7. O CONCEITO DE SEMIÓTIKA



5. A DEFINIÇÃO SEMÂNTICA PARA A ELABORAÇÃO DE GLOSSÁRIOS

Maria de Fátima G. M. Tálamo

5.1 FORMULAÇÃO DE DEFINIÇÕES

5.2 A DEFINIÇÃO SEMÂNTICA

5.2.1 A definição como operação de expansão

5.2.1.1 A univocidade do conceito

5.2.1.2 A amplitude e a especificidade

BIBLIOGRAFIA

BIBLIOTECA
DO
IBICT

5. A DEFINIÇÃO SEMÂNTICA PARA A ELABORAÇÃO DE GLOSSÁRIOS

Maria de Fátima G. M. Tálamo

O glossário define-se como um instrumento de controle terminológico, cuja função específica é a de traduzir as linguagens técnica e científica numa linguagem sistêmica. É, portanto, um meio de conversão da linguagem-objeto (da Economia, das Artes, da Física etc...) para uma linguagem de dados e de informações. Sob esse ângulo, o glossário é uma metalinguagem técnica, usada de maneira unívoca tanto pelos produtores quanto pelos disseminadores da informação. Aos receptores, apresenta-se como o instrumento básico e uniforme, de conversão da linguagem sistêmica para a linguagem-objeto.

Para atingir seus objetivos, os descritores que integram a linguagem sistêmica devem ter seus significados fixados por meio de uma operação técnica conceitual, de forma que a definição de um termo exclua por completo as noções subjetivas que temos *a priori* a respeito do significado das palavras.

Na Idade Média, a palavra definição era usada como sinônimo de definição de uma coisa ou de uma idéia. Assim, ao definir um termo, relatava-se uma experiência ou uma investigação, de modo a delimitar o objeto real e concreto. Confundiam-se duas realidades: o mundo exterior e o universo da linguagem.

Com efeito, as unidades lingüísticas fundamentam-se apenas no seu emprego lingüístico: a decupagem da realidade extra-lingüística em unidades lingüísticas não considera as coisas em si mesmas mas manifesta o aspecto arbitrário da linguagem. Essa decupagem varia de uma língua para outra e no interior de uma mesma língua, de um estado para outro, no eixo da diacronia. Prova deste fato é que a articulação de um mesmo campo nocional varia segundo as línguas

consideradas. Assim, sobre o continuum da cor, a língua russa articula duas unidades, |goluboj| e |sinij|, na zona que o português nomeia |azul|. Do mesmo modo, os esquimós dispõem de quatro unidades para articular um continuum, que em português designamos neve.

5.1 A FORMULAÇÃO DE DEFINIÇÕES

Os dicionários tradicionais não são, via de regra, elaborados segundo critérios científicos. Além de não disporem de mecanismos de atualização (registram as palavras e os significados já consagrados pelo uso), os dicionários tradicionais não propõem uma classificação do léxico e uma orientação da polissemia. Induzem, normalmente, o usuário ao erro, à medida que consideram a palavra isolada de suas relações.

A formulação de definições apropriadas deve seguir os seguintes princípios (HEGENBERG, 1974, p.27):

1. a definição deve referir-se à essência daquilo que se procura definir;
2. uma definição não deve ser circular;
3. uma definição deve ser apresentada na forma afirmativa;
4. uma definição não deve ser feita em linguagem metafórica.

Os três primeiros itens não apresentam grandes problemas de compreensão, pois, na verdade, dizem respeito mais ao modo de apresentar a definição do que aos critérios que regulam seu processo de elaboração, como é o caso do item 1.

Assim, o termo a ser definido não pode sê-lo por um termo que será, por sua vez, definido pelo primeiro. Exemplo: OMITIR vt preterir, postergar...; PRETERIR vt ir além, omitir.... O fenômeno da circularidade é bastante comum nos dicionários tradicionais, uma vez que eles se situam no âmbito dos conhecimentos comuns.

Na forma negativa, uma definição perde a sua característica principal, qual seja a de fornecer a especificidade do termo a ser definido, porque remete a um conjunto muito amplo. Exemplo: "belo é tudo que não é feio" - evidentemente, "tudo que não é feio" não equivale a "belo".

Do mesmo modo, o uso da linguagem metafórica cria mais problemas do que esclarece. "A beleza é a eternidade contemplando-se no espelho" (HEGENBERG, 1974, p. 27), é uma definição que implica a presença de elementos subjetivos para ser decodificada.

O primeiro princípio, no entanto, não é de fácil compreensão. Na verdade, o que significa "essência"?

5.2 A DEFINIÇÃO SEMÂNTICA

Embora desempenhe inúmeras funções, a linguagem apresenta-se sempre como uma organização. As unidades lingüísticas são determináveis pelas relações que mantêm no interior dessa organização. A linguagem não é, portanto, uma nomenclatura, não designa coisas ou conceitos pré-existentes.

Sob esse ângulo, toda unidade lingüística para ser interpretada exige a presença de dois mecanismos: a comparação com as unidades semelhantes (que poderiam substituí-la num determinado segmento, isto é, unidades que pertencem a um mesmo paradigma) e a combinação com as unidades que coexistem num mesmo sintagma.

A comparação consiste no estabelecimento simultâneo de diferenças e semelhanças. Assim, o falante seleciona uma unidade, dada sua especificidade no interior de uma classe. O termo "canoa", por exemplo, participa da classe "embarcações", razão pela qual está associado a "iate", "jangada", "veleiro", ... e é individualizado pelo traço "para pesca". Desse modo, a equivalência "canoa" =

embarcação para a pesca, articula a especificidade da unidade sobre um traço constante "embarcações", o qual, por sua vez, define uma classe.

Portanto uma definição apropriada "compreende necessariamente dois termos: um deles designa a compreensão mais abstrata e geral da palavra a ser definida e o outro a sua diferença específica" (POTTIER, 1977, p.23).

5.2.1 A definição como operação de expansão

A definição de um descritor consiste numa paráfrase, de modo que a relação termo definido/definição tenda à homossemia. Exemplo: "Aluno" = "pessoa matriculada em estabelecimento de ensino".

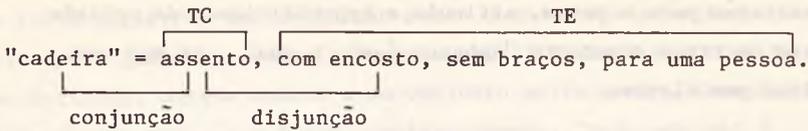
A definição se propõe, então, como tradução explícita do termo definido e, por essa razão, resulta de uma operação de expansão que explicita a diferença específica do termo em relação a um traço constante. Em resumo: a expansão articula a disjunção sobre a conjunção. Exemplo: relógio = máquina para marcar as horas

└──────────┘ + └──────────┘
conjunção disjunção

solteiro = homem que não se casou
└──────────┘ + └──────────┘
conjunção disjunção

A operação de conjunção delimita o domínio, ou campo semântico (CA) em que o descritor se inscreve, instaurando o traço comum (TC) a toda uma classe de termos, que estarão, por essa razão, associados. A disjunção, por sua vez, fornece os traços específicos (TE) do termo, de modo a individualizá-lo no interior de uma classe.

Assim, a definição do termo "cadeira" é que segue (baseado em LOPES, 1976, p. 279):



onde: o elemento conjuntivo, ou TC, estabelece o CA = "móvel para sentar-se, sobre pés" (= "assento"). Fazem parte desse CA os termos "banquinho", "sofá", "poltrona"... , à medida que todos eles manifestam o TC "assento". Portanto, os termos que compõem um mesmo CA estão associados e subordinados a um mesmo TC. Por outro lado, cada termo apresenta um ou mais TE que o singulariza. Assim, "com encosto" opõe "cadeira" à "banqueta"; "para uma pessoa" opõe "cadeira" a "sofá".

A operação de expansão é, portanto, uma operação de tradução entre signos e não envolve nenhum apelo ao objeto real. No exemplo acima, definiu-se "cadeira" como uma unidade lingüística, no interior de um determinado campo semântico e não o objeto cadeira.

5.2.1.1 A univocidade do conceito

A relação entre o termo definido e a definição deve ser unívoca, isto é, não deve levar a mais de uma interpretação.

Uma definição satisfatória deve ser unívoca e específica, do contrário permitirá o aparecimento da ambigüidade e da polissemia.

Uma definição genérica abre a possibilidade de uma série de interpretações, e não permite que ela funcione como um instrumento de conversão confiável.

Consideremos o descritor "carga horária semanal do professor" com a seguinte definição: "número de horas de trabalho do professor durante a semana" (BARBOSA, 1979, p.83). À primeira

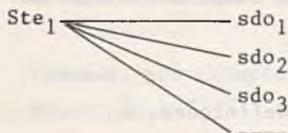
vista a definição é satisfatória, no entanto, basta inseri-la numa situação de comunicação, para que surjam os problemas de interpretação.

Na verdade, a definição apresentada leva, a pelo menos, três interpretações:

- a) Trabalho docente na escola
- b) Trabalho docente fora da escola
- c) Trabalho não docente na escola

todas elas decorrentes do alto grau de generalidade da definição.

Sempre que um descritor apresentar a relação significante/significado como se segue:



a definição deve ser revista, uma vez que esta situação é índice de ausência de especificidade.

Na linguagem comum, a palavra tem seu significado associado ao uso: o sistema linguístico pertence ao grupo social e apresenta uma organização de tal complexidade, que foge ao controle individual. A linguagem documentária, ao contrário, deve ser uniforme. Assim, cabe a um grupo determinar a sua forma de organização interna. Como bem observa Cintra (1983), as linguagens naturais são governadas por regras de FATO, não de direito, e as linguagens documentárias o são por regras de JURE, isto é, por grupos autorizados.

O glossário é, então, o parâmetro que impede a dispersão interpre-

tativa, devendo, portanto, propor conceitos ao mesmo tempo genéricos e unívocos, isto é, de uso universal e preciso.

Assim, desambigüizar o descritor "carga horária semanal do professor" equivale a propor uma definição que se aplique ao universo "carga horária docente" e, ao mesmo tempo, especifique a modalidade deste universo que está em jogo. Desse modo, as definições possíveis seriam:

- a) carga horária semanal que consta do contrato de trabalho do professor;
- b) a soma das horas/aula ministradas por cada professor, durante a semana;
- c) resultado da adição do número de horas/aula ministradas durante a semana com o número de horas semanais dedicadas à preparação de aulas, desde que essas sejam cumpridas dentro da escola.

É importante frisar que ao se fixar uma das definições, o descritor designa apenas e tão somente aquela selecionada. Caso seja necessário fixar duas delas, por exemplo, a e b, o descritor recebe uma modificação, de modo que $w = a$ e $w' = b$.

Do que foi dito, pode-se concluir que a desambigüização torna o campo mais analítico, o que não é, por outro lado, garantia de que a definição assim obtida seja pertinente ao universo do discurso tratado. Na verdade, a questão da pertinência resolve-se no âmbito das formulações de objetivos, ou seja, que tipo de informação será tratada e para quem ela se destina.

5.2.1.2 A amplitude e a especificidade

A descrição semântica deve, sempre que possível, obedecer uma ordem decrescente na seqüência dos traços, caminhando dos mais gerais para os mais específicos.

Assim, aos traços genéricos (TG) sucedem-se os traços específicos (TE). Os primeiros dão indicações da classe e podem ser subdivididos segundo seu grau de abrangência. Por exemplo, o TG "animado" é mais abrangente do que o TG "humano", mas ambos indicam uma classe. Os segundos, individualizam os termos, ou descritores, em relação ao traço genérico imediatamente anterior à seqüência de traços específicos. Por exemplo, o termo "cão" tem seus TE articulados em relação a "canino", que é, como o veremos, o último TG da série inicial, e não sobre "animal", que é uma classe, que comporta a subdivisão em outras.

Assim, denominaremos TC (traço comum) o último TG da série inicial, a partir do qual serão articulados os TE. Desse modo, o TC é o TG menos abrangente e definirá o campo semântico (CA), isto é, a classe em que o sentido específico de um conjunto de termos se define.

Tomemos, por exemplo, os termos "pato" e "cisne", com a seguinte descrição:

"pato" = $[M, C, A, A', P, S, V, A''; P', P'', D, C', G] = S_1$

"cisne" = $[M, C, A, A', P, S, V, A''; P', \bar{P}'', \bar{D}, \bar{C}', \bar{G}] = S_2$

(onde: M = material, C = contável, A = animado, A'' = animal, P = potente, S = sexuada, V = vertebrado, A'' = ave, P' = palmide, P'' = pequeno, D = doméstico, C' = de pescoço curto, G = gordo).

Desse modo: $D \in$ "pato" e $D \notin$ cisne, isto porque na descrição D = doméstico é um traço e não um domínio, embora possa sê-lo, em outra situação. O campo semântico CA será dado pelo TG menos abrangente, ou seja, na nossa terminologia TC, que corresponde a $S_1 \cap S_2 =$ "palmide". Os TE, por sua vez, articulam as diferenças específicas de cada termo no CA "palmide":

"pato" = palmide, doméstico, de pescoço curto

 conjunção disjunção

Evidentemente, tal descrição não considera a equivalência "pato" = "lentidão, bobice", uma vez que tal informação não é pertinente no universo de discurso considerado. O que está em jogo não é a linguagem comum, mas um universo de discurso limitado, que não cogita tratar essa informação.

Portanto, a redução da amplitude do campo semântico aumenta sua especificidade, isto é, "a amplitude do campo semântico definido por um termo é inversamente proporcional à especificidade semântica desse" (Paris, citado por BARBOSA, 1979, p. 90).

A questão da relação entre a amplitude e a especificidade pode ser assim visualizada:

$$\begin{array}{l}
 \text{"cão"} = \left[\overline{M}, C, A, A', P, S, V, M'; C', D, \overline{F}, \overline{A''} \dots \right] = S_1 \\
 \text{"lobo"} = \left[\overline{M}, C, A, A', P, S, V, M'; C', \overline{D}, F, \overline{A''} \dots \right] = S_2 \\
 \text{"raposa"} = \left[\overline{M}, C, A, A', P, S, V, M'; C', \overline{D}, \overline{F}, A'' \dots \right] = S_3
 \end{array}$$

\longleftarrow
 maior amplitude
 menor especificidade

\longrightarrow
 menor amplitude
 maior especificidade

(onde: M = material, C = contável, A = animado, A' = animal, P = potente, S = sexuado, V = vertebrado, M' = mamífero, C' = canino, D = doméstico, F = feroz, A'' = astuto).

Desse modo:

$$S_1 \cap S_2 \cap S_3 = \text{canino, que define o CA dos três termos.}$$

$$S_2 \cap S_3 = \text{canino + não doméstico, define um domínio menos abrangente, mais específico e que exclui "cão".}$$

A escolha do campo e a seleção dos traços são orientadas pelo critério de pertinência, que é o responsável pela definição do tipo específico de informação compatível com o universo. Portanto, uma definição pode ser altamente precisa mas não pertinente.

A definição do termo "quarto", por exemplo, variará segundo os universos de discursos considerados. Assim, a Engenharia Civil, o definirá, provavelmente, de acordo com a área, altura, ventilação, iluminação etc..., enquanto que a Arquitetura acrescentaria outros relativos à função, ao uso etc...

Do mesmo modo, um dicionário comum define "esmeralda" como pedra preciosa", traço que não será suficiente no domínio da Geologia, por exemplo. Nele, traços específicos, como: "minerais de brilho não metálico", "dureza alta", "fratura conchoidal", são pertinentes para o tratamento da informação.

Portanto, a formulação de uma definição apropriada não prescinde da formulação de objetivos, por mais rigoroso que seja o processo de análise dos termos.

Propositadamente, não se considerou, neste trabalho, os problemas decorrentes do tratamento da informação no universo dos discursos das Ciências Humanas e Sociais. Na verdade, nosso objetivo, neste momento, é apenas o de propor um procedimento de análise. De qualquer modo, vale lembrar que a análise dos discursos do referido universo é bastante complexa, uma vez que está sujeita à ausência de precisão do discurso original e às sucessivas revisões críticas dos sistemas teóricos. Por essa razão, apresentam-se campos semânticos contraditórios, termos pertencentes a campos semânticos complementares e traços ideológicos de difícil identificação (ver Capítulo 4)

BIBLIOGRAFIA

BARBOSA, M.A. (1979). Proposta de uma metodologia de análise estrutural e funcional de descritores de glossários técnico-científicos. *Acta Semiótica et Lingüística*, São Paulo, 3:69-96.

- CINTRA, A.M.M. (1983). Elementos de lingüística para estudos de indexação. *Ciência da Informação*, Brasília, 12 (1):5-22.
- ECO, U. (1974). *As formas do conteúdo*. São Paulo, Ed. Perspectiva/Ed. da Universidade de São Paulo.
- HEGENBERG, L. (1974). *Definições*. São Paulo, Cultrix/Ed. da Universidade de São Paulo.
- LOPES, E. (1976). *Fundamentos da lingüística contemporânea*. São Paulo, Ed. Cultrix.
- POTTIER, B. (1977). A definição semântica nos dicionários. In: LOBATO, L.M.P. (seleção, introdução e revisão). *A semântica na lingüística moderna: o léxico*. Rio de Janeiro, Francisco Alves, pp. 21-31.

6. A ANÁLISE DA IMAGEM: UM PRIMEIRO PLANO

Johanna W. Smit

6.1 O DOCUMENTO AUDIOVISUAL É ESPECÍFICO?

6.1.1 A "transparência" da imagem

6.1.2 A transcodificação

6.1.3 As informações técnicas

6.1.4 A "interpretação" da imagem

6.2 COMO ANALISAR A IMAGEM PARA EFEITOS DE DOCUMENTAÇÃO?

6.2.1 Como evitar a omissão?

6.2.2 Como evitar o excesso?

BIBLIOGRAFIA

6. A ANÁLISE DA IMAGEM: UM PRIMEIRO PLANO

Johanna W. Smit

O comportamento do usuário/pesquisador de imagens em fototecas é totalmente distinto do comportamento de alguém que procura um livro. Se esse não tem idéia de título ou autor do livro, mas sabe que está procurando informações sobre bases de dados bibliográficos, o título do livro de Murilo Bastos da Cunha "Bases de dados e bibliotecas brasileiras" lhe interessará e, pelo menos de imediato, o leitor considerará que obteve uma resposta à sua pergunta. Entretanto, se o usuário procura uma imagem de um pescador na beira de um rio, pode acontecer que ele encontre 30 ou 40 fotos que se inserem no assunto, mas que nenhuma resolva o problema... Por quê? A maioria dos pescadores está em pé e ele queria um pescador sentado. Mas as imagens de pescadores sentados mostram pessoas mais velhas, e o usuário queria um homem mais jovem. E, finalmente, naquela foto que mostra um pescador jovem, sentado, a expressão da menininha que brinca a seu lado não lhe agrada...

A descrição de uma imagem nunca é completa, como o exemplo acima mostra, e não se trata de uma questão de nível de especificidade. Senão, vejamos. Quem trabalha com imagens sabe que poderá se defrontar com um usuário procurando um rosto de velha, com cabelos grisalhos, presos, de óculos e muitas rugas. A velhinha precisa usar um colar clássico e trajar um vestido igualmente clássico, verde de preferência. A foto deve mostrar a velhinha meio de lado e esta deve transmitir uma impressão positiva. A tristeza da velhice não pode aparecer. Muito bem. Supondo que a fototeca encontre uma foto que corresponda perfeitamente à pergunta formulada, ainda é bastante provável que esta não seja aprovada, porque: a gola do vestido não agradou; os cabelos estão muito ralos; o sorriso é meio triste...

Como se depreende dos exemplos acima, quem trabalha com imagens trabalha com mais detalhes, mais informações e, principalmente, com informações menos evidentes: o que caracteriza um "sorriso meio triste"? e "cabelos meio ralos"? Como descrever e, sobretudo, analisar imagens? Como dar conta deste tipo de material? Com exemplos desta ordem, cabe até questionar se é possível "analisar" imagens...

6.1 O DOCUMENTO AUDIOVISUAL É ESPECÍFICO?

Quando a discussão chega neste ponto, há uma primeira divisão entre as opiniões: ou se considera que a análise da imagem não tem nada de específico e que as boas e comprovadas técnicas da análise documentária resolvem perfeitamente a questão, ou então já se parte do princípio de que as técnicas de análise de documentos escritos não são apropriadas para analisar imagens.

Esta dicotomia de opiniões atualiza algumas discussões que poderiam até ser chamadas de "semiológicas", e que giram ao redor de algumas questões que não são novas mas que, ao que tudo indica, ainda terão um futuro promissor.

6.1.1 A "transparência" da imagem

O principal problema, para quem quer trabalhar com imagens, é que a imagem tem aquela auréola de "espelho", a imagem mostra como a coisa "é", há "transparência" entre a imagem e o real, a imagem "é" o real. Este é o principal argumento invocado pelos defensores da aplicação das técnicas de análise documentária ao documento audiovisual, sem maiores adaptações. Afinal, se sabemos analisar o "real", e temos tesouros que dão conta dos fenômenos reais, saberemos e teremos como analisar a imagem. Todo este raciocínio é facilmente desmontado se lembrarmos que qualquer objeto pode ser fotografado de inúmeras formas diferentes... e que as imagens resultantes serão efetivamente diferentes. Os

semiólogos discutem amplamente esta pseudo "transparência" da imagem.

Em termos de análise documentária da imagem, esta questão tanto distingue quanto aproxima a imagem do documento escrito. Na introdução e no Capítulo 4 já foi discutida esta necessidade que hoje aparece como inevitável, ou seja, a ultrapassagem da acepção de documentação como técnica, para inseri-la no quadro mais amplo da produção e do consumo da informação.

A questão da produção da imagem, com a variedade de recursos técnicos que podem - e são - empregados para produzir certa imagem será analisada no item 6.1.3. Quanto ao consumo da imagem, é importante frisar sua distinção daquele que se processa frente ao documento escrito, dadas a flexibilidade e a adaptabilidade da imagem. Uma fotografia de uma praia (areia, mar azul, coqueiros, no fundo montanha com vegetação farta, ausência de pessoas, sol) pode servir, potencialmente, para ilustrar (ser consumida) muito mais contextos do que qualquer texto escrito que verse sobre a mesma praia. Há exemplos e mais exemplos de imagens que são consumidas deslocadas de suas condições originais de produção e, muitas vezes, este deslocamento passa totalmente despercebido. Um documentário sobre Getúlio Vargas fixa a câmara no então presidente discursando. No final de uma frase, ouvem-se muitos aplausos, o que leva muito "consumidor" à conclusão de que Getúlio Vargas, naquela fala, tinha o apoio da platéia. A trilha sonora faz "ver" uma platéia que o filme não "mostra". As condições de produção (a "técnica" e a cultura dentro da qual a técnica é exercida) e as condições de consumo (o poder de sugestão da imagem, e sua aparente transparência) devem ser, em todo caso, levadas em conta, se quisermos "bem" analisar uma imagem.

6.1.2 A transcodificação

Há os documentalistas que se preocupam com os resultados daquilo

que eles chamam de "problemas de transcodificação", ou seja, a tradução de um código para outro, com todos os problemas que toda tradução evoca automaticamente: perda de precisão, seleção de informação, possibilidade de erro etc. Afinal, analisar uma imagem significa, quer queiramos quer não, "traduzir" certos elementos desta imagem de um código icônico para um código verbal. Ex.: vejo uma "boina" e indexo "chapéu", porque meu tesouro não prevê maiores detalhes. Esta transcodificação leva a discussões bastante interessantes e riquíssimas do ponto de vista semiológico, mas por ora nos limitaremos à constatação que ela realmente ocorre e que, forçosamente, estas limitações da tradução têm consequências no rendimento e no procedimento da análise, a começar pela utilização dos tesouros. Os tesouros adequados para analisar documentos escritos dificilmente serão utilizados para analisar os documentos icônicos, pela simples razão de que estes demandam um número relativamente maior de termos concretos. Não há imagens de "agricultura" ou de "racismo", há plantações de soja ou milho, há cartazes em cima de portas com dizeres do tipo "for white only".

Desnecessário lembrar que a transcodificação que ocorre quando "traduzo" uma foto em palavras é uma operação relativamente simples quando comparada com a análise de filmes, compostos por imagem, trilha sonora e a interação dos dois (o som dos aplausos me faz ver a platéia).

6.1.3 As informações técnicas

O bibliotecário/documentalista habituado a trabalhar com documentos escritos não avalia, e nem pode, a importância que certas informações técnicas podem passar a ter no processo da análise documentária de imagens.

Não ocorre a nenhum usuário de bibliotecas diferenciar livros, que tratem do mesmo assunto, porque alguns foram compostos pela

tradicional tipografia de Gutenberg e outros foram impressos em *offset*, por exemplo. No caso das fotografias, estes detalhes "técnicos", referentes à produção do documento, mudam muita coisa e, portanto, devem ser, necessariamente, mencionados na análise. Basta lembrar o problema colocado pelo acabamento brilhante de uma foto, quando se quer utilizá-la diretamente numa produção de televisão: o acabamento apaco é muito mais adequado, porque não reflete a luz dos refletores.

Por enquanto, arrolamos três razões que levam à constatação de que o documento audiovisual é realmente diferente do escrito e que, como tal, demanda um tratamento documentário específico. A estas três razões, que eu chamaria de "condições de análise", eu acrescentaria uma outra, que concerne à "interpretação".

6.1.4 A "interpretação" da imagem

A associação entre a imagem e o real está de tal forma incorporada na leitura da imagem, que a percepção da imagem se torna difícil e demanda um certo treinamento. Conseqüentemente, na medida em que a dissociação entre a imagem e o real é difícil, é igualmente problemática a distinção entre o que seja descrição e interpretação, uma vez que a descrição da imagem, pela operação de tradução do código icônico para o código verbal (mencionada no item 6.1.2.), cria condições para sua interpretação.

O bibliotecário acostumado a analisar documentos escritos tende, invariavelmente, a analisar fotografias através de termos abstratos preponderantemente, esquecendo que a imagem dificilmente significa, de forma unívoca, um termo abstrato. Se, em documentação escrita, a utilização de termos abstratos geralmente significa uma medida de economia geral (um termo abstrato "resume" vários concretos), a imagem nos leva a uma realidade diametralmente oposta: o termo abstrato limita o significado de uma imagem, fixando uma leitura em detrimento de inúmeras outras. No caso, a

mesma preocupação com a economia geral do sistema o levará a descrever a imagem com termos concretos (p. ex.: peixes boiando na água, barriga para cima), deixando ao usuário a liberdade de atribuição dos termos abstratos que lhe parecem mais convenientes (p. ex.: poluição).

Para complicar mais ainda o quadro, vale lembrar que a imagem é, por definição, polissêmica.

Um exemplo passou para a história (não é o único, mas tem a vantagem de estar totalmente registrado, porque levou a um processo judicial): diz respeito a uma fotografia de Robert Doisneau, publicada na França com quatro legendas diferentes (a legenda, no caso, interpreta a imagem). A foto mostra o balcão de um bar, e duas pessoas de frente para este balcão, um homem de meia idade e, a seu lado, uma jovem que segura um copo que, na cultura francesa, todos associam a um copo de vinho. Vê-se outros copos, vazios, sobre o balcão.

As quatro legendas/interpretações são as seguintes:

- a) "uma jovem encantadora bebe vinho, encostada ao balcão de um bar, ao lado de um senhor de meia idade, que sorri".

Deixando de lado o julgamento de valor sobre a beleza da moça, a legenda descreve de forma bastante denotativa (ou "literal") a imagem, na ótica da análise documentária de imagens;

- b) "com um sorriso, um casal toma vinho e arruína sua saúde".

Esta legenda, de autoria da "Liga de Combate ao Alcoolismo", é próxima, em sua interpretação do "alcoolismo", da seguinte, enunciada por um serviço da Polícia parisiense:

- c) "apesar das novas regulamentações, uma infratora toma dois copos de vinho, ao invés de um único".

Os julgamentos de valor/interpretações ("casal", "arruinar a saúde", "infração à regulamentação") não são justificados pela própria imagem. Um homem e uma mulher na mesma foto não levam, de forma unívoca, ao conceito de casal, assim como a imagem de um copo de vinho não leva, forçosamente, à idéia da saúde arruinada. A contabilidade da polícia é mais discutível ainda, uma vez que ninguém prova, ao nível da imagem, que os outros copos vazios, no balcão, foram bebidos pela moça fotografada.

O conceito "alcoolismo" não está na imagem, mas pode ser interpretado a partir desta, em função de conceitos e valores de certos grupos sociais.

A 4ª legenda/interpretação é a mais "ousada":

d) "prostituição nos Champs-Élysées".

Além de constatar que a imagem do bar não fornece nenhuma informação que possa situá-lo na avenida dos Champs-Élysées, em Paris, mais temerária é a conclusão de que o consumo de um copo de vinho leva a uma relação de prostituição. A nível da imagem, nenhum elemento permite passar da análise denotativa para a interpretação/conotação, sendo que esta passagem se efetua em função de valores que não estão presentes na imagem. Pode-se ir além, e inverter a questão: para quem procura uma imagem para ilustrar o conceito "prostituição" (em contexto francês), será que esta imagem resolve, ou ele preferirá uma imagem de mulher em trajés provocantes, parada numa esquina?

A grande dificuldade na análise da imagem consiste nesta separação entre a denotação (o que a imagem mostra) e a conotação (o que a sociedade - e o bibliotecário - vêem, ou querem ver, na imagem), sabendo ainda que muitas vezes a legenda ou o contexto já nos desviam, subrepticamente, para a conotação.

6.2 COMO ANALISAR A IMAGEM PARA EFEITOS DE DOCUMENTAÇÃO?

A escolha de uma fotografia para "ilustrar" um conceito é um processo extremamente irracional, que não se presta a teoricizações. A prática indica que a seleção de uma única imagem que, oficialmente, responde integralmente ao pedido de um usuário, não é uma atitude eficiente, razão pela qual os documentalistas da imagem geralmente preferem entregar um grupo de imagens. As experiências realizadas por Ginette Bléry demonstraram que a escolha da imagem "boa" sempre se faz por comparação, sendo que a memória visual imediata do usuário, para comparação de imagens parecidas, não é muito desenvolvida: considera-se que 30 imagens constitui um campo suficiente para uma comparação eficiente. Mais de 30 imagens confundem o usuário, e menos de 30 imagens não forneceriam a necessária variedade de detalhes diferenciadores de imagens sobre o mesmo assunto.

Chegamos assim a uma regra básica na análise de imagens, e que parecerá descabida para o bibliotecário habituado a trabalhar com documentos escritos: a análise de imagens não precisa chegar a uma especificidade muito grande, mantendo portanto uma amplitude suficiente para que, a qualquer pergunta, se possa selecionar ao redor de 30 imagens que respondam à pergunta.

A exploração visual destas 30 imagens é suficientemente rápida (segundo cálculos de Abraham Moles, o olho abarca estas 30 imagens em meio segundo) para que uma seleção da "boa" imagem se faça com segurança e que uma série de detalhes não explicitados na pergunta possam, assim mesmo, ser levados em conta (a gola do vestido da velhinha, por exemplo).

O equilíbrio a ser encontrado entre a análise (e conseqüente indexação) dos detalhes de uma imagem que são importantes, e o descarte dos detalhes "insignificantes", para ser preciso sem ser específico demais: este é o dilema cotidiano do documentalís-

ta de imagens. Diz o provérbio que uma imagem vale por mil palavras, mas o provérbio não diz como priorizar estas palavras e, principalmente, **quando parar a descrição**, descartando, portanto, certas palavras. Num acervo de fotos históricas de um centro urbano, p. ex., tem-se tendência a chamar a atenção para o poste de luz ou então o chapéu-coco do senhor sentado naquele bonde...

O conhecimento do usuário e do acervo são imprescindíveis para resolver a dúvida, mas não descartam a utilização de um recurso extremamente prático (o recurso se aplica quando se trabalha com diapositivos unicamente): olhar o diapositivo mui rapidamente contra a luz - o detalhe não percebido desta forma provavelmente também não será percebido pelo usuário (excetuando-se certos detalhes que, apesar de ínfimos, não podem ser descartados porque mudam o conjunto da imagem: placas com nomes de ruas ou cidades, por exemplo).

Neste contexto, a pergunta que o documentalista de imagens se coloca com maior frequência sempre diz respeito ao "o que descrever" e "quando parar a descrição". Em outras palavras: como descrever uma imagem sem pecar por omissão, esquecendo de descrever detalhes importantes, e tampouco pecar por excesso, sobrecarregando o catálogo com informações desnecessárias?

6.2.1 Como evitar a omissão?

Certas categorias de informações são imprescindíveis na descrição da imagem. O Capítulo 4 já enumerou estas categorias e as ordenou de tal forma que uma hierarquização das informações, em função de sua relevância, para efeitos de documentação, pode ser determinada. A análise das imagens também recorrerá às categorias propostas, adaptando-as às suas necessidades específicas, a partir de uma listagem inicialmente elaborada pro G. Bléry:

- a. questões de "técnica" da produção da imagem, caso estes detalhes sejam "visíveis" (vista aérea, alto-contraste etc.);

- b. localização da imagem no espaço- termos geográficos ou então descrição do lugar: Pico da Bandeira ou interior de danceteria;
- c. localização da imagem no tempo - tempo histórico (anos 30) ou então, quando a imagem é atual, questões de época do ano, dia/noite, hora do dia etc.;
- d. quando a imagem focaliza seres vivos, estes devem ser descritos com certas precisões: idade, sexo, raça, atitude, tipo de roupa etc.;
- e. as ações destes seres vivos não podem ser esquecidas mas descritas em função daquilo que a imagem mostra e não em relação ao ato global no qual a imagem se insere. Neste sentido, não há fotos de "lançamento de nave espacial", há "astronauta, em traje especial, se dirigindo para a nave espacial". A fototeca pode até decidir que o ato global ("lançamento de nave espacial") também deve ser indexado, complementando a descrição da imagem propriamente dita: a documentação audiovisual leva a esta peculiaridade, ou seja, a "dupla indexação", também chamada "indexação em dois níveis";
- f. por último, deve-se descrever o ambiente no qual o ser vivo se encontra, quer seja elemento natural (praia, bosque, campo) ou artefato (casa, carro, cadeira). Caso não haja seres vivos na imagem, estes detalhes já foram descritos por ocasião do item b. acima.

Se a descrição responde às perguntas QUEM (seres vivos), ONDE (ambiente), QUANDO (tempo), ONDE (espaço), O QUE (ação) e COMO (técnica), poderemos supor que nenhum detalhe realmente importante tenha sido esquecido.

6.2.2 Como evitar o excesso?

O recurso da "vista d'olhos" já foi evocado, mas ainda pode ser lembrado outro recurso, mais técnico e, em certos casos, de

aplicação mais segura: o critério do 1º e 2º plano. Descreve-se o 1º plano, citando-se o 2º plano somente quando este modifica integralmente a leitura do 1º plano. Quando uma imagem traz uma atriz famosa em 1º plano, o 2º plano deverá ser mencionado caso este limite a localização da atriz no tempo ou espaço: um fotógrafo lambe-lambe ou então um longínquo Pão de Açúcar, por menos espaço que ocupem no conjunto da imagem, deverão ser descritos.

BIBLIOGRAFIA

- BARTHES, R. (1964). Rhétorique de l'image. *Communications*, Paris, (4):40-51.
- BLERY, G. (1981). La mémoire photographique. *Interphotothèque*, Paris. (41):9-33.
- BLERY, G. (1975). Une image vaut 1000 mots. *Le Photographe*, Paris, p. 20-27.
- CARVALHO, A.M. de F. (1986). *Fotografia como fonte de pesquisa: histórico, registro, arranjo, classificação e descrição*. Rio de Janeiro, Fundação Nacional Pró-Memória/Museu Imperial.
- CHAUMIER, J. (1977). *Le traitement linguistique de l'information documentaire: l'analyse documentaire*. Paris, Entreprise Moderne d'Édition.
- FOURNIAL, C. (1986). L'analyse documentaire des images animées. In: SAINTVILLE, D. *Panorama des archives audiovisuelles*. Paris, Documentation Française. p. 185-191.
- GARDIN, J.C. (1974). *Les analyses du discours*. Neuchâtel, Delachaux et Niestlé. Cap. 2.
- HUDRISIER, H. (1981). Relief, mouvement, cadrage et analyse de l'image. *Interphotothèque*, Paris, (41):69-82.
- LA BORDERIE, R. (1981). La documentation et l'audiovisuel. *Interphotothèque*, Paris, (41):59-67
- MOLES, A. (1981). La photographie, outil de connaissance de la vie sociale. *Interphotothèque*, Paris, (41):83-102.
- SHEFER, J.L. (1969). *Scénographie d'un tableau*. Paris, Seuil.

VIRUEZ, G.V. & FERREZ, H. (1985). *Tratamento técnico bibliográfico, catalogação, classificação e indexação de fotos*. Rio de Janeiro, FUNARTE. Trabalho apresentado no Seminário sobre Preservação e Conservação da Fotografia, Rio de Janeiro, RJ. mar. 1985.

1. The first part of the book is devoted to a general introduction to the subject of the history of the English language. It discusses the various influences that have shaped the language over time, including the contributions of Old English, Middle English, and Modern English. The author also touches upon the role of dialects and the process of language change.

CONTENTS

- CHAPTER I. THE HISTORY OF THE ENGLISH LANGUAGE. 1-100
- CHAPTER II. THE HISTORY OF THE ENGLISH ALPHABET. 101-150
- CHAPTER III. THE HISTORY OF THE ENGLISH GRAMMAR. 151-250
- CHAPTER IV. THE HISTORY OF THE ENGLISH VOCABULARY. 251-350
- CHAPTER V. THE HISTORY OF THE ENGLISH PROSE. 351-450
- CHAPTER VI. THE HISTORY OF THE ENGLISH POETRY. 451-550
- CHAPTER VII. THE HISTORY OF THE ENGLISH DRAMA. 551-650
- CHAPTER VIII. THE HISTORY OF THE ENGLISH NOVEL. 651-750
- CHAPTER IX. THE HISTORY OF THE ENGLISH ESSAY. 751-850
- CHAPTER X. THE HISTORY OF THE ENGLISH LETTER. 851-950
- CHAPTER XI. THE HISTORY OF THE ENGLISH JOURNAL. 951-1000

7. REVISÃO BIBLIOGRÁFICA

Isabel M. R. Ferin Cunha

Nair Yumiko Kobashi

Regina Keiko Obata F. Amaro

7.1 LINGÜÍSTICA GERAL

Saussure

Pottier

Fillmore

Greimas

Lyons

7.2 LÓGICA/FILOSOFIA DA LINGUAGEM

Apel

Pescador

Bronckart

Pêcheux

7.3 LINGÜÍSTICA/DOCUMENTAÇÃO

Gardin

Smit

Cintra

Cattenat e Paul

7. REVISÃO BIBLIOGRÁFICA

Isabel M. R. Ferin Cunha

Nair Yumiko Kobashi

Regina Keiko Obata F. Amaro

A análise documentária foi, durante muito tempo, feita em bases empíricas sem uma sistematização mais rigorosa de seus procedimentos. As necessidades de recuperação de informação em níveis cada vez mais refinados, a realização dessas atividades em bases industriais e o conseqüente uso de sistemas automatizados, passam a exigir maior rigor na representação e estruturação de conteúdos de documentos.

Diversos foram os caminhos trilhados em busca da sistematização de procedimentos que podem ser sintetizados em duas vertentes: a utilização dos métodos de ocorrência/frequência e os métodos de base lógico-semânticos. A revisão bibliográfica que ora apresentamos está na vertente lógico-semântica, e analisa os trabalhos mais significativos que se constituem em referencial para a construção de uma Teoria da Análise Documentária, tal como é proposta pelo Grupo Temma. Estão na base desta proposta a Lingüística, a Lógica e a Filosofia da Linguagem. A bibliografia analisada é apresentada aqui em três blocos, ordenados internamente de acordo com a relevância para a Biblioteconomia/Documentação:

1. Lingüística geral
2. Lógica/Filosofia da linguagem
3. Lingüística/Documentação.

7.1 Lingüística geral

SAUSSURE, Ferdinand de. *Curso de lingüística geral*. São Paulo, Cultrix, 1973.

Ferdinand de Saussure nasceu em 1857 e morreu em 1913 em Genebra, tendo exercido atividades de professor e pesquisador na sua cidade natal e em Paris.

O *Curso de Lingüística Geral* é uma obra póstuma (a 1ª edição é de 1916), resultado de três cursos de Lingüística Geral que Saussure ministrou na Universidade de Genebra entre 1906 e 1911. Considerada como obra clássica, constitui a base dos estudos lingüísticos modernos.

São apresentados e discutidos pela primeira vez conceitos fundamentais como língua/fala, signo/significado/significante, diacronia/sincronia, sintagma/paradigma, conceitos esses sempre presentes quer em estudos de seguidores ou contraditores de Saussure e que contribuíram também para as formulações estruturalistas em várias áreas do conhecimento como a Sociologia, Psicologia, Filosofia e Teoria da Informação, entre outras.

Em Biblioteconomia/Documentação essa obra interessa-nos pelas seguintes questões:

- 19) Definições conceituais que permitem à Análise Documentária uma maior precisão na análise, tradução e estruturação de informação.
- 20) Recuperação e adaptação dos conceitos de relações paradigmáticas e sintagmáticas que geram tanto a formação de palavras-chave como o controle de vocabulários e a sua adequação aos consumidores potenciais.

POTTIER, Bernard. *Linguistique générale: théorie et description*. Paris, Klincksieck, 1974.

Bernard Pottier é professor da Universidade de Paris, Sorbonne. Seus trabalhos "Linguistique générale: théorie et description" (1974) e "A definição semântica nos dicionários" (1977) trazem

elementos significativos para a construção das bases teóricas da Análise Documentária.

Em "Linguistique générale", Pottier apresenta a noção de "caso", um dos aspectos teóricos e descritivos da Linguística indispensáveis ao estudo da comunicação e da expressão.

Segundo Pottier, os "casos" existem em todas as línguas e podem se dividir em três zonas:

Zona I (Zona central; de participação primária), onde, na relação atributiva, não há geralmente distinção de caso; a base está no Nominativo. Na relação ativa, cria-se uma oposição entre o termo da base: Ergativo (dotado de poder) e o Acusativo (elemento de não-poder).

Zona II (Zona de participação secundária), onde os elementos de participação situam-se antes ou depois de um ponto de vista cronológico. É subdividido em: Zona IIa, onde encontramos os casos Causal, Instrumental e Agentivo; e, a Zona IIb, com os casos Dativo, Benefactivo e Final.

Zona III (Zona de dependência), onde a dependência se distribui em três Zonas de experiência: a) Zona de experiência espacial (E), b) Zona de experiência temporal (T) e c) Zona de experiência nacional (N). O caso vinculado a esta Zona é chamado de Locativo, sendo especificado em cada um desses domínios.

Além dos casos vinculados às três Zonas acima, existe um caso marginal chamado de Sociativo.

Adequações a essas noções de caso, possibilitam uma aplicação na Análise Documentária em função da:

- 19) proposta de procedimentos metodológicos para a estruturação do vocabulário de base de área específica, através da identificação das categorias formais as quais pertencam as palavras do vocabulário de base.
- 29) formação do campo semântico da área específica em que o vocabulário deverá ser estruturado.

POTTIER, Bernard. A definição semântica nos dicionários. In: *A semântica na lingüística moderna: o léxico*; seleção, introdução e revisão de Lúcia Maria Pinheiro Lobato. Rio de Janeiro, F. Alves, 1977. pp. 21-31.

Neste artigo, Pottier trabalha com a questão do conteúdo semântico de uma palavra. Segundo o autor, este conteúdo semântico pode ser analisado distinguindo-se quatro tipos de definidores: *semema*, traço semântico mínimo; *classema*, conjunto de classes conceituais gerais; *arquissemema*, subconjunto comum a um conjunto de *sememas*; e, o *virtuema*. Neste sentido, todo o *lexema* terá como significado a definição independente ou *semantema* (*semema* + *classema*) e a definição relativa (*arquissemema*).

Essas noções são importantes para a Análise Documentária tendo em vista:

- 19) a construção do glossário de área específica do vocabulário a ser controlado, possibilitando a identificação dos traços semânticos mínimos e distintivos das palavras.
- 29) a passagem das palavras da linguagem natural para a linguagem documentária em função da estruturação do vocabulário controlado.

FILLMORE, Charles J. Em favor do caso. In: *A semântica na lingüística moderna: o léxico*; seleção, introdução e revisão técnica de Lúcia Maria Pinheiro Lobato. Rio de Janeiro, F. Alves, 1977. pp. 277-365.

Charles J. Fillmore é um lingüísta norte-americano que, com base na teoria chomskyana, trabalha uma gramática de casos. Analisa os universais lingüísticos, propostos em estudos de Chomsky, a partir de universais sintáticos formais e materiais lingüísticos, onde a noção gramatical de "caso" está na base da gramática de todas as línguas.

É feita uma modificação na teoria da gramática transformacional com uma reintrodução da interpretação dos conceitos que baseiam os sistemas de caso através de uma compreensão da diferença entre Estrutura Profunda e Estrutura Superficial.

Na estrutura básica das sentenças encontramos o que poderia ser chamado de "proposição", um conjunto de relações, incluindo verbos e substantivos, separado do constituinte de "modalidade" que incluirá, nas sentenças tomadas como um todo, modalidades como a negação, o tempo, o modo e o aspecto. As noções de caso incluem um conjunto de conceitos universais que identificam certos tipos de julgamentos que somos capazes de fazer acerca dos acontecimentos e assuntos tais como "quem fez", "com quem aconteceu" e "o que foi mudado".

Os casos que parecem ser necessários incluem: Agentivo, o caso do instigador da ação identificada pelo verbo, instigador esse percebido tipicamente por ser animado; Instrumental, o caso da força ou objeto inanimado, causalmente implicado na ação ou estado identificados pelo verbo; Dativo, o caso do ser animado afetado pelo estado ou ação identificados pelo verbo; Factivo, o caso do objeto ou ser resultante da ação ou estado, identificado pelo verbo, ou compreendido como parte do significado do verbo; Locativo, o caso que identifica a localização ou orientação espacial do estado ou ação identificados pelo verbo; Objetivo, o caso semanticamente mais neutro, no qual qualquer coisa representada por um substantivo cujo papel na ação ou estado nomeado pelo verbo é identificado pela interpretação semântica do próprio

verbo.

Na Análise Documentária, esses conceitos são subsídios para:

- 19) a análise de textos/documentos. Foram utilizados por Austin para o sistema de indexação PRECIS.
- 20) através de adaptações, possibilitar o estabelecimento de procedimentos metodológicos de análise das categorias dos vocabulários de área específica para a formação do campo semântico e estruturação do vocabulário controlado.

GREIMAS, Algirdas J. Da Modalidade. In: *Semiótica do discurso científico*. São Paulo, Difel/Sociedade Brasileira de Professores de Linguística, 1976. pp. 57-86.

Algirdas Julien Greimas é professor da École des Hautes Études en Sciences Sociales de Paris. Linguísta e semiólogo, é autor de obras como ensaios, artigos, textos teóricos e de aplicação. O seu trabalho "Da modalidade" tem como objetivo propor elementos para a elaboração de uma teoria das modalidades discursivas, constituindo um instrumental para as pesquisas semio-lingüísticas.

A primeira parte trata das estruturas modais simples: o ato; os enunciados elementares, a transformação, a junção; performance e competência; as modalizações translativas, as modalidades verdictórias, as modalidades factivas; encadeamento das estruturas modais simples.

A segunda parte trata das sobremodalizações: a competência e suas sobremodalizações; inventário provisório, categorização e denominação; modalizações do sujeito e do objeto; abordagem sintagmática; organização da competência pragmática.

Na terceira parte discute as confrontações modais: modalizações

aléticas, deônticas e bulênticas (volitivas); os sistemas de regras e aptidões dos sujeitos.

Este trabalho interessa à Biblioteconomia/Documentação pelas seguintes razões:

19) Propõe um modelo de análise

Em Biblioteconomia/Documentação necessita-se de métodos de análise de textos/documentos. A utilização pragmática das "modalidades" poderá trazer, num primeiro momento, a compreensão da lexicalização do discurso assim como a apreensão dos atos ilocucionais, contribuindo para o reconhecimento das "constantes" que regem implicitamente o "bom-senso" na identificação da "informação significativa".

20) Discute na "teoria das modalidades" a inter-relação linguagem-lógica-metalinguagem.

Em Biblioteconomia/Documentação a inter-relação linguagem-lógica-metalinguagem é fundamental para a análise de textos/discursos/documentos. Neste sentido, a identificação através da "teoria das modalidades" das funções sujeito, objeto e verbo permitem-nos levantar hipóteses de trabalho centradas na coincidência entre tipos de modalidades e maior relevância informativa de textos/documentos.

LYONS, John. *Semântica I*. Lisboa, Editorial Presença/Martins Fontes, 1977.

John Lyons, professor de Linguística Geral da Universidade de Sussex, apresenta nesta obra uma abordagem de Semântica, extraindo das várias Ciências (Etnologia, Psicologia, Filosofia, Antropologia e Linguística) as discussões acerca do significado e da comunicação.

Os estudos da Semântica com essas abordagens são relativamente recentes, tendo surgido a partir do final de década de 60 e início da década de 70.

São discutidos os conceitos básicos da Semântica como as distinções entre Sentido e Significado; Referência e Denotação, bem como aspectos da Semântica Estrutural como a teoria dos campos semânticos e as relações de sentido, fundamentais para a análise da linguagem.

Este trabalho contribui para a Biblioteconomia e Documentação:

- 1º) clarificando os conceitos envolvidos nas diferentes etapas de Análise Documentária.
- 2º) na explicitação de procedimentos e métodos de análise de textos e estruturação de vocabulários.

7.2 Lógica/Filosofia da linguagem

APEL, Karl-Otto. *Towards a transformation of philosophy*. London, Routledge & Kegan Paul, 1980.

Karl-Otto Apel é professor de Filosofia da Universidade de Frankfurt. O seu trabalho engloba tanto as tradições Anglo-Saxônicas de filosofia analítica da linguagem e do pragmatismo como as tradições filosóficas do continente europeu da fenomenologia, existencialismo e hermenêutica.

O seu objeto centra-se na linguagem como construtora de pensamento e significações. É portanto, um estudo de filosofia da linguagem. Neste sentido analisa: o conhecimento hermenêutico proposto por Wittgenstein; a passagem de uma teoria da ciência para uma ciência cognitiva de base antropológica; a transformação da lógica transcendental de Kant e Peirce e seus desdobramentos metodológicos nas Ciências Sociais; os problemas do sujeito

cognitivo na interpretação dos signos na semiótica pragmática; a relação entre a teoria da linguagem de Chomsky e a filosofia contemporânea.

As suas propostas visam um melhor reconhecimento das possibilidades humanas de "CONHECER" através da tomada de consciência do processo de formação do pensamento, da linguagem e da Comunicação em Sociedade, o que implica em última instância uma nova metodologia centrada no Homem para as Ciências, principalmente as Sociais.

Este trabalho interessa à Biblioteconomia/Documentação por três razões básicas:

1º) Discute o processo de aquisição do Conhecimento e relaciona-o com o pensamento e a linguagem.

Na Biblioteconomia/Documentação estas questões surgem-nos na análise de textos/documentos e na tentativa de determinar as construções lógicas realizadas pelo autor/produzidor.

2º) Discute a Semiótica e suas implicações com o sujeito cognitivo.

Na Biblioteconomia/Documentação este problema está presente no processo de representação de textos/documentos e na seleção individual de conceitos/palavras-chave realizada pelos analistas.

3º) Discute uma nova metodologia para as Ciências centrada no Homem cognitivo.

Na Biblioteconomia/Documentação esta questão leva-nos a pensar quem é esse Homem cognitivo, que metodologia adotar e quais suas implicações na nossa incipiente metodologia.

PESCADOR, José Hierro S. *Principios de Filosofía del Language.*

Madrid, Alianza Editorial, 1980, 2 vols.

José Hierro Pescador é professor da Universidade Autónoma de Madrid e da Universidade Complutense. O seu trabalho apresenta-se como um manual que tem como objetivo a filosofia da linguagem vista através da evolução das doutrinas lingüísticas, semióticas e psicolingüísticas.

O volume I tem como objetivos: os aspectos mais genéricos da linguagem como sistema de signos; os aspectos comuns à linguagem humana e aos sistemas de comunicação animal; a gramática como mecanismo gerador, principalmente o modelo Chomskiano, as críticas surgidas ao modelo Chomskiano e seus desdobramentos; a aquisição da linguagem e sua relação com o Conhecimento.

O volume II tem como objetivo geral a teoria do significado e envolve as seguintes questões: as idéias precursoras de Stuart Mill, a fundação da teoria do significado por Frege e o atomismo lógico de Russell de Wittgenstein; a teoria dos atos de fala de Austin e Searle e a concepção pragmática do significado em Grice; a teoria verificacionista de Carnap prosseguida por Tarski e Quine e seus desdobramentos.

Como método, o próprio autor afirma seguir no volume I uma abordagem generalista e sistemática enquanto no volume II preferiu optar por um maior detalhamento e argumentação, relevando as conexões históricas.

As propostas do estudo encontram-se no final do volume II resumindo-se: viabilidade de uma teoria completa e unitária do significado; relações entre ideologia e linguagem.

Este trabalho interessa à Biblioteconomia/Documentação pelas seguintes razões:

19) Discute a linguagem enfocando a interpenetração da Lingüística, Lógica, Semiótica e Psicolingüística.

Em Biblioteconomia/Documentação encontram-se envolvidas várias linguagens: a Linguagem Natural (e especificidades dos autores/produtores, analistas e consumidores); a Linguagem Natural (e especificidades das áreas); a Linguagem Documentária (e especificidades da sua artificialidade). Um estudo interdisciplinar como esta obra, alerta-nos para os problemas envolvidos e para os subsídios teóricos e metodológicos a serem retirados.

20) Discute a inter-relação Linguagem-Conhecimento e Linguagem-Lógica.

Na Biblioteconomia/Documentação esta inter-relação não se encontra muito clara apesar de nos surgir na análise de textos/documentos, na estruturação de instrumentos de trabalho e nas definições que temos de adotar a nível de vocabulários.

30) Discute a viabilidade de uma teoria completa e unitária do significado e da interferência da ideologia na linguagem.

Em Biblioteconomia/Documentação ambas as questões são problemáticas na medida em que analisamos textos/documentos e os traduzimos em conceitos/palavras-chave, isto é buscamos significados. Nestes reflete-se o problema da transferência da ideologia da linguagem natural, especializada e documentária.

BRONCKART, Jean Paul. *Le fonctionnement des discours*. Paris Delachaux & Niestlé, 1985.

Jean Paul Bronckart é doutor em psicologia genética e trabalha atualmente com psicologia da linguagem. É desde 1976 professor de psicopedagogia das línguas na Universidade de Genebra.

Neste trabalho, seu objeto é a atividade de linguagem considerada em condições sócio-culturais determinadas. É seu objetivo estabelecer os fundamentos de uma descrição explícita de conjunto de entidades (parâmetros, operações, unidades) suscetíveis de dar conta dos comportamentos presentes no processo de produção de discursos. Considerando que as condutas verbais são orientadas por propósitos comunicacionais determinados, propõe um modelo teórico de produção do discurso com base na filosofia da linguagem, lingüística e psicologia, aplicando esse modelo na análise de discursos concretos em francês contemporâneo. Definidas as intenções comunicativas como as de informar, clarificar, ativar e criar canais de interação, propõe quatro tipos de discursos fundamentais que, por obedecerem a essas condições de produção diferentes, caracterizar-se-ão como tipos de textos também diferentes em termos morfossintáticos. São eles: Discurso em Situação, Discurso Teórico, Discurso Dialógico, Discurso Narrativo.

Para precisar, operacionalizar e quantificar as unidades lingüísticas características dessas diferentes categorias de textos, monta dicionários de análise compostos por classes lexicais, morfemáticas e verbos.

A importância deste trabalho para a Análise Documentária localiza-se:

- 19) na proposta de análise do discurso, balizada pela interação social e o ato de produção de textos, superando portanto o patamar da simples análise da palavra (ocorrência e frequência) no texto.
- 20) na identificação de unidades significativas no texto e a subsequente elaboração de dicionários, o que permite um avanço considerável na análise automática de textos, considerados os limites de operacionalização apontados pelo próprio autor.

32) na possibilidade de, a partir desse modelo trabalhar-se em textos arquetípicos em língua portuguesa, tendo em vista a construção de dicionários análogos.

PÊCHEUX, Michel. *Analyse automatique du discours*. Paris, Dunod, 1969.

Michel Pêcheux, pesquisador francês de formação filosófica, tem se preocupado com as questões relacionadas com a teoria das ideologias e a história das ciências.

A partir dessas preocupações, Pêcheux identifica a necessidade de uma teoria do discurso, que implica, por sua vez, na tentativa de construção de um instrumento de análise do discurso.

Este trabalho, originalmente sua tese de doutoramento de 1968, é o resultado de pesquisa interdisciplinar realizada no Laboratório de Psicologia Social da Sorbonne, com a concorrência de várias áreas do conhecimento: Linguística, Psicologia-linguística, Pragmática, Retórica, Semântica.

A partir da definição de "discurso" como um processo determinado pelo conjunto de condições de produção sobre uma língua natural dada, Pêcheux propõe uma análise de conteúdo que não utiliza esquemas sintático-semânticos propostos tradicionalmente. A sua proposta consiste na transformação da seqüência discursiva em um enunciado elementar de composição e dimensão fixas, sendo as seqüências discursivas determinadas pelas condições de produção idênticas (zonas de similaridade). A partir desses componentes teóricos, faz a aplicação da análise automática do discurso no campo da investigação sociológica e da história das ciências.

A contribuição deste trabalho para a Biblioteconomia e Documentação está em:

- 19) propor a aplicação de uma metodologia de análise do discurso para o estabelecimento de procedimentos de análise de conteúdo com a finalidade de recuperar a informação.
- 20) dar subsídios para a aplicação de metodologias que permitam avançar na automação da Análise Documentária.

7.3 **Linguística/Documentação**

GARDIN, Jean-Claude. *Les analyses de discours*. Neuchâtel, Delachaux et Niestlé, 1974.

Jean-Claude Gardin é diretor do Centre National de Recherches Archéologiques (CRA) pertencente ao Centre National de la Recherche Scientifique (CNRS) e professor na École Pratique des Hautes Études em Paris. Os seus trabalhos visam a Análise do Discurso em textos de caráter científico em Ciências Humanas, com vista à sistematização da informação neles contida. Neste sentido propõe métodos de análise para a informatização utilizando preferencialmente estudos de caso em Arqueologia. Seus trabalhos mais significativos para a Análise Documentária são: "Les analyses de discours" (1974), "La logique du plausible: essais d'épistémologie pratique" (1981), "Informática e Arqueologia" (1985). "Systèmes experts et sciences humaines: les cas de l'Archéologie" (1986).

"Les analyses de discours" é composto de quatro capítulos que tratam basicamente de discursos das áreas de Ciências Humanas e de seus respectivos procedimentos. Faz a distinção epistemológica entre as Ciências do discurso e o Discurso da Ciência visando detectar as particularidades existentes nessas construções. Compara os vários tipos de procedimentos e métodos da construção de discursos nas Ciências Humanas com a Análise Documentária - seus procedimentos, métodos e finalidades. Propõe que a Análise Documentária, dado o conjunto de informação e necessidades crescentes de

informatização, proceda a uma formalização dos seus procedimentos. Para tal, propõe instrumentais lógico-lingüísticos que permitam tanto a análise do discurso como a racionalização da transferência da informação.

Para a Biblioteconomia/Documentação este trabalho interessa em razão das seguintes questões:

- 19) Proposta de métodos e procedimentos para a Análise de texto/discurso que possibilitam análises "objetivas" e "universais".
- 20) Proposta de interdisciplinaridade lógico-lingüística, realizando recortes pragmáticos com vista à recuperação da informação.
- 30) Sistematização de métodos e procedimentos lógico-lingüísticos de forma a permitir a operacionalização da recuperação e disseminação da informação.

SMIT, Johanna. *Les langages documentaires comme métalangages du discours scientifique*. Paris, École Pratique des Hautes Études, 1973. Mémoire.

Johanna Smit é professora de graduação e pós-graduação do Departamento de Biblioteconomia e Documentação da ECA-USP. Criou na ECA a linha de pesquisa denominada Análise Documentária, área na qual vem pesquisando desde 1971. Fez mestrado e doutorado na França, tendo sido orientada de Jean-Claude Gardin. O trabalho "Les langages documentaires comme métalangages du discours scientifique" é sua dissertação de mestrado, de cuja banca examinadora participaram Roland Barthes e A.J. Greimas. Trata-se de um estudo teórico do fazer documentário, tendo como parâmetros os estudos semânticos e semiológicos. Sua hipótese de trabalho desenvolve-se na abordagem da natureza do discurso científico e das linguagens documentárias enquanto metalinguagens desses discursos. Recupera as noções mais importantes na lingüística para a análise do discurso científico, compara as operações de

descrição de significação com as etapas da Análise Documentária, analisa sistemas documentários de critérios semânticos e não-semânticos e os métodos de tradução sintática.

Resumo desse trabalho foi publicado em artigo intitulado "Documentação e lingüística: inter-relação e campos de pesquisa" na *Revista Brasileira de Biblioteconomia e Documentação*, 11 (1/2), pp. 33-42, jan./jun. 1978.

Para a Biblioteconomia e Documentação este trabalho interessa em razão de:

- 1º) sistematizar as questões de lingüística importantes para a análise do discurso.
- 2º) sistematizar as questões da lingüística importantes para a análise documentária.
- 3º) evidenciar os problemas de linguagem que se colocam para a análise documentária frente a discursos de Ciências Exatas e Ciências Humanas.

CINTRA, Anna Maria M. Elementos de lingüística para estudos de indexação, *Ci. Inf.*, Brasília, 12(1):5-22, 1983.

Anna Maria Marques Cintra é doutora em Letras (Lingüística) pela USP, professora de graduação e pós-graduação da PUC-São Paulo e professora convidada do Curso de Pós-graduação da ECA-USP. Foi professora do Curso de Mestrado em Ciência da Informação do IBICT.

No trabalho "Elementos de lingüística para estudos de indexação" o objetivo é o de buscar dentro da Lingüística geral elementos auxiliares para indexação em língua portuguesa. Assim, levanta as questões de leitura envolvidas nos procedimentos de indexação e elementos de semântica, sintaxe e morfologia que permitem dar bases teóricas mais sólidas para a teoria da análise documentária. Discute os conceitos de língua e fala, arbitrariedade e

linearidade do signo lingüístico e as questões referentes à linguagem natural e linguagens documentárias. Trata também dos vocabulários especializados, dos campos semânticos e da organização de núcleos semânticos.

Este trabalho interessa à Biblioteconomia/Documentação por:

- 1º) Evidenciar a leitura como questão importante no processo de indexação.
- 2º) Fazer recortes na teoria que possibilitam, a partir de categorias embasadas nos casos lingüísticos, a operacionalização de procedimentos para indexação e construção de vocabulários especializados.

GARDIN, Jean-Claude e outros. *La logique du plausible: essais d'épistémologie pratique*. Paris, E. Maison des Sciences de l'Homme, 1981.

Em "La logique du plausible" J.C. Gardin propõe uma epistemologia prática em Ciências Humanas. Essa epistemologia centra-se em duas questões básicas: como se organizam os raciocínios em Ciências Humanas; quais os sistemas simbólicos utilizados nesses discursos. Dadas estas questões, propõe uma maior formalização na construção dos discursos de Ciências Humanas. Essa formalização é condição necessária tanto à sua consolidação como à eficácia da Análise Documentária.

Neste sentido, a Análise Documentária não seria apenas a representação do texto mas a reconstrução do percurso epistemológico da construção do discurso, reconstrução que visa tanto a identificação do corpus em análise (objetivos, métodos de construção/descrição/ordenação) e validação, como a possibilidade de simular construções análogas. Nesta proposta de simulação entra-se no campo da inteligência artificial através de sistemas lógico-lingüísticos e da fixação de uma metalinguagem.

Para a Biblioteconomia/Documentação este trabalho interessa em razão das seguintes questões:

- 19) Formalização, explicitação e validação de procedimentos de análise de discurso.
- 20) Fixação de metalinguagens.
- 30) Reconhecimento de procedimentos epistemológicos na construção de discursos científicos em Ciências Humanas e possibilidade de simulação de construções análogas em computador.

GARDIN, Jean-Claude. *Informática e arqueologia*. Lisboa, Inst. Nac. de Invest. Científica, 1985.

Em "Informática e Arqueologia", tomando como estudo de caso a Arqueologia, retoma as questões colocadas em "Les analyses de discours" e "La logique du plausible". Avança no sentido de propor novos parâmetros para a construção de bases de dados tendo como vetor a qualidade e não a quantidade. Esses parâmetros visariam a reconstrução do circuito de produção e consumo da informação assim como a explicitação dos princípios que lhe deram origem. Na reconstrução desses circuitos propõe simulações através da Inteligência Artificial, as quais permitirão, dadas as alternativas apresentadas, equacionar núcleos informacionais produtores e consumidores potenciais de informação. Dessa equação e suas alternativas resultarão possibilidades de escolha de produtos documentários mais adequados.

Para a Biblioteconomia/Documentação este trabalho interessa em razão das questões já levantadas em seus trabalhos anteriores e por demonstrar a aplicabilidade de simulações em Ciências Humanas com fins documentários.

GARDIN, Jean-Claude e outros. *Systèmes experts et sciences humaines: les cas de l'Archéologie*. Paris, Eyrolles, 1986.

Através de exemplos de aplicação na área de Arqueologia, o autor procura identificar o desenvolvimento da Inteligência Artificial e os problemas teóricos que surgem, tanto nesta área como em outras áreas das Ciências Humanas. Este trabalho tem como objetivo uma reflexão sobre o uso adequado de novas tecnologias para fazer avançar as próprias ciências do homem através de uma expressão mais clara das regras que comandam as interpretações dos fatos e dos comportamentos humanos.

Em Documentação, este trabalho contribui para a aplicação dos resultados obtidos com o uso das novas tecnologias, em particular da Inteligência Artificial, nas questões relacionadas à análise automática de textos que possam ser utilizadas na Análise Documentária.

CATTENAT, Annette e PAUL, Gérard. *Intelligence artificielle*. IN: *LES NOUVELLES TECHNOLOGIES DANS L'INFORMATION SCIENTIFIQUE ET TECHNIQUE*. Valbonne, Institut National de Recherche en Informatique Automatique, 1984.

Annette Cattenat e Gérard Paul são documentalistas e professores da Universidade de Paris VIII. O seu trabalho apresenta-se como uma introdução ao uso da Inteligência Artificial na Informação Científica e Técnica.

O artigo tem os seguintes objetivos: fornecer noções gerais sobre a IA no tratamento da Documentação; apresentar as linhas de pesquisa existentes em IA, seus objetivos, procedimentos e suportes teóricos.

O artigo propõe a utilização dos sistemas de compreensão da linguagem natural, desenvolvidos pela IA, no tratamento da

Este trabalho interessa à Biblioteconomia/Documentação pelas seguintes razões:

1º) Discute os sistemas de representação dos sentidos.

Em Biblioteconomia/Documentação trabalha-se com "sentidos" nas análises de textos/documentos, na atribuição de conceitos/palavras-chave, na estruturação e controle de vocabulários, na construção e uso de tesauros. Dado isto, é necessário conhecer como se processa essa representação e quais os instrumentos utilizados. Por exemplo, quais os indicadores gramaticais, sintáticos, semânticos etc.

2º Discute os sistemas de representação dos conhecimentos e dos raciocínios.

Em Biblioteconomia/Documentação tem-se necessidade de entender como se organiza o conhecimento e a que princípios presidem determinados raciocínios, principalmente quando se trabalha com textos/documentos científicos. Neste sentido, torna-se fundamental entender as estruturas lógicas e epistemológicas dos circuitos de produção/consumo da informação tanto ao nível das instituições como dos autores/produtores e consumidores.



ICI
O
ECV