

CONCEITOS DE ORGANIZAÇÃO E REPRESENTAÇÃO DO CONHECIMENTO NA ÓTICA DAS REFLEXÕES DO GRUPO TEMA

CONCEPTOS DE ORGANIZACIÓN Y REPRESENTACIÓN DEL CONOCIMIENTO BAJO LA PERSPECTIVA DE LAS REFLEXIONES DEL GRUPO TEMMA

Marilda Lopes Ginez de Lara - larama@usp.br
Doutora em Ciências da Comunicação, Prof. Associada do
Depto de Biblioteconomia e Documentação da ECA-USP.

Resumo

Relaciona a visibilidade da ciência com sua terminologia e apresenta elementos para a delimitação de conceitos de Organização e Representação do Conhecimento na ótica das reflexões do Grupo Temma, recuperando os termos Análise Documentária, Leitura Documentária, Informação Documentária, Linguagem Documentária e Linguística Documentária. Justifica a seleção argumentando sobre a necessidade de se reportar às origens dos termos, a maioria proveniente da vertente francesa de Documentação. Conclui ressaltando a importância de maior pesquisa sobre a formação dos conceitos da área, quer para proporcionar maior visibilidade às pesquisas, quer para encaminhar projetos de harmonização de conceitos.

Palavras-chave

Organização e representação do conhecimento. Análise documentária. Linguística documentária. Leitura documentária. Informação documentária. Linguagem documentária. Grupo temma.

1 INTRODUÇÃO

"[...] a trama de todo pensamento e de toda pesquisa são os símbolos", sendo a linguagem a própria essência do pensamento, assim como a precisão, a contribuição de outras 'mentes' e a 'liberdade mental'. Mas é indispensável haver um "acordo geral acerca do uso de termos e notações", acordo esse não como imposição arbitrária, "mas por força de princípios racionais sobre a conduta dos homens" (PEIRCE, 1977, p. 39-40).

As condições de visibilidade da Ciência dependem de seu vocabulário. Isso coloca a Ciência da Informação em situação delicada, uma vez que este ainda se encontra em processo de construção: debate-se sobre o objeto dessa ciência, sobre a pertinência de classificá-la desse modo, ou desloca-se a questão a partir da identificação do campo

como interdisciplinar, o que não resolve o problema. Há profusão de expressões que são aproximadas na prática, mas há também conceitos emprestados de outras disciplinas sem efetivo sinal de apropriação. Essa situação também caracteriza o subcampo da Organização e Representação do Conhecimento, acarretando dificuldades para o entendimento dos conceitos ou para estabelecer uma harmonização entre eles.

Como observa Benveniste,

[...] a história particular de uma ciência se resume na de seus termos específicos. Uma ciência só começa a existir ou consegue se impor na medida em que faz existir e em que impõe seus conceitos, através de sua denominação. Ela não tem outro meio de estabelecer sua legitimidade senão por especificar seu objeto denominando-o, podendo este constituir uma ordem de fenômenos, um domínio novo ou um modo novo de relação entre certos dados" [...] "Denominar, isto é, criar um conceito, é, ao mesmo tempo, a primeira e a última operação da ciência (BENVENISTE, 1989, p. 252).

A legitimidade de termos e conceitos, no entanto, não depende apenas de sua proposição. Por um lado, ela é assegurada pelos discursos que propõem reflexões de natureza teórica a partir de generalizações oriundas da prática; por outro, pela sua validação por parte da comunidade científica que constitui a área de especialidade em causa. A comunidade científica reconhece ou rejeita os textos – e o conjunto de conceitos e termos nele veiculados – submetendo as propostas ao debate e atribuindo-lhes valor. Essa é uma etapa para o ‘acordo’ a que se refere Peirce, acordo esse que, por sua vez, se relaciona à existência de um campo científico (BOURDIEU, 2004), lugar de duas formas de poder. O campo científico envolve:

- o capital científico e social ligado, principalmente, à ocupação de posições de destaque nas instituições científicas;
- o capital específico, que repousa sobre o reconhecimento pelos pares, mais exposto a contestações.

Assim, o acordo diz respeito ao reconhecimento, pelos pares, de uma terminologia (termos e conceitos, como forma e conteúdo indissociáveis) a partir da análise de proposições e argumentos que caracterizam a produção técnico-científica (artigos, teses etc.). A visibilidade e a fixação terminológica depende, por esse motivo, do jogo entre as instâncias institucionais e as de propriedade científica. Ao poder argumentativo do discurso que propõe uma série de conceitos se associa a inserção e reconhecimento dos proponentes no campo científico ou profissional. Se existe equilíbrio entre esses campos, de modo corolário, existe condição de fixação terminológica. Por sua vez, a fixação terminológica conduz à visibilidade.

Do ponto de vista do capital científico, podem ocorrer flutuações – reconhecimento, esquecimento, reconhecimento – porque esse é um campo de forças políticas. Já do ponto de vista do capital específico, trata-se de garantir uma produção propositiva, sustentada em argumentos, apoiada em metodologias sólidas. Este último aspecto solicita a exposição frequente das propostas ao debate por meio de publicações.

2 O SUBCAMPO DE ORGANIZAÇÃO E REPRESENTAÇÃO DO CONHECIMENTO

A dificuldade em traçar um quadro panorâmico dos estudos de Organização e Representação do Conhecimento decorre do estágio de seu desenvolvimento: mesmo que existam diferentes vertentes teóricas e, conseqüentemente, diferentes formas de observar os problemas, é necessário que seus termos e conceitos sejam explícitos e que correspondam a sistemas de conceitos coerentes. A dispersão terminológica (conceitual e denominativa) dificulta a comunicação e compreensão do quadro de propostas existentes.

A literatura brasileira usa simultaneamente os termos ‘Organização da Informação’ e ‘Organização do Conhecimento’: ora esses termos aparecem associados, ora são utilizados para falar de coisas distintas. Também é frequente a utilização do termo Representação, muitas vezes com um sentido próximo ou sinônimo de Organização da Informação.

Há um acordo ainda recente que se expressa na denominação atual do GT2, Grupo de Trabalho da ANCIB: ‘Organização e Representação do Conhecimento’. De 1994 a 2003 (cinco ENANCIBs – Encontro Nacional de Pesquisa e Pós-Graduação em Ciência da Informação), o Grupo foi identificado por uma justaposição de expressões – ‘Representação do Conhecimento/Indexação/Teoria da Classificação’ – cuja alteração ocorreu em 2005, quando se passou a utilizar a denominação ‘Organização do Conhecimento e Representação da Informação’. Em 2006, o Grupo passou a adotar o termo ‘Organização e Representação do Conhecimento’ (MARTELETO; LARA, 2008), denominação que se manteve nos encontros anuais subsequentes até o presente momento, em 2011.

A utilização de denominações paralelas no primeiro período dos ENANCIBs sinaliza fortemente a dispersão terminológica existente. Essa dispersão, no entanto, não afetou a reunião, nos eventos, de trabalhos com preocupações semelhantes. Resta saber os impactos provocados pela criação de novos Grupos como, por exemplo, em 2008, do GT8 – Informação e Tecnologia e, em 2009 e 2010, respectivamente, do GT9 – Museu,

Patrimônio e Informação e do GT10 – Informação e Memória, uma vez que suas ementas mostram níveis de superposição temática com o GT2¹.

Pode-se conjecturar que as alterações recentes das denominações do GT2 têm sua origem nos trabalhos da International Society for Knowledge Organization – ISKO para identificar o campo científico, nomeado ‘Knowledge Organization – KO’. Como ressalta Fujita (2008), a despeito de suas origens remotas na Teoria do Conhecimento, desde a Antiguidade, o reconhecimento da identidade e da conformação do campo científico ganharam destaque e fortalecimento a partir da criação daquela instituição.

Com a estruturação do capítulo brasileiro da ISKO, em 2007 (GUIMARÃES, 2008), abre-se a possibilidade, simultaneamente, de maior integração internacional e de reconhecimento da importância da pesquisa brasileira. A existência do capítulo brasileiro é também oportuna para projetar as pesquisas nacionais, muitas das quais têm origem francesa (compartilhando princípios com a pesquisa espanhola de origem semelhante), que por esse motivo diferencia-se, em muitos aspectos, da linha anglo-saxônica que marca a ISKO Internacional, o que pode significar a ampliação do panorama de tendências na área.

Se as iniciativas de integração são importantes, a identificação do quadro do vocabulário em torno do tema Organização e Representação do Conhecimento no Brasil é um passo indispensável para conhecer o estágio da terminologia sobre o tema. Sem esgotar o assunto, buscas preliminares realizadas na base BRAPCI – Base de Dados Referencial de Artigos de Periódicos em Ciência da Informação, que referencia principalmente a literatura brasileira, revelam grande variação terminológica na literatura coberta, conforme se pode observar a seguir:

Quadro 1 – Seleção de termos relativos à ‘Organização e Representação do conhecimento’, período 1970-2011, Base BRAPCI

TERMO	BUSCA POR PALAVRA-	BUSCA POR PALAVRA-CHAVE, RESUMO E
-------	--------------------	-----------------------------------

¹ A análise dos prováveis impactos não será objeto deste trabalho.

	CHAVE	TÍTULO
'Análise de assunto'	11	16
'Análise documental'	18	35
'Análise documentária'	4	23
'Classificação' (termo isolado e com qualificador)	87	190
'Comunicação documentária'	1	3
'Indexação'	119	193
'Informação documentária'	9	14
'Leitura documentária'	4	7
'Linguagem de indexação'	12	16
'Linguagem documental'	1	4
'Linguagem documentária'	32	41
'Linguística documentária'	7	7
'Organização da informação'	40	68
'Organização do conhecimento'	46	66
'Organização do conhecimento' e 'Organização da informação'	4	5
'Organização e representação da informação'	1	4
'Organização e representação do conhecimento'	2	14
'Representação documentária'	5	6
'Representação da informação'	20	46
'Representação do conhecimento'	36	58
'Representação do conhecimento' e 'Representação da Informação'	2	6
'Representação temática'	9	15
'Tratamento temático'	2	2
'Vocabulário controlado'	12	19

3 TERMOS E CONCEITOS ORIGINÁRIOS DA LINHA DE PESQUISA ANÁLISE DOCUMENTÁRIA

No âmbito das pesquisas desenvolvidas inicialmente pela Escola de Comunicações e Artes da Universidade de São Paulo, que deu origem ao Grupo Temma, fundado em 1986, as pesquisas que apresentam convergência com a Organização e Representação do Conhecimento, para usar o termo adotado pelo GT2, podem ser analisadas em torno dos termos a seguir: Análise Documentária, Leitura Documentária, Informação Documentária, Linguagem Documentária e Linguística Documentária. Dada a limitação de espaço no presente artigo, destacaremos apenas esses termos procurando reunir elementos para sua melhor compreensão.

Justificamos a escolha dos termos² por considerar que eles estão na base do desenvolvimento das pesquisas brasileiras desenvolvidas por integrantes do Grupo Temma³. Procuraremos nos reportar às origens de sua utilização acreditando na importância da recuperação do pensamento da vertente francesa, pouco conhecida pelos estudantes brasileiros, inclusive por problemas de idioma.

Salientamos, de início, que a preocupação que marca a vertente de origem francesa enfatiza os aspectos lógicos, linguísticos, semióticos, comunicacionais e terminológicos da atividade documentária, afirmando que é primordialmente na linguagem que se expressam os valores sociais, econômicos e culturais, ou em suma, simbólicos, que marcam as diferentes práticas das sociedades.

3.1 O termo Análise Documentária

O termo Análise Documentária⁴ tem sua origem nos trabalhos de Coyaud (1966) e Gardin (1966, 1968a; 1968b, entre outros) para designar as operações semânticas que transformam um texto original em uma ou várias palavras-chave, ou ainda, paráfrases, visando facilitar a representação de 'conteúdos' e a recuperação da informação. Para Coyaud (1966, p. 18), "a análise documentária é primeiramente uma operação de reconhecimento das unidades lexicais que representam as noções importantes de um documento".

Nos trabalhos iniciais de Gardin, a Análise Documentária é identificada com um tipo de análise descritiva cuja função é a de fornecer uma "representação sistemática de certos fatos que se supõem parcialmente ordenados" (GARDIN, 1968b, p. 84-86). Trata-se de:

[...] uma operação semântica, mesmo que ela não obedeça a uma regra precisa, e que cada organismo, cada analista [...] se limite a ver na ocorrência uma certa regularidade interna fundada mais na experiência ou hábito do que sobre algum procedimento explícito [...] (GARDIN, 1970, p. 631).

Gardin privilegiou a Análise Documentária destacando sua importância para a expressão do conteúdo de textos científicos (GARDIN, 1970), caracterizando a atividade como uma operação de extração de significados de textos que, por sua vez, são designados por símbolos que não necessariamente os encontrados nos textos de origem

² O presente artigo tem como ponto de partida nossa tese de Livre-Docência intitulada 'Linguística Documentária: seleção de conceitos', apresentada ao Depto de Biblioteconomia e Documentação da ECA-USP em 2009.

³ Grupo Temma no Diretório dos Grupos de Pesquisa no Brasil: <http://dgp.cnpq.br/buscaoperacional/detalhegrupo.jsp?grupo=0067607UVP2I76>

⁴ Variação no Brasil: Análise documental

(GARDIN, 1973a). O objetivo da Análise Documentária é isolar o sentido dos textos visando permitir pesquisas retrospectivas da informação a partir de seus conteúdos ou significação (GARDIN, 1973b).

Os investimentos metodológicos realizados pela Análise Documentária partiram do princípio de que a formalização dos procedimentos nessas operações poderia alterar o quadro empírico de representação no âmbito dos sistemas documentários. As operações de substituição de textos em Linguagem Natural por representações supõem necessariamente a existência de uma metalinguagem, que é constituída por unidades lexicais (símbolos designando noções ou conceitos) e convenções sintáticas (para expressão das relações lógicas presentes nos textos em Linguagem Natural).

Sob o plano semântico são consideradas as relações que são sempre válidas em um certo domínio de estudo, ou seja, as relações mais banais, aquelas que, nesse domínio, são consideradas adquiridas – o fato, por exemplo, que o cavalo é um animal... Sobre o plano sintático, ao contrário, aparecem sobretudo as relações novas, aquelas que fazem o objeto de um discurso porque não são estabelecidas no mesmo grau (GARDIN, 1973a, p. 76).

As funções normalizadoras dessa metalinguagem permitem operar semanticamente no tratamento de sinonímias, homotaxias, homografias, polissemias etc (GARDIN, 1970, p. 632).

Ao conceber a Análise Documentária como uma atividade eminentemente metalinguística, o autor confere 'autonomia' à linguagem de representação relativamente à Linguagem Natural, mesmo que seu estatuto seja distinto das metalinguagens definidas pela Linguística. Esse entendimento corrobora à caracterização da Linguagem Documentária como linguagem artificial cujo plano de conteúdo apresenta semelhança com a sua linguagem objeto (LARA, 2009).

A Análise Documentária se constituiu numa linha de pesquisa na Escola de Comunicações e Artes a partir de iniciativa de Johanna Wilhelmina Smit. A publicação do livro *'Análise Documentária: a análise de síntese'* (SMIT, 1987) oficializou o Grupo Temma reunindo pesquisadores em torno de discussões que mobilizavam a interface entre Documentação e Linguística, bem como subscreveu formalmente a denominação Análise Documentária.

As reflexões do Grupo acabaram por afirmar a condição da Análise Documentária como disciplina metodológica que sugere procedimentos para a análise de textos com o objetivo de selecionar conteúdos informativos que poderão ser representados, recuperados e disseminados (TÁLAMO, M.F.G.M.; LARA, M.L.G.; KOBASHI, N.Y. &

LIMA, V.M.A.). Embora a proposta original de Gardin dirigia-se explicitamente à análise de textos técnico-científicos, na apropriação brasileira esses procedimentos foram generalizados para outros tipos de textos, como por exemplo para a análise de imagem (SMIT, 1987) e de dados estatísticos (ROCHA, 2006), com adaptações.

Nos desenvolvimentos dos ensinamentos de Coyaud e Gardin, a proposta brasileira destacou o privilégio conferido ao texto frente ao discurso como objeto de análise, quer por uma razão econômica – seria impossível realizar uma análise do discurso, operação de expansão do texto, num contexto documentário – quer por considerar que é no texto que se manifestam as estruturas informacionais, elementos que respondem pela coesão e progressão textual e pelo seu fechamento e autonomia. Os elementos relativos às condições de produção são, nessa perspectiva, codificados pela representação descritiva, e não temática. “Enquanto no discurso se expressa o sentido – apropriação individual da significação articulada pelo sistema (BENVENISTE, 1989), no texto manifesta-se a significação como veículo de informação” (LARA, 1993, citado em LARA, 2009). Além disso, diferentemente da análise do discurso, a Análise Documentária persegue a síntese, como meio de diminuir a dispersão frente ao volume da produção.

No esforço de síntese, a Análise Documentária, observando políticas de indexação institucionais, procura identificar, no texto, as unidades ‘informativas’ do documento a partir da observação da articulação das formas estruturais, qualquer que seja o suporte físico em que tais unidades podem se manifestar: objetivos, metodologia, análise, conclusões etc. Nessa operação são mobilizadas a noção de tema (o que, quem, como, quando, onde) e a noção de estrutura temática dos documentos (KOBASHI, 1994), ou seja, sua forma de organização como orientadora para a localização de informações a serem destacadas.

A operação de Análise Documentária imprime nos produtos documentários uma redução crescente:

do texto para o resumo, do resumo para o enunciado, do enunciado para a unidade de tradução via código documentário. A atividade de Análise Documentária caracteriza-se, portanto, como uma sucessão de processos de transformação do texto original, observando-se, a cada etapa, graus crescentes de generalização. Nesse sentido, alguns procedimentos documentários podem ser considerados como uma modalidade específica de tradução (LARA, 1993, p. 41).

A redução é obtida por um processo de desestruturação do texto original seguida de uma operação de generalização: por condensação, se obtém o resumo, o enunciado, as palavras-chave; submetendo as reduções à intermediação da Linguagem Documentária são obtidas notações classificatórias ou descritores normalizados. Esse

processo sofre influências de variáveis que interferem na formulação da informação documentária, a saber (LARA, 1993, p. 42-43):

- 1) A estrutura do texto original;
- 2) As informações brutas presente no texto, ou informações de base (Vanoye, 1991);
- 3) O estado de sistematização metodológica e terminológica da área;
- 4) A instituição na qual o processo se desenvolve;
- 5) O usuário da informação documentária;
- 6) O estoque de conhecimento anterior do analista e sua formação ideológica.

A Análise Documentária caracterizou grande parte das pesquisas desenvolvidas na pós-graduação da ECA-USP, influenciando, posteriormente, a formação de graduação e de pós-graduação em outras instituições, como a UNESP-Marília. Focalizando modalidades textuais diferentes, áreas do conhecimento específicos ou partes do processo de análise propriamente ditos, sob a denominação Análise Documentária foram abrigados diferentes caminhos de pesquisa relativos ao tratamento da informação para a constituição de sistemas informacionais.

Atualmente, a Análise Documentária faz parte de um vasto conjunto dos estudos que podem ser incluídos no subcampo da Linguística Documentária que tende a integrar, por sua vez, o conjunto dos estudos de Organização e Representação do Conhecimento.

3.2 Leitura Documentária

O conceito de Leitura Documentária surge como proposta de aperfeiçoamento da Análise Documentária. A Leitura Documentária é uma modalidade específica no processo global de leitura a partir da qual se iniciam as operações de Análise Documentária propriamente dita.

Parte-se do princípio que, genericamente, o texto é, simultaneamente, “*um jogo de estratégia*” que fornece instruções para uma imagem de ‘*leitor modelo*’, e uma “*máquina preguiçosa*” que deixa ao leitor a tarefa de preencher os buracos dos ‘*não-ditos*’, ou seja, uma parte da construção do próprio texto (ECO, 1984, p. 97-99). No processo de leitura, o leitor negocia com o texto, aceitando alguns ‘*ditos*’, descartando outros, e preenchendo os

‘não-ditos’ a partir de sua experiência. Como um leitor-modelo genérico, ele interpreta o texto a partir de sua enciclopédia particular.

Na visão de Grice, desenvolve-se uma espécie de cooperação entre interlocutores que levam à uma comunicação mais eficiente dependendo da observação de máximas conversacionais: quantidade (volume de informação fornecida pelo locutor ao ouvinte); qualidade (veracidade das informações); relação (propósito da informação) e modalidade (forma pela qual se apresenta o dito quanto à clareza, ausência de obscuridade ou de ambiguidade, brevidade e método) (GRICE, citado por TÁLAMO, 1982).

Conforme Vanoye (1991), o processo de análise de um texto põe em relevo as informações referenciais – informações enumeráveis, sintetizáveis, brutas, objetivas – cujo reconhecimento na estrutura de organização do texto auxiliaria a interpretação das mensagens. Verón (1980), por outro lado, mostra que ao se analisar um texto analisam-se pelo menos dois, com isso querendo ressaltar que a significação não se reduz à investigação do conteúdo da mensagem. Pêcheux (1969), por sua vez, se refere a mecanismos linguísticos e não-linguísticos mobilizados na fala, que levariam à consideração do referente, do contexto, da situação na qual se insere o discurso, ou seja, suas condições de produção, ou seja, mecanismos por meio dos quais se daria a significação e o entendimento.

Considerando o processo global de leitura, é possível propor uma aproximação entre as noções de *‘estratégia de negociação leitor-texto’*, *‘princípio de cooperação’*, *‘identificação de elementos referenciais a partir de informações brutas’* e *‘condições de produção’*, para compreender o processo geral de leitura.

No entanto, a leitura para fins documentários se organiza de forma diferente. Frente ao documento, o leitor-documentalista não se caracteriza como um leitor-modelo, uma vez que não dispõe, necessariamente, de condições para estabelecer com o texto uma negociação. Do mesmo modo, enfrenta, em seu trabalho, textos que não obedecem as máximas conversacionais, e não raramente, como leitor não especialista, tem dificuldades para localizar rapidamente as informações ‘brutas’ do texto. O leitor-documentalista realiza uma leitura que se enquadra num processo de produção industrial de textos (parafraçando GARDIN), não podendo dedicar mais tempo à leitura do que aquele previsto na atividade de indexação de um grande volume de publicações.

Em suma, o leitor-documentalista tem de realizar uma leitura para localizar proposições no texto usando estratégias que, além das restrições de sua enciclopédia

particular (seu conhecimento sobre o tema do texto que está lendo), remete a propósitos oriundos da inserção de seu trabalho num espaço institucional determinado. Segundo Cintra (1983), concorrem à leitura documentária o nível de estruturação do texto (sua qualidade), o conhecimento prévio e as estratégias de leitura, o que vincula o processo de leitura à uma habilidade, à que se somam, segundo nosso ponto de vista, regras específicas relacionadas à normalização dos produtos documentários.

O processo de leitura documentária deve, então, ser apoiado em quadros de referência que permitam reconhecer o vocabulário das áreas e as formas mais usuais de organização textual. Conforme sugere Cintra (1983), a leitura na indexação também demanda o domínio de estratégias de leitura (estratégias cognitivas que respondem pelo processamento automático e inconsciente, e estratégias metacognitivas, relativas ao processamento consciente). Como lembra a autora, o autor não prevê o receptor documentalista como seu leitor, razão pela qual o princípio de cooperação, previsto por Grice, é rompido.

Em outras palavras, é necessário que o profissional receba treinamento específico para conhecer estratégias de leitura, para identificar tipologias textuais e para reconhecer a terminologia dos textos. Essas competências, somadas ao conhecimento dos objetivos institucionais e dos perfis de seus usuários, habilita-o a desempenhar a produzir a síntese documentária, meio para construir a informação documentária.

3.3 Informação documentária

Pesquisa preliminar mostra que o termo Informação Documentária aparece, em francês – *information documentaire* – em textos de Chaumier (1971; 1977), mas já havia sido utilizado pelos russos Mijailov, A.I., Chernii, A.I. e Quiliarevskii, R.S., em obra de 1968, conforme versão espanhola (*información documentaria*).

A informação documentária é o produto da Análise Documentária, ou seja, a representação de um texto para integrar sistemas documentário-informacionais. Pode assumir diferentes formas, caracterizando produtos documentários distintos, desde os mais simples – listas de palavras retiradas dos textos, palavras-chave, unitermos – até os mais elaborados como resumos e índices. Os produtos obtidos por extração, sem tratamento, não caracterizam verdadeiramente processos de Análise Documentária, ao contrário dos resumos e índices cuja elaboração requer o uso de procedimentos metodológicos específicos (LARA, 1993).

O resumo é uma representação construída a partir de condensação do texto original (resumo indicativo, informativo, crítico) que apresenta uma relação de contiguidade e semelhança com o texto que lhe deu origem valendo-se, em princípio, dos mesmos elementos do sistema semiótico do texto original (sua forma de organização). O índice é uma representação obtida por intermediação de uma Linguagem Documentária, espécie de código comutador que tem uma função nos níveis gramaticais e semânticos. Neste caso, a condensação é expressa pelos elementos do código de comutação exterior ao texto submetido à conversão e não apresenta necessariamente uma relação de contiguidade e semelhança com o texto original, envolvendo ao menos dois sistemas semióticos distintos: o presente no texto original e o estabelecido pela Linguagem Documentária.

A função do resumo é de substituir documentariamente o texto original; a função do índice obtido por intermediação de uma Linguagem Documentária, além de permitir a substituição da atribuição aleatória de termos de representação, orienta a construção dos pontos de acesso aos documentos num sistema documentário-informacional.

Em ambos os casos – e em especial quando se usa uma Linguagem Documentária – há redução semântica. Os resumos situam-se entre a generalização e a individualização, expressando uma tensão entre os dois pólos: procuram evidenciar o que, no texto, é informação específica, garantindo sobre a generalidade, ou ao que é comum, o particular, o individual. As representações via Linguagem Documentária caracterizam-se por uma normalização crescente do texto original: são construções desnaturalizadas (GARCÍA GUTIÉRREZ, 1990), sem a articulação que caracteriza a linguagem natural. As representações obtidas via Linguagem Documentária não recuperam informação, mas conjuntos de documentos. Por meio delas busca-se substituir a atividade aleatória de atribuição de termos por unidades de representação cujo significado esteja garantido pelo jogo mútuo de relações internas ao sistema (LARA, 1993).

3.4 Linguagem Documentária

O termo Linguagem Documentária corresponde a diferentes denominações que sinalizam distintas ênfases. A enumeração feita por Dodebei (2002) permite mostrar a variação de designações: linguagens documentárias (GARDIN, COYAUD e, acrescentaríamos, CHAUMIER); linguagens de indexação (MELTON); linguagens descritoras (VICKERY); codificações documentárias (GROLIER); linguagens de

informação (SOERGEL); vocabulários controlados (LANCASTER); lista de assuntos autorizados (MONTGOMERY); linguagens de recuperação de informação; linguagens de descrição da informação. Como salienta a autora, às variações correspondem diferentes orientações teórico-metodológicas relacionadas ao entendimento do que é informação, à delimitação de seus objetos de análise (assunto, informação, descritor), à estrutura de representação (código, linguagem, vocabulário, lista). Tais diferenças terminológicas também podem ser relacionadas aos níveis de abordagem da representação da informação ou ao foco privilegiado: objeto de análise, processos, produtos, instrumentos de representação.

Seguindo a linha francesa, os pesquisadores da ECA-USP e do Grupo Temma utilizam o termo Linguagem Documentária⁵, procurando ressaltar o caráter de ‘linguagem’ dos instrumentos de intermediação. Para resgatar as origens do termo, recuperamos as reflexões de Gardin.

Para Gardin, a Linguagem Documentária é um conjunto de termos, providos ou não de regras sintáticas, utilizadas para representar conteúdos de documentos técnico-científicos com fins de representação ou busca retrospectiva de informações. Ela integra três elementos fundamentais: um léxico ou lista de elementos descritores devidamente filtrados e depurados; uma rede paradigmática para traduzir certas relações essenciais e, geralmente estáveis, entre descritores (rede lógico-semântica que corresponderia ao que, *lato sensu*, poderíamos chamar de classificação); e uma rede sintagmática, para expressar as relações contingentes entre os descritores, relações essas que, diferentemente das relações paradigmáticas, são válidas apenas no contexto particular onde aparecem. As relações ditas sintagmáticas entre os termos permitem coordená-los para expressar um tema (GARDIN et al., 1968).

Coyaud (1966, p. 127) ressalta o papel das Linguagens Documentárias salientando sua função de “ferramentas mediatas entre usuários e autores dos documentos”, bem como de instrumento que opera a tradução dos textos escritos em linguagem natural para a linguagem do documento.

Gardin (1973a) caracteriza a Linguagem Documentária como metalinguagem, uma vez que ela reúne unidades lexicais (símbolos designando noções ou conceitos) e convenções sintáticas para a expressão das relações lógicas presentes nos textos em linguagem natural (GARDIN, 1970). Assim, ela exerce também a função de normalização,

⁵ Dentro do próprio Grupo Temma há pesquisadores que utilizam o termo Linguagem Documental.

inexistente como tal caracterização nas linguagens naturais, operando semanticamente no tratamento de sinonímias, homotaxias, homografias, polissemias etc.

As relações semânticas (que Gardin também denomina analíticas) exprimem correspondências entre as unidades do léxico, sendo que a diferença entre elas e as relações sintáticas (ou lógicas) só tem fins práticos: sob o plano semântico manifestam-se as relações comuns e estabilizadas; sob o plano sintático, as relações novas, objeto do discurso, portanto não estabilizadas. No campo da ciência passa-se progressivamente do plano sintático para o semântico, quando as relações sintáticas ganham estabilidade (GARDIN, 1973a; 1973b).

Gardin distingue entre as representações obtidas por métodos de 'tabulação' das obtidas por meio de uma metalinguagem. As primeiras não constituem representações propriamente ditas, porque são o resultado de escolhas realizadas diretamente sobre os textos. Segundo o autor, ao operar diretamente sobre os textos em linguagem natural esbarra-se no problema da língua (1973b), a menos que se postule a existência de uma linguagem semântica universal e, conseqüentemente, de características universais, "o que não deixa de ser uma ficção interessante sob o plano especulativo" (GARDIN, 1970, p. 633), em sua opinião.

Enquanto metalinguagem, a Linguagem Documentária é um:

[...] conjunto dos símbolos através dos quais serão expressas as equivalências e as diferenciações introduzidas na fraseologia de um corpus em LN dado, para que apareçam as aproximações desejadas entre as palavras, grupos de palavras, proposições etc. (GARDIN, 1970, p. 633).

Mas, para Gardin, diferentemente do que propomos atualmente, essa metalinguagem, que ele equivale a *information retrieval language*, é constituída no momento da Análise Documentária:

o jogo próprio da análise documentária engendra, à medida que se desenvolve, a constituição de um sistema de símbolos necessários para nomear, de uma maneira ou outra, os produtos mesmos do jogo, sob o plano lexical ou sintático" (p. 634) o que determina, em larga medida, a forma e o conteúdo dessas análises (GARDIN, 1970, p. 636).

Para Gardin, portanto, o conteúdo e a estrutura dessas metalinguagens são, em geral, inferidas dos textos em LN, o que, segundo ele, implica que metalinguagens diferentes apresentem semelhanças quando relacionadas a um mesmo universo de discurso, o que também acontece com os resultados da Análise Documentária feita por analistas diferentes (GARDIN, 1973a).

É de se notar que, já em 1973, Gardin (1973a; 1973b) refere-se aos estudos da semântica e da sintaxe em sua reflexão sobre tais metalinguagens, priorizando primeiramente os aspectos semânticos em detrimento dos sintáticos, para depois considerá-los de forma mais integrada. Como já se viu, para o autor a distinção entre sintaxe e semântica é contingente, sendo que as ferramentas relacionais das Linguagens Documentárias devem ser definidas independentemente dessa distinção. No âmbito da Linguística, o autor não confere importância à gramática de Chomsky que, segundo ele, não trouxe resultados práticos para a Documentação. Destaca, no entanto, os estudos de Katz e Fodor no início da década de 60 do séc. XX, que mostraram a impossibilidade de resolver certas dificuldades sem fazer apelo a dados semânticos, afirmando que enquanto os algoritmos gramaticais traziam resultados ambíguos para a interpretação das palavras dos textos, os ‘marcadores semânticos, combinados a ‘regras de seleção’ permitiam a quebra da ambiguidade.

Gardin também ressalta a importância do aparelho semântico na análise linguística citando as ‘árvores semânticas’, de Postal (1971), o ‘dicionário combinatoire-raisonné da língua russa’, de Mel’Chuk (1970) e as experiências com as ‘redes semânticas’ (GARDIN, 1973b). Para o autor, essas propostas têm muitas semelhanças com aquelas desenvolvidas quinze anos antes para a construção de Linguagens Documentárias, afirmando que para a

[...] passagem de um texto natural à sua representação metalinguística [...] os documentalistas [...] construíram redes semânticas ad hoc, onde se pode ver uma pré-figuração das redes semânticas da Linguística atual” (GARDIN, 1973b, p. 79).

Entre outras, o autor se referia ao SYNTOL – *Syntagmatic Organization Language*, desenvolvido por ele e sua equipe no início da década de 60 do séc.XX. A importância das redes semânticas também se relaciona à estruturação do enunciado, quando se substituem as categorias gramaticais por outras de natureza lógica (mais do que linguística), tais como Estado, Processo, Ação (GARDIN, 1973a). As funções gramaticais, no mesmo sentido, são substituídas por funções lógicas a partir, por exemplo, dos ‘casos’ de Fillmore. Na mesma linha, Gardin aponta a utilização de categorias – Entidade, Processo, Ação – a noção de ‘função’ (*role*) e os ‘operadores de função’ para as operações sintáticas nas linguagens documentárias utilizadas pela Documentação antes das propostas da Semântica (GARDIN, 1973b). Importante observação de Gardin relativa à ordem relativa da semântica e da sintaxe na análise se remete à Postal:

[...] a análise de textos, a análise Linguística, deve começar em um nível notadamente anterior às estruturas profundas”, um nível pré-lexical (formacional, cognitivo, para os autores), que se reporta ao plano metalinguístico remetendo não às palavras, mas às noções metalinguísticas (POSTAL, citado por GARDIN, 1973b, p. 81).

Como Gardin não separa o processo de construção da Linguagem Documentária da operação de Análise Documentária, suas analogias com as propostas de análise semântica e de redes semânticas dizem respeito ao processo global de representação. Ao ressaltar, por exemplo, a categorização semântica, preocupa-se, simultaneamente, com a construção de enunciados e com a própria Linguagem Documentária, o que apresenta semelhanças com as propostas de Ranganathan, na década de 30. A esse respeito, Gardin cita explicitamente as Classificações Facetadas elaboradas na década de 50 pelos ingleses do *Classification Research Group* com base nos trabalhos do pesquisador indiano, citando o uso de ‘funções universais’ para léxicos relativos a determinados universos de discurso – tal como Agente, Paciente, Instrumento, Objetivo, etc. -, bem como de limitadores funcionais de campo (*field bound roles*) que expressariam as relações sintáticas entre os enunciados.

As contribuições de Gardin para a reflexão sobre a Linguagem Documentária não se esgotam no que foi exposto. O arqueólogo francês, que se ocupou, entre a década de 60 a meados da década de 80, com problemas documentários – em interface com a Arqueologia – investiu, em seguida, na pesquisa sobre Sistemas Especialistas.

As primeiras discussões de pesquisadores brasileiros sobre a Linguagem Documentária no âmbito da Análise Documentária são realizadas por Smit (1973), ainda na França, sob orientação de Gardin. Ela estabelece um paralelo entre as definições de metalinguagem científica (Hjelmstev) e as de Análise Documentária, reportando-se à Gardin e à Coyaud, propondo que tal metalinguagem

[...] é um sistema construído, onde o plano de conteúdo é constituído por outro sistema, aquele da significação. Os termos empregados para descrever o sistema de significação constituem um corpo de definições coerente, isto é, eles foram previamente estudados em seguida (SMIT, 1973, p. 15-16).

A autora se pergunta se as Linguagens Documentárias podem operar como metalinguagens e se elas cumpririam, efetivamente, as tarefas de uma metalinguagem. A resposta à questão é remetida às leis e características do discurso científico, matéria-prima da Documentação. A condensação dos documentos, via processo de denominação, similar à codificação compreensiva das mensagens discursivas apontadas por Greimas – permitiria identificar “o movimento oscilatório entre a expansão, a condensação, a definição e a denominação (GREIMAS, 1976, p. 101) e o “*vai e vem entre a denominação*

e a *definição*” da atividade científica e da Documentação (SMIT, 1973, p. 22). Isso permitiria afirmar que na Documentação

[...] a equivalência entre grupos de lexemas pode existir e ser reconhecida, o que levaria a dizer que a operação denominativa é central na Documentação, podendo tomar como partida as diferentes unidades do discurso – “do tema à frase, passando pelo motivo, o enunciado ou as palavras[...]

sendo que essa escolha é que determina o sucesso ou fracasso de uma Linguagem Documentária (idem, p. 23-24). Smit também afirma existir uma ligação que une a linguagem utilizada na Documentação à sua metalinguagem, mas a associa com a análise, ou mais especificamente, o modelo de análise de Kristeva para quem elas seriam “*sistemas formais cuja estrutura é isomorfa ou análoga à estrutura de um outro sistema*” (KRISTEVA, 1969, citada por SMIT, 1973, p. 16).

Posteriormente, Smit (1976), utilizando o modelo de análise de Gardin, afirma que da ‘*descrição*’ e ‘*ordenação*’, operações básicas de representação, passa-se à ‘*interpretação*’, sendo esta última apoiada na identificação das relações lógico-semânticas existentes entre conceitos, operação que viabiliza a construção progressiva da metalinguagem de tradução. Verifica-se, portanto, que mesmo que se enuncie a questão da metalinguagem, ela ainda é muito centrada na sua ligação com o processo de Análise Documentária.

No início da década de 90, o tema específico da Linguagem Documentária é retomado por integrantes do Grupo Temma e, na ECA⁶. Inaugura-se nesse período a busca de referenciais da Terminologia e um maior investimento na Linguística Documentária, o que leva a concretizar o projeto de separação dos estudos para a construção da linguagem, daqueles relacionados ao processo de Análise Documentária. Com a Terminologia foram buscadas referências mais concretas para a organização das linguagens documentárias considerando áreas de conhecimento ou de atividade específicas.

A forte relação entre a Linguagem Documentária e as áreas técnico-científicas é, de fato, uma preocupação de Gardin, em especial no trabalho sobre o SYNTOL – *Syntagmatic Organization Language*. Mas as propostas para sua construção ainda se desenvolveram de modo concomitante à análise. Partimos do pressuposto, no entanto,

⁶ TÁLAMO, M. F. G. M.; LARA, M. L. G.; KOBASHI, N. Y.; LIMA, V. M. A. Instrumentos de controle terminológico: limites e funções. In: SIMPÓSIO LATINO-AMERICANO DE TERMINOLOGIA, 2.; ENCONTRO BRASILEIRO DE TERMINOLOGIA TÉCNICO-CIENTÍFICA, 1., 1990, Brasília, DF, 1992. *Anais*. Brasília, DF : Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT). p. 364-370.

que o *corpus*, ao ser constituído à medida em que os textos vão sendo indexados, pode não ser suficientemente representativo da organização conceitual da área.

Posição intermediária é sugerida por Fujita (1990) que, em pesquisa experimental para a construção de um tesauro de Odontologia sugere, apoiando-se em Sudarsham, que o processo combine duas frentes de trabalho: a primeira, quando uma equipe desenvolve e propõe um tesauro preliminar que é, em seguida, aplicado por uma segunda equipe que, por sua vez, fornece elementos para as adaptações necessárias.

De qualquer modo, García Gutiérrez já havia afirmado que a “Documentação apresenta dois *corpus* que se imbricam com a linguística: a Análise Documentária e a Linguística Documentária” (GARCÍA GUTIÉRREZ, 1990, p. 24-25), com isso sugerindo provavelmente que a segunda proveria elementos para a primeira.

Ao propor a separação do processo de Análise Documentária do processo de construção da Linguagem Documentária passa-se a defender, considerando experiências práticas, que as etapas de análise documentária e de construção de linguagens documentárias mobilizam procedimentos diferentes. Evidência dessa perspectiva é confirmada pela publicação, em 1986, da norma ISO 2788 - ‘*Guidelines for the establishment and development of monolingual thesauri*’, na qual foram propostos princípios específicos para a elaboração de um gênero sofisticado de Linguagem Documentária, desenvolvendo-se, a partir daí, inúmeros trabalhos sobre construção de tesouros.

A separação dos processos de análise documentária e de construção de linguagens documentárias pode ser justificada pela própria natureza do índice quando comparado ao resumo. Resumo e índice, ambos produtos da Análise Documentária, envolvem sistemas semióticos diferentes, sendo este último um produto altamente generalizante, conforme já observamos. O texto original, relativamente ao índice, sofre um processo de desautenticação, já que é representado por idéias, conceitos, dispostos em classes categoriais, não mais na sua individualidade (LARA, 1993). Por meio do índice, portanto, não são recuperados os textos originais, mas um conjunto de textos que compartilham de conceitos semelhantes (CINTRA, A.M.M.; TÁLAMO, M.F.G.M.; LARA, M.L.G. de & KOBASHI, N.Y, 1994). Como afirmou García Gutiérrez (1990), as Linguagens Documentárias tradicionais promovem, via de regra, construções desnaturalizadas. Assim, o índice, isolado do seu universo original de referência, pode não garantir convenientemente sua interpretação.

Para recuperar a condição de interpretação (e significação), além de relacionar os termos entre si, como recomenda a norma ISO 2788, há que se procurar referências para a significação. Não é raro verificar que nas Linguagens Documentárias tradicionais as noções (símbolos de classificação, cabeçalhos de assunto ou descritores) são fragilmente definidas: ou está ausente o princípio de arranjo (sistema de relacionamento entre as unidades), ou não definições, ponto de partida para compreender e justificar os relacionamentos. Por esse motivo, buscou-se a Terminologia como meio de rederendar a significação a partir dos universos de discurso onde os termos foram utilizados.

Outro argumento que fundamenta a proposta de separação dos processos acima apontada é a natureza da Linguagem Documentária. Desde Gardin, opõe-se a Linguagem Documentária à Linguagem Natural. Esse caráter salienta seu aspecto construído, o que leva muitos autores a caracterizá-la como Linguagem Artificial, o que é verdade, muito embora preferimos identificá-la como intermediária, posto que compartilha simultaneamente elementos das duas linguagens: da Linguagem Natural e da Linguagem Artificial.

As distinções entre a Linguagem Natural e a Linguagem Artificial permitem verificar os limites da Linguagem Documentária enquanto sistema de significação e de comunicação e identificar os principais investimentos necessários para o aperfeiçoamento das metodologias para sua construção. Os quadros a seguir (LARA, 2009) ressaltam alguns aspectos que as aproximam e as afastam.

Quadro 2: Principais diferenças entre a Linguagem Natural, a Linguagem Artificial e a Linguagem Documentária

LING. NATURAL	LING. ARTIFICIAL	LING. DOCUMENTÁRIA
Formas escritas secundárias em relação às formas orais.	Formas escritas fundamentais	Formas escritas fundamentais
Os lexemas não têm significação unívoca.	Os significados são altamente padronizados. Univocidade de significação.	Os significados devem ser fixados visando a univocidade de interpretação, embora não se possa garanti-la totalmente.
Dinâmica	Estática	Estática
Usa-se a língua para falar dela mesma.	Necessário usar a LN para falar da LA.	Necessário usar a LN para falar da LD.
Tem regras de fato. Não é necessário conhecer as regras da língua para usá-la.	Regras de jure. Fundamental obedecer às regras. Linguagem altamente simbólica.	Regras de jure. Fundamental obedecer às regras, embora reconheça-se semelhanças com as palavras da LN. Pode integrar elementos da LA.
Não tem uma função específica, mas funciona em muitos contextos e para diferentes objetivos.	Elaborada para desempenhar funções específicas.	Elaborada para desempenhar uma função particular: representar a informação visando sua recuperação.
Tem inúmeras funções: função expressiva (estética), descritiva (informativa, referencial), valorativa, prescritiva.	Desempenha principalmente a função descritiva (informativa, referencial) e a função prescritiva.	Desempenha principalmente a função descritiva (informativa, referencial) e a função prescritiva.
Tem produtividade (1ª. e 2ª. articulações).	Não tem produtividade.	Não tem produtividade.
Pouco formalizada.	Altamente formalizada.	Relativamente formalizada.
Alto poder combinatório.	Médio poder combinatório.	Baixo poder combinatório.

Fontes: CINTRA, A. M. M. Elementos de linguística para estudos de indexação. *Ciência da Informação*, Brasília, v. 12, n. 1, p. 5-22, 1983.

LARA, M.L.G. de. *A representação documentária: em jogo a significação*. São Paulo : Escola de Comunicações e Artes - USP, 1993. (Dissertação de Mestrado)

HUTCHINS, W.J. *Languages of indexing and classification*. Herts : Peter Peregrinus, 1975.

Quadro 3 - Principais semelhanças entre a Linguagem Natural, a Linguagem Artificial e a Linguagem Documentária

LING. NATURAL	LING. ARTIFICIAL	LING. DOCUMENTÁRIA
Lexemas têm forma e significado.	Símbolos têm forma e significado.	Descritores têm forma e significado.
Estrutura paradigmática e sintagmática.	Estrutura paradigmática e sintagmática.	Estrutura paradigmática e sintagmática.
Paralelismo entre lexemas.	Símbolos funcionam como espécies de lexemas.	Unidades correspondem a lexemas ou combinação de lexemas (sintagmas).

Fontes: CINTRA, A. M. M. Elementos de linguística para estudos de indexação. *Ciência da Informação*, Brasília, v. 12, n. 1, p. 5-22, 1983.

LARA, M.L.G. de. *A representação documentária: em jogo a significação*. São Paulo : Escola de Comunicações e Artes - USP, 1993. (Dissertação de Mestrado)

HUTCHINS, W.J. *Languages of indexing and classification*. Herts : Peter Peregrinus, 1975.

As diferentes características das linguagens justificam investimentos metodológicos diferentes. Enquanto instrumento comutador nas operações de representação para compor sistemas informacionais, a Linguagem Documentária tem a função primordial realizar a intermediação entre linguagens. Por esse motivo, suas unidades devem: ter significado relativamente fixado e estabelecido *por 'regras de jure'* (CINTRA, 1983); ser formuladas com a intenção de funcionar como meio de comunicação nos sistemas documentários; ser razoavelmente formalizadas, dispondo de mecanismos que, na ausência de articulação e de produtividade, permitam serem geradas 'frases documentárias' que possam ser interpretadas.

A comparação entre a Linguagem Natural e a Linguagem Artificial permite compreender que o caráter intermediário da Linguagem Documentária não porque apenas porque funciona como ponte entre sistema documentário e usuário, mas principalmente porque conjuga elementos de ambos os tipos de linguagem. Essa complexa natureza da Linguagem Documentária faz com que seus interpretantes corram o risco de remeter, ora à Linguagem Natural, ora à Linguagem Artificial.

A Linguagem Natural é dinâmica: os lexemas, ou signos linguísticos, não têm significação unívoca e dependem da inserção contextual para o desenvolvimento de interpretantes; já na Linguagem Artificial, que é estática, os significados são altamente padronizados e unívocos, ou seja, seus interpretantes são mais localizados ou, a rigor, não seria pertinente falar em interpretantes. Por sua vez, na Linguagem Documentária os significados devem ser fixados ou controlados, mas como não raras vezes seus descritores são expressos de modo idêntico à Linguagem Natural, não se pode falar em univocidade estrita.

Na base do conceito de Linguagem Documentária está o conceito de ordem e de ponto de vista. Os critérios de ordenação devem partir do pressuposto de que a interpretação e a compreensão são adesivas. Os vínculos de adesão são de ordem lógico-cognitiva, linguístico-comunicacional, pragmática e terminológica.

Segundo critérios lógico-cognitivos, considera-se que a compreensão mobiliza o reconhecimento por semelhança (vínculos estáveis) – dedução, indução, generalização e categorização – e por associação espaço-temporal (vínculos circunstanciais).

Na ótica dos critérios linguístico-comunicacionais, enfatiza-se que a significação é contextual, devendo ser observada a partir do sistema em que é gerada. A noção linguística de valor permite explicar a diferença entre uma lista de palavras e um tesouro, por exemplo, em que os termos são mutuamente relacionados. O arranjo, ou sistema de relações entre os termos, confere estrutura a partir da qual as unidades de uma Linguagem Documentária podem significar e comunicar. Na base do arranjo estão conceitos relativos à linguagem – signo, sintagma/paradigma, forma/conteúdo, sincronia/diacronia – a partir dos quais a Linguagem Documentária funciona enquanto sistema de significação e de comunicação. “A Linguagem Documentária representa a informação como consequência das relações instituídas entre suas unidades” (TÁLAMO, 1997).

Do ponto de vista dos critérios pragmáticos, fica em evidência a função da informação e, de modo correlativo, o reconhecimento de princípios compartilhados, de enciclopédias locais (conhecimento especializado) e da experiência colateral (conhecimento dos indivíduos). Os critérios pragmáticos se associam aos terminológicos quando são observadas as referências de linguagem e de terminologia das comunidades discursivas (organização dos universos temáticos em referências mais consensuais: linguagens de especialidade). Como afirmou Sager (CABRÉ, 1993, p. 14), *‘estudiar una*

material equivale a aprender los lenguajes de esa materia". A linguagem de uma disciplina inclui sua terminologia e, por sua vez, a organização conceitual da área de conhecimento ou de atividade. Essa é a razão pela qual o apoio nas terminologias (produtos terminológicos como dicionários, glossários) é fundamental para a construção das Linguagens Documentárias, já ela fornece referências para a materialização dos vínculos de linguagem, de significação e adesão. Na Linguagem Documentária, o conhecimento da linguagem de especialidade se expressa não só nas designações como na forma de organização e estruturação dos descritores.

Toda ordenação têm origem em hipóteses de organização. A explicitação dessa ordenação no plano de classificação da Linguagem Documentária é realizada a partir de estruturas de relacionamento. O fundamento do relacionamento dos termos, quer do ponto de vista lógico, linguístico, pragmático ou terminológico, é que só o relacional é explicativo. Por meio dos relacionamentos os termos são reunidos, ordenados por superordenação e subordinação, coordenação e associação. Ao relacionar os termos parte-se do princípio que um termo não comporta significação nele mesmo. A significação supõe necessariamente a existência de relações.

Quanto mais os instrumentos de classificação e indexação têm as características acima, mais se aproximam da caracterização de uma linguagem, muito embora comumente todos eles sejam agrupados sob a denominação 'Linguagem Documentária'. Nessa perspectiva, pode-se identificar basicamente três tipos de instrumentos documentários: as classificações bibliográficas, as listas de cabeçalho de assunto e os tesouros. Ficam fora dessa classificação as listas alfabéticas de assunto, bem como as feitas por extração por não possuírem qualquer ordenação. Cada Linguagem Documentária determina, a seu modo, a leitura dos documentos a serem indexados.

3.5 Linguística Documentária

A Linguística Documentária constitui um subcampo da Ciência da Informação que tem como objetivo estudar os problemas que caracterizam a Linguagem Documentária como uma forma específica de linguagem inscrita no universo da linguagem geral. Originalmente, o termo Linguística Documentária foi o proposto por García Gutiérrez (1984; 1990) para estudar a formação convencional das linguagens fechadas destinadas à transmissão da mensagem documentária por meio de elementos significantes. Sob sua perspectiva, interessa à Linguística Documentária não a língua ou a linguagem na comunicação em geral, mas a comunicação no âmbito dos processos científicos e

informativos estabelecidos por meio de documentos. No âmbito dos estudos da Linguística Documentária, o interesse social prevalece sobre o individual.

Para identificar os problemas da Linguística Documentária e enunciá-los, García Gutiérrez (1984; 1990) reconhece a importância do sistema linguístico documentário, ou seu '*corpus* linguodocumental', buscando mobilizar conhecimentos da Linguística, da Terminologia, da Semântica, da Gramática aplicada à gestão da informação, como em campos relacionados, como a análise do discurso, a análise do conteúdo e, de modo geral, das Ciências Cognitivas.

O trabalho de García Gutiérrez também tem sua origem nas propostas de Jean-Claude Gardin. O ponto comum entre o seu trabalho e o desenvolvido por pesquisadores do Grupo Temma é a procura do entendimento do funcionamento da linguagem para o tratamento da informação, quer no aspecto metodológico de construção do instrumento mediador, quer na elaboração de produtos documentários que funcionem na comunicação documentária, ou comunicação em ambientes documentário-informacionais.

A proposição da Linguística Documentária como campo de estudos no Brasil conheceu seus desenvolvimentos específicos. Mantendo-se os objetivos de substituir a atividade empírica de representação de conteúdos por outra, via Análise Documentária e uso da Linguagem Documentária, esse desenvolvimento tem focalizado, com maior ênfase, a construção do instrumento de mediação, mobilizando para isso as referências, linguístico-semióticas e estabelecendo um diálogo mais frequente com a Terminologia. Reconhece-se que o tratamento e a disseminação requer o apoio da linguagem construída para tratar a informação e disponibilizá-la para o acesso e apropriação, considerando os problemas relacionados ao caráter simbólico da informação, as particularidades da mediação documentária e as possibilidades de prover a circulação social da informação.

A hipótese que baliza os estudos da Linguística Documentária funda-se no reconhecimento da noção de estrutura linguística, do caráter semiótico da linguagem e dos processos de estruturação da mensagem documentária, que associam os dados da produção informacional e os quadros de referência fornecidos pelas terminologias das áreas de especialidade e de atividade que caracterizam os discursos dessas comunidades. A Linguagem Documentária, concebida a partir desses parâmetros, constitui-se em meio para o estabelecimento da cultura informacional que, na sociedade da informação, mais do que exercer um papel codificador, se propõe como representação

e funciona como insumo do processo social de geração de sentido (TÁLAMO; LARA, 2006). Sua organização se articula, também, com hipótese sobre o modo de organização dos objetos, uma decupagem que articula o conjunto de referências da produção, do contexto e do uso. A construção da Linguagem Documentária persegue a construção de um discurso documentário que visa a interação, supondo uma recepção em que os sujeitos produzem sentido.

Assim, cabe à Linguística Documentária

[...] compor os quadros de referência para a análise, avaliação e construção da linguagem documentária entendida como linguagem de informação, associando os níveis sintático-semântico-pragmático para identificar com clareza a inserção do signo documentário no pano sistêmico e no plano funcional, objetivando-o no tempo, no espaço e na cultura (TÁLAMO; LARA, 2006, p. 206).

As questões próprias da Linguística Documentária são relacionadas às especificidades da linguagem nos fluxos informacionais promovidos pelas atividades da Ciência da Informação. Todas essas questões têm origem no modo como se concebe a informação no interior do campo da Ciência da Informação, cujos traços distinguem o conceito de outros usos mais genéricos, ou ainda específicos, ligados a campos do conhecimento. O conceito de informação (TÁLAMO; LARA, 2006) se caracteriza como:

- algo intencionalmente construído a partir da análise da produção do conhecimento e dos objetivos institucionais;
- algo que se apresenta sob forma específica – uma mensagem documentária, como um produto documentário-informacional;
- algo que instaura uma relação comunicativa particular, a comunicação documentária.

Parte-se do princípio que a informação não é um dado, mas uma construção, sendo sua transmissibilidade condicionada às condições de aderência definidas como vínculos de significação (TÁLAMO; LARA, 2006). Decorrem dessa definição parâmetros metodológicos para a abordagem da produção técnico-científica e sócio-cultural e sua representação que buscam observar as referências sócio-cognitivas e de linguagem das comunidades a que se destinam os produtos informacionais. Operacionalmente, isto compreende a seleção da documentação, uma linguagem de representação adequada baseada nas terminologias das áreas, políticas de indexação que resultem na proposta de indicadores de conteúdo, bem como a definição das formas pelas quais são apresentadas as informações e os recursos para facilitar a busca.

A abordagem interpretativa da produção, como ato que não esgota seu significado, é proposta como modelização que combina objetivos das políticas institucionais que compreendem, simultaneamente, as referências da emissão e da recepção, e que se segue à primeira modelização promovida pela língua natural. Por sua vez, a perspectiva da dimensão reflexiva da atividade linguístico-comunicacional qualifica as mensagens documentárias como signos documentários que visam desencadear processos de negociação do sentido. A informação documentária, por sua vez, como produto documentário, é um indicador que encaminha o usuário à produção propriamente dita.

Em termos de seus fundamentos, da Linguística e das teorias da linguagem, a Linguística Documentária destaca os aspectos relacionados ao caráter relacional da linguagem, sua natureza arbitrária e sua autonomia. Operacionalmente, a noção de valor, de diferença e de identidade parametrizam a construção da Linguagem Documentária. Da Semiótica, destaca o processo onde algo funciona como signo no processo de significação e no processo de comunicação. São mobilizados os conceitos de signo, interpretante, objeto, contexto e intérprete. Operacionalmente, a Linguística Documentária propõe o signo documentário e ressalta as características da semiose documentária. Da Terminologia, se apropria das reflexões teóricas e metodológicas, bem como observa as terminologias concretas dos domínios. Por essa via, observa indiretamente os discursos produzidos pelas áreas do conhecimento e de atividade. Do ponto de vista operacional, a Terminologia auxilia a identificação e a seleção dos termos e conceitos que compõem universos temáticos, suas definições, o mapeamento do sistema de conceitos segundo perspectivas compartilhadas e as condições de seu relacionamento decorrentes das definições. Permite, assim, construir mapas conceituais, identificar e decidir sobre designações bem como realizar equivalências na linguagem de especialidade em face à linguagem natural, quando esses universos forem muito distintos.

4 CONSIDERAÇÕES FINAIS

Este artigo não esgota, em absoluto, a recuperação de características que respondem pela formação dos conceitos selecionados. Seria preciso empreender um grande esforço de análise da produção francesa, brasileira, e também espanhola, para contemplar um número maior de aspectos que permitiriam melhor caracterizar os conceitos e justificar as denominações utilizadas. Não foram referidos de modo suficiente quer os trabalhos dos pesquisadores da USP, como os da UNESP, que integram o Grupo Temma.

Acreditamos, porém, que o panorama oferecido pode contribuir para um mapeamento do estado da arte relativo aos conceitos de 'Organização e Representação do Conhecimento' no Brasil, tema deste número da revista, mapeamento esse importante para proporcionar maior visibilidade das pesquisas, bem como encaminhar projetos de harmonização de conceitos.

Ao concluir o artigo, acreditamos ser importante assinalar o que consideramos ser a característica mais marcante desse universo de pesquisa: a preocupação com o desenvolvimento de metodologias específicas para a atividade documentária. O delineamento dos conceitos e a proposição de sua terminologia se apóia em apropriações interdisciplinares, buscando fundamentar o trabalho a partir da consideração da impossibilidade de uma linguagem universal para a representação dos textos, da rejeição da noção de representação como reprodução, do questionamento da noção de conteúdo como algo objetivamente identificável, da necessidade de observar a diversidade da produção e dos públicos da informação e, conseqüentemente, dos seus distintos universos simbólicos e condições de apropriação.

REFERÊNCIAS

BENVENISTE, É. *Problemas de lingüística geral, II*. Campinas: Pontes; Ed. UNICAMP, 1989.

BOURDIEU, P. *Os usos sociais da ciência: por uma sociologia clínica do campo científico*. Trad. Denice B. Catani. São Paulo: Editora UNESP, 2004.

BRAPCI: Base de Dados Referencial de Artigos de Periódicos. Disponível em: <<http://www.brapci.ufpr.br/index.php>>. Acesso em: 02 de jul. de 2011.

CABRÉ, M.T. *La terminología: teoría, metodología, aplicaciones*. Barcelona: Ed. Antártida; Empúries. 1983.

CINTRA, A.M.M. Estratégias de leitura em documentação. In: SMIT, J.W. (org.). *Análise documentária: a análise da síntese*. Brasília: IBICT, p. 28-35, 1983.

CINTRA, A.M.M. Elementos de linguística para estudos de indexação. *Ciência da Informação*, Brasília, v.12, n.1, pp. 5-22, 1983. Disponível em: <<http://revista.ibict.br/ciinf/index.php/ciinf/article/view/1526/1144>>. Acesso em: 10 de jul. de 2011.

CINTRA, A.M.M.; TÁLAMO, M.F.G.M.; LARA, M.L.G. de & KOBASHI, N.Y. *Para entender as Linguagens Documentárias*. São Paulo: APB/Polis, 1994.

COYAUD, M. *Introduction a l'étude des langages documentaires*. Paris: C. Klincksieck, 1996.

DODEBEI, V. L. D. L. M. *Tesouro: linguagem de representação da memória documentária*. 1. ed. Rio de Janeiro : Interciência: Niterói :Intertexto, 2002.

ECO, U. *Conceito de texto*. Trad. de Carla de Queiroz. São Paulo: T.A. Queiroz; Ed. Universidade de São Paulo, 1984

FUJITA, M. Organização e representação do conhecimento no Brasil: análise de aspectos conceituais e da produção científica do ENANCIB no período de 2005 a 2007. *Tendências da Pesquisa Brasileira em Ciência da Informação*: v. 1, n.1, 2008. Disponível em: <<http://inseer.ibict.br/ancib/index.php/tpbci/article/viewFile/4/1325>>. Acesso em: 09 de mar. 2009.

FUJITA, M.S.L. Linguagem documentária em odontologia: uma aplicação do sistema de indexação "PRECIS". *Actas do II Simpósio Latino-americano de Terminologia*, 1990. Disponível em: <<http://www.riterm.net/actes/2simposio/indice90.htm>>. Acesso em 09 de mar. de 2009.

GARCÍA GUTIÉRREZ, A. L. G. *Linguística documental: aplicación a la documentación de la comunicación social*. Barcelona: Ed. Mitre, 1984.

GARCÍA GUTIÉRREZ, A. L. G. *Estructura lingüística de la documentación: teoría y método*. Murcia: Ed. Universidad de Murcia, 1990.

GARDIN J-C. Eléments d'un modèle pour la description des lexiques documentaires. *Bulletin des Bibliothèques de France*, v. 11, n. 5, mai 1966, p. 171-182.

GARDIN, J.-C. Análise documentária e análise estrutural em arqueologia; trad. de J. T. Coelho Netto. In: CORDIER, S., ed. *Lévi-Strauss*. São Paulo : Documentos, 1968a.

GARDIN, J.-C. Procédures d'analyse sémantique dans les sciences humaines. In: POUILLON, J. et MARANDA, P. (orgs.). *Échanges et communications: mélanges offerts à Claude Lévi-Strauss à l'occasion de son 60ème. anniversaire*. (Tirage à part). Paris: Mouton, p. 628-657, 1970.

GARDIN, J.-C. Document analysis and linguistic theory. *The Journal of Documentation*, v. 29, n. 2, pp.137-168, June, 1973^a.

GARDIN, J.-C. Linguistique et documentation. *Bollettino d'Informazioni*, v.13, n. 2/3, pp. 67-85, 1973b.

GARDIN, J.-C. e outros. *L'automatisation des recherches documentaires: un modèle général "Le SYNTOL"*. 2. ed. revue et augmentée. Paris: Gauthier-Villars, 1968b

GREIMAS, A.-J. (1976). *Semântica estrutural: pesquisa de método*. São Paulo: Cultrix; Universidade de São Paulo, 1976.

GUIMARÃES, José Augusto Chaves. A dimensão teórica do tratamento temático da informação e suas interlocuções com o universo científico da International Society for Knowledge Organization (ISKO). *Revista Ibero-americana de Ciência da Informação (RICI)*, v.1 n.1, p. 77-99, jan./jun. 2008. Disponível em: <<http://www.red.unb.br/index.php/RICI/article/view/2761/2331>>. Acesso em: 09 de mar. de 2011.

HUTCHINS, W.J. *Languages of indexing and classification*. Herts : Peter Peregrinus, 1975.

ISO 2788. *Documentation - Guidelines for the establishment and development of monolingual thesaurii*. Génève: International Standard Organization, 1986.

KOBASHI, N.Y. *A elaboração de informações documentárias: em busca de uma metodologia*. 1994. Tese (Doutorado)-Escola de Comunicação e Artes, Universidade de São Paulo, São Paulo, 1994.

LARA, M.L.G. *A representação documentária: em jogo a significação*. 1993. Dissertação (Mestrado)-Escola de Comunicação e Artes, Universidade de São Paulo, São Paulo, 1993.

LARA, M.L.G. *Linguística Documentária: seleção de conceitos*. 2009. Tese (Livre-Docência apresentada ao Depto de Biblioteconomia e Documentação)- Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2009.

MARTELETTO, R. M.; LARA, M. L. G. de. Os Grupos de Trabalho – GTs da ANCIB e a promoção da pesquisa em Ciência da Informação. In: FUJITA, M. S. L. F. ; MARTELETTO, R. M.; LARA, M. L. G. de (Orgs.). *A dimensão epistemológica da ciência da informação e suas interfaces técnicas, políticas e institucionais nos processos de produção, acesso e disseminação da informação*. São Paulo: Cultura Acadêmica; Marília: Fundepe, p. 3-15, 2008.

PEIRCE, C.S. *Semiótica*. Trad. José Teixeira Coelho Netto; rev. J. Guinsburg. São Paulo: Perspectiva, 1977. (Estudos; 46).

PÊCHEUX, M. *Analyse automatique du discours*. Paris: Dunod, 1969.

ROCHA, S. G. *A representação documentária da informação estatística: quando a ordem dos fatores altera o produto*. 2006. Dissertação (Mestrado)- Escola de Comunicação e Artes, Universidade de São Paulo, São Paulo, 2006.

SAGER, J.C. Prólogo: la terminología, puente entre varios mundos. In: CABRÉ, M.T. *La terminología: teoría, metodología, aplicaciones*. Barcelona: Ed. Antártida; Empúries. p. 11-17, 1993.

SMIT, J.W. *Les langages documentaires comme metalangages du discours scientifique*. Paris: École Pratique des Hautes Études, 1973. (Mémoire)

SMIT, J.W. (NATALI) *De l'analyse documentaire à l'analyse poétique: a propos des critiques des "Chats" de Baudelaire*. Paris: École Pratique des Hautes Études en Sciences Sociales, 1976. (Thèse de 3e. Cycle)

SMIT, J.W., (Coord.) *Análise Documentária: a análise da síntese*. Brasília: IBICT, 1997.

TÁLAMO, M.F.G.M. *Linguagem documentária*. São Paulo: APB, 1997 (Ensaio APB; 45)

TÁLAMO, M.F.G.M. As máximas conversacionais de Grice: In: _____. *Comunicação e funcionamento da linguagem*. 1982. Dissertação (Mestrado)-Escola de Comunicação e Artes, Universidade de São Paulo, São Paulo, 1982.

TÁLAMO, M. F. G. M. ; LARA, M. L. G. O campo da Linguística Documentária. *Transinformação*, v. 18, p. 203-211, 2006. Disponível em: <<http://revistas.puc-campinas.edu.br/transinfo/viewissue.php?id=14>>. Acesso em: 09 de mar. de 2011.

TÁLAMO, M.F.G.M.; LARA, M.L.G.; KOBASHI, N.Y. & LIMA, V.M.A. Instrumentos de controle terminológico: limites e funções. *Actas do II Simpósio Latino-americano de*

Terminologia, 1990. Disponível em: <<http://www.riterm.net/actes/2simposio/indice90.htm>>. Acesso em 09 de mar. de 2011.

VANOYE, F. *Usos da linguagem: problemas e técnicas na produção oral e escrita*. Trad. e adapt. de Clarisse Madureira Sabóia e outros. 8. ed. São Paulo: Martins Fontes, 1991.

VERÓN, E. *A produção do sentido*. Trad. De Alceu Dias Lima e outros. São Paulo : Cultrix, 1980. Brasilat, 1980.

Title

Organization Concepts and Knowledge Representation from the perspective of Temma Group's reflections

Abstract

It relates the visibility of science and its terminology, and presents evidence for the delimitation of the concepts of Knowledge Organization and Representation from the perspective of Temma Group's reflections, retrieving the terms Documentary Analysis, Documentary Reading, Documentary Information, Documentary Language and Documentary Linguistics. It justifies the selection arguing about the need to refer to the origins of the terms, most of them from the French strand of Documentation. It concludes emphasizing the importance of further research on the formation of concepts in the area to provide greater visibility to the research or to forward projects of harmonization of concepts.

Keywords

Knowledge organization and representation. Documentary analysis. Documentary linguistics. Documentary reading. Documentary information. Documentary language. Tema group.

Título

Conceptos de organización y representación del conocimiento bajo la perspectiva de las reflexiones del Grupo Temma

Resumen

Este trabajo relaciona la visibilidad de la ciencia con su terminología y presenta elementos para delimitación de conceptos de Organización y Representación del Conocimiento según la óptica de las reflexiones del Grupo Temma, recuperando los términos Análisis Documental, Lectura documental, Información Documental, Lenguaje Documental y Lingüística Documental. Justifica la selección argumentando respecto a la necesidad de que sean reportados los orígenes de los términos, en su mayoría provenientes de la vertiente francesa de Documentación. Concluye resaltando la importancia de aumentar las investigaciones sobre la formación de los conceptos en el área, tanto para proporcionarle mayor visibilidad a las mismas como para encaminar proyectos que armonicen conceptos.

Palabras clave

Organización y representación del conocimiento. Análise documental. Lingüística documental. Lectura documental. Información documental. Lenguaje documental. Grupo temma.

Recebido em: 15/08/2011

Aceito em: 07/12/2011