

O Unicórnio (o Rinoceronte, o Ornitorrinco ...), a Análise Documentária e a Linguagem Documentária*The Unicorny (the Rhinoceros, the Ornithorhynchus ...), the Documentary Analysis and the Documentary Language*por [Marilda Lopes Ginez de Lara](#)

Resumo: Os processos de análise documentária e de construção de linguagens documentárias apresentam semelhanças relativamente ao que ocorre no processo de conhecimento. Ambos são caracterizados pela analogia e generalização, operações realizadas a partir de hipóteses inferenciais formuladas em face a referências anteriores disponíveis em nossa cultura. Tendo como ponto de partida a discussão do texto Marco Polo e o unicórnio, de Umberto Eco, propomos-nos a discutir as semelhanças e diferenças que caracterizam, de um lado, o processo de conhecimento e, de outro, a construção da informação documentária, quer no processo de análise dos textos, quer na construção de linguagens intermediárias de representação da informação. Para efeito de maior clareza, introduzimos alguns conceitos da Terminologia, da Lingüística e da Semiótica, que permitem caracterizar as operações de recorte de conteúdo e mostrar que a informação é uma construção, não um dado apriorístico. Procuramos mostrar que, também na atividade documentária, a noção de conteúdo se relaciona, em larga medida, aos objetivos funcionais e pragmáticos do sistema documentário. Salientamos que a especificidade da atividade documentária de representação de conteúdos está no seu propósito - organizar para transferir, transferir para permitir a apropriação da informação - razão pela qual é necessário estabelecer vínculos que permitam instaurar uma situação de comunicação documentária. A construção desses vínculos apoia-se em referências concretas fornecidas pela terminologia do domínio em foco e pelas referências de uso, que combinadas, permitem viabilizar a interpretação e a apropriação das informações documentárias.

Palavras-chave: análise documentária; linguagem documentária; informação documentária; terminologia; lingüística; semiótica

Abstract: The documentary analysis process, as well as the documentary languages construction process, have many similarities with the knowledge process. Both of them use analogy and generalization, operations that are carried out by inferential hypothesis formulated from cultural references. Starting with the discussion of Umberto Eco's Marco Polo e o unicórnio, we intend to identify the similarities and differences between the processes of documentary analysis and construction of intermediate representation languages and the knowledge process. To ease comprehension, we introduce some concepts of Terminology, Linguistics and Semiotics, to characterize the operations of contents segmentation and to show that the information is a construction, not a apriorist fact. We intend to show that, also at the documentary activity, the contents notion is related to the functional and pragmatic goals of the documentary system. We emphasize that the documentary activity specificity is related to its aims - organize to transfer and transfer to lead to information appropriation - and that's why it is necessary to identify the interactivity conditions that allow documentary communication. This operation must be supported by concrete terminological references of the area in focus and by references of usage that, together, may make the information interpretation and appropriation possible.

Keywords: documentary analysis; documentary language; documentary information; terminology; linguistics; semiotics

Introdução

O texto *Marco Polo e o unicórnio*, do livro "Kant e o ornitorrinco", de Umberto Eco (ECO, 1998), como tantos outros do autor, nos permite realizar uma leitura exploratória tecendo um paralelo entre o processo de conhecimento e o processo de representação da informação documentária. A idéia básica discutida por Eco refere-se à experiência de Marco Polo, em uma de suas viagens, quando se depara com um animal desconhecido e tenta classificá-lo (e nomeá-lo) a partir de sua experiência.

Vale ressaltar que, um texto, na concepção do autor, é uma máquina preguiçosa que solicita a participação do leitor na compreensão dos ditos e no preenchimento dos não ditos. Conseqüentemente, toda leitura é parcial, corroborando a esse processo inúmeros fatores, dentre os quais motivação e os objetivos da leitura, as experiências anteriores do leitor e o seu estoque prévio de conhecimento. Na leitura que faremos objetivamos reter alguns aspectos que nos parecem pertinentes para discutir os processos de classificação, de representação, de construção e uso da linguagem documentária que contribuem à construção da informação. Mais do que realizar uma análise do texto, pretendemos utilizá-lo como ponto de partida para uma discussão sobre os processos de Análise Documentária e de construção de uma Linguagem Documentária.

1. A analogia na construção das representações

O processo de conhecimento se realiza fundamentalmente através da analogia. Diante de um fenômeno desconhecido "reagimos por aproximação, procuramos aquele recorte de conteúdo, já presente na nossa enciclopédia, que bem ou mal parece prestar contas do novo fato" (ECO, 1998, p.55). Quando Marco Polo se defronta, em Java, com animais nunca vistos (hoje, rinocerontes), realiza um esforço para identificá-los (e nomeá-los) consultando seu estoque de conhecimentos anterior: lançando mão de características de animais que já conhece e a partir das descrições disponíveis em sua cultura, tenta estabelecer relações e, por aproximação, classifica o novo a partir de referências existentes. Se de imediato identifica-os aos unicórnios - seu corpo, as

quatro patas, um chifre acima do nariz, etc., - rapidamente verifica que a classificação feita não é inteiramente própria, uma vez que, nessa comparação, constata a existência de diferenças relativamente ao modelo anterior.

De fato, os animais que se lhe apresentam agora "não são brancos e ágeis, mas têm 'pêlo de búfalos e pés de elefantes', o chifre é negro e desgracioso, a língua espinhosa, a cabeça semelhante àquela de um javali: 'É uma besta muito feia de se ver. Não é, como se diz por aqui, que ela se deixe pegar como uma donzela, mas acontece ao contrário' (Milione 143, apud ECO, op.cit.). Tais características não condizem exatamente àquelas do unicórnio presentes em sua enciclopédia, motivo pelo qual precisa prosseguir na sua tentativa de classificação. Neste momento, deve tomar uma decisão: ou segmenta novamente o conteúdo sugerindo um novo animal, ou modifica a intensão dos unicórnios alterando sua descrição anterior para acomodar os novos animais.

Se Marco Polo optasse por modificar a intensão dos unicórnios para poder dar conta dos animais que agora conhecia, acrescentaria características ao conceito original. Esse acréscimo resultaria, entretanto, na perda de especificidade dos unicórnios. A intensão do conceito seria um amálgama indiferenciado.

Processo similar experimentamos ao construir representações de conteúdo na área da Ciência da Informação. Operamos sempre por analogia e generalização, procurando reunir os conceitos a partir de seus traços comuns, numa tentativa de organizar a informação e de garantir sua recuperação posterior. Ao classificar (e indexar) segmentamos o conteúdo a partir de referências que já possuímos, formando agrupamentos em função de suas propriedades comuns, ou mais exatamente, das características que julgamos pertinentes para os nossos propósitos, já que nunca esgotamos as propriedades das coisas e fenômenos. Frente ao que desconhecemos, ou não conhecemos suficientemente, encontramos dificuldades para realizar segmentações e formar grupos de termos.

Para que as coisas fiquem mais claras, retomemos alguns conceitos acima verificando seu significado sob a ótica da Terminologia[1], antes de prosseguir na análise :

- . objeto: tudo o que pode ser percebido ou concebido;
- . conceito: unidade de conhecimento criada pela combinação de características únicas;
- . característica: propriedade abstrata de um objeto ou conjunto de objetos que servem para descrever os conceitos;
- . intensão (ou compreensão): conjunto de características constituintes de um conceito
- . extensão: totalidade dos objetos aos quais corresponde um conceito

FONTE: ISO 1087-1:2000(E/F)

Também é útil tentar precisar alguns conceitos lingüísticos e semióticos:

. recorte de conteúdo:

Segundo Benjamim Whorf, cada língua "recorta a realidade de um modo particular", segundo sua cultura. Essa afirmação complementa a proposta anterior de Saussure, segundo a qual a língua é forma, não substância. Conseqüentemente, as distintas línguas não são um decalque ou rotulação da realidade, nem variações de expressões com significados universalmente válidos e estáveis (LOPES, 1987). Isso explica porque, enquanto no português temos sete termos para expressar as cores do arco-íris, em bassa, língua indígena da Libéria, existem apenas duas. Explica, também, porque o esquimó tem várias formas para designar a neve, quando temos apenas uma.

. continuum da realidade:

Expressão utilizada por Eco para referir-se ao que Hjelmslev nomeou "sentido", fator comum entre as línguas, pensamento provisoriamente apresentado como massa amorfa, grandeza não analisada. Lembre-se que Hjelmslev postulou a solidariedade entre as funções de expressão e conteúdo, afirmando que é impossível existir uma sem a outra, o que implica dizer que não há pensamento sem linguagem. A proposta de Eco visa dirimir as confusões que o termo sentido poderia gerar. O continuum, na visão do autor, seria essa grandeza provisoriamente não analisada - a realidade - que cada língua segmenta a seu modo, segundo seus padrões culturais.

. enciclopédia / dicionário:

. dicionário:

Para compreender o conceito de enciclopédia é necessário, anteriormente, compreender o que é dicionário. A interpretação em termos de dicionário concerne apenas ao conhecimento lingüístico (ou do léxico da língua - mais as regras desse código lingüístico). O conhecimento lingüístico remete aos sentidos registrados em dicionário. (Por exemplo, a palavra "missionário", no Dicionário Aurélio, tem pelo menos três acepções). O sentido dicionarial permite dar conta de fenômenos como a sinonímia, paráfrase, semelhança e diferença entre termos, antonímia,

hiponímia e hiperonímia (subordinação e superordenação), como também identificar anomalias semânticas, redundância, ambigüidade provocada por homônímia (expressões idênticas com significados distintos). Permite, também, constatar verdades analíticas, concisão, inconsistência inclusão e implicitação semântica, a partir da relação entre afirmações efetivamente presentes num enunciado (ECO, 1991, p.84-85).

. enclopédia:

A interpretação em termos de enclopédia ultrapassa os limites do dicionário, uma vez que mobiliza o conhecimento extralingüístico, o conhecimento adquirido a partir do uso de determinadas expressões ou enunciados em determinados contextos e circunstâncias, ou o conhecimento resultante da experiência vivida. Por exemplo, o conhecimento enclopédico permitiria associar, ao sentido dicionarial de "missionário", todas as conotações conferidas à palavra em função de seu uso em situações específicas, como qualificar o trabalho voluntário como "missionário", já que ele pode apresentar várias das características que permitem aproxima-lo do trabalho religioso.

A enclopédia é o conjunto de "todas as interpretações, concebíveis como a biblioteca das bibliotecas, onde uma biblioteca é também um arquivo de toda a informação não verbal de algum modo registrada, das pinturas rupestres às cinematecas" (ECO, 1991, p.113). Mas ela é um postulado semiótico, ou seja, não é passível de descrição na sua totalidade, já que as possibilidades interpretativas são infinitas, às vezes contraditórias e implicam continuamente novas segmentações no continuum da realidade. A enclopédia é uma espécie de competência global constituída de diversas enclopédias parciais que são ativadas, a medida do necessário e individualmente, no processo interpretativo. Enquanto conjunto não ordenado de marcas é passível de ser ordenado parcialmente, sob forma de dicionário, "toda vez que queremos circunscrever a área de consenso dentro da qual um discurso se move (idem, p.133). É o conhecimento enclopédico, e não apenas lingüístico, que nos permite interpretar convenientemente um enunciado do tipo "O cão do meu vizinho uivou a noite toda".

Retomando nossa análise, verificamos que as informações presentes na enclopédia de Marco Polo não permitiam dar conta adequadamente da descrição dos novos animais, porque as características do unicórnio ali retidas não eram suficientemente extensas. Ao tentar corrigir a descrição inicial dos unicórnios, Marco Polo "modifica a intensão[2] deixando a extensão sem juízo", ou seja, altera a definição original incluindo nela novas características, mas descaracteriza a classe que permitiria acomodá-los. A modificação da intensão do termo unicórnio pelo acréscimo (ou eliminação) de características, interfere na extensão do termo, do que resulta que nem os unicórnios, nem os novos animais, podem pertencer, de direito, à nova classe resultante, pois os traços não são específicos nem de um, nem de outro e não servem mais para descrevê-los num mesmo conjunto.

A única alternativa que resta é a de proceder a uma nova segmentação do conteúdo: a partir da observação dos novos animais Marco Polo pode acrescentar um novo animal ao universo dos seres vivos.

Caso mais difícil de resolver teria enfrentado Marco Polo se tivesse chegado à Austrália e tivesse visto um ornitorrinco. "O ornitorrinco é um estranho animal, que parece concebido para desafiar qualquer classificação...: tendo em média uns cinquenta centímetros, cerca de dois quilos, o corpo chato coberto por uma pelagem marrom-escura, sem pescoço, e com uma cauda de castor, bico de pato, de cor azulada por cima e rosa ou matizada por baixo, sem pavilhões auditivos, as quatro patas terminam com cinco dedos espalmados, mas com garras; fica bastante tempo debaixo d'água (e ali come) para ser considerado um peixe ou um anfíbio, a fêmea põe ovos, mas amamenta os próprios filhotes, mesmo se não vemos nenhuma teta (além disso, não vemos no macho sequer os testículos, que são internos)." (p.55-56). Frente a esse animal tão estranho, além da dificuldade de classificá-lo quer como mamífero, como anfíbio ou mesmo peixe, as características do unicórnio, mesmo um pouco modificadas, não se prestariam à descrição. O estranho animal fugiria aos vários modelos.

Fica claro a partir deste texto que os problemas relacionados ao conhecimento derivam, por raciocínio hipotético, do conhecimento de fatos externos e dos acontecimentos precedentes (PEIRCE apud ECO, 1991, p.56). Eco sugere que se Peirce analisasse o caso talvez tivesse afirmado que Marco Polo, antes de propor um novo animal, teria realizado uma bricolage de noções precedentes, descrevendo-o a partir de entidades já conhecidas. Com efeito, o trabalho classificatório procede por sucessivas aproximações, num processo inferencial e a partir de hipóteses. As inferências se desenvolvem por sucessivas aproximações e reformulações à regra, até o ponto em que se verifica que elas são insuficientes para dar conta do novo: é quando nasce um novo modelo. No esforço de compreensão do mundo buscamos a estabilidade. Se as coisas não cabem nas categorias que já dispomos experimentamos um profundo desconforto, pois é da natureza humana a necessidade de ordenar, catalogar e hierarquizar o mundo. Qualquer novo modelo é, também, uma hipótese e, embora no processo classificatório procuremos invariavelmente pela segurança da exclusão e da inclusão, tal como num leito de Procuste, é necessário admitir a instabilidade e a possibilidade de coexistência de modelos.

O processo de conhecimento é, por natureza, semiótico. A semiose caracteriza-se como um percurso interpretativo que se define pela passagem contínua de signo a signo. A interpretação mobiliza informações conhecidas num processo hipotético e inferencial. Sob este ponto de vista, deve-se recusar a idéia de que a realidade está dada e refletida num espelho imperfeito, solicitando-nos seu contínuo polimento (RORTY, 1988). Ao contrário, o conhecimento (e a interpretação) se desenvolve a partir de nossa experiência, o que leva a afirmar que está sempre em construção. Enquanto processo de natureza hipotética é evidente que selecionamos algumas qualidades (propriedades) que nos são mais salientes (pela experiência ou necessidade) e, a partir delas, fazemos generalizações.

Com efeito, o processo de representação em Documentação se desenvolve de forma semelhante, pois é fruto de segmentações do conteúdo: ou fazemos uso de classificações existentes (se elas forem consideradas "adequadas" para acolher mais membros), ou, ainda, segmentamos mais uma vez o conteúdo propondo uma nova hipótese de organização. Se tentarmos modificar a intensão dos objetos ou fenômenos anteriormente classificados, podemos incorrer no erro da descaracterização das classes ou na produção de amalgamas indiferenciados, como ocorre muitas vezes na acomodação de novos conceitos em classificações rigidamente hierárquicas. De qualquer modo, o processo de organização persegue invariavelmente "modelos" orientadores que permitam comunicar e compartilhar nossas idéias.

2. A noção de conteúdo na representação da informação

Afirmamos correntemente que no processo de representação documentária representamos conteúdos, mas já é possível perceber, a partir do que foi dito anteriormente, que a noção de conteúdo é absolutamente imprecisa. É preciso verificar o que entendemos por conteúdo e que conteúdo representamos no contexto documentário quando utilizamos uma linguagem documentária. Assim como a segmentação de conteúdo feita por Marco Polo se fazia a partir de referências anteriores, no processo documentário "representar conteúdos" se relaciona a propósitos e concepções definidas de antemão.

No universo documentário tradicional a informação era considerada um dado, fato que marcou fortemente a história das linguagens documentárias. A concepção de representação subjacente remetia à crença de que poderia existir uma relação de fidedignidade entre o conteúdo de um documento e a expressão utilizada para "representá-lo" (LARA, 1991). Uma concepção do gênero tem várias consequências perigosas:

- a) Ao considerar a informação como um dado, a classificação só pode ser reprodução. Essa crença está fundada na preocupação com a busca da "essência" ou, dito de outro modo, da informação objetiva. Tal procedimento remeteria ao "polimento do espelho", já que a informação estaria dada aprioristicamente, requerendo um esforço para resgatá-la do documento;
- b) A ativação daquilo que é considerado "conteúdo do documento" é feita a partir de uma combinatória universal - uma classificação única válida para quaisquer contextos. Os sistemas bibliográficos universais reproduziriam a realidade e permitiriam, por esse motivo, que se representasse fielmente os conteúdos dos textos.

Já a partir de uma perspectiva mais pragmática, a informação é vista sob outro ângulo, para o que corrobora o texto de Umberto Eco. A informação nunca é um dado, ao contrário, é sempre uma construção, porque:

- a) As generalizações a que procedemos são um produto cultural, fruto de nossa experiência com a realidade ;
- b) Enquanto produtos culturais, as classificações (e generalizações) partem necessariamente de hipóteses; podem existir, concomitantemente, distintas hipóteses; enquanto hipóteses, privilegiam determinadas características em detrimento de outras. Sob esse aspecto, uma classificação não pode ser avaliada como certa ou errada, mas mais, ou menos adequada para determinados propósitos, para o que se selecionam determinados aspectos e não outros. Dito de outra forma, não se trata de polir o espelho para aceder à informação, mas de construí-la em função de objetivos. Sob essa perspectiva, compreendemos que qualquer proposta de classificação universal é datada, ou seja, parcial, momentânea e centrada num ponto fixo de enunciação. Isso explica porque as classificações bibliográficas de natureza encyclopédica^[3] e quase que exclusivamente hierárquicas oferecem dificuldade para dar conta de distintos aspectos não contemplados de início. Tais propostas classificatórias têm uma articulação dura entre o plano do conteúdo e o plano de expressão e funcionam exclusivamente a partir de operações de encaixe.
- c) A informação envolve sempre uma relação de pertença ao que nos rodeia (GADAMER, apud NUNES, 1998, p.10). Para haver informação é preciso que ela faça sentido para alguém, ou seja, que exista algum nível de compartilhamento a partir do qual se estabeleça um vínculo para a interpretação e o sentido (BAITELLO, 1988).

No âmbito documentário, portanto, representar conteúdos para constituir informação significa recortar, segmentar. Como qualquer segmentação, ela não é neutra, mas tributária de hipóteses de organização. Além disso, conscientes ou não, desde que usamos a língua estamos utilizando convenções: a língua é o primeiro sistema modelizante a partir do qual todos os outros sistemas derivam. Também não é necessariamente unívoca a interpretação do usuário, sujeito que também tem suas convicções e está apto a formular hipóteses, tanto quanto o produtor do sistema de representação. Com efeito, para que possamos falar de transmissão de informação documentária é necessário formular hipóteses (e explicitá-las) que sejam razoavelmente compartilhadas pela comunidade de usuários do sistema que propomos.

A noção de conteúdo, portanto, não é simples. Sob a ótica do enunciador, a constituição de sistemas de informação para transferência é uma operação carregada de intencionalidade. A representação não tem um fim em si mesma, ao contrário, é uma construção que persegue objetivos específicos. Num determinado contexto uma vaca é um tipo de mamífero; noutro, um bicho sagrado e etc. Contrariamente à visão tradicional que distingue as informações pelas características dos documentos (documentos de bibliotecas, de museus etc.), são suas funções informativas desses documentos que devem ser privilegiadas (SMIT, 1988, p.4-5) para a constituição de sistemas. Necessidades pragmáticas, portanto, determinam os recortes (segmentações) do conteúdo, de forma a procurar viabilizar um fluxo informativo efetivo. Conseqüentemente, um mesmo universo documental pode ser organizado e representado de diversas maneiras, pautando-se nos distintos propósitos a que se destinam.

3. A função da Terminologia (e das terminologias) na determinação das hipóteses de organização da informação nas Linguagens Documentárias

Se se aceita o que foi dito, resta afirmar que a construção de hipóteses de organização dos sistemas de informação, ou mais especificamente, das Linguagens Documentárias, não pode ser um processo aleatório. Embora se possa afirmar que as propostas de organização têm ingredientes empíricos, isso só pode ser entendido como fundado em necessidades pragmáticas. Dito de outro modo, uma linguagem documentária que será utilizada para "representar conteúdos" não pode ser formulada a partir de opiniões ou referências particulares e isoladas. Ao contrário, se o objetivo dos sistemas documentários é transmitir informação, tais linguagens devem ser organizadas a partir de hipóteses que façam sentido para determinados objetivos.

No caso de Marco Polo, a hipótese utilizada para proceder à classificação se baseava num conjunto de formulações definidas pela história de sua cultura à época. Nos dias de hoje, quais as referências para formular hipóteses para a organização da informação? Durante muito tempo, a referência exclusiva foi a garantia literária. Acreditava-se que o uso dos termos presentes nos textos a serem indexados seria condição suficiente para garantir a representatividade dos textos em foco. As experiências mostraram que os inventários construídos com base na freqüência e ocorrência de palavras, mesmo se submetidos a algum tratamento semântico para depurá-los dos problemas decorrentes dos processos de extração, não se traduziam como conjuntos organizados. Verifica-se ainda nos dias de hoje que o tratamento semântico característico, por exemplo, das listas de cabeçalhos como a LCSH - Library of Congress Subject Headings, não é formulado rigorosamente, quer em relação ao controle da sinonímia, quer na distinção entre relações hierárquicas e associativas (WELLISCH, H., 1991, p.388). Problema não menos grave caracteriza muitos inventários construídos empiricamente pelas bibliotecas (que imaginam possuir uma base de referência constituída por uma série de cabeçalhos que devem apenas ser traduzidos), onde a noção de controle é restrita à verificação da sinonímia e ao estabelecimento de referências cruzadas (os veja também). Hoje, entretanto, a Terminologia teórica e concreta permitem fundamentar a seleção dos termos e o estabelecimento de redes de relacionamento entre eles a partir de referências concretas de determinados universos conceituais.

É preciso salientar que o simples uso dos termos de uma área não constitui um trabalho terminológico e não garante, por conseqüência, que o sistema conceitual do domínio do saber ou da área de atividade (que se sustenta por uma hipótese) esteja representado. Tal procedimento resulta da visão ingênua que associa o significado dos termos de uma Linguagem Documentária com algo que está fora dela, como se o simples agrupamento de palavras típicas fosse suficiente para caracterizar uma linguagem (TÁLAMO, 1997). Se uma linguagem é um sistema (ou estrutura), a determinação do sentido de cada termo se estabelece na relação mútua entre os termos do mesmo conjunto. Uma linguagem pressupõe a noção de arranjo, e esse arranjo se desenha relativamente a uma proposta de organização.

A Terminologia trabalha a partir do conceito. Um conceito, como já vimos, é uma unidade de conhecimento formada por uma combinação de características. A formação de um conceito é um processo de abstração que tem como base a seleção de propriedades (características) de um conjunto de um ou mais objetos. O conceito, conseqüentemente, não tem vida independente, mas é o resultado de uma proposição feita dentro de um domínio

do saber. Cada conceito é determinado pela posição que ocupa no conjunto estruturado de conceitos, denominado sistema conceitual.

O trabalho terminológico é normalmente desenvolvido para a construção de dicionários terminológicos, glossários, bancos de dados terminológicos etc. , mas pode servir a distintos propósitos: é referência fundamental à tradução, como também pode ser muito útil como apoio à estruturação das linguagens documentárias.

A unidade básica da Terminologia é o termo, designação verbal de um conceito dentro de um domínio específico. Uma designação é uma representação de um conceito por um signo que o denota (um símbolo, um nome ou um termo). Para

elaborar um sistema de conceitos, enquanto conjunto estruturado de acordo com as relações que os unem, o trabalho terminológico parte de definições. A definição é a representação de um conceito por um enunciado descritivo que permite diferenciá-lo de conceitos relacionados dentro de um domínio de especialidade. Um domínio é um campo especializado do conhecimento (áreas do saber ou de atividades) expresso por uma língua de especialidade (LSP[4] , ou língua utilizada pelo domínio e caracterizada pelo uso de meios de expressão lingüística particulares). Enquanto subconjunto do sistema lingüístico, as línguas de especialidade refletem um modo mais ou menos normalizado de se comunicar num dado domínio do saber, o que seria impossível na língua geral onde os sentidos são variáveis.

As definições podem ser intensionais ou extensionais. A definição intensional (ou por compreensão) descreve a intensão do conceito indicando o conceito superordenado e as características distintivas do conceito em foco. Dito de outro modo, faz menção ao conceito genérico mais próximo, já definido ou supostamente conhecido, e às características distintivas que delimitam o conceito a ser definido. Já a definição extensional (ou definição por extensão) descreve o conceito enumerando todos os conceitos subordinados segundo um critério de subdivisão[5]

O trabalho terminográfico é realizado sobre um conjunto de documentos que, criteriosamente, são selecionados como representativos do universo do domínio enfocado: dicionários e glossários já existentes, textos produzidos na área dos quais são criteriosamente coletados termos-candidatos. Normalmente, a coleta é realizada utilizando-se um suporte formal - fichas terminológicas - onde se transcrevem literalmente as definições formais presentes na documentação, ou os contextos onde aparecem os termos-candidatos. As fichas terminológicas são, em seguida, analisadas, comparando-se as características presentes nas definições formais ou nas ocorrências de uso dos termos. Esse procedimento fundamenta a seleção dos termos considerados mais apropriados para representar os conceitos, possibilita a identificação de variantes, e apoia a construção ou reconstrução de definições. Embora nem todos os trabalhos terminológicos cheguem efetivamente a apresentar mapas conceituais dos domínios (muitas vezes o trabalho se encerra na elaboração de um dicionário ou de um glossário), é a partir do processo definicional que são criadas as condições para se estabelecer uma rede de relações entre os termos[6].

Os dicionários terminológicos constituem, assim, uma ferramenta imprescindível para a construção da linguagem documentária, já que permitem conferir referência aos descriptores. Se antes a Documentação trabalhava empiricamente a partir de palavras (através dos processos de extração baseados em freqüência ou ocorrência), ou da seleção empírica de 'unidades significativas' (sem a definição daquilo que poderia caracterizar uma unidade significativa), a partir da Terminologia pode-se contar com instrumentos que trabalham com o termo, unidade que representa o conceito dentro de um domínio ou área de atividade.

Lembre-se que a palavra é uma unidade do léxico de uma língua; a representação do léxico constitui o dicionário da língua. O empreendimento de organizar sistematicamente o léxico de uma língua é muito difícil, senão inviável, pelo seu tamanho e pela dificuldade em dar conta das várias possibilidades de sentido[7] . Já o termo é uma unidade da terminologia. Seu significado é fixado no interior de um domínio, ou seja, corresponde à palavra contextualizada no uso (no discurso) de um determinado campo do saber (LE GUERN, 1989). O significado de um termo é delimitado pelas relações que estabelece com os outros termos do mesmo domínio (fixado no processo definicional, para facilitar a comunicação).

O uso do termo (e de mapas conceituais de domínios) permite à Documentação transformar uma atividade anteriormente empírica, de categorização, num trabalho teórica e metodologicamente sedimentado em referências concretas dos domínios de especialidade. O uso da Terminologia teórica e das terminologias concretas cria condições, conseqüentemente, para formular propostas de organização das linguagens documentárias à medida que provê fundamentos para a seleção e estruturação dos termos.

4. A linguagem documentária não é terminologia: objetivos do sistema de informação e referências de uso para a formulação das hipóteses de organização

Embora o uso da Terminologia seja crucial para referendar o processo de construção de linguagens documentárias - e com isso possibilitar melhor representação e recuperação da informação -, ela não garante por si só o sucesso da comunicação documentária. Isso significa dizer que as hipóteses de organização de tais linguagens são mais complexas do que poderia parecer inicialmente, uma vez que a atividade de informar requer, além dos parâmetros de um sistema conceitual, a observação das necessidades de uso ou recepção da informação.

A linguagem documentária, portanto, não é a simples reprodução da terminologia de uma determinada área. Num tesouro, que dentre as linguagens documentárias se destaca por apresentar uma boa solução para estruturação do eixo paradigmático^[8] , as unidades são, antes, preferenciais. Os descritores são termos^[9] de um tesouro que devem ser empregados para representar uma noção (ou conceito) contida(o) num texto ou numa equação de busca; os não-descritores, termos que não devem ser empregados, mas incorporados à linguagem como referência de acesso a um descritor a ser utilizado em seu lugar.

O termo preferencial é escolhido como descritor entre um conjunto de termos equivalentes (ISO 5127/6-1983-E/F). A definição apontada na norma não esclarece quais critérios devem ser utilizados para a preferência, porém pode-se considerar que eles remetem simultaneamente à garantia literária (como o conceito é designado com maior freqüência na literatura) e à garantia do uso (como os usuários procuram por informação^[10]). Embora não existam orientações precisas para preferir ou preterir um termo, é lícito supor que o critério pragmático (uso pela literatura e pela comunidade de usuários) deve ser perseguido, para o que deveria existir metodologia. A existência de um critério pragmático, embora venha sendo tratado de maneira empírica, é expressão evidente de que os sistemas documentários, construídos para transmitir informação, não podem se descolar das referências de uso. Convém investigar, dentre outras possibilidades, como a Socioterminalogia poderia contribuir para definir procedimentos de identificação e incorporação de tais referências à organização das linguagens documentárias.

A linguagem documentária também não é terminologia porque ambas se propõem tarefas distintas. Se a Terminologia visa estudar a estrutura, a formação, o desenvolvimento, o uso e a gestão de terminologias em diferentes domínios, a Ciência da Informação, quando constrói linguagens documentárias tem como objetivo assegurar a organização e a transferência da informação. Uma vez que a informação é uma construção, a linguagem documentária, embora se aproprie das terminologias e dos sistemas conceituais por ela mapeados, se propõe como modo de organização da informação, para o que articula temas, subtemas e termos orientados para dar conta de conjuntos documentários e para transferir informação para determinados públicos segundo objetivos específicos. Com efeito, segmenta, a seu modo, o universo enfocado segundo perspectivas pragmáticas. A partir dessa segmentação que se apresenta como proposta organizadora, uma linguagem documentária configura-se como em instrumento facilitador da comunicação em contextos documentários específicos.

Portanto, se para o senso comum as referências são genéricas e não necessariamente localizadas, podendo remeter a enciclopédias variadas, para a transferência da informação é necessário conjugar critérios terminológicos e pragmáticos que permitam, por um lado, permitir a identificação da porção da enciclopédia a ser ativada (como interpretar um termo) e, por outro, permitir o estabelecimento de vínculos de comunicação efetivos nos contextos informacionais. Se Marco Polo estivesse interessado, ao invés de classificar os novos animais, em saber para que servem, certamente teria selecionado características diferentes.

Comentários finais

Os processos de tratamento e recuperação de informações apresentam características similares às que ocorrem no processo de conhecimento em geral. Não é sem razão que as metodologias de organização da informação fazem uso de referenciais da Lógica, da Semiótica, da Lingüística, da Semântica, da Pragmática e das Ciências Cognitivas. Procura-se, a partir desses referenciais, operacionalizar com maior rigor os procedimentos de identificar e registrar as conjunções (superordenação), disjunções (subordinação), associações não-lineares, equivalências entre os termos, à semelhança do que ocorre, grande parte das vezes de forma inconsciente, em qualquer processo de conhecimento.

O que particulariza a atividade documentária se determina exatamente a partir de sua função: organizar para transferir, transferir para viabilizar a apropriação da informação. Nessa perspectiva, compreender como se desenvolve o processo interpretativo e identificar quais são as condições mínimas para que ele se desenvolva com eficácia em contextos informacionais, pode significar a diferença entre simplesmente estocar e transmitir para o

A exploração do texto de Eco permite verificar que os problemas enfrentados para a representação (ou melhor, construção) da informação não são simples. À semelhança do que ocorre na relação do homem com a realidade, a atividade documentária se caracteriza por segmentações ou recortes sobre os conteúdos, visando imprimir uma certa organização a um continuum indiferenciado. Na vida cotidiana, as distintas formas de organização são condição para a vida em sociedade; em contextos informacionais, escolhemos modos de organização específicos para alcançar objetivos específicos. Nos dois contextos é necessário existir compartilhamento, o que se torna possível a partir da presença de vínculos de significação. Nos sistemas informacionais, porém, onde a busca de graus de univocidade de significação é premente, faz-se necessário lançar mão de referenciais que permitam que esses vínculos tornem possível o estabelecimento de contextos de interpretação específicos. Sob esse aspecto, já não basta recorrer indiferenciadamente às enciclopédias disponíveis em nossa cultura: é preciso ir além, determinando qual dessas enciclopédias permite referendar um processo interpretativo determinado, ao qual deve se seguir a possibilidade de apropriação e uso efetivo da informação. No contexto documentário, portanto, a classificação do rinoceronte ou do ornitorrinco deverá ser feita a partir de um critério pragmático, determinando-se, a partir da função do sistema informacional, quais os aspectos a serem retidos em prol da criação de boas condições de recepção da informação.

Notas

[1]- Terminologia: ciência que estuda a estrutura, a formação, o desenvolvimento, o uso e a gestão de terminologias em diferentes domínios (ISO 1087-2000).

[2] - Na tradução publicada no Brasil, o termo intensão está grafado com "ç", erroneamente. Trata-se de dois conceitos distintos.

[3] - O entendimento de "enciclopédicas", neste contexto, refere-se à pretensão universalizante de abarcar, num só sistema, todo o universo do conhecimento. Vimos, entretanto, que a enciclopédica é um postulado semiótico: o que conseguimos registrar é sempre parcial frente à infinita possibilidade de interpretação nciclopédica.

[4] - LSP: Language for special purposes.

[5] - Lembremos aqui as tentativas de classificação feitas por Marco Polo para identificar o "rinoceronte", primeiramente, a partir de uma definição intensional. A impossibilidade de generalização motivou-o a modificar a definição original, alteração abandonada em seguida pelo prejuízo que causou à definição extensional.

[6] - Alguns trabalhos terminográficos partem, inicialmente, de um mapa conceitual construído mais ou menos empiricamente, que é refeito à medida que o trabalho de coleta e organização se desenvolve.

[7] - Os tesouros de língua, como por exemplo, o Thesaurus de Roget, são exemplos de léxicos organizados por associação de idéias. Representam, todavia, uma proposta de organização orientada pela seleção de determinados vértices de sentido.

[8] - Uma linguagem documentária, para Gardin, é constituída de um léxico, uma rede paradigmática e uma rede sintagmática (CINTRA et al., 1994). O eixo paradigmático das LDs é representado pela estruturação vertical (conjunções e disjunções) e horizontal (associações não-hierárquicas) do conjunto de termos do léxico. Já o eixo sintagmático da maior parte das LDs é muito precário se comparado ao das Línguas Naturais, pois os recursos de combinação entre os termos são muito restritos (nos tesouros, por exemplo, a sintaxe é realizada através de operadores booleanos: and, or, not). O investimento nas linguagens documentárias sintagmáticas, como foi o caso do SYNTOL - Sintagmatic Organization Language (Gardin), do Semantic Code ou WRU (Kent, Perry e Melton) e da Análise Codificada (Pagès), todos na década de 60-início de 70, precisaria ser retomado.

[9] - Termo, neste contexto, é uma designação genérica e não se confunde com o termo da Terminologia.

[10] - As principais referências de uso são consideradas através das relações de equivalência, muito embora não sejam as únicas. A função das relações de equivalência é facilitar o acesso do usuário ao sistema. Medida de economia em sistemas que ocupavam muita memória em disco até bem pouco tempo, interessa, nos nossos dias, apenas como medida de controle: a proposta de uma entrada única já não é imperativa quando se trata de permitir o acesso ao sistema, porém permanece a necessidade de dar conta da variabilidade das formas de busca.

Referências Bibliográficas

- BAITELLO JR., N. (1998). Comunicação, mídia e cultura. **São Paulo em Perspectiva**, v.12, n.4, p.11-16, out./dez.
- CINTRA, A.M.M.; TÁLAMO, M.F.G.M.; LARA, M.L.G. de & KOBASHI, N.Y.(1994). **Para entender as Linguagens Documentárias**. São Paulo: APB/Polis.
- ECO, U. (1988). Marco Polo e o unicórnio. In: _____. **Kant e o ornitorrinco**. Rio de Janeiro : Record, 1998.
- ECO, U. (1991). **Semiótica e filosofia da linguagem**. Trad. de Mariarosaria Fabris e José Luiz Fiorin. São Paulo : Ática, 1991.
- ISO 704 (2000). **Principles and methods of terminology**. 2.ed.
- ISO 1087-1:2000(E/F). **Terminology work - principles and methods. Part 1: theory and application; Travaux terminologiques - vocabulaire. Partie 1: théorie et application**. 1.ed.
- ISO 5127/6(1983-E/F). **Documentation and information - vocabulary - part 6: Documentary languages; Documentation et information -vocabulaire - partie 6: Langages documentaires**. 1.ed.
- LARA, M. L. G. de (1991) **Representação e linguagens documentárias: bases teórico-metodológicas**. São Paulo, ECA-USP (Tese de doutorado).
- LE GUERN, M. (1989). Sur les relations entre terminologie et lexique. **Meta**, v.34, n.3, p.340-343.
- LOPES, E. (1987). **Fundamentos da lingüística contemporânea**. São Paulo : Cultrix, 1987.
- NUNES, B. (1998). Nós somos um diálogo. São Paulo, Jornal de Resenhas, **Folha de São Paulo**, 08.ago., Cad. Especial, p.10 (Caderno Discurso Editoria/USP/Unesp)
- RORTY, Richard,(1988). **A filosofia e o espelho da natureza**. Lisboa: Dom Quixote.
- SMIT, J. W. (1998). **A perspectiva interdisciplinar da informação no contexto da Ciência da Informação**. Congresso Brasileiro de Arquivologia, 12, João Pessoa, 15-19 jun.
- TÁLAMO, M.F.G.M. (1997). **Linguagem documentária**. São Paulo: APB. (Ensaios APB, n.45).
- WELLISCH, h. h. (1991). Thesauri. In: _____. **Indexing from A to Z**. New York, The H. W. Wilson Company.

Sobre a autora / About the author:

Marilda Lopes Ginez de Lara

larama@usp.br

Profa. Doutora

Dept. de Biblioteconomia e Documentação

Escola de Comunicações e Artes, USP

Av. Lúcio Martins Rodrigues, 443

CEP 05508-900 - Cidade Universitária - São Paulo, SP