

LINGUAGENS DOCUMENTÁRIAS*: NÚCLEO BÁSICO DE CONHECIMENTO PARA SEU ESTUDO

*Maria Luiza de Almeida Campos***

Os alunos de escolas de biblioteconomia são incentivados a usar instrumentos de recuperação da informação já consagrados, sendo que pouca atenção vem sendo dada à criação de novos instrumentos e o uso crítico dos mesmos. Questões são levantadas quanto à criação de um núcleo básico de conhecimento necessário às atividades de pesquisa, ensino e elaboração de linguagens documentárias, como o tesauro e a tabela de classificação. Tal núcleo constitui nos elementos comuns para a formação de um sistema de conceitos – os conceitos e as relações – tendo por base as teorias: da terminologia, da classificação facetada e do conceito.

1 INTRODUÇÃO

A ausência de estudos teórico-metodológicos na área de representação do conhecimento/informação no País pode ser comprovada, de modo geral, nas próprias escolas de biblioteconomia, através da análise de um grande número de disciplinas, oferecidas nessa área, que privilegiam apenas a prática*** (1). Assim sendo, verifica-se

* O termo "linguagem documentária" tem sido usado nos sistemas de documentação para designar a "linguagem artificial" ali usada para fins de indexação, armazenamento e recuperação.

** Professora do Departamento de Documentação da Universidade Federal Fluminense.

*** Mestre em Ciência da Informação - IBICT/UFRJ

Em recente pesquisa sobre avaliação do uso de material didático nos cursos de graduação de biblioteconomia/ciência da informação, a professora Nice Figueiredo identifica nas disciplinas afins à área de representação da informação a deficiência de estudos teóricos nesta área.

na formação dos alunos dessas escolas um incentivo ao uso de instrumentos de recuperação da informação já consagrados, não os preparando para a criação de instrumentos que correspondam à realidade brasileira, ou mesmo para um uso crítico de tais instrumentos.

Neste trabalho pretendemos discutir algumas questões que vimos estudando há algum tempo e que são objeto de nossa dissertação de mestrado (2). O trabalho pretende tornar evidente um núcleo básico de conhecimento, necessário às atividades de pesquisa sobre ensino e elaboração de linguagens documentárias. Atualmente, não é mais possível dissociar a teoria da prática. Nesse sentido, não mais interessa a profissionais que lidam com a informação somente o uso de instrumentos de recuperação de informação/documents, elaborados, por vezes, para atender a realidades completamente adversas ao sistema de informação em que se está inserido. Interessam, também, as possibilidades de elaboração de tais instrumentos que venham atender às necessidades específicas de um dado sistema. Esse fato, porém, só poderá acontecer a partir do domínio das bases teóricas das linguagens documentárias.

2 INSTRUMENTOS DE REPRESENTAÇÃO/RECUPERAÇÃO DA INFORMAÇÃO E SUA CARACTERIZAÇÃO

Todo o movimento existente nos sistemas de recuperação da informação tem, por princípio geral, possibilitar a seu usuário o acesso à informação. Tendo-se em vista esse acesso, têm sido criadas, desde o século XVI (3), classificações bibliográficas para organizar informações documentárias nos diversos domínios do conhecimento.

Atualmente, vários são os instrumentos utilizados para se organizar e representar o conhecimento de uma dada área do saber, com a finalidade de permitir a seus usuários o acesso à informação/documents. Esses instrumentos são denominados, de forma geral, linguagens documentárias, como o tesauro e a tabela de classificação, para citar apenas os mais utilizados.

A tabela de classificação e o tesauro, na maioria das vezes, apresentam-se, sob duas formas: a alfabética e a sistemática. A forma sistemática, que será o objeto de nosso estudo, se apresenta de modo a evidenciar uma estrutura de conceitos, que formam entre si uma rede. Essa estrutura permite ao usuário visualizar as relações

entre os conceitos na área de abrangência do instrumento, facilitando a comunicação entre ele e a base de dados.

Quando o usuário procura o sistema para encontrar informação sobre determinado assunto, este se apresenta em sua mente como o agrupamento de conceitos¹ que estão relacionados entre si e com outros conceitos, formando assim uma estrutura conceitual. Também nos instrumentos como o tesauro e as tabelas de classificação, os conceitos se apresentam segundo uma estrutura conceitual.

Para que os conceitos sejam manipulados com a finalidade de representação/comunicação, utiliza-se de termos (nomes e/ou notações). A grande tendência, atualmente, é a representação através de termos. Estudando-se a teoria da terminologia, verifica-se que ela possui características semelhantes à teoria da classificação facetada e ao tesauro-com-base-em-conceito.

A terminologia como disciplina científica, a teoria da classificação facetada e a teoria do tesauro-com-base-em-conceito têm características semelhantes, porque todas objetivam a elaboração de instrumentos de representação do conhecimento. Entretanto, esses instrumentos possuem funções diferentes. Citaremos aqui pelo menos dois momentos em que essas diferenças se apresentam: 1 – o ambiente de origem, a saber, os espaços comunicacional e informacional; 2 – a função dos instrumentos.

Quanto ao primeiro – o lugar onde se originam a tabela de classificação, o tesauro e a terminologia – é necessário que se afirme que todos se estabelecem em um ambiente de produção do conhecimento, embora em espaços diferentes desse mesmo ambiente, tendo-se em vista a finalidade de uso/função de cada instrumento.

Como vimos anteriormente, todos esses instrumentos têm em comum “o conceito”. DROZD (5), em um de seus trabalhos, procura mostrar os níveis de formação de conceitos. Seu modelo contribui para o nosso entendimento dos espaços de origem da tabela de classificação, do tesauro e da terminologia, pois esses instrumentos nada mais são do que redes conceituais. Dessa forma, tentaremos estabelecer um paralelo entre o seu modelo e o espaço de produção do conhecimento.

¹ Entende-se, aqui, por conceito, segundo Dahlberg (4), a menor unidade de conhecimento que a mente humana consegue manipular.

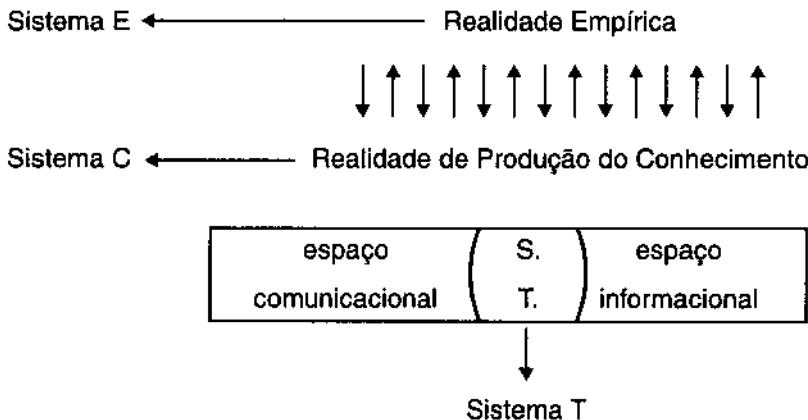
As tabelas de classificação, os tesouros e as terminologias são elaboradas em um espaço onde se instaura uma realidade reconstruída pelas palavras – a realidade da produção de conhecimento. DROZD denominou esse espaço de "Sistema C" – o sistema de conceitos, abstraído do nível teórico de consciência social, controlado pelas leis do pensamento.

Assim, esses instrumentos não estariam inseridos em uma realidade empírica – aquela em que se dá a ação dos homens e dos fenômenos naturais, também denominado – por DROZD como o "Sistema E", um sistema de entidades extralingua, a realidade do mundo, controlada pelas leis da natureza –, mas na realidade reconstruída pelas palavras (o Sistema C). Entretanto, essa realidade de produção de conhecimento influencia a realidade empírica e é por ela influenciada.

No ambiente de produção do conhecimento, notamos que existem dois espaços imbricados, mas de natureza diferente. São eles: a) o espaço comunicacional, onde as descobertas tornam-se registros, através do intercâmbio entre o "gerador" de conhecimento e o próprio conhecimento e entre o "gerador" e seus pares; e b) o espaço informacional, onde existe um "necessitador" de informação e um sistema possuidor de documentos/informação. Entre esses dois espaços encontra-se, também, a nosso ver, o que DROZD resolveu chamar de "Sistema T" – o sistema terminológico –, nomeado para um dado Sistema C. O Sistema T ocorre em um ambiente contextualizado de produção de conhecimento, controlado pelas leis da linguagem e da lógica. Neste sistema é que nos parece originar-se a terminologia, pois ela é criada para permitir ao próprio pesquisador pensar sobre a área e também para comunicar os próprios resultados.

No espaço informacional, mencionado anteriormente, verifica-se a necessidade de criação de instrumentos que possibilitem a criação não mais entre os pares, mas entre o usuário de um sistema de informação e o próprio sistema, que seria o espaço do tesouro e da tabela de classificação. Porém, esses instrumentos, para serem criados, necessitam de uma estrutura terminológica/conceitual que será buscada em um sistema terminológico.

O quadro a seguir apresenta de forma esquemática essas relações:



Quanto à segunda questão, ou seja, à função dos instrumentos, recuperaremos COYAUD (6) através de seu estudo sobre os níveis de comunicação de cada um deles. COYAUD classifica os instrumentos em a) aqueles que permitem acesso direto, e b) aqueles que permitem o acesso indireto ao documento, ou seja, aqueles em que primeiro se acessa a informação, necessitando-se ainda de um segundo movimento para acessar o próprio documento. Segundo ele, seria possível classificar o tesouro como sendo um instrumento de comunicação indireta, pois ele possibilita ao usuário recuperar a informação referencial e não necessariamente o documento. A tabela de classificação, que tem por função principal a ordenação física dos documentos, seria considerada como de acesso direto, apesar de termos conhecimento de que a Tabela de Classificação Universal de OTLET e LA FONTAINE tenha sido construída com a finalidade de organização bibliográfica, o que a caracterizaria como de comunicação indireta. Podemos dizer que, enquanto tabela de classificação para ordenação de documentos, ela foi utilizada posteriormente. As terminologias aqui não estão incluídas, pois não poderiam ser classificadas como um instrumento de recuperação da informação, e sim, como um instrumento de comunicação entre indivíduos de um mesmo universo discursivo. Mas elas são as fontes dos termos dos tesouros.

A tabela de classificação, o tesouro e as terminologias são, no entanto, instrumentos de comunicação, e, apesar de possuírem finalidades diferentes, todos lidam com conceitos e suas relações, for-

mando uma estrutura sistemática, o que possibilita a comparação entre as teorias que estão na base de construção de tais instrumentos.

3 CONTRIBUIÇÕES TEÓRICAS

Na década de 30, pautado na "teoria dinâmica do conhecimento" RANGANATHAN (7) elabora o método de facetas, que consiste na categorização de uma área de conhecimento para a organização das classes de conceitos e não mais de assunto, como ocorria até então, na elaboração das classificações bibliográficas. Assim, a sua unidade de trabalho passa ser o conceito definido como um isolado, unidade combinatória que tem por função facilitar a formação da notação. Por exemplo, na *Colon Classification* – a tabela de classificação elaborada por RANGANATHAN (8), conhecida também como "classificação de dois pontos" – a formação da notação para "psicologia infantil" dá-se a partir do agrupamento da notação correspondente aos conceitos psicologia (código S) e criança (código 1), este último definido como um isolado na teoria da classificação apresentada por RANGANATHAN. Pode-se dizer que o representante do conceito nesse tipo de tabela é a notação e não o termo. A notação resultante é S1. Com esses princípios RANGANATHAN consegue representar conceitos que não se encontram nomeados na língua. Por exemplo: psicologia + pré-adolescente, psicologia + idade madura, psicologia + menino.

Para a organização das classes, em sua tabela, RANGANATHAN (9) desenvolveu uma metodologia que levou em consideração as características dos conceitos como base de divisão (características de divisão) para a formação das sub-classes, das cadeias, dos conjuntos dentro das classes, dos renques,...(9) A teoria da classificação apresentada por ele continuou a ser desenvolvida pelo *Classification Research Group* – CRG na Inglaterra. VICKERY e seus colegas ampliaram as categorias/facetas e desenvolveram diversas tabelas de classificação já na década de 50. (10)

Embora Ranganathan tenha publicado formalmente os princípios da "teoria dinâmica do conhecimento" somente em 1957, em seu livro *"As Cinco Leis da Biblioteconomia"*, ele, já na década de 30, aplicava os princípios da teoria, na *Colon Classification*.

Com o surgimento, cada dia maior, do material bibliográfico especializado e dos serviços de informação, a elaboração e a constante atualização dos esquemas classificatórios passaram a se constituir num problema real. AITCHISON (11), percebendo o fato, utilizou os princípios ranganathianos de classificação para elaborar tesouros. Estes são instrumentos que possuem, além de uma abordagem alfabética, uma abordagem sistemática que facilita os processos de comunicação. Unindo a tabela de classificação e o tesouro num único instrumento, ela cria o *thesaurofacet*. Neste instrumento, ela inclui outras relações conceituais não-hierárquicas sem apresentar, porém, as bases teóricas.

Os estudos desenvolvidos, até então, pelo próprio RANGANATHAN forneceram as bases teóricas para a organização hierárquica dos conceitos, mas os tesouros revelaram que havia outras relações entre conceitos, como as relações associativas, embora suas bases teóricas não fossem desenvolvidas (12, 13, 14, 15, 16).

Na década de 30, um engenheiro austríaco, WUESTER (17) organizou a terminologia de eletrotécnica, com o objetivo de garantir comunicação precisa nesse campo da ciência. Esta experiência levou-o à criação de uma nova disciplina científica, a saber, a ciência da terminologia; WUESTER desenvolveu uma série de princípios que chamou de TGT, teoria geral da terminologia. Segundo essa teoria, a terminologia se ocupa dos conceitos de uma área do conhecimento. Para ela, o termo é a representação do conceito e um dos postulados da terminologia é que os termos de uma área do conhecimento formam entre si um sistema de conceitos. Seu ponto de partida, para essa sistematização, é o conceito, que se estabelece a partir da análise de suas características, que ele e seus seguidores categorizam (18, 19, 20, 21).

WUESTER avança, assim, para a elaboração das bases teóricas do conceito e do relacionamento entre conceitos, identificando relacionamentos lógicos (ou de abstração) e ontológicos (partitivos, causa/efeito, sucessão...), na formação do sistema. Nesse, ele introduz o conceito de relação hierárquica para as relações genéricas e partitivas.

Ao considerar que os termos formam um sistema, ele propõe que as terminologias se apresentem de forma sistemática e não alfabética, como os tradicionais dicionários especializados. Ora, seus princípios de sistematização apresentam-se bastante semeihantes e até certo ponto complementares aos da teoria apresentada por RANGANATHAN.

Se, por um lado, a teoria da classificação avançou no estabelecimento de bases que poderão auxiliar na categorização de uma dada área de conhecimento, por outro, a teoria da terminologia ampliou as bases para o estabelecimento das relações entre os conceitos e para a formação do termo.

A comparação entre essas metodologias poderá auxiliar na construção de um núcleo comum de conhecimento necessário à formação de profissionais que estejam interessados em trabalhar, tanto na elaboração de linguagens documentárias, como na criação de instrumentos de representação do conhecimento.

4 CONSIDERAÇÕES FINAIS

Na formação do profissional que irá lidar com os processos de elaboração de linguagens e de indexação, é necessário, além do conhecimento específico ligado à área das linguagens documentárias – como a teoria da classificação, a teoria da terminologia, a teoria do conceito, entre outras –, ter informação sobre a natureza do conhecimento em que se originam os documentos que vão ser objeto de representação/recuperação.

Até aqui, nos limitamos às questões específicas das linguagens documentárias, mas, para a criação de instrumentos eficazes, é necessário ter uma visão sistêmica, ou seja, compreender as relações que existem entre os diversos elementos de um sistema de recuperação de informação, e entre o sistema e a área de conhecimento no qual ele está inserido.

A linguagem documentária é um elemento do SRI; os outros elementos, como a coleção, a indexação, os usuários e suas necessidades fazem parte deste todo que é o sistema e que interagem entre si. Então, não é possível dissociar o estudo das linguagens documentárias de questões como o espaço discursivo do documento, também chamado por RANGANATHAN de universo do documento, isto é, os diferentes conteúdos influenciando as diferentes formas de suporte de conhecimento, pois a forma que um documento apresenta está diretamente ligada a um tipo de conhecimento produzido. Um exemplo dessa questão diz respeito às diferentes tabelas de forma nas diferentes áreas do conhecimento, que fazem parte de instrumentos como os tesouros e as tabelas de classificação.

Consideramos, devido a fatores como esses, que o estudo da natureza do conhecimento – isto é, a especificidade de cada área de saber, sua ligação com as outras áreas, a própria definição de área de conhecimento, a dinâmica do próprio conhecimento – é de vital importância, à medida que encaramos os documentos como suportes desses conhecimentos.

RANGANATHAN, já na década de 30, percebeu a importância desse estudo e elaborou a “teoria dinâmica do conhecimento”, que tem como pressuposto principal que o “universo do conhecimento é dinâmico e está em um eterno transformar-se”, e a linguagem documentária deveria possuir uma metodologia que pudesse acompanhar essa evolução. Elabora, assim, o método de faceta e define as categorias conceituais – o PMEST (personalidade, matéria, energia, espaço e tempo).

Na Inglaterra, o estudo da natureza do conhecimento já é uma realidade nos cursos de biblioteconomia. KEMP (22) discute todas essas questões, fazendo uma ponte entre questões filosóficas da área da epistemologia e da teoria do conhecimento para o contexto do profissional da informação. O estudo de questões como essas permitirá, ao profissional da informação, uma visão mais ampla da natureza e função dos diferentes instrumentos de representação/comunicação de uma área de conhecimento, ou de representação/comunicação entre a base documental e/ou informacional e o usuário.

DOCUMENTARY LANGUAGES: CORE OF KNOWLEDGE FOR THEIR STUDY

Library school students are induced to use information retrieval tools which are widely known, but they are not being prepared to a critical usage, or to the development of new ones. Questions are raised as to creating a body of knowledge, needed in research and teaching activities and in the development of indexing languages such as thesauri and classification schemes. Such a body of knowledge consists of common elements for the organization of a system of concepts, i. e. concepts and relations, based on terminology, classification and concept theories.

REFERÊNCIAS BIBLIOGRÁFICAS

- 1 FIGUEIREDO, Nice M. de. **Avaliação do uso de material didático nos Cursos de material didático nos Cursos de Biblioteconomia/Ciência da Informação no País, a nível de graduação.** Rio de Janeiro: CNPq/IBICT, 1991. 40p.
- 2 CAMPOS, Maria Luiza de A. **Em busca de princípios comuns na área de representação da informação: uma comparação entre o método de classificação facetada, o método de tesauro-baseado-em-conceito e a teoria geral da terminologia.** Rio de Janeiro: UFRJ/ECO/IBICT, 1974. 190p. (Dissertação de mestrado).
- 3 KAULA, P. N. Rethinking on the concepts in the classification. **Herald of Library Science**, v. 23, n. 1/2, p. 30-44 Jan./Apr. 1984.
- 4 DAHLBERG, I. **Ontical strutures and international classification.** Bangalore: Sarada Ranganathan Endowment for Library Science, 1978. 64p.
- 5 DROZD, M. L. Terminological synonyms and the function of definition. Thesis. In: ASSOCIATION INTERNATIONALE DE TERMINOLOGIE. **Problèmes de la définition et de l' synonymie en terminologie.** Quebec: Université Laval, 1982. p. 85-99.
- 6 COYAUD, M. **Introduction a l' étude des languages documentaires.** Paris: Librairie C. Klinchksieck, 1966. 148p.
- 7 RANGANATHAN, S. R. **The five laws of library science.** Madras: Library Association, 1957. 456p.
- 8 _____ **Colon Classification.** Bombay: Asia Publishing House, 1933.
- 9 _____ **Prolegomena to library classification.** Bombay: Asia Publishing House, 1963. 640p.
- 10 VICKERY, R. C. **Classificação e indexação nas ciências.** Rio de Janeiro: BNG/Brasilart, 1980.
- 11 AITCHISON, J., GILCHIRIST, A. **Manual para elaboração de tesouros.** Rio de Janeiro: BNG/Brasilart, 1979.
- 12 SOERGE, D. **Indexing languages and thesauri construction and maintenance.** Los Angeles: Melville, 1974.
- 13 NEELAMEGHAN, A. Non-hierarquical associative relationships: their types in computer generation of RT links. In: **Seminar on Thesaurus in Information Systems.** Bangalore, 1975. Section A, A1-AB.
- 14 WILLETS, M. An investigation of the nature of relation between terms in thesauri. **J. Doc.** v. 31, n. 3, p. 258-84 Sept. 1975.
- 15 RANJAN, T. N. "Related Terms" in thesauri. In: **Seminar on Thesaurus in Information Systems.** Bangalore, 1975. Section A, A13-A21.

- 16 CHERNYI, A. I. General procedure for construction thesauri. **Automatic Documentation and Mathematical Linguistics**, v. 2, n. 2, p. 13-51, 1968.
- 17 WUESTER, E. L'Étude scientifique générale de la terminologie, zone frontalier entre la linguistique, la logique et les sciences des choses. In: RONDEAU, G., FELBER, H. **Textes Choisis de Terminologie**. Québec: Université Laval, 1981. p.55-108.
- 18 DAHLBERG, I. Terminological definitions: characteristics and demands. In: ASSOCIATION INTERNATIONALE DE TERMINOLOGIE. **Problèmes de la définition et de la synonymie en terminologie**. Québec: Université Laval, 1982. p.15-21.
- 19 _____, Les objets, les notions, les définitions, et les termes. In: RONDEAU, G., FELDER, H. **Textes choisis de terminologie**. Québec: Université Laval, 1981. p. 221-55.
- 20 LOTTE, D. S. Principes d'établissement d'une terminologie scientifique et technique. In: RONDEAU, G., FELBER, H. **Textes choisis de terminologie**. Québec: Université Laval, 1981. p. 1-38.
- 21 KANDELAKI, T. L. Les sens des termes et les systèmes de sens des terminologies scientifiques et techniques. In: RONDEAU, G., FELBER, H. **Textes choisis de terminologie**. Québec: Université Laval, 1981. p. 133-57.
- 22 KEMP, D. A. **The nature of knowledge; an introduction for librarians**. London: Clive Bingley, 1976.