

A REPRESENTAÇÃO DOCUMENTÁRIA DE ARTIGOS CIENTÍFICOS EM EDUCAÇÃO ESPECIAL: ORIENTAÇÃO AOS AUTORES PARA DETERMINAÇÃO DE PALAVRAS CHAVES
THE DOCUMENTARY REPRESENTATION OF SCIENTIFIC ARTICLES IN SPECIAL EDUCATION: GUIDANCE TO THE AUTHORS TO ESTABLISH KEYWORDS

Mariângela Spotti Lopes FUJITA¹

RESUMO: A palavra-chave é uma representação do conteúdo significativo do texto e também é utilizada para representar uma necessidade de informação na estratégia de busca. O tratamento temático de conteúdo para sistemas de recuperação da informação oferece subsídios teóricos e metodológicos para a determinação de palavras-chaves tendo em vista a estrutura e a tipologia textual do artigo científico. Propõe-se a explicitação do processo de determinação de palavras-chaves em artigos científicos da área de Educação Especial com o objetivo a representação documental do seu conteúdo para atendimento de demanda de recuperação pelo leitor especialista da área. O texto científico é apresentado como variável importante em metodologia de indexação para a determinação de palavras chaves utilizando como estratégia de análise de assunto a localização do tema principal pela exploração da estrutura textual combinada com a análise conceitual por questionamento e representação dos termos obtidos pela linguagem documental do sistema que indexa o periódico. Conclui-se pela orientação de que a palavra-chave é resultado da representação documental realizada mediante compatibilização com linguagens documentais após identificação e seleção de termos durante a análise de assunto do artigo pelo autor-leitor.

PALAVRAS-CHAVE: indexação; palavras chaves; linguagens documentais; artigo científico.

ABSTRACT: the keyword is a representation of the text's significant content and is also used to represent a need of information in the search strategy. The thematic treatment of content to systems of information retrieval offers theoretical and methodological subsidies to establish keywords, considering the structure and textual typology of the scientific article. We propose to explicit the process of establishing the keywords in scientific articles in the area of Special Education, aiming the documentary representation of their content regarding the retrieval demand by the area expert reader. The scientific text is presented as an important variable in the indexing methodology to establish keywords, using as strategy of subject analysis the localization of the main subject by exploiting the text structure matched to the conceptual analysis by questioning and representing the terms obtained by the documentary language of the system that indexes the periodical. The conclusion is that the keyword is a result of the documentary representation accomplished by means of compatibility with documentary languages after the identification and selection of terms, by the author-reader during the subject analysis of the article.

KEYWORDS: indexing; keywords; documentary languages; scientific article.

INTRODUÇÃO

O significado de uma palavra-chave é intrínseco ao conteúdo de um texto, da mesma forma como se pudéssemos encapsular todo o conteúdo de um texto em uma palavra-chave. Desse modo, determinar palavras-chaves de modo a representar o conteúdo de um texto para uma futura recuperação é, certamente, uma operação que envolve conhecimento sobre o assunto do texto.

O autor do texto detém o conhecimento sobre seu conteúdo e isso é algo automaticamente esperado por todos. Também é entendido que o autor do

¹ Profa. Adjunta do Departamento de Ciência da Informação da UNESP – Campus de Marília - goldstar@flash.tv.br

texto previu os leitores de seu texto e o redigiu colocando em seu conteúdo o conhecimento previsto para aquela determinada área de assunto, de forma a propiciar condições de legibilidade para a assimilação de novos conhecimentos.

Mesmo com o conhecimento sobre o assunto do texto, o autor, ao determinar palavras-chaves, deverá ter em mente, não somente o assunto principal do texto, mas principalmente o objetivo de representação documentária do conteúdo textual para recuperação por um leitor cuja demanda de informação seja compatível ao assunto representado pela palavra-chave. De outro modo, também, nos referimos à um processo de intermediação entre os significados de um texto e os significados de uma demanda informacional por meio de palavras-chaves. A palavra, nesse contexto, é uma chave comutadora de significados entre o texto e um leitor que o necessita.

Portanto, para a determinação de palavras-chaves o autor deverá ter uma visão sobre o texto e a organização de seu conteúdo de modo a representa-lo conforme seu assunto principal e, em outro plano, o objetivo de representação do conteúdo documentário, a visão sobre a demanda do leitor previsto. Nesse sentido, o artigo propõe a explicitação do processo de determinação de palavras-chaves em artigos científicos da área de Educação Especial tendo em vista a representação documentária do seu conteúdo para atendimento de demanda de recuperação pelo leitor especialista da área.

O artigo terá como pressuposto teórico o tratamento temático de conteúdo para sistemas de recuperação da informação que oferecerá subsídios metodológicos para a determinação de palavras-chaves tendo em vista a estrutura e a tipologia textual do artigo científico.

A REPRESENTAÇÃO DOCUMENTÁRIA EM PALAVRAS CHAVES

O conceito de palavra-chave está muito ligado às buscas de informação em bases de dados onde as ferramentas de busca solicitam as palavras-chaves que deverão apresentar compatibilidade com as palavras-chaves representativas dos textos e, a partir disso, recuperar o documento demandado pela busca por palavras-chaves. Historicamente as palavras-chaves ou *keywords* surgiram como *termos de acesso* nos primeiros índices automatizados existentes nas Bibliografias especializadas, hoje denominadas de bases de dados.

A palavra-chave é uma representação do conteúdo significativo do texto e também é utilizada para representar uma necessidade de informação na estratégia de busca. A palavra-chave pode ser indicada por meio do conteúdo do texto ou escolhida em vocabulários livres e/ou controlados. Assim, a determinação de palavras-chaves constitui-se como representação documentária no processo de indexação documentária e na recuperação da informação e, nesse sentido, este artigo tratará da palavra-chave no contexto da indexação para orientar os autores de artigos científicos em Educação Especial quanto à determinação de palavras –chaves.

A comunicação humana somente será durável se registrada em suportes documentários (livros, fotos, imagens, memórias digitais, etc). Especificamente, a comunicação registrada em suportes documentários é objeto das atividades de organização da informação desenvolvida por instituições informacionais a partir de um conjunto de operações básicas.

Segundo Kobashi (1994, p. 15), "a circulação de informações e documentos comporta operações básicas" que podemos considerar como um ciclo de operações documentárias constituídas de: Coleta de documentos, Tratamento de documentos e Difusão de documentos para uma determinada comunidade usuária. As mesmas operações são indicadas por Guinchat e Menou (1994, p.30), com uma pequena variação terminológica, ao referir-se à coleta de documentos, tratamento intelectual e difusão da informação.

Cada uma dessas operações – coleta, tratamento e difusão - desdobram-se em atividades dotadas de política e procedimentos metodológicos bem definidos:

- A *Coleta* compreende toda a operação de seleção e aquisição de documentos convencionais e não-convencionais;
- O *Tratamento* executará o processamento dos documentos coletados com relação, tanto ao suporte material quanto a seu conteúdo;
- A *Difusão* será realizada através dos produtos e serviços do sistema documentário planejados de acordo com a demanda da comunidade usuária: levantamentos bibliográficos retrospectivos e atualizados, consultas bibliográficas, empréstimo de documentos, comutação documentária, entre outros.

Nesse contexto, a Organização da informação compreende as atividades e operações do tratamento da informação, envolvendo para isso, o conhecimento teórico e metodológico disponível quanto ao tratamento descritivo do suporte material da informação e ao tratamento temático de conteúdo da informação. Esta dicotomia que se apresenta no tratamento da informação é explicada, de um lado, pelo desenvolvimento teórico e metodológico distintos alcançados pelas duas áreas e, de outro, pela diferença existente entre os aspectos da informação – o material e o conteúdo, que exigem tratamento diferenciado.

Dessa forma, o tratamento do suporte material do documento é denominado, segundo Kobashi (1994), Representação Descritiva, porquanto visa à descrição normalizada dos aspectos físicos dos documentos (título, autoria, edição, publicação, data de publicação, local de publicação) para, justamente, propiciar uma acessibilidade física. A descrição física dos aspectos físicos dos documentos é desenvolvida mais especificamente, segundo Guinchat e Menou (1994) pela descrição bibliográfica ou catalogação no contexto da representação descritiva.

O tratamento quanto ao conteúdo do documento objetiva, por outro lado, uma descrição do conteúdo para uma representação condensada do que está expresso no texto com vistas à uma acessibilidade temática.

A descrição de conteúdo é denominada de *Análise documentária*, cujas operações principais são a análise, síntese e representação. A descrição de cada operação indica uma seqüência de procedimentos (KOBASHI, 1994, p.23):

ANÁLISE: leitura e segmentação do texto para identificação e seleção de conceitos;

SÍNTESE: Construção do texto documentário (enunciado de assunto composto de termos ou resumo) com os conceitos selecionados:

REPRESENTAÇÃO: a representação possui duas naturezas distintas: na primeira, a representação é construída por meio de um processo de condensação intensiva do texto original gerando os diferentes tipos de resumo e a determinação de palavras-chaves como produtos documentários; na segunda, a representação é realizada com o uso de um código comutador, ou seja, uma Linguagem documentária que poderá ser um tesouro ou lista de cabeçalho de assunto especializados da área de assunto.

É correto observarmos que, em termos de representação, o texto é uma primeira representação escrita do conhecimento gerado por um autor. O resumo e a palavra-chave, por sua vez, são representações condensadas do texto. Quando utilizamos uma linguagem documentária, tal como o tesouro, para verificarmos se a palavra-chave determinada é autorizada e está adequada à recuperação da informação a ser realizada por um leitor interessado, estamos realizando, assim, uma terceira representação. Esta seqüência de operações constitui-se na representação documentária que possui um grau de condensação cada vez maior em cada nível.

Pelo nível de descrição do conteúdo a Análise documentária poderá realizar a Classificação, a Indexação e a elaboração de resumos, obtendo diferentes resultados quanto ao nível de condensação. O resumo, como demonstrado na Fig. 1, apresenta um nível de condensação menor do que as palavras-chaves para a indexação ou o número de classificação na operação de classificação, entretanto, o resultado da indexação, da classificação e da elaboração de resumos é uma representação do conteúdo do documento para sua acessibilidade temática.

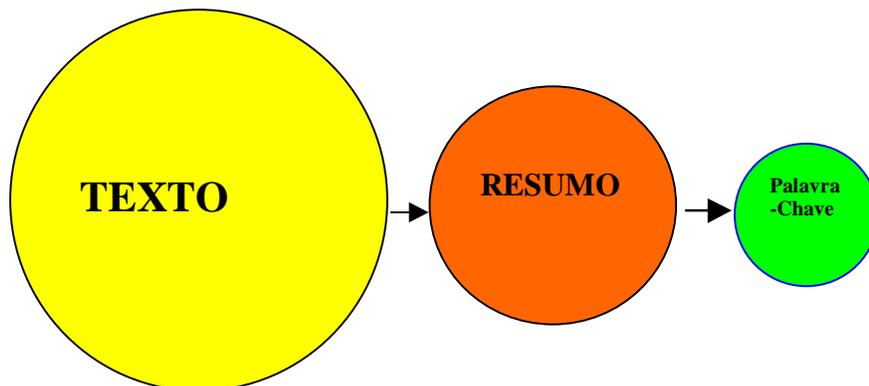


Figura 1: Níveis de condensação documentária entre o resumo e a palavra-chave

O grau de redução do texto original é maior na determinação da palavra-chave resultando em generalização crescente e conseqüente perda de informação semântica que poderá ser minimizada com uma representação totalmente voltada para o conteúdo do texto.

A determinação de palavras-chaves faz parte do processo de indexação documentária. A indexação é vista por Chaumier (1980, p. 42) como a "...parte mais importante da análise documentária", é uma combinação metodológica altamente estratégica entre o tratamento do conteúdo de documentos e sua recuperação por um usuário, demonstrando uma relação estreita entre o processo e a finalidade da indexação.

Portanto, seria correto supormos, a partir de uma perspectiva macro-analítica da área, que a Indexação, a Classificação como processo e a elaboração de Resumos são atividades de Análise Documentária.

Dessa forma, compreendemos a Análise Documentária como área teórica e metodológica com o objetivo de tratamento temático de documentos que abrange as atividades de Indexação, Classificação e elaboração de resumos, considerando suas diferentes finalidades de recuperação da informação científica.

A recuperação de informações científicas é providenciada por sistemas de informação especializados que coletam, tratam e divulgam documentos. Um sistema de recuperação da informação, segundo Lancaster (1968), pode recuperar textos completos de documentos, textos indicativos (como resumos) ou ainda nomes e endereços de documentos, ou seja, referências bibliográficas completas. Lopes (1985, p.242), complementa que esses sistemas realizam, também, o armazenamento e a organização de dados sobre documentos originais "[...] podendo registrar informações mais específicas ou mais abrangentes de uma determinada área do conhecimento, geradas em um país ou em vários."

Assim, os serviços de recuperação de informação exercem várias atividades classificadas conforme as funções de entrada e saída da informação:

⇒ *entrada*

- Os documentos são adquiridos por meio de uma política de seleção baseada nas necessidades do usuário;
- Os documentos são organizados e controlados através da catalogação, classificação e indexação;
- Armazenagem dos documentos de acordo com a organização adotada pelo sistema;
- Construção das bases de dados bibliográficas que servirão como suporte e indicadores na recuperação da informação;

⇒ *saída*

- O serviço de informação atende às demandas dos seus usuários mediante elaboração de estratégias de busca de informação que devem ser compatíveis com o tipo de indexação adotada pelo sistema.

Nos sistemas de informação, a recuperação está condicionada às condições de armazenagem, de tal forma que, segundo Cesarino (1985), a eficiência de um sistema de recuperação de informação depende muito da *qualidade* da análise conceitual tanto dos documentos quanto das questões. Segundo Houghton e Convey (1977) as funções de armazenamento e recuperação consistem das seguintes operações:

Armazenamento

- A análise do assunto para determinação de conceitos;
- A tradução dos conceitos analisados em palavras-chaves da linguagem documentária do sistema;
- A organização dos arquivos pelos quais a base de dados é composta.

Recuperação

- A análise da questão da busca;
- A tradução da questão em palavras-chaves em linguagem documentária do sistema;
- A formulação da estratégia de busca.

No Brasil e no mundo, existem, atualmente, sistemas que cobrem áreas de assunto especializadas com serviços de análise para tratamento temático e descritivo da literatura publicada por pesquisadores brasileiros para geração de bases de dados que serão utilizadas para a recuperação e divulgação da informação científica.

Considerando-se a necessidade de recuperação e divulgação científica dos artigos publicados por meio de bases de dados, é importante que as revistas científicas em áreas especializadas cuidem do aprimoramento formal e de mérito dos artigos orientando os autores a gerar adequadamente os produtos documentários do texto original: o resumo e as palavras-chaves, que deverão representar o conteúdo do artigo para a recuperação da informação em bases de dados elaboradas pelos sistemas de recuperação da informação.

Portanto, o item seguinte apresentará o texto científico como variável importante na análise de assunto e orientará o autor do artigo científico com a metodologia de indexação para a determinação de palavras-chaves utilizando como estratégia de análise de assunto a localização do tema principal pela exploração da estrutura textual combinada com a análise conceitual por questionamento (FUJITA, 2003).

A METODOLOGIA DE ANÁLISE DE ASSUNTO DE TEXTOS CIENTÍFICOS PARA DETERMINAÇÃO DE PALAVRAS CHAVES

Na análise de assunto para determinação de palavras chaves, o texto é variável importante e precisa ser compreendido como objeto de análise pelo leitor-autor no que diz respeito à exploração de sua estrutura textual característica de artigos científicos para localizar e identificar as palavras significativas que representem o seu conteúdo.

O texto, além de uma estrutura lingüística, possui uma estrutura de significado que somente "aparece" quando o leitor faz uma leitura compreensiva. Então, também o texto está sujeito a uma interpretação cognitiva e não somente descritiva.

No que diz respeito à estrutura do texto, afirma-se estar associada ao modo com o qual as idéias são organizadas no texto; com relação ao conteúdo, ao tema e aos conceitos tratados no texto. Como a estrutura do texto se articula ao seu conteúdo, o autor de um texto escolhe determinada estrutura textual que venha coincidir com o conteúdo que quer transmitir.

Por isso, uma parte importante do processo de compreensão de leitura é justamente essa habilidade de reconhecer o gênero do texto, bem como os diferentes tipos de textos.

Para Van Dijk (1992), o que o leitor procura durante a leitura é a informação importante, podendo esta variar de um leitor para outro. Em vista disso, são consideradas duas categorias de informação importantes, a saber:

- Informação textualmente importante considerada pelo autor;
- Informação contextualmente importante considerada pelo leitor mediante a sua intenção de leitura.

Cavalcanti (1989), ao descrever a centralidade do princípio de relevância na comunicação produtiva e receptiva, chama isso de saliência-autor/relevância-leitor. No que tange à saliência-autor, entenda-se a de saliência textual (idéias que o autor salienta no texto) e à relevância-leitor, as idéias que o leitor escolhe para interagir durante a leitura.

Para Giasson (1993), a informação importante pode ser tanto a idéia principal apresentada no texto com mais ênfase pelo autor dentro do contexto de determinado assunto narrado, como também o assunto tratado de forma global.

Pelo exposto, observando-se a variável texto do processo de leitura, numa perspectiva macro, é possível notar que os textos apresentam uma estrutura com as partes informacionais que os compõem organizadas numa seqüência lógica diferenciada de uma tipologia textual para outra.

A idéia principal varia de acordo com a estrutura textual, por exemplo: num texto narrativo a idéia principal pode ser um acontecimento ou a sua interpretação; num texto informativo pode ser uma regra, um conceito, ou uma generalização.

Quando a idéia principal aparece implícita, o leitor deve inferi-la com base nas informações fornecidas pelo texto e no seu conhecimento prévio sobre o assunto. No texto técnico-científico, Tálamo (1987) reconhece a identificação do tema no "objetivo" do trabalho.

Esse tipo de conhecimento prévio pelo leitor (de estruturas textuais) possibilita-lhe identificar a parte do texto que traz a idéia principal, fato que o auxilia a compreender, de forma global, o texto e a realizar uma leitura mais objetiva, pois já conhece as partes que tem a explorar e os conceitos pertencentes a cada parte, chegando, dessa forma, ao tema do texto.

Uma dica importante para identificação do tema é fazer um questionamento por categorias temáticas: o que? (categoria essencial); quando?, onde?, como? (categorias acessórias).

Para Kobashi (1994), a extração de informação documentária para produzir resumos e índices pode ser eficiente se o indexador conhecer a superestrutura textual, ou seja, como o texto está organizado. E, enquanto paradigma de organização textual, o esquema, ou superestrutura, fornece uma base para a interpretação do texto.

Kato (1986) salienta que a superestrutura textual é fundamental para a compreensão da leitura e que, em vista disso, tanto o leitor quanto o autor devem conhecer o formato de texto – o leitor, para buscar a compreensão do mesmo, identificando assim, o tipo de informação que deverá encontrar; - o autor, para optar por esta ou aquela estrutura textual na exposição de suas idéias.

Ao abordar a estrutura dos textos, Van Dijk (1992) enfatiza a estrutura narrativa, a argumentativa e a do discurso científico, considerando que os textos narrativos são formas básicas muito importantes da comunicação textual, enquanto as argumentativas são as mais utilizadas em filosofia e teoria da lógica e as últimas servem de base ao discurso científico.

Em textos científicos, Van Dijk (1992) aponta variantes especiais das superestruturas argumentativas, ou seja, a estrutura básica do discurso científico não só consiste de uma *conclusão* e sua *justificativa*, mas também da apresentação de um *problema* e uma *solução*.

Segundo Van Dijk (1992), pode-se perceber isso em um artigo científico que trata de informe experimental, cujas partes ou estrutura do texto seriam *observação*, para encontrar uma explicação para cada circunstância e se formular a *hipótese*; dessa derivam-se *predições* que podem ser comprovadas por *teste*; temos ainda as *peças envolvidas no ensaio*, a *estruturação do experimento*, *diferentes condições de investigações* e a *execução do experimento*, *resultado*, *conclusão(ões)* e *solução*.

Para caracterizar o texto científico, Pinto e Gálvez ([1996], p. 23), indicam os seguintes aspectos:

- uma estrutura esquemática estereotipada, que contém objetivo, metodologia, resultados e conclusões;

- um estilo cuidadoso e altamente formalizado;
- objetividade do conteúdo, de acordo com a realidade científica;
- emprego de uma linguagem científica diferente da linguagem habitual;
- prioridade ao implícito, ao conhecido, à informação acumulada durante o desenvolvimento da humanidade e
- a própria essência como criação humana.

As autoras ressaltam ainda que o texto científico pressupõe grande quantidade de informação e que a pressuposição é importante fator a considerar no exercício das tarefas analíticas.

Assim, o artigo científico possui uma estrutura textual característica e esquematizada, cujo tema principal poderá ser identificado pelos seus objetivos, localizados, via de regra, na Introdução. A determinação de palavras-chaves no artigo científico dependerá do conhecimento dessa estrutura textual combinado com uma metodologia de indexação.

O processo de indexação, segundo os “Princípios de indexação” e a Norma ABNT 12.676, compreende dois estágios: o analítico, em que é realizada a compreensão do texto como um todo, a identificação e a seleção de conceitos válidos para a indexação e o estágio de tradução, que consiste na representação de conceitos por termos de uma linguagem de indexação:

- Determinação do assunto: estabelecimento dos conceitos tratados num documento;
- Representação de conceitos por termos de uma linguagem documentária: a tradução dos conceitos nos termos da linguagem documentária determinando as palavras chaves para a recuperação da informação.

O primeiro estágio, a análise de assunto, é subdividido em outros três estágios:

- compreensão do conteúdo do documento;
- identificação dos conceitos que representam este conteúdo;
- e seleção dos conceitos válidos para recuperação.

Para que o indexador (leitor-autor) não negligencie nenhuma informação relevante, é preciso examinar partes importantes do texto que merecem especial atenção, durante sua leitura: título, introdução e as primeiras frases de capítulos e parágrafos; ilustrações, tabelas, diagramas e suas explicações; conclusão; palavras ou grupos de palavras sublinhadas ou impressas com tipo diferente. Os primeiros itens do texto geralmente apresentam as intenções do autor, enquanto que as partes finais comunicam o alcance dessas intenções. Por isso, não se recomenda a determinação de termos somente pelo título e/ou pelo resumo.

Na *identificação de conceitos*, o indexador, após o exame do texto, passa a abordá-lo de uma forma mais lógica a fim de selecionar os termos que melhor representem seu conteúdo. Para isso, a norma internacional recomenda que a identificação de conceitos obedeça a um esquema de categorias existente na área coberta pelo documento, como, por exemplo, o fenômeno, o processo, as propriedades, as operações, o material, o equipamento, etc.

Na metodologia proposta por Tálamo (1987), o processo de indexação consiste em identificar os termos representativos de um documento por meio de um mecanismo de perguntas e respostas agrupadas por generalidades e que respondem a cada uma das seguintes questões fundamentais: Quem? (ser), O que? (tema), Como? (modo), Onde? (lugar) e Quando? (tempo). Conforme a autora, identificando essa estrutura temática encontra-se o objetivo principal do texto, isto é, as informações relevantes separando-as das acessórias.

Kobashi (1994) destaca que a categoria "Quem?" não foi identificada em textos técnicos científicos, enquanto que a categoria "O que?" é essencial por ser o *elemento nuclear da estrutura temática*. Segundo a autora, as categorias Quando?, Onde? e Como? são categorias acessórias da principal "O que?", podendo assim aparecer, ou não, no texto, independente da "ordem de precedência de entre elas".

O tema, portanto, possui uma estrutura temática composta por conceitos ou categorias ou facetas cuja identificação decorrerá da análise conceitual do documento. A composição das categorias identificadas formularão o tema do documento em questão.

A respeito de "onde" localizar conceitos depende da identificação da estrutura temática, contudo, conforme a legibilidade e a estrutura textual do documento, o tema poderá estar formulado de forma clara no "objetivo" do trabalho e, quando isso não acontecer, será preciso a identificação dos conceitos dentro da estrutura textual do documento.

Em seguida, a *seleção de conceitos* deve ter em vista os objetivos para os quais as informações são indexadas, deixando claro que nem todas os conceitos identificados serão necessariamente selecionados.

Na representação de conceitos por termos de uma linguagem documentária é preciso realizar a terceira representação, a que nos referimos no item anterior, e obtermos as palavras-chaves do artigo que serão utilizadas para sua recuperação em sistemas de buscas. Dessa forma, o conceito é expresso por termos significativos, considerados representativos do conteúdo do artigo e torna-se palavra-chave quando, após sua extração do conteúdo pela análise de assunto do artigo, é traduzido por uma linguagem documentária na etapa de representação.

Como se constata, embora existam algumas divergências com relação ao número de etapas operacionais, o processo de indexação essencialmente consiste de três etapas básicas: análise, síntese e representação.

Portanto, a metodologia de indexação para determinação das palavras-chaves consiste em:

- a) examinar, pela leitura do texto científico, as partes de sua estrutura textual (Objetivo, Introdução, Problema, Hipótese, Metodologia, Resultado e Conclusão);
- b) localizar na introdução do artigo o objetivo do trabalho;
- c) Responder, se possível, todas as perguntas:
 - *o que aconteceu? (ação)*
 - *a que ou a quem isto aconteceu? (objeto da ação - sistema chave)*
 - *o que ou quem fez isto? (agente da ação)*
 - *onde aconteceu? (local)*
- d) identificar os conceitos expressos por termos pelas respostas às perguntas;
- e) verificar se outras partes do texto como metodologia, resultados e conclusão, contém termos adicionais;
- f) selecionar os conceitos expressos por termos mais representativas do conteúdo do artigo.

Para melhor compreensão da metodologia exposta, é preciso observar que os itens b), c) e d) são combinados de forma a sistematizar a exploração da estrutura textual do artigo científico com a identificação de conceitos por meio de questionamento e, assim facilitar, para o autor-leitor, a determinação de palavras significativas para a representação mais adequada do conteúdo do artigo. No quadro a seguir, está a visualização da combinação proposta na metodologia.

| Partes do texto | Conteúdo pertinente | Conceitos de |
|----------------------------------|--|--------------------------------------|
| Introdução (objetivos) | - Referencial teórico - Tema: objetivos* - Objetivos | objeto agente ação |
| Metodologia | - Descrição de materiais, métodos, processos e técnicas utilizados. | métodos local físico materiais |
| Resultados | - Compatibilidade com objetivos enunciados e materiais e métodos utilizados, mostrados, às vezes em tabelas; | |
| Discussão dos resultados | - Verificação dos resultados a partir do referencial teórico utilizado | |

* observe que o tema, geralmente, está expresso através do objetivo.

Quadro 1: Identificação de conceitos mediante exploração da estrutura textual

Os conceitos colocados na terceira coluna à direita do quadro têm significados distintos, como demonstrado a seguir, e terão a função de determinar as palavras-chaves referentes ao seus significados:

Objeto: é algo ou alguém que está sob estudo do pesquisador.

- *Ação*: processo sofrido por algo ou alguém
- *Agente*: aquele ou algo que realizou a ação
- *Métodos*: métodos utilizados para realização da pesquisa
- *Local físico ou ambiência*: local físico onde foi realizada a pesquisa

A seguir, temos um exemplo que demonstra o uso do questionamento e a obtenção de termos como resposta à identificação dos conceitos estabelecidos.

⇒ **ARTIGO:**

LOPES, A. W. de A.; VALDÉS, M. T. Formação de professores de educação física que atuam com alunos com necessidades educacionais especiais (deficiência auditiva): uma experiência no ensino fundamental da rede pública. *Revista Brasileira de Educação Especial*, v. 9, n.2, 2003.

⇒ **EXPLORAÇÃO DA ESTRUTURA TEXTUAL DO ARTIGO**

Objetivo do artigo localizado na introdução:

“O objetivo deste estudo consistiu em delinear um programa de formação de professores de Educação Física que atuam com alunos com necessidades educacionais especiais (deficiência auditiva) do ensino fundamental da rede pública, com o intuito de ampliar os conhecimentos acerca da Educação Física Adaptada, através de elementos teóricos-práticos que possibilitam um atendimento de forma responsável e competente, minimizando as barreiras e percalços do movimento inclusivista na escola.”

⇒ **QUESTIONAMENTO PARA IDENTIFICAÇÃO DE CONCEITOS REPRESENTADOS POR TERMOS SIGNIFICATIVOS**

- *O que aconteceu? (ação)*: formação de professores
- *A que ou a quem isto aconteceu? (objeto da ação - sistema chave)*: professores de educação física
- *O que ou quem fez isto? (agente da ação)* educação especial
- *Onde aconteceu? (local)* rede pública de ensino fundamental

⇒ **IDENTIFICAÇÃO DE OUTROS TERMOS SIGNIFICATIVAS**

Em Metodologia:

- práticas pedagógicas,
- Fortaleza (local geográfico),

- Alunos deficientes auditivos (sujeitos da pesquisa)

SELEÇÃO DE CONCEITOS REPRESENTADOS POR TERMOS SIGNIFICATIVOS

A partir da *identificação de conceitos*, serão selecionados os termos considerados importantes para uma representação mais pertinente ao conteúdo do artigo e baseada na demanda do sistema de informação que o indexa.

Considerando-se, também, o grau de especificidade do vocabulário utilizado no artigo científico, é preciso realizar a seleção de conceitos em função do significado específico da palavra-chave dentro da área de assunto, evitando-se, dessa forma selecionar, palavras tais como “estudo”, “análise”, “avaliação” entre outras.

A seleção de termos deve ser realizada a partir da lista de termos identificados no conteúdo do artigo, evitando-se a seleção de outras palavras propostas a partir da interpretação do autor-leitor e que não existem no conteúdo do artigo.

No quadro abaixo, todos os termos identificados foram selecionados, entretanto, dependendo do contexto específico da área poderão haver escolhas entre os termos identificados.

Exemplo:

| Termos identificados | Termos selecionados | Palavras-chaves * |
|-------------------------------------|-------------------------------------|--------------------------------------|
| Programa de Formação de professores | Programa de Formação de professores | Programa de Formação de professores |
| Professores de Educação Física | Professores de Educação Física | Professores de Educação Profissional |
| Educação especial | Educação especial | Educação especial |
| Rede pública de ensino fundamental | Rede pública de ensino fundamental | Rede pública de ensino |
| Práticas pedagógicas | Práticas pedagógicas | Práticas educativas |
| Fortaleza | Fortaleza | Fortaleza |
| Alunos deficientes auditivos | Alunos deficientes auditivos | Educação dos deficientes auditivos |

*Palavras-chave selecionadas após consulta ao Thesaurus Brasileiro da Educação do INEP no site www.inep.gov.br

Quadro 2: Termos identificados e selecionados com representação por palavras-chaves pela linguagem documentária

REPRESENTAÇÃO DOCUMENTÁRIA EM PALAVRAS CHAVES DA LINGUAGEM DOCUMENTÁRIA (TRADUÇÃO)

Para a representação documentária é preciso consultar a linguagem documentária adotada pelo sistema de informação que indexa os artigos da revista, promovendo a garantia de uso do documento e evitando, na medida do possível, palavras-chaves que não compatíveis com o léxico especializado da linguagem

documentária. No exemplo acima, observa-se que algumas palavras-chaves identificadas no conteúdo do artigo foram modificadas em função do uso autorizado pela linguagem documentária adotada, como por exemplo a palavra "Professores de Educação Física" que no Thesaurus Brasileiro de Educação encontra-se como "Professores de Educação Profissional", assim como "Rede pública de ensino fundamental" colocada como "Rede pública de ensino".

CONSIDERAÇÕES FINAIS

A indexação em análise documentária, sob o ponto de vista dos sistemas de recuperação de informação, é reconhecida como a parte mais importante porque condiciona os resultados de uma estratégia de busca. *O bom ou o mau desempenho da indexação reflete-se na recuperação da informação feita através de estratégias de busca.* Isso nos leva a considerar que a recuperação do documento mais pertinente à questão da busca é aquela cuja indexação proporcionou a identificação de conceitos mais pertinentes ao seu conteúdo, produzindo uma correspondência precisa com o assunto pesquisado em estratégias de buscas em bases de dados que indexa periódicos.

Assim, o processo de indexação que resultará na determinação de palavras-chaves está vinculado a uma metodologia de análise de assunto que combina a exploração da estrutura textual do artigo com uma sistemática de identificação de conceitos por questionamento proposta por Fujita (2003) para orientar os autores de artigos da Revista Brasileira de Educação Especial na determinação de palavras-chaves em artigos científicos da área de Educação Especial tendo em vista a representação documentária do seu conteúdo para atendimento de demanda de recuperação pelo leitor especialista da área.

A metodologia indicada para orientação dos autores tem fundamentação teórica no tratamento temático de conteúdo para sistemas de recuperação da informação com enfoque na variável texto e na análise conceitual para determinação de palavras-chaves tendo em vista a estrutura e a tipologia textual do artigo científico.

Neste contexto, ressalta-se o conceito de palavra-chave, cujo significado advém do uso de linguagens documentárias utilizadas como instrumentos de representação documentária na indexação e na recuperação da informação científica em bases de dados. A palavra-chave, portanto, é resultado da representação documentária realizada mediante compatibilização com linguagens documentárias submetida aos termos identificados e selecionados durante a análise de assunto do artigo pelo autor-leitor.

Conscientes da necessidade de determinação de palavras-chaves pertinentes à representação do conteúdo do artigo e da estratégia de busca, recomendamos ao Conselho Editorial da Revista Brasileira de Educação Especial a adoção da metodologia de indexação indicada neste artigo bem como o uso de linguagem documentária adequada à área de Educação Especial.

REFERÊNCIAS

- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. *NBR 12676: Métodos para análise de documentos - determinação de seus assuntos e seleção de termos de indexação*. Rio de Janeiro, 1992. 4 p.
- CARNEIRO, M. V. Diretrizes para uma política de indexação. *R. Esc. Bibliotecon. UFMG*, Belo Horizonte, v. 14, n. 2, p. 221-241, 1985.
- CAVALCANTI, M. C. *Interação leitor-texto: aspectos de interpretação pragmática*. Campinas: UNICAMP, 1989. 271 p.
- CESARINO, M.A.N. Sistemas de recuperação da informação. *Revista da Escola de Biblioteconomia da UFMG*, Belo Horizonte, v.14, n.2, p.157-168, set. 1985.
- CHAUMIER, J. *Travail et methodes du/de la documentaliste: connaissance du problème*. Paris : ESF/Libraries Techniques. Exposé 3, Chap.3: L'indexation, p. 42-7, 1980.
- FUJITA, M. S. L. *A leitura documentária do indexador: aspectos cognitivos e lingüísticos influentes na formação do leitor profissional*. 2003, 321f. Tese (Livre-Docência em Análise Documentária e Linguagens Documentárias Alfabéticas) – Faculdade de Filosofia e Ciências da UNESP.
- GIASSON, J. (1993) *A compreensão na leitura*. Lisboa: Asa, 1993. 317 p.
- GUINCHAT, C.; MENOU, M. *Introdução geral às ciências e técnicas da informação e documentação*. 2.ed. rev. aum. Brasília: IBICT, 1994.
- HOUGHTON, B., CONVEY, J. *On-line information retrieval systems: an introductory manual to principles and practice*. London: Clive Bingley, 1977.
- KATO, M. *No mundo da escrita: uma perspectiva psicolingüística*. São Paulo: Ática, 1986. 144p.
- KOBASHI, N. Y. *A elaboração de informações documentárias: em busca de uma metodologia*. 1994. 195 f. Tese (Doutorado em Ciências da Comunicação) – Escola de Comunicação e Artes, Universidade de São Paulo, São Paulo.
- LANCASTER, F.W. *Information retrieval systems: characteristics, testing and evaluation*. New York: John Wiley & Sons, 1968 apud CARNEIRO, M.V. Diretrizes para uma política de indexação. *Revista da Escola de Biblioteconomia da UFMG*, Belo Horizonte, v.14, n.2, p.221-241, set. 1985.
- LOPES, E. F. Avaliação de serviços de indexação e resumo: critérios, medidas e metodologia. *R. Esc. Bibliotecon. UFMG*, Belo Horizonte, v.14, n.2. p.242-256, set. 1985.
- PINTO, M., GÁLVEZ, C. *Análisis documental de contenido: procesamiento de información*. Madrid : Síntesis, [1996]. 158p.
- TÁLAMO, M. F. G. M. *Elaboração de resumos*. Escola de Comunicação e Artes, 1987. 14 f. Datilografado.
- VAN DIJK, T. A *La ciência del texto: um enfoque interdisciplinar*. Tradução de Sibila Hunzinger. Barcelona: Paidós, 1992. 309 p. Tradução de: Tekstwetenschap. Een Interdisciplinaire inleiding.

Recebido em 17/11/2004

Aceito em 21/12/2004

FUJITA, M. S. L.