

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE BIBLIOTECONOMIA E COMUNICAÇÃO
DEPARTAMENTO DE CIÊNCIAS DA INFORMAÇÃO

Verônica da Silva Ezequiel

TERMINOLOGIA DA SUBÁREA DE REPRESENTAÇÃO TEMÁTICA:

uma análise em *corpus* textual

Porto Alegre
2004

Verônica da Silva Ezequiel

TERMINOLOGIA DA SUBÁREA DE REPRESENTAÇÃO TEMÁTICA:

uma análise em *corpus* textual

Monografia elaborada como requisito para obtenção do título de Bacharel em Biblioteconomia da Faculdade de Biblioteconomia e Comunicação, da Universidade Federal do Rio Grande do Sul

Orientadora: Prof. Dra. Regina H. Van der Laan

Porto Alegre
2004

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
Reitor Prof. Dr. José Carlos Ferraz Henemann
Vice-Reitor: Pedro Cezar Dutra Fonseca

FACULDADE DE BIBLIOTECONOMIA E COMUNICAÇÃO
Diretora: Dr^a Márcia B. Machado
Vice-Diretor: Prof. Ricardo S. da Silva

DEPARTAMENTO DE CIÊNCIAS DA INFORMAÇÃO
Chefe: Prof. Dr. Valdir José Morigi
Chefe Substituta: Prof^a Ms Itália Maria Falceta da Silveira

E98 t Ezequiel, Verônica da Silva

Terminologia da Subárea de Representação Temática:
uma análise em corpus textual / Verônica da Silva Ezequiel;
professora orientadora Regina H. Van der Laan. - Porto
Alegre, 2004.
72f.

Monografia para Trabalho de Conclusão de Curso (TCC) na
Universidade Federal do Rio Grande do Sul (UFRGS).

1. Representação Temática 2. Terminologia
I. Van der Laan, Regina H. II. Título

Departamento de Ciências da Informação
Rua Ramiro Barcelos, 2705
CEP: 90035-007
Tel: (51) 3316 5146
Fax: (51) 3316 5435
E-mail: fabico@ufrgs.br

AGRADECIMENTOS

À minha família pelo apoio e paciência.

À Inês M. de Gasperin e Márcio Ezequiel pelas observações.

À Maria Luíza F. de Campos pelo apoio durante a apresentação.

À incansável professora Regina H. Van der Laan pela orientação.

“Pensar é esquecer diferenças, é generalizar, abstrair”.

Jorge Luis Borges (Funes, o Memorioso)

RESUMO

Estudo da terminologia empregada na subárea de Representação Temática mediante a coleta de termos em *corpus* textual. Tem como pressuposto a existência de variações terminológicas empregadas na referida subárea. Conta com o aporte teórico da Teoria Comunicativa da Terminologia (TCT). A metodologia para execução do estudo constitui-se das seguintes etapas: seleção dos artigos referentes à Representação Temática que compuseram o *corpus* textual da pesquisa; coleta dos termos empregados em Representação Temática; identificação de variantes terminológicas; análise dos dados obtidos. A análise dos dados permitiu a classificação dos termos e variantes terminológicas em categorias específicas. Os resultados obtidos apontam que algumas dessas variações podem interferir na compreensão da terminologia empregada na subárea bem como a representação e recuperação da informação. Ressalta a importância do uso de vocabulário controlado e remissivas para o processo de representação e recuperação da informação. Sugere a elaboração de glossários de termos empregados em Representação Temática.

Palavras-chave: Terminologia. Linguagem Especializada. Biblioteconomia. Representação Temática.

ABSTRACT

A study of the terminology used in the field of Thematic Representation, through the collection of terms in a textual *corpus*. It is based on the presupposition that there are terminological variations employed in this field. The Communicative Theory of Terminology (TCT) grounds it. The methodology used for this study has been constituted by the following steps: selection of the articles referring to Thematic Representation that have made up the textual *corpus* of the research; collection of the terms used in Thematic Representation; identification of terminological variants; analysis of the data obtained. Data analysis has permitted the classification of the terms and terminological variants in specific categories. The results obtained signal that some of these variants may interfere in the understanding of the terminology used in this field as well as the information representation and retrieval. It emphasizes the importance of using controlled vocabulary and remissives in the process of information representation and retrieval. It suggests making up glossaries of the terms used in Thematic Representation.

Key words: Terminology. Specialized Language. Library Science. Thematic Representation.

SUMÁRIO

1 INTRODUÇÃO	9
1.1 Definição do Problema	11
1.2 Objetivo Geral	11
1.3 Objetivos Específicos	11
2 PROCEDIMENTOS METODOLÓGICOS	13
2.1 <i>Corpus</i> Textual da Pesquisa	14
2.2 Instrumento de Coleta de Dados	15
3 REFERENCIAL TEÓRICO	16
3.1 Ciências da Informação e Representação Temática	16
3.2 Terminologia	19
3.2.1 Conceito	21
3.2.2 Termo	24
3.2.3 Teoria Geral da Terminologia (TGT)	27
3.2.4 Teoria Comunicativa da Terminologia (TCT)	28
4 ANÁLISE DOS DADOS	31
4.1 Flexão de Número	37
4.2 Variação Concorrente Lingüística	38
4.2.1 Variantes Terminológicas Sintáticas	39
4.2.2 Variantes Terminológicas Lexicais	40
4.2.3 Variação Terminológica Gráfica	42
4.3 Variantes Coocorrentes	43
4.3.1 Descritor(es), Termo(s) e Termos de Indexação	44
4.3.2 Linguagem Controlada, Linguagens Controladas, Linguagem de Indexação...	45
4.3.3 Organização da Informação e Organização do Conhecimento	48
4.3.4 Indexação, Análise Documentária...	50
4.3.5 Sistema de Recuperação da Informação e Sistemas de Busca	54
4.3.6 Sistema de Classificação, Sistemas de Classificação...	55
4.3.7 Tratamento da Informação, Processamento da Informação...	57
4.4 Ausência de Variação Terminológica	58

4.5 Termos Pertencentes a Outras Áreas Empregados em Representação	64
5 CONCLUSÕES	66
REFERÊNCIAS	69
APÊNDICE A – Modelo de Ficha Terminológica	72

1 INTRODUÇÃO

A análise da literatura especializada permite observar a existência de ambigüidades na linguagem empregada em Biblioteconomia. Tais ambigüidades são geradas pelo uso de diferentes denominações para a representação dos mesmos conceitos, o que acarreta dificuldade na compreensão da terminologia empregada na área, bem como no processo de representação e recuperação da informação.

Pode ser aqui citado um exemplo bastante evidente do uso de diferentes termos para designação de um único conceito. Em *Introdução à Catalogação* (MEY, 1995, p.68) encontramos o termo *indexação* como sendo equivalente à “análise e representação do conteúdo de um item, ou representação temática”. Já em *Princípios de Indexação* Lancaster (1993, p.17) afirma que a “distinção entre as expressões *catalogação de assuntos* e *indexação de assuntos* [...] é artificial, enganosa e incongruente”.

Temos, com este exemplo, cinco termos: *análise do conteúdo*, *representação do conteúdo*, *representação temática*, *catalogação de assuntos* e *indexação de assuntos* para a representação de um mesmo conceito - *indexação*. A percepção dessas variações na representação dos conceitos instigou o estudo da terminologia empregada na área.

A ocorrência de variações terminológicas pode estar relacionada à existência de termos ainda não consolidados empregados em Biblioteconomia. Essa problemática em torno da consolidação terminológica tem sido abordada por pesquisadores em Ciência da Informação. Smit et al. (2004, p.2) cita a “ [...]

ausência de consolidação do corpo conceitual utilizado na análise e produção discursiva da área [Ciência da Informação]” como um importante obstáculo à constituição deste campo científico.

O termo técnico científico é o objeto primordial de estudo da Terminologia enquanto disciplina (KRIEGER, 2004). Foi com aporte teórico da Terminologia que se fez este estudo, mediante a coleta de termos da área, identificação de possíveis variações terminológicas e análise dos dados obtidos.

Durante o processo de delimitação do tema de pesquisa, houve uma preocupação quanto à abrangência da área de especialidade a ser estudada, por isso se fez necessária a restrição do tema de pesquisa, garantindo assim a qualidade do trabalho. Para tanto foi utilizado como referência o *Currículo do Curso de Biblioteconomia do Ano de 2000* (COMISSÃO DE GRADUAÇÃO DE BIBLIOTECONOMIA, 2000) da Universidade Federal do Rio Grande do Sul (UFRGS). Neste documento as disciplinas do curso estão organizadas em quatro áreas, a saber:

- a) área 1: Fundamentos das Ciências da Informação;
- b) área 2: Organização e Tratamento da Informação;
- c) área 3: Recursos e Serviços de Informação;
- d) área 4: Gestão de Sistemas de Informação.

Para fins desse estudo a área de abrangência foi delimitada na área 2: Organização e Tratamento da Informação. Sendo essa área ainda muito ampla, a pesquisa deteve-se ao estudo da terminologia empregada na subárea de Representação Temática.

1.1 Definição do Problema

O problema de pesquisa foi sintetizado na seguinte pergunta: a terminologia empregada no discurso dos especialistas da subárea de Representação Temática é consistente, não ocorrendo ambigüidades que possam dificultar a compreensão dos conteúdos da área e a recuperação da informação?

1.2 Objetivo Geral

O objetivo norteador deste estudo foi o seguinte: coletar termos, em situação discursiva, empregados em Representação Temática, verificando a existência de variações terminológicas.

1.3 Objetivos Específicos

- a) Selecionar os artigos referentes à Representação Temática que constituirão o *corpus* textual para coleta de termos;
- b) coletar, em *corpus* textual, termos empregados na área de Representação Temática;
- c) identificar variantes terminológicas;

d) analisar os dados obtidos.

2 PROCEDIMENTOS METODOLÓGICOS

Essa pesquisa parte do pressuposto de que ocorrem variações terminológicas empregadas no discurso referente à subárea de Representação Temática.

Tem como embasamento teórico a Teoria Comunicativa da Terminologia (TCT), que se apresenta como uma nova proposta teórica em relação à Teoria Geral da Terminologia (TGT), permitindo compreender o termo em uma dimensão comunicativa.

A TCT oferece fundamentação teórica para este estudo já que “[...] inclui também a apreciação da variação conceitual e denominativa, tendo-se em conta as dimensões comunicativa e discursivo-textuais” (KRIEGER, 2004).

A coleta dos termos da área foi possível mediante a aplicação de alguns procedimentos essenciais à pesquisa em Terminologia. Arntz (1995) apresenta algumas etapas para realização da pesquisa terminológica, dentre as quais, foram consideradas relevantes para este estudo as seguintes:

- a) considerações preliminares de organização – definição da área de estudo, importância do trabalho, decisão sobre a forma de apresentação do produto final;
- b) delimitação da área especializada;
- c) aquisição e análise do *corpus* textual;
- d) compilação e ordenação dos termos – elaboração de uma ficha para registro dos termos coletados (instrumento de coleta de dados);

e) análise terminológica.

As etapas referentes ao *corpus* textual e ao instrumento de coleta de dados serão tratadas separadamente a seguir.

2.1 *Corpus* Textual da Pesquisa

O *corpus* textual foi constituído por artigos referentes à subárea de Representação Temática. Esses artigos foram selecionados nos seguintes periódicos: *Revista Ciência da informação* e *Data Grama Zero – Revista de Ciência da Informação*. O período abrangido pela pesquisa foi de cinco anos, foram selecionados artigos entre 1999 e 2003.

Dois critérios foram determinantes para a escolha desses periódicos: ambos estão disponíveis gratuitamente na internet, o que em muito facilitou o trabalho de coleta de termos, e ambos foram avaliados pela Capes, no programa *Qualis*, com nível A.

O periódico *Ciência da Informação* é publicado desde 1972, sob a responsabilidade do Instituto Brasileiro de Informação em Ciência e Tecnologia – IBICT, com periodicidade quadrimestral. Esse periódico, devido a sua qualidade, é consagrado na área de Biblioteconomia. Desde 1996, além da versão impressa, está disponível na internet a versão eletrônica do periódico.

O periódico *Data Grama Zero – Revista de Ciência da informação* é de responsabilidade do Instituto de Adaptação e Inserção na Sociedade da Informação - IASI, uma organização não-governamental sem fins lucrativos. Seus fascículos são

disponibilizados bimestralmente na internet desde 1999. O periódico não possui versão impressa.

Foram identificados treze artigos referentes à subárea de Representação Temática, publicados por esses dois periódicos entre os anos de 1999 e 2003, que constituíram o *corpus* textual desta pesquisa. Quatro desses artigos foram publicados pelo periódico *Ciência da Informação* e nove pelo *Data Grama Zero – Revista de Ciência da Informação*.

2.2 Instrumento de Coleta de Dados

Para o registro dos dados coletados foi elaborada uma ficha terminológica. Essa ficha foi estruturada nos seguintes campos: definição, contexto, siglas, sinônimos e variações. Os dados registrados nesses campos subsidiaram a contextualização e análise dos termos.

No campo *definição* foi registrado o significado do termo, expresso pelo autor, no texto. Já no campo *contexto* foram compilados todos os aspectos relevantes para o entendimento do emprego do termo nos artigos. No campo de *variações* foram registradas as variações terminológicas encontradas nos artigos.

Outros dados tais como *siglas*, *sinônimos*, *área* e *subárea* referentes ao termo coletado, *fonte* de cada artigo, espaço para *observações*, registros do nome do *compilador* e *data* da coleta também constam nessas fichas, cujo modelo encontra-se em apêndice (APÊNDICE A).

3 REFERENCIAL TEÓRICO

Com o objetivo de obter embasamento teórico para a elaboração deste trabalho, foi feita uma revisão da literatura sobre o assunto. Esta revisão não pretende ser exaustiva, mas sim capaz de definir conceitos básicos de interesse desta pesquisa.

3.1 Ciências da Informação e Representação Temática

Em *Terminologia Relacionada com los Estudios de Usuarios y la Formación de Usuarios de la Información* (1997, p.1) temos que o termo *informação* pode ser definido como “ [...] conjunto de símbolos com a capacidade de significar, estar registrada em diferentes suportes e com potencial para ser recuperada.”

As ciências que têm por objeto de estudo a informação, procurando tratá-la para posterior recuperação, são denominadas por Smit (2000) de *Três Marias*: Arquivologia, Biblioteconomia e Museologia.

De acordo com a autora, é necessário que se faça um recorte, já que nem toda a informação interessa às *Três Marias*. A informação registrada em um suporte tem a possibilidade de ser armazenada em um tipo de acervo, e desse modo nos remete a idéia de arquivos, bibliotecas e museus.

Porém, não basta que as informações sejam registradas para que sejam preservadas, é papel das Ciências da Informação avaliar o valor potencial de tais informações, tratando e recuperando aquilo que é de fato importante para determinada comunidade usuária ou de importante valor para a sociedade em geral.

Têm sido constantemente discutido na literatura os limites entre as Ciências da Informação e suas diferenças e especificidades. Smit (2000, p.122) afirma que demarcar os limites entre uma ciência e outra, de acordo com o tipo de documento tratado, não é suficiente e explica que:

[...] tradicionalmente, as bibliotecas guardam livros e periódicos (e alguns outros documentos, desde que considerados pertinentes aos objetivos da instituição), os arquivos, por sua vez, as informações geradas pelas instituições no cumprimento de suas atividades (os documentos 'administrativos') os museus guardam objetos... Duas categorias de documentos – os eletrônicos e os audiovisuais – bastam para demonstrar a fragilidade da divisão por documentos e sua inadequação aos tempos atuais.

A autora segue argumentando que a ênfase que deve ser dada a cada uma das Ciências da Informação não está relacionada ao tipo de documento, mas à função pela qual as informações são armazenadas. Assim, os arquivos preservam as informações ligadas à história, funcionamento e relação das instituições entre si e com a sociedade. As bibliotecas preocupam-se com a disponibilização da informação propriamente dita, e os museus preservam registros daquilo que a sociedade produziu ou utilizou.

Este trabalho irá se deter à ótica biblioteconômica e sua terminologia, é portanto oportuno fazer aqui algumas considerações sobre esta ciência em especial. A Biblioteconomia tem passado por constantes transformações e mudanças de paradigmas. De acordo com Le Coadic (1996, p.15):

À biblioteca tradicional, que conservava apenas livros, sucedeu a biblioteca que reúne acervos muito mais diversificados, tanto por seus suportes como por sua origem: imagens, sons, textos. Transformou-se em mídioteca. Ademais, ao acolher não somente as obras de patrimônio legado pelo passado, mas as informações veiculadas por redes comerciais atuais e em tempo real, ela passou a ser um sistema de informações.

Temos com isto a reafirmação de que a Biblioteconomia, bem como as outras Ciências da Informação, não pode ser caracterizada pelo tipo de documento que preserva. O autor aponta ainda o papel dinâmico desempenhado pelas bibliotecas atuais, oferecendo aos usuários informações precisas e em tempo real, ajustando-se a velocidade com que as informações são veiculadas na atualidade (através dos meios de comunicação e da internet).

A terminologia empregada em Biblioteconomia foi o foco deste estudo. A delimitação da área de estudo se fez necessária, portanto, foram coletados e analisados termos empregados na área de Organização e Tratamento da Informação, mais especificamente na subárea de Representação Temática.

De acordo com Kobashi (1996, p.11) “‘representação’ é um conceito mediador entre o emissor e o receptor. É entendido, também, como o processo e o produto social da construção de sentidos, por meio de qualquer sistema de significação”.

Van der Laan et al. (2002, p.6) explica que:

A mediação entre quem deseja a informação e quem a tem é um complexo processo comunicativo duplamente codificado. O bibliotecário codifica a informação utilizando uma linguagem de indexação. [...] O usuário utiliza seu código lingüístico e seu conhecimento para construir as questões de busca de informação. Nessa dupla codificação, é necessário que ambos os códigos utilizados sejam coincidentes para o sucesso nas buscas de informação.

Assim, temos que a informação é representada por um sistema de códigos cujo objetivo é a recuperação rápida e precisa da informação. Com a finalidade de

representar as informações, as unidades de informação valem-se de vocabulários controlados, ou linguagens de indexação, que padronizam os termos utilizados na indexação de documentos. Os tesouros e as listas de cabeçalhos de assunto são exemplos de controle de vocabulário para representação e recuperação da informação.

Portanto, a representação da informação, ou representação temática, pode ser considerada um processo que, mediante um sistema de significação, é capaz de servir de elo entre a mensagem e seu receptor; entre a informação e o usuário.

3.2 Terminologia

Embora o desenvolvimento de estudos teóricos acerca da Terminologia tenha se evidenciado a partir do século XX, não é possível afirmar que este campo de estudos seja recente. Rondeau (1984 apud KRIEGER; FINATTO, 2001, p.34)¹ explica que:

A Terminologia não é um fenômeno recente. Com efeito, tão longe quanto se remonte na história do homem, desde que se manifesta a linguagem, nos encontramos em presença de línguas de especialidade, é assim que se encontra a terminologia dos filósofos gregos, a língua de negócios dos comerciantes cretas, os vocábulos especializados da arte militar, etc.

A terminologia enquanto linguagem especializada acompanha o homem e sua história. Van Hoof (1989 apud ALVES, 2004)² aponta oportunamente alguns títulos:

a) Idade Média:

¹ RONDEAU, 1984 apud KRIEGER; FINATTO, 2001, p.34 .

² VAN HOOFF, 1989. apud ALVES, 2003.

- *Explicação das Palavras Gregas em Siríaco*, de Hunayn Ibn Ishâq, médico e filósofo;

- *Léxico Siríaco-Árabe*, de Ibn Bahlûl;

- *O Grande Colecionador*, de Rhazès;

- *O Livro da Explicação das Designações de Drogas*, de Maimonide (1139-1204).

b) Renascimento:

- *Glossário Árabe-Latino de Termos Médicos*, do médico italiano Andrea Alpago (?-1520);

- *Livro dos Segredos da Agricultura* (1617), do religioso espanhol Miguel Agustí.

Nos séculos XVIII e XIX, surge a preocupação dos cientistas quanto à disseminação de novos termos e a decorrente dificuldade na comunicação entre os mesmos. Surge então, no século XX, a Terminologia Moderna, com os estudos de Eugen Wüster e a Teoria Geral da Terminologia (TGT), que será tratada em um item específico mais adiante neste trabalho.

A terminologia é considerada um termo polissêmico, já que possui três diferentes acepções:

a) enquanto prática – coleta de termos de uma área de especialidade de acordo com metodologias pré estabelecidas;

b) enquanto produto – conjunto de termos específicos de uma área de conhecimento;

c) enquanto disciplina – estudo que designa os conceitos de uma área de especialidade.

Temos com isto que a importância da Terminologia se dá nos mais diferentes âmbitos, interessando de modo geral, a todos os integrantes (técnicos, especialistas, cientistas) de uma área de conhecimento, à medida que, para se comunicar, os mesmos fazem uso de uma linguagem de especialidade.

3.2.1 Conceito

Termo e conceito são elementos essenciais ao estudo das linguagens especializadas, portanto é imprescindível que o significado destas expressões sejam compreendidas. Segundo Van der Laan (2003, p.1) “[...] conceito é uma abstração, uma representação da realidade, que os indivíduos elaboram mentalmente”. Para Dubuc (1999, p.58) conceito é a “[...] reunião dos traços característicos do objeto designado pelo termo. Supõe, portanto, uma referência à realidade”.

Considerando as afirmações acima, podemos concluir que conceito é uma elaboração mental feita pelo indivíduo para representar a realidade que o cerca, podendo assim, melhor compreendê-la. Para formar um conceito, cada indivíduo observa seu objeto de análise, procurando em seu repertório, experiência de vida, características já conhecidas para descrevê-lo.

Esse conjunto de características é formado por aspectos inteligíveis que tornam o conceito único e impossível de ser confundido. Este grupo de caracteres ou notas denomina-se *intensão* ou *compreensão conceitual*. (Maritain, 1989).

Ao considerarmos o conceito *homem* podemos tomar como conjunto de características que o diferenciam de qualquer outro as seguintes: animal, mamífero,

bípede, racional. Esta é a *intensão* do conceito, ou seja, é a amplitude do conceito em relação às suas características.

Ao nos referirmos ao conceito *brasileiro* notamos que este é de maior intensão do que o conceito *homem*, já que para *brasileiro* são atribuídas as mesmas características que para *homem* e ainda as suas próprias: homem, latino americano, natural de um país chamado Brasil.

Já a *extensão conceitual* refere-se à amplitude do conceito em relação aos indivíduos aos quais se aplica ou se relaciona. Assim estende-se o conceito *homem* a seres mamíferos, bípedes e racionais; o conceito *brasileiro* a homem, latino americano, natural de um país chamado Brasil. Aqui podemos então concluir que o conceito *brasileiro* tem menor extensão que o conceito *homem*. Sendo a intensão do conceito *brasileiro* maior que a do conceito *homem* (já que o conceito *brasileiro* pressupõe as características do conceito *homem* e ainda as suas próprias) e a extensão do conceito *brasileiro* menor que a do conceito *homem* (o número de indivíduos aos quais o conceito *brasileiro* se aplica é menor do que os relacionados ao conceito *homem*), percebemos, então, uma relação inversa entre *intensão* e *extensão* conceitual: quanto maior a extensão de um conceito menor sua intensão ou compreensão.

Ao identificarmos características em comum entre conceitos devemos considerar a existência de relações entre os mesmos. Para Dahlberg (1978) estas relações podem ser lógicas, hierárquicas, partitivas, de oposição e funcionais. Segundo a autora as relações lógicas podem ocorrer por:

- a) identidade - designam equivalência entre dois ou mais conceitos, possuindo estes o mesmo conjunto de características, mesma intensão ou compreensão;

- b) implicação – um conceito está contido em outro;
- c) intersecção – um determinado elemento coincide entre dois conceitos;
- d) disjunção – dois ou mais conceitos não possuem qualquer característica em comum;
- e) negação – um conceito possui determinada nota cuja negação encontra-se na intensão de outro dado conceito.

As relações hierárquicas pressupõem que um elemento pode estar subordinado a outro possuindo todas as características do elemento mais genérico e pelo menos mais uma que o diferencia. Por exemplo: biblioteca - biblioteca pública. Uma relação partitiva entre conceitos se dá entre um todo e suas partes. Por exemplo: livro - capa, folhas, lombada.

A relação de oposição conceitual ocorre por contradição: branco/preto; numérico/não-numérico; vertebrado/invertebrado; racional/irracional, ou por gradação: alto - igual - baixo (DODEBEI, 2002, p.94). Finalmente, as relações funcionais, ou de intersecção, procuram expressar conceitos referentes a processos. Dahlberg (1978) exemplifica da seguinte forma: produção – produto – comprador.

As relações entre conceitos podem ser classificadas também como intrínsecas ou extrínsecas. Na relação intrínseca temos que: “[...] o conceito A inclui o conceito B sem ser idêntico a este último” (DUBUC, 1999), esta relação parte do genérico para o específico, por exemplo: *profissional da informação e bibliotecário*, o primeiro conceito inclui o segundo, porém não equivale ao mesmo.

De acordo com o autor, uma relação extrínseca pode ser expressa pelo “[...] agrupamento dos conceitos segundo as funções da área de especialidade estudada”. Assim, se pretendemos classificar o conceito *organização e tratamento da informação* podemos fazê-lo com base em cada uma das funções desta área

sejam elas: representação temática (indexação, classificação), representação descritiva (catalogação) e geração de produtos de recuperação da informação (elaboração de tesouros, por exemplo).

Existem ainda as relações termo-conceito, porém antes de elucidá-las é conveniente tratarmos deste outro elemento constitutivo do estudo da Terminologia – o termo.

3.2.2 Termo

Termos são unidades lexicais que representam conceitos de uma determinada área de especialidade, respeitando a estrutura do sistema lingüístico no qual estão inseridos.

Sager (1993, p. 91) afirma que “[...] os termos são as representações lingüísticas dos conceitos”. Segundo Cabré, (1993, p.170) termos são “[...] unidades de forma e conteúdo que pertencem ao sistema de uma língua determinada”.

Para que uma unidade lexical adquira estatuto de termo é necessário que seu sentido independa de um contexto. Krieger (2004, p.77) se refere a esta propriedade do termo como *invariabilidade semântica* e explica que:

[...] enquanto o significado que uma palavra adquire é, em larga medida, dependente do contexto discursivo em que se insere, as unidades terminológicas não sofrem esses efeitos porquanto se limitam a expressar conteúdos das ciências e das técnicas.

Ainda a este respeito Fedor de Diego (1995, p.54, tradução nossa) afirma que:

[...] a palavra é um símbolo lingüístico que admite variações semânticas e depende do contexto; o termo apresenta um grau de precisão muito mais elevado e pertence a um sistema de conceitos determinado.

Quanto à relação existente entre termo e conceito, a autora afirma que o ideal é a univocidade absoluta, ou seja, “[...] cada conceito se relaciona unicamente a um termo e cada termo refere-se a um só conceito”. (1995, p.57, tradução nossa). Entretanto, algumas variações podem surgir durante o discurso entre falantes de um mesmo idioma, serão citados aqui: a homonímia, a polissemia e a sinonímia.

A homonímia é uma relação possível entre termo e conceito, trata-se de um termo que designa dois ou mais conceitos que semanticamente não tem relação entre si. Por exemplo: cadeira – objeto e cadeira – disciplina acadêmica (usado no sul do país). A polissemia assim como a homonímia também se refere a um termo que designa dois ou mais conceitos, porém estes conceitos possuem uma relação semântica. Por exemplo: asa do pássaro e a asa do avião (FEDOR DE DIEGO, 1995).

A sinonímia é, segundo a autora, o fenômeno menos bem visto na terminologia, já que indica a existência de dois ou mais termos do mesmo idioma se referindo a um só conceito, provocando os maiores problemas de ambigüidade na comunicação.

Conforme Cabré (1993) os termos podem ser classificados de acordo com sua *forma, função, significado e procedência*. Pela forma os termos podem ser classificados, de acordo com o número de morfemas, em simples e complexos, assim: sangue (termo simples) – sangüíneo (termo complexo); câncer (termo simples) – cancerígeno (termo complexo). Os termos complexos podem ainda, segundo os tipos de morfema que intervém na formação dos mesmos, ser classificados em derivados (formados mediante a junção de afixos a bases lexicais)

e compostos (formados pela combinação de bases lexicais, atuais ou históricas, existindo ainda a possibilidade de acréscimo de afixos às mesmas).

Ainda segundo a autora os termos complexos podem ser formados por uma combinação de palavras que seguem uma determinada estrutura sintática, trata-se dos sintagmas terminológicos. Tais sintagmas, de acordo com Van der Laan (2002), são geralmente formados pela combinação de: substantivo e adjetivo; substantivo, preposição e substantivo; substantivo e substantivo.

Cabré (1993) indica, do ponto de vista da *função*, que os termos podem ser classificados como: substantivos, adjetivos, verbos e advérbios, sendo a ocorrência de substantivos muito maior.

Segundo o *significado* os termos podem ser classificados conforme a classe de conceitos que denominam. Os conceitos podem ser agrupados em classes e subclasses de acordo com características que possuem em comum e de acordo com as relações estabelecidas entre eles.

Finalmente, de acordo com a *procedência lingüística*, os termos podem ser construídos a partir da aplicação de regras do seu próprio código lingüístico ou por empréstimo de outro código.

Deste modo, percebemos que, os termos são as unidades básicas para a formação da terminologia de uma área. Os termos representam os conceitos de uma linguagem especializada, sendo estes últimos representações abstratas da realidade.

3.2.3 Teoria Geral da Terminologia (TGT)

Nos séculos XVIII e XIX, surge a preocupação por parte da comunidade científica quanto à disseminação de novos termos e a decorrente dificuldade na comunicação entre os especialistas. Surge então, no século XX, a terminologia moderna, com os estudos do engenheiro austríaco Eugen Wüster que publica em Viena a obra *A Normalização Internacional da Terminologia Técnica* (1931) que tratava dos termos empregados em eletro-técnica.

A obra de Wüster, fruto da sua tese de doutorado, teve por objetivo eliminar as ambigüidades existentes na comunicação científica através do estabelecimento de metodologia e normalização dos termos. Nela são expostos os princípios que devem nortear o estudo de termos e é a partir da mesma que surge a Teoria Geral da Terminologia (TGT)

A TGT visa a normalização dos termos, estabelecendo o uso de um único termo, em detrimento de outros, para representar um conceito. Entende-se que a univocidade auxilia no processo de comunicação de especialistas de uma mesma área do conhecimento.

Wüster aponta algumas diferenças entre a Lingüística em geral e a Terminologia, indicando os três fatores a seguir como responsáveis por tais diferenças:

- a) quanto ao ponto de partida – a Terminologia parte dos conceitos, procurando delimitá-los claramente, enquanto a Lingüística ocupa-se do significado das palavras. Isto se explica pelo fato de que o significado de uma palavra poder conter várias acepções, dependendo do contexto na qual a mesma está

inserida, já os conceitos devem ser representados, em princípio, por um único termo. A univocidade e a monorreferencialidade são preceitos básicos da TGT;

- b) limitação do léxico – a Terminologia preocupa-se com o léxico, desprezando o estudo da morfologia flexiva e da sintaxe;
- c) enfoque sincrônico – para a terminologia interessam os termos usuais na atualidade. Termos em desuso (diacrônicos) não interessam (VAN DER LAAN, 2003).

Ainda conforme a autora, para os teóricos da TGT, o desenvolvimento de uma linguagem de especialidade acontece de modo consciente e não espontâneo, diferentemente da língua em geral. De acordo com a TGT a normalização da terminologia empregada nas linguagens especializadas é necessária para que sejam evitadas confusões e ambigüidades que poderiam ser geradas através da livre criação terminológica.

3.2.4 Teoria Comunicativa da Terminologia (TCT)

Maria Teresa Cabré, juntamente com o grupo de pesquisadores do Instituto de Lingüística Aplicada da Universidade Pompeu Fabra em Barcelona, propôs a Teoria Comunicativa da Terminologia (TCT).

Diferentemente da TGT, a TCT prioriza os aspectos comunicativos da linguagem especializada, considerando que o conteúdo de um termo não é fixo e varia, portanto, de acordo com o “[...] cenário comunicativo em que se inscreve”

(KRIEGER, 2004, p.35). Deste modo, uma unidade lexical assume o estatuto de termo em um contexto específico.

Conforme a TCT as unidades lexicais sofrem as interferências comuns à linguagem natural comportando assim sinonímias e variações. Van der Laan (2002, p.49) explica que: “[...] é possível observar [...] quer no discurso oral, quer no escrito dos especialistas, a existência de uma expressiva criação terminológica”. Esta criação terminológica enriquece a linguagem de especialistas e é parte da constante evolução das áreas de especialidade.

Faulstich (2002) aponta dois grupos de variantes terminológicas: concorrentes e coocorrentes. As variantes concorrentes ocorrem simultaneamente no discurso especializado e concorrem entre si até que uma dessas variantes se consolide na área. As variantes concorrentes podem ser lingüísticas ou de registro.

As variantes terminológicas lingüísticas classificam-se em:

- a) fonológicas – surgem a partir da fala;
- b) morfológicas - apresentam alternância morfológica sem que o conceito se altere;
- c) sintáticas – apresentam alternância entre duas construções sintagmáticas funcionando como predicação de uma *Unidade Terminológica Comunicativa*;
- d) lexicais – um item da estrutura lexical é apagado sem que o termo seja alterado;
- e) gráfica – apresentam mais de uma forma gráfica.

As variantes terminológicas de registro são classificadas pela autora em:

- a) geográficas – termos diferentes que designam o mesmo conceito utilizados por falantes da mesma língua, porém de regiões diferentes;

- b) de discurso – termos utilizados de acordo com os níveis de discurso, estes níveis podem ser científico, técnico ou de divulgação.
- c) temporais – termos diferentes que remetem ao mesmo conceito e que são utilizados simultaneamente até que um deles caia em desuso.

As variantes coocorrentes são aquelas referentes à ocorrência de sinonímia no discurso especializado, isto é, um conceito pode ser representado por vários termos e um termo pode remeter a mais de um conceito.

Esse estudo contou com suporte teórico da TCT, já que pressupõe a existência de variações terminológicas na subárea de Representação Temática.

4 ANÁLISE DOS DADOS

Através da leitura dos artigos selecionados foram coletados termos representativos na área de Representação Temática. Todos os termos coletados foram validados por um especialista da área. Cada termo foi coletado e registrado em uma ficha terminológica. Ao todo foram registradas cento e cinquenta e sete fichas terminológicas. Algumas delas serão apresentadas ao longo desta análise para ilustrar o uso de determinados termos em situação discursiva.

Para manter a impessoalidade dos comentários e análises aqui apresentadas, fez-se necessária a omissão da autoria e título dos artigos analisados. As fichas que ilustraram esse trabalho apresentam apenas a referência dos periódicos nos quais os artigos foram disponibilizados.

Para a coleta dos termos em cada artigo fez-se necessário o estabelecimento de alguns critérios para determinação de qual palavra ou expressão seria considerada *termo* e quais seriam *variantes* desse termo. Isso ocorreu uma vez que em um mesmo artigo, muitas vezes, o autor empregava diferentes termos para referir a mesma idéia ou conceito. Por exemplo, de um mesmo artigo foram coletados os termos *recuperação da informação*, *recuperação de informações*, *recuperação*, *recuperação de informação* e *recuperação de conteúdo*, todos representando a mesma idéia.

Para determinar qual dessas representações seria *termo* e quais seriam *variantes* foram aplicados os seguintes critérios:

- a) ocorrência – palavra ou sintagma que foi empregado com maior freqüência pelo autor no artigo foi considerado termo, os demais foram considerados como variações terminológicas;
- b) uso do singular – quando o autor empregou a representação de um conceito ora no singular, ora no plural, foi considerado *termo* a forma no singular.
- c) primeira ocorrência no texto – para os casos em que os termos tiveram o mesmo número de ocorrências no singular, foi considerado *termo* aquele que primeiro aparece no texto.

Desse modo, no artigo exemplificado anteriormente foi considerado *termo* o sintagma *recuperação da informação*, já que o mesmo foi empregado pelo autor com maior freqüência no texto. *Recuperação de informações*, *recuperação*, *recuperação de informação* e *recuperação de conteúdo* foram consideradas variantes terminológicas e registradas no campo *variante* da ficha, assim:

TERMO: RECUPERAÇÃO DA INFORMAÇÃO
Fonte: Data Grama Zero – Revista de Ciência da Informação , Rio de Janeiro, v.3, n.1, 2002. Disponível em: < http://www.dgz.org.br >. Acesso em: 25 set. 2004.
Definição:
Contexto: “(...) as pesquisas realizadas na área de recuperação da informação concentram-se, de maneira geral, no desenvolvimento de ferramentas que possibilitem a extração do conteúdo diretamente dos textos completos dos documentos disponibilizados eletronicamente”.
“A ambigüidade causa ruído na recuperação da informação , pois, sob um mesmo termo, o usuário encontrará informação relevante e irrelevante”.
“Cabe aos pesquisadores da área de Ciência da Informação acompanhar os desenvolvimentos dessas áreas e avaliar a possibilidade de aplicação e a adequação de novos métodos e técnicas à recuperação de informação ”.
Sigla:
Sinônimo:
Variante: recuperação de informações; recuperação; recuperação de informação; recuperação de conteúdo.
Área: Organização e Tratamento da Informação
Subárea: Representação Temática
Observações:
Compiladora: Verônica Ezequiel
Data: 12.10.2004

Após a coleta dos termos as fichas foram analisadas e agrupadas por conceitos, ou seja, todas as fichas cujos termos representavam o mesmo conceito foram reunidas.

Por exemplo, o termo *recuperação da informação* foi empregado em sete dos artigos analisados, nos demais foram encontrados os termos *recuperação das informações*, *recuperação de informação* e *recuperação de informações*. Embora os autores tenham empregado termos diferentes, foi possível, através da análise dos dados coletados, concluir que estes se referiam ao mesmo conceito.

Para fins deste estudo foi necessária a determinação de qual dessas expressões deveria ser considerada *termo* e quais deveriam ser entendidas como *variantes terminológicas*. Para tanto, foram obedecidos os mesmos critérios de ocorrência e uso singular, anteriormente citados, e ainda o critério de atualidade como vemos a seguir:

- a) ocorrência;
- b) uso do singular;
- c) atualidade – nos casos em que as ocorrências foram numericamente iguais, considerou-se *termo* aquele empregado no artigo mais recentemente disponibilizado.

Assim, a expressão com maior número de ocorrências no conjunto de artigos analisados, com o emprego do singular foi considerada como *termo*, enquanto as demais foram entendidas como variantes terminológicas. No caso das expressões exemplificadas anteriormente, *representação da informação* foi considerado termo uma vez que o mesmo foi empregado em sete artigos diferentes. *Recuperação das informações*, *recuperação de informação* e *recuperação de informações* foram consideradas variantes terminológicas.

Todos os termos e variantes coletados foram organizados em uma listagem que permite visualizar com clareza a ocorrência dos mesmos nos artigos analisados. Temos, portanto, na listagem a seguir, os quarenta e três termos coletados em *corpus* textual e suas respectivas variantes terminológicas. Os *termos* aparecem em destaque e precedem as *variações terminológicas*. Cada *termo* e *variante* aparece seguida do número de artigos nos quais foram empregados, como vemos abaixo:

1. **armazenamento da informação (2)**;
2. **busca da informação (2)** - busca de informação (1), busca de informações(1);
3. **cabeçalho de assunto (1)** - cabeçalhos de assuntos (1);
4. **classificação (6)** - classificação de assuntos (1);
5. **Classificação Decimal de Dewey (2)** – CDD (1);
6. **Classificação Decimal Universal (2)** – CDU (1);
7. **Classificação de Dois Pontos (1)**;
8. **conceito (2)** – conceitos (2);
9. **descriptor (2)** - descritores (4), termo (2), termos (2), termos de indexação(1);
10. **disseminação da informação (1)** - disseminação de informação (1);
11. **especificidade (1)**;
12. **esquemas de classificação bibliográfica analítico-sintéticos (1)**;
13. **esquemas de classificação bibliográfica enumerativos (1)**;
14. **esquemas de classificação bibliográfica semi-enumerativos (1)**;
15. **esquemas de representação (1)** - esquemas de representações do conhecimento (1);
16. **estratégia de busca (1)** - estratégias de busca (1);

17. **indexação (5)** - indexação por assunto (1), análise documentária (1), representação da informação (3), representação da informação documentária(1), representação de informações (1), representação do conteúdo dos documentos (1), representação do conteúdo temático (1), representação do conhecimento (1);
18. **indexação automática (2)** - indexação automatizada (1);
19. **indexação pós – coordenada (1)**;
20. **indexação pré – coordenada (1)**;
21. **instrumento de recuperação da informação (1)**;
22. **linguagem natural (6)**;
23. **listas de cabeçalhos de assunto (3)**;
24. **não-descritores (1)**;
25. **organização da informação (2)** - organização do conhecimento (1);
26. **política de indexação (1)** - políticas de indexação (1);
27. **pontos de acesso à informação (1)**;
28. **precisão (5)**;
29. **recuperação da informação (7)** - recuperação das informações (1), recuperação de informação (2), recuperação de informações (1);
30. **relevância (1)**;
31. **revocação (4)**;
32. **ruído (1)**;
33. **sistema de recuperação da informação (3)** - sistema de recuperação (1), sistemas de recuperação da informação (1), sistema de recuperação de informação (1), sistemas de recuperação de informação (2), sistemas de busca (1);

34. **sistema de classificação (1)** - sistemas de classificação (3), esquemas de classificação bibliográfica (1), esquemas de classificação (1);
35. **tabela de classificação (1)** - tabelas de classificação (1);
36. **termo geral (1)** ;
37. **termos associados (1)**;
38. **termos específicos (1)**;
39. **tesauro (1)**, tesouros (6), thesaurus (2);
40. **tesauro eletrônico (1)**;
41. **tesouros facetados (1)**;
42. **tratamento da informação (4)** - tratamento de informações (1), tratamento informacional (1), processamento da informação (1);
43. **vocabulário controlado (2)** - vocabulários controlados (3), linguagem controlada (1), linguagens controladas (1), linguagem de indexação (1), linguagens de indexação (2), linguagem documentária (2), linguagens documentárias (1).

Dos dados consolidados nessa listagem emergiram as seguintes categorias de análise: *flexão de número (singular e plural)*; *variação concorrente lingüística (sintática, lexical e gráfica)*; *variação coocorrente e ausência de variação*. Foram coletados ainda termos de outras áreas que tiveram significativa ocorrência nos artigos analisados, esses termos reunidos formaram uma última categoria: *termos pertencentes a outras áreas empregados em Representação Temática*. Cada uma dessas categorias de análise será tratada separadamente a seguir.

4.1 Flexão de Número

As flexões de número referem-se ao uso do singular e plural nos artigos analisados. Tais flexões de número fazem parte da estrutura discursiva nos textos, não implicando em alteração de sentido.

Os termos foram coletados e registrados nas fichas terminológicas preferencialmente no singular. Desse modo, se determinado artigo apresentava o termo ora no singular, ora no plural, o termo era coletado na sua forma singular. Na situação em que o termo foi empregado unicamente no plural por determinado autor, esse foi considerado *termo*, não se fazendo a flexão para o singular.

A lista abaixo apresenta tais flexões de número, permitindo observar a ocorrência de termos no singular e no plural nos artigos analisados.

- a) busca de informação (1) / busca de informações (1);
- b) cabeçalho de assunto (1) / cabeçalhos de assuntos (1);
- c) conceito (2) / conceitos (2);
- d) descritor (2) / descritores (4);
- e) estratégia de busca (1) / estratégias de busca (1);
- f) linguagem controlada (1) / linguagens controladas (1);
- g) linguagem de indexação (1) / linguagens de indexação (2)
- h) linguagem documentária (2) / linguagens documentárias (1);
- i) vocabulário controlado (2) / vocabulários controlados (3);
- j) política de indexação (1) / políticas de indexação (1);
- k) recuperação da informação (7) / recuperação das informações (1);
- l) recuperação de informação (2) / recuperação de informações (1);

- m) sistema de recuperação da informação (3) / sistemas de recuperação da informação (1);
- n) sistema de recuperação de informação (1) / sistemas de recuperação de informação (2);
- o) sistema de classificação (1) / sistemas de classificação (3)
- p) tabela de classificação (1) / tabelas de classificação (1);
- q) termo (2) / termos (1);
- r) tesouro (1) / tesouros (6).

A ocorrência de termos no plural nos artigos refere-se a concordância, uma vez que o autor pretende expressar mais de um elemento. Por exemplo, quando o autor empregou o termo *descriptor* entende-se que se trata de uma única unidade ou de uma referência genérica. Já quando foi utilizado o termo *descriptores* temos que o autor estava expressando a idéia de mais de um, ou de um conjunto de descritores. Nos dois casos – emprego de *descriptor* ou *descriptores* – o conceito ao qual o autor se refere é o mesmo, não havendo portanto alteração de sentido devido ao uso de singular ou plural.

4.2 Variação Concorrente Lingüística

A categoria de variações concorrentes lingüísticas trata do fenômeno propriamente lingüístico existente nas variações. Os seguintes tipos de variantes lingüísticas foram coletados durante a pesquisa: variantes terminológicas sintáticas,

variantes terminológicas lexicais e variantes terminológicas gráficas. Cada uma dessas categorias será tratada a seguir com os respectivos termos coletados.

4.2.1 Variantes Terminológicas Sintáticas

Como já vimos anteriormente, as variações terminológicas sintáticas caracterizam-se pelo emprego alternado de duas construções sintagmáticas funcionando como predicação de uma unidade terminológica (FAULTISCH, 2002, p.81). As seguintes variantes terminológicas sintáticas foram coletadas dos artigos analisados:

- a) busca da informação (2) / busca de informação (1);
- b) disseminação da informação (1) / disseminação de informação (1);
- c) indexação automática (2) / indexação automatizada (1)
- d) recuperação da informação (7) / recuperação de informação (2);
- e) representação da informação (3) / representação de informações (1);
- f) sistema de recuperação da informação (3) / sistema de recuperação de informação (3);
- g) tratamento da informação (4) / tratamento de informações (1) / tratamento informacional (1).

A maior parte dos termos e variações acima listados apresenta diferenças na preposição utilizada para formação do sintagma. Este é o caso dos sintagmas *busca da informação* e *busca de informação*, *disseminação da informação* e *disseminação*

de informação, recuperação da informação e recuperação de informação, representação da informação e representação de informações, sistema de recuperação da informação e sistema de recuperação de informação.

No sintagma *busca da informação*, por exemplo, foi utilizado o substantivo, a preposição, um artigo definido e outro substantivo, desta forma:

SUBSTANTIVO	+ PREPOSIÇÃO	+ ARTIGO DEFINIDO	+ SUBSTANTIVO
(busca)	(de)	(a)	(informação)

Assim, o sintagma expressa uma idéia de algo mais definido pelo uso do artigo. Já a expressão *busca de informação* sinaliza para algo mais indefinido ou geral.

Os sintagmas *tratamento da informação, tratamento de informações e tratamento informacional* apresentam diferença de preposição na formação dos sintagmas (*tratamento da informação e tratamento de informações*) e ainda o emprego da forma adjetivada para o substantivo *informação (tratamento informacional)*.

A formação sintagmática de *indexação automática e indexação automatizada* é constituída de substantivo (*indexação*) e adjetivo (*automática e automatizada*). Em *indexação automatizada* temos o particípio do verbo *automatizar* adjetivando o substantivo *indexação*.

As diferenças na constituição dos sintagmas apontadas acima, embora não interfiram na compreensão dos conceitos por esses representados, precisam ser percebidas no momento da indexação dos documentos. A possibilidade de que os documentos sejam indexados, em uma unidade de informação, ora por *tratamento da informação* ora por *tratamento informacional*, por exemplo, poderia interferir na posterior recuperação dessas informações.

Isso justifica a importância do uso de vocabulários controlados, que padronizem a linguagem utilizada na indexação, e o estabelecimento de uma expressiva rede de remissivas que permitam a recuperação da informação.

4.2.2 Variantes Terminológicas Lexicais

As variantes terminológicas lexicais, como já foi mencionado no Referencial Teórico (seção 3.2.4), apresentam um item da estrutura lexical apagado sem que o conceito seja alterado. Esses apagamentos, quando inseridos no discurso especializado, não acarretam na perda de sentido dos sintagmas.

A listagem a seguir apresenta os termos, na forma em que eles foram coletados dos artigos, acrescidos do elemento apagado entre colchetes. Esta complementação aos termos coletados que aparece entre colchetes foi inferida pela autora desta pesquisa através da leitura dos textos analisados. Não foi possível definir a preposição que cada autor teria usado em cada artigo, por isso foi apenas indicada a necessidade do uso da mesma, como vemos abaixo:

- a) classificação **[preposição + assuntos]** (5);
- b) esquemas de representação **[preposição + informação]** (1);
- c) estratégia(s) de busca **[preposição + informação]** (2);
- d) sistema de recuperação **[preposição + informação]** (1);
- e) sistemas de busca **[preposição + informação]** (1);
- f) sistema(s) de classificação **[preposição + assuntos]** (4);
- g) esquemas de classificação **[preposição + assuntos]** (1);

h) tabela (s) de classificação **[preposição + assuntos]** (2).

No grupo de termos coletados acima, uma parte do sintagma terminológico fica implícita. Esses itens que foram apagados dos termos estavam subentendidos no contexto dos artigos, portanto sua ausência não compromete o entendimento do leitor.

Porém, a omissão de parte do sintagma em uma situação de tratamento de informação pode prejudicar a recuperação dos documentos indexados. O termo expresso pelo autor que apresenta omissão de parte do sintagma, ao ser extraído do discurso pode perder seu significado. Os sintagmas que sofreram apagamento de uma unidade lexical perderiam o sentido fora do contexto: Representação Temática.

Por isso cabe ao bibliotecário, ao elaborar instrumentos de recuperação da informação tais como tesouros, cabeçalhos de assuntos ou mesmo para subsidiar o processo de indexação automática, tornar explícita aquela unidade lexical que estava oculta nos termos empregados pelos autores.

4.2.3 Variação Terminológica Gráfica

Foi ainda coletada uma variação terminológica gráfica. Em seis artigos foi encontrado o termo *tesauro* na sua forma plural *tesouros*, em dois artigos o mesmo termo foi grafado pela sua forma latina *thesaurus*. Deve ser aqui observado que o termo *thesaurus*, do latim, foi empregado em dois artigos diferentes, porém da mesma autora. Inicialmente, na literatura da área, o termo era empregado em latim,

com o decorrer do tempo esse passou a ser empregado no seu equivalente em português: *tesauro*.

4.3 Variantes Coocorrentes

Retomando a definição de variante coocorrente, já explicitada no capítulo Referencial Teórico, temos que tais variações formalizam a sinonímia no discurso especializado. Faulstich (2002, p.83) explica que “[...] entendemos por formas em coocorrência as que têm presença simultânea em textos que tratam de assunto da mesma natureza”.

As variantes coocorrentes coletadas em *corpus* textual estão listadas abaixo:

- a) descritor (2), descritores (4) - termo (2), termos (1) - termo de indexação (1);
- b) vocabulário controlado (2), vocabulários controlados (3) - linguagem controlada (1), linguagens controladas (1) - linguagem de indexação (1), linguagens de indexação (2) - linguagem documentária (2), linguagens documentárias (1);
- c) organização da informação (2) - organização do conhecimento (1);
- d) indexação (5) - análise documentária (1) - representação da informação (3) - representação da informação documentária (1) - representação do conteúdo dos documentos (1) - representação do conteúdo temático (1) - representação do conhecimento (1);
- e) sistema de recuperação da informação (3) - sistemas de busca [preposição + informação] (1);

- f) sistema de classificação (1) - esquemas de classificação bibliográfica (1) - esquemas de classificação (1) ;
- g) tratamento da informação (4) - processamento da informação (1).

Cada um desses casos será comentado em tópicos separados a seguir.

4.3.1 Descritor(es), Termo(s) e Termos de Indexação

De acordo com Gomes (1990, p.15) os descritores são “[...] termos escolhidos para nomear um conceito”. A autora relaciona ao longo do seu texto conceitos de elaboração de tesouros com a prática da pesquisa terminológica. A aproximação entre *termo* e *descriptor* é discutida também por Van der Laan (2002). A autora explica que os descritores “[...] poderiam adquirir maior representatividade se tivessem um tratamento de unidades lexicais terminológicas”. (VAN DER LAAN, 2002, p.62).

Essa relação existente entre os termos *descriptor*, *termo* e *termos de indexação* na literatura, pode ser comprovada nos artigos que compõem o *corpus* textual desta pesquisa. Do campo *definição* de uma das fichas terminológicas foi retirada a seguinte afirmação: “Os descritores são termos de um tesouro que devem ser empregados para representar uma noção (ou conceito) contida(o) num texto ou numa equação de busca”. Neste artigo temos, portanto, uma definição de *descritores* como unidades que representam conceitos, reforçando assim a idéia de aproximação entre *termo* e *descriptor*.

Em outro artigo foi empregado o termo *termos de indexação*. A análise do contexto permite inferir que o autor refere-se ao mesmo conceito representado pelos

termos *descriptor* e *termo*. O termo *descritores* foi utilizado por um terceiro autor que ao longo do texto emprega também o termo *palavras-chave* com o mesmo sentido, como podemos observar no campo *variante* da ficha:

TERMO: DESCRITORES
Fonte: Ciência da Informação , Brasília, v.29, n.3, p.50-70, 2000. Disponível em: < http://www.ibict.br >. Acesso em: 25 set. 2004.
Definição:
Contexto: “Esta problemática de uso da lingüística na formação de descritores abrange várias áreas do conhecimento humano, especificamente dentro da área informacional, tais como a terminologia, a tradutologia e, mais recentemente, a socioterminologia”.
Sigla:
Sinônimo:
Variante: palavras-chave. “O uso de palavras-chave é fundamental para esta recuperação, e a lingüística entra como processo natural na recuperação dessa informação, resolvendo problemas de representação de termos no sistema utilizado pelos usuários”.
Área: Organização e Tratamento da Informação
Subárea: Representação Temática
Observações:
Compiladora: Verônica Ezequiel
Data: 20.09.2004

Logo, é possível inferir que os autores dos artigos analisados empregam os termos *descritores*, *termos de indexação* e *palavras-chave* para a representação de um mesmo conceito.

4.3.2 Linguagem(ns) Controlada(s), Linguagem de Indexação...

Os termos *linguagem controlada*, *linguagens controladas*, *linguagem de indexação*, *linguagens de indexação*, *linguagens de recuperação*, *linguagem documentária*, *linguagens documentárias*, *vocabulário controlado* e *vocabulários controlados* foram empregados com o mesmo sentido nos textos. Todos estes

termos referem-se à linguagem padronizada que tem por objetivo representar a informação para sua posterior recuperação. De acordo com Lara (1993, p.72) temos que:

A representação (a construção de índices) é realizada através do uso de um código comutador, ou seja, uma Linguagem Documentária (LD) que tem como função a normalização das unidades significantes ou conceituais presentes no texto original, a partir de elementos que constituem, de alguma forma, uma condensação de áreas de assunto. A condensação nesse caso, é expressa pelos elementos do código de comutação, sendo portanto, exterior ao texto submetido à conversão.

Com isto a autora explica que as linguagens documentárias são instrumentos, externos ao documento analisado, elaborados para padronização dos descritores que representarão o conteúdo do mesmo.

A seguir serão apresentados alguns exemplos que mostram o uso desses termos nos artigos que constituíram o *corpus* textual dessa pesquisa.

O campo *contexto* de uma das fichas de coleta apresenta o seguinte registro: “[...] a Ciência da Informação, quando constrói linguagens documentárias tem como objetivo assegurar a organização e a transferência da informação”.

Com isso temos que as linguagens documentárias têm a função de representar os conteúdos dos documentos com o objetivo de organizar e transferir a informação.

Podemos observar na ficha de coleta de dados a seguir que o autor define o termo *linguagem controlada* e aponta o termo *vocabulário controlado* como seu sinônimo como vemos no campo *definição* da ficha. Foram empregados ainda os termos *linguagens documentárias* e *linguagem de indexação controlada* com o mesmo sentido de *linguagem controlada*, como registra o campo *variante* da ficha:

TERMO: LINGUAGEM CONTROLADA
Fonte: Ciência da Informação , Brasília, v.31, n.1, p.41-52, 2002. Disponível em: < http://www.ibict.br >. Acesso em: 25 set. 2004.
Definição: “[...] a linguagem controlada caracteriza-se como a que é utilizada apenas nos campos de descritor, termos de indexação e identificadores”.
“(...) Esta, denominada também vocabulário controlado, pode ser definida como um conjunto limitado de termos autorizados para uso na indexação e busca de documentos”.
“ Linguagem controlada ou vocabulário controlado pode ser definido como um conjunto de termos organizados de forma hierarquizada e/ou alfabética, com o objetivo de possibilitar a recuperação de informações temáticas, reduzindo substancialmente a diversidade de terminologia. São também conhecidos como linguagens documentárias ou linguagens controladas”.
Contexto:
Sigla: LC.
Sinônimo:
Variante: vocabulário controlado; linguagens documentárias; linguagem de indexação controlada.
Área: Organização e Tratamento da Informação
Subárea: Representação Temática
Observações:
Compilador: Verônica Ezequiel
Data: 07.09.2004

O emprego do termo *linguagens de indexação* também foi registrado durante a coleta de termos. Em um dos textos o autor cita as listas de cabeçalhos de assuntos, os sistemas de classificação e os *thesauri* como exemplos de linguagens de indexação. Ao longo do texto foi empregado ainda o termo *linguagens documentárias* com o mesmo sentido do termo *linguagens de indexação*.

O termo *vocabulário controlado* foi coletado de dois artigos pertencentes ao *corpus* textual da pesquisa. Em um desses artigos o autor aponta as funções deste instrumento. Novamente encontramos a padronização de termos para posterior recuperação da informação como função desses instrumentos.

Dessa forma podemos perceber que para os autores dos artigos analisados nessa pesquisa, os termos *linguagem documentária*, *linguagem controlada*, *linguagem de indexação* e *vocabulário controlado* representam o mesmo conceito.

4.3.3 Organização da Informação e Organização do Conhecimento

No segundo grupo de variantes terminológicas os termos *organização da informação* e *organização do conhecimento* são empregados nos textos com o mesmo sentido. De um dos artigos foi coletado o termo *organização da informação*. Durante o discurso o autor empregou o termo *organização do conhecimento* como uma forma variante. Como vemos a seguir:

TERMO: ORGANIZAÇÃO DA INFORMAÇÃO
Fonte: Data Grama Zero – Revista de Ciência da Informação , Rio de Janeiro, v.2, n.5, 2001. Disponível em: < http://www.dgz.org.br >. Acesso em: 25 set. 2004.
Definição:
Contexto: “A eficácia dos processos de representação da informação reside na medida em que os elementos contidos no estoque informacional encontram-se ‘perceptíveis’ pelo usuário através do sistema, de forma que o modo de organização da informação contida no estoque apresente condições de acessibilidade, otimizada em função dos meios utilizados para sua representação”.
Sigla:
Sinônimo:
Variante: organização do conhecimento.
Área: Organização e Tratamento da Informação
Subárea: Representação Temática
Observações:
Compiladora: Verônica Ezequiel
Data: 21.10.2004

Em um segundo artigo foi encontrado o termo *organização do conhecimento*, como registra a ficha a seguir:

TERMO: ORGANIZAÇÃO DO CONHECIMENTO
Fonte: Data Grama Zero – Revista de Ciência da Informação , Rio de Janeiro, v.2, n.2, 2000. Disponível em: < http://www.dgz.org.br >. Acesso em: 25 set. 2004.
Definição:
Contexto: “Nos sistemas tradicionais, ao longo do tempo foram desenvolvidos instrumentos de organização do conhecimento com a finalidade de arranjo físico e arquivamento e, métodos e técnicas de representação da informação com a finalidade de descrever o conteúdo intelectual dos documentos para fim de recuperação por assunto”.
“A organização do conhecimento por si só já é reconhecida como um campo autônomo de conhecimento. Como bem aponta DAHLBERG, a necessidade de organizar o conhecimento sempre foi reconhecida. Durante muito tempo este interesse foi terreno quase que exclusivo de Bibliotecários e Filósofos. Mais tarde passou a ser também área de interesse dos Autores de Enciclopédias e Educadores, seguidos pelos Documentalistas, Cientistas da Informação e Linguistas”.
“A evolução histórica dos instrumentos de organização do conhecimento teve momentos marcantes, caracterizados por diferentes metodologias de construção e diferentes técnicas e métodos de indexação por assunto”.
Sigla:
Sinônimo:
Variante:
Área: Organização e Tratamento da Informação
Subárea: Representação Temática
Observações:
Compiladora: Verônica Ezequiel
Data: 09.10.2004

Embora os termos *organização da informação* e *organização do conhecimento* tenham sido empregados com o mesmo sentido nos textos, deve-se observar que *informação* e *conhecimento* são termos referentes a conceitos distintos. A este respeito Le Coadic (1996, p.9-10) explica que:

Nosso estado (ou nossos estados) de conhecimento sobre determinado assunto, em determinado momento, é representado por uma estrutura de conceitos ligados por suas relações: ‘nossa imagem’ do mundo. Quando constatamos uma deficiência ou uma anomalia desse(s) estado(s) de conhecimento, encontramos-nos em estado anômalo de conhecimento. Tentamos obter uma informação ou informações que corrigirão essa anomalia. Disso resulta um novo estado de conhecimento.

Com isto temos que a informação não é um fim, mas sim um meio. Um conjunto de informações, a interpretação das mesmas e a relação destas informações com as demais já elaboradas pelo indivíduo poderão levá-lo a um novo estágio de conhecimento.

O conhecimento tem natureza mais ampla que a informação. A este respeito, Cintra (1994, p.14) afirma que:

- enquanto o conhecimento é estruturado, coerente e freqüentemente universal, a informação é atomizada, fragmentada, particular;
- enquanto o conhecimento é de duração significativa, a informação é temporária, transitória, talvez mesmo efêmera;
- enquanto o conhecimento é um estoque, a informação é um fluxo de mensagens.

Assim, os termos *organização da informação* e *organização do conhecimento* não deveriam ser empregados para a representação do mesmo conceito.

Em uma situação de escolha de descritores durante o processo de indexação desses artigos, o bibliotecário deveria optar pelo termo *organização da informação* em detrimento do termo *organização do conhecimento*. Porém, seria razoável que houvesse uma remissiva para o termo *organização do conhecimento*. Essa atitude evitaria perda de informação ou silêncio no momento da recuperação da informação.

4.3.4 Indexação, Análise Documentária...

Foram coletados ainda os termos *indexação*, *análise documentária*, *representação da informação*, *representação da informação documentária*, *representação do conteúdo dos documentos*, *representação do conteúdo temático* e *representação do conhecimento* que remetiam para um mesmo conceito.

O termo *indexação* já está consagrado na área de Biblioteconomia, portanto, não é possível afirmar que tenha acontecido apagamento de um item lexical, como ocorre com as já citadas variantes terminológicas lexicais.

Embora não tenha ocorrido nenhum apagamento neste caso, é importante mencionar que o termo *indexação* transita em diferentes áreas do conhecimento. Portanto, ao ser extraído do seu contexto de origem, torna-se fundamental o uso de um qualificador que indique o contexto do qual o termo foi extraído. Desse modo, no processo de extração do termo *indexação* de um documento e de atribuição de status de descritor ao mesmo, é essencial o uso de um qualificador.

Gomes (1990, p.19) explica que “[...] nos termos homonímicos é imprescindível o uso de um contextualizador na medida em que ele precisará o significado do termo em questão. Exemplo: Indexação (Economia); Indexação (Documentação)”.

Quanto ao significado do termo *indexação* cabe fazer aqui algumas considerações. A expressão *indexação*, em Biblioteconomia, possui duas acepções, a primeira mais ampla, ou seja, entende que o ato de indexar consiste na determinação de todos os pontos de acesso de um documento para posterior recuperação, sejam estes: autor, título e assunto. A segunda é mais restrita e entende por *indexação* “[...] apenas o processo de determinação de assuntos de um documento e sua tradução para uma linguagem de *indexação*”. (VAN DER LAAN, 2002, p.11-12).

A leitura e interpretação dos artigos que compuseram o *corpus* textual desta pesquisa permitem inferir que os autores empregaram o termo *indexação* com o sentido da segunda acepção apresentada acima, ou seja, *indexação* como determinação de assunto(s) e tradução deste(s) assunto(s) para uma linguagem documentária. Nos exemplos que seguem temos o uso do termo *indexação* em dois artigos analisados:

TERMO: INDEXAÇÃO
Fonte: Data Grama Zero – Revista de Ciência da Informação , Rio de Janeiro, v.3, n.1, 2002. Disponível em: < http://www.dgz.org.br >. Acesso em: 25 set. 2004.
Definição:
Contexto: “Segundo Michel Le Guern (1991) ‘a passagem da indexação elaborada por especialistas, indexadores, para a indexação automática não modifica a natureza dos descritores, mas ele obriga a não mais se contentar com uma abordagem intuitiva e empírica.’ [6]. A reflexão apresentada mostra que a indexação automática, mais do que simplesmente viabilizar economicamente a indexação dos documentos de uma base, não poderia ter trocado os descritores por uma outra unidade desprovida de informação. Além disso, a extração de informações de um conjunto de documentos deveria ser sistemática e não intuitiva e empírica como é realizado o processo de indexação por parte dos técnicos de informação especializados nesse processo”.
“Diante do exposto, entendo que o processo de indexação deveria extrair dos documentos informações que possam facilitar a sua recuperação e não símbolos sem referência como o são as palavras”
Sigla:
Sinônimo:
Variante: processo de indexação.
Área: Organização e Tratamento da Informação
Subárea: Representação Temática
Compiladora: Verônica Ezequiel
Data: 16.10.2004

Desse modo, o termo *indexação* foi empregado pelos autores dos artigos com o mesmo sentido de *análise documentária*, *representação da informação*, *representação da informação documentária*, *representação do conteúdo dos documentos*, *representação do conteúdo temático* e *representação do conhecimento*.

O termo *análise documentária* é definido na literatura da área como “um tratamento documentário de conteúdo com a finalidade de elaborar representações condensadas do que está nos textos”. (FUJITA, 1998, p.21). A análise documentária trata-se, portanto, do estudo e identificação dos assuntos tratados pelos documentos.

O uso concomitante dos termos *análise documentária* e *representação do conteúdo* no discurso especializado pode ter originado o termo *análise de conteúdo*.

Na ficha abaixo o autor empregou o termo *análise documentária*. Em

determinado momento foi empregado o termo *análise de conteúdo* com o mesmo sentido de *análise documentária*, como vemos no campo *variante* da ficha:

TERMO: ANÁLISE DOCUMENTÁRIA
Fonte: Ciência da Informação , Brasília, v.29, n.3, p.50-70, 2000. Disponível em: < http://www.ibict.br >. Acesso em: 25 set. 2004.
Definição:
Contexto: “A ciência da informação, que investiga as propriedades e comportamento da informação, as forças governantes dos fluxos e os meios de processar a informação, tendo como objetivo a sua organização, armazenamento, recuperação e disseminação, tem estreita ligação com a lingüística pela intermediação da análise documentária , que se utiliza de métodos e processos para descrever o conteúdo dos documentos”.
Sigla:
Sinônimo:
Variante: análise de conteúdo. “(...) o acesso à informação, a análise de conteúdo , a estruturação de conceitos, a representação e a geração de novos conhecimentos implicam uma série de linguagens que são dependentes entre si, pelas suas partes (...)”.
Área: Organização e Tratamento da Informação
Subárea: Representação Temática
Observações:
Compiladora: Verônica Ezequiel
Data: 20.09.2004

O termo *análise de conteúdo* empregado em Biblioteconomia está relacionado à idéia de *conteúdo dos documentos* ou *tema tratado nos documentos*. Esse termo é próprio da área de Metodologia Científica o que pode gerar confusão, já que o mesmo termo possui significados distintos nessas duas áreas. Em Biblioteconomia o termo adquiriu o mesmo sentido de análise documentária.

O termo *representação do conteúdo temático* apresenta uma redundância já que, neste contexto, *conteúdo* e *tema* referem-se ao mesmo conceito. Desse modo, o termo *representação do conteúdo dos documentos*, também empregado em um dos artigos, é mais apropriado do que *representação do conteúdo temático*.

Aqui novamente aparecem os termos *informação* e *conhecimento* como sendo equivalentes já que os termos *representação da informação* e *representação*

do *conhecimento* foram empregados com o mesmo sentido. Como já vimos, os conceitos *informação* e *conhecimento* não se equivalem.

4.3.5 Sistema de Recuperação da Informação e Sistemas de Busca

Foram coletados os termos: *sistema de recuperação da informação* e *sistemas de busca [preposição + informação]* remetendo para o mesmo conceito. Nos textos analisados estes termos referem-se aos sistemas automatizados de armazenamento e recuperação da informação, como as bases de dados. No campo *contexto* da ficha que segue, o autor indica o termo *banco de dados* como sinônimo de *sistemas de recuperação de informação*:

TERMO: SISTEMAS DE RECUPERAÇÃO DE INFORMAÇÃO
Fonte: Ciência da Informação , Brasília, v.31, n.2, p.60-71, 2002. Disponível em: < http://www.ibict.br >. Acesso em: 25 set. 2004.
Definição:
Contexto: “O acesso aos grandes sistemas de recuperação de informação , também denominados de bancos de dados e, conseqüentemente, às suas bases de dados veio ampliar significativamente a qualidade das buscas bibliográficas, visto que essas bases proporcionam diversificados pontos de acesso à informação”.
Sigla:
Sinônimo:
Variante: sistemas de recuperação da informação; sistema de recuperação.
Área: Organização e Tratamento da Informação.
Subárea: Representação Temática.
Observações:
Compiladora: Verônica Ezequiel
Data: 23.09.2004

Com o surgimento de novas tecnologias, que oferecem suporte à organização e tratamento da informação, parece haver uma tendência à associação do conceito

sistemas de recuperação da informação (SRI's) com sistemas automatizados de recuperação da informação.

Na literatura da área temos que a definição de sistemas de recuperação da informação é mais ampla. Van der Laan (2002) explica que “[...] o principal objetivo de um sistema de recuperação da informação (SRI), como próprio nome sugere, é recuperar as informações de modo a atender às necessidades do usuário”.

Portanto, um sistema de recuperação da informação pode ser manual e ainda assim cumprir com a sua função. Porém, o emprego do termo *sistema de recuperação da informação* relacionado a sistemas automatizados é compreensível, já que há uma tendência muito forte à automatização de acervos e disponibilização de catálogos *on-line*.

4.3.6 Sistema de Classificação, Sistemas de Classificação...

Os termos *sistema de classificação*, *sistemas de classificação*, *esquemas de classificação bibliográfica* e *esquemas de classificação* são também empregados com mesmo sentido nos textos. Em geral estes termos se referem, nos textos, às tabelas de classificação: Classificação Decimal de Dewey (CDD) e Classificação Decimal Universal (CDU). Como mostra a ficha terminológica adiante:

TERMO: ESQUEMAS DE CLASSIFICAÇÃO BIBLIOGRÁFICA
Fonte: Data Grama Zero – Revista de Ciência da Informação , Rio de Janeiro, v.2, n.2, 2000. Disponível em: < http://www.dgz.org.br >. Acesso em: 25 set. 2004.
Definição:
Contexto: “Os esquemas de classificação bibliográfica começaram sendo gerais, ou seja, abrangendo todas as áreas do conhecimento. Os primeiros foram enumerativos (ex. Classificação Decimal de Dewey - CDD) , os que se seguiram foram semi-enumerativos ou semi-facetados (ex. Classificação Decimal Universal - CDU)”.
Sigla:
Sinônimo:
Variante:
Área: Organização e Tratamento da Informação
Subárea: Representação Temática
Observações:
Compiladora: Verônica Ezequiel
Data: 10.10.2004

De acordo com Cintra (1994) os sistemas de classificação e o tesauro são as mais conhecidas linguagens documentárias. Ainda de acordo com a autora, a CDD e a CDU são sistemas de classificação bibliográfica. Em um dos artigos foi empregado o termo *linguagem de indexação*, que como já vimos é empregado nos textos analisados com o mesmo sentido de *linguagem documentária*, e cita como exemplo de *linguagens de indexação* os *sistemas de classificação*, como vemos a seguir:

TERMO: SISTEMAS DE CLASSIFICAÇÃO
Fonte: Data Grama Zero – Revista de Ciência da Informação , Rio de Janeiro, v.2, n.5, 2001. Disponível em: < http://www.dgz.org.br >. Acesso em: 25 set. 2004.
Definição:
Contexto: “É o caso da criação/manutenção de linguagens e códigos, como as linguagens de indexação (listas de cabeçalhos de assuntos, sistemas de classificação , thesauri) e os códigos de catalogação”.
Sigla:
Sinônimo:
Variante:
Área: Organização e Tratamento da Informação
Subárea: Representação Temática
Observações:
Compiladora: Verônica Ezequiel
Data: 19.10.2004

É possível depreender dessas leituras que o termo *linguagens documentárias* é mais geral que *sistemas de classificação*, estabelecendo-se assim uma hierarquização entre esses dois termos. Há, portanto, consenso entre os autores (Cintra na literatura da área e demais autores cujos textos pertencem ao *corpus* textual da pesquisa) no que se refere à hierarquia entre esses termos.

4.3.7 Tratamento da Informação, Processamento da Informação...

Os termos *tratamento da informação* e *processamento da informação* são empregados nos textos com o mesmo sentido ao referirem-se ao processamento técnico da informação. Um dos artigos analisados apresenta uma definição bastante consistente para o termo *tratamento da informação*. Explica que o tratamento da informação refere-se à descrição física e temática dos documentos, como registrou o campo *definição* de uma das fichas de coleta:

TERMO: TRATAMENTO DA INFORMAÇÃO
Fonte: Data Grama Zero – Revista de Ciência da Informação , Rio de Janeiro, v.2, n.5, 2001. Disponível em: < http://www.dgz.org.br >. Acesso em: 25 set. 2004.
Definição: Nos sistemas de informação e de recuperação da informação, o tratamento da informação é definido como a função de descrever os documentos, tanto do ponto de vista físico (características físicas dos documentos) quanto do ponto de vista temático (ou de descrição do conteúdo).
Contexto:
Sigla:
Sinônimo:
Variante: tratamento.
Área: Organização e Tratamento da Informação
Subárea: Representação Temática
Observações:
Compiladora: Verônica Ezequiel
Data: 19.10.2004

Os termos *tratamento informacional* e *processamento da informação* empregados por outros dois autores também possuem o mesmo sentido da definição apresentada no exemplo acima.

4.4 Ausência de Variação Terminológica

Os termos listados abaixo foram coletados em mais de um artigo pesquisado sem que tenha ocorrido variação terminológica. As únicas variações registradas nestes termos foram as flexões de número, que como já vimos, não interferem no sentido dos termos.

1. Armazenamento da Informação (2)
2. Cabeçalho de Assunto (1) / Cabeçalhos de Assuntos(1)
3. Classificação Decimal de Dewey (2) / CDD (1)
4. Classificação Decimal Universal (2) / CDU (1)
5. Conceito (2) / Conceitos (1)
6. Estratégia de Busca (1) / Estratégias de Busca (1)
7. Linguagem Natural (6)
8. Listas de Cabeçalhos de Assunto (3)
9. Política de Indexação (1) / Políticas de Indexação (1)
10. Precisão (5)
11. Revocação (4)
12. Tabela de Classificação (1) / Tabelas de Classificação (1)

A ausência de variações terminológicas no emprego dos termos acima listados, em *corpus* textual, sinaliza que esses podem já estar consolidados na área. Cada um desses termos será comentado a seguir.

O termo *armazenamento da informação* foi empregado nos artigos analisados com o mesmo sentido de acomodação e guarda da informação para posterior recuperação.

De acordo com Castilho (1994, p.112) armazenamento da informação é uma “operação que consiste em guardar documentos nos seus devidos lugares, em áreas que lhes são próprias”. Com isso temos que há coincidência entre a idéia a que os autores se referem com a utilização desse termo nos artigos e a definição de Castilho.

O termo *cabeçalho de assunto* é definido pelo *Glossário Ala de Bibliotecologia y Ciencias de la Información* (1988, p.335, tradução nossa) como “ponto de acesso a um registro bibliográfico contido por uma palavra ou frase designando o tema do estudo ou estudos contidos em uma obra”.

Uma lista de cabeçalhos de assuntos determina, portanto, o conjunto de termos que serão utilizados no momento da indexação. Os termos *cabeçalho de assunto* e *listas de cabeçalhos de assunto* aparecem nos artigos analisados com o mesmo sentido das definições apresentadas acima.

Os termos *Classificação Decimal Universal (CDU)* e *Classificação Decimal de Dewey (CDD)* são empregados pelos autores nos artigos analisados. Essas tabelas, como já vimos, são sistemas de classificação.

Em um dos artigos os termos *Classificação Decimal Universal* e *Classificação Decimal de Dewey* foram indicados unicamente por suas siglas: *CDU* e *CDD*. Esses termos estão consagrados na área de Biblioteconomia. O emprego das siglas

desses termos na literatura da área é comum e não prejudica o entendimento dos textos.

As siglas ou “[...] signos verbais reduzidos” (KRIEGER, 2004, p.82) são definidos por Cabré (1993, p.178) como: “[...] unidades formadas pela combinação das iniciais de várias palavras que constituem uma expressão maior”.

Entretanto, seria adequado que essas expressões fossem empregadas, pelo menos na sua primeira ocorrência nos textos, por extenso. A *Associação Brasileira de Normas Técnicas (ABNT)* indica como obrigatório o registro da expressão por extenso antes do uso da sigla entre parênteses, quando esta aparece no texto pela primeira vez. (ASSOCIAÇÃO BRASILEIRA..., 2002).

Outro termo coletado que não apresenta variação terminológica no *corpus* textual pesquisado é *conceito*. Como já foi mencionado, *conceito* é uma elaboração mental, uma abstração que representa a realidade. No exemplo que segue, o autor emprega esse termo e faz uma citação onde é expressa a relação entre o termo *conceito* da área de Terminologia com termo *conceito* trabalhado em Representação Temática, como ilustra o campo *contexto* retirado de uma das fichas:

TERMO: CONCEITO
Fonte: Ciência da Informação , Brasília, v.31, n.1, p.41-52, 2002. Disponível em: < http://www.ibict.br >. Acesso em: 25 set. 2004.
Definição:
Contexto: “ ‘ Conceitos devem ser representados consistentemente para os propósitos de recuperação, por substantivos ou frases substantivadas; os indexadores devem trabalhar com um vocabulário de termos preferidos, designando-se um dos sinônimos de um determinado conceito como o termo mais adequado para uso; a opção pelo singular ou plural dos conceitos e suas exceções devem ser registradas claramente nesse vocabulário, visando à consistência da indexação, sua fidedignidade e posterior uso na recuperação (Austin, 1986, p.8)”.
Sigla:
Sinônimo:
Variante:
Área: Organização e Tratamento da Informação
Subárea: Representação Temática
Observações:
Compiladora: Verônica Ezequiel
Data: 20.09.2004

O termo *estratégia(s) de busca* teve duas ocorrências durante a coleta dos termos. Grogan (1995, p.114) define *estratégia de busca* como “a tradução da questão do usuário para a terminologia aceitável pelo sistema”.

As definições encontradas nos artigos para esse termo coincidem com a definição de Grogan, como vemos a seguir:

TERMO: ESTRATÉGIAS DE BUSCA
Fonte: Ciência da Informação , Brasília, v.31, n.2, p.60-71, 2002. Disponível em: < http://www.ibict.br >. Acesso em: 25 set. 2004.
Definição: “No âmbito da recuperação da informação, a estratégia de busca pode ser definida como uma técnica ou conjunto de regras para tornar possível o encontro entre uma pergunta formulada e a informação armazenada em uma base de dados”.
“Bates (1987, 1988) conceituou a estratégia de busca como o ‘estudo da teoria, princípios e prática de planejar e executar táticas e estratégias de busca’”.
Contexto: “(...) planejamento e a execução de uma estratégia de busca é uma ‘arte de escolher onde, quando e com que investigar cuidadosamente’ a fonte de informação para alcançar os objetivos específicos do solicitante”.
“Esses sistemas [de informação] possibilitam o planejamento de estratégias de busca com maior nível de complexidade envolvendo vários conceitos na mesma estratégia”.
“Tanto as bases de dados referenciais, quanto as textuais vão exigir um planejamento acurado da estratégia de busca , visando a uma recuperação de informação de acordo com as necessidades dos usuários”.
“(…) o alcance da qualidade na informação recuperada requer o planejamento de estratégias de busca específicas para cada base de dados”.
Sigla:
Sinônimo:
Variante:
Área: Organização e Tratamento da Informação
Subárea: Representação Temática
Observações: também é empregado no plural pela autora (estratégias de busca).
Compiladora: Verônica Ezequiel
Data: 24.09.2004

O termo *linguagem natural* teve seis ocorrências nos artigos analisados. A linguagem natural é aquela empregada no discurso comum, ou seja, na fala e na escrita onde não há um controle de vocabulário. (LANCASTER, 1993).

No campo *definição* de uma das fichas de coleta temos uma citação de Lancaster onde o autor apresenta o termo *texto livre* como sinônimo de *linguagem natural*, como vemos a seguir:

TERMO: LINGUAGEM NATURAL
Fonte: Ciência da Informação , Brasília, v.31, n.1, p.41-52, 2002. Disponível em: < http://www.ibict.br >. Acesso em: 25 set. 2004.
Definição: "(. . .) a linguagem natural abrange os termos do título e do resumo dos documentos referenciados."
"A linguagem natural (LN) pode ser definida como a linguagem do discurso técnico-científico, e, no contexto da recuperação da informação, Lancaster (1993, p. 200) afirma que "a expressão normalmente se refere às palavras que ocorrem em textos impressos, considerando-se como seu sinônimo a expressão 'texto livre'.
Contexto:
Sigla: LN.
Sinônimo: Texto livre.
Variante:
Área: Organização e Tratamento da Informação
Subárea: Representação Temática
Observações:
Compilador: Verônica Ezequiel
Data: 07.09.2004

Política de indexação pode ser definida como um guia para tomada de decisões que leva em consideração diversos fatores, tais como as características e objetivos da organização, o tipo de serviço a ser oferecido, identificação dos usuários, recursos humanos, materiais e financeiros. Os elementos que influenciam diretamente na elaboração de políticas de indexação são: a cobertura de assuntos, seleção e aquisição de documentos, processo de indexação, estratégia de busca, forma de saída, tempo de resposta do sistema e avaliação do sistema. (CARNEIRO, 1985).

Pela análise dos dados coletados pode-se afirmar que os artigos que compõem o *corpus* textual dessa pesquisa empregam o termo *política(s) de indexação* com mesmo significado expresso na literatura.

Os termos *revocação* e *precisão* foram empregados nos textos apresentando uma relação inversa. Essa relação de inversão pode ser percebida na literatura da área. A *revocação* refere-se a todas as informações recuperadas pelo sistema, enquanto a *precisão* está relacionada ao grau de aproveitamento dessas

informações pelo usuário, ou seja, o grau de especificidade dessas informações. (CARNEIRO, 1985).

De acordo com Lancaster o ideal é manter o equilíbrio entre revocação e precisão, isto é, “obter o máximo de revocação, porém mantendo um nível aceitável de precisão”. (LANCASTER, 1993, p.76).

Nesse caso novamente os textos que compõem o *corpus* textual encontram respaldo no restante da literatura da área.

Os termos *tabela de classificação* e *tabelas de classificação* foram também empregados nos artigos pesquisados. Como já vimos, tais tabelas são sistemas de classificação e é neste sentido que os autores se referem a estes termos.

Durante o processo de coleta verificou-se que vários termos tiveram uma única ocorrência. Isso não implica afirmar que os mesmos encontram-se consolidados inexistindo variantes terminológicas.

A ocorrência única de cada um desses termos talvez se justifique devido ao reduzido número de artigos analisados. Esses termos encontram-se listados abaixo:

1. classificação de dois pontos (1);
2. especificidade (1);
3. esquemas de classificação bibliográfica analítico-sintéticos (1);
4. esquemas de classificação bibliográfica enumerativos (1);
5. esquemas de classificação bibliográfica semi-enumerativos (1);
6. indexação pós – coordenada (1);
7. indexação pré – coordenada (1);
8. instrumento de recuperação da informação (1);
9. não-descritores (1);
10. pontos de acesso à informação (1);

11. relevância (1);
12. ruído (1);
13. termo geral (1);
14. termos associados (1);
15. termos específicos (1);
16. tesouro eletrônico (1);
17. tesouros facetados (1).

4.5 Termos Pertencentes a Outras Áreas Empregados em Representação Temática

Durante o processo de coleta de termos verificou-se a ocorrência de terminologia não pertencente à área de Biblioteconomia. Essa ocorrência de termos de outras áreas explica-se devido ao caráter interdisciplinar da área. Esses termos encontram-se listados a seguir:

- a) ambigüidade (2);
- b) banco de dados (3);
- c) base de dados / bases de dados (4);
- d) conceitos compostos (1);
- e) conceitos simples (1);
- f) conceitos terminológicos (1);
- g) homonímia (1);
- h) lingüística (1);

- i) neologismos (1);
- j) polissemia (4);
- k) sinonímia (2);
- l) Terminologia (2);
- m) termos lingüísticos (1).

Nota-se que os termos empregados com maior freqüência são *bancos e base(s) de dados*, pertencentes à área de informática e *ambigüidade, polissemia, sinonímia, e Terminologia*, pertencentes à Lingüística e Terminologia.

Smit et al. (2004) também aborda em sua pesquisa o empréstimo de outras áreas do conhecimento. A autora exemplifica essa categoria com o termo *base de dados* também encontrado nessa pesquisa.

Os termos referentes às áreas de Lingüística e Terminologia reafirmam a importância da aproximação e estudo dessas disciplinas com a área de Representação Temática. Essa aproximação justifica-se posto que o signo empregado no processo de indexação alfabética é o signo lingüístico.

5 CONCLUSÕES

O trabalho de coleta e análise de termos em *corpus* textual permitiu a verificação de variações terminológicas na subárea de Representação Temática.

Foi constatado que as variações terminológicas sintáticas e lexicais identificadas nesta pesquisa não prejudicaram a compreensão do conteúdo dos textos. Apesar de não dificultar o entendimento das idéias expressas pelos autores, essas variações podem prejudicar a recuperação da informação.

Durante o processo de indexação essas variações devem ser identificadas e tratadas adequadamente. Por exemplo, a variação terminológica sintática existente entre os sintagmas *tratamento da informação* e *tratamento informacional* não interfere na compreensão do conceito para o qual esses se referem. Porém, se determinado documento for indexado sob o descritor *tratamento da informação* e outro sob *tratamento informacional* sem o uso de remissivas, então certamente haverá perda de informação no resultado de uma posterior busca de informação.

Em um processo de indexação automática essas variações podem ser muito prejudiciais, à medida em que serão indexados diferentes termos que representam os mesmos conceitos. Desse modo, o usuário poderá não receber como resposta todos os documentos que tratam do assunto de seu interesse existentes em determinada unidade de informação.

Quanto às variantes lexicais, temos que toda a unidade lexical que é apagada dos sintagmas, no discurso especializado, deve ser explicitada para extração de tais estruturas do seu contexto de origem.

Assim, o indexador ao deparar-se com o termo *recuperação* em um documento da subárea de Representação Temática, deverá identificar, através do contexto, qual é a unidade lexical que está oculta. Nesse caso o termo é *recuperação [preposição + informação]*. Identificado o sintagma completo, poderá ser estabelecido qual o descritor autorizado será utilizado na indexação mediante a utilização de instrumentos de controle de vocabulário.

Os casos de variação coocorrentes coletados nessa pesquisa parecem ser os maiores geradores de imprecisão na linguagem empregada na subárea de Representação Temática. Por exemplo, o termo *indexação*, predominantemente empregado nos artigos, apresentou como variantes outros sete termos - *indexação por assunto, análise documentária, representação da informação, representação da informação documentária, representação do conteúdo dos documentos, representação do conteúdo temático e representação do conhecimento*. Um número tão expressivo de variações terminológicas pode induzir o leitor a acreditar que existe distinção entre esses termos.

Quanto ao processo de indexação, nesses casos de existência de variação coocorrente, se repetem os problemas já mencionados. A percepção dessas variações pelo agente indexador, o uso de vocabulário controlado e remissivas podem minimizar o ruído e o silêncio no momento da recuperação da informação.

O caráter interdisciplinar dos termos empregados em Biblioteconomia ficou evidente na coleta dos termos. Esses empréstimos terminológicos demonstram a interface existente entre a Biblioteconomia e as demais áreas do conhecimento.

Inúmeros autores têm trabalhado com o estudo da Terminologia e Lingüística aplicada à Biblioteconomia: Kobashi, Smit, Lara, Tálamo entre outros. O que se

propõe aqui é que os estudos acerca da terminologia empregada na subárea de Representação Temática prossigam.

A elaboração de glossários de termos empregados em Representação Temática poderia elucidar o significado dos mesmos e permitir aos leitores um maior entendimento da terminologia empregada na subárea. O estudo da linguagem dos especialistas e a geração de produtos tais como glossários e dicionários especializados certamente poderão contribuir para a consolidação da terminologia empregada em Biblioteconomia e conseqüentemente na subárea de Representação Temática.

REFERÊNCIAS CONSULTADAS

ALVES, Ieda Maria. **A Pesquisa em Terminologia**: algumas considerações. Disponível em <http://www.letras.ufrj.br/abralin/boletim/boletim21_tema08.html>. Acesso em 02 de mar. 2004.

ARNTZ, Reiner; PICHT, Heribert. **Introducción a la Terminología**. Madrid: Fundación Germán Sánchez Rupérez, 1995. 384 p.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 6023: Informação e documentação: referências: elaboração**. Rio de Janeiro, 2002.

CABRÉ, Maria Teresa. **La Terminologia**: teoria, metodologia, aplicaciones. Barcelona: Antártica / Empúries, c1993. 528 p.

CARNEIRO, Marília Vidigal. Diretrizes para uma Política de Indexação. **Revista da Escola de Biblioteconomia da UFMG**, Belo Horizonte, v.14, n.2, p.221-241, set. 1985.

CASTILHO, Ataliba Teixeira de (org.). **A Sistematização de Arquivos Públicos**. Campinas, SP: Editora do Livro, 1994.

CINTRA, Anna Maria M. et al. **Para Entender as Linguagens Documentárias**. São Paulo: Polis: APB, 1994. 72 p.

COMISSÃO DE GRADUAÇÃO DE BIBLIOTECONOMIA (org.). **Curso de Biblioteconomia da UFRGS**: currículo 2000. Porto Alegre: Faculdade de Biblioteconomia e Comunicação da UFRGS, 2000. 57p.

DODEBEI, Vera Lúcia Doyle. **Tesauro**: linguagem de representação da memória documentária. Rio de Janeiro: Interciência, 2002. 119 p.

DUBUC, Robert. Término y Concepto. In:_____. **Manual de Terminologia**. Santiago (Chile): Ril Editores, 1999. P. 55-64.

FAULSTICH, Enilde. Variação em Terminologia. Aspectos de Socioterminologia. In: GUERRERO RAMOS, Glória; PÉREZ LAGOS, Manuel F. (coord.). **Panorama Actual de la Terminología**. Granada: Editorial Comares, 2002. P. 65-91.

FEDOR DE DIEGO, Alicia. Metodologia para la Elaboración de Terminologias. In:_____. **Terminologia**: teoria y practica. Caracas: Equinoccio, 1995. P.99-106.

FUJITA; NARDI; SANTOS. A Leitura em Análise Documentária. **Transinformação**, Campinas, v.10, n.3, p.20-31, set/dez. 1998.

GOMES, Hagar Espanha (coord.). **Manual para Elaboração de Tesouros Monolíngües**. Brasília: Programa Nacional de Bibliotecas das Instituições de Ensino Superior, 1990. 77p.

GROGAN, Denis. **A Prática do Serviço de Referência**. Brasília: Briquet de Lemos, 1995. 196 p.

KOBASHI, Nair Yumiko. Análise Documentária e Representação da Informação. **INFORMARE**: Cadernos do Programa de Pós-Graduação em Ciência da Informação, Rio de Janeiro, v.2, n.2, p.5-23, jul/dez. 1996.

KRIEGER, Maria da Graça; FINATTO, Maria José Bocorny. **Introdução à Terminologia**: teoria e prática. São Paulo: Contexto, 2004. 223 p.

KRIEGER, Maria da Graça; FINATTO, Maria José Bocorny (org.). **Temas de Terminologia**. Porto Alegre / São Paulo: ed. Universidade / UFRGS / Humanitas / USP, 2001. 454p.

LANCASTER, F. W. **Indexação e Resumo**. Brasília: Briquet de Lemos, 1993. 347p.

LARA, Marilda L. G. de. **Linguagens Documentárias, Instrumentos de Mediação e Comunicação**. 1993.

LE COADIC, Yves-François. **A Ciência da Informação**. Brasília, DF: Briquet de Lemos, 1996. 119p.

MARITAIN, Jacques. **Elementos de Filosofia II**: a ordem dos conceitos – lógica menor (lógica formal). Rio de Janeiro: Agir, 1989.

MEY, Eliane Serrão Alves. Pontos de Acesso. In: _____. **Introdução à Catalogação**. Brasília: Briquet de Lemos, 1995. p.56-77.

NOVELLINO, Maria Salet F. A Linguagem como Meio de Representação ou de Comunicação da Informação. **Perspectivas em Ciência da Informação**, Belo Horizonte, v.3, n.1, p.137-146, 1998.

SAGER, Juan C. La Dimensión Lingüística. In: _____. **Curso Prático Sobre el Procesamiento de la Terminología**. Madrid: Fundación Germán Sánchez Ruipérez, Pirámide, 1993. p.89-97.

SMIT, Johanna. O Profissional da Informação e Sua Relação com as Áreas de Biblioteconomia / Documentação, Arquivologia e Museologia. In: VALENTIM, M. L. P. (org.). **O Profissional da Informação: formação, perfil e atuação profissional**. São Paulo: Polis, 2000. p. 119-134.

SMIT, Johanna W.; TÁLAMO, Maria de Fátima G. M.; KOBASHI, Nair Y. A Determinação do Campo Científico da Ciência da Informação: uma abordagem terminológica. **Data Grama Zero – Revista de Ciência da Informação**, Rio de Janeiro, v.5, n.1, fev. 2004. Disponível em: <<http://www.dgz.org.br>>. Acesso em: 02 de mar. 2004.

TERMINOLOGIA Relacionada com los Estudios de Usuarios y la Formación de Usuarios de la Información. In: SEMINÁRIO Latinoamericano Sobre Formación de Usuarios de la Información y los Estudios de Usuarios. México: UNAM, Centro Universitario de Investigaciones Bibliotecológicas, 1997. P.1-10.

VAN DER LAAN, Regina H et al. **Avaliação dos descritores relativos a Ciências da Informação, empregados no processo de indexação alfabética pela Biblioteca Setorial da Faculdade de Biblioteconomia e Comunicação / UFRGS**. Porto Alegre, 2002. 11f.

VAN DER LAAN, Regina H. **Terminologia Clássica**: alguns pontos Wusterianos. Porto Alegre, 2003. Polígrafo de aula.

_____. **Tesouro e Terminologia**: uma inter-relação lógica. 2002. 262 f. Tese (Doutorado em Letras) – Instituto de Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre.

YOUNG, Heartsill (Ed.). **Glossario ALA de Bibliotecología y Ciencias de la Información**. Madrid: Ediciones Díaz de Santos, c.1988. p.335-336.

Apêndice A – Modelo de Ficha Terminológica

TERMO: Fonte:
Definição:
Contexto:
Sigla:
Sinônimo:
Variante:
Área: Organização e Tratamento da Informação Subárea: Representação Temática
Observações:
Compiladora: Verônica Ezequiel
Data: