



UNIVERSIDADE ESTADUAL PAULISTA  
CAMPUS DE MARÍLIA  
Faculdade de Filosofia e Ciências

BRÍGIDA MARIA NOGUEIRA CERVANTES

**A CONSTRUÇÃO DE TESAUROS COM A INTEGRAÇÃO DE  
PROCEDIMENTOS TERMINOGRÁFICOS**

MARÍLIA  
2009

BRÍGIDA MARIA NOGUEIRA CERVANTES

**A CONSTRUÇÃO DE TESAUROS COM A INTEGRAÇÃO DE  
PROCEDIMENTOS TERMINOGRÁFICOS**

Tese apresentada ao Programa de Pós-Graduação em  
Ciência da Informação da Faculdade de Filosofia e  
Ciências, da Universidade Estadual Paulista - UNESP,  
*Campus* de Marília, para a obtenção do título de  
Doutor em Ciência da Informação.

Área de Concentração: Informação, Tecnologia e  
Conhecimento.

Linha de Pesquisa: Produção e Organização da  
Informação.

Orientadora: Profa. Dra. Mariângela Spotti Lopes Fujita

MARÍLIA  
2009

### Dados Internacionais de Catalogação-na-Publicação (CIP)

C419c Cervantes, Brígida Maria Nogueira.

A construção de tesouros com a integração de procedimentos terminográficos / Brígida Maria Nogueira Cervantes. – Marília, 2009. 209 f.: Il ; 30 cm.

Tese (Doutorado em Ciência da Informação) – Universidade Estadual Paulista, Faculdade de Filosofia e Ciências, Marília, 2009. Bibliografia: f. 177-198.

Orientadora: Mariângela Spotti Lopes Fujita.

1. Construção de tesouros. 2. Terminografia. 3. Linguagens documentárias alfabéticas. I. Autor. II. Fujita, Mariângela Spotti Lopes. III. Universidade Estadual Paulista. Faculdade de Filosofia e Ciências. Marília. IV. Título.

CDD 025.49

CDU 025.43.05

BRÍGIDA MARIA NOGUEIRA CERVANTES

## **A CONSTRUÇÃO DE TESAUROS COM A INTEGRAÇÃO DE PROCEDIMENTOS TERMINOGRÁFICOS**

Tese apresentada ao Programa de Pós-Graduação em Ciência da Informação, da Faculdade de Filosofia e Ciências da Universidade Estadual Paulista Júlio de Mesquita Filho – UNESP, *Campus* de Marília, como requisito parcial para a obtenção do título de Doutor em Ciência da Informação.

Área de concentração: Informação, Tecnologia e Conhecimento  
Linha de pesquisa: Produção e Organização da Informação  
Orientadora: Profa. Dra. Mariângela Spotti Lopes Fujita

Marília, 25 de setembro de 2009.

### **BANCA EXAMINADORA**

---

Professora Livre-docente Mariângela Spotti Lopes Fujita (Orientadora)  
Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília - SP

---

Professora Doutora Maria de Fátima Gonçalves Moreira Tálamo  
Escola de Comunicação e Artes, Universidade de São Paulo - SP

---

Professor Doutor João Batista Ernesto de Moraes  
Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília - SP

---

Professora Livre-docente Marta Lígia Pomim Valentim  
Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília - SP

---

Professora Doutora Vera Regina Casari Boccato  
Centro de Educação e Ciências Humanas, Universidade Federal de São Carlos - SP

*Ó Deus! Dedico-lhe este trabalho por dar-me força, amparo para enfrentar as adversidades... durante este período e por tornar-me capaz de alcançar tão acalentado Sonho!*

*À Santa Teresinha do Menino Jesus por mostrar-me que Deus nunca nos inspira desejos que não possam ser satisfeitos, mesmo que eles possam, às vezes, parecer inatingíveis!*

*Aos meus familiares especiais:*

*Meu Esposo Cervantes por se dedicar integralmente ao nosso Lar, mantendo-se firme, amoroso, apoiando e tornando os meus, em Nossos Sonhos!*

*Meus Filhos amados por se manterem fiéis aos compromissos de vida em família e em comunidade pautadas no amor e na fé, na honestidade e solidariedade.*

*Parabéns à bióloga Vivian e ao químico Thiago!*

*Minha Mãe Terezinha (fortaleza) por suas palavras de confiança, pelo reforço de sempre e pelas muitas e muitas orações!*

## AGRADECIMENTOS

*O sucesso não é uma conquista solitária. Ao buscar o êxito, todo profissional sempre está rodeado de amigos e colaboradores, verdadeiros “Anjos”, que Deus apresenta em nossas vidas e desse modo contribuem para a realização de nossas conquistas. Estes agradecimentos sinceros e profundos são dirigidos às pessoas que, de muitas formas ou, de alguma forma, contribuíram para a realização dessa grande e especial Jornada! Todas estão em meu coração. Dentre as quais:*

*À orientadora: Profa. Dra. Mariângela Spotti Lopes Fujita por ser uma figura exemplar de professora e de pesquisadora, por seu dinamismo, seus ensinamentos, suas orientações, acompanhamento e contribuições para o desenvolvimento desta pesquisa, além das oportunidades e dos incentivos constantes.*

*Aos membros da banca de defesa: Prof. Dr. João Batista Ernesto de Moraes pela confiança em nosso trabalho, por suas sugestões pontuais no exame de qualificação e na defesa, pelo acolhimento e amizade desde o mestrado. À Profa. Dra. Maria de Fátima Gonçalves Moreira Tálamo pela forma como compartilha seu conhecimento, pelo proficiente apoio e valiosas contribuições durante o exame de qualificação e na defesa. À Profa. Dra. Marta Lígia Pomim Valentim pelo profissionalismo brilhante e sabedoria em suas palavras, pelo incentivo à minha formação em pesquisa, pela amizade e jornada compartilhada. À Profa. Dra. Vera Regina Casari Boccato pela amizade e colaboração. Aos suplentes da banca de defesa: Prof. Dr. José Augusto Chaves Guimarães por seus ensinamentos, pelo apoio e oportunidade de acesso à preciosa bibliografia. À Profa. Dra. Silvana Drumond Monteiro e ao Prof. Dr. Miguel Luiz Contani pelo companheirismo, pela cooperação e amizade.*

*Aos professores do doutorado por suas importantes contribuições. Aos colegas do doutorado pela convivência harmoniosa e solidária. Aos amigos da Biblioteca e da Seção de Pós-Graduação da UNESP de Marília pelos anos compartilhados, apoio e amizade. Aos colegas e amigos, muito especiais do Departamento de Ciência da Informação, do Departamento de Letras Estrangeiras Modernas e do Sistema de Bibliotecas da UEL pelo apoio, compartilhamento, companheirismo e o incentivo à minha capacitação.*

### **O Líder Vai Adiante**

*"A este o porteiro abre; e as ovelhas ouvem a sua voz; e Ele chama pelo nome as suas ovelhas, e as conduz para fora. Depois de conduzir para fora todas as que lhe pertencem, vai adiante delas, e as ovelhas o seguem, porque conhecem a sua voz."*

*(João 10:3, 4)*

*"[...] o sentido da vida é o conhecimento que, desse modo, é ilimitado pela amplitude da pergunta, e é, ao mesmo tempo, limitado e útil pelo alcance de nossa capacidade de resposta."*

*Carlos Vogt*

CERVANTES, Brígida Maria Nogueira. **A construção de tesouros e a integração de procedimentos terminográficos**. 2009. 209 f. Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2009.

## RESUMO

Investiga a integração da Terminografia para a construção de tesouros na busca de procedimentos terminográficos que podem ser aplicados em conjunto com procedimentos metodológicos existentes de análise de assunto, para o aprimoramento da representação de conceitos na construção de tesouros. Realiza um estudo teórico-metodológico da construção de tesouro, com enfoque na identificação de conceitos em áreas de especialidade para a organização e recuperação temática da informação. Apresenta como objetivo geral enunciar um modelo metodológico para a construção de tesouro com a integração de procedimentos terminográficos. Como objetivos específicos: analisar e sintetizar referenciais teórico-metodológicos sobre construção de tesouros; identificar os principais aspectos teórico-metodológicos da Terminologia/Terminografia contribuintes para a construção de tesouros; e apresentar proposta de um modelo metodológico terminográfico para a construção de tesouros. A metodologia da pesquisa qualifica-se por sua natureza bibliográfica, descritiva e exploratória, concentrando-se na abordagem temática do vocabulário de áreas de especialidade. Enfatiza como um resultado do trabalho aplicado o “Tesouro Terminográfico Preliminar em Gestão da Informação”, disponível na web. Conclui que o aprimoramento de etapas da construção de tesouro, aliado a contribuições de procedimentos terminográficos, produz uma representação de conceitos, por meio de termos, tendo em vista a obtenção de um vocabulário consistente, que compõe a base para a organização e recuperação temática da informação, e compatível com a demanda de áreas de especialidade.

**Palavras-chave:** Construção de tesouros. Terminografia. Tesouro Terminográfico. Organização e representação da informação e do conhecimento.



CERVANTES, Brígida Maria Nogueira. **Thesauri construction and terminographic procedures integration**. 2009. 209 f. Thesis (Doctorate degree in Information Science) – Philosophy and Science Faculty. Universidade Estadual Paulista, Marília, 2009.

## **ABSTRACT**

It investigates the terminographic integration for the thesauri construction in the search of terminographic procedures that may be used together with existing methodological procedures of subject analysis, for the improvement of the concepts representation in the thesauri construction. It is a theoretical-methodological study of the thesauri construction, focusing in the concepts identification in specialized area for the representation and thematic information retrieval. The general purpose of this research is to conceive a methodological model for the thesauri construction with the terminographic procedures integration. The specific purposes are to analyze and synthesize theoretical-methodological framework on thesauri construction, identify the main theoretical-methodological aspects of Terminology/Terminography which contribute to the thesauri construction, and present a proposal of a terminographic methodological model for the thesaurus construction. The research methodology is bibliographical, descriptive and exploratory, focusing on the thematic approach of vocabulary of speciality areas. It emphasizes as a result of this work, the “Preliminary Terminographic Thesauri in Information Management”, available in the web. It concludes that the improvement of thesauri construction stages, with the contributions of terminographic procedures, produce a concepts representation, by means of terms, having in mind the acquisition of a consistent vocabulary, which forms the basis for the organization and thematic information retrieval, and compatible with the demand of specialized areas.

**Key words:** Thesaurus construction. Terminography. Terminographic Thesauri. Information and Knowledge Organization and Representation.

## LISTA DE QUADROS

<b>Quadro 1</b>	– Abreviaturas utilizadas em tesouros como símbolos dos termos e seus equivalentes em inglês e francês.....	50
<b>Quadro 2</b>	– Relações básicas de um tesouro e sua simbolização.....	51
<b>Quadro 3</b>	– Tipologias de vocabulários controlados e suas aplicações em áreas de especialidade.....	61
<b>Quadro 4</b>	– O Tesouro - definições conforme Normas e/ou Autores.....	66
<b>Quadro 5</b>	– Demonstração da estrutura interna das Diretrizes IBICT (1984); UNESCO (1993) e ANSI/NISO (2005).....	75
<b>Quadro 6</b>	– Síntese das Etapas de Construção de um Tesouro segundo as Diretrizes IBICT (1984); Diretrizes UNESCO (1993); e Diretrizes ANSI/NISO (2005).....	206
<b>Quadro 7</b>	– Métodos de Compilação de termos: terminologias e definições.....	209
<b>Quadro 8</b>	– Síntese das etapas de construção de tesouros segundo autores.....	113
<b>Quadro 9</b>	– Tesouros e seus métodos de construção e o uso de equipamento de informática.....	118
<b>Quadro 10</b>	– Sistematização de etapas da construção de tesouros.....	163
<b>Quadro 11</b>	– Ficha terminológica-guia de registro de termos.....	169

# SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b> .....	11
<b>2</b>	<b>A REPRESENTAÇÃO DE CONCEITOS PARA A CONSTRUÇÃO DE TESAuros</b> .....	27
2.1	Organização e Representação do Conhecimento registrado.....	27
2.2	Para entender as linguagens documentárias alfabéticas.....	39
2.3	Tesauros: uma perspectiva diacrônica de parâmetros e modelos para sua construção.....	54
2.3.1	Possíveis influências a partir da Indexação Alfabética de Assuntos.....	56
2.3.2	Possíveis influências a partir da Classificação Bibliográfica.....	58
2.4	Etapas da construção de um tesauro.....	72
2.4.1	Etapas da construção de tesauro segundo as diretrizes.....	76
2.4.2	Etapas da construção de tesauro segundo os autores.....	104
2.5	Tesauros e seus métodos de construção.....	114
<b>3</b>	<b>ASPECTOS CONTRIBUINTES DA TERMINOLOGIA E DA TERMINOGRAFIA NA CONSTRUÇÃO DE TESAuros</b> .....	124
3.1	Tendências teóricas e metodológicas da Terminologia: de Wüster a Temmermann.....	124
3.1.1	Wüster: a Teoria Geral da Terminologia, ou terminologia clássica.....	126
3.1.2	Teoria Comunicativa da Terminologia: Maria Teresa Cabré e Maria das Graças Krieger .....	128
3.1.3	Teoria Sociocognitiva da Terminologia: Rita Temmermann.....	130
3.1.4	Teoria Socioterminológica: Enilde Faulstich.....	132
3.1.5	Sobre a Rede Ibero-americana de Terminologia (RITerm).....	134
3.1.5.1	Algumas considerações acerca da Terminologia no Brasil.....	135
3.2	Do Conceito ao Termo: a natureza metodológica da Terminografia... ..	139
3.3	Metodologia da Terminografia: atividade de pesquisa terminológica.. ..	142
3.3.1	Abordagem onomasiológica e semasiológica.....	144
3.3.2	Etapas da pesquisa terminológica.....	146
<b>4</b>	<b>A INTEGRAÇÃO DA TERMINOGRAFIA NA CONSTRUÇÃO DE TESAuros</b> .....	159
4.1	Para um modelo metodológico para construção de tesauros.....	159
4.2	Etapas da pesquisa terminológica para construção de tesauros.....	161
<b>5</b>	<b>CONSIDERAÇÕES FINAIS</b> .....	174
	<b>REFERÊNCIAS</b> .....	177
	<b>APÊNDICES</b> .....	199
	APÊNDICE A.....	200
	APÊNDICE B.....	207

## 1 INTRODUÇÃO

Nosso estudo apresenta como tema a integração da Terminografia na construção de tesouros. A problematização foi constituída a partir de investigações que tratam de metodologias propostas para a construção de tesouros fundamentada na abordagem da pesquisa terminológica de áreas de especialidade. O interesse pelo estudo da construção de linguagens para propósitos especiais tem crescido muito nas últimas décadas, motivado principalmente pelos avanços científicos e tecnológicos.

Torna-se, portanto, indispensável que a ciência, seja ela básica ou aplicada, mantenha sempre atualizado o seu discurso metalinguístico com vistas às necessidades que se renovam. Ao produzir novos conhecimentos, os especialistas em uma determinada matéria criam novos conceitos que necessitam de denominação e de identificação. As unidades resultantes do processo de denominação são os termos científicos e técnicos que, uma vez identificados e sistematizados em instrumentos terminológicos, veiculam os saberes e os fazeres consolidados por uma ciência ou tecnologia.

Nesse sentido, o tesouro é, ao lado de vocabulários técnico-científicos e outros recursos terminológicos, um importante instrumento de apoio às pesquisas científicas nas áreas de conhecimento. Em unidades de informação, o tesouro é considerado uma modalidade de linguagem documentária e tem a função de subsidiar, ao mesmo tempo, os processos de representação do conteúdo do documento e de recuperação temática da informação.

Fujita (1992, p. 24) evidencia que a “função de controle de vocabulário do tesouro, tanto na entrada (representação do conteúdo do documento) quanto na saída (representação da questão de busca), servirá igualmente ao bibliotecário-indexador e ao usuário do sistema de informação”. Desse modo, o tesouro pode ser visto como um instrumento de mediação capaz de permitir que bibliotecários indexadores, bibliotecários de serviço de referência e informação e usuários de um sistema de informação especializado compartilhem de um mesmo vocabulário.

Esta pesquisa amparou-se em estudos no âmbito da área da Ciência da Informação e para apresentar precedentes da Ciência da Informação, como campo de estudos, observou-se, conforme alguns autores, que esse campo científico é relacionado por, pelo menos, dois acontecimentos históricos marcantes:

O primeiro demarca a sua origem a partir do aparecimento da Documentação no fim do Século XIX, e o outro, a começar da explosão informacional ocorrida após a Segunda Guerra Mundial e que culminou com a publicação de um artigo de Vannevar Bush (1945), o qual não somente delimitou o problema que preocupava a todos, organizar a informação para a sua disseminação e transferência aos sistemas de informação, mas envidou esforços para encontrar uma solução com a utilização da tecnologia existente na época chegando a propor “uma máquina com capacidade de ‘associar idéias’, que duplicaria os ‘processos mentais artificialmente’, o *Memex*” (OLIVEIRA, 2005, p. 12).

Recorreu-se a Rayward (1997, p. 290, grifo do autor) ao declarar que esse campo de estudos tem sua origem na “Documentação”, termo estabelecido por Paul Otlet e Henri La Fontaine, dois advogados belgas, no final do Século XIX, em Bruxelas, e que envolveu a divulgação de novas idéias, novas estruturas de informação, além do desenvolvimento de novas técnicas e ferramentas para o acesso e uso de informações.

Nessa direção, observamos Tálamo et al. (2002, p. 7) quando destacam Paul Otlet como o criador de estruturas informacionais pela paz mundial e argumentam que “Otlet conferiu à documentação um papel social de maior relevância, uma vez que o mesmo priorizava em suas análises a informação como processo, gerando o conhecimento, diminuindo a insegurança e a incompreensão entre os homens”.

Na visão de Rayward (1997, p. 290), os principais conceitos de Ciência da Informação considerados na atualidade, “os sistemas técnicos e atividades profissionais sobre os quais se sustentam, já estavam implícitos e operacionalizados, em 1895 e nas décadas subsequentes”, quando foram criados no âmbito do Instituto Internacional de Bibliografia (IIB), atualmente denominado Federação Internacional de Documentação (FID). Essas idéias e práticas, na terminologia atual, correspondem às denominações de “informação tecnológica, recuperação de informação, estratégias de busca, centros de informação, serviços de informação pagos, bases de dados *on-line*, [...]” (RAYWARD, 1997, p. 290).

Saracevic (1996) afirma que o conceito de Ciência da Informação como um campo de estudos surgiu no início dos anos 60, após discussões feitas na época e que foram sintetizadas por Borko (1968). Robredo (2003, p. 55, grifo do autor) corrobora essa argumentação e complementa que “a primeira formulação do que seria a ‘ciência da informação’ surgiu em 1961 e 1962, como resultado dos trabalhos realizados no quadro das conferências do *Georgia Institute of Technology*” (conhecido como *Georgia Tech*).

Na visão de Smit, Tálamo e Kobashi (2004) a Ciência da Informação “é um campo científico em constituição e não apresenta consenso quanto ao seu objeto e campo abrangido”. Por essa razão, ao conceituarmos Ciência da Informação<sup>1</sup>, valemo-nos da afirmação de Borko (1968) que a apresenta como “a disciplina que investiga as propriedades e o comportamento da informação, as forças que regem o fluxo informacional e os meios de processamento da informação para a otimização do acesso e uso”.

O autor supracitado esclarece que se trata daquele corpo de conhecimentos relacionados à produção, coleta, organização, armazenamento, recuperação, interpretação, transmissão, transformação e utilização da informação (BORKO, 1968, p. 1). O uso da informação visa a atender a uma necessidade oriunda da vida social, a qual leva uma pessoa a buscar informação, tanto para resolver um problema, como para alcançar um objetivo específico.

Segundo Robredo (2003, p. 56), a definição formulada por Borko (1968) guarda semelhança “à que saíra das conferências de 1961-1962, com que se identifica, mas vai além, assinalando que a biblioteconomia e a documentação são componentes aplicados da ciência da informação”.

A Ciência da Informação contém tanto um componente de ciência pura, que investiga o objeto sem consideração para com sua aplicação, quanto um componente de ciência aplicada, que desenvolve serviços e produtos. Autores como Borko (1968), Saracevic (1996) consideram a Ciência da Informação como uma ciência interdisciplinar e que a interdisciplinaridade é parte constituinte e determinante de sua gênese.

---

<sup>1</sup>Com base em definições de Robert S. Taylor apresentadas no *Annual Review of Information Science and Technology*, em 1966, devidamente refletidas e sintetizadas por Harold Borko em seu artigo *Information Science - what is it?*, publicado em 1968.

Dessa forma, explica Borko (1968), mantém relações interdisciplinares importantes com a Matemática, a Lógica, a Lingüística, a Psicologia, a Tecnologia da Computação, a Pesquisa de operações, as Artes Gráficas, as Comunicações, a Administração, a Biblioteconomia e Documentação.

Ao relacionar uma área do saber com o fazer profissional dessa mesma área, a Ciência da Informação significa ainda “nomear os pontos de intersecção, ou convergência, entre os conceitos de uma área científica e os procedimentos práticos constitutivos do fazer profissional dessa mesma área” (SMIT; BARRETO, 2002, p. 9).

A esse respeito, Smit (1993, p. 82, grifo da autora) chama a atenção para a aproximação de profissões que compartilham de documentos, enquanto suportes de informação, objetivos próximos e técnicas semelhantes, a saber: Museologia, Arquivística e Biblioteconomia e Documentação, identificando-as como irmãs da organização da informação e cunhando-as com a expressão ‘**3 Marias**’.

A Biblioteconomia, como um ramo de saber, abrange um conjunto de conhecimentos teóricos e práticos essenciais à identificação, seleção e organização das informações contidas nos mais diversos suportes informacionais, documentos, que na concepção de Paul Otlet são os livros, as revistas, os jornais; as peças de arquivos, as estampas, as fotografias, as medalhas, as músicas; são, também, os filmes, os discos e toda a parte documental que precede e sucede a emissão radiofônica.

Além dos textos e das imagens, existem objetos documentais por si mesmos, as réalias (amostras de espécimes, protótipos, ou seja, as representações em três dimensões). Em síntese: “representação da realidade sob uma forma literária (escrito, texto), ou gráfica, ou plástica (ícone, imagem)” (OTLET, 1934, p. 372). Em complemento, a Documentação “é uma disciplina que se propõe tratar a informação para fins de recuperação. Tem como princípio geral garantir o acesso do usuário aos conteúdos informacionais, no menor tempo possível, com o menor custo” (TÁLAMO, 2001, p. 142).

Em razão disso, a Documentação constitui-se por uma série de operações distribuídas entre pessoas e organismos diferentes: o autor, o copista, o impressor, o editor, o livreiro, o bibliotecário, o documentalista, o bibliógrafo, o crítico, o analista, o compilador, o leitor, o pesquisador, o trabalhador intelectual (BORKO, 1968).

Na visão de Borko (1968), a Biblioteconomia e a Documentação são disciplinas da Ciência da Informação e, por essa razão, as técnicas e procedimentos usados por bibliotecários e documentalistas são, ou deveriam ser, baseados nas descobertas teóricas da Ciência da Informação e, inversamente, o teórico deveria estudar as técnicas testadas por aplicações práticas. “Além de ter um forte aspecto operacional, é conceitualmente afetada por uma tecnologia intensa, com elevado teor de inovação e em contínua mutação” (SMIT; BARRETO, 2002, p. 18).

Por conseguinte, a pesquisa que H. Borko se refere inclui a investigação de representações da informação tanto no sistema natural quanto no artificial, o uso de códigos para transmissão eficiente de mensagens e o estudo de dispositivos e técnicas de processamento de informação tais como computadores e seus sistemas de programação (BORKO, 1968).

São incontáveis as melhorias nas várias instituições e procedimentos dedicados ao acúmulo e transmissão de conhecimento, incluindo-se, nesse caso, livros para registrar o conhecimento; escolas para ensinar o conhecimento acumulado de muitas gerações; bibliotecas para armazenar e disseminar o conhecimento; filmes e televisão para a exibição visual do conhecimento; periódicos para a comunicação escrita dos últimos avanços técnicos em áreas especializadas; e conferências para a comunicação oral da informação (BORKO, 1968), com base nas perspectivas de soluções de problemas de representação de conceitos, na construção de tesouros, para a organização e comunicação da informação.

Nossa pesquisa estabeleceu vínculo com a Documentação para favorecer “como-melhor-representar-a-informação” em razão da exigência “cada vez maior da sociedade humana em organizar cada vez mais adequadamente a informação visando sua pronta recuperação em períodos de tempo cada vez mais breves”, argumenta Fujita (1992, p. 7). Concordando com a observação de Benveniste (1976), a linguagem só funciona como instrumento de comunicação se funcionar, simultaneamente, como meio para a construção do saber comunicado.

É fato que o relacionamento do homem com a informação antecede à primeira biblioteca. Mas é a partir desta instituição que se criam modos sociais próprios de recepção do conhecimento registrado: “foi a biblioteca o lugar [...] das obras raras, [...] da seleção de uma ordem, [...]. Através deste equipamento cultural,



e de suas variações, estabeleceram-se ordens e formas de socialização e de uso da informação armazenada”, argumentam Smit, Tálamo e Kobashi (2004).

É possível que profissionais de outras áreas, em razão de uma terminologia variada, também sintam os efeitos dessas mudanças terminológicas vivenciadas por bibliotecários, sejam eles indexadores, ou de serviços de referência e informação, pois, para atender as necessidades de informação do usuário, deparam-se com um número crescente de novas terminologias as quais, com certa frequência, são causadoras de incertezas referentes ao assunto em questão.

Por isso, sobreveio a necessidade de se criarem meios e instrumentos documentários que pudessem prover modos de organização para facilitar o acesso dos usuários às informações de que necessitam. Com base nesse contexto, a Terminologia exerce papel fundamental, uma vez que, ao proceder à sistematização dos conceitos de cada área do conhecimento, determina-se igualmente a condição de referência dos termos que irão compor instrumentos terminológicos para fins documentários. Tálamo, Lara e Kobashi (1992) argumentam: com isso se assegura que o conteúdo lexical das linguagens documentárias não se apresenta desprovido de indicações referenciais específicas, mas funcionam antes com termos cujos interpretantes remetem-se às significações estruturadas e registradas pelos discursos da Ciência.

A esse respeito, Hjørland declara que, em sua proposta de abordagem orientada ao domínio (HJØRLAND; ALBRECHTSEN, 1995), um ponto fundamental dessa visão é a alegação de que as “ferramentas, conceitos, significados, estruturas de informação, necessidades de informação e critérios de relevância são moldados em comunidade de discursos” (HJØRLAND, 2002, p. 258). O autor esclarece referindo-se a disciplinas científicas, que são partes da divisão do trabalho feita pela sociedade. Trata-se de uma comunidade de discursos, na qual ocorre um processo de comunicação ordenado e delimitado (HJØRLAND, 2002).

Hjørland (2002) publicou um artigo<sup>2</sup> em que expõe onze abordagens de estudos da Ciência da Informação que poderão beneficiar-se da análise de domínio, a saber: 1) produção de guias de literatura ou portais temáticos; 2) construção de classificações especiais e tesouros; 3) indexação e recuperação de especialidades; 4) estudos empíricos de usuários; 5) estudos bibliométricos; 6)

---

<sup>2</sup> No periódico *Journal Documentation*, intitulado: *Domain analysis in information science: eleven approaches, traditional as well as innovative*.

estudos históricos; 7) estudos de documentos e gêneros; 8) estudos críticos e epistemológicos; 9) estudos terminológicos, LSP (linguagens para propósitos especiais), semântica de base de dados e estudos do discurso; 10) estudos de estruturas e instituições em comunicação científica e 11) cognição científica, conhecimento especializado, e inteligência artificial.

Em razão de nossa pesquisa versar sobre a construção de tesouros em sistemas de informação especializados, no âmbito das abordagens destacadas por Hjørland (2002) interessa-nos particularmente a segunda abordagem: construção de classificações especiais e tesouros, a qual contempla a construção de tesouros. Para Hjørland (2002) as onze abordagens citadas não se exaurem nem se excluem mutuamente. O autor sugere que, em vez de aplicar qualquer abordagem isolada, seria melhor combinar abordagens. Conforme o exposto pelo autor, por exemplo, neste estudo entendemos que seria possível combinar a abordagem dois com a nove que trata de “estudos terminológicos, LSP (linguagens para propósitos especiais), semântica de base de dados e estudos do discurso” (HJØRLAND, 2002, p. 451).

Concordamos com o autor supracitado quando argumenta “combiná-las poderia servir para fortalecer a identidade da Ciência da Informação e corroborar a relação entre pesquisa e prática na área. E [...] podem representar uma importante abordagem para a análise de domínio na Ciência da Informação”, provendo informação sobre os estudos terminológicos em áreas de especialidades, identificados em comunidade de discursos, desde que fundamentados pela própria teoria (HJØRLAND, 2002, p. 451).

Além disso, observamos que, no âmbito desses estudos, há uma forte tendência das pesquisas em considerar não somente os aspectos cognitivos dos indivíduos, mas também os aspectos sócio-cognitivos, ou seja, os aspectos culturais, históricos, sociais, terminológicos e organizacionais que representam o contexto no qual o indivíduo está inserido. Esse indivíduo, produtor de conhecimento, “é um sujeito ativo que em sua relação com o mundo, com seu objeto de estudo, reconstrói (no seu pensamento) este mundo. O conhecimento envolve sempre um fazer, um atuar do homem” (REGO, 2008, p. 98). Fujita (2003) concorda com essa abordagem e reforça que é preciso “conhecer os objetivos do sistema de informação e a demanda da comunidade usuária”.

Essa preocupação encontra-se fundamentada em projetos coordenados pela Profa. Dra. Mariângela Spotti Lopes Fujita intitulados: “O contexto da leitura documentária de indexadores de bibliotecas universitárias em perspectiva sócio-cognitiva para a investigação de estratégias de ensino”, projeto integrado financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq); e “O contexto sócio-cognitivo do catalogador em bibliotecas universitárias: perspectivas para uma política de tratamento da informação documentária”, com financiamento para coleta de dados do Fundo de Pesquisa da Faculdade de Filosofia e Ciências da Universidade Estadual Paulista – UNESP (RUBI, 2008).

De acordo com Fujita (2006), “entender a perspectiva cognitiva e sócio-cognitiva no domínio da Organização da Informação em unidades de informação é de suma importância”. Sobre esse ângulo, a autora supracitada argumenta, ainda, que “o objetivo primeiro do tratamento de conteúdo documentário é a recuperação desse conteúdo pelo usuário”. Ademais, com isso, é possível ter uma melhor visão desse domínio para realizar processos de representação de termos em linguagens documentárias (FUJITA, 2006).

Outros estudos confirmam o exposto: Gomes e Campos (2004) verificaram que há “[...] muitos problemas relativos à unidade de representação, ou seja, ao termo de indexação”, e Araújo (2006) enfatizou a necessidade da construção de uma linguagem documentária consistente que possibilite a representação e a recuperação de temas por meio de termos/descriptores e que considere o uso de remissivas para o caso das variações denominativas, além de Boccato (2009) que verificou em sua tese de doutorado: “falta de vocabulário especializado e atualizado; de estrutura sintático-semântica inconsistente [...]; falta de remissivas e cabeçalhos específicos; e incompatibilidade entre a linguagem do sistema e a de busca do usuário”, presenciada em três áreas do conhecimento analisadas.

Além disso, é preciso considerar os “relatos verbais” de pesquisadores e profissionais usuários de unidades de informação acadêmicas, pois Cervantes (2004) constatou “a ausência de termos usuais/conhecidos que não são considerados na linguagem do sistema de recuperação da informação”. É fato que o uso de termos técnico-científicos confirmados pelo especialista deve ser a base de um sistema de recuperação de informação (CERVANTES; VALENTIM, 2007), contudo é importante considerar que a variação está presente em qualquer área do

conhecimento e, como tal, tem que compor a linguagem de um sistema de recuperação de informação. Mesmo quando se trata de uma linguagem de especialidade, peculiar de um grupo de especialistas, ocorre o processo de dinamização que gera novos termos, ressignificando-se termos confirmados e promovendo-se variações para conceitos já denominados.

A dinâmica dessas temáticas que estão em crescimento contínuo, apresenta diferentes problemas terminológicos (BARITÉ, 2005, p. 25). A organização da informação é o que precede a busca, ela permite a elaboração de repositórios estruturados de informação e desenvolve técnicas que fornecem subsídios para evitar a criação de redes de conceitos confusas onde os usuários gastam muito tempo buscando sem encontrar o que precisam.

Essa condição influencia sobremaneira o trabalho do pesquisador que, ao evitar informações desnecessárias, logrará maior tempo para trabalhar com informações que poderão trazer diferencial para sua pesquisa. Organização da informação e organização do conhecimento são terminologias muito próximas e que estão sendo usadas com o mesmo enfoque, embora apresentem características que as distinguem.

Para Brascher e Café (2008 p. 6), a organização da informação aplica-se às ocorrências individuais de objetos informacionais; e a organização do conhecimento, aos conceitos. Desse modo, a organização da informação preocupa-se com a organização de um objeto físico, material. Já a organização do conhecimento refere-se à organização da cognição humana. Pode-se dizer que o objeto investigativo da organização do conhecimento reside naquilo que Dahlberg (1993, p. 214) denomina como “conhecimento em ação”, ou seja, algo acerca do qual existe consenso social, um conhecimento registrado e socializado, cuja organização e representação desenvolvidas de modo que, a partir dele, possa ser gerado novo conhecimento.

Para Dahlberg (2006 apud FUJITA, 2008a), essa combinação simples de conceitos, na qual o objeto e sua própria atividade já são indicados, cobrem o âmbito e o objeto da área de Organização do Conhecimento, ou seja, “conhecimento” no sentido de “conhecer” e “organização” no sentido de ordenação de objetos e não de coletividades relacionadas às pessoas. Essa distinção é importante, por considerar que o nome Organização do Conhecimento refere-se ao objeto e atividade da área. Com a mudança de título do periódico *International*

*Classification para Knowledge Organization*, Dahlberg (2006) introduziu a conceituação de Organização do conhecimento como “[...] os objetos e atividades da teoria do conceito, classificação e indexação e representação do conhecimento [...]”, argumenta Fujita (2008a).

No entanto, é preciso considerar que, a partir de diferentes contextos, a organização do conhecimento e da informação requer a adoção de metodologias específicas respaldadas teoricamente, respeitando-se o ambiente onde foi produzida e os recursos informacionais disponíveis. Inclui-se, nesse processo, a representação documentária vista sob o prisma do conteúdo e da sua forma. A representação do conteúdo de um documento é denominada representação temática e abarca os processos de classificação e indexação e estabelece interfaces amigáveis com diferentes as áreas de estudos, em especial com a Linguística, a Terminologia e a Ciência da Computação.

No que se refere à forma e ao conteúdo a representação descritiva e a representação temática, auxiliam significativamente na recuperação da informação a partir da determinação e disponibilização de pontos de acesso, apoiadas pela inovação tecnológica. Em razão disso, atentar às características do ambiente e de seus usuários são fatores determinantes no processo de organização da informação e do conhecimento. Nesse contexto, faz-se necessário refletir a respeito de modelos metodológicos de organização da informação e do conhecimento que atendam as reais necessidades de informação de usuários no âmbito de um sistema de informação especializado.

A informação indica um conteúdo documentário que se encontra disponível nos mais variados meios e suportes e, quando incorporada aos sistemas de informação, essa se acumula e se agrega em uma estrutura ou repositório. Para a recuperação da informação, é importante considerar que o sistema de informação que a disponibiliza, esteja amparado por um vocabulário capaz de operar os dois modos de comunicação, ou seja, a linguagem do sistema de informação e a linguagem de busca do usuário de áreas especializadas.

Na visão de Fujita (2003, p. 176), “é preciso usar estratégias de exploração textual e ter domínio da estrutura textual complementado pelo conhecimento lingüístico. A autora argumenta que “é preciso conhecer os objetivos do sistema de informação e a demanda da comunidade usuária para a seleção dos conceitos”. Nossa pesquisa sobre a Construção de Tesouros com a integração da

Terminografia apresenta-se como uma continuação do estudo da contribuição da Terminologia na construção de linguagem documentária e a aplicabilidade do Protocolo Verbal, defendido em 2004, quando estudamos o processo de categorização, identificação e confirmação de conceitos/termos para a construção de Linguagem Documentária que atuasse como instrumento de organização e recuperação da informação documentária (CERVANTES, 2004).

Naquela investigação, procuramos explicitar a importância do estudo da interface da Terminologia e Ciência da Informação, por meio da aplicação convergente de metodologias da Terminologia e de metodologia introspectiva de coleta de dados, com a aplicação, nesse caso, do Protocolo Verbal Interativo, nos moldes de Nardi (1999). A interação entre o usuário e o sistema de recuperação da informação compreende o ponto para onde converge a pesquisa voltada para uma abordagem cognitiva e sócio-cognitiva em Ciência da Informação.

A abordagem cognitiva refere-se, de forma mais simples, aos estudos que consideram o conhecimento humano, tanto sob ponto de vista de processamento quanto de representação, como parâmetro para análise e elaboração de teorias e metodologias. O foco residiu, portanto, na cognição – o processo de conhecer humano que oferece uma perspectiva de investigação baseada na compreensão, no processamento e na representação (FUJITA; CERVANTES, 2005).

Com os resultados obtidos no estudo sobre “a pesquisa terminológica no contexto do processo de inteligência competitiva”, verificamos que a metodologia do protocolo verbal proposta contribuiu favoravelmente para o processo de categorização e, posteriormente, no processo de confirmação dos termos no subdomínio estudado; exigiu maior contato com pesquisadores e profissionais altamente qualificados; permitiu o processo de aprendizagem visando à construção do conhecimento e embasou a formulação de uma proposta de estabelecimento da estrutura conceptual ou categorização de subdomínio. Podemos afirmar que a abordagem cognitiva, mediante princípios teóricos e metodológicos em consonância com a aplicação do Protocolo Verbal, modalidade interativo, obteve respostas muito satisfatórias, como técnica de coleta para acesso ao conhecimento e processos de conhecimentos dos especialistas e profissionais de áreas especializadas envolvidos (CERVANTES; FUJITA; NARDI, 2003; FUJITA; CERVANTES, 2005; CERVANTES; FUJITA; RUBI, 2008).

Vale lembrar que para o trabalho “Contribuição para a terminologia do processo de inteligência competitiva: estudo teórico e metodológico” (CERVANTES, 2004), a área curricular de “Gestão da Informação”, no âmbito da área maior, Ciência da Informação, foi a área de especialidade selecionada como campo de trabalho empírico que necessita de instrumentos adequados que atendam ao desenvolvimento dessa área de estudo e às questões terminológicas de seus usuários. Além disso, com os resultados obtidos, colaboramos para a fixação de terminologias que podem ser usadas para compor um vocabulário de especialidade para o estabelecimento de relações entre os conceitos visando à recuperação da informação em de sistema de informação especializado.

Evidenciamos que, no decorrer de todo o processo da pesquisa, ou seja, desde a escolha da área de especialidade a ser estudada até a apresentação dos resultados, a Terminologia resultou em importantes contribuições. Por esse motivo, podemos afirmar que o estudo para aproximação da Terminologia, em especial da Terminografia, em seus diferentes aspectos, revelou-se fértil para a Ciência da Informação.

Cabe esclarecer que utilizamos a Terminologia para a identificação de conjunto de termos, contudo, não chegamos à proposta de um modelo metodológico para a construção de tesouros com a integração da Terminografia. Continuando os estudos sobre Terminologia no contexto da Ciência da Informação, investigamos, de 2002 a 2005, no âmbito do Grupo de Pesquisa “Informação, Conhecimento e Inteligência Organizacional”, Terminologias em Inteligência Competitiva (CERVANTES; VALENTIM, 2006; 2007); pesquisamos, desde 2002, no espaço do Grupo de Pesquisa “Análise Documentária”, Terminologia: estudo teórico e metodológico (CERVANTES; MORAES; FUJITA, 2002); (CERVANTES; FUJITA; NARDI, 2003); (CERVANTES, 2004); (FUJITA; CERVANTES, 2005); terminologias em Organização da Informação (CERVANTES; FUJITA; RUBI, 2008); além desses, participamos, desde 2004, do Grupo de Estudos Terminológicos da UEL, colaboramos no estudo de terminologias da Moda (Vestuário, Área Têxtil) (RAIMUNDO; CERVANTES; PEREIRA, 2006); e Pesquisa terminológica temática trilingue (SANTOS; CERVANTES, 2006).

A partir do conhecimento teórico-prático adquirido nesses estudos, considerando a pertinência das contribuições da Terminologia/Terminografia, apoiada em importantes fundamentações teórico-metodológicas e, considerando

que a construção de tesouros envolve certo grau de complexidade por exigir que os métodos empregados para a sua construção levem em conta, principalmente, a terminologia da área de especialidade em questão. Há a necessidade de harmonizar os vocabulários entre integrantes de um sistema de recuperação de informação especializada, ou seja, equilibrar os vocabulários dos usuários, do sistema de informação, com o utilizado na área especializada. Em razão disso, é importante que se tenha um vocabulário compatível que opere os dois modos de comunicação.

Nesse contexto, situa-se a presente pesquisa que nos permitiu estabelecer a integração de procedimentos terminográficos à complexidade da construção de tesouros como nosso problema de pesquisa. Consideramos, destarte, as seguintes premissas de pesquisa:

- O aprimoramento de etapas de construção de tesouros contribui para melhorar a representação temática da informação;
- A Terminografia possibilita a harmonização dos conceitos para o acesso à informação desejada em áreas especializadas.

O nosso objeto de estudo é, portanto, a integração da Terminografia na construção de tesouros. Os pressupostos acima são expandidos com a idéia de que os tesouros são construídos com a função de organizar e representar o conhecimento registrado em áreas especializadas, uma vez que a necessidade de representação e recuperação dos conceitos harmoniza-se com as necessidades informacionais da sociedade, para atingir sua finalidade, qual seja, a de produzir e transmitir conhecimento.

A hipótese que buscamos comprovar é a de que há a necessidade de combinar referenciais teórico-metodológicos da Documentação no que tange à análise de assunto, e aliá-los ao referencial da Terminografia para identificação de conceitos, visto que são dêixis que norteiam a identificação de terminologias que compõem a base para a organização do conhecimento de áreas de especialidade. A nossa tese de pesquisa é que o aprimoramento de tesouros como instrumentos terminológicos necessários à organização e representação de conceitos em sistemas de informação de áreas especializadas, depende de estudo teórico-metodológico que integre procedimentos terminográficos no processo de identificação de conceitos para construção de tesouros.



Nossa proposição de pesquisa é investigar a integração da Terminografia no contexto da Organização e Representação da Informação para apresentação de proposta de um modelo metodológico de construção de tesouros, especialmente no que se refere à identificação de conceitos em áreas de especialidade, com a integração de procedimentos terminográficos. Nossa pesquisa apresenta como objetivo geral enunciar um modelo metodológico para a construção de tesouro com a integração de procedimentos terminográficos. Como objetivos específicos definimos:

- Analisar e sintetizar referenciais teórico-metodológicos sobre construção de tesouros;
- Identificar os principais aspectos teórico-metodológicos da Terminologia/Terminografia contribuintes para a construção de tesouros;
- Apresentar proposta de um modelo metodológico para a construção de tesouros com a integração de procedimentos terminográficos.

Esse direcionamento da pesquisa se justifica pela crescente preocupação de estudiosos de linguagens documentárias alfabéticas, especialmente de tesouros, na busca por maior harmonização terminológica, por meio de suas representações do campo conceptual ao qual se direciona e por uma representação do conhecimento registrado mais compatível aos conteúdos documentários dos diferentes segmentos de usuários. A preocupação é organizar formas de melhorar a aproximação entre os sistemas de recuperação da informação e os seus usuários.

Nessa direção, nosso estudo investiga procedimentos terminográficos que podem ser aplicados em conjunto com procedimentos metodológicos existentes de análise de assunto, para o aprimoramento da representação de conceitos na construção de tesouros. Institucionalmente, pretende-se contribuir com a linha de pesquisa “Organização da Informação” do Programa de Pós-Graduação da UNESP – Câmpus de Marília, com referencial teórico-metodológico da Terminografia para a construção de tesouros.

A nossa pesquisa iniciou-se com o levantamento de referências sobre os temas Terminologia/Terminografia e Ciência da Informação/Documentação, com ênfase na construção de tesouros no intuito de conceitualizar cada um deles,

evidenciando semelhanças e diferenças. Nossas buscas às fontes de informação foram realizadas tendo como foco os seguintes termos preferidos: *Thesaurus Construction, Methodology and Construction Thesaurus, Compilation*. A partir de bases de dados internacionais, tais como: *Library Information Science Abstract (LISA)*; *Library, Information Science and Technology Abstracts (LISTA)*, *Wilson Web, Emerald*, e outras, disponíveis pelos Sistemas de Bibliotecas da Universidade Estadual Paulista e da Universidade Estadual de Londrina. Buscamos também em portais de revistas eletrônicas: Portal de Periódicos da CAPES, SciELO, Dialnet e Terminômetro, além do sítio da Biblioteca Digital de Teses e Dissertações, vinculado ao IBICT, disponível em: [<http://bdtd2.ibict.br/>].

Nessa busca, verificamos aspectos históricos, conceituação, função e importância da Terminologia/Terminografia, procedimentos terminográficos e aproximações com a construção de tesouros, considerando os idiomas francês, inglês e espanhol, além do português. Vale dizer que este estudo tenciona contribuir para os propósitos do Grupo de Pesquisa “Análise Documentária”, coordenado pela Profa. Dra. Mariângela Spotti Lopes Fujita. Para realizá-lo, empregamos como *corpus* de análise, publicações utilizadas em estudos do referido grupo. Esse *corpus* representativo contribuiu na identificação de elementos temáticos de um texto que descrevem o teor dos assuntos abordados pelos autores.

A nossa pesquisa qualifica-se por sua natureza bibliográfica, descritiva e exploratória, concentrando-se na abordagem temática da terminologia de áreas de especialidade. Os procedimentos metodológicos efetuados no estudo dos referenciais bibliográficos possibilitaram a elaboração dos capítulos teórico-metodológicos com a finalidade de apresentar uma visão a respeito do tema desta pesquisa, evidenciando as idéias, teorias, métodos e aplicações fundamentados em pesquisas científicas, como respaldo à comprovação ou não da hipótese proposta nesta tese. Em atenção aos objetivos propostos, apresentamos:

No capítulo 2, a discussão da representação de conceitos para a construção de tesouros, sua contextualização no âmbito da Organização e Representação do Conhecimento e da informação, além de destacar referencial teórico sobre as linguagens documentárias alfabéticas com foco em perspectiva diacrônica de parâmetros e modelos para sua construção. Discutimos, ainda, as etapas da construção de tesouros sob a égide de normas e autores, com a

finalidade de contribuir para a seleção de uma proposta de modelo metodológico a ser apresentado como principal objetivo dessa investigação.

No capítulo 3, refletimos sobre aspectos contribuintes da Terminografia na construção de tesouros, por meio de tendências teóricas e metodológicas da terminologia de Wüster a Temmermann; refletimos sobre a perspectiva onomasiológica, ou seja, do conceito ao termo, evidenciando a natureza metodológica da Terminografia. Explicitamos, com base em recomendações da normalização da Terminologia, os procedimentos terminográficos realizados na atividade de pesquisa terminológica para áreas de especialidade que podem ser aplicados em consonância com procedimentos existentes na Documentação para construção de tesouro. No capítulo 4, destacamos proposta de um modelo metodológico de integração da Terminografia na construção de tesouro. No capítulo 5, apresentamos as considerações finais.

## 2 A REPRESENTAÇÃO DE CONCEITOS PARA A CONSTRUÇÃO DE TESAUROS

Neste capítulo, analisamos e sintetizamos algumas das principais concepções sobre a representação de conceitos para a construção de tesouros no âmbito da Ciência da Informação, a Representação do conhecimento registrado e sobre áreas com as quais esta desenvolve relações interdisciplinares, como a Lingüística, a Semiótica, a Terminologia e outras. Abordamos em Para entender as Linguagens Documentárias Alfabéticas alguns aspectos observados sobre a evolução das linguagens documentárias, com ênfase na construção de linguagens documentárias alfabéticas. Destacamos em Tesouros: parâmetros, modelos para sua construção, o estudo sistemático de metodologias de construção de tesouros para a identificação de um modelo metodológico para a construção de tesouros. Finalizamos esse capítulo com a apresentação de uma síntese dos principais pontos estudados.

### 2.1 Organização e Representação do Conhecimento registrado

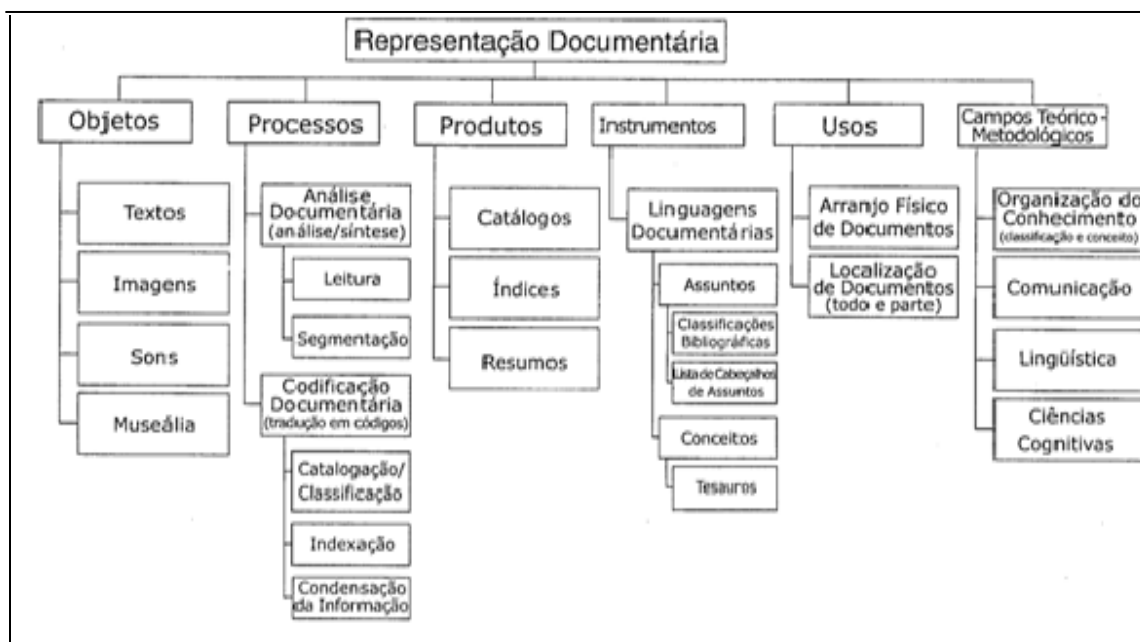
Na busca do conceito de “representação”, percebemos que a sua abordagem não é homogênea. Essa percepção tem como base a literatura observada, não de forma exaustiva, em diferentes áreas do conhecimento com as quais a representação estabelece vínculos. O conceito de representação relaciona-se com a Psicologia, Linguística, Semiótica, Ciência da Informação, especialmente, Documentação e outras. Em razão disso, concordamos com Bentes Pinto (2002) quando argumenta que sua compreensão conceptual é dada por meio da sua utilização, conforme evidenciamos nas definições apresentadas a seguir:

No senso comum, entre as definições para o vocábulo “representação” encontram-se quatro acepções, a saber: “1. Ato ou efeito de representar(-se). 2. Exposição escrita de motivos, de queixas, etc., a quem de direito. 3. Coisa que se representa. 4. Reprodução daquilo que se pensa” (FERREIRA, 2004). Destacamos a acepção 3: “Coisa que se representa”, e a

acepção 4: “Reprodução daquilo que se pensa”. Na Psicologia, o conceito “representação” encontra-se vinculado a essas duas acepções e, conseqüentemente, a dois objetos: o representado e o representante. O significado do termo representação, conforme o *Grand Dictionnaire de la Psychologie* (1991): “Designa de uma parte o processo de colocar em correspondência dois elementos de maneira que um (o representante) repete, substitui ou apresenta de outra maneira o outro (o representado)”. Na outra parte “um aspecto: do resultado desse processo, em ocorrência o único elemento representante de qualquer natureza que seja” (GRAND DICTIONNAIRE..., 1991, p. 596).

Na Linguística, representação é um conceito associado ao aparecimento da imagem verbo-mental no falante, quando a linguagem está no lugar de outra coisa: “a função do signo é a de representar, de tomar lugar de outra coisa evocando-a a título de seu substituto” (BENVENISTE, 1974). Na Semiótica, representar alcança um sentido bem próximo ao do conceito lingüístico de representação, definido por Benveniste. Peirce afirma que representar é “estar em lugar de, isto é, estar numa tal relação com outro que, para certos propósitos, é compreendido por alguma mente como se fosse a outra coisa” (PEIRCE, 1977, p. 61).

A representação sob o ponto de vista da “construção” conforme em GONZALEZ DE GOMEZ (1993, p. 220) ou com o significado de intermediação, considerando que a “representação dos vínculos que dão coesão à sociedade” (MARTÍN-BARBERO, 2008, p. 12), levando em conta que as representações se colocam entre o meio e o homem, concordamos com Dodebei (2002, p. 32) ao reconhecer que “as acepções de representação, representação da informação e representação do conhecimento são inúmeras”. Na Ciência da Informação, a autora supracitada apresenta, na Figura 1, um modelo de organização conceptual de Representação Documentária em que esboça uma das possibilidades de entendimento desse campo de estudo (DODEBEI, 2002, p. 42).



Fonte: DODEBEI (2002, p. 43).

**Figura 1** - Organização conceitual da Representação Documentária.

Na Ciência da Informação, a representação, no sentido de representação do conhecimento, é como a estrutura lógica da representação conceitual e, também, o resultado da identificação de conceitos por termos determinados em função da terminologia utilizada (DAHLBERG, 2006). A esse respeito, podemos dizer que a representação do conhecimento, nessa área, possui “dois aspectos distintos: o resultado da representação de conteúdo pela identificação de conceitos e a representação da estrutura lógica do conhecimento. Esse último como resultado da Organização do Conhecimento”.

A partir da mudança de denominação do periódico *International Classification*, publicado pela ISKO, para *Knowledge Organization: international journal devote to concept theory, classification, indexing, and knowledge representation*, em 1993, estabeleceu-se a conceituação para a Organização do Conhecimento como “[...] os objetos e atividades da teoria do conceito, classificação e indexação”, contudo, permanece a correspondência com o subtítulo da revista e admite artigos sobre pesquisas teóricas e aplicadas realizadas no contexto da organização e representação do conhecimento, de sistemas de representação do conhecimento e de sistemas de recuperação de documentos (ISKO, 2008)

Em razão disso, Guimarães (2001, p. 68) estrutura a área de Organização e Representação do Conhecimento em três núcleos básicos: 1)

Fundamentos de organização e representação do conhecimento: estudo da inserção da área no universo do conhecimento, com ênfase nas disciplinas que lhe são de interface; 2) Organização do conhecimento: estudo da base científica da área (princípios teóricos e metodologias); 3) Representação do conhecimento: estudo dos instrumentos (ferramentas) ou produtos da área.

Os sistemas de organização do conhecimento (*knowledge organization systems*), inseridos no contexto da Organização e Representação do Conhecimento, possuem aplicabilidade em ambientes informacionais impressos e eletrônicos, abrangendo todos os tipos de esquemas que possibilitam a organização do conhecimento, a gestão, além de sua disseminação. Esses sistemas, de acordo com Hodge (2000), têm por finalidade a organização dos documentos de um acervo e sua recuperação em sistemas de recuperação da informação. São caracterizados pelos sistemas de classificação e pelas linguagens documentárias alfabéticas, exemplificadas pelas listas de cabeçalhos de assunto e tesouros e, igualmente, por sistemas empregados em ambientes digitais como as redes semânticas e as ontologias (BOCCATO; RAMALHO; FUJITA, 2008).

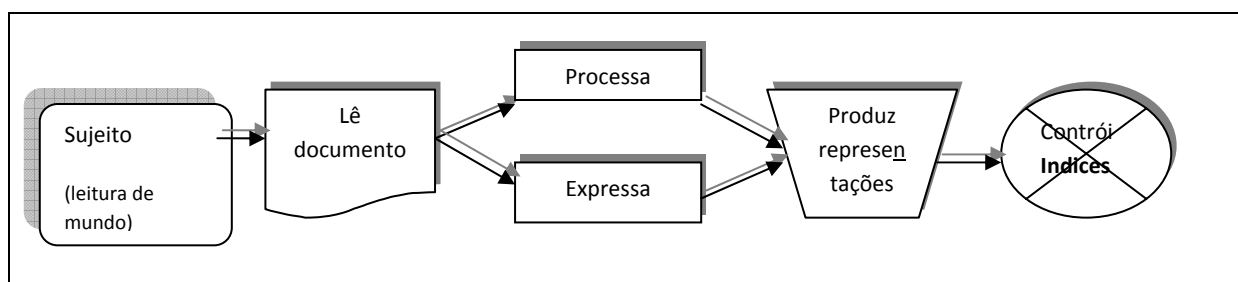
Por seu turno, os sistemas de organização do conhecimento sendo utilizados como instrumentos de comunicação e de mediação da informação, apoiados pelas linguagens documentárias alfabéticas, possibilitam a representação sintética das ideias dos autores de conteúdos documentários, por termos. A esse respeito, Moraes e Guarido (2008, p. 196) corroboram “a Terminologia, com suas regras e conceitos, contribui para que esse processo ocorra de forma precisa, uma vez que as Linguagens Documentais são elaboradas em áreas específicas para atender usuários com interesses específicos”. Desse modo, tendem viabilizar a obtenção de resultados precisos e relevantes que atendam às necessidades de buscas dos usuários de sistemas de informação de áreas especializadas.

Na visão de Dahlberg (2006, p. 12), o conhecimento é a “[...] certeza subjetiva ou objetivamente conclusiva da existência de um fato ou do estado de um caso. [...] não é transferível e só pode ser construído por meio da reflexão.” A autora considera que o conhecimento possui, fundamentalmente, uma natureza subjetiva, individual e que não é transferível, podendo ser somente elaborado por uma reflexão pessoal de alguém. Nesse sentido, evocamos Fujita (2008a) para argumentar que, em Guimarães (2001), diferencia-se a concepção de conhecimento da ISKO, enquanto processo individual, fundamentada em Dahlberg (2006), da concepção de

conhecimento enquanto “algo sobre o qual existe certo consenso social, trabalhando-se aqui com o conhecimento registrado e divulgado”.

O conceito de representação do conhecimento registrado “constitui-se na atividade de reprodução da percepção do(s) tema(s) abordado(s) em um documento, independentemente do suporte e da forma em que o conhecimento tenha sido registrado” (BENTES PINTO, 2002, p.119). Desse modo, a mensagem gerada por um sistema de informação, o qual, por sua vez, representa a informação contida nos documentos, é denominada representação documentária ou representação da informação codificada por meio da utilização de uma Linguagem Documentária (LIMA, 1998).

Ao refletirmos sobre a questão da representação de conceitos/termos para a construção de tesouros, distanciando-nos de possíveis divergências conceituais, referimos como significados similares as expressões representação da informação, representação do conhecimento registrado e representação documentária, uma vez que a representação, no âmbito da Documentação, visa a promover o acesso à forma de organização da informação para uso e posterior contribuição na construção de conhecimentos. Cabe ressaltar que ditas expressões são representações com foco na compreensão do conteúdo, das mensagens do documento e expressas em forma de conceitos sintagmáticos. Os conceitos acima expostos podem ser sintetizados por meio de uma visualização, tal como a apresentada a seguir:



Fonte: BENTES PINTO (2002, p. 119).

**Figura 2** – Representação do Conhecimento registrado em Ciência da Informação.

Explicando melhor o fluxo acima, podemos dizer que o sujeito (indexador) realiza suas leituras com base em seu conhecimento de mundo, lê o documento e o processa, expressando a sua leitura mediante a produção de outras representações e construindo índices para possibilitar a recuperação da informação



no âmbito dos sistemas de recuperação da informação de um modo mais eficaz. A representação e a indexação temática da informação constituem a síntese por meio da qual o conteúdo de um documento é identificado por conceitos/termos que se agrupam e se relacionam com os conteúdos de outros documentos de temas análogos ou semelhantes para atender aos usuários que buscam recuperá-los.

Fujita (1992) declara que, para existir a recuperação da informação desejada, é preciso que tanto a linguagem de busca quanto a fonte de informação estejam adequadamente representadas por intermédio de uma linguagem documentária. Em razão disso, a função das linguagens documentárias, no processo de representação para a recuperação da informação, é imprescindível, uma vez que proporciona a conexão entre as linguagens utilizadas tanto na entrada de conteúdos como na busca a esses conteúdos, num sistema de recuperação da informação.

A representação da informação por meio de linguagem é o resultado da operação de indexação da informação mediante o uso de termos que representam seus conteúdos temáticos para recuperar, no menor tempo possível, um tema específico em um conglomerado de documentos que compõem um sistema de informação (NAUMIS PEÑA, 2007). A representação, a informação está, fundamentalmente, ligada à atividade de indexação, uma vez que nesse fazer busca-se, por meio do uso do vocabulário de especialidade o estabelecimento de relações entre a linguagem do gerador da informação com a linguagem utilizada pelo usuário do sistema de informação. Essa visão considera a indexação como um processo ativo que trabalha o conceito, a mensagem. Trata-se de um processo carregado de valores, crenças, opiniões.

Portanto, constitui-se uma forma de representação social que reconstrói um novo documento na forma de índice, modificando-se o texto original ou anterior. Vale dizer que esse índice gerado é o objeto representante dos elementos integrantes do documento, ou seja, o objeto representado, dado que nele seus elementos são repetidos. Ainda que sejam identificadas duas características da representação, objeto representado e representante, sobre a atividade de indexação, é oportuno esclarecer que os índices não têm a função de substituir o documento, mas a de oferecer indicações para que seja possível ter acesso ao documento original, imprimindo maior agilidade e qualidade no processo.

Com respeito à qualidade da indexação, consoante com a NBR12676 (ABNT, 1992) “depende da hospitalidade da linguagem de indexação

utilizada”. Nessa direção, entendemos que, na construção de tesouro, devemos prever a inclusão de novos termos e atentar para as mudanças na terminologia, em razão da própria dinamicidade de uma linguagem. No que se refere à abordagem sobre a linguagem em seu sentido mais amplo, observou-se, no decorrer dos tempos, que a concepção de linguagem foi se transformando. “Até o Século XVIII predominou uma concepção teleológica que colocava em destaque sua origem e as regras universais da sua lógica”, conforme Cintra et al. (2002, p. 27).

A respeito desse período, mencionamos a contribuição da Gramática de *Port-Royal* de Claude Lancelot e Antoine Arnauld, de 1660. Para os gramáticos de *Port-Royal*, a língua é um sistema de signos. “As palavras e expressões são invólucros das idéias. Apenas as idéias ligam-se aos objetos. O nível mais elaborado é o nível lógico das idéias, a língua exterioriza essa lógica, [...] daí a gramática fundir-se com a lógica” (ARAÚJO, 2004, p. 21). Manifesta-se ainda (reflexões da Lógica de *Port-Royal*) a partir da publicação do livro *La logique ou l’art de penser*, de autoria de Antoine Arnauld e Pierre Nicole, uma obra de base aristotélica e cartesiana. No Século XIX, evidenciou-se uma concepção historicista que via a linguagem como um processo em evolução através dos tempos. Atualmente, predominam as concepções da linguagem como sistema em funcionamento (CINTRA et al., 2002, p. 27).

Para Koch (2006, p. 7, grifos da autora), a linguagem humana, pode ser “sintetizada com base em três concepções principais: a) como representação, ‘espelho’ do mundo e do pensamento; b) como instrumento, ‘ferramenta’ de comunicação; c) como forma, ‘lugar’ de ação ou interação”. Na primeira concepção, o homem representa para si o mundo através da linguagem. Na segunda, considera-se a língua como um código através do qual um emissor comunica a um receptor determinadas mensagens. Já na terceira concepção, analisa-se a linguagem como atividade, como forma de ação, como lugar de interação que possibilita aos membros de uma sociedade a prática de diferentes tipos de atos (KOCH, 2006).

Nesse sentido, Berlo (1999, p. 30) assevera que “toda a comunicação humana tem alguma fonte, uma pessoa ou um grupo de pessoas com um objetivo, uma razão para empenhar-se em comunicação”. Podemos dizer que essa condição se encontra presente em todos os domínios da atividade humana, em instituições que têm o dever de comunicar um saber como a família, o sistema de ensino, os sistemas profissionais e a administração. Algumas dessas instituições

especializaram-se no tratamento funcional da informação científica e técnica, desde a sua fonte até o usuário.

Morin (1999, p. 133) argumenta que a linguagem é tão fundamental “à constituição, à perpetuação, ao desenvolvimento da cultura quanto à inteligência, ao pensamento e à consciência do homem; tão consubstancial ao humano do humano que se pode dizer que a linguagem faz o homem”. Considera-se que a linguagem e sua expressão nos diálogos que acontecem diariamente, nos textos e nas imagens que construímos, nas referências que tornam viável a existência da memória, permitem que nos identifiquemos como membros deste ou daquele grupo social. Maturana e Varela (2001, p. 233) argumentam que “é a linguagem que faz com que existamos num mundo sempre aberto de interações lingüísticas recorrentes”. Com ela, “[...] é possível descrever, imaginar, relacionar”. Explica Morin (1999, p.134) que a linguagem “é ao mesmo tempo individual, comunicacional e comunitária”. Desse modo, na visão de Morin (1999), graças à linguagem:

a) toda operação cognitiva, toda aquisição, toda fantasia pode ser nomeada, classificada, estocada, rememorada, comunicada, logicamente examinada, conscientizada; b) as palavras, noções, conceitos operam como fatores de discriminação, seleção, polarização relativas a todas as atividades do espírito; c) o espírito pode combinar ao infinito palavras e frases e assim explorar ao infinito as possibilidades do pensamento (MORIN, 1999, p.134).

A linguagem traduz e transfere em enunciados sequenciais o que se manifesta como simultaneidade superposta no cérebro e no real. Assim, utilizamos a língua e outros sistemas de significação socialmente construídos para elaborar os significados, as representações que dão sentido à nossa vida. Salientamos, com isso, que é na linguagem que se constroem as culturas humanas, ou seja, que se elaboram os discursos e as narrativas que direcionam nossas ações.

Concordando com argumentações de Tálamo (2001, p. 142), admitimos que os processos informacionais se realizam em universos simbólicos que, por sua vez, se estabelecem por mecanismos lógico-lingüísticos e terminológicos, transformando-se em processos comunicacionais socialmente constituídos.

Segundo uma das premissas básicas em Organização do Conhecimento, defendidas por Barité (2001, p. 42), “o conhecimento se realiza a

partir da informação, e ao socializar-se se transforma em informação”. Esse conhecimento produzido pelo indivíduo e seu grupo, quando compartilhado, transforma-se em informação. A informação - que resulta do processo de compartilhamento do conhecimento – se convenientemente absorvida, altera o conteúdo informacional do indivíduo e seu grupo, provocando-lhes inquietações que conduzem à construção de conhecimentos.

Nesse sentido, “os instrumentos de representação da informação para indexação, armazenamento e recuperação de documentos, considerados como linguagens documentárias por Gomes (1990), Fujita (1992), Lara (1993c), Tálamo (1997b), Campos (2001), Dodebei (2002) e outros, cumprem importante papel na organização e recuperação da informação no âmbito de um sistema de informação especializado.

A denominação linguagem documentária, de acordo com Lara (2004, p. 232), identifica diferentes tipos de instrumentos especializados no tratamento dos conteúdos informacionais, tais como: sistemas de classificação enciclopédicos ou facetados e tesouros. A linguagem documentária é considerada como uma linguagem intermediária, por ser um recurso fundamental e muito utilizado para a representação temática da informação, dado que a sua função comunicativa não ultrapassa os limites do contexto documentário.

Tálamo (1997b) argumenta que “uma linguagem documentária é, simultaneamente, um modo de organização e uma forma de comunicação da informação”. Para realizar a função do modo de organização e, ao mesmo tempo, cumprir o papel de instrumento de comunicação, a linguagem documentária deve: “a) funcionar como código inteligível e fonte para interpretação do sentido, b) caracterizar-se como metalinguagem, c) incorporar o usuário como integrante do processo” (LARA, 2004, p. 232-233).

Sendo assim, a linguagem documentária vem a ser um instrumento que se destina a estabelecer a ligação entre a linguagem construída para o tratamento dos conteúdos informacionais e a linguagem utilizada pelo usuário para a recuperação dos conteúdos no âmbito de um sistema de recuperação da informação (LARA, 2004).

As tipologias das linguagens documentárias podem ser classificadas (ou hierárquicas) e alfabéticas (de indexação ou combinatórias) no que se refere à forma de apresentação dos conceitos e, quanto ao princípio da coordenação, elas

podem ser pré-coordenadas ou pós-coordenadas (GUIMARÃES, 1990, VAN SLYPE, 1991, LANCASTER, 1995). No âmbito das linguagens de estrutura hierárquica, os termos relacionam-se entre si a partir da subordinação e “co-subordinação”, constituindo-se como uma relação assimétrica entre dois elementos, em que um é superior ao outro por caráter normativo, isto é, pela correspondência exata existente entre eles (VIZCAYA ALONSO, 1997).

Dessa maneira, em geral, cada classe do sistema se identifica com um código numérico, alfabético ou alfa-numérico, sendo exemplos de estrutura hierárquica os sistemas de classificação, a saber: Classificação Decimal de Dewey (CDD), Classificação Decimal Universal (CDU), *Library of Congress Classification* (LCC); *Colon Classification* (CC) e outros. Em consonância com as classificações das tipologias documentárias formuladas pelos autores, Guimarães (1990) e Van Slype (1991), preferimos usar a expressão “linguagens documentárias alfabéticas” consoante aos sistemas de organização do conhecimento, utilizados para fins de indexação e recuperação da informação em sistemas de informação de áreas especializadas. Em especial, por tratar-se da tipologia de linguagem documentária a ser estudada nesta tese.

Uma linguagem documentária alfabética é constituída por vocabulário e sintaxe. O vocabulário corresponde à relação dos termos utilizados para a identificação temática dos documentos, e a sintaxe refere-se à combinação dos termos com o objetivo de expressarem o conteúdo dos documentos (FUJITA, 2005).

No âmbito das linguagens documentárias, o vocabulário deve ser entendido como o conjunto de ocorrências provenientes de várias áreas, da terminologia (linguagens de especialidades) e das palavras utilizadas pelos usuários (linguagem natural). Pelo fato de serem linguagens construídas, o vocabulário utilizado para a sua concepção é originário de discursos diferentes, representado por palavras preferenciais, numa relação combinatória de significados e significantes (CINTRA et al., 2002).

As linguagens documentárias são consideradas linguagens pré-coordenadas ou pós-coordenadas. Em geral, constituem-se de duas partes: uma apresenta os termos estruturados de modo sistemático, a outra parte apresenta os termos dispostos em ordem alfabética. Evidenciamos que uma linguagem pré-coordenada é aquela que requer a coordenação dos termos no momento da

representação dos conteúdos documentários, enquanto uma linguagem pós-coordenada possibilita a coordenação dos termos no momento da busca e recuperação da informação.

As linguagens documentárias alfabéticas apresentam os seus termos ordenados alfabeticamente, dentro de uma estrutura hierárquica, de acordo com as características de cada instrumento. São exemplos representativos dessa estrutura as listas de cabeçalhos de assunto e os tesouros.

As listas de cabeçalhos de assunto são linguagens documentárias alfabéticas pré-coordenadas de estrutura associativa ou combinatória de palavras ou expressões, com a finalidade de representar os conteúdos documentários. Possuem regras específicas para as formas de entrada dos cabeçalhos, do uso de abreviaturas e, geralmente, coletam palavras significativas correspondentes de diversas áreas do conhecimento.

O indexador determina quais são os assuntos de um documento, estabelece a ordem do cabeçalho no momento da indexação e procura reuni-los sob formas pelas quais deduz que o usuário irá buscá-los. Nesse caso, o usuário pode ter dificuldades para recuperar a informação se não souber exatamente como esse cabeçalho foi elaborado pelo bibliotecário. Ainda que se estabeleçam relações entre os cabeçalhos, muitas vezes essas relações são indicadas conforme o documento indexado naquele instante, não refletindo a organização conceitual da área do conhecimento.

Nas listas de cabeçalhos de assunto, a representação de seus cabeçalhos indiretos é feita por meio do uso de traço, vírgula e parênteses, sendo as mais utilizadas e conhecidas a *Sears List of Subject Headings* (SEARS) e a *Library of Congress Subject Headings* (LCSH). Ainda que, em sua natureza, seja considerada uma linguagem pré-coordenada, a LCSH apresenta atualmente os seus cabeçalhos dentro de uma estrutura lógico-semântica de relações hierárquicas, minimizando os problemas semânticos que ocorrem advindos das polissemias, sinonímias e homonímias existentes na linguagem natural.

Os cabeçalhos de assuntos podem ter sido a solução para representar os diferentes pontos de vista de um mesmo documento quando se trata de catálogos de assuntos construídos manualmente como ocorria (e ainda ocorre) com a construção dos cabeçalhos de assuntos, mais recentemente denominados como Lista de Autoridades de Assuntos. Um exemplo dessa linguagem é a Lista de

Autoridades de Assuntos, desenvolvida pela Rede de Catalogação Cooperativa - Bibliodata, ou Rede Bibliodata em parceria com os integrantes da Rede.

Há algum tempo, contudo, os Sistemas de Informação utilizam *softwares* específicos para a geração de catálogos automatizados que possibilitam a busca por operadores booleanos *and*, *or*, *not* e permitem utilizar outras linguagens, a saber: as linguagens pós-coordenadas.

As Linguagens documentárias alfabéticas pós-coordenadas são aquelas em que o usuário combina estrategicamente os assuntos no momento da busca de uma informação específica. Um assunto pode ser formado por conceitos, ao se indexar um documento, tendo como princípio a pós-coordenação - o indexador representa separadamente cada conceito. Dessa forma, possibilita-se que o usuário realize múltiplas combinações no momento da busca, como no caso dos tesouros.

O tesouro constitui-se no exemplo mais característico de linguagem documentária alfabética, limita-se, em princípio, ao vocabulário de especialidade. Oportuno salientar que consideramos “especialidade” toda ação humana que exige conhecimento específico. Enfatizarmos que a linguagem documentária alfabética, em destaque, ainda que se aproprie das terminologias e dos sistemas conceituais por ela mapeados, se propõe como modo de organização da informação, articulando temas, subtemas e termos orientados para dar conta de conjuntos documentários e para transferir informação para determinados públicos segundo objetivos específicos (LARA, 2001).

A Linguagem Documentária, contudo, “não se define em relação ao acervo. Não se concebe uma Linguagem Documentária para tratar conjuntos de registros, mas sim para organizar conhecimento” (TÁLAMO, 1997b, p. 10). Com base no exposto, até o momento, em consonância com as argumentações apresentadas por Fujita (1992), Lara (2001) e Tálamo (1997b), julgamos que a construção de uma linguagem documentária alfabética, foco deste estudo, requer observação cuidadosa de normas e princípios norteadores, o cumprimento de procedimentos metodológicos adequados para a sua construção. Além disso, torna-se profícuo observar experiências de construção de autores.

Desse modo, acreditamos que as linguagens documentárias alfabéticas prestam contribuições fundamentais em sistemas de organização do conhecimento registrado em áreas especializadas. Diante disso, consideramos importante investigar, nos próximos itens, a contextualização das linguagens

documentárias alfabéticas em seus diferentes aspectos, em especial o normativo, para aprimoramento de parâmetros à sua construção.

## 2.2 Para entender as Linguagens Documentárias Alfabéticas

Jean Claude Gardin pode ser considerado um dos primeiros pesquisadores a utilizar parâmetros linguísticos para a organização de um modelo de Linguagem Documentária. Valendo-se de texto escrito em 1965 e publicado em artigo de periódico no ano seguinte, apresentou elementos de um modelo para a descrição de léxicos documentários. O autor supracitado incluiu, nesse modelo, os instrumentos para uso documentário e também aqueles que apoiam a construção de Linguagens Documentárias (GARDIN, 1966, p. 171).

Gardin (1973) mencionado em estudo de Lara (1999, p. 52-54), considera que a atividade de representação documentária se desenvolve no âmbito da linguagem. Com isso, oferece grandes contribuições para as pesquisas sobre representação documentária, tornando-se uma das principais referências em estudos sobre o tema. Atribui-se também a Gardin a responsabilidade de por em voga a expressão Análise Documentária na literatura da Documentação (LARA, 1999, p. 54).

Apresentamos, a seguir, um epítome sobre a área de estudo Análise Documentária com o objetivo de destacar o contexto de uso das linguagens documentárias, em âmbito nacional. Na década de 1980, a pesquisadora e professora Johanna W. Smit, iniciou no Brasil uma linha teórica de pesquisa em Análise Documentária - com a criação posterior do Grupo Temma, na Escola de Comunicações e Artes da Universidade de São Paulo - por meio de uma inter-relação entre Documentação, Lógica e Ciências da Linguagem.

As autoras Cunha (1989), Kobashi (1989) e Amaro (1991) evidenciaram que foi principalmente a partir dos estudos de Gardin, nas décadas de 1960 e 1970, que a Análise Documentária deixou de ser um produto implícito do bom-senso para se tornar - ou ao menos buscar sê-lo - uma operação que possa ser explicitada, por meio de uma atividade interdisciplinar.

Seguindo essa tendência, pelos trabalhos do grupo supracitado, ficou manifesto não apenas a necessidade de a Documentação explicitar os



processos inerentes à atividade de análise como também para, ao fazê-lo, recorrer ao aparato teórico de outras áreas do conhecimento, conforme Guimarães (1994).

Nessa direção, a Documentação tem dedicado especial atenção à Representação Documentária, seja no que se refere às chamadas linguagens hierárquicas (ou notacionais) de indexação, neste caso, denominados sistemas de classificação bibliográfica; seja quanto às linguagens alfabéticas, conhecidas como listas de cabeçalhos de assunto e tesouros. Juntam-se a esses, mais recentemente, os bancos de dados terminológicos para a construção automática de tesouros e aos estudos sobre a interface Terminologia e Documentação (GUIMARÃES, 1994).

Para Kobashi (1994, p. 15), Análise Documentária “é uma disciplina de natureza teórica e prática que compõe o domínio conhecido por Biblioteconomia e Documentação, mais modernamente denominado Ciência da Informação”. Na visão de Lara (1993c, p. 4), a Análise Documentária “define-se como uma atividade específica [...] respondendo pela análise e tratamento da informação com o objetivo de recuperá-la e disseminá-la”. Dessa forma, apresenta-se como disciplina de natureza metodológica que propõe processos de organização da informação visando a sua recuperação e comunicação.

A Análise Documentária estabelece relação entre o conceito de representação do resultado das operações de análise e síntese do conteúdo de textos com o objetivo de transferência de informação (LARA, 1999). Nesse sentido, quando se codifica o conteúdo informacional de um documento, esse é representado de maneira condensada para facilitar a circulação de informações. O resultado dessa representação, segundo Kobashi (1994, p. 50), deve manter com o texto original uma relação de similaridade e, da mesma forma, ser equivalente a esse texto do ponto de vista do conteúdo informacional.

Com isso, permite-se que o sentido do texto original, elaborado em linguagem natural e que se aplica a contextos e situações específicas, seja representado por meio de uma Linguagem Documentária. Assim, expõe-se, homologamente, a compreensão científica de estrutura do conhecimento, elaborando metodologias para a análise de assunto e para a geração de sistemas de classificação e linguagens documentárias (NOVELLINO, 1996).

Sob o ponto de vista prático, Kobashi (1994, p. 15-19) ensina que a Análise Documentária é uma atividade realizada dentro de sistemas de informação. Os sistemas de informação, por sua vez, estão inseridos em instituições

informativas que, basicamente, desenvolvem um ciclo de operações documentárias que corresponde ao processo global da circulação de informações e documentos constituídos de: coleta, tratamento e difusão de documentos para uma determinada comunidade usuária.

Cada uma dessas operações desdobra-se em atividades dotadas de política de processo global da circulação de informações, já anteriormente apresentadas e sistematizadas por Carneiro (1985) e retomadas por Rubi (2008) sob as seguintes categorias:

1. Necessidade do usuário: parte-se do princípio geral de que, para cada usuário, é possível fornecer produtos e serviços específicos, os quais resultam, por sua vez, de modos de tratamento igualmente específicos;
2. Domínio tratado: cada instituição tem um perfil próprio que se manifesta, entre outras coisas, nos tipos de materiais que compõem seu acervo. Cada tipo de material deve receber, evidentemente, um tratamento diferenciado, segundo a natureza original do documento e o uso a que se destina;
3. Recursos disponíveis: diferentes tratamentos implicam, por sua vez, em habilidades, competências e recursos igualmente específicos;
4. Produtos e serviços: cada tipo de produto ou serviço requer operações de duração e complexidade variáveis. Desse modo, cada instituição planeja produtos e serviços de acordo com os recursos de que dispõe;
5. Relação entre custo e desempenho: parte-se do princípio de que os recursos devem ser utilizados de modo a satisfazer o maior número de demandas com menor custo.

Nesse caso, um sistema de informação propõe-se, então, a responder de maneira duradoura a um conjunto específico de usuários. Sendo assim, a coleta de documentos deve ser baseada no conhecimento dos organismos que produzem documentos convenientes aos interesses dos usuários do sistema. O tratamento deve ser regido por parâmetros claros e, na difusão os produtos e serviços devem ser planejados e executados de modo a responder às

especificidades da demanda. Como parte do tratamento da informação, a Análise Documentária, sob o ponto de vista metodológico, se define como “um conjunto de procedimentos utilizados para exprimir o conteúdo dos documentos científicos sob formas destinadas a facilitar a sua localização ou consulta” (GARDIN, 1973 apud KOBASHI, 1994, p. 15).

Esses procedimentos caracterizam-se por meio de três operações: 1) Análise - faz-se leitura e segmentação do texto para identificação e seleção de conceitos; 2) Síntese - constrói-se o texto documentário com os conceitos selecionados; 3) Representação - constitui-se nas modalidades de representações condensadas: o resumo e o índice.

Assim, os textos são, primeiramente, analisados e sintetizados. Em seguida, as informações resultantes são submetidas a um processo de combinação, de modo a converterem-se em novos textos, ditos resumos, ou em símbolos de uma linguagem documentária (indexação) (KOBASHI, 1994, p. 23-24). No resumo, o texto base é sintetizado de modo a permitir que certos fragmentos reconstruam potencialmente o sentido do original. Na indexação, procura-se obter um grau ainda maior de compactação do texto, instaurando-se, assim, uma nova ordem de sentido, mediada por um instrumento comutador - a linguagem documentária que generaliza as informações do texto, integrando-o a "classes" já previstas no referido instrumento (KOBASHI, 1994, p. 43).

Desse modo, o conteúdo dos documentos expresso em linguagem natural seria convertido em metalinguagem definida por Gardin (1974 apud KOBASHI, 1989, p. 48), como “um sistema simbólico” que faz a mediação entre o texto e a sua representação a fim de assegurar que indexadores de um mesmo sistema ou sistemas afins possam utilizar os mesmos conceitos para representar e comunicar conteúdos semelhantes.

Vale dizer que se utiliza a expressão metalinguagem, nesse momento, somente para evidenciar uma de suas particularidades, ou seja, salientar “o caráter simbólico do vocabulário organizado para gerar as representações de textos” (VOGEL, 2007, p. 13).

A Análise Documentária procede ao tratamento da informação para que ela possa ser recuperada, permitindo ao Sistema de Informação e ao seu usuário a utilização de um código comum, mediado pela Linguagem Documentária, para que se estabeleça o processo de comunicação documentária.

Garcia Gutierrez (1990a apud TÁLAMO, 2001, p. 145) afirma que “[...] as linguagens documentárias constituem o centro da discussão sobre a análise da informação documentária, pois condicionam efetivamente a qualidade e o valor dos produtos documentários obtidos por tradução”. A autora supracitada argumenta que tais afirmações remetem inevitavelmente à “elaboração de interfaces entre a Linguística Documentária e a Terminologia, pois desse modo será estabelecida a relação complementar entre dois pontos de vista: o do codificador e o do consumidor da informação”. Em razão disso, a autora enfatiza que “cabe à Linguística Documentária estabelecer relação entre mensagens e contextos” (TÁLAMO, 2001, p. 145).

Convém esclarecer que, durante certo período, a Linguagem Documentária foi considerada, ao mesmo tempo, um produto e um instrumento da Análise Documentária. Como produto, porque era elaborada durante processo da Análise Documentária, conforme a seleção dos termos. Ou seja, no momento em que se analisavam documentos para determinação de seus assuntos, selecionavam-se termos desses assuntos para a construção da Linguagem Documentária. Essa Linguagem, posteriormente, seria utilizada como instrumento de representação e de recuperação de novos documentos.

A Análise Documentária compreende fazer uso de Linguagens Documentárias para o tratamento temático das informações, contudo, o desenvolvimento dessas não corresponde, propriamente, a um procedimento de Análise Documentária. Concordamos com Garcia Gutierrez (1990a, p. 24-25) quando afirma que a "Documentação apresenta dois *corpora* que se imbricam com a Lingüística: a Análise Documentária e a Linguística Documentária", mas alerta que não formam uma mesma área. Com respeito à abordagem atual, uma linguagem documentária “[...] deve estruturar, *a priori*, os conceitos da área do conhecimento que irá representar, a fim de que posteriormente possa ser utilizada como instrumento de representação da informação nesta área, efetuando ajustes [...]” à proporção que evolui o conhecimento (GARCIA GUTIERREZ 1990a, p. 96).

Garcia Gutierrez (1990a, p. 96) enfatiza que nem tudo o que é produzido pela Análise Documentária é válido como base léxica da Linguagem Documentária, do mesmo modo, nem todo vocabulário de uma Linguagem Documentária provém da Análise Documentária ou é utilizado diretamente por ela, ainda que, uma vez constituída, a primeira sirva como referente para a segunda.

Dessa forma, entende-se que uma Linguagem Documentária somente cumpre a função de representação da informação quando são instituídas relações entre as unidades informacionais, ou seja, os descritores que a compõem.

Essas unidades informacionais são denominadas de termos preferidos, conforme a Norma documentária NBR 12676 (ABNT, 1992), algumas vezes conhecidas como descritores. Os termos preferidos são estabelecidos pela relação entre termos dispostos no vocabulário de especialidade a que pertencem os conteúdos a serem representados e entre os termos utilizados pelos usuários.

Para a norma ANSI/NISO Z39.19 (2005, p. 18), a expressão termo preferido é considerada para a identificação dos termos autorizados que compõe o tesauro, enquanto o descritor, um termo conhecido, é usado como termo não-preferido. Segundo a Norma NBR 12676 (ABNT, 1992), termo não-preferido é “sinônimo ou quase-sinônimo do termo preferido (descritor); também conhecido como “não-descritor”. O Termo não-preferido não é atribuído aos documentos, mas utilizado como remissiva no índice, para instruir o usuário (Use ou Ver) na busca do termo preferido.

Para a norma ISO 2788 (1986), o termo descritor vem a ser “a representação de um conceito, preferencialmente, em forma de substantivo ou frase nominal”. Além disso, o descritor também “pode ser uma palavra simples, composta ou uma frase que representa o conteúdo do documento” (NAUMIS PEÑA, 1998). A autora apresenta, ainda, uma nova proposta de definição do termo descritor: “um termo normalizado e relacionado no marco de um tesauro, que descreve determinado conteúdo documental e se associa e agrupa com outros conteúdos no sistema de informação de um contexto especializado” (NAUMIS PEÑA, 2000).

Contudo, a ANSI/NISO Z39.19 (2005, p. 176) apresenta em seu glossário a seguinte nota: “*NOTE: In the previous version of this Standard, a preferred term was known as a descriptor*”. Cabe esclarecer que concordamos com o uso do termo descritor, mas, nesse estudo, ao apresentarmos definições e posicionamentos de normas e autores visitados respeitamos a terminologia adotada pelos mesmos.

Tálamo (1997b, p. 6) enfatiza que “a linguagem representa algo, afirma-se que, por força da significação, ela se propõe no lugar de algo, o substitui”. Entendemos que uma linguagem documentária representa a informação como uma consequência das relações que foram instituídas entre as unidades informacionais

no interior dessa mesma linguagem e não como decorrência da mera presença dessas mesmas unidades, tal como afirma Tálamo (1997b).

Por esse motivo, a Linguagem Documentária, ao organizar ou classificar a estrutura conceptual de um campo científico, técnico ou especializado, tem por objetivo compatibilizar a linguagem utilizada na entrada do sistema – quando o documento é analisado para registro e seu conteúdo é identificado e traduzido conforme os termos da linguagem documentária, segundo a política de indexação estabelecida, como a utilizada à saída do sistema – quando a partir da solicitação da informação pelo usuário, é feita a representação para a busca.

Desse modo, sua solicitação é analisada, seu conteúdo é identificado e traduzido nos termos da linguagem documentária utilizada. Assim, com o uso dessa linguagem, pretende-se reduzir ao mínimo o índice de ruído entre a comunicação da informação contida nos documentos e o usuário que dela necessita (GARCIA GUTIERREZ; LUCAS FERNANDEZ, 1987, p. 67; TÁLAMO, 1997b, p. 3; LIMA, 1998, p. 24). Em outras palavras, uma linguagem documentária deve organizar os conteúdos segundo os interesses de seus usuários a fim de lhes possibilitar a comunicação com o sistema de informação.

A partir desse reconhecimento, Garcia Gutierrez (1990a) propôs a criação de um subdomínio da Documentação ao qual denominou Linguística Documentária que compreenderia o estudo dos meios de representação da informação com foco nas linguagens de processamento e produção para fins de circulação do conteúdo informacional.

Do mesmo modo, Lara e Tálamo (2006) afirmam que a expressão Linguística Documentária foi inicialmente utilizada (GARCIA GUTIERREZ, 1990a), para sugerir uma disciplina que, partindo do reconhecimento de que os problemas relacionados à informação são problemas de linguagem, busca, de um modo geral, seus fundamentos nas ciências da linguagem. Recorre, tanto, à linguística, à semântica, à gramática aplicada, à gestão da informação como aos campos afins: a análise do discurso, a análise do conteúdo e, de modo geral, as ciências cognitivas, orienta Garcia Gutierrez (1998) evocado por Lara e Tálamo (2006).

Ao estabelecerem que a Linguística Documentária responde pelo desenvolvimento de parâmetros para a elaboração das Linguagens Documentárias, Lara e Tálamo (2006) fundamentam-se na Linguística Estrutural, na Semiótica, na Terminologia e na Lógica Formal. As autoras afirmam que o

campo da Linguística Documentária constitui um subdomínio da Ciência da Informação que se preocupa em agregar os problemas decorrentes dos processos simbólicos do tratamento e da recuperação da informação, em pesquisá-los e buscar soluções que minimizem a distância entre os estoques e o uso da informação (LARA; TÁLAMO, 2006, p. 207).

A seguir, discutimos os diferentes aspectos da linguagem documentária, porquanto é a partir dela que o subdomínio Linguística Documentária se organiza (LARA; TÁLAMO, 2006). As denominadas linguagens documentárias caracterizam-se como instrumentos que possibilitam uma intermediação entre textos e usuários. A esse respeito, Gardin et al. (1968 apud CINTRA et al., 2002) esclarecem que a linguagem documentária constitui um sistema de representação do conteúdo de documentos técnico-científicos com finalidade de classificação ou a busca retrospectiva de informações. Deve integrar três elementos básicos:

[...] um léxico - identificado como uma lista de elementos descritores, devidamente filtrados e depurados; uma rede paradigmática - para traduzir certas relações essenciais e, geralmente estáveis, entre descritores. Essa rede lógico-semântica corresponde à organização dos descritores numa forma que, *lato sensu*, poder-se-ia chamar classificação; e uma rede sintagmática - destinada a expressar as relações contingentes entre os descritores, relações que são válidas no contexto particular onde aparecem. A construção de 'sintagmas' é feita por meio de regras sintáticas destinadas a coordenar os termos que dão conta do tema (GARDIN et al., 1968 apud CINTRA et al., 2002, p. 35).

Entendemos o léxico como o conjunto de palavras de um idioma ou área de especialidade, ou seja, o conjunto de vocábulos da língua de um dado contexto. Por sua vez, uma rede paradigmática conhecida, genericamente, como relações associativas, refere-se a uma padronização da linguagem utilizada como forma de manter uma relação entre as palavras cujo significado é de senso comum entre os especialistas da área. Já uma rede sintagmática refere-se às relações que podem ser determinadas entre os termos, isto é, estabelecem quais as unidades que os constituem e qual a ordem em que sucedem, ou, dito de outra forma, o que se segue e o que precede cada unidade constitutiva.

Para Cintra et al. (2002, p. 49), as áreas especializadas da experiência humana “devem ter seu universo nocional devidamente identificado a partir de um dado ponto de vista, para que seja possível organizá-la de forma

sistemática [...]”. A organização dos conceitos de uma área possibilita o uso de instrumentos eficazes para o tratamento e recuperação da informação. O conceito, segundo a Norma ISO 1087-1, é “[...] unidade de conhecimento constituída por abstração, com base em um conjunto de traços ou características comuns, atribuídas a uma classe de objetos, de relações ou de entidades”.

Para a Norma NBR 12676 (ABNT, 1992, p. 1), o conceito é “qualquer unidade de pensamento. O conceito pode ter o seu conteúdo semântico reexpresso pela combinação de outros conceitos, que podem variar de uma língua ou de uma cultura para outra”. O sistema nocional de uma área do conhecimento constitui-se em arcabouço fundamental para a construção de tesouros, à medida que possibilita a materialização das relações entre as noções (CINTRA et al., 2002, p. 51).

As noções expressas por meio dos termos e símbolos são influenciadas pelo contexto sociocultural. Uma linguagem focada em um contexto, ou seja, em uma área especializada, naturalmente elimina as possíveis diferenças de significado de uma palavra. Na visão de Cintra et al. (2002, p. 71), “[...] não podem existir, por exemplo, duas ou mais palavras que se refiram a um mesmo conceito ou uma palavra para designar vários conceitos”.

Nesse caso, destacam-se alguns fenômenos da Linguística, reconhecidos pela Linguagem Documentária no controle de vocabulário para tratamento e recuperação da informação: a polissemia, que decorre ou da organização sintático-semântica com a finalidade de distinguir a plurissignificação como fenômeno geral ou da ambiguidade, que ocorre quando há mais de uma interpretação em razão da plurissignificação que, como a polissemia, poderá também levar à ambiguidade (CINTRA et al., 2002, p. 71-72).

Em sistemas de informação, a plurissignificação não é permitida, porque a presença de termos com vários significados ou várias interpretações contribui para o aumento do ruído e lacunas na recuperação das informações, reduzindo-se o nível de eficiência do sistema. Nessa direção, Cintra et al. (2002, p. 72-74) afirmam que “[...] numa linguagem documentária, tanto a polissemia quanto a ambigüidade devem ser neutralizadas, para que seja garantida a monossemia [...]. A monossemia é obtida por meio de redes relacionais e definições dos termos”. As autoras esclarecem “não se deseja que um termo se enriqueça. Exige-se que ele expresse conceitos determinados”. Para neutralizar a polissemia, as autoras



recomendam a “elaboração de redes relacionais e estabelecimento de definições e nota de escopo, sempre que as redes se mostrarem insuficientes para a interpretação unívoca da significação”.

Nos tesouros são reconhecidas “três classes de relações básicas entre os termos: a relação hierárquica; a relação associativa; a relação de equivalência” (UNESCO, 1993, p. 41). Conforme Cintra et al. (2002, p. 53), as *relações hierárquicas* revelam os níveis de superordenação e subordinação entre as noções, e elas podem ser *genéricas*, *específicas* ou *partitivas* e constituem a estrutura principal de um tesouro no sentido vertical (esquema de árvore).

As árvores semânticas são, em geral, elaboradas segundo as cadeias conceptuais que partem dos termos mais gerais ou amplos, em direção aos termos mais específicos ou restritos. Na *relação genérica*, a superordenação avança das diferenças para as semelhanças, ou seja, da espécie para o gênero, enquanto que, na *relação específica*, a subordinação, no sentido inverso, avança das semelhanças para as diferenças, ou seja, do gênero para a espécie.

Já na *relação partitiva*, segundo Cintra et al. (2002, p. 53), que é um tipo de relação hierárquica, a noção de superordenação refere-se a um objeto considerado como um todo e as noções subordinadas a objetos considerados como suas partes. Por exemplo: em relação a *navio*, a noção de *casco* é uma noção específica partitiva e denota que *navio* é uma noção que se refere ao todo e que *casco* é uma noção que se refere à parte.

Já as *relações associativas* são estabelecidas pela dependência entre conceitos, sem, no entanto, estabelecer uma superordenação. As relações associativas podem ser divididas em dois grupos: *relação sequencial*, que é a relação de dependência estabelecida entre conceitos que se referem a objetos que têm contigüidade espaço-temporal como, por exemplo: “causa-efeito”, “produtor-produto”, “etapas de um processo”; e *relação pragmática* que é a relação entre conceitos que podem ser estabelecidas por conexões temáticas.

Para Cintra et al. (2002, p. 54), “as relações que não se submetem a uma hierarquia são aquelas que apresentam entre si contigüidade espacial ou temporal entre termos que não possuem relacionamento hierárquico e nem de equivalência”. Por este motivo, as *relações de equivalência* são utilizadas em uma linguagem documentária para o controle do vocabulário. Essas relações operam no nível de sinonímia e da polissemia, estabelecem as remissivas com a finalidade de

encaminhar o usuário para os termos preferidos pelo sistema. Na linguagem natural, vários termos, palavras ou expressões podem designar uma única ideia ou conceito. Já na Linguagem Documentária apenas uma denominação é selecionada para figurar como descritor. Os demais termos são considerados como não-descritores, fazendo-se destes, remissivas para aquele (GOMES, 1990, p. 47).

Concordamos com as autoras supracitadas e evidenciamos que a fixação de rede de relações, a nota de escopo ou de aplicação compreendem o princípio organizador essencial e básico na construção de tesouros. Em razão disso, a sinonímia é uma relação de equivalência, que trata da relação entre designações de uma mesma língua de especialidade que representam o mesmo conceito. Os sinônimos, na construção de tesouros, permitem “estabelecer as relações de equivalência entre termos preferidos e não-preferidos, funcionando como meio de controle da variação de significado e como recurso para o acesso à linguagem do sistema” (LARA, 1999, p. 105).

No caso da sinonímia utilizada na construção de tesouros, consideramos que são sinônimos quando dois termos carregam a possibilidade funcional de se substituírem um ao outro. Cintra et al. (2002, p. 77) complementam que pode “compreender tanto a sinonímia absoluta como a quase-sinonímia” e esclarecem ainda que a sinonímia nas linguagens documentárias “é de caráter eminentemente preferencial e visa [a] remeter o usuário de um termo não preferencial, para um termo selecionado [...]” (CINTRA et al., 2002, p. 77)

Há ainda a hiponímia que, segundo Lyons (1987, p. 116), é “[...] a relação existente entre um lexema mais específico e um mais geral (entre ‘tulipa’, ‘rosa’ etc., e ‘flor’; entre ‘honestidade’, ‘castidade’ etc., e ‘virtude’)”. Dito de outro modo, a categoria denominada hiponímia opera com a noção de inclusão, a mesma noção permite reunir unidades numa classe. A hiponímia refere-se aos “termos mais específicos abrangidos por um termo mais geral, por exemplo: rosa e tulipa, são hipônimas de flor” (LYONS, 1976, p. 309). A relação de hiponímia/hiperonímia ou (subordinação/superordenação) permite verificar que um termo pertence ou subordina-se a um outro mais geral, o gênero, mas não permite identificar em que os termos subordinados se diferenciam entre si.

Além disso, a relação de hiponímia representa uma operação de conjunção em face do termo superordenado e de disjunção, tomando-se a série de termos obtidos a partir da divisão realizada. A relação de hiponímia situada pela

linguística “permite explicar, nas LDs, vários tipos de relacionamentos tomados como hierárquicos que não cabem dentro da classificação gênero/espécie (e tão pouco nas relações todo/parte, parte/parte)” (CINTRA et al., 2002, p. 80).

De acordo com as Diretrizes da UNESCO (1993, p. 25), apresentamos uma lista de abreviaturas e símbolos que são utilizadas em tesouros como prefixos dos termos. Cada abreviatura e símbolo apresentada em português com uma explicação sobre seu significado, além de seus equivalentes em inglês e em francês, é usada para indicar a relação ou função do termo ou nota que se segue, conforme explicitamos no Quadro 1.

PORTUGUÊS	INGLÊS	FRANCÊS
<b>TG = Termo Genérico</b> O termo que segue refere-se a um conceito com conotação mais ampla. Superordenado.	<b>BT = Broader Term</b>	<b>TG = Terme Générique</b>
<b>TGM = Termo Genérico Maior</b> O termo que segue é o nome da classe mais ampla à qual pertence o conceito específico usado também, às vezes, na seção alfabética de um tesouro.	<b>BTG = Broader Term (Generic)</b>	-----
<b>TGP = Termo Genérico Partitivo</b> O termo que segue representa o todo em relação à parte.	<b>BTP = Broader Term (Partitive)</b>	<b>TGP = Terme Générique Partitive</b>
<b>TE = Termo Específico</b> O termo que segue refere-se a um conceito com conotação mais específica. Subordinado.	<b>NT = Narrower Term</b>	<b>TS = Terme Spécifique</b>
<b>TEP = Termo Específico Partitivo</b> O termo que segue representa a parte em relação ao todo..	<b>NTP = Narrower Term (Partitive)</b>	<b>TSP = Terme Spécifique Partitive</b>
<b>TR = Termo Relacionado</b> O termo que segue está associado, mas não é nem sinônimo, nem termo genérico ou termo específico.	<b>RT = Related Term</b>	<b>VA = Voir Aussi</b>
<b>NE = Nota Explicativa (ou Nota de Escopo)</b> Nota que se junta a um termo para indicar seu significado específico dentro de uma linguagem de indexação.	<b>SN = Scope Note</b>	<b>NE = Note Explicative</b>
<b>UP = Usado Para</b> O termo que segue é um sinônimo ou um quase-sinônimo do termo preferido.	<b>UF = Used For</b>	<b>EM = Employé pour</b>
<b>USE</b> O termo que segue é o termo preferido quando se deve escolher entre sinônimos ou quase-sinônimos.	<b>USE</b>	<b>EM = Employer</b>

**Quadro 1** - Abreviaturas utilizadas em tesouros como símbolos dos termos e seus equivalentes em inglês e francês.

As abreviaturas acima listadas são convenções reconhecidas que aparecem em diversos tesouros publicados e possuem valor mnemônico óbvio, embora se reconheça que também sejam dependentes do idioma (UNESCO, 1993,

p. 79). Em suma, ancorados nessa Norma podemos dizer que as relações básicas de um tesouro - relações de equivalência, relação hierárquica e relação associativa - expressam-se por meio das abreviaturas e símbolos evidenciadas a seguir:

RELAÇÕES DE EQUIVALÊNCIA	RELAÇÃO HIERÁRQUICA	RELAÇÃO ASSOCIATIVA
<b>USE</b> = precede o termo preferido	<b>TGM</b> = Termo genérico Maior de uma hierarquia	<b>TR</b> = Termo Relacionado
<b>UP</b> = precede termo não-preferido	<b>TG</b> = Termo Genérico	
	<b>TGP</b> = Termo Genérico Partitivo	
	<b>TE</b> = Termo Específico	
	<b>TEG</b> = Termo Específico Genérico	
	<b>TEP</b> = Termo Específico Partitivo	

**Quadro 2** - Relações básicas de um tesouro e sua simbolização.

Importante salientar que as Diretrizes da UNESCO (1993, p. 16-17) orientam as seguintes convenções para a escrita de termos preferidos e não-preferidos: Os termos preferidos são impressos em maiúscula ou em negrito.

Por exemplo:

**CARROS**

**ANIMAIS**

Os termos não-preferidos são impressos com a letra inicial em maiúscula, exceto quando se referir a uma abreviatura ou sigla que deve ser descrita totalmente em maiúscula. Por exemplo:

**Animais**

**UP:** Fauna

**Carros**

**UP:** Automóveis

Cabe lembrar que é necessário definir quantas hierarquias deverão ser organizadas para a apresentação dos termos. Em alguns casos, é possível reunir todos os termos relativos a um determinado assunto, em uma hierarquia. Já em outros, pode haver necessidade de várias hierarquias. Por exemplo:

### **Produtos agrícolas**

Tipos Ou Por propriedades

Subtipos Por tipo de solo requerido

Para as relações associativas, devemos observar a natureza daquelas que são frequentes no domínio, para apresentá-las de forma consistente em todo o tesouro. Por exemplo:

### **Pássaros**

TR: Ornitologia

### **Ornitologia**

TR: Pássaros

Notas de escopo ou de aplicação - Apresentam instruções de uso de alguns termos preferidos, explicitando como devem ser aplicados. Por meio dessas notas, pode-se ampliar ou restringir o significado de um descritor, conforme a política de indexação utilizada.

Tálamo (1997c, p. 12) afirma que as notas de escopo permitem um controle efetivo de todas as operações documentárias relativas ao tesouro. Podem incluir data de inclusão/exclusão de descritores, documentando todo o processo. Por exemplo:

### **Países do terceiro mundo**

**NE** Indexa-se também pelos nomes dos países.

### **Aditivo**

**NE** Substâncias adicionadas a produtos ou processos em proporções relativamente pequenas para melhorar propriedades desejáveis e excluir propriedades indesejáveis.

Apresentação do tesouro – Utiliza-se, com frequência, a apresentação sistemática acompanhada de apresentação alfabética. Na parte sistemática, os termos são relacionados conforme as categorias ou classes e apresentados sob a forma de listas das classes em ordem alfabética.

Na apresentação alfabética, todos os termos de indexação, preferidos ou não-preferidos, são organizados em uma só sequência alfabética. Os termos não-preferidos, em geral, são acompanhados apenas pela indicação USE, referindo-se aos equivalentes preferidos. Por exemplo:

OMS

**USE Organização Mundial da Saúde**

Já o termo preferido aparece com aqueles aos quais se encontra relacionado. Para as Diretrizes da UNESCO (1993, p. 60), a informação auxiliar que acompanha os termos preferidos deve ser disposta na seguinte ordem:

1. Notas explicativas **NE**;
2. Referências **UP** a termos equivalentes não-preferidos;
3. Referências **TGM** a termos genéricos maiores, se necessário;
4. Referências **TG** a termos genéricos;
5. Referências **TE** a termos específicos;
6. Referências **TR** a termos relacionados.

De acordo com as recomendações da Norma supracitada, a seguir apresentamos uma simulação com o termo preferido “Bibliotecas”, conforme Lancaster (1987, p. 71).

**BIBLIOTECAS UNIVERSITÁRIAS**

**TG** Bibliotecas

**TE** Bibliotecas de faculdade

Bibliotecas de pós-graduação

**TR** Educação superior

As diferentes séries formadas a partir de um mesmo termo podem ser vistas como um dos resultados de diferentes modos de realizar a conjunção, procedente de diferentes pontos tomados como origem da subdivisão e/ou das diferentes características observadas para a construção de cada hierarquia (CINTRA et al., 2002, p. 83). Significa dizer que é a inclusão do significado de um termo no significado de outro, em outras palavras, pode ser uma relação entre espécie e gênero que constitui uma classe. Com isso, evidencia-se que a informação tem a sua existência atrelada aos sistemas de significado e que a operação, nesse universo, é fundamental para a sua identificação, análise, tratamento e disseminação.

### 2.3 Tesouros: uma perspectiva diacrônica de parâmetros e modelos para sua construção

No âmbito biblioteconômico, encontramos o termo tesouro representado com diferentes grafias *thesaurus*, *tesaurus*, *thesauri*, *thesauros* e *tesauros*. O vocábulo tesouro, em sua origem *thesaurus* ou *thesauros*, era o nome dado à *sala onde se guardava o tesouro*. Possui raiz semântica grega como um – *bem material ou de conhecimento*. Posteriormente, foi usado para designar certas obras como dicionários com a mesma conotação de lugar onde se guarda o mais importante que, neste caso, é menos corpóreo porque se trata do saber humano.

Historicamente, o termo tesouro está associado a obras de estruturas diversificadas, como dicionários, índices, repertórios de termos ou vocabulários. Entre 1262-1268 - no *Livre du Trésor* - Florentino Brunetto Latini usou pela primeira vez o vocábulo nessa Enciclopédia composta em francês (época em que se começava a escrever em línguas vulgares).

No Séc. XV, o termo aparece no *Thesaurus adversus haereticos de Cirilo de Alejandria*, cópia do original (s.d). Ainda no Séc. XV, durante o período do Renascimento, aparecem os dicionários de uma só língua: *Thesaurus linguae latinae* - Robert Estienne, 1532, e *Thesaurus graecae linguae* - Henri Estienne, 1572 (NAUMIS PEÑA, 2007).

Já no Séc. XIX, especificamente no ano de 1852, o termo se popularizou a partir do dicionário analógico *Thesaurus of english words and phrases*, de Peter Mark Roget, publicado inicialmente em Londres, que, na atualidade, soma 90 reedições, aproximadamente. A obra de Roget é considerada um marco teórico-metodológico em razão de que as palavras não foram agrupadas segundo a ordem alfabética, tal como ocorre com os dicionários linguísticos, mas “de acordo com as idéias que elas exprimem. [...] O que se almeja [...] – Tendo-se a idéia, encontrar a palavra, ou as palavras, pelas quais a idéia possa ser expressa [...] de acordo com seu significado” (ROGET, [1852] apud GOMES, 1990, p. 13). A originalidade de Roget, quanto à forma de apresentação de seu dicionário, ensejou que a palavra ‘thesaurus’ ficasse, na área da documentação, associada à forma de organização do vocabulário de indexação/recuperação (GOMES, 1990).

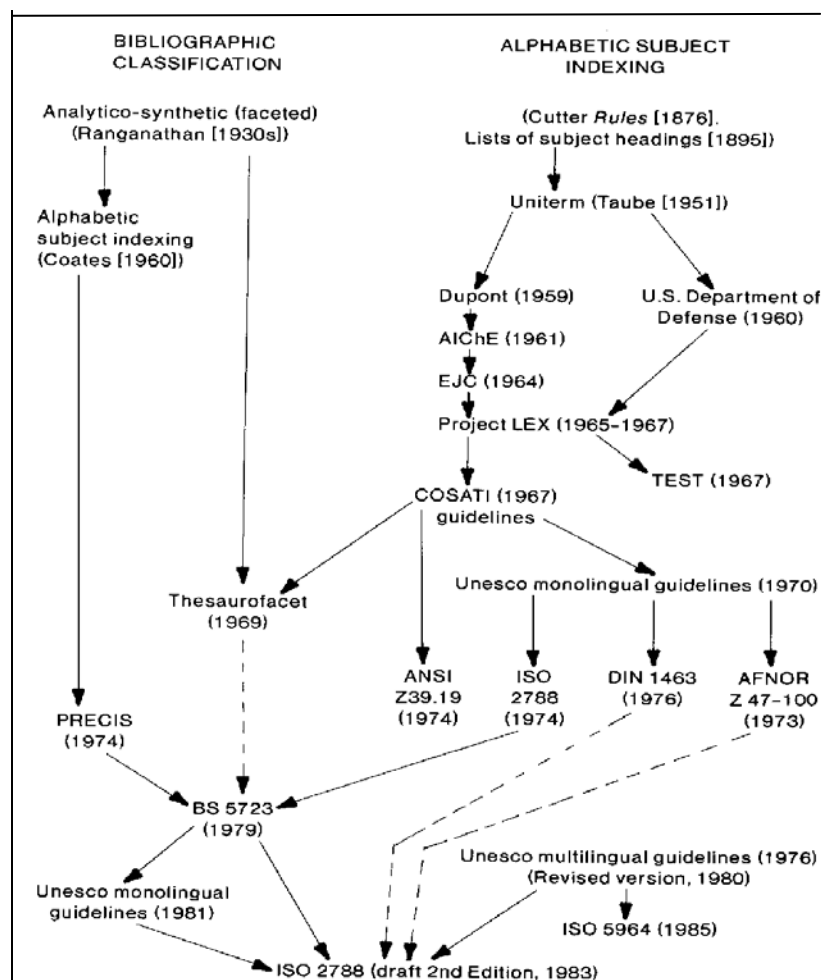
Em 1957, o termo tesouro foi utilizado como um tipo de instrumento linguístico na indexação dos documentos. Há certa divergência quanto à sua autoria: atribui-se a Helen Brownson em apresentação em Conferência, defendida por Vickery (1997). Todavia, credita-se a autoria para Hans Peter Luhn em texto publicado no mesmo ano, segundo Lancaster (1995).

Como o foco desta pesquisa refere-se ao processo de construção de tesouros, buscamos referenciais que tratam de métodos para a construção de tesouros com o objetivo de ter uma ideia sobre a origem e traços da evolução histórica dos tesouros.

Em contato com publicações de Lancaster (1986; 1987; 1995) observamos que essa evolução manifesta-se sob duas linhas de influência: uma, que toma por base a abordagem alfabética de assuntos provenientes da América do Norte, mais especificamente os Estados Unidos, e a outra que se orienta pela classificação bibliográfica da Europa, mais particularmente o Reino Unido.

No esboço que apresentamos a seguir, visualiza-se a evolução das normas sobre tesouros e as possíveis influências das abordagens: alfabética de assuntos e classificação bibliográfica, refletidas na sistematização realizada por Lancaster (1986; 1995), separando-as da seguinte forma: a) As possíveis influências a partir da Indexação Alfabética de Assuntos e b) As possíveis influências a partir da Classificação Bibliográfica.





Fonte: Lancaster (1995, p. 50; 1986 p. 30).

**Figura 3** - Evolução das Normas e Padrões sobre Tesauros.

### 2.3.1 Possíveis influências a partir da Indexação Alfabética de Assuntos

Inicia-se em 1876, com Charles Ammi Cutter, considerado o primeiro a propor normas para a construção de cabeçalhos alfabéticos de assuntos com a elaboração de *Rules for a Dictionary Catalogue*. No ano de 1895, surge o primeiro vocabulário controlado para catálogos alfabéticos de assuntos, com a publicação da *List of Subject Headings for Use in Dictionary Catalogs de La American Library Association* (LANCASTER, 1995, p. 49).

No Séc. XX, mais precisamente em 1951, o sistema Unitermo, introduzido por Mortimer Taube, pode ser considerado muito relacionado ao aparecimento do tesauro. A principal característica desse sistema é que representa os assuntos com palavras simples, extraídas do texto dos documentos, sem nenhum

tipo de controle. Em razão da simplicidade desse procedimento, as ideias de Taube foram aceitas e aplicadas por importantes organizações governamentais e industriais dos Estados Unidos nos anos de 1950.

No início, foram utilizados cartões escritos à mão ou à máquina, posteriormente, utilizaram-se sistemas de cartões perfurados, e mais tarde, tornou-se necessária a aplicação de computadores em razão do grande número de termos manipulados para recuperação dos assuntos. Por ser um sistema sem nenhum controle de vocabulário e com base em termos simples, o primeiro tesouro foi utilizado para estabelecer um controle terminológico em um sistema Unitermo.

Concordamos com Lancaster (1995, p. 51) quando reforça que as contribuições de Mortimer Taube podem ter influenciado na redação das primeiras normas sobre construção de tesouros. O primeiro tesouro com a função de controlar o vocabulário em sistemas de recuperação da informação foi elaborado, em 1959, por *Engineering Information Center* de E. I. Dupont de Nemours nos Estados Unidos (CHAUMIER, 1986, p. 12). O principal centro de informação do Departamento de Defesa dos Estados Unidos, atualmente denominado *Armed Services Technical Information Agency* (ASTIA), elaborou seu primeiro tesouro em 1960.

Já em 1961, o *American Institute of Chemical Engineers* (AIChE) publicou o *Chemical Engineering Thesaurus* que se fundamentava no trabalho de Dupont e tornou-se o primeiro tesouro de ampla difusão pública. Em 1964, surgiu o *Thesaurus of Engineering Terms*, publicado pelo *Engineering Joint Council* (EJC), fundamentado no tesouro da AIChE, mas com cobertura de toda a área de Engenharia. Com isso, observamos que as primeiras recomendações para a construção de tesouros foram estabelecidas tendo como base os trabalhos: Dupont de 1959, AIChE de 1961 e EJC de 1964 (LANCASTER, 1995, p. 50).

A realização de trabalho em conjunto do Projeto LEX entre os anos 1965 a 1967, desenvolvido pelos segmentos: Defesa e Indústria respectivamente ASTIA e EJC, resultou na publicação do *Thesaurus of Engineering and Scientific Terms* (TEST), em 1967. As diretrizes sobre construção de tesouros emanadas do Projeto LEX foram aprovadas e publicadas, ainda em 1967, pelo *Committee on Scientific and Technical Information* (COSATI), um comitê oficial ligado ao Conselho Federal para Ciência e Tecnologia dos Estados Unidos.

Como recomendação a todas as agências governamentais dos Estados Unidos a adoção das diretrizes COSATI para a construção de tesouros

(LANCASTER, 1995, p. 51). A norma nacional americana para a construção de tesouros, ANSI Z39.19 *American National Standards Institute* de 1974, pode ser considerada como um dos resultados das experiências obtidas nos projetos anteriormente citados.

Além disso, concordando com Lancaster (1995), podemos considerar como outro resultado dessa experiência a primeira versão das *Guidelines for the Establishment and Development of Monolingual Scientific and Technical Thesauri* da UNESCO, em 1970, em razão de ser muito semelhante e compatível com a norma ANSI Z39.19-1974.

As recomendações da UNESCO constituíram a base para a primeira normativa internacional para a construção de tesouros da *International Organization for Standardization* (ISO 2788), publicada em 1974, assim como para as normas francesas da *Association Française de Normalization* (AFNOR Z47-100), em 1973, e a alemã *Deutsches Institut für Normung* (DIN 1463), de 1976. A norma espanhola surgiu em 1990, argumenta Alejandro de la Cueva Martín em nota de tradução na obra de Lancaster (1995, p. 51).

### 2.3.2 Possíveis influências a partir da Classificação Bibliográfica

Inicia-se, em 1930, com o método de classificação analítico-sintético (por facetas) idealizado por Shiyali Ramamrita Ranganathan. Na perspectiva do autor supracitado, existem em todo universo de assuntos cinco idéias fundamentais que são utilizadas para a divisão do Universo, a saber: *Tempo, Espaço, Energia, Matéria e Personalidade*, essas idéias, representadas pela sigla, inglesa, *PMEST*, são denominadas: Categorias Fundamentais. Salientamos que desde então, são esses princípios, valiosos para uma série de esquemas de classificação especializadas, em diversas áreas temáticas, produzindo efeito na indexação alfabética de assuntos com o trabalho de E. J. Coates em 1960.

As influências levam ao surgimento do Sistema PRECIS, desenvolvido por Derek Austin, em 1974. A confluência da classificação por facetas e do tesouro convencional cederam espaço ao tesouro com facetas (*Thesaurofacet*), desenvolvido pelo *Classification Research Group* (CRG), em 1969, com o propósito

de estabelecer um instrumento que combinasse o melhor de ambos os procedimentos.

O tesouro facetado, publicado em 1969 por Jean Aitchinson, mostra as vantagens de indexar e classificar os assuntos com uma correspondência entre facetas e descritores. O levantamento realizado por Lancaster (1986) tornou evidente a contribuição de todas essas fontes na elaboração das normas britânicas sobre construção de tesouros, *British Standards* (BS 5723), publicadas em 1979.

Ainda que PRECIS e ISO consistissem as normas que mais as influíram, Lancaster (1995) argumenta que a experiência e as normas norte-americanas aliadas à parte da filosofia e à experiência do trabalho britânico, sobre indexação alfabética de assuntos, foram importantes no surgimento da segunda edição das Diretrizes monolíngues Unesco em 1981.

Lancaster (1986), enfatiza, que mesmo seguindo por um caminho diferente, as recomendações da Unesco presentes nas Diretrizes multilíngues Unesco, publicada em 1976, e edição revista em 1980, cederam lugar à norma ISO 5964, publicada em 1985. Além disso, podem ter influenciado as recomendações da Unesco para tesouros monolíngues e multilíngues, presentes na norma ISO 2788 para tesouros monolíngues, publicada em 1983 (ISO 2788, 1986).

Na sistematização de seu estudo, Lancaster (1986) expõe sua visão sobre o processo de evolução das normas de construção de tesouros. Nele podemos observar as duas importantes influências explicitadas pelo autor, ou seja, a Indexação Alfabética de Assuntos e a Classificação Bibliográfica.

Lancaster (1995) aponta para a dificuldade de expor essa evolução de uma forma precisa, argumentando que não há clareza sobre quem influenciou quem. Nesse sentido, alerta que nem todos os passos traçados, na figura 3, podem ser considerados definitivos. Esclarece, ainda, que as linhas tracejadas que aparecem no referido quadro indicam pontos em que há poucas influências.

No Brasil, em 1984, as Diretrizes para elaboração de tesouros monolíngues publicada pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) define o tesouro como “vocabulário controlado e dinâmico abrangendo área específica do conhecimento”. Em complemento, evidencia que “sua estrutura patenteia as relações, vigentes entre os termos ou descritores, sinonímicas, hierárquicas e outras, que, no conjunto, constitui a linguagem de indexação” (IBICT, 1984, p. 5).

No ano de 1986, as *Guidelines to establishment and development of monolingual thesauri* (ISO, 2788), da *International Standard Organization* (ISO), apresentam o estabelecimento de normas para a construção de tesouro. A norma ISO 2788 (1986) - define o tesouro no âmbito das linguagens de indexação como: “vocabulário de uma linguagem de indexação controlado e organizado formalmente com objetivo de explicitar as relações *a priori* entre conceitos (por exemplo, mais genérico que... ou mais específico que...)”.

Para isso, descreve a linguagem de indexação como: “conjunto controlado de termos extraídos da linguagem natural e utilizados para representar de forma breve os assuntos dos documentos” (ISO 2788, 1986). Para Van Slype (1991), o tesouro é “uma lista estruturada de conceitos destinados a representar de maneira unívoca o conteúdo dos documentos e das consultas dentro de um sistema documental determinado [...]”. O autor complementa que o tesouro “inclui descritores, não-descritores, relações hierárquicas e de associação e eqüivalências lingüísticas” (VAN SLYPE, 1991, p. 23-24).

As Diretrizes da Unesco (1993, p. 14) definem Tesouro como: “O vocabulário controlado de uma linguagem de indexação, formalmente organizado para explicitar as relações *a priori* entre conceitos” (por exemplo, como genéricas e específicas). Currás (1998) concorda com Van Slype (1991) e acrescenta que o tesouro é “um vocabulário especializado, normalizado, pós-coordenado, usado com fins documentários, onde os elementos lingüísticos que o compõem, termos simples ou compostos, se encontram relacionados entre si sintática e semanticamente”.

As normas atuais da ANSI/NISO Z39.19 e da BSI 5723, sinalizam para a importância da construção de vocabulários controlados. No que se refere a vocabulário controlado que apresenta características multidisciplinares, que corresponde ao conjunto de áreas especializadas de uma instituição, enfatizamos como modelo bem sucedido de construção o Vocabulário Controlado do Sistema de Bibliotecas da Universidade de São Paulo, SIBi-USP (Lima et al., 2006).

No início do Séc. XXI – mais precisamente em 25 de julho de 2005 - a *American National Standards Institute* divulgou a norma Z39.19 - *Guidelines for the construction, format, and management of monolingual controlled vocabularies* – proposta desenvolvida pela organização norte-americana *National Information Standards Organization*. A ANSI/NISO Z39.19 (2005) propõe linhas gerais para a construção, formatação e manutenção de vocabulários controlados monolíngues.

Nesse sentido, a norma supracitada exprime que “um vocabulário controlado é uma lista de termos organizados explicitamente. Os termos não podem ser ambíguos e redundantes e devem ser controlados por um registro de autoridade”.

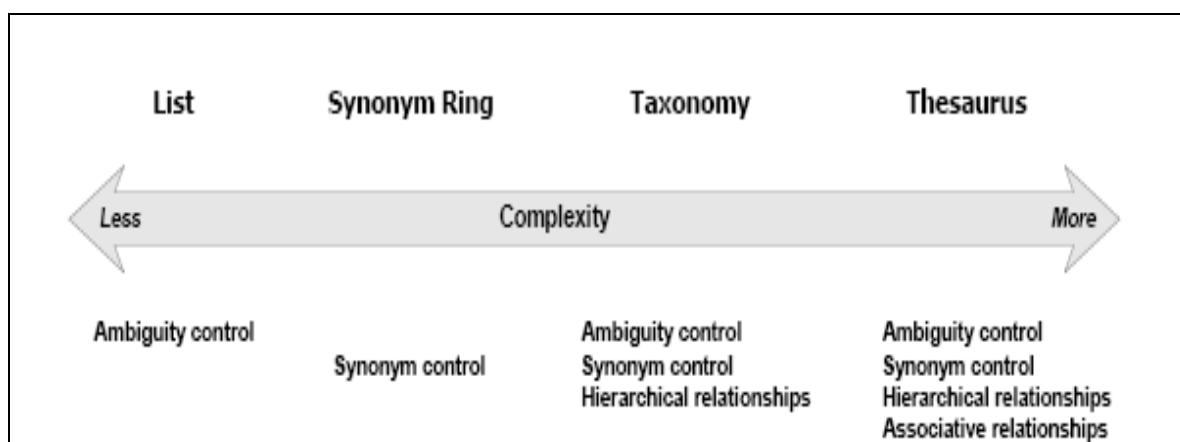
A Norma ANSI/NISO Z39.19 (2005, p. 5) define, também, o tesouro como “um vocabulário controlado organizado em uma ordem conhecida e estruturada de modo que os vários relacionamentos entre os termos sejam identificados e indicados claramente por meio de orientações normativas” (ANSI/NISO Z39. 19, 2005, p. 9).

Para melhor visualização, elaboramos no quadro 3, conforme explicitado na norma supracitada, tipologias de vocabulário controlado como: listas, redes de sinônimos, taxonomias, tesouros e suas respectivas aplicações em áreas de especialidade:

<b>TIPOLOGIAS DE VOCABULÁRIOS CONTROLADOS E SUAS APLICAÇÕES EM ÁREAS DE ESPECIALIDADE</b>	
<b>TIPOLOGIAS</b>	<b>APLICAÇÕES</b>
<b>Listas</b>	Listas são usadas para exibir pequenos conjuntos de termos que são utilizados para finalidades estritamente definidas, como uma lista de opções ou uma lista de comandos para baixar conteúdos da Web.
<b>Redes de sinônimos</b>	Redes de sinônimos são com frequência utilizadas como “pano de fundo” para melhorar a recuperação, especialmente em um ambiente em que a indexação não utiliza um controle de vocabulário e/ou não há indexação para busca por texto completo.
<b>Taxonomias</b>	Taxonomias são criadas e utilizadas, muitas vezes, em aplicações de indexação para navegação na Web. Devido a sua (normalmente simples) estrutura hierárquica, são eficazes na condução dos usuários até os termos mais específicos em um determinado domínio.
<b>Tesouros</b>	Tesouros são instrumentos de controle terminológico utilizados em sistemas de informação para representar a linguagem natural de documentos, de indexadores e de usuários, num sistema de linguagem documentária, usada na indexação e recuperação de informações de um determinado ramo do conhecimento.

**Quadro 3** - Tipologias de vocabulários controlados e suas aplicações em áreas de especialidade.

A abrangência de vocabulário controlado compreende estruturas de complexidade crescente que vão das menos complexas como listas e redes de sinônimos, até estruturas de maior complexidade como taxonomias e tesouros. A diferença entre essas estruturas é a complexidade envolvida no processo de construção, conforme enfatizado na figura 4:



Fonte: ANSI/NISO Z39.19 (2005, p. 17).

**Figura 4** - Complexidade estrutural crescente entre vocabulários controlados.

Observamos, na figura anterior, as evidências de uma estrutura crescentemente complexa de vocabulários controlados, ditada pelas exigências dos tipos de relações que cada um tem que acomodar. Também mostra que os vocabulários mais complexos (taxonomias, *thesauri*), incluem as estruturas mais simples (listas, redes de sinônimos). Por exemplo, um tesouro inclui dispositivos explícitos por controlar sinônimos, organizar hierarquias, e criar relações associativas. Enquanto uma lista é um conjunto de termos simples que não contém relação de qualquer tipo, não distingue rigorosamente “termo” de “palavra”, também não apresenta relações estruturais entre os elementos (GOMES, 1990).

Palavra é a menor unidade léxica que, tomada isoladamente, pode ter vários significados. O uso de palavra na indexação/recuperação é inadequado pela ambiguidade que elas carregam. Concordamos com Gomes (1990) quando afirma que “neste ponto começa a artificialidade da linguagem documentária: a partir de alguns princípios, escolhe-se uma determinada palavra ou expressão para representar um único conceito, ou idéia”. Quando isso ocorre, têm-se não mais uma “palavra”, mas sim um “termo”. Dado que um termo equivale a um conceito (+) uma designação. Essa designação pode ser constituída por uma ou mais unidades léxicas.

O tesouro é considerado por Naumis Peña (2007; 2000) um “sistema de classificação temática ou facetada, cuja estrutura básica está conformada por uma relação de descritores que representam ou descrevem autoridades ou conteúdos temáticos”. A autora complementa a definição acrescentando que é “por meio de unidades lingüísticas, semânticas e suas relações, extraídas da linguagem

formal de uma disciplina ou área específica do conhecimento” que o tesouro se torna um instrumento de representação e recuperação da informação (NAUMIS PEÑA, 2007; 2000).

Ainda, conforme Naumis Peña (2007), o *corpus* ou texto do tesouro encontra-se integrado por: unidades linguísticas unívocas de comunicação representadas por descritores; unidades linguísticas não unívocas de comunicação representadas por não-descritores; estrutura conceptual de uma área do conhecimento em um sistema de informação especializado, que atende a determinados usuários; relações hierárquicas, de equivalência e de afinidade, assim como as que explicam os termos por meio da língua em uso; unidades linguísticas que representam os conteúdos temáticos expressos nos documentos de um sistema, para buscar e recuperar informação.

Nessa direção, Boccato (2009) substancia que os tesouros são constituídos por meio de uma base léxica (descritores e não-descritores) estruturada em relações hierárquicas (termos genéricos e específicos), não-hierárquicas (associativas – termos relacionados) e de equivalência (não-descritores – sinônimos ou quase-sinônimos).

De acordo com Boccato, Ramalho e Fujita (2008, p. 201):

Os tesouros são linguagens de estruturas combinatórias e pós-coordenadas, constituídas de termos - unidades linguísticas provenientes da linguagem de especialidade e da linguagem natural - denominados de descritores, providos de relações sintático-semânticas, referentes a domínios científicos especializados, possibilitando a representação temática do conteúdo de um documento, bem como a recuperação da informação.

Barité (2008), no Dicionário de Organização do Conhecimento, apresenta um conjunto de características para o termo tesouro:

*Tipo de lenguaje documental que se integra con términos analizados y normalizados que guardan entre sí relaciones semánticas y funcionales. El tesouro se organiza bajo fuerte control terminológico, con objeto de proporcionar un instrumento idóneo para el almacenamiento y la recuperación de la información en áreas especializadas. Puede ser monolingüe, monolingüe con equivalencias o multilingüe, conforme a la cobertura idiomática que proponga. En ciertos casos, agrega una notación (BARITÉ, 2008).*



Com essa definição, o autor supracitado evidencia que o tesouro é um “tipo de linguagem documentária composta de termos analisados e normalizados que mantêm entre si relações semânticas e funcionais”. Reforça que o tesouro “é organizado sob rigoroso controle terminológico, com o objetivo de fornecer um instrumento idôneo para o armazenamento e recuperação de informações em áreas especializadas”. Para facilitar o acesso ao conhecimento pode ser apresentado em uma ou várias línguas, ou seja, monolíngue, monolíngue com equivalências, ou multilíngue, em conformidade à língua de cobertura definida na proposta de construção do tesouro. Esclarece ainda que, o tesouro, em certos casos, agrega notação (BARITÉ, 2008).

Gomes (1990, p. 16) define o tesouro de documentação como “[...] Linguagem documentária dinâmica que contém termos relacionados semântica e logicamente, cobrindo de modo compreensivo um domínio do conhecimento”. Tálamo; Lara; Kobashi (1992) afirmam que o tesouro documentário é um vocabulário estruturado segundo um léxico hierarquizado, conforme as relações semânticas entre os seus termos os quais contribuem para o exercício da representação. Sua elaboração apoia-se, essencialmente, em dois referenciais: de um lado, no conhecimento categorizado em assuntos e, de outro, em um *corpus* discursivo do qual são compilados os termos considerados significativos.

No primeiro conjunto, pretende assegurar, de forma classificatória e preditiva, a organização do universo objeto de representação; no segundo, a garantia literária procede à seleção dos conceitos mais frequentes para representação da informação. Nesse aspecto, Lancaster (1987) apresenta referencial sobre garantia do usuário, expressão usada por Lancaster em 1972, também conhecida como Endosso do usuário (DODEBEI, 2002).

Entendemos que essa garantia é uma confirmação direta dos produtores e utilizadores das informações de um campo de atividades que os termos escolhidos são, efetivamente, aqueles utilizados na comunidade e, desse modo, serão eles utilizados nas solicitações de busca no sistema de recuperação da informação (LANCASTER, 1987, DODEBEI, 2002).

Nesse sentido, Tálamo (2001) sintetiza: o tesouro “é uma linguagem documentária que representa de forma normalizada os conceitos de uma área específica através de termos que se manifestam em estruturas lógico-semânticas”. O tesouro tem sido reconhecido como um instrumento apropriado para transmitir

conceitos e as relações recíprocas desses, semelhantemente ao que ocorre com os termos expressos na linguagem dos documentos (DODEBEI, 2002). A autora sobredita esclarece que, pelo fato de a sua documentação possibilitar, através do agrupamento dos termos, o acesso a uma idéia, a sua função é indexar e recuperar informação.

Robredo (2005, p. 157-158) considera que um tesouro pode ser definido levando em consideração sua função ou sua estrutura. De acordo com a sua função – “um tesouro é um instrumento de controle terminológico que permite traduzir a linguagem natural dos documentos, dos indexadores e dos usuários, numa ‘linguagem sistêmica’ mais rígida (linguagem documentária, linguagem do sistema de informação)”. Considerando a sua estrutura – “um tesouro é um vocabulário controlado e dinâmico de termos relacionados semântica e genericamente, que cobre um campo específico de conhecimentos”.

Desse modo, manifestam-se as duas principais características do tesouro: os conceitos, representados por termos e as relações entre eles. Com isso, asseguram-se a contiguidade e a semelhança entre o vocabulário a ser utilizado para representação e o universo do conhecimento expresso em textos (TÁLAMO; LARA; KOBASHI, 1992). Concordamos com as argumentações apresentadas pelos autores supracitados e em consonância com Fujita (1992), entendemos que o tesouro é uma linguagem documentária alfabética, construída por meio de unidades conceituais, extraídas da linguagem formal de uma área específica do conhecimento científico ou técnico.

Sua estrutura sugere a idéia de sistema, visto que os conceitos relacionam-se entre si e são representados por termos. Cada termo, por sua vez, possui vinculação com outro termo, por meio de relação de equivalência, de hierarquia ou de associação. O tesouro, utilizado para a organização e recuperação da informação, em diferentes áreas de especialidade, constitui-se em importante “[...] instrumento de apoio às pesquisas científicas nas áreas de conhecimento” (FUJITA, 1992). Em razão das diversas definições apresentadas a respeito do tesouro, a seguir, elaboramos o quadro 4 com uma síntese das idéias, dispostas cronologicamente, para uma melhor visualização das definições de tesouro segundo normas e/ou autores.

<b>TESAURO</b>	
<b>DEFINIÇÕES SEGUNDO NORMAS E/OU AUTORES</b>	<b>ANO</b>
“[...] vocabulário controlado e dinâmico abrangendo área específica do conhecimento. Em sua estrutura patenteia as relações vigentes entre os termos ou descritores – sinonímicas hierárquicas e outras – que, no conjunto, constitui a linguagem de indexação”. (IBICT, 1984, p. 5).	<b>1984</b>
“[...] vocabulário de uma linguagem de indexação controlado e organizado formalmente com objetivo de explicitar as relações <i>a priori</i> entre conceitos (por exemplo, mais genérico que... ou mais específico que...)”. Para isso, descreve a linguagem de indexação como: “conjunto controlado de termos extraídos da linguagem natural e utilizados para representar de forma breve os assuntos dos documentos”. (ISO 2788, 1986).	<b>1986</b>
“[...] Linguagem documentária dinâmica que contém termos relacionados semântica e logicamente, cobrindo de modo compreensivo um domínio do conhecimento”. (GOMES, 1990, p. 16).	<b>1990</b>
“[...] lista estruturada de conceitos destinados a representar de maneira unívoca o conteúdo dos documentos e das consultas dentro de um sistema documental determinado [...] inclui descritores, não-descritores, relações hierárquicas e de associação e equivalências lingüísticas”. (VAN SLYPE, 1991, p.23-24).	<b>1991</b>
“[...] linguagem documentária, construída por meio de unidades conceituais, extraídas da linguagem formal de uma área específica do conhecimento científico ou técnico. Sua estrutura sugere a idéia de sistema, visto que os conceitos relacionam-se entre si e são representados por termos. Cada termo, por sua vez, possui vinculação com outro termo, por meio de relação de equivalência, de hierarquia ou de associação. O tesouro, utilizado para a organização e recuperação da informação, constitui-se em importante “[...] instrumento de apoio às pesquisas científicas nas áreas de conhecimento”. (FUJITA, 1992, p. 23-24).	<b>1992</b>
“[...] linguagem documentária que representa de forma normalizada os conceitos de uma área específica através de termos que se manifestam em estruturas lógico-semânticas”. (TÁLAMO; LARA; KOBASHI,, 1992).	<b>1992</b>
“[...] vocabulário controlado de uma linguagem de indexação, formalmente organizado para explicitar as relações <i>a priori</i> entre conceitos (por exemplo, como genéricas e específicas). (UNESCO, 1993, p. 14).	<b>1993</b>
“[...] vocabulário especializado, normalizado, pós- coordenado, usado com fins documentários, onde os elementos lingüísticos que o compõem, termos simples ou compostos, se encontram relacionados entre si sintática e semanticamente”. (CURRÁS, 1998).	<b>1998</b>
“[...] sistema de classificação temática ou facetada, cuja estrutura básica está conformada por uma relação de descritores que representam ou descrevem autoridades ou conteúdos temáticos”. [...] por meio de unidades lingüísticas, semânticas e suas relações, extraídas da linguagem formal de uma disciplina ou área específica do conhecimento que [...] se torna um instrumento de representação e recuperação da informação”. (NAUMIS PEÑA, 2000).	<b>2000</b>
“[...] instrumento apropriado para transmitir conceitos e as relações recíprocas desses, semelhantemente ao que ocorre com os termos expressos na linguagem dos documentos”. (DODEBEI, 2002, p. 67).	<b>2002</b>
“[...] como função – “um instrumento de controle terminológico que permite traduzir a linguagem natural dos documentos, dos indexadores e dos usuários, numa ‘linguagem sistêmica’ mais rígida (linguagem documentária, linguagem do sistema de informação)”. Como estrutura – “um vocabulário controlado e dinâmico de termos relacionados semântica e genericamente, que cobre um campo específico de conhecimentos”. (ROBREDO, 2005, p. 157-158).  “[...] vocabulário controlado organizado em uma ordem conhecida e estruturada de modo que os vários relacionamentos entre os termos sejam identificados e indicados claramente por meio de orientações normativas”. (ANSI/NISO-Z39.19, 2005, p. 9).	<b>2005</b>
“[...] linguagens de estruturas combinatórias e pós-coordenadas, constituídas de termos - unidades lingüísticas provenientes da linguagem de especialidade e da linguagem natural -, denominados de descritores, providos de relações sintático-semânticas, referentes a domínios científicos especializados, possibilitando a representação temática do conteúdo de um documento, bem como a recuperação da informação”. (BOCCATO; RAMALHO; FUJITA, 2008, p. 201).  “[...] Tipo de linguagem documental composta de termos analisados e normalizados que guardam entre si relações semânticas e funcionais. O tesouro se organiza sob rigoroso controle terminológico, com objetivo de proporcionar um instrumento idóneo para o armazenamento e a recuperação da informação em áreas especializadas [...]”. (BARITÉ, 2008).	<b>2008</b>

**Quadro 4 – O Tesouro - definições conforme Normas e/ou Autores.**

Com base nos referenciais acima destacados, podemos considerar que o tesouro é um importante instrumento e dos mais adequados para representar e recuperar a informação em áreas especializadas. Sua importância apoia-se no controle de terminologias que compõem os documentos de um sistema de informação especializada e no emprego dos mesmos termos utilizados para representar os documentos que incluem os mesmos conceitos, quando realizada a busca de um assunto.

A respeito da função do tesouro, Fujita (1992) apresenta-o sob dois pontos de vista: no primeiro, recomenda-o para o controle de vocabulário no processo de indexação para representar a linguagem de conteúdos de documentos, de indexadores e de usuários de áreas especializadas, para integrar os diversos vocabulários existentes em sistemas de informação; já no segundo ponto de vista, indica-o para auxílio na busca e recuperação de informação pelo usuário.

Na visão de Fujita (1992, p. 23), essas funções são complementares. Argumenta que “para existir recuperação da informação desejada é preciso que, tanto a linguagem de busca quanto a fonte de informação estejam adequadamente representadas através de uma linguagem documentária”. O uso dos tesouros nas atividades de indexação e recuperação de informações, “tenta resolver o problema da alocação de documentos em classes de assuntos [...] porque é um instrumento que relaciona os descritores/termos de forma mais consistente”, corrobora Dodebei (2002).

Em geral, a condição principal a ser cumprida por um tesouro é a de restringir-se a um assunto especializado, no entanto, construíram-se alguns de âmbito geral. Assim, numa primeira classificação serão identificados como: gerais ou especializados (CURRÁS, 2005).

E, ainda, quanto ao assunto podem ser classificados como: tesouros multidisciplinares – são os que incluem termos de diversas disciplinas que inferem num problema ou tesouros monodisciplinares - voltados apenas a uma disciplina científica (GOMES, 1990).

Além disso, podemos classificá-los pelo nível de especificidade dos termos como: macrotresouros – quando apresentam termos com conceitos mais ou menos amplos ou microtesouros – quando os termos representam conceitos bastante específicos, restritos a uma área ou subárea do conhecimento.

Quando consideramos a ordenação do assunto ou assuntos tratados, as classificações são diversificadas. Desse modo, será necessário considerar tesouros como: alfabéticos ou sistemáticos. Nos tesouros sistemáticos, os assuntos ali compreendidos podem ser ordenados hierarquicamente, distribuídos em facetas, dispostos em gráficos apropriados ao caso. Assim, os tesouros poderão ser: hierárquicos, facetados ou gráficos.

O tesouro apresenta uma estrutura sintética, precisa e uma complexa rede de remissivas que tem alguma ligação ou nexos entre si, realidade que permite ao profissional da informação localizar com mais facilidade o termo preferido para uma busca. Ele organiza “um relacionamento lógico e hierárquico dos descritores, o que contribui para a indexação dos documentos ao nível específico e/ou genérico (modulação)” (DODEBEI, 2002, p. 67).

Naumis Peña (2007) reforça que o tesouro apresenta-se sob dois aspectos: estrutural e funcional. No primeiro, o estrutural, revela-se seu caráter de linguagem controlada de relações semânticas e hierárquicas que cobrem uma determinada área do conhecimento. No segundo aspecto, o funcional, considera-se seu uso como linguagem documentária para traduzir a linguagem natural dos documentos desde os indexadores até aos usuários de uma linguagem especializada.

Para Chaumier (1986), as funções do tesouro apresentam-se de grande utilidade: na entrada é usado para a descrição do documento; no tratamento é usado para a elaboração de catálogos e, na saída – usado para formulação de perguntas em um sistema de informação. Concordamos com o autor quando esclarece a importância do tesouro: na entrada – a) classifica os conceitos; b) proporciona o vocabulário que traduz os conceitos; c) apresenta um vocabulário unívoco; d) favorece a coerência da análise documentária e e) apresenta o entorno semântico dos descritores escolhidos para traduzir os conceitos; no tratamento – a) facilita o armazenamento da informação; b) transforma a informação em dados manipuláveis; c) favorece o controle e a validação dos dados e na saída – a) estabelece os conceitos da demanda; b) proporciona o vocabulário dos conceitos da demanda; c) facilita o diálogo entre homem e sistema; d) focaliza ou amplia a pergunta (CHAUMIER, 1986).

Lara (1999, p.138) ensina que “o enfoque da mediação em Documentação [...] acontece mais exatamente entre a linguagem do sistema e a

linguagem do usuário”. Entendemos, nesse contexto, que os esforços envidados no sentido de promover a mediação decorrem por meio da aplicação de métodos e técnicas no processo de construção do tesouro, com o objetivo de facilitar a comunicação e contribuir para o processo de geração do conhecimento.

Diante do exposto, destacamos que, nos dias atuais, existem linguagens documentárias que apresentam certo rigor metodológico na sua construção. Para tanto, reportamos a Svenonius (2000) ao argumentar que “essas linguagens quando utilizadas nos sistemas de recuperação da informação podem provê-lo de qualidade adicional de modo que por seu intermédio possamos transformar informação em conhecimento.” A autora supracitada enfatiza que, quando isso acontece, “a linguagem de indexação torna-se análoga ao próprio conhecimento” (SVENONIUS, 2000).

Entendemos que as linguagens documentárias são utilizadas nos sistemas de informação para apoiar os trabalhos do indexador na seleção dos termos mais significativos para a representação do(s) assunto(s) de um documento e para guiar os pesquisadores, e, até mesmo, os usuários que não se encontram bem familiarizados com a temática, tampouco com a terminologia da área especializada, como escolher os termos indexados que melhor representam no sistema o assunto solicitado.

Nessa direção, Moura et al. (2005, p. 57) destacam mais algumas funções pertinentes às linguagens documentárias: recuperar documentos com conteúdo semelhante, recuperar documentos relevantes sobre um assunto específico, recuperar documentos por grandes áreas de assunto, possibilitar a conversão dos termos de indexação entre diferentes linguagens, auxiliar na escolha do termo adequado para a estratégia de busca, representar o assunto de maneira consistente permitindo a compatibilidade e o diálogo entre a linguagem do autor, do indexador e a do pesquisador.

Partindo do princípio de que não existe um instrumento adequado, compatível à área especializada para a qual se precisa organizar e representar tematicamente uma coleção, entendemos ser indispensável, no processo de construção de tesouros, que se aproxime a representação dos dois segmentos integrantes do sistema de recuperação de informação em áreas especializadas: o segmento das demandas por informação e o segmento dos conteúdos temáticos documentários disponíveis, em razão da necessidade de interação entre o

bibliotecário e o usuário de um sistema de informação em áreas especializadas para promover o acesso ao conteúdo do documento/informação (CHAUMIER, 1986; GOMES, 1990; FUJITA, 1992; DODEBEI, 2002; NAUMIS PEÑA, 2007).

Acreditamos que a linguagem documentária, principalmente, o tesauro, exerce um papel fundamental no processo de recuperação da informação, em razão de ser um instrumento construído para atuar tanto nos processos de indexação como nos processos de recuperação de informação de um sistema de recuperação da informação em áreas especializadas. As alegações de Fujino e Jacomini (2007) podem corroborar nossa posição:

[...] em geral, o usuário tem muita dificuldade para explicitar sua necessidade de informação ao sistema. É conhecido pelos profissionais que atuam com informação tecnológica, o pressuposto de que todo problema apresentado pelo usuário, traz outro problema oculto que, na maioria das vezes, é o foco da questão (FUJINO; JACOMINI, 2007, p. 87).

Julgamos importante a integração entre o bibliotecário e o usuário de um sistema de informação especializado, e o envolvimento participativo desse último, em atividades de estudos voltadas para conhecer e identificar as reais necessidades de informação dos usuários. Isso pressupõe uma efetiva participação do usuário em estudos de avaliação do uso de produtos e serviços disponíveis no sistema de informação. Um sistema de informação tem como principal objetivo responder às necessidades de informação demandadas pelos seus usuários, com pertinência, velocidade de resposta e atualidade, conforme declara Melo (1994).

E “vai ao encontro de seus objetivos em atender às necessidades de informação solicitadas por seus usuários com competência e rapidez, visando à satisfação e à confiabilidade desses usuários nos seus serviços prestados” (BOCCATO, 2005). Surge, com isso, a necessidade de ter bons instrumentos de apoio, tais como: catálogos, índices, sistemas de recuperação *on line* e outros (MELO, 1994).

Neste momento, não podemos deixar de considerar as crescentes mudanças no ambiente de um sistema de informação relacionadas às tecnologias de informação e comunicação, centradas no uso intensivo de redes de computadores. Assim, utilizar a informática para acesso aos recursos de informação disponíveis em um sistema de informação, neste caso, em especial os catálogos

tornou-se fundamental nos dias atuais. “Não é possível pensar em dado, informação, conhecimento e posterior aprendizado sem pensar nas ferramentas que facilitam sua busca e identificação” (MIRANDA; FORCELLINI, 2007, p. 97).

Esses instrumentos podem e necessitam ser avaliados no sentido de se ter um controle sobre a qualidade da recuperação de informação (MELO, 1994). Ainda que não se configure como foco de nosso estudo, nesse momento, consideramos oportuno observar estudos realizados sobre recuperação de informação e avaliação de linguagens documentárias, tendo como base de mediação, o uso de linguagens de busca. Isso posto, o tesouro torna-se um dos principais componentes de catálogos disponíveis *on line* para acesso ao sistema de informação de áreas especializadas.

O catálogo, de modo geral, informa o que existe de determinado autor, o que determina a existência ou não de certa obra conhecida pelo título, ou sobre determinado assunto e a sua localização (ZAMBEL, 1978). É a relação de todo e qualquer documento, arranjada de acordo com algum plano definido. É também a enumeração, inventário de listas de pessoas, coisas ou acontecimentos postos em uma determinada ordem e onde podem ser encontrado.

Enquanto o catálogo de assunto, reúne as entradas das obras existentes na biblioteca segundo o conteúdo de cada uma delas. Contém uma lista por ordem alfabética de assuntos com as remissivas cruzadas necessárias para relacioná-las. É também chamado de catálogo ideográfico ou catálogo metódico (ZAMBEL, 1978). O catálogo de assunto é um instrumento de apoio para o leitor que não sabe exatamente qual obra procura, nem conhece o nome do autor, porém lhe interessa ler sobre determinado tema. Pode-se dizer que este é talvez o catálogo mais difícil de elaborar, mas é um dos mais procurados pelos usuários.

Diante do exposto, destacamos alguns resultados de estudos realizados sobre avaliação de catálogos disponíveis *on line* que consideram as linguagens documentárias como um dos aspectos analisados no processo de recuperação de informação, conforme relatos de Markey (1983), Anderson (1998), e Játiva Miralles (2004), mencionados na pesquisa realizada por Boccato (2009). Esclarecemos, contudo, que extraímos dos trabalhos analisados, unicamente, os resultados significativos para o reforço de nosso estudo.

Os estudos de Markey (1983), realizados em bibliotecas públicas de escolas de ensino médio, além de bibliotecas universitárias, localizadas em Ohio,



Estados Unidos, revelaram a insatisfação dos usuários na recuperação da informação, tendo como base de mediação o uso das linguagens de busca disponíveis nos respectivos catálogos.

Com base nos resultados das pesquisas de Anderson (1998) da *University of Huddersfield*, no Reino Unido, em catálogos *on line* de bibliotecas universitárias, evidenciamos que: o usuário deve conhecer a linguagem do sistema visando à otimização de seu uso na busca por assunto, necessidade de programa de treinamento para os novos usuários na utilização do sistema, necessidade de compatibilidade das linguagens dos catálogos com a linguagem do usuário.

A pesquisa de Játiva Miralles (2004), na Espanha realizou estudos em catálogos *on line* de bibliotecas universitárias em cinco universidades espanholas. Os resultados correspondem à apresentação de indicadores de qualidade aplicáveis para a análise, avaliação e comparação de interfaces disponíveis para acesso aos catálogos, pontuando-se requisitos de problemas na construção da linguagem documentária adotada pelo sistema e características do processamento técnico.

Os resultados apresentados nesses estudos, no que se refere à questão das linguagens documentárias usadas nos sistemas como instrumentos de apoio nos processos de representação e recuperação de informação, conduzem particular atenção às etapas que envolvem o processo de construção de tesouros, especialmente, quanto à etapa de compilação de termos.

## 2.4 Etapas da construção de um tesouro

No intuito de prover o embasamento teórico-metodológico desta pesquisa, recorreremos às publicações normativas, traduzidas e adaptadas às peculiaridades da língua portuguesa que orientam o processo de construção de tesouros. As recomendações para a construção de tesouros e as técnicas descritas nas publicações elencadas para esta investigação, a saber, Diretrizes IBICT(1984) e Diretrizes UNESCO (1993), fundamentam-se em Princípios Diretivos gerais que podem ser aplicáveis a qualquer área de assunto (IBICT, 1984; UNESCO, 1993), além de fornecer subsídios para a compatibilização de tesouros que estão em

construção ou que serão elaborados no futuro (UNESCO, 1993). Em razão disso, apresentamos, a seguir, um extrato dos principais estudos, que identificamos ao longo da pesquisa que tratam do processo de construção de tesouros, em especial, no que diz respeito aos métodos e etapas da construção de tesouro.

As normas, manuais e outros trabalhos sobre construção de tesouros, na sua maioria, fundamentam-se na norma 2788 - *Guidelines for the Establishment and Development of Monolingual Thesauri*, da *International Standardization Organization* (ISO). Por esse motivo, os Princípios em que se fundam essa Norma podem ser considerados, independentemente, da língua e da cultura. Desse modo, eles têm sido anuídos também como a base para construção de tesouros multilíngües, conforme UNESCO (1993).

A seguir, identificamos as normas analisadas nesse estudo. Inicialmente destacamos:

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA. **Diretrizes para elaboração de tesouros monolíngües**; projeto coordenado por Hagar Espanha Gomes. Brasília: IBICT, 1984. 70 p. (Publicação baseada na BS 5723: 1979).

Esclarecemos que a primeira das Diretrizes indicadas para a construção de tesouros, identificada nesse estudo como Diretrizes IBICT (1984), fundamenta-se nos Princípios Diretivos da *British Standard Institution* BSI 5723: 1979 – Diretrizes para o Estabelecimento e Desenvolvimento de Tesouros Monolíngües que, por sua vez, tem fulcro na publicação 2788 da ISO. Por isso, reforçamos que essa norma reúne conceitos e diretrizes adotados em nível internacional.

Em seguida, distinguimos a outra norma analisada:

UNESCO. **Diretrizes para o estabelecimento e desenvolvimento de tesouros monolíngües**. [2ª ed. rev. por Derek Austin e Peter Dale]; traduzido por Bianca Amaro de Melo e revisto por Lígia Maria Café de Miranda. Brasília: IBICT/SENAI, 1993. 86 p. (PGI/81/WS/15).

Essa publicação é uma versão revista e ampliada da versão preparada por Derek Austin e Peter Dale, do *Bibliographic Systems and Standards Office*, da *Bibliographic Services Division*, da *British Library*, Londres, como uma

base para a nova norma ISO 2788. Publicada em 1981 pela UNESCO, foi traduzida por Bianca Amaro de Melo, revista por Lúgia Maria Café de Miranda e impressa no Brasil pelo IBICT em 1991. Em 1993, foi impressa pelo IBICT em parceria com o Serviço Nacional de Aprendizagem Industrial (SENAI). Indica um conjunto de regras para a preparação e desenvolvimento de tesouros, seja por meio manual ou mecânico, sem considerar a área do conhecimento. Esse documento é identificado nesse estudo por Diretrizes UNESCO (1993).

Combinamos às normas analisadas, a **Z39.19** - *Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies*, da ANSI/NISO, edição de 2005, que propõe linhas gerais para a construção, formato e gestão de vocabulários controlados monolíngues, de um modo geral, abrangendo os tesouros e também estruturas menos complexas como listas, redes de sinônimos e taxonomias. Além de constituir-se em referência normativa atualizada, especialmente, no que se refere a orientações sobre construção de tesouros eletrônicos, essa publicação encontra-se identificada, neste estudo, como ANSI/NISO (2005). A seguir, apresentamos sua referência completa:

AMERICAN NATIONAL STANDARDS INSTITUTE; NATIONAL INFORMATION STANDARDS ORGANIZATION. **Z39.19** - *Guidelines for the construction, format, and management of monolingual controlled vocabularies*. Bethesda, Ma: NISO, 2005. 184 p. (tradução com propósito acadêmico).

Para efeito de demonstração, no quadro 5, apresentamos o sumário que corresponde à estrutura interna das Normas supracitadas, além de ser mantida a grafia apresentada na descrição de cada estrutura interna.

<b>E S T R U T U R A S</b>	IBICT. <b>Diretrizes para elaboração de tesouros monolíngües</b> ; projeto coordenado por Hagar Espanha Gomes. Brasília: IBICT, 1984. 70 p.	UNESCO. <b>Diretrizes para o estabelecimento e desenvolvimento de tesouros monolíngües</b> . [2ª ed. rev. por Derek Austin e Peter Dale]; Trad. de Bianca Amaro de Melo; rev. de Lígia Maria Café de Miranda. Brasília: IBICT/ SENAI, 1993. 86 p.	ANSI/NISO. <b>Z39.19 - Guidelines for the construction, format, and management of monolingual controlled vocabularies</b> . Bethesda, Ma: NISO, 2005. 184 p.
--	<b>SUMÁRIO</b>	<b>SUMÁRIO</b>	<b>Sumário</b>
--	INTRODUÇÃO	PREFÁCIO	<b>Prefácio</b>
1	ÂMBITO	INTRODUÇÃO E CAMPO DE APLICAÇÃO	<b>Introdução</b>
2	REFERÊNCIAS	REFERÊNCIAS	<b>Escopo</b>
3	DEFINIÇÕES	DEFINIÇÕES	<b>Normas Referenciadas</b>
4	ABREVIATURAS E SÍMBOLOS	ABREVIATURAS E SÍMBOLOS	<b>Definições, Abreviaturas e Símbolos</b>
5	CONTROLE DE VOCABULÁRIO	CONTROLE DE VOCABULÁRIO	<b>Vocabulários Controlados (Propósitos, Conceitos, Princípios e Estrutura)</b>
6	TERMOS DE INDEXAÇÃO	TERMOS DE INDEXAÇÃO	<b>Seleção de Termos, Aplicação e Formas de apresentação</b>
7	FATORAÇÃO	TERMOS COMPOSTOS	<b>Termos compostos</b>
8	RELACIONAMENTOS BÁSICOS ENTRE OS TERMOS DE UM TESAURO	RELAÇÕES BÁSICAS EM UM TESAURO	<b>Relacionamentos</b>
9	APRESENTAÇÃO GRÁFICA DOS TERMOS E DE SEUS RELACIONAMENTOS	APRESENTAÇÃO DOS TERMOS E SUAS RELAÇÕES	<b>Apresentação de Vocabulários Controlados</b>
10	ALGUNS ASPECTOS GERENCIAIS DA ELABORAÇÃO DE TESAURUS	ASPECTOS GERENCIAIS DA ELABORAÇÃO DE TESAUROS	<b>Interoperabilidade</b>
11	----	----	<b>Construção, Teste (Avaliação), Manutenção, e Sistemas de Gestão</b>
--	APÊNDICE A – Referências	APÊNDICE: Simbolização das Relações de um Tesouro	<b>Apêndice A - Síntese da Norma Requisitos / Recomendações</b>
--	APÊNDICE B – Páginas de exemplos de tesouros	ÍNDICE	<b>Apêndice B - Comparação dos Tipos de Vocabulários</b>
--	----	----	<b>Apêndice C - Características e Usos do Vocabulário Controlado Opções de Apresentação</b>
--	----	----	<b>Apêndice D - Métodos para obtenção da interoperabilidade</b>
--	----	----	<b>Apêndice E - Exemplo de Formulários de Registro de Termo Candidato</b>
--	----	----	<b>Apêndice F - Referências</b>
--	----	----	<b>Bibliografia</b>
--	----	----	<b>Glossário / Índice</b>

**Quadro 5** – Demonstração da estrutura interna das Diretrizes IBICT (1984); UNESCO (1993) e ANSI/NISO (2005).

#### 2.4.1 Etapas da construção de tesouro segundo as diretrizes

A seguir, separadamente, descrevemos a sequência de etapas da construção de um tesouro, conforme indicado pelas Normas selecionadas para este estudo. Iniciamos com as Diretrizes para Elaboração de Tesouros Monolíngues, publicada pelo IBICT (1984). Em seguida, expomos as Diretrizes para o Estabelecimento e Desenvolvimento de Tesouros Monolíngues, da UNESCO (1993). Por último, contudo, não menos importante, destacamos as instruções das Diretrizes para a Construção, Formato e Gestão de Vocabulários Controlados Monolíngues, da ANSI/NISO Z39.19 (2005, p. 90).

##### a) Diretrizes para Elaboração de Tesouros Monolíngues (IBICT, 1984)

As Diretrizes têm por objetivo “garantir a uniformidade metodológica: a) dentro de um único serviço ou em uma rede de serviços cooperantes e b) entre diferentes serviços de indexação” (IBICT, 1984, p. 4). Essa publicação, além da introdução, encontra-se organizada em 10 seções. Nas seções enumeradas de 1 a 10, nessa ordem, observamos: Âmbito, Referências, Definições, Abreviaturas e símbolos, Controle de vocabulário, Termos de indexação, Fatoração, Relacionamentos básicos entre os termos de um tesouro; Apresentação gráfica dos termos e de seus relacionamentos e Alguns aspectos gerenciais da elaboração de tesaurus. Para os propósitos deste estudo, nesta publicação, interessa-nos atentar à Seção 10.

A Seção 10 - “Alguns Aspectos Gerenciais da Elaboração de Tesaurus” apresenta uma sequência de oito etapas que orientam o processo de construção de tesouro. A descrição das etapas tem início na primeira etapa com as orientações preliminares, a saber:

1) Trabalho preliminar – como orientações preliminares ao processo de construção de tesouros indicam duas recomendações: a) “Comunicação de intenção – quando um serviço decide a preparar um tesouro novo, essa intenção deve ser comunicada em periódicos apropriado (por ex., num periódico de

documentação, de classificação)”; b) Não duplicação de trabalho - a verificação da existência de tesouro que cobre a mesma área de conhecimento ou parte dele, pode ser útil como ponto de partida. Todavia, orientam que “cada tesouro deve refletir as necessidades de uma comunidade específica de usuários” (IBICT, 1984, p. 52).

2) Métodos de compilação – como orientações gerais acentuam que, tanto quanto possível, a forma do tesouro, lista alfabética, apresentação sistemática ou planigráfica, deve ser determinada antes que os termos sejam considerados como candidatos à inclusão. Quanto aos métodos de compilação, tais diretrizes, destacam duas abordagens de compilação possíveis:

- a) método dedutivo: os termos são extraídos da literatura durante o estágio preliminar de indexação. Entretanto, até que um número suficiente de termos tenha sido coletado, todos os termos são considerados por uma comissão, formada de preferência por indexadores e especialistas que decidem primeiro quais termos representam as classes mais amplas, fixando os demais termos nessas classes e relacionando-os logicamente com base em suas conceituações;
- b) método indutivo: novos termos são admitidos no tesouro logo que são encontrados na literatura, sendo cada termo designado como membro de uma ou mais classes estabelecidas em bases *ad hoc* durante o ato de indexação. A elaboração do tesouro é considerada, desde o início, como uma operação contínua e, embora se busque assistência de especialistas sempre que for preciso, tais profissionais não são membros de uma comissão formal. Na oportunidade, ainda que tenham sido evidenciadas as duas abordagens descritas acima, nessa publicação aparece mais uma abordagem de compilação, denominada;
- c) método relacional: em notas de rodapé, na publicação, consta a seguinte explicação: “\*Dahlberg aponta um terceiro tipo de abordagem – a abordagem relacional – que parte do aspecto formal, categorial” (IBICT, 1984, p. 52). Para Dahlberg, método relacional é o método analítico que consiste tanto na análise das definições dos conceitos integradores do sistema quanto na estruturação destes, em vista das relações existentes entre eles

(DAHLBERG, 1978b). O método relacional faz parte do aspecto formal e categorial no qual o elemento material é o conceito, e todos os conceitos de determinada área de assunto são vistos como elementos do sistema, ou seja, relacionados entre si (MOTTA, 1987). É um método que pretende ser instrumento para o estabelecimento de relações em tesouros de forma mais objetiva porque parte do aspecto formal, categorial e fundamenta-se na Teoria do Conceito, a qual, por sua vez, compreende entre outras coisas, o conceito e sua definição.

O conceito, conforme propõe Dahlberg (1978a), é entendido como unidade de conhecimento, pois o conceito é o ponto de partida para estabelecer as relações conceptuais e determinar a forma verbal. A definição é o estabelecimento de uma equivalência entre o termo [...] e as características necessárias do referente de um conceito [...] com o objetivo de delimitar o uso do termo no discurso (DAHLBERG, 1978a). Sendo assim, a definição é ferramenta básica para a aplicação do método relacional no estabelecimento de relações em um tesouro (MOTTA, 1987). O modelo desse método envolve três elementos fundamentais: o conceito, a análise conceitual e a definição do conceito. O método relacional “parece ser útil em várias circunstâncias [...]: para determinação de sinônimos, atualização do sistema, inclusão de novos termos em uma só categoria e mapeamento de áreas de assunto [...]” (MOTTA, 1987, p. 9).

Nas Diretrizes, encontramos, ainda, no item dedicado aos métodos de compilação, as seguintes orientações: durante a sequência de etapas de construção do tesouro, empregam-se, de modo igual, os métodos dedutivo e indutivo. Por exemplo: as classes de termos que tenham sido estabelecidas de início de forma indutiva podem ser revistas mais tarde por uma comissão de indexadores e especialistas da área do tesouro. Ambas as técnicas são empíricas. Nesse sentido, decisões tomadas nos primeiros estágios da compilação podem ser revistas, à medida que os indexadores ganham mais experiência. A inclusão de novos termos, a conferência dos relacionamentos entre os termos, a fatoração, as notas de aplicação devem seguir uma linha lógica e consistente. A assessoria de especialistas na escolha entre termos sinônimos e outras formas variantes é

recomendada. Entendemos que as recomendações apresentadas podem constituir-se em mais outra abordagem – Combinação de métodos.

3) Registro de termos – manter um registro individual, em ficha, para cada termo admitido no tesouro. Nesse registro, deve-se identificar a fonte do termo, o nome das autoridades consultadas e a data de inclusão, indicar também os termos sinônimos, os mais genéricos, os mais específicos e outros tipos de relacionamentos.

4) Verificação do termo – verificar a forma correta de um termo antes de sua inclusão no tesouro. Revisar os relacionamentos entre cada novo termo com os demais termos e hierarquias já registrados no tesouro. Para os termos candidatos à inclusão no tesouro, consultar as fontes autorizadas como: dicionários e enciclopédias especializadas, tesouros e esquema de classificação, índice de periódicos especializados, tratados e manuais atualizados e especialistas no assunto, especialmente os que possuem algum conhecimento de indexação ou documentação.

5) Especificidade – empregar terminologia específica ao núcleo central do assunto coberto. Se um serviço de informação compreende mais de uma área de conhecimento, pode tornar-se necessário desenvolver tesouros específicos para as áreas, ligados a um tesouro geral.

6) Admissão e exclusão de termos – a) assegurar que os termos e seus relacionamentos reflitam a linguagem usada pelos especialistas da área do tesouro, (fazer sempre uma remissiva da forma não-preferida do termo); b) alguns termos podem ser admitidos sob o argumento de que podem ser pontos de acesso úteis ao tesouro ou de que possuem valor potencial como descritores, embora não tenham ainda sido usados na indexação; c) termos muito usados e termos que não são usados frequentemente devem ser reconsiderados com vistas a sua exclusão, ambos os tipos são em geral ineficazes na recuperação. Em alguns casos, o termo muito utilizado poderia ser substituído por dois ou mais termos de maior especificidade; d) se um termo for retirado do tesouro, mas tiver figurado anteriormente na indexação, seus registros devem receber a seguinte anotação: “Apenas para fins de indexação” e a data do registro da anotação; e) antes do tesouro ser publicado, recomenda-se que ele seja testado por meio de uma experiência-piloto. Por exemplo: uma versão preliminar deve ser distribuída a um



grupo seletivo de usuários, e qualquer sugestão de modificação dos termos e/ou de seus relacionamentos devem ser consideradas pelos compiladores.

O tesouro deve ser submetido a uma revisão técnica após determinado período de uso e, a partir daí, a intervalos regulares, levando em conta mudanças no uso de termos dentro do campo de cobertura do tesouro.

7) Uso de equipamento automático de processamento de dados – a) a seleção de termos, a escolha da forma e outras atividades semelhantes, devem envolver decisões intelectuais por parte dos indexadores, mas pode-se utilizar o computador para tarefas como: 1) seleção automática de termos candidatos extraídos de títulos e resumos; 2) determinação automática (nos sistemas automatizados de indexação) da frequência com que os termos são usados na indexação e nas perguntas de busca. Termos com índice excepcionalmente elevados ou baixos devem ser considerados como candidatos à exclusão; b) o uso do computador sofre restrições impostas por limitações do equipamento disponível ou do software. As características mais prováveis de serem afetadas são: 1) número de caracteres permitido em um termo; 2) conjunto de caracteres, inclusive sinais de pontuação, disponíveis na impressora; 3) número de níveis de subordinação permitido em um tesouro sistemático. Sempre que possível, os indexadores devem ter acesso aos equipamentos e aos programas capazes de lidar com as necessidades especiais de um tesouro. Não é recomendável adaptar o tesouro para ajustá-lo a uma máquina inadequada.

8) Forma e conteúdo do tesouro – 1) em razão da variedade de apresentações disponíveis, não é possível estabelecer um formato padrão de apresentação de tesouros impressos. Entretanto, os seguintes elementos devem ser destacados: a) página de rosto; b) sumário; c) introdução; d) apresentação sistemática ou planigráfica (quando for possível) e e) parte alfabética; 2) todos os tesouros devem conter uma introdução abrangente que informe de maneira clara: a) finalidade; b) área do conhecimento que abrange, com as áreas de assunto central e periférica identificadas separadamente; c) significados de todas as convenções e abreviaturas; d) número total de termos (com os totais de descritores, não descritores, indicados separadamente); e) regras adotadas para o estabelecimento dos termos de indexação e para a decisão de seus interrelacionamentos; f) regras de alfabetização adotadas e os significados de quaisquer sinais que não sejam usados de acordo com as normas; g) informações sobre a política de atualização,

nome e endereço do serviço responsável ao qual devem ser enviados comentários e sugestões e h) data em que foram incluídos os últimos termos. Sempre que apropriado, esses pontos devem ser ilustrados com exemplos (IBICT, 1984).

b) Diretrizes para o Estabelecimento e Desenvolvimento de Tesouros Monolíngues (UNESCO, 1993)

Essas Diretrizes indicam um conjunto de regras destinadas a facilitar a preparação e desenvolvimento de tesouros, seja por meio manual ou mecânico, sem considerar a área do conhecimento para a qual tenham sido estabelecidos. Têm como objetivo maior “fornecer subsídios para a compatibilização de tesouros que estão em elaboração ou que serão elaborados no futuro” (UNESCO, 1993, p. 5). Do mesmo modo que a Norma do IBICT (1984), citada anteriormente, essa publicação encontra-se organizada em 10 seções. Nas seções enumeradas de 1 a 10, nessa ordem, observamos: Introdução e Campo de Aplicação; Referências; Definições; Abreviaturas e símbolos; Controle de vocabulário; Termos de indexação; Termos compostos; Relações básicos em um tesouro; Apresentação dos termos e suas relações e Aspectos gerenciais da elaboração de tesouros.

A Seção 10 - “Aspectos Gerenciais da Elaboração de Tesouros”, expõe uma sequência de oito etapas referentes ao processo de construção de tesouro, iniciando essas etapas pelos Métodos de Compilação, a saber:

1) Métodos de compilação – a forma do tesouro, lista alfabética, apresentação sistemática ou gráfica, deve ser decidida, se possível, antes de se reunirem os termos e considerá-los candidatos à inclusão. No processo de construção de tesouros, dois métodos iniciais podem ser aplicados:

- a) Método dedutivo - os termos são extraídos de documentos durante uma etapa preliminar de indexação. Não é feito controle do vocabulário, nem mesmo para determinar as relações entre os termos, a não ser que um número suficiente de termos tenha sido coletado. Nesse caso, todos os termos são revistos por um grupo de especialistas no assunto. Primeiro, devem ser identificados os termos que representam categorias genéricas.

Os termos restantes devem ser relacionados nessas categorias conforme as suas relações lógicas. Desse modo, as categorias tendem a ser estabelecidas sobre uma base de termo genérico à específico. O controle do vocabulário deve ser efetuado no estágio em que as categorias são estabelecidas, observando os princípios descritos nas seções de 5 a 7, dessa Norma.

Na seção 5 – Controle de Vocabulário – empregam-se dois meios principais para a sua obtenção: a) a aplicação dos termos se restringe deliberadamente aos significados selecionados. Cada termo de um tesouro, em geral está restrito a um único significado mais adequado às necessidades de um sistema de indexação. A estrutura de um tesouro, especialmente sua representação de relações hierárquicas, indica o significado que se pretende atribuir a um termo. Deve-se acrescentar ao termo uma definição ou nota explicativa. Essa nota deve especificar o significado escolhido e também pode indicar outros significados utilizados na linguagem natural, mas que foram rejeitados para fins de indexação; b) quando o mesmo conceito puder ser expresso por dois ou mais sinônimos, selecionar-se-á um desses como o termo preferido – termo utilizado consistentemente para representar conceito no processo de indexação; também conhecido como descritor – que será utilizado na indexação. Em qualquer sinônimo que também funcione como ponto de acesso do usuário, deve-se fazer referência ao termo preferido. Deve-se distinguir os termos preferidos dos termos não-preferidos – sinônimos ou quase-sinônimos de termos que aparecem como pontos de entrada em um tesouro ou índice alfabético, remetendo o usuário, por meio de uma instrução (por exemplo, USE e VEJA), ao termo preferido apropriado também conhecido como não-descritor.

Nas seções 6 e 7 - Termos de Indexação e Termos Compostos, respectivamente, indicam-se outros meios para se obter o controle de vocabulário. Orientam a escolha de formas singulares ou plurais, a seleção do termo preferido quando existirem sinônimos e quando um termo composto deve ser conservado em sua forma pré-coordenada ou ser dividido em componentes distintos, sendo cada um expresso como um substantivo e utilizado independentemente como um termo de indexação.

- b) Método indutivo - novos termos são admitidos no tesouro logo que aparecem nos documentos. O controle do vocabulário é feito desde a etapa inicial, e cada termo, à medida que é admitido, é designado como membro de uma ou mais categorias genéricas constituídas sobre uma base *ad hoc*. O tesouro é, portanto, organizado sobre uma base de termo específico a genérico. A elaboração do tesouro é, desde o princípio, uma operação contínua e, apesar de ser necessária a assistência de técnicos especialistas, eles não precisam fazer parte de um grupo editorial formal ou comitê. Vale dizer que, nessas Diretrizes, encontramos, também, a seguinte recomendação;
- c) Combinação de Métodos - na prática, é possível empregar-se tanto o método dedutivo como indutivo em uma ou outra etapa da elaboração do tesouro. Com base no exposto, essa orientação poderia ser considerada mais uma abordagem. Por exemplo, um grupo editorial composto por indexadores e especialistas pode primeiramente estabelecer as categorias dos termos indutivamente e depois examiná-las pelo ponto de vista dedutivo. Ambas as técnicas são essencialmente empíricas e, desde o início, deve-se ter em mente que algumas decisões tomadas durante a etapa de compilação podem ter que ser revisadas à medida que se adquire maior experiência. A adição de novos termos a uma linguagem de indexação frequentemente esclarece as decisões tomadas anteriormente quanto aos termos e suas inter-relações. Os compiladores devem revisar frequentemente os termos e as hierarquias para assegurar que as relações entre os termos, a fatoração de termos compostos etc., seguem os princípios gerais que têm sido aplicados. Recomenda-se particularmente que se busque o auxílio de especialista na escolha dos termos sinônimos e formas variantes dos termos.

2) Registro dos termos - para cada termo admitido no tesouro, deve ser mantido o registro individual, por exemplo, em uma ficha. Deve-se fazer o registro no momento da admissão do termo e nele identificar a fonte do termo (especialmente no caso de neologismos ou termos não-familiares), os nomes de

todas as autoridades consultadas e a data da inclusão, como também referências e sinônimos, termos genéricos, específicos e relacionados.

3) Verificação dos termos - quando possível, deve-se verificar se o termo é correto antes de admiti-lo no tesauro. O compilador deve também revisar as relações entre cada novo termo e os outros termos da hierarquia à qual está relacionado. Devem ser examinadas as seguintes fontes de referência para os termos candidatos à inclusão: a) dicionários e enciclopédias especializadas; b) tesouros existentes; c) esquemas de classificação; d) índices de periódicos especializados e periódicos de resumos; e) tratados e manuais atuais; e) especialistas no assunto, principalmente, aqueles que possuem algum conhecimento de indexação ou documentação.

4) Especificidade – o uso de termos altamente específicos deve ser estrito à área principal do campo temático coberto pelo tesauro, uma vez que o uso excessivo em áreas adjacentes pode resultar em um tesauro difícil de manejar e desbalanceado. Em um serviço de indexação que trabalhe com documentos relacionados com mais de um campo de conhecimento, pode ser necessário desenvolver vários tesouros especializados relacionados e compatíveis a um tesauro geral que tenha um nível mais baixo de especificidade e seja produzido pelo mesmo serviço.

5) Inclusão e exclusão de termos – a) em geral, os termos e suas relações devem refletir a maneira pela qual os especialistas da matéria utilizam a linguagem no campo coberto pelo tesauro. Se um termo de uso comum é fatorado tendo em vista a consistência lógica, uma referência deve ser feita ao termo não-fatorado, mas que não é sua forma preferida; b) quando as hierarquias estão sendo estabelecidas, especialmente durante as etapas iniciais do método indutivo, são frequentemente incluídos no tesauro termos que ainda não tenham sido usados na indexação, devido ao fato de poderem proporcionar pontos úteis de acesso (por exemplo, como termos genéricos) e porque também têm valor potencial como termos preferidos. Esses termos devem ser indicados através de um símbolo especial no arquivo de termos ou por uma frase como “ainda não utilizado”. Esse símbolo ou frase deve ser suprimido quando o termo for utilizado na indexação de um documento; c) termos muito abrangentes, sem conteúdo semântico bem definido, e termos que a existência mostra não ser muito frequente, devem ser candidatos à exclusão, pois ambas as classes de termos são geralmente ineficazes

na recuperação. Em alguns casos, os termos muito abrangentes podem ser proveitosamente substituídos por dois ou mais termos de maior especificidade. Se um termo que aparece anteriormente em registro de indexação for excluído, a data de exclusão do termo deve ser registrada.

#### 6) Uso de equipamentos de processamento automático de dados –

a) nesta norma sustentou-se que a seleção de termos, a escolha da forma etc. requerem decisões intelectuais por parte dos indexadores. Pode-se empregar a ajuda de máquinas para as seguintes tarefas: 1) identificar automaticamente os termos candidatos a partir de textos legíveis à máquina, por exemplo, títulos ou resumos. O número de termos potenciais deve primeiramente ser reduzido pelo uso de uma *stop list* [palavras vazias]. Os termos restantes devem ser comparados com aqueles já registrados no tesouro. Os termos não registrados podem ser considerados como candidatos potenciais à inclusão; 2) nos sistemas mecanizados de indexação, pode-se registrar automaticamente com que frequência um termo tem sido utilizado na indexação (e também na formulação de perguntas). Termos com índices excepcionalmente altos ou baixos devem ser considerados como candidatos à exclusão; b) o equipamento e o suporte lógico disponíveis podem ocasionalmente limitar a reprodução de um tesouro a partir de arquivos conservados em máquina. As seguintes características têm mais probabilidade de serem afetadas: 1) o número de caracteres permitidos em um termo; 2) o conjunto de caracteres, inclusive sinais de pontuação, disponíveis na impressora; 3) o número de níveis hierárquicos permitido em um tesouro sistemático. Os indexadores devem ter acesso a equipamentos e programas capazes de satisfazer as necessidades especiais de um tesouro. Não deveria ser necessário adaptar o tesouro às limitações de uma máquina inadequada.

#### 7) Forma e conteúdo de um tesouro –

a) não é possível estipular uma forma-padrão de diagramação em um tesouro impresso devido a uma série de apresentações possíveis. Entretanto, as seguintes partes devem ser claramente distinguidas: 1) página de rosto, 2) sumário, 3) introdução, 4) apresentação sistemática ou gráfica (quando for o caso), 5) seção alfabética, b) todos os tesouros devem conter uma introdução abrangente que mencione claramente: 1) o propósito do tesouro; 2) o(s) campo(s) temático(s), identificando separadamente as áreas centrais e relacionadas; 3) os significados de todas as convenções e abreviaturas; 4) o número total de termos, com subtotais de termos preferidos e não-preferidos; 5) os

critérios adotados para selecionar as formas preferidas dos termos de indexação e para estabelecer suas inter-relações; 6) as regras de registro empregadas, seguindo (e citando), quando possível, uma norma nacional ou internacional apropriada; 7) os significados de todos os sinais de pontuação utilizados de maneira não convencional; 8) um resumo da política de atualização e o nome e endereço do serviço responsável ao qual devem ser enviados comentários e sugestões; 9) a data de inclusão do último termo. Sempre que possível esses pontos devem ser ilustrados com exemplos.

8) Outras questões editoriais – a) notificação da intenção. Quando um serviço decidir publicar um novo tesouro, deve anunciar tal intenção em uma publicação profissional apropriada; b) não duplicação de esforços. O compilador deve determinar mediante referência a uma das centrais de distribuição indicadas a seguir, se um tesouro cobre uma só área de conhecimento ou se há superposição de áreas. Uma completa duplicação do campo temático coberto é rara, mas pode ser útil consultar preliminarmente um ou mais tesouros em campos relacionados; c) depósito em uma área central de distribuição. Uma cópia da primeira edição e das edições subsequentes deve ser depositada na central nacional apropriada e em uma das centrais internacionais abaixo indicadas: 1) Para tesouros em língua inglesa, incluindo tesouros multilíngues que contenham seções em inglês: *Thesaurus Clearinghouse*; 2) Para tesouros em outras línguas: *Centralny Instytut Informacji Naukowo Technicznej*; 3) Para tesouros em língua portuguesa, incluindo tesouros multilíngues que contenham seções em português: Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT); d) procedimentos de revisão. Recomenda-se realizar uma prova piloto do tesouro antes de sua publicação. Uma versão provisória deve ser distribuída a um grupo seletivo de usuários, e qualquer sugestão para mudanças de termos e/ou suas inter-relações devem ser consideradas pelo editor. Também deve-se fazer uma revisão completa do tesouro depois de um período de uso e, posteriormente, a intervalos regulares, levando em conta todas as mudanças no uso dos termos dentro do campo coberto pelo tesouro (UNESCO, 1993).

c) *Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies* (ANSI/NISO Z39.19, 2005)

Essa Norma propõe linhas gerais para a Construção, Formato e Gestão de Vocabulários Controlados Monolíngues e constitui uma referência normativa importante, atualizada, especialmente, a orientações sobre a questão de tesouros na Web. A construção de vocabulários controlados é um processo intenso e demorado, especialmente se o domínio a ser coberto for amplo e a terminologia em uso for rica e complexa. O trabalho envolvido é justificado porque o uso de vocabulários controlados contribui para assegurar consistência em indexação e promover uma recuperação mais satisfatória (ANSI/NISO, 2005).

A publicação encontra-se organizada em 11 seções. Nas seções enumeradas de 1 a 11, nessa sequência, verificamos: Introdução; Escopo; Normas Referenciadas; Definições, Abreviaturas e Símbolos; Vocabulários Controlados (Propósitos, Conceitos, Princípios e Estrutura); Seleção de Termos, Aplicação e Formas de apresentação; Termos compostos; Relacionamentos (ligações semânticas, e relacionamentos de equivalência, hierárquicos e associativos); Apresentação de Vocabulários Controlados; Interoperabilidade; e Construção, Teste (Avaliação), Manutenção, e Sistemas de Gestão. Vale destacar que essa Diretriz contempla na seção 10, um conteúdo importante sobre a interoperabilidade, desde a questão da necessidade de Interoperabilidade, fatores que afetam a interoperabilidade até o processo de fusão de vocabulários controlados.

Entendemos oportuno, nesse momento, enfatizar que, embora, as Diretrizes analisadas se refiram, de um modo geral, à construção, formato e gestão de vocabulários controlados/tesouros monolíngues, julgamos que tais Diretrizes apresentam normas e recomendações com grande similaridade às consideradas como fundamentais para a construção de um tesouro geral. Isso posto, consideramos que pode ser aplicável a tesouros que apresentem seus termos, de modo equivalentes em mais de uma língua.

Em razão disso, identificamos na Seção 11, que trata da “Construção, Teste, Manutenção, e Sistemas de Gestão”, no item I, dedicado à Construção, uma sequência de etapas que se refere, especificamente, ao processo de construção de vocabulários controlados/tesouros. Além disso, na seção 11,



verificamos: no item II, Teste e Avaliação - orientações sobre métodos de teste e critérios de avaliação de vocabulários controlados; no item III, Manutenção – recomendações e sugestões para a atualização e manutenção de vocabulário controlado, considerando sua aplicação na web; item IV, Sistemas de Gestão apresenta possibilidades de aplicação de um sistema de gerenciamento de vocabulário controlado.

Para efeito desse estudo, abordamos, especificamente, o **item I**, que trata de Construção de vocabulário, da **seção 11** que apresenta uma sequência de oito etapas com similaridade às etapas de construção de tesouros das Diretrizes IBICT (1984) e UNESCO (1993). Apresentamos a seguir uma síntese a partir do item **I. Construção:**

1) Evitar duplicação de vocabulários - Antes de iniciar a construção de um vocabulário controlado (tesauro), deve-se averiguar se um vocabulário controlado existente cobre o mesmo domínio ou um domínio de conhecimento coincidente. Consultar as referências, listas e portais da Web sobre vocabulários controlados para ter acesso a um ou mais vocabulários controlados em campos relacionados, o que pode frequentemente servir como um ponto de partida útil. Por exemplo, um desenvolvedor de vocabulário controlado no campo da Lei de Malversação deveria iniciar por consultar vocabulários controlados existentes na área de medicina e direito que irão, com certeza, conter muitos dos termos e relações necessários para um vocabulário controlado (tesauro) mais especializado e interdisciplinar.

2) Determinar a estrutura e formatos de apresentação – Se possível, a estrutura do vocabulário controlado/tesauro (estrutura genérica e específica, apresentação hierárquica e/ou gráfica) e formato de apresentação (impresso, online ou disponível na web) deveriam ser decididos antes que os termos fossem coletados e considerados como candidatos para inclusão. O formato de apresentação do vocabulário controlado afeta os tipos de remissões e indicadores de relação que são fornecidos. Por exemplo, se uma estrutura de árvore com notação de classificação for adotada, o vocabulário controlado pode não incluir referências BT/NT. Essa Norma orienta que o formato de apresentação deveria ser estabelecido antes que as relações entre termos fossem construídas, embora seja possível converter um formato de exibição de vocabulário controlado para outro.

3) Métodos de construção – São recomendadas três abordagens iniciais para a construção de vocabulário controlado/tesauro: a) Abordagem de Comitê; b) Abordagem Empírica; c) Combinação de Métodos. Contudo, as Diretrizes indicam mais uma abordagem: d) Assistência da Máquina.

a) Abordagem de Comitê - especialistas no domínio de assunto do vocabulário controlado fazem uma lista de termos chave na área e indicam as relações entre eles com assistência de especialistas em criação de vocabulário controlado. Listas de termos podem ser retiradas de várias fontes ou submetidas a vários usuários/comunidades. Dois métodos principais para criar vocabulários controlados por Comitê têm sido utilizados: 1) *Top down* (Descendente) – os termos mais genéricos são identificados primeiro e então termos mais específicos são selecionados para atingir o nível desejado de especificidade. As estruturas hierárquicas e relações necessárias são estabelecidas à medida que o trabalho evolui; 2) *Bottom up* (Ascendente) – esse caso ocorre frequentemente quando listas de termos forem derivadas de um *corpus* de objetos de conteúdo e devem então ser incorporadas em um vocabulário controlado. Como no caso acima, as estruturas e relações hierárquicas necessárias são criadas à medida que o trabalho avança, mas inicia-se com os termos que têm uma abrangência mais específica e segue para os termos mais genéricos. Se um novo vocabulário controlado/tesauro for construído, a abordagem “descendente” é preferida. No caso de um vocabulário controlado/tesauro já publicado, a abordagem “ascendente” é mais usada para o acréscimo de novos termos para cobrir novos conceitos.

b) Abordagem Empírica - a abordagem empírica evidencia dois métodos básicos: 1) Método Dedutivo – termos são extraídos de objetos de conteúdo (por humanos ou computadores), opcionalmente durante um estágio preliminar de indexação, mas nenhuma tentativa é feita para controlar o vocabulário, nem para determinar relações entre termos, até que um número suficiente de termos tenha sido coletado. Todos os termos são revisados por um grupo de especialistas, preferencialmente, composto de especialistas em informação e no assunto tratado. Eles deveriam primeiro identificar termos que representam as classes mais amplas, e então alocar os termos remanescentes para aquelas classes segundo suas relações lógicas, a fim de que as hierarquias tendam a ser estabelecidas sobre uma base da mais genérica para a mais específica. O controle de vocabulário deveria ser desenvolvido no estágio em que hierarquias e outras relações são estabelecidas,

seguindo os princípios descritos nas seções que tratam de Termos compostos e de Relacionamentos entre os termos; 2) Método Indutivo - novos termos são selecionados para potencial inclusão no vocabulário controlado/tesauro como eles forem encontrados em objetos de conteúdo. O controle de vocabulário é aplicado desde o início. Se o vocabulário em construção tiver algum tipo de arranjo hierárquico, cada termo admitido é designado como um membro de uma ou mais classes genéricas que são construídas sobre uma base *ad hoc* em um estágio anterior. O vocabulário controlado é, portanto, estabelecido sobre uma base de termo do mais específico para o mais genérico. A construção de vocabulário controlado é considerada do início como uma operação contínua. Assistência de especialistas no assunto é altamente recomendada. Esses especialistas podem servir como membros de uma comissão ou comitê editorial/formal;

c) Combinação de Métodos - na prática, é possível empregar mais de uma dessas abordagens em um estágio ou outro durante a construção de um vocabulário controlado. Por exemplo, hierarquias e outras relações entre termos que foram primeiro estabelecidos indutivamente poderiam mais tarde ser examinadas a partir de um ponto de vista dedutivo. Ambas as técnicas são essencialmente empíricas. Deveria ser aceito, desde o início, que algumas decisões com relação aos termos e suas inter-relações que foram feitas durante os primeiros estágios de compilação, possam ser revisadas como uma experiência posterior. Os compiladores deveriam verificar os termos e hierarquias frequentemente para assegurar uma aplicação consistente de princípios em tais procedimentos como: o estabelecimento de inter-relações entre termos e a divisão de termos compostos;

d) Assistência da Máquina – assume-se de acordo com essa Norma que a construção de vocabulário controlado/tesauro requer decisões intelectuais. A assistência da máquina pode ser empregada, no entanto, para tarefas de identificação de termo tais como as seguintes:

- i) Identificação de termos candidatos – termos candidatos podem ser identificados automaticamente a partir de texto legível por máquina (por exemplo: títulos e/ou resumos). O número de termos potenciais deveria ser reduzido pelo uso de palavras vazias ou uma lista vazia (palavras funcionais – artigos, conjunções, e preposições tais como “um/uma”, “e” e “para” – além de outras palavras consideradas sem valor para a

recuperação). Remover as palavras vazias da consideração é geralmente o primeiro passo para a indexação auxiliada pela máquina. Os termos remanescentes deveriam ser comparados com aqueles já registrados no vocabulário controlado. Termos que não foram registrados poderiam ser considerados como candidatos à inclusão;

- ii) Registro da frequência de designação (indicação) de termo – Em sistemas de indexação computadorizados, a frequência com a qual um termo é usado em indexação, pode ser registrada automaticamente. Termos com resultados altos ou baixos podem ser considerados candidatos à modificação ou exclusão;
- iii) Registro de termos a partir das consultas do usuário – Termos encontrados em consultas de usuário que não combinam com um ou mais termos de entrada também podem ser considerados para inclusão especialmente quando um dado termo ocorre em múltiplas consultas. Para assegurar privacidade, os usuários não deveriam ser identificados quando se rastreia e coleta consultas de usuário. Termos que ocorrem em linguagem natural e são formas comuns de descrever um conceito, são incluídos no vocabulário controlado/tesauro como termos de entrada. Um termo de entrada é uma palavra ou frase que é um nome comumente usado ou rótulo para um conceito, mas não é o termo preferido selecionado para representar o conceito.

Por exemplo, referência cruzada auxiliada por máquina para termo preferido: Síndrome da imunodeficiência adquirida USE AIDS

NOTA: No exemplo, a síndrome da imunodeficiência adquirida é um termo de entrada. Enquanto **AIDS** é o termo preferido.

4) Registros de termos – um registro individual deveria ser criado para todo termo e opcionalmente para todo termo de entrada, assim que ele for admitido no vocabulário controlado/tesauro. Registros para termos de entrada podem incluir notas de fonte bem como a data de admissão no vocabulário controlado/tesauro. Para termos, o registro pode conter quaisquer ou todos os seguintes elementos: termo; fonte(s) consultada(s) para termos e termos de entrada.

NOTA: Este campo é especialmente importante para neologismos ou termos não familiares. Pode incluir citações para fontes publicadas ou os nomes de autoridades pessoais consultadas; nota de escopo; referência UP (Usado Para) – para indicar quais sinônimos, sinônimos aproximados e outras expressões cobertas pelo termo e variações que não são apresentadas. Por exemplo: erros comuns de ortografias, termos genéricos, termos específicos, termos relacionados, relações estabelecidas localmente, categoria ou número de classificação, nota histórica, incluindo, no mínimo a data da inclusão do termo, bem como o registro das mudanças, se houver.

5) Verificação de termo – antes de um termo ser admitido em um vocabulário controlado/tesauro, ele deveria ser validado de acordo com as regras propostas, nas seções 6 e 7, das Diretrizes, com relação à abrangência, forma, e escolha de termos. O compilador deveria também revisar as relações entre cada novo termo e outros termos na hierarquia para a qual ele for designado. Os seguintes tipos de autoridade dentro do domínio de assunto deveriam ser verificados antes que os termos candidatos sejam aceitos para inclusão: a) dicionários técnicos, glossários, monografias eruditas, textos de referência e enciclopédias; b) vocabulários controlados existentes; c) esquemas de classificação; d) em casos problemáticos, o conselho de especialistas deveria ser consultado sobre a seleção de um termo a partir de formas variantes de termos.

6) Termos candidatos – termos candidatos (também chamados de termos provisórios) são termos propostos que não passaram por todos os procedimentos de validação. Esses termos deveriam ser marcados por um símbolo especial ou frase no registro do termo. Assim que um termo candidato é aprovado como termo, o símbolo ou frase deve ser excluído. Em um sistema *on line* no qual o vocabulário controlado está integrado a uma base de dados única, os termos candidatos, geralmente, não são mostrados ao usuário. Em vocabulários controlados que não estão ligados dessa forma, os termos candidatos podem ser exibidos.

7) Níveis de especificidade – a adição de termos específicos é normalmente restrita à área central do campo de assunto coberto por um vocabulário controlado/tesauro, porque a proliferação de tais termos em áreas periféricas é susceptível de conduzir a um vocabulário controlado que é difícil de gerenciar. Embora o custo de armazenagem em computador de um vocabulário controlado muito grande possa ser insignificante, o custo humano para estabelecer

relações entre inúmeros termos na periferia do domínio é alto. Em uma organização que lida com objetos de conteúdo que cobrem mais de um domínio de conhecimento, pode ser necessário desenvolver um número maior de vocabulários controlados especializados, cada um, ligado a, e compatível com, um vocabulário controlado geral que tem um nível inferior de especificidade, produzido pela mesma organização. Um modelo para isso pode ser encontrado no Macrotesauro OECD – Vocabulário das Nações Unidas para o Processamento da Informação no Campo do Desenvolvimento Econômico e Social, ou no Metatesauro UMLS – Sistema Unificado de Linguagem Médica.

8) Termos não designados – quando hierarquias são estabelecidas em um vocabulário controlado/tesauro, os termos que ainda não foram usados na indexação são, frequentemente, admitidos no vocabulário controlado com o fundamento de que eles são necessários para completar uma hierarquia (por exemplo, como termos genéricos) e que têm valor potencial como termos de indexação.

Diante do exposto, apresentamos o estudo realizado para verificar quantas e quais são as etapas do processo de construção de tesouros à luz das Diretrizes IBICT (1984); Diretrizes UNESCO (1993) e Diretrizes ANSI/NISO (2005). Para melhor visualização, elaboramos um quadro síntese, demonstrativo das etapas de construção de tesouros, conforme as diretrizes acima mencionadas (Vide Quadro 6 no Apêndice A).

Vale dizer que as Diretrizes analisadas compreendem, igualmente, um conjunto de oito etapas, inerentes ao processo de construção de um tesauro que, supomos pela proximidade de seus conteúdos e para efeito dessa apresentação, neste instante, foram agrupadas em 7 categorias temáticas, denominadas conforme consta em cada etapa e a(s) sua(s) norma(s) correspondente(s), indicada(s) entre parênteses, a saber:

**1.** Trabalho preliminar/orientações gerais; **2.** Métodos de compilação; **3.** Registro de termos; **4.** Verificação de termos/Admissão e exclusão de termos; **5.** Especificidade; **6.** Uso de equipamento automático de processamento de dados e **7.** Forma e conteúdo de um tesauro.

## CATEGORIAS TEMÁTICAS

### 1. TRABALHO PRELIMINAR / ORIENTAÇÕES GERAIS

As orientações sobre essa atividade encontram-se expostas na **etapa 1** das Diretrizes IBICT (1984) que expõe duas recomendações: a) “Comunicação de intenção – quando um serviço decide a preparar um tesouro novo, essa intenção deve ser comunicada em periódicos apropriados (por ex., num periódico de documentação, de classificação)”; b) Não duplicação de trabalho – a verificação da existência de tesouro que cobre a mesma área de conhecimento ou parte dele, pode ser útil como ponto de partida. Todavia, orientam que “cada tesouro deve refletir as necessidades de uma comunidade específica de usuários” (IBICT, 1984, p. 52). Ainda na **etapa 1**, nas Diretrizes ANSI/NISO (2005), ao alertar que deve-se “Evitar duplicação de vocabulários” - antes de iniciar a construção de um vocabulário controlado (tesouro), deve-se averiguar se um vocabulário controlado existente cobre o mesmo domínio ou um domínio de conhecimento coincidente. Enquanto que, nas Diretrizes UNESCO (1993), essa preocupação é sinalizada na **etapa 8**, denominada de “Outras questões editoriais” com as seguintes recomendações: a) notificação da intenção - em uma publicação profissional apropriada; b) não duplicação de esforços – consultar preliminarmente um ou mais tesouros em campos relacionados. Além disso, fazer uma revisão completa do tesouro depois de um período de uso e, depois, a intervalos regulares, levando em conta todas as mudanças no uso dos termos dentro do campo coberto pelo tesouro.

### 2. MÉTODOS DE COMPILAÇÃO (vide Quadro 7, Apêndice B)

As Diretrizes UNESCO (1993) iniciam as etapas de construção do tesouro com a indicação, na **etapa 1**, de três métodos para o processo de compilação de termos, denominados respectivamente de: Método dedutivo, Método indutivo e Combinação de métodos. No que se refere ao primeiro, “Método dedutivo” - os termos são extraídos de documentos durante uma etapa preliminar de indexação. Não é feito controle do vocabulário, nem mesmo para determinar as relações entre os termos, a não ser que um número suficiente de termos tenha sido coletado. Nesse caso, todos os termos são revistos por um grupo de especialistas no assunto. Primeiro, devem ser identificados os termos que representam categorias genéricas, e os termos restantes devem ser relacionados nessas categorias

conforme as suas relações lógicas. Deste modo, as categorias tendem a ser estabelecidas sobre uma base de termo genérico a específico. O controle do vocabulário deve ser efetuado no estágio em que as categorias são estabelecidas, observando os princípios descritos nas seções de 5 a 7 dessa Norma. A seção 5 trata do Controle de Vocabulário. As seções 6 e 7 abordam Termos de Indexação e Termos Compostos, respectivamente. No segundo, “Método indutivo” - novos termos são admitidos no tesouro tão logo apareçam nos documentos. O controle do vocabulário é feito desde a etapa inicial. Cada termo, à medida que é admitido, é designado como membro de uma ou mais categorias genéricas constituídas sobre uma base *ad hoc*. O tesouro é, portanto, organizado sobre uma base de termo específico a genérico. A elaboração do tesouro é, desde o princípio, uma operação contínua e, apesar de ser necessária a assistência de técnicos especialistas, eles não precisam fazer parte de um grupo editorial formal ou comitê. No terceiro, “Combinação de Métodos” - “na prática, é possível empregar-se tanto o método dedutivo como indutivo em uma ou outra etapa da elaboração do tesouro”. Por exemplo, um grupo editorial composto por indexadores e especialistas pode primeiramente estabelecer as categorias dos termos indutivamente e depois examiná-las pelo ponto de vista dedutivo. Ambas as técnicas são essencialmente empíricas e, desde o início, deve-se ter em mente que algumas decisões tomadas durante a etapa de compilação podem ser revistas, à medida que se adquire maior experiência. A adição de novos termos a uma linguagem de indexação frequentemente esclarece as decisões tomadas anteriormente quanto aos termos e suas inter-relações. Os compiladores devem revisar frequentemente os termos e as hierarquias para assegurar que as relações entre os termos, a fatoração de termos compostos, e outros, seguem os princípios gerais que têm sido aplicados. Recomenda-se particularmente que se busque o auxílio de especialista na escolha dos termos sinônimos e formas variantes dos termos.

As Diretrizes IBICT (1984) destacam os “Métodos de compilação” **na etapa 2**, identificando os seguintes métodos: “Método dedutivo”, os termos são extraídos da literatura durante o estágio preliminar de indexação; entretanto, até que um número suficiente de termos tenha sido coletado, todos os termos são considerados por uma comissão formada de preferência por indexadores e especialistas que decidem primeiro quais termos representam as classes mais amplas, fixando os demais termos nessas classes e relacionando-os logicamente



com base em suas conceituações; “Método indutivo”, novos termos são admitidos no tesauro logo que são encontrados na literatura, sendo cada termo designado como membro de uma ou mais classes, estabelecidas em bases *ad hoc* durante o ato de indexação. A elaboração do tesauro é considerada, desde o início, como uma operação contínua e, embora se busque assistência de especialistas sempre que for preciso, tais profissionais não são membros de uma comissão formal. Na oportunidade, ainda que tenham sido evidenciadas as duas abordagens descritas acima, nessa publicação aparece mais uma abordagem de compilação, denominada “Método relacional”. Em notas de rodapé, consta a seguinte explicação: “\*Dahlberg aponta um terceiro tipo – a abordagem relacional – que parte do aspecto formal, categorial” (IBICT, 1984, p. 52). Além de “Combinação de métodos”. Durante a sequência de etapas de construção do tesauro, empregam-se, de modo igual, os métodos dedutivo e indutivo. Por exemplo: as classes de termos que tenham sido estabelecidas de início de forma indutiva podem ser revistas mais tarde por uma comissão de indexadores e especialistas da área do tesauro. Ambas as técnicas são empíricas. Nesse sentido, decisões tomadas nos primeiros estágios da compilação podem ser revistas, à medida que os indexadores ganham mais experiência. A inclusão de novos termos, a conferência dos relacionamentos entre os termos, a fatoração, as notas de aplicação devem seguir uma linha lógica e consistente. A assessoria de especialistas na escolha entre termos sinônimos e outras formas variantes é recomendada. As Diretrizes ANSI/NISO (2005) apresentam reunidos, **na etapa 3**, os “Métodos de construção/compilação”. Nesses Métodos de construção, são recomendadas três abordagens iniciais para a construção de vocabulário controlado/tesauro: a) Abordagem de Comitê; b) Abordagem Empírica; c) Combinação de Métodos. Contudo, as Diretrizes indicam mais uma abordagem d) Assistência da Máquina.

No que se refere à “Abordagem de Comitê”, especialistas no domínio de assunto do vocabulário controlado fazem uma lista de termos chave na área e indicam as relações entre eles, com assistência de especialistas em criação de vocabulário controlado. Listas de termos podem ser retiradas de várias fontes ou submetidas a vários usuários/comunidades. Dois métodos principais para criar vocabulários controlados por Comitê têm sido utilizados: 1) *Top-Down* (Descendente) – os termos mais genéricos são identificados primeiro e então termos mais específicos são selecionados para atingir o nível desejado de especificidade.

As estruturas hierárquicas e relações necessárias são estabelecidas à medida que o trabalho evolui; 2) *Bottom up* (Ascendente) – esse caso ocorre frequentemente quando listas de termos forem derivadas de um *corpus* de objetos de conteúdo e devem então ser incorporadas em um vocabulário controlado. Como no caso acima, as estruturas e relações hierárquicas necessárias são criadas à medida que o trabalho avança, mas inicia-se com os termos que têm uma abrangência mais específica e segue para os termos mais genéricos. Se um novo vocabulário controlado/tesauro será construído, a abordagem “descendente” é preferida. No caso de um vocabulário controlado/tesauro já publicado, a abordagem “ascendente” é mais usada para o acréscimo de novos termos para cobrir novos conceitos.

Na “Abordagem Empírica”, evidenciam-se dois métodos básicos:

1) Método Dedutivo – termos são extraídos de objetos de conteúdo (por humanos ou computadores), opcionalmente durante um estágio preliminar de indexação, mas nenhuma tentativa é feita para controlar o vocabulário, nem para determinar relações entre termos, até que um número suficiente de termos tenha sido coletado. Todos os termos são revisados por um grupo de especialistas, preferencialmente, composto de especialistas em informação e no assunto tratado. Eles deveriam primeiro identificar termos que representam as classes mais amplas, e então alocar os termos remanescentes para aquelas classes segundo suas relações lógicas, a fim de que as hierarquias tendam a ser estabelecidas sobre uma base da mais genérica para a mais específica. O controle de vocabulário deveria ser desenvolvido no estágio em que hierarquias e outras relações são estabelecidas, seguindo os princípios descritos nas seções que tratam de Termos compostos e de relacionamentos entre os termos;

2) Método Indutivo - novos termos são selecionados para potencial inclusão no vocabulário controlado/tesauro como eles forem encontrados em objetos de conteúdo. O controle de vocabulário é aplicado desde o início. Se o vocabulário em construção tiver algum tipo de arranjo hierárquico, cada termo admitido é designado como um membro de uma ou mais classes genéricas que são construídas sobre uma base *ad hoc* em um estágio anterior. O vocabulário controlado é, portanto, estabelecido sobre uma base de termo do mais específico para o mais genérico. A construção de vocabulário controlado é considerada do início como uma operação contínua. Assistência de especialistas no assunto é

altamente recomendada. Esses especialistas podem servir como membros de uma comissão ou comitê editorial/formal.

Na “Combinação de Métodos,” é possível empregar mais de uma dessas abordagens em um estágio ou outro durante a construção de um vocabulário controlado. Por exemplo, hierarquias e outras relações entre termos que foram primeiro estabelecidos indutivamente poderiam mais tarde ser reexaminadas a partir de um ponto de vista dedutivo. Ambas as técnicas são essencialmente empíricas. Deveria ser aceito, desde o início, que algumas decisões com relação aos termos e suas inter-relações, que foram feitas durante os primeiros estágios de compilação, possam ser revisadas como uma experiência posterior. Os compiladores deveriam verificar os termos e hierarquias frequentemente para assegurar uma aplicação consistente de princípios em tais procedimentos como: o estabelecimento de inter-relações entre termos e a divisão de termos compostos.

Na seqüência, com “Assistência da Máquina”, assume-se de acordo com essa Norma que a construção de vocabulário controlado/tesauro requer decisões intelectuais. A assistência da máquina pode ser empregada, no entanto, para tarefas de identificação de termo tais como as seguintes:

i) Identificação de termos candidatos – termos candidatos podem ser identificados automaticamente a partir de texto legível por máquina, (por exemplo: títulos e/ou resumos). O número de termos potenciais deveria ser reduzido pelo uso de palavras vazias ou uma lista vazia (palavras funcionais – artigos, conjunções, e preposições tais como “um/uma”, “e”, e “para” – além de outras palavras consideradas sem valor para a recuperação). Remover as palavras vazias da consideração é geralmente o primeiro passo para a indexação auxiliada pela máquina. Os termos remanescentes deveriam ser comparados com aqueles já registrados no vocabulário controlado. Termos que não foram registrados poderiam ser considerados como candidatos à inclusão;

ii) Registro da freqüência de designação (indicação) de termo – Em sistemas de indexação computadorizados, a freqüência com a qual um termo ter sido usado em indexação, pode ser registrada automaticamente. Termos com resultados altos ou baixos podem ser considerados candidatos à modificação ou exclusão;

iii) Registro de termos a partir das consultas do usuário – Termos encontrados em consultas de usuário que não combinam com

um ou mais termos de entrada também podem ser considerados para inclusão, especialmente quando um dado termo ocorre em múltiplas consultas. Para assegurar privacidade, os usuários não deveriam ser identificados quando se rastreia e coleta consultas de usuário. Termos que ocorrem em linguagem natural e são formas comuns de descrever um conceito, são incluídos no vocabulário controlado/tesauro como termos de entrada. Um termo de entrada é uma palavra ou frase que é um nome comumente usado ou rótulo para um conceito, mas não é o termo preferido selecionado para representar o conceito.

Por exemplo: referência cruzada auxiliada por máquina para termo preferido. Síndrome da imunodeficiência adquirida USE **AIDS**

NOTA: No exemplo, a síndrome da imunodeficiência adquirida é um termo de entrada. Enquanto **AIDS** é o termo preferido.

### 3. REGISTRO DE TERMOS

Para cada termo admitido no tesauro deve ser mantido um registro individual em ficha. É o que evidencia a **etapa 3** das Diretrizes IBICT (1984). Nas Diretrizes UNESCO (1993) essa exigência encontra-se descrita na **etapa 2**. Enquanto nas Diretrizes ANSI/NISO (2005) semelhante recomendação encontra-se sinalizada na **etapa 4**.

### 4. VERIFICAÇÃO DE TERMOS/ADMISSÃO E EXCLUSÃO DE TERMOS

Essas etapas apresentam conteúdo que se imbricam em duas etapas nas Diretrizes IBICT (1984). Observamos na **etapa 4** – “Verificação de termos”, que a forma correta de um termo deve ser verificada antes de sua admissão ao tesauro. O indexador deve rever também os relacionamentos entre cada novo termo e os termos e as hierarquias já registradas. Para os termos candidatos à inclusão no tesauro, consultar as fontes autorizadas como: dicionários e enciclopédias especializadas, tesouros e esquema de classificação, índice de periódicos especializados, tratados e manuais atualizados e especialistas no assunto, especialmente aqueles que possuem algum conhecimento de indexação ou

documentação. A **etapa 6** – “Admissão e exclusão de termos” evidencia que é importante assegurar que os termos e seus relacionamentos reflitam a linguagem usada pelos especialistas, além de outras orientações como: “o tesauro deve ser submetido a uma revisão técnica após determinado período de uso e, a partir daí, a intervalos regulares, levando em conta mudanças no uso de termos dentro do campo de cobertura do tesauro”. Nas Diretrizes UNESCO (1993), as orientações com respeito a termos candidatos à inclusão no tesauro, consultar as fontes autorizadas como: dicionários e enciclopédias especializadas, tesouros e esquema de classificação, índice de periódicos especializados, tratados e manuais atualizados e especialistas no assunto, principalmente aqueles que possuem algum conhecimento de indexação ou documentação encontram-se descritas na **etapa 3**. Ainda de acordo com as Diretrizes UNESCO (1993), a recomendação para a “Inclusão e exclusão de termos”, tais como: “os termos e suas relações devem refletir a maneira pela qual os especialistas da matéria utilizam a linguagem no campo coberto pelo tesauro,” são descritas na **etapa 5**. Nas Diretrizes ANSI/NISO (2005), a **etapa 5**, Verificação de termo, evidencia que “antes de um termo ser admitido em um vocabulário controlado/ tesauro, ele deveria ser validado de acordo com regras propostas”. Além de, na **etapa 6**, tratar sobre “Termos candidatos”, que são os termos propostos que não passaram por todos os procedimentos de validação. Esses termos deveriam ser marcados por um símbolo especial ou frase no registro do termo. Assim que um termo candidato é aprovado como termo, o símbolo ou frase deve ser excluído. Em um sistema online no qual o vocabulário controlado está integrado a uma base de dados única, os termos candidatos, geralmente, não são mostrados ao usuário. E na **etapa 8**, “Termos não designados” refere-se a quando hierarquias são estabelecidas em um vocabulário controlado/tesauro. Os termos que ainda não foram usados na indexação são, frequentemente, admitidos no vocabulário controlado com o fundamento de que eles são necessários para completar uma hierarquia (por exemplo, como termos genéricos) e de que têm valor potencial como termos de indexação.

## 5. ESPECIFICIDADE

Na **etapa 5**, as Diretrizes IBICT (1984) advertem “se um serviço de informação compreende mais de uma área de conhecimento, pode tornar-se necessário desenvolver tesouros específicos para as áreas, ligados a um tesauro

geral”. Nas Diretrizes UNESCO (1993), esse alerta encontra-se descrito, de modo semelhante, na etapa 4. Enquanto as Diretrizes ANSI/NISO (2005) apontam para a questão de “Níveis de especificidade”, na **etapa 7**, evidenciando que a adição de termos específicos é normalmente restrita à área central do campo de assunto coberto por um vocabulário controlado/tesauro, porque a proliferação de tais termos em áreas periféricas é susceptível de conduzir a um vocabulário controlado que é difícil de gerenciar. Em uma organização que lida com objetos de conteúdo que cobrem mais de um domínio de conhecimento, pode ser necessário desenvolver um número de vocabulários controlados especializados, cada um, ligado a, e compatível com, um vocabulário controlado geral que tem um nível inferior de especificidade, produzido pela mesma organização.

## 6. USO DE EQUIPAMENTO AUTOMÁTICO DE PROCESSAMENTO DE DADOS

As Diretrizes IBICT (1984), na **etapa 7**, recomendam o uso de equipamentos informáticos na: 1) seleção de termos: para a seleção automática de termos candidatos extraídos de títulos e resumos e determinação automática (nos sistemas automatizados de indexação) da frequência com que os termos são usados na indexação e nas perguntas de busca, 2) o uso do computador sofre restrições impostas por limitações do equipamento disponível ou do software: o número de caracteres permitido em um termo; o conjunto de caracteres, inclusive sinais de pontuação, disponíveis na impressora, o número de níveis de subordinação permitido em um tesauro sistemático. Nas Diretrizes UNESCO (1993), o uso de equipamento automático de processamento de dados encontra-se detalhado na **etapa 6**. Ainda que, apresente conteúdo semelhante, optamos por apresentá-los em razão de evidenciar pontos importantes para o processo de construção de tesouros. Quanto à seleção de termos: a escolha da forma etc. requer decisões intelectuais por parte dos indexadores. O equipamento e o suporte lógico disponíveis podem ocasionalmente limitar a reprodução de um tesauro a partir de arquivos conservados em máquina. Os indexadores devem ter acesso a equipamentos e programas capazes de satisfazer as necessidades especiais de um tesauro. Não deveria ser necessário adaptar o tesauro às limitações de uma máquina inadequada.

## 7. FORMA E CONTEÚDO DE UM TESAURO

As Diretrizes IBICT (1984), na **etapa 8**, orientam para as seguintes situações: 1) não é possível estabelecer um formato padrão de apresentação de tesouros impressos. Os seguintes elementos devem ser destacados: a) página de rosto; b) sumário; c) introdução; d) apresentação sistemática ou planigráfica (quando for possível) e e) parte alfabética; 2) todos os tesouros devem conter uma introdução abrangente que informe de maneira clara: a) finalidade; b) área do conhecimento que abrange, com as áreas de assunto, central e periférica, identificadas separadamente; c) significados de todas as convenções e abreviaturas; d) número total de termos e outros.

Nessa direção, as Diretrizes UNESCO (1993), na **etapa 7**, reforçam a forma de apresentação de um tesouro: a) as seguintes partes devem ser claramente distinguidas: 1) página de rosto; 2) sumário; 3) introdução; 4) apresentação sistemática ou gráfica (quando for o caso); 5) seção alfabética; b) todos os tesouros devem conter uma introdução abrangente que mencione claramente: 1) o propósito do tesouro; 2) o(s) campo(s) temático(s), identificando separadamente as áreas centrais e relacionadas; 3) os significados de todas as convenções e abreviaturas; 4) o número total de termos, com subtópicos de termos preferidos e não-preferidos; 5) os critérios adotados para selecionar as formas preferidas dos termos de indexação e para estabelecer suas inter-relações; 6) as regras de registro empregadas, seguindo (e citando), quando possível, uma norma nacional ou internacional apropriada; 7) os significados de todos os sinais de pontuação utilizados de maneira não convencional; 8) um resumo da política de atualização e o nome e endereço do serviço responsável, ao qual devem ser enviados comentários e sugestões; 9) a data de inclusão do último termo. Sempre que possível esses pontos devem ser ilustrados com exemplos.

Enquanto que, nas Diretrizes ANSI/NISO (2005), para determinar a estrutura e formatos de apresentação: a estrutura do vocabulário controlado/tesouro (estrutura genérica e específica, apresentação hierárquica e/ou gráfica) e formato de apresentação (impresso, online, ou disponível na web) deveriam ser decididos antes que os termos fossem coletados e considerados como candidatos para inclusão, por esse motivo, apresentam essa recomendação, logo, na **etapa 2**.

- **Síntese**

Indicamos, a seguir, as etapas da construção de tesouros obtidas com base no estudo realizado à luz das Diretrizes IBICT (1984); Diretrizes UNESCO (1993); e Diretrizes ANSI/NISO (2005). Apuramos um conjunto de oito etapas, na construção de um tesouro, conforme adiantamos anteriormente que, para efeito desse estudo, foram agrupadas em 7 categorias temáticas.

Segundo as normas documentárias para construção de tesouros:

1. Trabalho preliminar (orientações gerais);
2. Métodos de compilação;
3. Registro de termos;
4. Verificação de termos (Admissão e exclusão de termos);
5. Especificidade;
6. Uso de equipamento automático de processamento de dados;
7. Forma e conteúdo de um tesouro.

Em suma, o agrupamento por similaridades de etapas de construção de tesouros, conforme posto nas Diretrizes acima descritas, foi possível em razão de que as duas primeiras Diretrizes são fundamentadas em princípios da Norma ISO 2788 e são traduções, com adaptações para a língua portuguesa. Do mesmo modo, as Diretrizes da ANSI/NISO que indicam conteúdo mais abrangente e atualizado, ainda que de origem norte-americana, preserva seus direcionamentos pautados, também, nos preceitos da ISO 2788 (vide fls. 50-51).

Com respeito a essa diretriz, o primeiro argumento sustenta-se em razão de apresentar conteúdo sobre a complexidade estrutural da modalidade vocabulário controlado, abrangendo, além de tesouros, as estruturas consideradas menos complexas como: listas de palavras-chave, rede de sinônimos e taxonomias. Podemos perceber que a diferença entre essas estruturas reside na complexidade envolvida no processo de sua construção. O segundo argumento tem relação com o conteúdo atualizado. Entendemos que essa Norma compreende um período de aproximadamente duas décadas de mudanças tecnológicas e de terminologias.



#### 2.4.2 Etapas de construção de tesouro segundo os autores

Para uma melhor compreensão do processo de construção de tesouro, faz-se necessário o estudo de referências fundamentais sobre o desenvolvimento desse processo, para análise dos métodos existentes, dentre os quais o que se mostra mais adequado ao processo de construção de tesouros, além da identificação das etapas desenvolvidas para a sua construção. Desse modo, evidenciamos estudos realizados por diversos autores para a construção de um tesouro. Como um ponto de referência ao processo de identificação das etapas de construção de tesouros, destaca-se inicialmente:

Fujita (1992), em conformidade com as Normas ISO 2788 (1986), argumenta que o processo de construção de tesouros “compreende a observância de uma sequência lógica de etapas, comuns à maioria dos métodos utilizados para construção de diferentes tesouros”. A autora supracitada pontua e denomina uma sequência de etapas para a construção de tesouros, a saber: “a) registro e seleção dos termos compilados; b) estruturação de conceitos com controle terminológico dos termos; c) estruturação automática das partes alfabética e sistemática do tesouro: uso do programa TECER” (FUJITA, 1998, p. 109).

Aitchison; Gilchrist (1979) apresentam a compilação de termos após o estabelecimento das divisões básicas da estrutura de categorias. Cabe salientar que esses autores, embora não explicitem métodos de compilação, sugerem que a seleção de termos seja realizada após o estabelecimento das categorias, portanto recomendam o método indutivo. O tesouro, entendido em finais da década de 70 como um sistema de conceitos, agrega esse princípio através da elaboração do *Thesourofacet* em 1969.

Devadason (1985 apud FUJITA, 1992) define o CLASSAURUS como um esquema hierárquico facetado de termos com controle de vocabulário. É um sistema de termos que possui lista hierárquica em separata com as categorias elementares: disciplina, entidade, propriedade e ação, juntos com suas respectivas espécies/tipos, partes e modificadores especiais. Também são apresentadas tabelas separadas para os modificadores comuns: forma, tempo, meio e lugar. Cada um dos termos dessas listas hierárquicas está enriquecido com sinônimos, quase sinônimos etc. A lista hierárquica constituindo a parte sistemática está suplementada pelo

índice alfabético das cadeias de entrada. CLASSAURUS é usado na formulação de cabeçalhos de assunto em geral e, em particular, cabeçalhos de assunto de acordo com o postulado baseado na linguagem de indexação de assunto permutado (POPSI). Devadason (1985) indica três aspectos básicos para a construção de tesouros: 1) seleção de palavras chave; 2) estabelecimento de categorias elementares; 3) estabelecimento de interrelações das palavras chave. A seleção de termos é feita antes da categorização. Concluímos ser o método dedutivo o sugerido pelo autor.

Robredo (2005) cita dois métodos que são utilizados para a construção de vocabulários controlados e, mais especificamente, de listas de autoridade: a) Método Evolutivo - consiste na observação da realidade terminológica e em certo estágio, detém a evolução do vocabulário: o material reunido pelos indexadores, depois de trabalhar certo tempo com uma quantidade importante de documentos, é analisado, depurado e organizado para elaborar a lista de descritores; b) Método Consensual - consiste na elaboração do vocabulário a partir de um consenso entre especialistas. As duas abordagens citadas por Robredo (2005), evidenciam que o método evolutivo e o método consensual correspondem às denominações propostas na normalização formulada pela UNESCO para a construção de tesouros monolíngues.

O autor se refere também às duas abordagens: o método analítico e o método *Gestalt*. Explica que esses dois métodos são considerados empíricos, porque não dependem de uma análise linguística minuciosa. “O método analítico implica a análise do conteúdo dos documentos e a seleção dos termos significativos. É aplicado com vantagem em áreas especializadas do conhecimento”. Já “o método *Gestalt* utiliza a experiência e o conhecimento de um grupo de especialistas que selecionam e ordenam os termos a partir de elementos já estruturados, como índices [...] e outras fontes de referência”. É aplicado com frequência em áreas do conhecimento que cobrem várias disciplinas (ROBREDO, 2005, p. 163).

Lancaster (1987, p. 19, 72) destaca que na construção de um tesouro “podem ser identificadas quatro etapas fundamentais: 1. Coleta de termos; 2. Ordenação dos termos; 3. Produção de uma estrutura final; e 4. Impressão e apresentação”. O autor dedica atenção especial à primeira delas - a coleta de termos - que consiste na coleta da matéria-prima, ou seja, a terminologia do assunto a cobrir.

Nessa direção, Lancaster apresenta duas abordagens: 1) A abordagem de cima para baixo (*top down*) ou dedutiva – essa abordagem, de modo frequente, exige o trabalho em equipe. As categorias principais de termos são identificadas e subdivididas de cima para baixo. Por exemplo: um grupo de bibliotecários, construindo um tesouro sobre Ciência da Informação, reconhece a necessidade de uma categoria tipos de bibliotecas, adotando bibliotecas como o cabeçalho mais amplo, começa a subdividir essa categoria; 2) A abordagem empírica, de baixo para cima (*bottom up*) ou indutiva – essa abordagem “é o oposto e tende a ser mais confiável”. Os termos que ocorrem na área são coletados de fontes diferentes. Por exemplo: uma subdivisão de uma categoria brinquedos por cores, azul, verde, ainda que teoricamente possível, pode não significar nada para o fabricante, para quem o tesouro está sendo construído (LANCASTER, 1987, p. 20).

Ainda, com relação à abordagem empírica, podem ser identificados dois princípios: 1) Garantia literária (*Literary warrant*) também conhecida como Garantia bibliográfica (*Bibliographic warrant*) foi expresso por E. Wyndhan Hulme em 1911, ao argumentar, contrariamente, sobre o estabelecimento das classes na construção de linguagens documentárias, tendo como fundo a classificação do conhecimento. Na oportunidade, Hulme mostrou-se favorável à sua determinação, embasada nas classes em que existe literatura. Dito de outro modo, “as características da literatura em si é que determinam as classes definidas no sistema” (HULME, 1950 apud LANCASTER, 1987, p. 20; DODEBEI, 2002, p. 71). Como um exemplo de aplicação desse princípio, destacamos o sistema de classificação da Biblioteca do Congresso Americano, desenvolvido de modo pragmático a partir do estudo de seu acervo bibliográfico (DODEBEI, 2002); 2) Garantia do usuário (*User warrant* ou *Personal warrant*), expressão usada por Lancaster em 1972, também conhecida como Endosso do usuário (DODEBEI, 2002). Essa garantia é uma confirmação direta dos produtores e utilizadores das informações de um campo de atividades que os termos escolhidos são, efetivamente, aqueles utilizados na comunidade e, desse modo, serão eles utilizados nas solicitações de busca no sistema de recuperação da informação, de acordo com Lancaster (1987) e Dodebei (2002).

Concordando com os autores supracitados, que a seleção de termos a partir do princípio da garantia literária é de grande importância, entendemos, igualmente importante que os termos coletados representem as necessidades dos

usuários do serviço de informação, e, com isso, adotem um tesouro, também a partir da garantia do usuário. Nesse sentido, o compilador do tesouro “faria muito bem em identificar os termos que traduzem os interesses temáticos dos usuários, ao mesmo tempo em que os coletam da literatura”, conforme sugere Lancaster (1987, p. 22).

Se o serviço de informação mantém uma “coletânea de registros de buscas de literatura realizadas para os usuários”, tais registros são, provavelmente, uma rica fonte de terminologia, além de questionários que também podem ser úteis. Esses questionários podem ser enviados a todos os usuários, ou a uma amostragem deles. No caso de “uma comunidade de pesquisa, os títulos e resumos de suas produções constituem valiosa fonte de termos” (LANCASTER, 1987).

Lancaster (1987) apresenta ainda uma terceira abordagem que poderia ser considerada “combinação de princípios de garantia”. O autor supracitado explica: “É possível combinar em uma única operação a garantia literária com a garantia de uso. Usuários representativos recebem cópias das publicações que refletem seus atuais interesses profissionais. Esses usuários sublinham os termos” que identificam e tornam essas publicações interessantes para eles. Nos estudos realizados, percebemos, ainda, a existência de outra forma de garantia, denominada Garantia de acesso a qual se supõe ser decorrente de aplicações tecnológicas.

Batty (1989) mostra que, para a construção de tesouros, é necessário observar duas fases: a de planejamento e a de procedimento de construção. Na fase do planejamento, são considerados alguns aspectos, como tipo de usuário, suas necessidades, abrangência e nível de especificidade do tesouro que dizem respeito à delimitação de fronteiras da área específica do tesouro como área núcleo e área de ramificações (intensidade e números de termos). O autor descreve como fases de procedimentos: 1) identificar fontes de procedimentos e de coleta de termos; 2) coletar e registrar o vocabulário básico; 3) organizar os termos básicos em categorias tendo como critério a afinidade semântica; 4) definir subcategorias; 5) arranjar os termos dentro das categorias, hierarquicamente; 6) considerar até onde especificar, dependendo da complexidade do vocabulário; 7) usar notações; 8) estabelecer prioridade na ordenação das categorias, se necessário; 9) estabelecer termos compostos ou pré-coordenados; 10) refinar o vocabulário base: (a) confirmar o nível de especificidade; (b) harmonizar o vocabulário básico com a linguagem do usuário; (c) definir o significado de palavras que tenham múltiplos significados; (d) esclarecer a maneira em que um termo

deverá ser usado; 11) Fazer referências / estabelecer relações (BT/NT, RT). Com base no relato do autor sobre os procedimentos para construção de tesouros percebe-se uma orientação pelo método indutivo.

Gomes (1990) explica no “Manual de elaboração de tesouros monolíngues”, no que se refere à compilação de termos, que há vários métodos de compilação. Apresenta ainda, dois procedimentos para a pesquisa terminológica para a construção de tesouros, identificando-os: 1) coleta de termos via literatura – a coleta é feita na literatura especializada e permite obter o termo com o significado que lhe atribuem os especialistas da área. Em seguida, o levantamento deve ser validado pelos especialistas da área do tesouro para a obtenção, além da garantia literária, da garantia do usuário ou 2) coleta de termos via tabela de classificação – a coleta é feita nos sistemas de classificação com notação. Nesses sistemas de classificação, os termos são as notações e não as palavras ou expressões correspondentes. Existem dois tipos de sistemas de classificação: enumerativo ou facetado. Nos sistemas de classificações enumerativas, os termos são estruturados em assuntos, e os assuntos são reunião de conceitos. Cada assunto é analisado a fim de identificar os conceitos que o formam. Nos sistemas de classificação facetada, estruturam-se conceitos e não assuntos, e cada conceito é organizado em uma faceta ou categoria, o levantamento dos termos é menos trabalhoso do que nos sistemas enumerativos, vez que nesses é preciso usar o processo de análise. Gomes (2004) em seu tutorial sobre elaboração de tesouro documentário, disponível em: <<http://www.conexaorio.com/bitit/tesauro/>> destaca as etapas gerais para a construção de tesouros: 1) planejamento; 2) levantamento do vocabulário; 3) organização dos conceitos e 4) apresentação final.

Crouch (1990), em revisão de estudos anteriores que investigaram a construção de tesouros por meios manuais, semi-automáticos e automáticos, examinou as várias abordagens de construção de tesouros, descreveu uma abordagem da geração de tesouros globais (macrotesouros) baseada no modelo de discriminação do termo com base no valor (modelo de Salton, Yang e Yu) e baseada no algoritmo de compilação apropriada. Esse método foi implementado e aplicado a duas coleções de documentos. Preliminarmente, os resultados indicaram que esse método proporciona melhoria de 10 a 15% no desempenho da recuperação nas coleções testadas, é variável e merece maiores investigações. Para a elaboração do tesouro, segue-se um modelo já descrito por Salton, Yang e Yu, que se baseia nos

termos selecionados pelo seu valor e importância e não por sua frequência. Os termos seriam compilados de documentos estreitamente relacionados entre si e, a partir destes termos de baixa frequência mas de valor significativo, seriam construídas as classes (categorias). Portanto, utilizou-se o método dedutivo.

Fujita (1992) orienta que é fundamental refletir sobre algumas questões para a construção de tesouros como: A construção do tesouro deverá ser realizada: antes ou depois da indexação dos documentos? Qual o momento certo para construir ou adaptar o tesouro?

Para responder a essas questões, recorreremos à autora supracitada, que propõe três alternativas fundamentadas por sua vez em Suddarshan (1985 apud FUJITA, 1992): 1) indexar, certo número de documentos sem controle de vocabulário para compilação de termos e, juntamente, com os termos coletados de dicionários, diretórios, bases de dados e outras, construir o tesouro. Quando o tesouro estiver pronto, será usado para indexar os outros documentos e novamente os já indexados; 2) construir um tesouro completo coletando termos de várias fontes: dicionários, diretórios, periódicos, livros, outros tesouros relacionados com a área e outras, e depois usá-lo para a indexação dos documentos; 3) indexar os documentos e construir o tesouro simultaneamente.

Concordando com Fujita (1992), entendemos que, no primeiro método, a qualidade do tesouro é superior uma vez que se utilizará dos termos da própria coleção de documentos. Por outro lado, julgamos que terá a séria desvantagem de duplicação de trabalho na reindexação da mesma coleção. No segundo método não existe o problema de re-trabalho, contudo, o breve período de coleta de termos e a quantidade limitada de fontes terminológicas consultadas prejudicarão a qualidade do tesouro. Além disso, haverá um tempo de espera antes da indexação da coleção. Já, o terceiro método abriga as seguintes vantagens: “reunir as qualidades e eliminar as limitações dos dois métodos destacados anteriormente” (FUJITA, 1992, p. 87).

Diante do exposto, para a adesão ao terceiro método e para acionarmos o funcionamento simultâneo da indexação e construção do tesouro, embasamo-nos nas orientações de Fujita (1992) que alerta sobre a necessidade de formação de “dois grupos: tesouro e indexação. O grupo do tesouro será formado antecipadamente e iniciará a construção preliminar do tesouro [...]. O grupo de indexação será formado após a construção do tesouro preliminar” (FUJITA, 1992, p.

87). Assim, o segundo grupo, de indexação, desde o início do processo de indexação, manterá “contato permanente com o grupo de construção do tesauro, alimentando a estrutura do tesauro com termos de indexação pertinentes aos documentos analisados” (FUJITA, 1992, p. 88).

Em sua pesquisa, a autora supracitada reconheceu as inúmeras vantagens do terceiro método explicitado, entretanto optou pela construção do Tesauro Preliminar de Odontologia fundamentado no primeiro método exposto. Fujita (1992) apresenta os dois motivos que contribuíram para a sua tomada de decisão: o primeiro - porque é objeto de uma pesquisa experimental em que está sendo proposta a indexação dos documentos pelo sistema PRECIS como método de compilação, o que pressupõe a indexação de um bom número de documentos sem controle de vocabulário; e o segundo motivo - porque esse tesauro é produto de um esforço isolado relatado em um trabalho monográfico, destinado à sua capacitação em nível de doutorado (FUJITA, 1992, p. 88).

Amat e Alcocer (1993) apresentam um procedimento que tem permitido a criação de vários tesouros em torno do Sistema de Recuperação da Informação. Esses tesouros compreendem 7.500 termos correspondentes da geografia mundial e controlam as bases de dados bibliográficas e audiovisuais de informação da atualidade. A estruturação do arquivo de fichas e a delimitação das relações terminológicas têm sido realizadas empregando um sistema de gestão de base de dados em microcomputador. O alcance semântico dos termos tem se representado mediante códigos classificadores alfanuméricos que permitem identificar as relações de equivalência (igualdade de códigos) e hierárquicas. O procedimento é igualmente aplicável ao estabelecimento de relações associativas.

Os programas detalhados ao longo do trabalho baseiam-se na relação de um arquivo geral em que os termos se ordenam alfabeticamente e num outro arquivo em que os termos referentes se dispõem em ordem hierárquica segundo seus códigos. Os arquivos adicionais contêm os termos não preferidos e os associados. Como resultado, obtêm-se tanto as edições alfabéticas e hierárquica simples como a combinação de ambas, apta para a ficha posterior do arquivo no modelo do tesauro BRS e outros grandes sistemas. Os autores não esclarecem o método de compilação utilizado, entretanto sua descrição acerca da representação dos conceitos através dos códigos classificatórios elaborados para o sistema

utilizado, proporcionou uma estrutura dos registros nos arquivos de bases de dados, o que significa dizer que o método de compilação utilizado foi o dedutivo.

Autores como Guinchat e Menou (1994, p. 153) asseveram que existem dois métodos de construção de tesouros: a) o “método *a priori*, que consiste em recensear a terminologia existente, a partir de índices, dicionários, outras linguagens documentais, taxionomias e listas de termos fornecidas por especialistas”. b) o “método *a posteriori*, ou analítico, consiste em extrair os termos de um conjunto representativo de documentos e de questões tratadas durante as operações documentais, em um período determinado”. Em seguida, os autores orientam que se reúnam os dados lexicográficos e se faça a seleção. Além disso, são estabelecidas as relações entre os descritores em potencial e o reagrupamento dos conjuntos importantes. Depois disso, selecionam-se os descritores e escolhe-se sua forma verbal.

- **Síntese**

A seguir, apresentamos uma síntese dos métodos de construção de tesouros utilizados ou recomendados pelos autores analisados: Aitchison; Gilchrist (1979) recomendam o método indutivo (*Thesaurofacet*); Devadason (1985) sugere método dedutivo (Classaurus) Robredo (2005), método evolutivo, método consensual correspondem às Diretrizes UNESCO para tesouros monolíngues; nomeia ainda como método analítico ou método gestalt; Lancaster (1987) reconhece duas abordagens: dedutiva e indutiva, mas acentua que a última “tende a ser mais confiável”; Batty (1989) em relato sobre os procedimentos para construção de tesouros, percebe-se uma orientação pelo método indutivo; Gomes (1990; 2004) pela descrição da etapa de compilação e no tutorial acima mencionado, poderia ser considerada as abordagens dedutiva/indutiva. Crouch (1990), em revisão de estudos anteriores, que investigaram a construção de tesouros, por meios manuais, semi-automáticos e automáticos, considerando as diferentes abordagens de construção de tesouros, revelou-se favorável às atuais quanto às antigas e descreve uma abordagem da geração de tesouros globais baseada no método dedutivo. Fujita (1992), em sua tese de doutorado, explicita tanto os métodos dedutivo, indutivo



como também sinaliza para uma terceira abordagem, identificada como combinação de métodos. Contudo, em seu Tesouro Preliminar em Odontologia optou pelo método dedutivo.

Desse modo, Gomes (1990), Fujita (1992), Guinchat e Menou (1994) e Robredo (2005) reforçam a importância dos métodos: dedutivo e indutivo na construção de tesouros, embora com designações diferentes às duas abordagens metodológicas. Após a análise das posições e argumentos apresentados pelos autores elencados nessa pesquisa, apoiamos-nos nas fundamentações de Fujita (1992) para chegar ao entendimento de que a principal divisão é estabelecida da seguinte forma: Quando o método é dedutivo, a preocupação não reside apenas na coleta dos termos, mas também no estabelecimento de sua estrutura conceptual ou categorização. Mapeiam-se o domínio e o subdomínio em estudo, antes do início da coleta de termos. Quando o método é indutivo, não existe uma preocupação inicial com a estrutura conceptual ou categorização, porque se organizam o domínio e o subdomínio em estudo, após a coleta dos termos (FUJITA, 1998).

Além disso, os autores Fujita (1992), Guinchat e Menou (1994), além de Robredo (2005), argumentam que, na prática, quando se trata da construção de um tesouro, geralmente, combinam-se os dois métodos. O que contribui para colocar em evidência outra abordagem, a de combinação de métodos. Nessa direção, indicada anteriormente por Fujita (1992), incorporam-se “os princípios do método dedutivo em determinados momentos da etapa de construção de tesouros, em outros, orientam-se seguindo os fundamentos do método indutivo”. Concordando com a autora supracitada, entendemos como possível a integração de outras combinações, como os procedimentos terminográficos que venham a contribuir para o aprimoramento de etapas na construção de tesouros.

Em outra análise, percebemos, ainda, que autores como: Aitchison; Gilchrist (1979), Devadason (1985), Lancaster (1987), Batty (1989), Gomes (1990; 2004), além de Fujita (1992; 1998) indicam uma sistematização de processos para a construção de tesouro com as quais elaboramos o quadro 8, com a composição de cada etapa e breve resumo de seu conteúdo, seguido do(s) autor(es) e o ano da publicação analisada, conforme apresentamos a seguir:

ETAPAS DE CONSTRUÇÃO DE TESAURO SEGUNDO AUTORES	
CATEGORIAS TEMÁTICAS	AUTORES
<b>1) Fase do planejamento</b> tipo de usuário, suas necessidades; abrangência e nível de especificidade do tesauro; identificações de fontes de procedimentos e de coleta de termos.	Batty (1989) Gomes (1990) Fujita (1992) Gomes ([2004])
<b>2) Formas/métodos de Compilação de termos</b> Dedutivo Indutivo Combinação de Métodos (Dedutivo/Indutivo)	Aitchison; Gilchrist(1979) Lancaster (1987) Batty (1989) Gomes (1990) Fujita (1992) Gomes ([2004])
<b>3) Compilação de termos</b> a) coleta - registro e seleção dos termos compilados e b) validação – registro do vocabulário básico; coleta e validação de termos.	Aitchison; Gilchrist (1979) Lancaster (1987) Batty (1989) Gomes (1990) Fujita (1992) Gomes ([2004])
<b>4) Estabelecimento de relações entre termos/ Categorização</b> estruturação de conceitos com controle terminológico dos termos; ordenação dos termos; estabelecimento de categorias elementares; organização dos termos básicos em categorias (critério a afinidade semântica); definição de subcategorias; estabelecimento de relações entre termos.	Aitchison; Gilchrist (1979) Lancaster (1987) Batty (1989) Gomes (1990) Fujita (1992) Fujita (1998) Gomes ([2004])
<b>5) Especificidade</b> Estabelecimento de limites de especificação/ dependendo da complexidade do vocabulário.	Lancaster (1987) Gomes (1990) Gomes ([2004])
<b>6) Uso de equipamento informático para processamento de dados</b> estruturação automática das partes alfabética e sistemática do tesauro (etapa 7); produção de uma estrutura final (etapa 7)	Lancaster (1987) Gomes (1990) Fujita (1992) Gomes ([2004])
<b>7) Formas de Apresentação</b> alfabética; sistemática; alfabética/classificada; facetada. estruturação automática das partes alfabética e sistemática do tesauro; produção de uma estrutura final	Lancaster (1987) Gomes (1990) Fujita (1992) Gomes ([2004])

**Quadro 8** - Síntese das etapas de construção de tesouros segundo autores.

O quadro 8 apresentou-nos uma síntese das etapas de construção de tesauro, fundamentada nos autores analisados. Com a sistematização realizada, foi possível visualizar melhor os resultados obtidos e verificar que, em sua totalidade, os autores analisados reconhecem as etapas de compilação de termos e de estabelecimento de relações entre termos, como essenciais ao conjunto das etapas de construção de tesauro. Na etapa de compilação de termos, constatamos que Gomes (1990) menciona a pesquisa terminológica para a construção de

tesauros, argumentando que a coleta feita, na literatura especializada, permite obter o termo com o significado que lhe atribuem os especialistas da área. Além disso, a autora supracitada orienta que, em seguida, “o levantamento do vocabulário deve ser validado pelos especialistas da área do tesouro para a obtenção, além da garantia literária, a garantia do usuário”; Observamos também que os autores Lancaster, Fujita e Gomes estão presentes na maior parte das etapas identificadas.

Apresentamos, a seguir, conciso relato sobre tesauros e seus métodos de construção.

## 2.5 Tesauros e seus métodos de construção

Destacamos, neste breve relato, que foram identificados inúmeros tesauros de diferentes áreas. Contudo, selecionamos, para síntese, tesauros construídos mediante o uso de aplicações tecnológicas e, preferencialmente da área da Ciência da Informação, se possível, disponíveis na Web.

Destacamos ainda duas exceções aos critérios indicados para seleção e análise dos tesauros: no primeiro caso, tratamos do Tesouro de Ciência da Informação (IBICT) por julgarmos ser uma das primeiras iniciativas brasileiras de elaboração de programa de computador para geração de tesouro - TECER; e no segundo, evidenciamos o Tesouro Preliminar de Odontologia, construído com o uso de recursos da informática para a elaboração de indexação automática de documentos e para a geração automática de tesauros, utilizando o programa TECER.

**Tesouro de Ciência da Informação** do Instituto Brasileiro de Informação em Ciência e Tecnologia, publicado em 1989, foi desenvolvido para atender às necessidades de indexação dos documentos existentes no acervo do Centro de Documentação e Informação desse Instituto. O acervo desse Centro, à época, especializado em Biblioteconomia, Documentação e Ciência da Informação, compreendia, portanto, as áreas cobertas pelo Tesouro (IBICT, 1989a). A estrutura temática do Tesouro foi orientada em sete grandes categorias que também podem ser entendidas como indicadores de facetas: Informação; Documento; Unidade de informação; Planejamento; Processos e serviços de informação; Transferência e uso

da informação e Profissão. Nessa estruturação do tesouro, observou-se certa articulação de suas facetas em torno do conceito Informação, a despeito das muitas atualizações ocorridas na Biblioteconomia e na Ciência da Informação. Foi utilizado o programa de elaboração de tesouro – TECER (IBICT, 1989b). Apresenta uma lista alfabética estruturada e uma lista por categorias com os termos dispostos alfabeticamente (IBICT, 1989a).

**Tesouro Preliminar de Odontologia** produzido por Fujita (1992) constitui-se de uma iniciativa e um esforço isolado de construção de uma linguagem documentária alfabética como um experimento em sua pesquisa de doutorado. Foi construído para atender aos interesses da comunidade docente, dos indexadores e discentes, usuários do sistema de informação da Faculdade de Odontologia de Bauru da Universidade de São Paulo. O Tesouro foi construído tendo como fundamento os princípios diretivos da ISO 2788 (1986), com base na abordagem do método dedutivo e na metodologia de indexação do Sistema PRECIS a fim de viabilizar sua aplicação automática utilizando qualquer tipo de documento. Apresentou proposta inédita, no Brasil, de indexação dos documentos pelo sistema PRECIS como método de compilação de termos para a construção de linguagem documentária alfabética. Além disso, no processo de construção do Tesouro foi utilizado o programa para geração automática de tesouro – TECER desenvolvido pelo IBICT. O processo de construção, passo a passo, encontra-se explicitado, brilhantemente, em sua tese de doutorado. O SPI – Sistema Precís de Indexação é destinado a profissionais da informação que sejam competentes em relação ao PRECIS (*PREserved Context Index System*). O *corpus* foi constituído por publicações científicas e considerou o direcionamento das pesquisas realizadas pela professora doutora Mariângela Spotti Lopes Fujita, da Universidade Estadual Paulista, Campus de Marília que averiguou o funcionamento do Sistema de Indexação PRECIS na língua portuguesa. O resultado desse estudo encontra-se publicado no livro “PRECIS na língua portuguesa: teoria e prática de indexação” (FUJITA, 1988).

**ASIS Thesaurus of Information Science and Librarianship** da *American Society for Information Science*, em segunda edição, segue as orientações propostas na edição anterior. Orienta-se pela *National Information Standards Organization* (NISO) da *American National Standards Institute*, ANSI/NISO Z39.19 (1993). Abrange os campos da Ciência da Informação e Biblioteconomia para a

indexação nestas áreas. O *ASIS Thesaurus* descreve e incorpora o caráter interdisciplinar da Ciência da Informação e suas interfaces mais evidentes. O tesauro encontra-se delimitado a partir de assuntos tópicos. De acordo com a política de seleção dos termos, não foram incluídos nomes próprios de organizações, pessoas, programas e outros, com algumas exceções como, por exemplo, os nomes de ferramentas, tendo em vista sua representatividade na prática da indexação e catalogação. Outra exceção aberta diz respeito a termos fundamentais no contexto da Web, já que por sua disseminação e uso comum podem ser considerados nomes próprios. Nessa edição, o *ASIS Thesaurus* inclui descritores e não-descritores, distribuídos em trinta e seis facetas. Evidencia equipe de construção do tesauro e as dificuldades para sua atualização, frente ao dinamismo existente na Ciência de Informação cujo campo de ação sofreu a velocidade de mudanças, principalmente após a Web, com a incorporação de novos termos e o desuso de muitos outros. O conjunto de programas The Liu-Palmer *Thesaurus Construction System* (TCS), versão 3.3, foi utilizado para a construção do tesauro. O TCS é um sistema para auxílio à construção de tesouros que apresenta flexibilidade, permite a construção de hierarquias e relacionamentos, inclusive distinguindo os tipos de relacionamentos entre os termos (MILSTEAD, 1998).

***Tesauro de Biblioteconomía y Documentación*** construído por Mochón Bezares e Sorli Rojo, em 2002, para ser um instrumento para a análise de conteúdo e a recuperação dos documentos incluídos na Base de Dados sobre Biblioteconomia e Documentação, ISOC-DC, do *Consejo Superior de Investigaciones Científica* (CSIC) desde 1975. Para a sua criação orientou-se na norma UNE 50-106 (1990), que corresponde à norma ISO 2788 (1986). Além do conjunto de termos indexados na base de dados ISOC-DC, utilizou-se de termos compilados de outras fontes de informação especializadas como dicionários e tesouros de Biblioteconomia e Documentação, e outros. Este tesauro consta de três índices: alfabético, hierárquico e permutado KWOC. A implantação do tesauro possibilitou uma indexação mais homogênea dos documentos incorporados na base de dados, permitindo uma recuperação fácil e exaustiva, enriquecendo o trabalho dos documentalistas e ampliando o campo de busca do usuário. O tesauro proporciona à comunidade científica um conjunto estruturado de termos sobre a base de um sistema de conceitos para a organização do conhecimento biblioteconômico. (MOCHÓN BEZARES; SORLI ROJO, 2003).

**DOCUTES - Tesouro de Ciencias de la Documentación**, foi desenvolvido pela área de Biblioteconomia e Documentação da *Universidad de León* por meio de um projeto de investigação científica patrocinado pela *Junta de Castilla e León* em 2000. Esse Tesouro foi elaborado em razão de fazer parte de um objetivo maior, de potencializar práticas e ações na área por meio de tutoriais desenvolvidos a partir de material adequado. O tesouro *DOCUTES* é utilizado pelos professores da área de Biblioteconomia e Documentação como recurso didático para as aulas. Contém relação completa dos termos distribuídos nas seguintes facetas: Ciência da documentação: História, Teorias, Sistemas, Informação, Documentos, Fontes de informação, Investigação e metodologia documental, Profissionais e usuários, Representação e recuperação de informação, Sistemas de informação e Tecnologias da informação. Para a construção do tesouro, utilizou-se do programa *MultiTes*, desenvolvido pela empresa *MULTYSYSTEMS*. Trata-se de um *software* específico que possibilita e otimiza o tratamento automatizado dos termos e posterior edição de tesouros eletrônicos. Encontra-se disponível para acesso em: <<http://www3.unileon.es/dp/abd/tesouro/pagina/tesdocumentacion/docutes.htm>> (RODRÍGUES BRAVO et al, 2004). .

**TCI - Tesouro em Ciência da Informação** foi construído a partir dos seguintes tesouros: *ASIS Thesaurus of Information Science and Librarianship* da *American Society for Information Science* (MILSTEAD, 1998), o *DOCUTES - Tesouro de Ciencias de la Documentación*, da *Universidad de León* (Espanha), o *Tesouro de Biblioteconomía y Documentación*, do *Consejo Superior de Investigaciones Científica* (CSIC) e o Tesouro em Ciência da Informação, do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), conforme relatos de Moreira e Moura (2006). A organização dos conceitos foi conduzida pelo processo de indução/dedução. Os autores argumentam que, à medida que as relações foram estruturadas, novos conceitos foram acrescentados, complementando a organização de cada faceta do futuro tesouro. De forma genérica, três tipos de relacionamentos direcionaram a elaboração de tesouros: a) relacionamento de equivalência; b) relacionamento hierárquico; c) relacionamento associativo. A versão acadêmica do *Thesaurus Construction System Version 8* (TCS-8) (WEBCHOIR, 2008) foi utilizada para a construção do tesouro a partir dos termos e das relações hierárquicas já estabelecidas. O TCS-8 é um sistema para auxílio à construção de tesouros que apresenta flexibilidade e facilidades de

manuseio e permite aos usuários criar tesouros e outros vocabulários. Orienta-se pelo padrão Z39.19 da ANSI/NISO, que permite ao usuário criar suas próprias regras de estrutura (MOREIRA; MOURA, 2006).

Apresentamos, no quadro 9, uma síntese dos tesouros e seus métodos de construção. Além disso, evidenciamos os usos de recursos de informática para a sua construção:

TESAUROS E SEUS MÉTODOS DE CONSTRUÇÃO/USO DE RECURSOS DE INFORMÁTICA	
TESAUROS	MÉTODOS DE CONSTRUÇÃO/USO DE RECURSOS DE INFORMÁTICA
<b>Tesouro de Ciência da Informação</b> (IBICT, 1989b)	Apresenta uma lista alfabética estruturada e uma lista por categorias com os termos dispostos alfabeticamente. Foi utilizado o programa de geração automática de tesouro – TECER.
<b>Tesouro Preliminar de Odontologia</b> (FUJITA, 1992)	Com base na abordagem do método dedutivo e na metodologia de indexação do Sistema PRECIS a fim de viabilizar sua aplicação automática, utilizando qualquer tipo de documento. Apresentou proposta inédita, no Brasil, de indexação dos documentos pelo sistema PRECIS como método de compilação de termos para a construção de linguagem documentária alfabética. Além disso, no processo de construção do Tesouro foi utilizado o programa para geração automática de tesouro – TECER desenvolvido pelo IBICT.
<b>ASIS Thesaurus of Information Science and Librarianship</b> (MILSTEAD, 1998)	O conjunto de programas The Liu-Palmer <i>Thesaurus Construction System</i> (TCS), versão 3.3, foi utilizado para a construção do tesouro. O TCS é um sistema para auxílio à construção de tesouros que apresenta flexibilidade, permite a construção de hierarquias e relacionamentos, inclusive distinguindo os tipos de relacionamentos entre os termos.
<b>DOCUTES - Tesouro de Ciencias de la Documentación –</b> (RODRÍGUES BRAVO et al, 2000)	A Equipe do Projeto DOCUTES utilizou-se do programa MultiTes, desenvolvido pela MULTSYSTEMS, que possibilita e otimiza o tratamento automatizado dos termos e posterior edição de tesouros eletrônicos. No intuito de potencializar práticas e ações na área, por meio de tutoriais, o DOCUTES foi desenvolvido por professores da área de Biblioteconomia e Documentação da <i>Universidad de León</i> como recurso didático para as aulas. Disponível em: < <a href="http://www3.unileon.es/dp/abd/tesouro/docutes.htm">http://www3.unileon.es/dp/abd/tesouro/docutes.htm</a> >. Apresenta o conjunto dos termos distribuídos por facetas: Ciência da documentação; Fontes de informação; Profissionais e usuários; Representação e recuperação de informação; Sistemas de informação e Tecnologias da informação.
<b>Tesouro de Biblioteconomía y Documentación</b> (MOCHÓN BEZARES; SORLI ROJO, 2003).	Desenvolvido para ser uma linguagem controlada para a análise de conteúdo e recuperação de documentos incluídos na <i>Base de Datos de Biblioteconomía, Documentación y Política Científica</i> ISOC-DC, produzida pelo CINDOC, do <i>Consejo Superior de Investigaciones Científicas</i> (CSIC) desde 1975. Sua criação teve como alvo suprir a falta de léxicos documentais em espanhol abrangendo os campos semânticos representados nos textos técnicos e científicos publicados na Espanha.
<b>TCI - Tesouro em Ciência da Informação</b> (MOREIRA; MOURA, 2006)	Construído a partir de 4 tesouros existentes, teve a organização dos conceitos orientada pelo processo de dedução/indução. A versão acadêmica do <i>Thesaurus Construction System Version 8</i> (TCS-8) é um sistema para auxílio à construção de tesouros que apresenta flexibilidade e facilidades de manuseio, permite aos usuários criar tesouros e outros vocabulários. Orienta-se pelo padrão Z39.19, ANSI/NISO, que permite ao usuário, caso necessite, criar suas próprias regras de estrutura.

**Quadro 9 – Tesouros e seus métodos de construção e o uso de recursos de informática.**

A síntese apresentada no quadro 9, elaborada com base em relatos de pesquisa para a construção de tesouro, mostrou-nos resultados importantes, no tocante aos métodos empregados e o uso de equipamento de informática, em especial, pelos programas de computador, voltados para a construção e gestão de tesouros disponíveis na web para o acesso. Com base nos tesouros analisados, destacamos suas principais contribuições:

*Tesouro Preliminar em Odontologia* - explicitou o método dedutivo, aplicou e obteve resultados positivos com a metodologia de indexação do Sistema PRECIS para a compilação de termos e o uso do programa de computador TECER do IBICT, para geração automática do tesouro (FUJITA, 1992).

*TCl - Tesouro em Ciência da Informação* (MOREIRA; MOURA, 2006) indicou o modo de organização dos conceitos pelo processo de dedução/indução. A versão acadêmica do *Thesaurus Construction System, Version 8*, (TCS-8) foi o programa de computador utilizado para a construção do tesouro. Orienta-se pelo padrão Z39.19 da ANSI/NISO, que permite ao usuário, caso necessite, criar suas próprias regras de estrutura. Disponível na Web (MOREIRA, MOURA, 2006).

*ASIS Thesaurus of Information Science and Librarianship* - evidenciou as dificuldades para sua atualização, frente ao dinamismo existente no campo da Ciência de Informação que sentiu a velocidade das mudanças, tecnológicas, com a incorporação de novos termos e o desuso de muitos outros. O *Thesaurus Construction System* (TCS), versão 3.3, foi o programa de computador utilizado para a construção do tesouro. O TCS é um sistema para auxílio à construção de tesouros que apresenta flexibilidade, permite a construção de hierarquias e relacionamentos, inclusive distinguindo os tipos de relacionamentos entre os termos (MILSTEAD, 1998).

*DOCUTES - Tesouro en Ciencia de la Documentación* - recomendou o uso do programa de computador MultiTes da empresa MULTYSYSTEMS para a construção do tesouro. O MultiTes é um *software* específico que possibilita e otimiza o tratamento automatizado dos termos e posterior edição de tesouros eletrônicos. O DOCUTES é utilizado pelos professores da área de Biblioteconomia e Documentação como recurso didático para as aulas. Encontra-se disponível na Web para consulta.



- **Síntese**

Nesse capítulo, analisamos e sintetizamos os fundamentos teóricos e metodológicos da área de Organização da Informação, com foco na Organização e Representação do Conhecimento. Enfatizamos as suas contribuições para o desenvolvimento de estudos das linguagens documentárias alfabéticas, buscando seu entendimento como um sistema de organização do conhecimento, parte fundamental dos processos de construção de linguagens documentárias alfabéticas, em especial, o tesauro.

Evidenciamos alguns resultados importantes dos itens analisados, até o momento. A representação de conceitos e a indexação temática da informação constituem a síntese por meio da qual o conteúdo de um documento é identificado por conceitos e representado por meio de termos que, por sua vez, se agrupam e se relacionam com os conteúdos de outros documentos de temas análogos ou semelhantes para atender aos usuários que buscam recuperá-los no âmbito de um sistema de informação de áreas especializadas.

Destacamos que, para existir a recuperação da informação desejada, é preciso que tanto a linguagem de busca quanto a fonte de informação estejam adequadamente representadas e mediadas por uma linguagem documentária. Em razão disso, a função das linguagens documentárias no processo de representação para a recuperação da informação é imprescindível, uma vez que proporciona a conexão entre as linguagens utilizadas tanto na entrada de conteúdos documentários, como na busca a esses conteúdos, num sistema de recuperação da informação (GOMES, 1990); FUJITA, 1992; LARA, 1993b; TÁLAMO, 1997b; CAMPOS, 2001; DODEBEI, 2002).

Sinalizamos que o tesauro constitui o exemplo mais característico de linguagem documentária alfabética, limita-se, em princípio, ao campo de termos preferidos de uma área de especialidade. O tesauro ainda que se aproprie das terminologias e dos sistemas conceituais por ele mapeados, propõe-se como modo de organização da informação, articulando os temas, suas especificidades e termos orientados para dar conta de conjuntos documentários e para transferir informação para determinados públicos segundo objetivos específicos (FUJITA, 1992; TÁLAMO, 1997c; LARA, 2004).

Nesse contexto, observa-se que a mediação ocorre via aplicação de métodos e técnicas no processo de construção do tesouro, com o objetivo de facilitar a recuperação da informação e contribuir para o propósito maior de geração do conhecimento. Em especial, considerando que o sistema de informação de áreas especializadas não possui um instrumento adequado, compatível à área especializada para a qual se precisa organizar e representar tematicamente uma coleção.

Concordando com autoras supracitadas, entendemos como fator de qualidade indispensável que, na construção de tesouros, se aproxime à representação dos conceitos, os termos reconhecidos pelos dois segmentos integrantes do sistema de recuperação da informação: o primeiro, correspondente aos usuários, representados pelos estudantes, pesquisadores, profissionais e outros; e o outro segmento, correspondente ao conhecimento registrado organizado pelo sistema de informação, à coleção organizada dos conteúdos temáticos documentários disponíveis, independente do suporte físico e formas de acesso.

Após a análise e síntese dos procedimentos metodológicos, observadas as orientações e argumentos apresentados pelas normas, pelos autores e pelos tesouros e seus métodos de construção baseados no uso de equipamento de informática, decidimos prosseguir pela apresentação das etapas de construção de tesouros, fundamentadas de acordo com as normas documentárias e/ou autores analisados. A decisão por esse meio decorre por entendermos que cada programa de computador dedicado a essa finalidade, nesse caso, a construção de tesouros, guarda em si um procedimento metodológico específico. Todavia, o uso de equipamento de informática, para agilidade nos trabalhos, poderá ser utilizado em algumas das etapas de construção de tesouro, o que será exposto mais adiante.

No que se refere aos métodos de construção de um tesouro, as normas analisadas apresentam diferentes abordagens: a abordagem de comitê, a abordagem empírica e a abordagem combinação de métodos, conforme descrito no Quadro 6, APÊNDICE A. Em razão disso, buscamos entender de que modo tais abordagens podem interferir na construção de tesouro, especialmente, na realização das etapas de construção de tesouro e, ainda, de que forma essas abordagens podem ser percebidas.

As fundamentações de Fujita (1992; 1998) levam-nos ao entendimento de que a principal divisão é estabelecida da seguinte forma: quando o

método é dedutivo, a preocupação não reside apenas na coleta dos termos, mas também no estabelecimento de sua estrutura conceptual ou categorização. Mapeiam-se o domínio e o subdomínio em estudo, antes do início da coleta de termos; quando o método é indutivo, não existe uma preocupação inicial com a estrutura conceptual, ou categorização, porque se organizam o domínio e o subdomínio em estudo, após a coleta dos termos (FUJITA, 1998).

Os autores Aitchison; Gilchrist (1979), Lancaster (1987), Batty (1989), Gomes (1990; 2004), Fujita (1992), Guinchat e Menou (1994), além de Robredo (2005) argumentam que, na prática, quando se trata da construção de um tesouro, em geral, combinam-se os dois métodos. O que contribui para colocar em evidência outra abordagem, a de combinação de métodos. Nessa direção, indicada anteriormente por Fujita (1998), incorporam-se “os princípios do método dedutivo em determinados momentos da etapa de construção de tesouros, em outros, orientam-se seguindo os fundamentos do método indutivo”. Concordando com a autora supracitada, acreditamos como possível a integração de combinações de procedimentos, como dos procedimentos terminográficos para o aprimoramento de etapas específicas da construção de tesouros.

Apresentamos uma síntese das etapas da construção de tesouros com base no estudo das Diretrizes IBICT (1984), Diretrizes UNESCO (1993) e Diretrizes ANSI/NISO (2005), além da análise da sistematização das etapas de construção de tesouros realizada pelos autores analisados. Para apresentação desse estudo, foram agrupadas em 7 categorias temáticas, a saber:

1. Trabalho preliminar (Orientações gerais/Fase de planejamento);
2. Métodos de compilação (Formas/métodos de compilação de termos);
3. Registro de termos (Compilação de termos: coleta; validação);
4. Verificação de termos (Admissão e exclusão de termos/ Estabelecimento de relações entre termos/categorização);
5. Especificidade (Nível de especificidade);
6. Uso de equipamento automático de processamento de dados (Uso de equipamento informático para processamento de dados/utilização do computador);
7. Forma e conteúdo de um tesouro/Formas de apresentação.

Em conformidade com a investigação realizada, para apresentação de proposta de um modelo metodológico para a construção de tesouros com a integração de procedimentos terminográficos tendo em vista o aprimoramento de etapas da construção de tesouros. Cabe salientar que autores mencionam a pesquisa terminológica para a construção de tesouros, argumentando que a coleta feita, na literatura especializada, permite obter o termo com o significado que lhe atribuem os especialistas da área. Enfatizamos que, até o momento, não foram identificados e explicitados os procedimentos terminográficos em etapas de construção de tesouros.

Diante do exposto, no capítulo 3, tratamos dos aspectos contribuintes da Terminologia e da Terminografia na construção de tesouros.

### **3 ASPECTOS CONTRIBUINTES DA TERMINOLOGIA E DA TERMINOGRAFIA NA CONSTRUÇÃO DE TESAUROS**

Neste capítulo, identificamos os principais aspectos teórico-metodológicos da Terminologia/Terminografia contribuintes para a construção de tesauros. Apresentamos os principais aspectos teórico-metodológicos da Terminologia, dando ênfase aos princípios que regem a definição do termo e suas relações. Esboçamos uma breve visão diacrônica dos fundamentos da Terminologia como campo de estudo, citando as diferentes escolas de pensamento, os conceitos e a importância que se lhes atribui. Destacamos os aspectos análogos entre Terminologia/Terminografia e Documentação, enfatizando pontos de aproximações e semelhanças.

Estudamos sobre como integrar a Terminografia na construção de tesauros. Salientamos, desse modo, que a nossa interface neste trabalho corresponde às formulações da Terminografia, em razão, do interesse cada vez mais presente nas discussões entre pesquisadores, acerca da Terminologia e, por conseguinte, de seus componentes, como o estudo dos vocabulários e dos léxicos especializados, principalmente, devido ao desenvolvimento econômico, cultural, científico e tecnológico da sociedade atual. Torna-se indispensável que a ciência, seja ela básica ou a aplicada, mantenha atualizado o seu discurso metalinguístico para atender necessidades que se diversificam.

#### **3.1 Tendências teóricas e metodológicas da Terminologia: de Wüster a Temmermann**

Ao tratarmos sobre tendências da Terminologia na atualidade, optamos por apresentar, de um modo diacrônico, as abordagens terminológicas que discutem, complementam e se contrapõem ao estabelecido pela abordagem tradicional da terminologia. Em razão da diversidade de núcleos de influência com o decorrer dos anos, percebemos que não há uma unanimidade teórica e metodológica para a Terminologia, o que assevera Lara (2004).

Com aportes teóricos e metodológicos dos autores analisados, verificamos que as principais teorias, em tela, vão desde a Teoria da Terminologia considerada clássica, até àquelas que se apresentam de formas divergentes, das quais destacamos: a) Teoria Geral da Terminologia (TGT); b) Teoria Comunicativa da Terminologia (TCT); c) Teoria Sociocognitiva da Terminologia; além da, d) Teoria Socioterminológica, ou Socioterminologia. Entendemos que para a realização de um trabalho terminológico, alguma dessas abordagens deve ser considerada para conferir rigor e confiabilidade ao trabalho realizado.

Fundamentamos em pressupostos compartilhados por vários estudiosos de que a Terminologia, seja qual for sua temática ou o contexto em que se produz, cumpre sistematicamente duas funções: de representar o conhecimento especializado, e a de transmiti-lo. Evidenciamos, a Terminologia que focaliza-se nos estudos dos conceitos, definições e denominações. Tem sua atenção voltada, também, para a compilação, descrição, apresentação, criação e normalização de termos de áreas especializadas com a preocupação de promover a comunicação e o intercâmbio entre os especialistas e os profissionais. Nesse sentido, a Terminologia é uma disciplina que oferece, de forma sistemática, referencial do vocabulário de uma determinada especialidade e sua análise para a identificação de termos. Ao produzir novos conhecimentos, os especialistas em uma determinada matéria criam novos conceitos que necessitam ser identificados e compartilhados.

As unidades resultantes do processo de denominação são os termos técnicos por meio dos quais se veiculam os saberes e os fazeres. Com isso, a exigência de teorização passa a ser cada vez mais nuclear. Na sua função, a Terminologia não se apresenta como uma inovação. Sua origem remonta ao tempo em que o homem passou a sentir a necessidade de nomear seus primeiros instrumentos e atividades de trabalho.

Um dos mais antigos registros de um trabalho terminológico data do Século XVI, na área de Anatomia, elaborado por Versalius, entre 1514 e 1516. No Século XVIII, ressaltam-se os trabalhos de Lavoisier e Bertholet que colaboram para o estabelecimento da nomenclatura da área de Química e o trabalho de Carl B. Linnaeus, Lineu, na área de Botânica e de Zoologia. Essas obras são válidas ainda hoje, em suas respectivas áreas (FEDOR DE DIEGO, 1995, p. 14).

### 3.1.1 Wüster: a Teoria Geral da Terminologia, ou Terminologia clássica

A Terminologia moderna surge em 1931, com Eugen Wüster, engenheiro e linguísta austríaco, em razão da publicação de sua tese de doutorado sobre Normalização Internacional da Terminologia Técnica (*Internationale Sprachnorming in der Technik*) que demonstra preocupação, principalmente, com a questão metodológica e normativa da Terminologia. A Terminologia de Wüster era uma ferramenta capaz de eliminar as ambiguidades no âmbito das comunicações científicas e técnicas. Com isso, Wüster contribui para o estabelecimento da Teoria Geral da Terminologia (TGT) que, desde então, fundamentou a realização de trabalhos terminológicos (FEDOR DE DIEGO, 1995, p. 18).

Com a repercussão provocada pelos vocabulários especializados das primeiras décadas do Século XX, despertou-se o interesse pelos estudos terminológicos em diversos países da Europa, dando origem à criação de centros de base linguística voltados para os estudos terminológicos. Tal interesse decorre de uma série de fatores que representam influências positivas, especialmente no Canadá e na França, países que, naquele momento, passavam por constantes reformulações.

Dentre os centros de estudos terminológicos mais importantes, encontram-se a *Escola de Terminologia de Viena*, a *Escola de Terminologia de Praga* e a *Escola de Terminologia de Moscou* (FEDOR DE DIEGO, 1995, p. 17).

A *Escola de Terminologia de Viena*, fundada por Eugen Wüster, distingue-se sobremodo. Em razão disso, a publicação em livro, da tese de doutorado de Wüster *Internationale Sprachnorming in der Technik* (Normalização Internacional da Terminologia Técnica), em 1931, representa um marco para a evolução da ciência terminológica. Com base em sua tese, desenvolve-se a Teoria Geral da Terminologia (TGT) que trata da normalização internacional da linguagem técnica.

A Escola de Viena centra sua atenção nos estudos dos conceitos e direciona seus trabalhos para a normalização das noções e termos. Wüster, considerado o pai da Terminologia moderna e o principal representante da Escola de Viena, produziu mais de 500 trabalhos sobre o fazer terminológico. Faleceu em 1977. Felber, Lang e Wersig editam, em 1979, em sua memória, um livro intitulado

*La Terminologia como Ciência Lingüística Aplicada* (FEDOR DE DIEGO, 1995, p. 19).

Nesse panorama, o que se fazia no contexto da chamada Escola de Viena abasteceu a Terminologia também nos países ibero-americanos. As proposições teórico-metodológicas apresentadas com a chancela de Eugen Wüster constituíram-se na conhecida Teoria Geral da Terminologia (TGT) e promoveu as primeiras coordenadas “do saber e do fazer” em Terminologia. Além disso, alcançaram plena aceitação e foram bem recebidas tanto na Europa, como na América. Para Krieger (2004), “isso é incontestável [...] seus fundamentos epistemológicos, seus ideais padronizadores [são apropriados] para assegurar a intercomunicação profissional no plano internacional e do privilégio à dimensão cognitiva dos termos técnico-científicos”.

A *Escola de Moscou*, fundada em 1933 pelo terminólogo Lotte e pelo professor Caplygin, engenheiro e membro da Academia de Ciências, foi bastante influenciada pela obra de Wüster. Lotte é responsável pela publicação de vários trabalhos sobre os aspectos teóricos e metodológicos da Terminologia. Devido à situação da pluralidade linguística, instalada na antiga União Soviética essa Escola passa a interessar-se principalmente pela normalização dos conceitos e termos. Foi também responsável pela promoção de vários eventos, destacando-se o Simpósio Internacional sobre os Problemas Teóricos e Metodológicos da Terminologia realizado em 1979, pela Academia de Ciências da URSS.

A *Escola de Praga* combinou os preceitos de Wüster aos fundamentos do sistema funcionalista, principalmente pelas teorias estruturalistas de Ferdinand de Saussure, dedicava-se à investigação teórica e aplicada da Terminologia. Seu objetivo principal de investigação eram os aspectos funcionais da linguagem, a linguagem-padrão como instrumento de comunicação em todas as áreas da vida social, em particular na área da cultura, da civilização e da tecnologia. Seus estudos contribuíram para distinguir as diferentes tipologias de línguas para propósitos especiais ou línguas de especialidade como: língua falada, técnica, jornalística, poética e outras, além da elaboração de metodologias para assegurar sua estabilidade (FELBER, 1984).

Vale ressaltar que os trabalhos produzidos por essas Escolas colaboraram para a disseminação dos princípios teóricos e metodológicos da Terminologia em diversos países. A importância da Terminologia, enquanto



ferramenta linguística básica de comunicação, entre especialistas e, portanto, vital para troca de conhecimentos e transferência de tecnologias, conduziu à criação de organismos de cooperação internacional em Terminologia. Um destes organismos que tem como objetivo a coordenação das atividades terminológicas em nível internacional, é o Centro Internacional de Informação Terminológica (INFOTERM), criado em 1971. Em razão disso, salientamos a criação da Rede Internacional de Terminologia (TermNET) pelo centro supracitado, em 1977, do qual programa tem por objetivo: a) o desenvolvimento das bases científicas da Terminologia; b) o estabelecimento de uma colaboração mais estreita na preparação de terminologias e seu registro automatizado; e c) o estabelecimento de uma cooperação mais estreita na compilação, registro, processamento e disseminação dos dados e informação terminológica (FEDOR DE DIEGO, 1995, p. 29-30).

### 3.1.2 Teoria Comunicativa da Terminologia: Maria Teresa Cabré e Maria das Graças Krieger

A função dos termos para a Teoria Comunicativa da Terminologia é dupla: representar e transferir o conhecimento especializado em graus e modos distintos, como em situações diversas. As unidades terminológicas, de caráter poliédrico, desempenham funções cognitiva, social, comunicativa e linguística. Enquanto unidades linguísticas, com valor especializado, a Teoria Comunicativa da Terminologia admite a variação denominativa e conceitual. Diferentemente das propostas anteriores de Cabré (1993), a Teoria comunicativa atribui dimensão textual e discursiva à Terminologia (CABRÉ, 1999).

As propostas enunciadas evidenciam que os estudos terminológicos têm procurado refletir sobre as necessidades advindas das alterações no pensamento da ciência em geral, procurando verificar como enfrentar as restrições relativas aos seus fundamentos originais. No Brasil, as proposições teórico-metodológicas da Teoria Geral da Terminologia (TGT) que forneceram as primeiras coordenadas do saber e das práticas em Terminologia, são, agora, submetidas à revisão, sugerindo-se a valorização dos aspectos linguísticos e comunicacionais dos termos técnico-científicos.

[...] O paradigma científico tradicional e o modelo de organização disciplinar do conhecimento que o representava entraram em crise em meados do século XX com o surgimento de objetos de estudo ou situações que não se podiam explicar a partir de uma visão disciplinar. Este fenômeno representou e representa uma profunda mudança de perspectiva e de pensar a realidade [...] (LÓPEZ-HUERTAS, 2007).

A revisão teórica da Terminologia não é um modismo, mas deriva das novas necessidades contemporâneas. Para Krieger e Bevilacqua, a apreensão dos objetos de uma ciência que sofre grandes alterações se revela pela proliferação dos termos técnico-científicos, pela impossibilidade de supor fronteiras rígidas (no plano significante) entre léxico geral e léxico especializado e pelo consequente aumento da complexidade na tarefa de reconhecimento das unidades terminológicas especializadas. De fato, ocorre uma crise do conhecimento estruturado, uma vez que ele não mais se submete a um fechamento, mas se organiza de forma interdisciplinar e multidisciplinar, passando a exigir práticas que redirecionem as formas de abordagem e tratamento das terminologias em dimensões mais amplas, em consonância com a epistemologia das ciências.

Com base em princípios da Linguística, em diálogo com outras áreas de especialidade, desenvolve-se a linguística de *corpus*, experimentam-se teorias funcionalistas e cognitivas sobre a linguagem, passa-se a focalizar o texto e o discurso e a incluir a fraseologia especializada entre as unidades terminológicas (KRIEGER; BEVILACQUA, 2005). Do ponto de vista da Teoria Comunicativa da Terminologia (CABRÉ, 1998; 2000), não se nega a importância da teoria clássica e, embora não se coloque em oposição radical a ela, tende a ressaltar o papel da linguagem na caracterização das unidades terminológicas.

A Teoria Comunicativa da Terminologia, TCT como é conhecida, propõe ver os termos como unidades linguísticas, enfatizando a função da língua como instrumento de comunicação. Essa vertente constitui-se na confluência entre as teorias do conhecimento, as da comunicação e as da linguagem. Ao questionar a autonomia da Terminologia em relação às teorias da linguagem, recupera a noção de termo como unidade de forma e conteúdo que, em determinadas condições discursivas, adquirem valor especializado (CABRÉ, 2000; HJELMSLEV, 2006).

### 3.1.3 Teoria Sociocognitiva da Terminologia: Rita Temmermann

A Teoria Sociocognitiva da Terminologia questiona os princípios da teoria clássica da Terminologia. Rita Temmermann argumenta que esses princípios são orientados exclusivamente à padronização e que este é apenas um dos aspectos a serem analisados no âmbito de uma teoria da terminologia.

O principal objetivo de uma padronização é a unificação de conceitos e termos, o que é uma atividade deliberada, consciente e socioeconomicamente motivada. A padronização uniformiza a compreensão e a comunicação a fim de tornar mercadorias e informações de pronto intercambiáveis.

Em razão disso, Comitês de padronização, compostos de especialistas de uma determinada área, costumam encontrar-se para decidir, em conformidade, a definição precisa de um conceito. Essa definição estabelece sua posição na estrutura conceitual ao expor o termo no que se refere ao conceito superordenado e ao prover as características essenciais e aptas que o diferenciem de outros conceitos no sistema conceitual (TEMMERMANN, 2004).

Na Terminografia orientada à padronização, somente quando um conceito é definido, ou melhor, quando sua posição num sistema estruturado é claramente descrita, em geral ontológica ou logicamente, que o termo preferido é designado. O termo preferido deve ser monossêmico, ou seja, empregado para referir-se apenas a um único conceito (TEMMERMANN, 2004).

Na visão de Temmermann (2004), a Teoria Sociocognitiva da Terminologia reconhece que a padronização é uma atividade importante e necessária à sociedade numa situação específica de comunicação. Por exemplo, quando especialistas com conhecimento sobre uma área, “se reúnem para harmonizar seus conceitos e termos”. Para a NBR 13790 (ABNT, 1997b), harmonização de conceitos é a “atividade de redução ou eliminação de pequenas diferenças entre dois ou mais conceitos muito similares”. Enquanto a harmonização de termos é a “atividade que leva à designação de um conceito, nas diversas línguas, por termos que refletem as mesmas características ou características semelhantes ou têm as mesmas formas ou formas ligeiramente diferentes” (ABNT, 1997b, p. 2).

A base para a delimitação do conteúdo é, conforme perspectiva de Temmermann, o texto no qual está inserido o termo. Disso decorre que o conceito não é universal, nem imutável, “mas a expressão de um conjunto de elementos de natureza lingüística que se consubstanciam num texto que possui não apenas uma dimensão lingüística, mas também pragmática discursiva e comunicativa” (BARROS, 2006, p. 23). Desse modo, é o termo, e não o conceito, o ponto de partida da descrição terminológica.

Na perspectiva acima, dependendo do nível e tipo de especialização do emissor e do receptor, a informação considerada relevante para uma definição varia, o que torna mais flexível a estrutura conceptual, já que as descrições do significado dependem tanto do tipo de unidade de compreensão como dos participantes da comunicação. Consequentemente, a teoria admite a sinonímia e a polissemia no processo de compreensão e de comunicação justificadas pela funcionalidade, pela flexibilidade (os modelos cognitivos se alteram constantemente) e pela diversidade dos processos de categorização (TEMMERMANN, 2004).

A abordagem de Temmermann põe em relevo as relações entre os estudos da terminologia e da linguística textual, reforçando a idéia de que os termos não podem ser compreendidos fora de seu ambiente natural – os textos -, verdadeiros “conjuntos dinâmicos de elementos lingüísticos, pragmáticos, discursivos e comunicativos” (BARROS, 2006, p. 23, 25). A Teoria Sociocognitiva da Terminologia justifica que seu questionamento à Teoria Tradicional de Terminologia da Escola de Viena que tem seus princípios fundados em Wüster, considerado pai da Terminologia moderna, foi o equívoco de ter declarado “os princípios de padronização como [...] os de uma Teoria Geral da Terminologia. [...] que] não levam em conta o fato de que a terminologia também é importante para muitas outras situações comunicativas e cognitivas” (TEMMERMANN, 2004).

Com base nos estudos realizados, verificamos que a Teoria Sociocognitiva da Terminologia propôs um procedimento para a descrição terminológica com base na informação textual e fundamentada em dois parâmetros: a) o conteúdo das linguagens especializadas; b) o perfil dos usuários em potencial das informações terminológicas. Essa teoria reforça que os métodos para um projeto de terminografia também poderiam “ser determinados por esses dois parâmetros e não apenas pela diferenciação semasiológica/onomasiológica tradicionais” (TEMMERMANN, 2004). Dito de outro modo, “os usuários de terminologias

descritivas necessitam de informações que resultem da combinação de uma abordagem semasiológica e onomasiológica em terminografia” (TEMMERMANN, 2004, p. 31-36). Enquanto a primeira fornece informações sobre polissemia e transformações diacrônicas num processo de compreensão, a segunda concentra-se no tratamento de um conjunto de termos relacionados. Os elementos para uma descrição onomasiológica e semasiológica combinada estão postos em textos de linguagem especializada.

#### 3.1.4 Teoria Socioterminológica: Boulanger, Gaudin e Enilde Faulstich

Socioterminologia “é a disciplina que se ocupa da identificação e da categorização das variantes lingüísticas dos termos em diferentes tipos de situação de uso da língua” (FAULSTICH, 1995a, p. 281).

Alpizar Castillo, no evento *Véme Journée ERLA-GLAT*, de maio de 1994, realizado em *Brest*, enfatizou que um planejamento sociolinguístico da terminologia, sem se descuidar do imprescindível trabalho normalizador da terminologia, levando em conta as necessidades de uma comunicação científica e técnica mais fluída e eficiente, permite uma descrição mais consistente dos dados terminológicos (ALPIZAR CASTILLO, 1997).

Ainda, nessa direção, o autor supracitado argumenta que o termo (elemento para a comunicação) e o especialista que o emprega para emitir ou compreender mensagens científicas e técnicas (um falante, afinal de contas) não são quimeras, mas os componentes de uma realidade de multifacetadas. Uma Terminologia bem entendida deve avocar essa realidade como ponto de partida para suas reflexões teóricas e aplicações práticas, “*há de ser una socioterminología,*” assevera Alpizar Castillo (1997 apud SANTOS BORBUJO, 2001, p. 661).

Entretanto, posição semelhante já havia sido defendida por Jean-Claude Boulanger em 1981, em recensão de um livro sobre a normalização linguística. Outra contribuição para esse início refere-se a Yves Gambier que refletiu sobre o tema durante um Colóquio realizado em Paris, em 1986.

A Socioterminologia priorizou o uso, realizando uma ruptura mais incisiva com a Teoria Geral da Terminologia. Os primórdios da Socioterminologia

reportam às críticas ao caráter prescritivo e normatizador da Terminologia feitas por Boulanger, em 1991, e reforçadas por Pierre Auger, em 1993, conforme declara Faulstich (1995a, p. 281-282). Enfatiza as práticas sociodiscursivas, particularmente as profissionais, e o papel dessas na circulação social dos termos (FAULSTICH, 2006); (LARA, 2009, p. 120). Sem dúvida, a primeira obra importante sobre Socioterminologia consistiu a tese de doutorado de François Gaudin, em 1993 (SANTOS BORBUJO, 2001). A partir daí, a Socioterminologia se “firma como uma nova vertente, quando o autor ressalta a necessidade de considerar as práticas lingüísticas e sociais efetivas” (FAULSTICH, 2006).

Gaudin (1993) criticando a influência do pensamento positivista na teoria de Wüster, mostra que a relação entre o signo e a realidade não é direta, mas construída, ou mais exatamente, co-construída no quadro das interações verbais. A referência tem, portanto, um caráter social e, por essa razão, não é possível propor a biunivocidade entre conceito e termo. Para Gaudin, o fato de que as relações entre signo e realidade são relacionadas às situações sociais em que acontecem, justificaria a substituição do conceito de domínio pelo de ‘*episteme*’, mais adequado para descrever a realidade do trabalho científico e das práticas linguísticas que o sustentam, declara Gaudin (1993, grifo do autor).

A Socioterminologia destaca a importância das relações semânticas para um melhor entendimento da estrutura interna do léxico. As definições formais são preteridas em benefício das descrições mais versáteis do significado das palavras. A atitude descritiva prepondera, assim, sobre a prescritiva. Para Gaudin (1993), a Socioterminologia considera as metáforas e os contextos históricos de uso das palavras, o que implica o questionamento da monossemita dos termos (que tem por corolário a reabilitação da polissemia) e da variação como uma “perturbação da unidade lingüística”, na expressão de Wüster (1998, p. 150).

Essas relações entre conceitos e termos estabelecidas conforme a ISO 1087-1 (2000, p. 9-10) são as seguintes: monossemita: relação entre a designação e o conceito onde a designação representa um só conceito; mononímia: relação entre designação e conceito em que o conceito tem somente uma designação; homonímia: relação entre designações e conceitos em que uma designação representa conceitos diferentes; polissemia: relação entre designação e conceito em que uma designação representa dois conceitos ou mais com características comuns; sinonímia: relação entre termos representando o mesmo

conceito, antonímia: relação entre dois termos que representam conceitos opostos e equivalência: relação entre designações que representam o mesmo conceito em línguas diferentes.

Lara (2009, p. 121) ensina que “a Socioterminologia [...] evidencia a importância da história, conduzindo de modo diacrônico, a perspectiva lingüístico-social, a partir da qual se pode pesquisar” aspectos da abordagem onomasiológica. Além disso, “as estratégias da Socioterminologia são estruturadas na perspectiva da integração social dos discursos terminológicos: não só os termos”, mas também as relações sintagmáticas (LARA, 2009, p. 121). Em consonância com a autora supracitada, poder-se-ia supor uma abordagem híbrida entre os procedimentos onomasiológico e semasiológico para a realização de uma das etapas da construção de tesouros, denominada de compilação de termos.

### 3.1.5 Sobre a Rede Ibero-Americana de Terminologia (RITerm)

A RITerm, inicialmente representada pela sigla RIT, criada em 1988, é uma rede de intercâmbio e de trabalho na área da terminologia. Tem por objetivo estabelecer um canal de cooperação entre seus membros para consolidar as terminologias nos países hispanófonos e lusófonos.

Atualmente, a RITerm é um espaço consolidado de intercâmbio e um motor de inovação em Terminologia e desenvolve diversas atividades, entre as quais, conforme sítio da RITerm <<http://www.riterm.net/spip.php?rubrique3>>, acesso em 12 maio 2009, destacamos:

Coleta e intercâmbio de informação terminológica e factual entre os membros da Rede, com o objetivo de fomentar a compatibilização de sistemas, formatos, instrumentos metodológicos etc., na área da terminologia;

Cooperação no planejamento e desenvolvimento de projetos relacionados com a área da terminologia e/ou indústrias dos conhecimentos, bancos de dados terminológicos, lexicográficos, textuais etc., linguística informática, informática aplicada à terminologia, planejamento terminológico etc.;

Criação de um banco de dados terminológicos para a região hispanófona e lusófona;

Colaboração com os programas de formação em terminologia dos membros da Rede, a fim de promover o desenvolvimento e a difusão da terminologia como disciplina (por meio de programas acadêmicos, encontros, boletins informativos etc.);

Realização do Simpósio de Terminologia, com periodicidade de dois anos, e organizado ao mesmo tempo que a Assembléia Geral (Caracas-1988, Brasília-1990, San Millán de la Cogolla-1992, Buenos Aires-1994, México D.F.-1996, Havana-1998). O VII Simpósio Ibero-americano de Terminologia ocorreu em Lisboa, Portugal, em novembro de 2000, o VIII em Cartagena de Índias, Colômbia, em novembro de 2002, o IX em Barcelona, Espanha, em novembro de 2004, o X, em Montevideú, Uruguai, em novembro de 2006 e o XI ocorreu em Lima, Peru, em outubro de 2008.

#### 3.1.5.1 Algumas considerações acerca da Terminologia no Brasil

Nessa direção, ressaltamos a atuação do Grupo de Trabalho (GT) da Associação Nacional de Pós-Graduação em Letras e Linguística (ANPOLL) que, em 1988, incorporou a Terminologia ao Grupo de Trabalho de Lexicologia e Lexicografia, oficializando os estudos realizados por alguns integrantes do GT nessa área e colaborando para o desenvolvimento das pesquisas terminológicas em diversas universidades brasileiras.

Evidenciamos também, em 1987, a realização do I Encontro Brasileiro de Terminologia Técnico-Científica, juntamente com II Simpósio Ibero-Americano de Terminologia, em Brasília. Esses eventos consolidaram as pesquisas terminológicas no Brasil e na América Latina.

Em 1992, com a criação da Comissão de Estudo Especial Temporária de Terminologia (CEET), apoiada pela Associação Brasileira de Normas Técnicas (ABNT) e pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), as pesquisas terminológicas se consolidam e recebem um novo impulso para a divulgação e difusão dos estudos para fora do país (MELLO, 2002, p. 12). Dentre os trabalhos realizados pela CEET, evidenciamos a elaboração das seguintes Normas Brasileiras, a saber:



A NBR 13789 (ABNT, 1997a), intitulada Terminologia – Princípios e métodos – Elaboração e apresentação de normas de terminologia, fundamentadas na ISO 10241 de 1992, fornece os princípios e métodos para elaboração de terminologias normalizadas; e

A NBR 13790 (ABNT, 1997b), intitulada de Terminologia – Princípios e métodos – Harmonização de conceitos e termos, fundamentada na Norma ISO 860 de 1996, tem por objetivo fixar o enfoque metodológico adotado na harmonização dos conceitos, dos sistemas de conceitos e dos sistemas de termos. Aplica-se, principalmente, à elaboração de glossários nacionais e internacionais (ABNT, 1997b).

Cabe esclarecer que, as normas mencionadas, ambas válidas a partir de 30 de abril de 1997, se encontravam em vigor, conforme resposta à consulta feita ao escritório da ABNT em 08 de janeiro de 2009.

Em razão das contribuições dos autores visitados nos capítulos anteriores, podemos aferir que a apropriação das contribuições teórico-metodológicas da Terminologia leva ao aperfeiçoamento das metodologias de construção de tesouros. As normas terminológicas NBR 13789 (ABNT, 1997a) e MBR 13790 (ABNT, 1997b) complementam as Normas de construção de tesouro possibilitando um melhor entendimento dos conceitos, dos relacionamentos entre os conceitos, e também da modelagem dos sistemas conceituais (LARA, 2006). Cabe reforçar ainda que o uso da terminologia concreta, por meio de dicionários técnicos e especializados e de glossários, representa a possibilidade de conferir referencial sólido à interpretação dos descritores dos tesouros.

Para Faulstich (1995b, p. 279), outro fato importante que contribuiu para o desenvolvimento das pesquisas terminológicas no Brasil foi à criação do Mercado Comum do Sul (Mercosul). Em razão disso, o Mercosul colaborou para o desenvolvimento de terminologias nacionais dando enfoque às línguas espanhola e portuguesa que conquistaram espaço e atualmente, são focalizadas como línguas de intercomunicação da ciência e da cultura. Além disso, com a instalação do Banco Terminológico do Mercosul, surgiram projetos cooperativos de pesquisa, contribuindo para a expansão da Terminologia, tanto no nível de graduação como no de pós-graduação.

A esse respeito, Krieger (2004) assevera que dado o avanço percebido “talvez poucos campos de estudo tenham progredido tanto como a

Terminologia nos últimos 15 anos”. A autora supracitada explica que esse “período cronológico [...] toma por base o Simpósio da Rede Ibero-americana de Terminologia, mais conhecida como RITerm, ocorrido no Brasil em 1990”. Os trabalhos apresentados no evento davam conta de muitos aspectos dos termos e das interfaces da Terminologia, como a Documentação e a Tradução.

A autora supracitada esclarece que “o Brasil integra o mundo ibero-americano”, no qual o progresso da Terminologia é considerável. Associa-se ao reconhecimento de necessidades de organização sistêmica de repertórios terminológicos, como bancos de dados, entre outros produtos terminográficos, de modo a facilitar a comunicação entre especialistas, em nível nacional e internacional. Para tratar essa questão, torna-se importante focalizá-la na interface produzida pela terminografia.

Sob esse enfoque, Krieger (2004, grifo da autora) diz que houve várias propostas de realização de projetos, bem como de descrição de problemas para “lidar” com os termos técnico-científicos, evidenciando-se, porém, sempre, com raríssimas exceções, o papel normativo dos instrumentos terminológicos. Essa visão padronizadora motivava a maioria dos trabalhos apresentados, salientando-se que este enfoque não era exclusivo do Brasil. Em geral, essas perspectivas normalizadoras também conformavam as comunicações do mundo ibero-americano.

Na atualidade, o que se percebe, em se tratando de terminologia ibero-americana, é que essa área superou o plano das proposições, apresentou resultados de aplicações concretas, sem preocupações prescritivas, desenvolveu teorias, assumiu postulados e promoveu avanços importantes em investigações com luz própria. (KRIEGER, 2004). Em razão dessas evidências, verifica-se que esse é um dos espaços geográficos de maior representatividade do trabalho terminológico, sob o ponto de vista seja da investigação pura, seja dos resultados de aplicação, traduzidos sob diferentes formas de atuação.

A norma ISO 1087-1 (2000), define Terminologia como um conjunto de termos que representa o sistema de conceitos de um domínio particular. Essa definição refere-se à parte aplicada da Terminologia. A norma supracitada apresenta também a expressão Ciência da Terminologia que é definida como “estudo científico dos conceitos e dos termos em uso nas línguas de especialidade”, concedendo à Terminologia o status de ciência, devendo submeter-se ela a todo rigor exigido de uma ciência.

O termo, contudo, é objeto da Terminografia e um dos componentes das linguagens documentárias. É sob este ponto de vista que também aqui se focaliza a criação de novos termos, análise, processamento, apresentação e uso na comunicação especializada. A unidade terminológica, ou seja, o termo enquanto objeto de análise ou enquanto produto para elaboração de instrumentos terminológicos forma unidades (de termos) que correspondem, cada uma, a um nó cognitivo dentro de um campo de especialidade. O conjunto desses nós conectados por relações específicas (causa-efeito, todo-parte, contiguidade, anterioridade-posterioridade, entre outras) constitui a representação conceitual de dita especialidade (CABRÉ, 1999, p. 238).

O termo, do ponto de vista da norma ISO 1087-1 (2000, p. 24), é a “enunciação lingüística de um conceito, com ele identificando-se”. Dessa forma, pode ser constituído por uma ou mais palavras que representam um conceito geral em uma língua de especialidade. Em vista da sua importância, Aubert (2001) ressalta a necessidade de maior precisão no trabalho terminológico, uma vez que as terminologias constituem a base para: a) o ordenamento do conhecimento, ou seja, classificação conceptual de cada disciplina científica ou ramo de atividade humana; b) a transferência de conhecimentos e de *know-how* tecnológico; c) a formulação e disseminação de informações especializadas (redação e publicação científicas); d) a transferência de textos científicos para outros idiomas (pela tradução e pela interpretação); e) a armazenagem e recuperação de informação especializada (por meio de linguagem de busca, tesouros, índices, classificações, inclusive de dados eletrônicos).

Cabré (1999, p. 239) corrobora esses dizeres enfatizando que “a terminologia, seja qual for sua temática ou o contexto em que se produz, cumpre sistematicamente duas funções: a função de representar o conhecimento, especializado, e a de transmiti-lo”. Com isso, observa-se que a Terminologia exerce um papel de grande relevância social, contribuindo para uma comunicação mais precisa entre especialistas de determinada área. A Terminologia origina-se sob a influência de duas tendências: a interdisciplinaridade das ciências e as suas micro-especializações.

A primeira destaca como consequência a necessidade de normalização terminológica, pois a univocidade dos termos é condição essencial para a comunicação entre especialistas, seja no interior de uma mesma disciplina,

seja entre disciplinas diferentes. Já a segunda, acentua a criação de neologismos, termos novos ou antigos com nova significação, como por exemplo, o termo vírus, originalmente da área médica e que foi tomado por empréstimo pela área da informática (RONDEAU, 1984, p. 60; CABRÉ, 1991, p. 56).

A Terminologia constitui-se de estudos dinâmicos, porque se encontra a serviço da Ciência e da Tecnologia que se desenvolvem com muita rapidez, provocando constantemente o surgimento de novos conceitos, novos termos que precisam ser identificados e confirmados em benefício do progresso da própria Ciência e Tecnologia. A dinâmica destas temáticas que estão em crescimento contínuo, representa diferentes problemas terminológicos a tradutores (ausência de equivalências para neologismos na língua de chegada), comunicadores (normalização de neologia no texto jornalístico) e bibliotecários (desatualização de linguagens documentais divididas por disciplinas e carência de classificações para assuntos de áreas interdisciplinares para a indexação) (BARITÉ, 2005, p. 25), correspondem a algumas das dificuldades apontadas por essa dinamicidade.

### 3.2 Do conceito ao termo: a natureza metodológica da Terminografia

As unidades terminológicas fazem parte da linguagem natural, não se considerando os termos como entidades autônomas que formam um léxico especializado diferenciado. Os termos são unidades denominativo-conceptuais que têm capacidade de referência desde que ativados em seu uso em contextos e situações determinados (CABRÉ, 2000). Quando ativados, os conceitos de um mesmo âmbito especializado mantêm entre si relações que constituem a estrutura conceptual do campo de assunto, dependendo do valor de cada termo, de sua posição relativa, tal como acontece com as unidades da linguagem geral.

A terminologia entendida como “um conjunto de termos e como um sistema para a recuperação da informação documental, respectivamente, compartilham duas funções: a função de representação e a função de transferência do conhecimento especializado” (ADELSTEIN, FELIU, 2004, p. 115).

Para cumprir essas funções: de representação e de transferência do conhecimento especializado, o terminólogo, para desenvolver um trabalho terminográfico, necessita do auxílio do especialista da área em foco em diversos momentos, que vão desde a indicação de fontes de referência de onde se coletam os candidatos a termos até a validação de termos que comporão o instrumento terminológico, dicionário ou glossário, além da participação na elaboração/validação de suas definições. No entanto, nos momentos da localização das fontes de compilação de candidatos a termos, organização das informações coletadas da área de especialidade, estruturação do instrumento, o terminólogo atua sozinho (TELINÉ; ALMEIDA; ALUÍSIO, 2003)

Nessa direção, entendemos que como um vértice com efeito de integração que pode ser produzido pela aproximação e compartilhamento, destacado pelos vários autores citados, e também pelo fato de o terminólogo poder contar, também com a colaboração do profissional da Documentação (Ciência da Informação) em certos momentos, tais como: acesso a recursos informacionais, altamente especializados e atualizados, colaboração na categorização temática ou funcional, além disso, na organização de diferentes formas de acesso aos conteúdos informacionais.

Ainda que não seja ocupação da Ciência da Informação construir, glossários, com certa frequência, na atividade documentária, encontram-se universos temáticos para os quais não existem vocabulários formalizados. Por esse motivo, ocasionalmente, cabe recorrer ao uso de orientações terminológicas e terminográficas para a exploração de *corpus* representativo e para a identificação das redes relacionais entre os termos. Esse procedimento, na visão de Lara (2006), mostra-se fundamental para melhorar o conhecimento dos universos a serem trabalhados documentariamente “visando, seja a construção de instrumentos de intermediação como os tesouros, seja a organização direta de conjuntos informacionais, como é o caso dos conteúdos de sites” (LARA, 2006).

Conforme Lara (2004, p. 233) “quando os universos-focos são especializados - domínios do saber ou áreas de atividade - o papel da Terminologia é bem claro, fornecendo referencial concreto para a interpretação dos termos, [...]”, do modo como eles são definidos em cada um dos domínios de especialidade (em suas terminologias). A autora supracitada esclarece ainda que:

As definições desses termos respaldam a organização das redes de relacionamento entre eles. Porém, mesmo quando não se trata de universos especializados, ou seja, trata-se de universos cujos contornos não são bem determinados, os princípios terminológicos, associados aos documentários, são essenciais para referir o processo de organização espacial e visual dos elementos do repertório em jogo (LARA, 2004).

A relação da Terminologia<sup>3</sup> com a Organização e Representação da Informação envolve a possibilidade que ela tem de operacionalizar a sua organização sistêmica. De outro modo, a terminologia refere-se ao conjunto de termos de uma área de especialidade próprios de uma ciência, arte, técnica, profissão, por exemplo: a terminologia da Biblioteconomia, da Arquivologia, da Informática, do Jornalismo e outras. Enquanto a primeira dá apoio à organização estrutural de linguagens documentárias, a segunda é fundamental para a representação de sua macroorganização.

Em razão disso, as autoras Tálamo e Lara (2006; 2007), em diferentes estudos, argumentam que a Terminologia teórico-metodológica “contribui para o aperfeiçoamento das metodologias de construção de tesouros [...] porque fornece as bases para o entendimento do conceito e das unidades terminológicas, do sistema conceitual e das redes relacionais de natureza lógico-semântica e pragmática entre os termos que responde pela estruturação desses instrumentos”. Além disso, as autoras supracitadas também esclarecem que a terminologia concreta, por sua vez, “garante as referências para a interpretação dos descritores dos tesouros”.

Em confronto com a Terminologia, a Linguística Documentária propõe a questão que parece ser fundamental em nível de comunicação especializada, sobre o efetivo retorno social dos trabalhos descritivos levado a cabo sobre os textos de especialidade no que tange a sua contribuição para o desenvolvimento de efetiva cultura informacional para a sociedade e não apenas para segmentos delimitados (TÁLAMO; LARA, 2006; 2007).

Nessa direção, realiza-se o entendimento da “Linguística Documentária com a Terminologia que, enquanto campo de estudos, observa os discursos das áreas de especialidade propondo metodologias para a descrição de

---

<sup>3</sup> Terminologia com T maiúsculo abrange estudos da terminologia teórico-metodológica. Com t minúsculo diz respeito à terminologia concreta.

seus termos, com a finalidade de estruturar o campo nocional da especialidade” (LARA; TÁLAMO, 2007, grifos das autoras).

As autoras supracitadas esclarecem que os “discursos das áreas de especialidade” até bem pouco tempo eram identificados com as 'linguagens de especialidade', que se caracterizariam por 'peculiaridades especiais' como a temática, os tipos de interlocutores, as situações de comunicação e de intercâmbio, e outros (CABRÉ, 1993), são, atualmente, compreendidos como construções linguísticas em que os léxicos muitas vezes se confundem com o léxico comum (KRIEGER, 2001).

Por esse motivo, a Terminologia contemporânea propõe observar as terminologias *in vivo* por meio da análise de *corpora* discursivos (BOULANGER, 1995). Desse modo, Lara e Tálamo (2007) enfatizam que a atuação na “interface entre a Linguística Documentária e a Terminologia visa principalmente contribuir para o aperfeiçoamento das metodologias de organização da informação via linguagem documentária”.

### 3.3 Metodologia da Terminografia: atividade de pesquisa terminológica

A norma ISO 1087-1 (2000, p. 15) apresenta a seguinte definição para o termo Terminografia: “menção, tratamento e apresentação dos dados terminológicos obtidos pela pesquisa terminológica”. Dados terminológicos são “dados relativos a um conceito ou a sua designação”. Em nota, a Norma supracitada informa que o termo ‘terminografia’ substitui as expressões ‘lexicografia terminológica’ e ‘lexicografia especial’ caídas em desuso. Esclarecemos, também, que o termo Lexicografia refere-se a “um ramo aplicado da Lexicologia que se ocupa da elaboração de dicionários”, enquanto que a Terminografia “é o ramo da Terminologia que se ocupa, também, da elaboração de dicionários especializados, ou de glossários terminológicos” (CABRÉ, 2004, p. 20).

Terminografia é uma matéria fundamental em todas as atividades relacionadas ao conhecimento e, por esse motivo, representa uma interface chave para a Documentação. Para Alicia Fedor de Diego (1995), a Terminografia “é uma ciência eminentemente aplicada que conta com seus princípios e métodos. Estuda

os dados terminológicos segundo os preceitos teóricos e desenvolve instrumentos para sua ordenação e representação em sistemas de informação” (FEDOR DE DIEGO, 1995, p. 108). Alertamos que a definição para Terminografia exposta pela autora supracitada não está em total conformidade com o nosso entendimento, neste estudo. Vale esclarecer que concordamos com a afirmação “estuda os dados terminológicos segundo os preceitos teóricos e desenvolve instrumentos para sua ordenação e representação em sistemas de informação”, mas a refutamos quando menciona a condição de ciência da Terminografia. Por esse motivo, estamos em consonância às apresentadas por Cabré (2004), ou seja, a Terminografia é uma ramificação da Terminologia.

A Terminografia preocupa-se com a prática terminológica. Sua atividade consiste em recolher, constituir, organizar e difundir os termos e as noções de uma mesma área sob a forma de glossários, dicionários terminológicos e outros (BARBOSA, 1990; DIAS; 1999; MACIEL, 2001). Com efeito, investiga as unidades terminológicas que representam uma área específica, com dados procedentes de *corpus* especializado. A macroestrutura, ou seja as entradas das obras terminográficas costumam ser organizadas em ordem alfabética dentro dos campos conceituais que têm a função de cabeçalhos para facilitar e agilizar a consulta.

Conforme Barbosa (1990), a Terminografia tem a função de registrar a padronização para possibilitar uma comunicação profissional precisa, buscando validação em organizações e comissões de normalização e no parecer de especialistas. Toda atividade terminológica, desde a identificação de termos até a entrega de um produto final, pode ser feita manualmente (PAVEL; NOLET, 2002).

O desenvolvimento acelerado da Informática, contudo, contribui para a automatização propiciando maior rapidez na realização dos trabalhos terminológicos. Da ligação entre Terminologia e Informática, originou-se uma nova área de investigação, designada Terminótica ou Terminologia automatizada que pode ser entendida como o conjunto de operações automatizadas de tratamento dos termos (MACIEL, 2001).

Um dos requisitos da Terminótica é a elaboração de uma ficha terminológica, dotada de vários campos que podem conter informações de natureza documental, terminológica e linguística. A ficha terminológica vem a ser o primeiro passo para o registro terminológico por meio da inserção do termo num programa de



computador. Desse modo, tem-se como objeto o tratamento automático do termo, ou seja, da unidade terminológica. Constitui um componente da terminografia atual.

Retomamos a revisão de literatura a fim de delimitar os aspectos teóricos apresentados anteriormente que permitiram visitar diversas metodologias cuja contribuição oferece suporte à construção de tesouros. Percebemos, contudo, a necessidade de aprimoramento das metodologias existentes, relacionadas à seleção e identificação de termos para essa construção.

Incorporamos as recomendações metodológicas da Terminologia, especialmente as orientações da Terminografia, apoiadas, sobretudo na Teoria do Conceito que orienta a identificação de termos candidatos com base no contexto de uso, de áreas de especialidade a fim de verificar o processo de seleção e identificação de termos que podem ser utilizados para a construção de tesouros compatíveis com a terminologia de áreas de especialidade.

### 3.3.1 Abordagem onomasiológica e semasiológica

As pesquisas terminológicas, em geral, podem ser semasiológica ou onomasiológica. Na abordagem semasiológica, parte-se do significante para estudar o significado. Ou melhor, toma como ponto de partida a palavra, com seus vários significados. Enquanto na abordagem onomasiológica, parte-se do significado, o conceito, para estudar o significante, a denominação. A incorporação do referente representado pelo termo é significativa da abordagem denominada onomasiológica (CAMPOS, 2001).

A forma de abordagem do termo no tesouro e na terminologia diz respeito ao ato da compilação de termos em que os termos são tomados com um significado próprio, atribuído pelo contexto. Desse modo, “esse significado é que vai ser trabalhado, ou seja, o tratamento do termo é feito a partir desse significado assumido previamente” (CAMPOS, 2001, p. 117). Com efeito, é isso que determina a abordagem adotada, nesse caso, é a onomasiológica, que toma como ponto de partida do conceito para o termo.

Como os termos são tomados com um significado próprio, num dado contexto, destaca-se a indissociável relação entre linguagem, pensamento e

realidade. Essa evidência se realiza via conceitos, e não entre palavras, dada a relação direta entre o conceito e o termo, ou seja, um conceito é representado por um termo (símbolo) e este termo é usado para designar aquele conceito.

Wüster incorpora essa idéia ao afirmar que o conceito é uma unidade de pensamento. Do mesmo modo, Ranganathan identifica notação e termo. Contudo, “nos tesouros terminológicos, com a adoção da Teoria do Conceito, pela qual este é visto como uma tríade referente-características-termo, essa relação ocorre de forma mais consistente” (CAMPOS, 2001, p.118).

A metodologia utilizada procura associar procedimentos presentes nos trabalhos terminológicos, terminográficos e documentários, vinculando as operações terminológico-documentárias à idéia de que a organização do conhecimento tem como objetivo precípuo o estabelecimento de mediações para acesso, circulação, distribuição e reutilização da informação conforme enunciado pelo ciclo social da informação, um dos fundamentos da própria Ciência da Informação.

De modo concreto, a proposta didático-pedagógica visa a criar condições para que os alunos identifiquem os diferentes tipos de texto, compreendam o significado da linguagem de especialidade no conjunto da língua, sua importância na expressão do compartilhamento da linguagem como dos conteúdos informacionais, além do papel das unidades terminológicas e dos conceitos na estruturação dos significados, como formas de apoio à construção das redes relacionais entre os descritores da linguagem documentária.

Lara e Tálamo (2007) propõem, inicialmente, a análise de pequenos *corpora* que reúnem bibliografia de características especializada (de textos de divulgação científica para científicos propriamente ditos), em geral sobre uma temática atual. Após sua leitura, são identificados os termos (algumas vezes, em frases), considerados mais importantes para a compreensão do texto e, seguindo discussão coletiva, são selecionados os termos do vocabulário conceitual e do vocabulário funcional, sendo que, sobre os primeiros, incidirão os procedimentos de descrição terminológica.

Os termos são registrados em fichas terminológicas de coleta (simplificadas, tendo em vista os objetivos documentários) com o termo candidato, suas definições, (se existentes), contextos de uso (transcrições literais do termo no

enunciado em que aparecem), fonte, domínio, responsável pelo preenchimento e data.

Nessa direção, registra-se a importância do uso de procedimentos terminológicos e terminográficos para sistematizar os conceitos explorados na realização de monografias sob temas diversos. Rondeau (1984) apresenta ainda a abordagem mista com a possibilidade de combinar os tipos de abordagens. Do mesmo modo, para além desses benefícios, ressalta-se, por fim, que a Linguística Documentária pode contribuir efetivamente para a proposição de metodologias para a elaboração de produtos híbridos (LARA; TÁLAMO, 2007).

### 3.3.2 Etapas da pesquisa terminológica

O desenvolvimento de trabalho terminológico, em seu aspecto terminográfico, respalda-se nas orientações da Terminografia o que abrange um conjunto de etapas para a sua realização. Entendemos por pesquisa terminológica o conjunto de operações que compreende: a coleta, tratamento e difusão dos dados terminológicos (RONDEAU, 1984, p. 64). A pesquisa terminológica pode ser pontual monolíngue/plurilíngue ou temática monolíngue/plurilíngue. Concordamos que a pesquisa pontual visa à solução de problemas isolados de designação, se desenvolve num horizonte estreito, particular e tem como vantagem mais evidente proporcionar soluções rápidas para problemas específicos. Já a pesquisa temática se propõe a efetuar o levantamento do vocabulário terminológico de uma determinada atividade, especialidade, técnica (AUBERT, 2001, p. 47).

Tanto a pesquisa pontual como a temática, possuem sua própria metodologia de pesquisa que deve ser considerada quando da realização do trabalho terminológico. A pesquisa pontual tem como objetivo resolver um problema isolado de designação que chega até o terminólogo por meio de um usuário. Em geral, trata-se de um profissional relacionado com a linguagem, e que pode ser um redator técnico, ou um especialista de um domínio científico-técnico que necessita encontrar uma denominação precisa para um conceito ou, ao contrário, a que conceito se refere determinada denominação.

A pesquisa temática é marcada por duas características básicas: 1) o objetivo - investigação de forma exaustiva ou básica do conjunto de termos em relação a um ramo de atividades, e que pode explorar o universo de uma mesma língua ou de duas ou mais línguas; 2) o tempo de realização - que poderá ser mais ou menos prolongado, dependendo da extensão da pesquisa. Importante destacar, também, que um trabalho terminológico pode ter um caráter descritivo - cuja função é compilar um conjunto de termos pertencentes a um campo de especialidade e colocar à disposição dos usuários; ou um caráter prescritivo - que se propõe priorizar o uso de termos considerados recomendáveis com a finalidade de orientar os falantes sobre o uso da terminologia correta em uma determinada área (FELBER, 1987, p. 10-11).

Rondeau (1984, p. 70-77) apresenta o conjunto de etapas da pesquisa terminológica temática como resultado de sua investigação sobre métodos e procedimentos de pesquisa terminológica realizada com estudiosos de diversos países, a saber:

1. Escolha do domínio e da língua do trabalho – a escolha do domínio e da língua de trabalho, geralmente, são estabelecidas em função das necessidades dos usuários.

2. Delimitação do subdomínio – recomenda-se não desenvolver uma pesquisa terminológica sobre um domínio completo: por um lado devido à complexidade e amplitude que supõe uma tal tarefa; e por outro lado, porque em grande parte do tempo, um domínio compreende não somente uma rede nocional que lhe é própria, mas também numerosas redes nocionais conexas.

3. Consulta a especialista do domínio/subdomínio – a função do especialista do subdomínio nessa etapa consiste essencialmente em auxiliar o pesquisador na delimitação do subdomínio e em orientá-lo quanto à escolha do *corpus* do trabalho terminológico.

4. Coleta do *corpus* do trabalho terminológico – a etapa da coleta do *corpus* do trabalho terminológico tem a finalidade de reunir os documentos especializados necessários para o desenvolvimento da pesquisa terminológica. Rondeau (1984, p. 50-51) estabelece oito categorias de documentos de conteúdo terminológico, ou seja, nos quais se encontram termos: 1) normas internacionais ou nacionais; 2) manuais, catálogos, guias de utilização de produtos, entre outros; 3) livros e revistas especializados; anais de eventos científicos, relatórios de pesquisa,

teses, entre outros; 4) vocabulários, *thesaurus*, glossários, léxicos; 5) dicionários gerais e especializados, de língua ou enciclopédicos, enciclopédias, entre outros; 6) bancos de termos, fichários automatizados ou não; 7) consulta a especialistas da área; 8) bibliografias ou listas relacionadas com o domínio. Os materiais utilizados como fontes que dão origem ao *corpus* do trabalho terminológico devem respeitar os princípios da atividade terminológica no que tange à confiabilidade e representatividade.

Cabré (1993, p. 278, tradução nossa) indica algumas condições que julga relevantes com relação às fontes de pesquisa de termos: a) suficientemente representativa na área, de acordo com os objetivos do trabalho e a delimitação do tema, para que permitam elaborar uma lista prévia de unidades significativas dos conteúdos da matéria; b) atualidade, tanto no que se refere às denominações utilizadas pelos especialistas como à informação dos conteúdos que toda disciplina pode e costuma mudar continuamente; c) suficientemente explícitas, para que permitam recuperar a identificação e a informação de um documento em qualquer momento da compilação ou difusão dos dados terminológicos.

5. Estabelecimento da árvore de domínio – a árvore de domínio representa o conjunto nocional que tem a função de situar o campo nocional a ser estudado. Cabe alertar que, antes de estabelecer a árvore de domínio, o pesquisador deverá consultar os seguintes documentos: sistemas de classificação, glossários, entre outros. Alerta-se, ainda, que, em alguns domínios, esses instrumentos são até abundantes, mas, em outros domínios podem não existir.

6. Expansão da representação do domínio escolhido – a etapa de expansão do domínio escolhido normalmente decorre das etapas 2 (delimitação do subdomínio) e 5 (estabelecimento da árvore de domínio). Nesse momento, é necessário o auxílio dos especialistas do domínio para direcionar os trabalhos próprios dessa etapa e também para verificar os resultados obtidos na etapa precedente.

7. Estabelecimento dos limites da pesquisa terminológica – o limite da extensão da pesquisa terminológica, quanto ao número aproximado de termos, é estabelecido em função dos objetivos propostos, das disponibilidades de tempo e de meios financeiros. Desse modo, pode-se escolher um levantamento básico compilando em média de 300 termos, ou exaustivo, por volta de 2.500 termos.

8. Coleta e classificação de termos – a coleta de termos efetua-se a partir do *corpus* do trabalho terminológico selecionado. Consiste, geralmente, em fazer uma leitura do texto assinalando as unidades terminológicas a extrair. Essa operação requer da parte do pesquisador algum conhecimento metodológico do trabalho terminológico e também algum conhecimento sobre o domínio ou subdomínio. De acordo com a norma ISO 1087-1 (2000), que estabelece os critérios para a coleta dos termos e recorte do contexto de uso, o contexto é o “enunciado no qual figura o termo estudado” ou parte de um texto no qual ocorre o termo. Nesse sentido, o contexto tem um papel fundamental nas operações de coleta dos termos porque permite reduzir os riscos de erros no momento da identificação e recorte do termo.

De acordo com Rondeau (1984, p. 80), para identificar e recortar um contexto deve-se levar em conta os seguintes aspectos: o contexto deve ser conciso, mas completo; deve conter a unidade terminológica a ilustrar; deve ser colhido de uma fonte confiável; deve ser definitório ou descritivo a fim de apresentar elementos para definição do termo; deve colocar em evidência traços semânticos da unidade terminológica; deve ilustrar o comportamento sintático da unidade terminológica no texto.

Nessa direção, o autor supracitado apresenta três tipos de contextos que podem ser utilizados na prática terminológica: o contexto definitório, o contexto explicativo e o contexto associativo. O contexto definitório fornece dados precisos sobre o conceito do termo estudado. Já o contexto explicativo revela a natureza, o objetivo ou um aspecto do conceito estudado. Enquanto o contexto associativo se caracteriza pela ausência de descritores significativos do contexto, ele permite apenas retomar o termo estudado do campo de aplicação por associação com os termos que interage (RONDEAU, 1984, p. 80).

A norma ISO 1087-1 (2000) afirma que o vocábulo termo é definido como “enunciação lingüística de um conceito, com ele identificando-se.” Pode ser constituído por uma palavra ou grupo de palavras, de números ou até mesmo conter símbolos. A dificuldade reside no reconhecimento de um termo multivocabular (forma composta de duas ou mais palavras) como uma unidade terminológica.

Para auxiliar o pesquisador nesse reconhecimento, Aubert (2001) sugere algumas pistas: a) recursos gráficos e de leiaute: compreende um dos recursos usados pelo escritor para chamar a atenção do leitor para conceitos

básicos que são: aspas, negritos, itálicos e outros, ou a ocorrência de termos em posição de destaque no texto como títulos e subtítulos; b) frequência estatística - diz respeito à recorrência, com frequência do termo na mesma configuração; c) estratégias discursivas - quando o próprio texto esclarece, conceituando, explicando, definindo a unidade lexical, deixando claro que é uma unidade terminológica.

Cabe ao pesquisador verificar se o termo pertence à área do domínio. Contudo, é de suma importância recorrer a um especialista da área em estudo, porque somente ele poderá auxiliar na delimitação do sintagma e no reconhecimento desse como um termo empregado na área.

Para a elaboração das fichas de coleta de termos, ao se deparar com um termo, cabe ao pesquisador fazer a anotação do fato numa ficha. Essa ficha pode ser de dois tipos: ficha de detecção - em que aparece apenas o termo e o lugar onde foi identificado; ou ficha de citação - em que aparece o termo e seu contexto de uso (FELBER, 1987, p. 277). A ficha terminológica é um instrumento de anotação, a partir do qual se desenvolve o trabalho de análise terminológica.

Existem alguns componentes que são obrigatórios, outros são facultativos. Os componentes que constituem a ficha terminológica são: 1) de natureza documental: domínio(s), subdomínio(s), fonte, nome do autor; 2) de natureza terminológica: termo-entrada, nome científico, definição, contexto, entre outros; 3) de natureza linguística: categoria gramatical, variante gráfica, termos remissivos, sinônimos, nota(s), normalização.

Evidenciamos que existem diversos modelos de fichas terminológicas. Cada organismo, (e até cada pesquisador) estabelece sua própria ficha conforme suas necessidades. Os campos da ficha terminológica podem ser muito numerosos. Contudo, Rondeau (1984, p. 82-84) destaca os seguintes campos como principais: a) termo-entrada; b) contexto ou definição; c) fontes; d) domínio/subdomínio; e) informações adicionais. Dessa forma, sugere, para essa fase, uma ficha terminológica padrão para a realização dos registros terminológicos.

9. Verificação e classificação da noção/denominação – as operações da etapa anterior (item 8 – coleta e classificação de termos) levam a uma classificação provisória dos termos e a uma exploração sumária das noções que eles representam. Nessa etapa, cada noção é retomada com o objetivo de ser analisada. Essa análise refere-se: a) ao seu conteúdo através de comparações entre as definições e os contextos; b) ao seu lugar na rede nocional do domínio ou do

subdomínio. Destaca-se que as operações realizadas nessa etapa levam a um resultado triplo: a) delimitação mais precisa do termo com as referências dos documentos sobre os quais se embasou para esta delimitação; b) classificação definitiva dos termos; c) agrupamento dos sinônimos.

10. Trabalhos de apresentação de dados terminológicos – esta etapa refere-se à apresentação da ficha terminológica padrão que consiste em um conjunto de informações sobre os termos próprios de um determinado domínio. A começar da ficha terminológica padrão, devidamente preenchida, é possível partir para a construção de instrumentos terminológicos como glossários especializados e técnicos, ainda, tesouros terminológicos documentários.

- **Síntese**

Com base nos referenciais estudados e nas experimentações feitas em colaboração em Terminologia de áreas de especialidade em: Gestão da informação; Moda; Tecnologia de Alimentos, e outras. Acreditamos que o conjunto das etapas sistematizadas por Rondeau (1984) fundamentado em investigação realizada com diversos estudiosos da Terminologia como Pierre Auger, Alain Rey, C. Dubois, Juan Carlos Sager e outros, reúnem as condições essenciais e permitem a realização de análise terminológica em conformidade com os pressupostos estabelecidos pela pesquisa terminológica temática, orientada pela Terminografia.

Para efeito desse estudo, apresentamos uma síntese do conjunto de etapas da pesquisa terminológica temática, sugerido por Rondeau (1984, p. 70-77), composto de 10 fases/estágios conforme descrevemos a seguir:

- 1) escolha do domínio e da língua do trabalho;
- 2) delimitação do subdomínio;
- 3) consulta a especialista do domínio/subdomínio;
- 4) coleta do *corpus* do trabalho terminológico;
- 5) estabelecimento da árvore de domínio;
- 6) expansão da representação do domínio escolhido;
- 7) estabelecimento dos limites da pesquisa terminológica temática;



- 8) coleta e classificação de termos;
- 9) verificação, classificação e confirmação de termos;
- 10) trabalho de apresentação de dados terminológicos.

Apresentamos, a seguir, uma sequência de etapas, além de orientações sobre a construção de tesouros recomendadas por Tálamo (1997c) em um curso ministrado pela autora supracitada, sobre *Elaboração e Uso dos Tesouros*, a saber;

1) Definição do domínio – os tesouros são geralmente voltados para uma área de especialidade. Essa área pode ser uma disciplina (Física) ou uma atividade (Engenharia Civil). É necessário definir qual é o âmbito a ser coberto pelo tesouro, em vista da literatura a ser indexada e de acordo com as necessidades do usuário final. Para definir o domínio pode-se recorrer a instrumentos terminológicos como: glossários, vocabulários e dicionários técnicos já existentes ou consultar classificações sistemáticas. É importante, também, delimitar o número aproximado de termos preferenciais a serem incluídos no tesouro.

2) Coleta e seleção de termos – o trabalho de compilação visa a reunir uma lista provisória de termos a serem trabalhados no tesouro. Um sistema de noções e conceitos deve ser utilizado como ponto de partida para a coleta, mesmo que tal sistema seja reformulado, posteriormente, para a adequação do instrumento às necessidades informacionais. Na escolha dos termos candidatos ao tesouro, procuramos ter como base o quadro classificatório básico adotado, de maneira que as noções que se assemelham e são mais ligadas entre si sejam aproximadas. Assim, é necessário distinguir entre: *noções próprias do domínio*; *noções que ultrapassam o domínio*; *noções tomadas de empréstimo de outros domínios*; *noções de linguagem comum*. A coleta de termos deve ser feita na literatura da área, de modo a capturar o significado que lhes é atribuído pelos especialistas da área (garantia literária). O termo deve ser, na medida do possível, extraído com o contexto, para que o seu significado não se perca. Desse modo, extrai-se o termo na frase que o apresenta, ou seja, na frase que expressa de modo claro o seu significado. Devemos, também, selecionar como descritores os termos que expressem o conceito de modo claro. Por exemplo: *produção industrial* ao invés de *produção*; *campo de pouso* ao invés de *campo*. O processo de seleção dos termos responde a diferentes graus de exaustividade, que variam conforme o campo de

aplicação do tesouro e de acordo com o *corpus* que é utilizado. A literatura propõe três grandes fontes de coleta: a tabela de classificação, a literatura especializada e o conhecimento da equipe e dos usuários. Essas três formas engendrarão os modos dedutivos e indutivos de construção.

3) Definição dos termos – A definição dos termos básicos é fundamental para uma estruturação adequada do vocabulário. Ela deve sofrer alteração contínua durante o trabalho de construção, porque, à medida que o trabalho avança, a delimitação da área torna-se cada vez mais explícita e consistente. As definições não são simples definições de dicionário, pois devem ser elaboradas com base nos sistemas de conceitos e noções das áreas a que se referem. É necessário definir os termos de forma a privilegiar o enfoque predeterminado. Assim esse procedimento assegura a delimitação das relações a serem privilegiadas no instrumento.

4) Categorização – A partir da lista provisória de termos e a definição dos termos principais, deve-se reagrupá-los em categorias. No *método dedutivo*, as categorias são identificadas e subdivididas de cima para baixo. A delimitação das categorias principais pode basear-se nos mapas conceituais dos domínios, devendo-se contemplar, também, as facetas privilegiadas de acordo com as necessidades informacionais e do usuário. No *método indutivo*, o procedimento é o oposto. Os termos que ocorrem na área são coletados de fontes diferentes, e só é criada uma categoria de termos se ela for considerada importante e útil. Nesse segundo tipo de abordagem, abordagem empírica, recorre-se à garantia literária e à garantia de uso. Assim destacamos que os métodos dedutivo e indutivo apresentam suas vantagens e desvantagens. É aconselhável, portanto, combinar os dois procedimentos.

Fujita (1998, p. 33) alerta que, “para a estruturação de conceitos é necessária a definição prévia de categorias. Com as categorias definidas, os conceitos são classificados e seus relacionamentos determinados”. Para a definição de categorias no processo de categorização, a autora argumenta que podem ser utilizados “três modelos: o currículo, a estrutura departamental e o conhecimento do especialista”. Em pesquisa realizada na área de Odontologia, os resultados obtidos demonstraram que os modelos são mais bem utilizados em conjunto (FUJITA, 1992). Tálamo (1997c) complementa que a Terminologia fornece referencial concreto para o entendimento dos conceitos e noções de domínios de

especialidade. A categorização dela proveniente deve ser acoplada às facetas mais adequadas para dar conta da literatura existente e das necessidades do usuário.

5) Organização das relações – Definidas as categorias, procede-se à organização das relações entre os termos. Nos tesouros são reconhecidas “três classes de relações básicas entre os termos: a relação hierárquica; a relação associativa; a relação de equivalência” (AUSTIN; DALE, 1993, p. 41). Conforme Cintra et al. (2002, p. 53), as relações hierárquicas revelam os níveis de superordenação e subordinação entre as noções, e elas podem ser genéricas, *específicas* ou *partitivas* e constituem a estrutura principal de um tesouro no sentido vertical (esquema de árvore). As árvores semânticas são, em geral, elaboradas segundo as cadeias conceptuais que partem dos termos mais gerais, ou amplos, em direção aos termos mais específicos ou restritos. Na *relação genérica*, a superordenação avança das diferenças para as semelhanças, ou seja, da espécie para o gênero, enquanto que, na *relação específica*, a subordinação, no sentido inverso, avança das semelhanças para as diferenças, ou seja, do gênero para a espécie.

Já a *relação partitiva*, segundo Cintra et al. (2002, p. 53), que é um tipo de relação hierárquica, a noção de superordenação refere-se a um objeto considerado como um todo e as noções subordinadas a objetos considerados como suas partes. Por exemplo: em relação a *navio*, a noção de *casco* é uma noção específica partitiva e denota que *navio* é uma noção que se refere ao todo e que *casco* é uma noção que se refere à parte. Já as *relações associativas* são estabelecidas pela dependência entre conceitos, sem, no entanto, estabelecer uma superordenação. As relações associativas podem ser divididas em dois grupos: *relação sequencial*, que é a relação de dependência estabelecida entre conceitos que se referem a objetos que têm contigüidade espaço-temporal como, por exemplo: “causa-efeito”, “produtor-produto”, “etapas de um processo”; e *relação pragmática* que é a relação entre conceitos que podem ser estabelecidas por conexões temáticas.

Para Cintra et al. (2002, p. 54), “as relações que não se submetem a uma hierarquia são aquelas que apresentam entre si contigüidade espacial ou temporal entre termos que não possuem relacionamento hierárquico e nem de equivalência”. Por este motivo, as *relações de equivalência* são utilizadas em uma linguagem documentária para o controle do vocabulário. Estas relações operam no

nível de sinonímia e da polissemia; estabelecem as remissivas com a finalidade de encaminhar o usuário para os termos preferidos pelo sistema. Na linguagem natural, vários termos, palavras ou expressões podem designar uma única idéia ou conceito. Já na Linguagem Documentária apenas uma denominação é selecionada para figurar como descritor; os demais termos são considerados como não- descritores, fazendo-se destes, remissivas para aquele (GOMES, 1990, p. 47).

Para as relações associativas, devemos observar a natureza daquelas que são frequentes no domínio, visando-se apresentá-las de forma consistente em todo o tesouro.

6) Normalização – A normalização preocupa-se com a padronização das unidades do tesouro. Essa etapa tem a função de garantir o controle de vocabulário e a circulação de informações. Observamos que existem vários tipos de normalização, entre os quais destacam-se: o da normalização gramatical e o da normalização semântica dos termos. A normalização gramatical tem como objetivo uniformizar a apresentação das unidades informacionais do tesouro quanto aos aspectos relativos a gênero, número, grau; já a normalização semântica dos termos é garantida pela definição e pelas terminologias dos domínios e subdomínios de especialidade.

7) Notas de escopo ou de aplicação – Apresentam instruções de uso de alguns descritores, explicitando como devem ser aplicados. Por meio dessas notas, pode-se ampliar ou restringir o significado de um descritor, conforme a política de indexação utilizada. Tálamo (1997c, p. 12) afirma que as notas de escopo permitem um controle efetivo de todas as operações documentárias relativas ao tesouro. Podem incluir data de inclusão/exclusão de descritores, documentando todo o processo.

8) Apresentação do tesouro – Utiliza-se, usualmente, a apresentação sistemática acompanhada de apresentação alfabética. Na parte sistemática os termos são relacionados conforme as categorias ou classes e apresentados sob a forma de listas das classes em ordem alfabética. Na ordem alfabética, cada descritor aparece com aqueles com os quais se encontra relacionado, na seguinte ordem: equivalências, relações genéricas (descritor genérico e específico) e relações associativas. Conforme enfatizou Tálamo (1997c, p. 8-13) no referido curso, faz-se necessário o estabelecer os relacionamentos entre os termos, qualquer que seja a forma de apresentação do tesouro.

Em seguida, indicamos uma síntese das etapas para construção de tesouros, propostas por Tálamo (1997c), complementadas por Fujita (1998), como segue:

- 1) definição do domínio;
- 2) coleta e seleção dos termos;
- 3) definição dos termos;
- 4) categorização;
- 5) organização das relações;
- 6) normalização;
- 7) notas de escopo ou de aplicação; e
- 8) apresentação do tesouro.

De outro modo, relatamos uma experiência didática desenvolvida por Lara e Tálamo (2007) em que evidenciam um diálogo teórico e metodológico entre a Terminologia e a Linguística Documentária. Do ponto de reflexão da Linguística Documentária, procura-se oferecer ao aluno de graduação um conjunto de referências teóricas e de práticas experimentais que o levem a conhecer as possíveis contribuições da Terminologia ao fazer documentário.

Para as pesquisadoras, importa à formação a teorização, a interpretação e a proposição de metodologias de organização da informação orientadas para os diferentes ambientes, desenvolvendo o escopo do termo linguagem documentária para caracterizar a linguagem utilizada com a finalidade de organizar, permitir a navegação, a recuperação e a busca da informação em ambientes web, dentre outros (LARA; TÁLAMO, 2007).

A metodologia utilizada procura associar procedimentos presentes nos trabalhos terminológicos, terminográficos e documentários, vinculando as operações terminológico-documentárias à idéia de que a organização do conhecimento tem como objetivo precípua o estabelecimento de mediações para acesso, circulação, distribuição e reutilização da informação conforme enunciado pelo ciclo social da informação, um dos fundamentos da própria Ciência da Informação.

Desse modo, as pesquisadoras Lara e Tálamo (2007) orientam para a realização das seguintes ações:

- leitura e análise de pequenos corpora (que reúnem bibliografia de características crescentemente especializada, de textos de divulgação científica, para científicos propriamente ditos, geralmente sobre uma temática atual);
- identificação de termos (algumas vezes, frases), considerados mais importantes para a compreensão do texto;
- discussão coletiva;
- seleção dos termos dos vocabulários (conceitual e do vocabulário funcional, sendo que sobre os primeiros incidirão os procedimentos de descrição);
- descrição terminológica (vocabulário conceitual);
- registro de termos em fichas terminológicas de coleta (simplificadas, tendo em vista os objetivos documentários) com o termo candidato, suas definições, se existentes, contextos de uso (transcrições literais do termo no enunciado em que aparecem), fonte, domínio responsável pelo preenchimento e data;
- sistematizar os conceitos explorados na realização de monografias sob temas diversos (LARA; TÁLAMO, 2007).

Lara (2004, p. 239) argumenta que tanto a Terminologia teórica como a concreta contribuem para “a decupagem dos domínios e a organização da rede lógico-semântica entre os termos”. Recentemente, essas etapas foram ajustadas e melhor sistematizadas conforme roteiro apresentado por Lara (2009, p. 162), a saber:

- seleção de um *corpus* básico para a pesquisa e para apoio (dicionários e glossários já existentes);
- seleção de termos do vocabulário conceitual e do vocabulário funcional;
- registro da síntese a partir da observação das características, a formulação de uma definição;
- categorização e formulação [da] árvore conceitual de domínio aproximativa;
- reorganização da árvore com base nas informações das fichas terminológicas de coleta e de síntese;
- estruturação dos termos sob forma de tesouro a partir de uma hipótese de organização;

- construção do glossário para o tesouro terminológico (Lara; Tálamo, 2007).

No relato de experiência apresentado, observamos, nitidamente, as vantagens do procedimento metodológico para uso didático com alunos de graduação, as múltiplas possibilidades de aliar conhecimentos teóricos e práticos da área de Organização da Informação aos da área de Terminologia (em especial da Terminologia concreta) na busca de soluções para problemas vivenciados por profissionais em suas práticas cotidianas. Concordamos com as autoras quanto à contribuição efetiva para formulação de metodologias mistas para formulação de produtos e instrumentos híbridos.

Do mesmo modo, Rondeau (1984) asseverou a importância do uso de procedimentos terminológicos e terminográficos para sistematização de conceitos representados por termos e revelou o foco da abordagem mista, com a possibilidade de combinar métodos, conforme destacado anteriormente. Por tratar-se de um trabalho terminológico que exige rigor metodológico para a sua aplicação, apoiamos um conjunto de etapas que devem ser observadas para a realização da pesquisa terminológica temática.

Em razão disso, o desenvolvimento de proposta de um modelo metodológico, fundamentado em procedimentos terminográficos, para a construção de tesouros, no que tange ao processo de seleção e identificação de termos de áreas de especialidade, beneficia-se da interface com a Terminografia para aprimoramento dessas etapas, além de explicitar o modo de recorte dos termos em seu contexto de uso.

Diante do exposto, apresentamos, no capítulo seguinte, a proposta de um modelo metodológico para a construção de tesouros com a integração de procedimentos terminográficos para a apropriação e o aprimoramento de etapas da construção de tesouros tendo em vista a compatibilidade entre a linguagem documentária utilizada pelo sistema de informação de áreas especializadas e a linguagem de busca do usuário.

## **4 A INTEGRAÇÃO DA TERMINOGRAFIA NA CONSTRUÇÃO DE TESAUROS**

Neste capítulo, apresentamos proposta de um modelo metodológico para a construção de tesouros com a integração de procedimentos terminográficos. Para a formulação da proposta de um modelo metodológico para a construção de tesouros. Realizamos um percurso teórico e metodológico em busca de fundamentação para a nossa pesquisa cujo referencial destacamos nos capítulos anteriores. O modelo metodológico proposto será constituído por um conjunto de etapas de construção de tesouros com a integração de procedimentos terminográficos com base na sistematização dos resultados obtidos dos capítulos 2 e 3, observando pressupostos delineados para esse estudo no âmbito da Ciência da Informação/Documentação com a Terminologia/Terminografia.

### 4.1 Para um modelo metodológico de construção de tesouros

Em busca de um modelo metodológico para construção de tesouros fundamentados em nossas principais premissas de pesquisa, evidenciadas anteriormente, reiteradas nesse momento, a saber: a) o aprimoramento de etapas de construção de tesouros contribui para melhorar a representação temática da informação; e b) a Terminografia possibilita a harmonização dos conceitos para o acesso à informação desejada em áreas especializadas.

No esforço para alcançar ao que enunciamos para esta tese, estudamos no capítulo 2, intitulado representação de conceitos, a começar de aportes teóricos e metodológicos da área de Ciência da Informação/Documentação/Organização e Representação do Conhecimento registrado, com foco na construção de tesouros, particularmente, no que diz respeito às etapas da sua construção, os Métodos de construção de tesouros.

Nessa direção, entendemos que a principal divisão é estabelecida da seguinte forma: quando o método é dedutivo, a preocupação não reside apenas na coleta dos termos, mas também no estabelecimento de sua estrutura conceptual



ou categorização. Mapeiam-se o domínio e o subdomínio em estudo, antes do início da coleta de termos; de outro modo: quando o método é indutivo, não existe uma preocupação inicial com a estrutura conceptual ou categorização, porque se organizam o domínio e o subdomínio em estudo, após a coleta dos termos (FUJITA, 1998).

Entretanto, para admitir maior flexibilidade quando se trata da construção de um tesouro, em geral, combinam-se os métodos em determinados momentos, o método dedutivo ou o indutivo, das etapas de construção de tesouros. Isso contribui para colocar em evidência outra abordagem, a de combinação de métodos, em conformidade com Aitchison; Gilchrist (1979), Lancaster (1987), Batty (1989), Gomes (1990; 2004), Fujita (1992; 1998), Guinchat e Menou (1994), e outros. O estudo dos referenciais fundamentados em normas e autores, apresentados nesse capítulo permitiram-nos analisar diversas metodologias que orientam e dão sustentação na construção de tesouros.

Vale esclarecer que, em concordância com o método explicitado anteriormente, esta pesquisa considerou os princípios estabelecidos no primeiro método, o método dedutivo, para o desenvolvimento da proposta de um modelo metodológico para o processo de construção de tesouro. A principal justificativa para a adoção do referido método resulta do entendimento de que ele atenderia melhor aos propósitos de um sistema de informação especializado em razão de os termos preferenciais ou de indexação serem provenientes dos conceitos pertencentes à linguagem compatível com as áreas de especialidade do usuário.

Nessa direção, Fujita (1998, p. 33) alerta que, “para a estruturação de conceitos é necessária a definição prévia de categorias. Com as categorias definidas, os conceitos são classificados e seus relacionamentos determinados”. Para a definição de categorias no processo de categorização, concordamos com a autora supracitada quando argumenta que podem ser utilizados “três modelos: o currículo, a estrutura departamental e o conhecimento do especialista”. Em pesquisa realizada na área de Odontologia, os resultados obtidos demonstraram que os modelos são mais bem utilizados em conjunto (FUJITA, 1992).

Tálamo (1997c) complementa que a Terminologia fornece referencial concreto para o entendimento dos conceitos e noções de domínios de especialidade. A categorização dela proveniente deve ser acoplada às facetas mais adequadas para dar conta da literatura existente e das necessidades do usuário.

## 4.2 Etapas da pesquisa terminológica para construção de tesouros

Apresentamos uma síntese das etapas da construção de tesouros com base no estudo das Diretrizes IBICT (1984), Diretrizes UNESCO (1993) e Diretrizes ANSI/NISO (2005), além da análise da sistematização das etapas de construção de tesouros realizada pelos autores estudados. Para apresentação desse estudo, foram agrupadas em 7 categorias temáticas, a saber:

1. Trabalho preliminar (Orientações gerais/Fase de planejamento);
2. Métodos de compilação (Formas/métodos de compilação de termos);
3. Registro de termos (Compilação de termos: coleta; validação);
4. Verificação de termos (Admissão e exclusão de termos/Estabelecimento de relações entre termos/categorização);
5. Especificidade (Nível de especificidade);
6. Uso de equipamento automático de processamento de dados (Uso de equipamento informático para processamento de dados/utilização do computador; e
7. Forma e conteúdo de um tesouro/Formas de apresentação.

Contudo, percebemos a necessidade de complementação e o aprimoramento de metodologias existentes, no que tange à identificação e seleção de conceitos/termos para a construção de tesouros. Observamos que a combinação de métodos, mencionada anteriormente, poderia levar-nos à combinação de procedimentos, como a integração dos procedimentos terminográficos para o aprimoramento de etapas específicas da construção de tesouros.

Em busca dessa complementação, tendo em vista a formulação da proposta de modelo metodológico para o aprimoramento de etapas de construção de tesouro mais compatível com terminologia de área de especialidade, no capítulo 3, recorreremos às orientações teórico-metodológicas da Terminologia, particularmente à Terminografia considerada, nesse estudo, como um ramo da Terminologia, com foco nos procedimentos terminográficos que direcionam para a identificação e seleção de conceitos/termos candidatos com base no contexto de uso, em áreas específicas.

Nas referências analisadas sobre Terminologia, para a construção de instrumentos terminológicos (glossários) identificamos procedimentos terminográficos em manuais de Terminologia de estudiosos como Robert Dubuc (1999), Alicia Fedor de Diego (1995), Guy Rondeau (1984) e outros. Com base nos estudos optamos por um conjunto de etapas recomendadas por Rondeau (1984).

Cabe esclarecer que o conjunto de etapas da pesquisa terminológica temática refere-se ao resultado de investigação realizada por Rondeau (1984, p. 70-77) com pesquisadores de diferentes países sobre métodos e procedimentos de pesquisa terminológica, conforme descrevemos a seguir:

- 1) escolha do domínio e da língua do trabalho;
- 2) delimitação do subdomínio;
- 3) consulta a especialista do domínio/subdomínio;
- 4) coleta do *corpus* do trabalho terminológico;
- 5) estabelecimento da árvore de domínio;
- 6) expansão da representação do domínio escolhido;
- 7) estabelecimento dos limites da pesquisa terminológica temática;
- 8) coleta e classificação de termos;
- 9) verificação, classificação e confirmação de termos;
- 10) trabalho de apresentação de dados terminológicos.

Em consonância com a literatura apresentada, evidenciamos, a seguir, uma síntese das etapas para construção de tesouros, propostas por Tálamo (1997c), complementadas por Fujita (1998), como segue:

- 1) definição do domínio; (etapa 1 da sistematização);
- 2) coleta e seleção dos termos; (etapa 3 da sistematização);
- 3) definição dos termos; (etapa 4 da sistematização);
- 4) categorização; (etapa 4 da sistematização);
- 5) organização das relações; (etapa 4 da sistematização);
- 6) normalização; etapa 7 da sistematização;
- 7) notas de escopo ou de aplicação; (etapa 4 da sistematização);e
- 8) apresentação do tesouro; (etapa 7 da sistematização).

Por meio desses estudos, constatamos que a construção de tesouro é uma tarefa, essencialmente, intelectual que se realiza por meio de diversas etapas

que podem sofrer alguns ajustes e adaptações conforme as necessidades que se apresentam. É fato que as etapas para a construção de tesouros vão desde a definição da área de especialidade; passam por etapas que dizem respeito a método de compilação; registro; verificação de termos; até a etapa que trata da forma de apresentação do tesouro.

Certo de que o escopo deste estudo é o de propor um modelo metodológico no esforço de aprimoramento da identificação e seleção de conceitos, para a construção de tesouros, apresentamos uma sistematização de etapas identificadas para a construção de tesouros, fundamentada nas premissas principais da investigação sobre Documentação/Construção de tesouros e Terminologia/Terminografia orientadas por referenciais de normalização, literatura e tesouros, com ênfase nas abordagens indutiva e dedutiva e na combinação de métodos, como a própria Terminografia e seus procedimentos. Desse modo, consideramos as orientações onomasiológicas e semasiológicas no intento de corresponder às necessidades de terminologia de áreas de especialidade, conforme expomos no quadro 10, a seguir, para o respaldo da proposta de um modelo metodológico integrado para construção de tesouro.

<b>MODELO METODOLÓGICO INTEGRADO PARA CONSTRUÇÃO DE TESAURO</b>	
<b>Sistematização de etapas da construção de tesouros (normalização, literatura e tesouros) - Procedimentos terminográficos</b>	
<b>1. Trabalho preliminar</b> (Orientações gerais/Usos de equipamento automático de processamento de dados)	- escolha do domínio e da língua do tesouro; - delimitação do subdomínio; - estabelecimento dos limites da pesquisa terminológica temática; - consulta a especialista do domínio/subdomínio.
<b>2. Método de compilação</b> (Abordagem de compilação)	- coleta do <i>corpus</i> do trabalho terminológico; - estabelecimento da árvore de domínio; - expansão da representação do domínio escolhido.
<b>3. Registro de termos</b>	- coleta e classificação de termos.
<b>4. Verificação de termos</b> (Admissão e exclusão de termos /Especificidade)	- verificação, classificação e confirmação de termos; - elaboração de definições; - <b>uso do vocabulário de especialidade para o estabelecimento de relações entre os descritores e de relações entre descritores e não descritores.</b> - organização das relações entre descritores
<b>5. Forma de apresentação de um tesouro</b>	- trabalhos de apresentação do tesouro.

**Quadro 10** – Sistematização de etapas da construção de tesouros.

Cabe esclarecer que o quadro 10, composto da sistematização de etapas da construção de tesouros e de procedimentos terminográficos pertinentes a cada etapa da construção de tesouros, constitui-se na Proposta de integração da Terminografia às etapas de construção de tesouros. Tendo em vista a análise de literatura da área sobre construção de tesouros e os resultados obtidos com a sistematização dos referenciais fundamentados em normas, autores e tesouros, apresentamos nossa proposta de modelo metodológico com base nas premissas e na sistematização de etapas de construção de tesouros e de aspectos contribuintes da Terminografia, intitulada:

### **PROPOSTA DE INTEGRAÇÃO DA TERMINOGRAFIA ÀS ETAPAS DE CONSTRUÇÃO DE TESAUROS: um modelo metodológico**

Em sistemas de informação de áreas especializadas, o tesouro tem a função de apoiar, ao mesmo tempo, os processos de representação e de recuperação temática da informação. Os tesouros, em geral, são voltados para o vocabulário de especialidade. Conhecer os processos de identificação, caracterização e análise de um conjunto de termos de áreas de especialidade e suas aplicações tem grande importância no contexto das linguagens documentárias tendo em vista a organização e recuperação da informação.

A seguir, detalhamos o modelo metodológico proposto para a construção de tesouros para áreas de especialidade, com a apresentação de cada etapa acompanhada de sua descrição, com a integração dos elementos contribuintes da Terminografia para o aprimoramento de etapas da construção de tesouros tendo em vista à necessidade de adequação aos vocabulários presentes no sistema de informação de áreas especializadas.

Vale dizer que o modelo metodológico proposto para a construção de tesouros compreende uma sequência lógica de 5 etapas. Inicia-se pela etapa 1 Trabalho preliminar das etapas de construção de tesouros, que estabelece as orientações gerais para o desenvolvimento do trabalho de construção do tesouro; etapa 2 Método de compilação; etapa 3 Registro de termos; etapa 4 Verificação de termos; finaliza-se a sua construção com a etapa 5 Forma de apresentação de um

tesauro. Cada etapa responde a determinados procedimentos e integra tanto os procedimentos documentários como os terminográficos e outros, a saber:

## **I. TRABALHO PRELIMINAR DAS ETAPAS DE CONSTRUÇÃO DE TESAUROS**

Esta etapa é composta por:

### **- Orientações gerais**

Este item trata do estabelecimento das orientações gerais para o desenvolvimento das etapas de construção de tesauro, como: comunicar a intenção de construir um tesauro novo em periódicos apropriados; evitar duplicação de tesauros, e outras.

### **- Escolha da área de especialidade e da língua do tesauro**

Este tópico diz respeito à escolha da área de especialidade e a língua para a construção do tesauro que, em geral, define-se de acordo com as necessidades dos usuários, além da abrangência, nível de especificidade e a língua do trabalho. Define-se o público-alvo e a língua do trabalho levando-se em conta o envolvimento dos pesquisadores com a área de especialidade.

### **- Delimitação de subárea de especialidade do tesauro**

Este item refere-se à delimitação da subárea de especialidade para a construção de tesauro. Recomenda-se considerar o tipo de usuário, suas necessidades para delimitar a subárea de especialidade do tesauro, além do nível de especificidade.

### **- Estabelecimento dos limites da pesquisa terminológica para áreas de especialidades**

O limite da extensão da pesquisa terminológica quanto ao número aproximado de termos estabelece-se em função dos objetivos propostos, das disponibilidades de tempo e de meios financeiros. Desse modo, pode-se escolher um levantamento básico compilando uma média de 300 termos, ou exaustivo, por volta de 2.500 termos.

Nessa direção, recomenda-se delimitar um número aproximado de descritores a serem incluídos no tesauro, em razão da rede nocional que lhe é

própria, e das redes nocionais conexas. Para tanto, indica-se o uso do vocabulário de especialidade para o estabelecimento de relações entre descritores e de relações de descritores e não descritores.

- Consulta a especialista da área/subárea de especialidade

Este aspecto trata da consulta à especialistas e profissionais da área de especialidade para obtenção de contribuições quanto à escolha de *corpus* representativo para a coleta de termos e para conseguir o envolvimento do usuário neste trabalho.

## II. MÉTODO DE COMPILAÇÃO

Esta etapa é composta por:

- Abordagem de compilação

Este tópico diz respeito à forma de abordagem de compilação adotada para a construção do tesouro, com ênfase nas abordagens indutiva e dedutiva, combinação de métodos. Desse modo, consideram-se as orientações onomasiológicas e semasiológicas no intento de corresponder às necessidades de vocabulário de especialidade e para admitir maior flexibilidade quando se trata da construção de um tesouro, em conformidade com Aitchison; Gilchrist (1979); Lancaster (1987); Batty (1989); Gomes (1990; 2004); Fujita (1992; 1998); Tálamo (1997c); Guinchat e Menou (1994) e outros.

Em concordância com o método explicitado anteriormente, “combinação de métodos,” é possível empregar mais de uma dessas abordagens em um estágio ou outro durante a construção de um vocabulário controlado. Por exemplo, hierarquias e outras relações entre termos que foram primeiro estabelecidos indutivamente poderiam mais tarde ser reexaminadas a partir de um ponto de vista dedutivo. Ambas as técnicas (dedução/indução) são essencialmente empíricas.

Deve-se aceitar, desde o início da construção de tesouros, que algumas decisões com relação aos termos e suas inter-relações, que foram feitas durante os primeiros estágios de compilação, possam ser revisadas como uma experiência posterior. Torna-se necessário verificar os termos e hierarquias com

frequência para assegurar uma aplicação consistente de princípios em tais procedimentos como: o estabelecimento de inter-relações entre termos e a divisão de termos compostos.

- Coleta do *corpus* para compilação de termos do tesauro

Este item trata da coleta do *corpus* representativo com a finalidade de reunir a documentação necessária para a compilação de termos do tesauro: a documentação terminológica de cada área de especialidade – dicionários, glossários, vocabulários, tesouros, e outros; a documentação normativa – normas terminológicas, normas e diretrizes documentárias; e a documentação de referência ou autoral – textos da área de especialidade, indicados por especialistas da área, fundamentais para a compilação de termos do tesauro. Nessa direção, a literatura propõe três grandes fontes de coleta: a tabela de classificação, a literatura especializada e o conhecimento da equipe e dos usuários. Essas três formas engendrarão os modos dedutivos e indutivos de construção de tesouros (TÁLAMO, 1997c).

- Estabelecimento da estrutura conceitual/Categorização

Este tópico refere-se à categorização ou estabelecimento da estrutura conceitual da área/subárea de especialidade. Representa o conjunto nocional que tem a função de situar a área ou subárea de especialidade a ser estudada. Antes de estabelecer a estrutura conceitual, o pesquisador deverá consultar documentos como sistemas de classificação, glossários, tesouros, e outros. Em algumas áreas/subáreas de especialidade esses instrumentos são até abundantes, mas em outras áreas de especialidade podem encontrar-se desatualizados, ou não existir. Para a definição de categorias no processo de categorização, Fujita (1998, p. 33) argumenta que podem ser utilizados “três modelos: o currículo, a estrutura departamental e o conhecimento do especialista”. Em pesquisa realizada na área de Odontologia, os resultados obtidos demonstraram que os modelos são mais bem utilizados em conjunto (FUJITA, 1992). Tálamo (1997c) complementa que a Terminologia fornece referencial concreto para o entendimento dos conceitos e noções de domínios de especialidade. A categorização dela proveniente deve ser acoplada às facetas mais adequadas para dar conta da literatura existente e das necessidades do usuário.



- Expansão da representação da área de especialidade escolhida

Este aspecto diz respeito à expansão da representação da área de especialidade escolhida, normalmente, decorre dos itens 2, delimitação de subárea da especialidade, e 5, estabelecimento da estrutura conceitual/categorização. Nesse momento, é necessário o auxílio dos especialistas da área para direcionar os trabalhos próprios desta etapa e, também, para verificar os resultados obtidos no item precedente, categorização.

### III. REGISTRO DE TERMOS

Esta etapa é constituída por:

- Coleta e classificação de termos para construção de tesouro

Este tópico refere-se à coleta de termos a partir do *corpus* do selecionado. Em geral, consiste em fazer uma leitura do texto assinalando os conceitos/termos significativos a extrair. Para auxiliar o pesquisador no reconhecimento dos termos (algumas vezes fraseologias significativas) sugerem-se algumas pistas, tais como: recursos gráficos e de *layout* – como aspas, negritos, itálicos, entre outros, ou a ocorrência de termos em posição de destaque no texto como títulos e subtítulos; frequência estatística – a recorrência frequente do termo na mesma configuração; e estratégias discursivas - o próprio texto esclarece, conceitua, explica, define o termo.

Cabe ao pesquisador verificar se o termo pertence à área de especialidade (AUBERT, 2001, p. 64). Do mesmo modo, no momento da identificação do termo acompanhado de seu contexto de uso, deve-se ter em mente os três tipos de contextos citados por Rondeau (1984, p. 80), a saber: o definitório, o explicativo e o associativo: o contexto definitório - fornece dados precisos sobre o conceito do termo candidato; já o contexto explicativo - revela a natureza, o objetivo ou um aspecto do conceito estudado; enquanto o contexto associativo - se caracteriza pela ausência de termos significativos no contexto. O contexto é o “enunciado no qual figura o termo estudado”, ou significativo, ou parte de um texto no qual ocorre o termo (RONDEAU, 1984, p. 80).

Para o registro dos termos, procede-se da seguinte forma: após a leitura e identificação dos conceitos/termos significativos assinalados no momento da leitura, realiza-se a anotação dos termos em ficha terminológica de registro de termos a qual é composta pelos seguintes dados: termo, contextos de uso, fonte, domínio/área de especialidade, definições, responsável pelo registro e a data, conforme quadro 11.

Termos	Contextos de uso	Fonte(s) <i>corpus</i>	Definição	Relações:	Domínio/ Área de Especial//
Indicar o termo  Variante(s) gráfica(s)	Descrever o contexto em que o termo candidato aparece. (SOBRENOME, ano)	Sigla da fonte, v., n., art., p., ano	Descrever a(s) definição(ões) presentes(s) no <i>corpus</i> .  Pesquisar em dicionários, glossários, vocabulários, e outros.	Indicar o equivalente em outra(s) língua(s):  em inglês:  em espanhol:  Termos Relacionados:  Termos Sinônimos:	
<b>Notas:</b> As normalizações do termo ou observações devem ser anotadas neste local.				<b>Resp. Registro:</b> Indicar nome	<b>Data coleta:</b> dd/mm/aaaa

**Quadro 11** – Ficha terminológica-guia de registro de termos.

Esclarecemos que no quadro 11, a ficha terminológica apresentada, refere-se a uma ficha terminológica-guia de registro de termos que orienta o modo correto de preenchimento de cada componente elencado para o registro de termos.

#### IV. VERIFICAÇÃO DE TERMOS

Esta etapa é constituída por:

- Verificação e classificação da noção/denominação

Este item trata da verificação e classificação da noção/denominação em que cada conceito/termo é reexaminado e analisado para adequações, se necessárias. A análise refere-se: a) ao seu conteúdo - através de comparações entre as definições e os contextos; e b) ao seu lugar na rede nocional da área de especialidade/domínio ou subárea/subdomínio.

Entende-se que as operações dessa etapa atingem os seguintes resultados: 1) delimitação mais precisa do termo, com as referências aos documentos; 2) classificação definitiva dos termos; 3) Rede de sinônimos. Nesta direção, uma vez que o sistema de remissivas apresenta ao leitor um maior esclarecimento do termo, formando uma rede conceitual que se reflete tanto no nível de macroestrutura, pois cada remissão leva a uma nova entrada, como no nível da microestrutura do tesauro, pois é aí que ela se instala com o objetivo de interligar unidades de informação, descritores, que fazem parte do mesmo paradigma definicional.

- Notas de escopo ou de aplicação

Este aspecto refere-se à notas de escopo ou de aplicação, apresentam instruções de uso de alguns termos preferidos, explicitando como devem ser aplicados. Por meio dessas notas, pode-se ampliar ou restringir o significado de uma unidade de informação, descritor, conforme a política de indexação utilizada. As notas de escopo permitem um controle efetivo de todas as operações documentárias relativas ao tesauro. Podem incluir data de inclusão/exclusão de descritores, documentando a realização de todas as etapas de construção de tesauro (TÁLAMO 1997c, p. 12).

## **V. FORMA DE APRESENTAÇÃO DE UM TESAURO**

Este tópico é composto por:

- Trabalhos de apresentação do tesauro

Esta etapa trata das formas de apresentação de um tesauro que, em geral, consiste de uma apresentação sistemática acompanhada de apresentação alfabética. Na parte sistemática os descritores são relacionados conforme as categorias ou classes e apresentados sob a forma de listas das classes dispostas em ordem alfabética. Na ordem alfabética, cada descritor ou não-descritor aparece com aqueles com os quais se encontra relacionado, na seguinte ordem: relação equivalências, hierárquicas (genérico e específico) e relações associativas. Deve-se estabelecer os relacionamentos entre os termos, qualquer que seja a forma de apresentação do tesauro (TÁLAMO, 1997c).

Em razão disso, apresenta-se o uso da Terminografia como um aprimoramento às etapas de construção de tesouro que compõem normas e diretrizes consolidadas de construção de tesouros com o uso do vocabulário de especialidade para o estabelecimento de relações entre os descritores e de descritores e não descritores.

Desse modo, foram apresentadas as etapas de construção de tesouros com a integração de procedimentos terminográficos como proposta de um modelo metodológico para a construção de tesouros. Com base nas premissas e na sistematização de etapas de construção de tesouros reforçamos nossos referenciais com contribuições de Naumis Peña (2007) e Boccato (2009) que apresentam elementos contribuintes para a construção de tesouros.

Naumis Peña (2007) enfatiza que tais recomendações podem ser úteis para criar, expandir, revisar, reduzir ou aperfeiçoar um tesouro existente. A autora supracitada declara que, no tesouro, a seção mais importante é a que contém os termos em ordem alfabética com suas relações.

Para a apropriação e o aprimoramento ao uso da linguagem documentária, para a representação e recuperação da informação em sistemas de informação de bibliotecas universitárias de áreas especializadas, faz-se necessário promover algumas ações indicadoras que são apontados por Boccato (2009) tendo em vista a compatibilidade entre a linguagem documentária utilizada pelo sistema de informação e a linguagem de busca do usuário. A autora supracitada apresenta uma relação de 13 indicadores, conforme segue:

**a)** construção do vocabulário a partir das linguagens de especialidades das áreas científicas e da linguagem de busca do usuário, com vistas à compatibilidade entre a linguagem adotada pelo sistema e a de busca do usuário; **b)** incorporação de novos termos visando à atualização da linguagem que se fizer necessária, por meio de coleta em fontes de informação formais (dicionários especializados, glossários técnico-científicos, diretórios, entre outros) e informais (formulários de sugestões de assuntos preenchidos pelos usuários, catálogo e listas de assuntos locais elaborados pela biblioteca); **c)** eleição de termos expressivos visando à clareza na designação do assunto; **d)** revisão da tradução de termos existentes na linguagem, tendo em vista a devida correspondência conceitual que deve ocorrer em relação à terminologia das áreas científicas nacionais; **e)** eliminação das ambiguidades causadas pela homonímia e polissemia advindas da linguagem natural: adoção de termos qualificadores agregados ao termo preferido de modo a definir diferentes aspectos, conceitos e pontos de vista abordados

pelo autor sobre o assunto do documento. Os termos qualificadores possibilitam a especificidade na representação e na recuperação da informação. O uso de singular e plural também é um recurso auxiliar para a distinção entre termos homógrafos; **f)** incorporação de termos específicos, tendo em vista a especificidade exigida do tratamento de conteúdos documentários para a recuperação precisa da informação; **g)** controle de sinônimos: controle efetivo dos termos sinônimos, quase-sinônimos e das variantes lexicais (ortografia, singular - plural, nome completo divergente da abreviatura) evitando-se a dispersão temática e proporcionando maior exatidão na indexação e busca por assunto; **h)** estabelecimento das relações lógico-semânticas: a construção das relações hierárquicas, de equivalência e não-hierárquicas visa a indicação dos relacionamentos semântico-conceituais entre os termos, propiciando a consistência na representação e na recuperação. O incremento das relações não-hierárquicas (associativas) promove uma aproximação maior com a linguagem de busca do usuário, relevando a função comunicativa que a linguagem deve ter; **i)** verificação da sintaxe dos cabeçalhos de assunto compostos referentes à ordem das ideias que os compõem; **j)** incorporação de notas de escopos nos cabeçalhos de assunto; **k)** intensificação da função comunicativa da linguagem; **l)** representação de conceitos por meio de termos, visando à obtenção de cabeçalhos mais consistentes (BOCCATO, 2009, p. 227-228).

Diante do exposto, concordando com indicadores apontados, torna-se possível afirmar que esses aspectos, dentre outros, mostram-nos a importância do vocabulário de especialidade na sociedade atual, para ampliação do saber e do saber-fazer do profissional da Documentação. Nessa direção, decorre a relevância de aprimoramento de modelos metodológicos de tratamento, compilação, recuperação e comunicação de metalinguagens.

O discurso de uma ciência, representação e síntese das suas descobertas e do saber construído, se preciso e bem elaborado, conduz ao aperfeiçoamento da prática profissional de modo abrangente; essa prática realimenta tal discurso com novos fatos e atualização do vocabulário de especialidade, reafirmando o processo de alimentação e realimentação do discurso científico e técnico (BARBOSA, 2004). Nessa direção, a título de ilustração, temos como um exemplo do trabalho aplicado com o modelo metodológico proposto: "Tesauro Terminográfico Preliminar em Gestão da Informação". Disponível em: <<http://www.uel.br/portal/uel>>

**Tesouro Terminográfico Preliminar em Gestão da Informação**

Buscar

Lista sistemática  
Lista alfabética  
Sobre...  
Minha conta

português

**Gestão da informação (pt) –  
Information management (en) –  
Gestion de l'information (fr)**

Data de criação: 07-Ago-2009  
Termo aceptado: 07-Ago-2009

► INÍCIO ► Gestão (pt) – Management (en) – Gestion (fr) Gestão da informação (pt) –  
Information management (en) – Gestion de l'information (fr)

**Nota de escopo**

Gerenciamento de dados e informações que já estão consolidados em algum tipo de veículo de comunicação, desde o livro impresso até a rede internet. (VALENTIM, 2002) (DGZ, v.3,n.4, art.02,p.4, 2002)

TG Gestão (pt) – Management (en) – Gestion (fr)

**Gestão da informação (pt) – Information management (en) – Gestion de l'information (fr)**

TR Gerência de recursos informacionais (pt) – Informational resources management (en) –  
Gérence de recours informationnel (fr)  
TR Gerente de recursos informacionais (pt) – Informational resources manager (en) –  
Gérent de recours informationnel(fr)

**Figura 5** - Tesouro Terminográfico Preliminar em Gestão da Informação.

No capítulo seguinte, apresentamos as Considerações finais a respeito de nossa pesquisa, com ênfase para a importância da construção de tesouro com a integração de procedimentos terminográficos, no que tange a seleção e identificação de conceitos, para o estabelecimento de relações entre descritores no esforço de construir um vocabulário consistente e compatível com a demanda em áreas de especialidade.

## 5 CONSIDERAÇÕES FINAIS

A principal correlação buscada neste estudo foi a integração da Terminografia na construção de tesouros fundamentada na abordagem temática de áreas de especialidade. Nosso interesse em estudos de abordagem terminológica de áreas de especialidade esteve centrado na idéia de que um sistema de informação deve estar amparado por instrumentos capazes de compatibilizar linguagem adotada pelo sistema de informação, a linguagem de busca.

Frisamos a necessidade de harmonizar vocábulos entre integrantes de um sistema de informação de áreas especializadas, tendo como procedimento equilibrar os vocabulários dos usuários com o vocabulário utilizado nos textos especializados. Os seguintes princípios foram ressaltados: 1) ocorre a não-coincidência terminológica se o usuário desconhece o vocabulário de especialidade; 2) A redução dos registros das variações denominativas afeta a comunicação entre o sistema de recuperação da informação e seu usuário; 3) O mesmo processo pode ocorrer com as sinonímias existentes nas áreas especializadas.

O problema de pesquisa centrou-se na integração de procedimentos terminográficos à complexidade da construção de tesouros em que foram buscados procedimentos terminográficos aptos, em conjunto com procedimentos metodológicos existentes de análise de assunto, ao aprimoramento da identificação de conceitos em etapas da construção de tesouros.

A primeira premissa levantada foi de que o aprimoramento de etapas de construção de tesouros contribui para melhorar a representação temática da informação. Sustentamos esta premissa reconhecendo o trabalho implicado de leitura documentária e identificação de conceitos/termos, realizar a anotação dos termos numa ficha terminológica construída para este propósito. Adquire-se assim um recurso contendo: termo, contextos de uso, fonte, domínio/área de especialidade, definições, responsável pelo registro e data.

A segunda premissa foi de que a Terminografia possibilita a harmonização dos conceitos para o acesso à informação desejada em áreas especializadas. O modelo metodológico que se propôs apóia esta última premissa por meio da sequência de cinco etapas: 1. orientações gerais para o desenvolvimento do trabalho de construção do tesouro; 2. método de compilação;

etapa 3. registro de termos; 4. verificação de termos; 5. forma de apresentação de um tesouro.

O estudo foi delineado em termos de seu objeto, integração da Terminografia na construção de tesouros, partindo da hipótese de que há a necessidade de combinar e aliar referenciais teórico-metodológicos da Documentação ao referencial da Terminografia para identificação de conceitos. Derivou-se, como tese de pesquisa, que o aprimoramento de tesouros depende de estudo teórico-metodológico que integre procedimentos terminográficos na identificação de conceitos para a construção de tesouros - e, a partir dela a proposição de investigarmos essa integração no contexto da organização e representação da informação.

Cabe salientar, quando uma área de especialidade não apresenta uma documentação terminológica específica faz-se necessária a integração da Terminografia com o uso de fichas terminológicas na identificação de conceitos em etapas da construção de tesouros. Estabelecemos, desse modo, a proposta de um modelo metodológico de construção de tesouros para áreas de especialidade.

Foi realizado levantamento de referenciais sobre os temas Terminologia/Terminografia e Ciência da Informação/Documentação, com ênfase na construção de tesouros, verificando também aspectos históricos, conceituação, função e importância da Terminologia/Terminografia, procedimentos terminográficos e aproximações com a construção de tesouros. O *corpus* de análise foi constituído de publicações referenciadas em estudos do grupo de pesquisa “Análise Documentária”, coordenado pela Profa. Dra. Mariângela Spotti Lopes Fujita.

O capítulo 2 tratou do objetivo específico de analisar e sintetizar referenciais teórico-metodológicos sobre construção de tesouros. Incluiu a contextualização no âmbito da organização e representação do conhecimento e da informação, o referencial teórico sobre as linguagens documentárias alfabéticas com foco em perspectiva diacrônica de parâmetros e modelos. O capítulo 3 destinou-se a focalizar o objetivo específico de identificar os principais aspectos teórico-metodológicos da Terminologia/Terminografia contribuintes para a construção de tesouros, enriquecidos da reflexão sobre a perspectiva onomasiológica, das recomendações da normalização da Terminologia, dos procedimentos terminográficos realizados na atividade de pesquisa terminológica para áreas de especialidade. O capítulo 4 reservou-se ao objetivo de apresentar proposta de um



modelo metodológico para a construção de tesouros com a integração de procedimentos terminográficos.

A formação de pesquisadores ganha maior impulso com a instalação do Programa de doutorado em Ciência da Informação na Universidade Estadual Paulista, *Câmpus* de Marília-SP, em 2005, algo tão aguardado por professores e alunos da área. Evoluiu para sua consolidação, ampliando suas linhas de pesquisa, com o recente estabelecimento da linha de pesquisa intitulada “Gestão e Políticas da Informação e da Comunicação Midiática”. A produtividade tem condições de se propagar também pelo fato de que o Programa de Pós-Graduação em Gestão da Informação, Mestrado Profissional, da Universidade Estadual de Londrina-PR, prepara-se para dar início a sua segunda turma.

A construção do conhecimento por meio da pesquisa numa universidade pública, centro de excelência da pesquisa brasileira, é assegurar uma identidade científica. Gerado na pesquisa e difundido na extensão, os novos conhecimentos se integram a saberes teórico-metodológico, para elaboração de proposta de modelo metodológico para a realização de determinada atividade.

Embora complexa, a tarefa produz motivação pela busca (e encontro) de soluções, ainda que parciais e temporárias, para problemas existentes numa sociedade em evolução. Nesse caso, trata-se de substituir práticas, em geral, puramente empíricas de construção de linguagens documentárias, especialmente, de tesouros, pelo procedimento da observação orientada por um modelo metodológico que se apresenta como um esforço com finalidade a melhoria.

A melhoria da construção de tesouros além de aprimoramento de método é também um aperfeiçoamento da pessoa que lida com essa face da ciência. Ela adquire uma evolução de natureza cognitiva, obtém nova forma de pensar, passa a sistematizar o seu trabalho, valoriza a observação sistemática e apoiada em método. E ainda mais importante, percebe o valor de integrar procedimentos. Todas essas condutas serão inevitavelmente transportadas para outros setores de sua vida. Do ponto de vista do pesquisador, sem dúvida terá descoberto uma área de grande fascínio para investir suas energias e perseguir objetivos.

## REFERÊNCIAS

ACUERDOS y Recomendaciones. In: ENCUENTRO DE DIRECTORES Y DE DOCENTES DE ESCUELAS DE BIBLIOTECOLOGÍA Y CIENCIA DE LA INFORMACIÓN DEL MERCOSUR, 3., 2., Santiago, Chile, 1998. **Anais**. Santiago: UTEM, 1998. p.15 Apud ABECIN, 2003.

ADELSTEIN, A; FELIU, J. Relações semânticas entre unidades léxicas com valor especializado e descritores. In: KRIEGER, M. G.; ARAÚJO, L. (Orgs.) A Terminologia em Foco. **Cadernos de Tradução**, Instituto de Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre, v. 17, out-dez. 2004.

AITCHISON, J.; GILCHRIST, A. **Manual para construção de tesouros**. Rio de Janeiro: BNG/Brasilart, 1979.

AMARO, R. K. O. F. **Contribuição da análise do discurso para a análise documentária: o caso da documentação jornalística**. 1991. São Paulo: ECA-USP. (Dissertação de Mestrado).

AMAT, C.B.; ALCOCER, A.B. Generación de ficheros de *input* para la construcción y el mantenimiento de tesouros en BRS mediante un sistema de gestión de bases de datos en microordenador. **Rev. Española de Documentación Científica**, v.16, n.2, p.119-136. 1993.

AMERICAN NATIONAL STANDARDS INSTITUTE; NATIONAL INFORMATION STANDARDS ORGANIZATION. **ANSI/NISO Z39.19-2005**: guidelines for the construction, format, and management of monolingual controlled vocabularies. Bethesda, Ma: NISO Press, 2005. 184 p. Disponível em: <http://www.niso.org/standards/resources/Z39-19-2005.pdf> Acesso em: 15 fev. 2008. (tradução com propósito acadêmico),

AMERICAN NATIONAL STANDARDS INSTITUTE; NATIONAL INFORMATION STANDARDS ORGANIZATION. **Z39.19 - Guidelines for the construction, format, and management of monolingual controlled vocabularies**. Bethesda, Ma: NISO, 1993. 184 p. (tradução com propósito acadêmico).

AMERICAN SOCIETY FOR INFORMATION SCIENCE. **ASIS Thesaurus of Information Science and Librarianship**. 2. ed. Medford, NJ: ASIS, 1998. 169 p. (ASIS Monographs Series).

ANDERSON, S. A new horizon: an evaluation of a library online public access catalogue. **Library & Information Research News**, Sheffield, v. 22, n. 72, p. 15-24, Autumn/Winter, 1998.

AQUINO, M. A. (Org.) **O campo da ciência da informação: gênese, conexões e especificidades**. João Pessoa: UFPB, 2002.

ARAÚJO, I. L. **Do signo ao discurso: uma introdução à filosofia da linguagem**. São Paulo: Parábola Editorial, 2004.

ARAÚJO, V. M. P. de. **Documentação, Terminologia e Lingüística: uma interface produtiva**. Porto Alegre, 2006. Dissertação (Mestrado em Letras) – Universidade Federal do Rio Grande do Sul. 2006.

ARRUDA, S. M. de; CHAGAS, J. **Glossário de Biblioteconomia e Ciências afins:** português – inglês. Florianópolis, Cidade Futura, 2002.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 6028:** Informação e Documentação – Resumo - Apresentação. Rio de Janeiro: ABNT, 2003.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 12676:** Métodos para análise de documentos – Determinação de seus assuntos e seleção de termos de indexação. Rio de Janeiro: ABNT, 1992.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 13789:** Terminologia – Princípios e métodos – Elaboração e apresentação de normas de terminologia. Rio de Janeiro: ABNT, 1997a.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 13790:** Terminologia – Princípios e métodos – Harmonização de conceitos e termos. Rio de Janeiro: ABNT, 1997b.

ASSOCIAÇÃO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO. **GT 2 - Organização e Representação do Conhecimento.** Disponível em: <<http://www.ancib.org.br/grupos-de-trabalho/gt-2-organizacao-erepresentacao-do-conhecimento/>>. Acesso em: 29 out. 2008.

AUBERT, F. H. **Introdução à metodologia terminológica bilíngüe.** 2.ed. São Paulo: Humanitas, 2001.

AUSTIN, D.; DALE, P. **Diretrizes para o estabelecimento e desenvolvimento de tesauros monolíngües.** Brasília: IBICT/SENAI, 1993.

BARBOSA, A. P. Classificações facetadas. **Ciência da Informação**, Rio de Janeiro, v.1, n. 2, p. 73-81, 1972.

BARBOSA, M. A. Lexicologia, lexicografia, terminologia, terminografia: objeto, métodos, campos de atuação e de cooperação. In: SIMPÓSIO LATINO-AMERICANO DE TERMINOLOGIA, 2.; ENCONTRO BRASILEIRO DE TERMINOLOGIA TÉCNICO-CIENTÍFICA, 1., 1990, Brasília. **Anais.** Brasília: União Latina/CNPq/IBICT, 1990. p. 152-158.

BARBOSA, M. A. A Terminologia e o ensino da metalinguagem técnico-científica. In: ISQUERDO, A. N.; KRIEGER, M. das G. (Orgs.) **As ciências do léxico:** lexicologia, lexicografia, terminologia. Campo Grande: Ed. UFMS, 2004. v. 2. p. 311-325.

BARDIN, L. **Análise de conteúdo.** Lisboa: Edições 70, 1995, 2000.

BARITÉ, M. **Dicionário de Organização do Conhecimento:** Classificação, Indexação, Terminologia. Montevideo: EUBCA, 2008. Disponível em: <<http://www.eubca.edu.uy/barite/portugues/index.php>>. Acesso em: 10 jun. 2009.

BARITÉ, M. Relato de un proyecto interdisciplinario de terminología: banco de datos terminológico sobre Contratación Comercial Internacional – Proyecto “LEGISTERM”. In: Unión Latina. **El estado del arte de la Terminología y la Neología en el Uruguay:** ponencias y trabajos: jornada realizada en Unión Latina, Montevideo, el 24 de junio de 2004. Montevideo : Unión Latina, 2005.

BARITÉ, M. **Diccionario de organización y representación del conocimiento, clasificación, indización, terminología.** 2000a. Disponível em: <<http://eubca1.eubca.edu.uy/diccionario/diccion.htm>>. Acesso em: 20 set. 2008.

BARITÉ, M. Los conceptos y su representación: una perspectiva terminológica para el tratamiento temático de la información. **Scire**, Zaragoza, v. 6, n. 1, p. 31-53, 2000b.

BARITÉ, M. Organización del conocimiento: um nuevo marco teórico-conceptual em Bibliotecologia y Documentacion. In: In: CARRARA, K. (Org.). **Educação, universidade e pesquisa**. Marília: Unesp-Marília-Publicações; São Paulo: FAPESP, 2001. p. 35-60.

BARRETO, A. A. A questão da informação. **São Paulo em Perspectiva**: revista da Fundação SEADE, v.8, n.4, p.3-8, 1994.

BARROS, L. A. **Curso básico de Terminologia**. São Paulo: EdUSP, 2004.

BARROS, D. L. P. **Teoria semiótica do texto**. São Paulo: Ática, 2006.

BATTAGLIA, M. G. B. Tesouro de química uma ferramenta de recuperação importante para a inteligência competitiva. In: WORKSHOP BRASILEIRO DE INTELIGÊNCIA COMPETITIVA & GESTÃO DO CONHECIMENTO, 2., 2001, Florianópolis. **Anais eletrônicos**. Florianópolis: FIESC/SENAI/IEL, 2001.

BATTY, D. Thesaurus construction and maintenance: a survival kit. **Database**, v.12, n.1, feb., p.13-20, 1989.

BENTES PINTO, V. Representação do conhecimento em Ciência da Informação através dos esquemas de tarefa (ETAF). In: AQUINO, Mirian de Albuquerque. **O campo da Ciência da Informação: gênese, conexões e especificidade**. João Pessoa: Ed. Universitária da UFPB, 2002.

BENVENISTE, E. **Problemas de lingüística geral**. São Paulo : Companhia Editora Nacional, 1976.

BENVENISTE, E. *Problèmes de linguistique générale*, II. Paris, Gallimard, 1974.

BERLO, D. K. **O processo de comunicação**. 21.ed. Rio de Janeiro: Fundo de Cultura, 1999.

BIDERMAN, M. T. C. A definição lexicográfica. **Cadernos do Instituto de Letras**, Porto Alegre, n. 10, p. 23-43, jul. 1993.

BOCCATO, V. R. C. **Avaliação de linguagem documentária em Fonoaudiologia na perspectiva do usuário**: estudo de observação da recuperação da informação com protocolo verbal. 2005. 239 f. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2005. Disponível em: <<http://www.marilia.unesp.br/>>. Acesso em: 20 set. 2008.

BOCCATO, V. R. C. **Avaliação do uso de linguagem documentária em catálogos coletivos de bibliotecas universitárias**: um estudo sociocognitivo com protocolo verbal. Marília, 2009. 301f. Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, 2009.

BOCCATO, V. R. C.; RAMALHO, R. A. S.; FUJITA, M. S. L. A contribuição dos tesouros na construção de ontologias como instrumento de organização e recuperação da informação em ambientes digitais. In: García Marco, F. J. (Ed.). **Avances y perspectivas en sistemas**

**de información y documentación** - *IBERSID*, 2008. Zaragoza: Universidad de Zaragoza, 2008. p. 199-209.

BORKO, H. Information science: what is it? *American Documentation*, v.19, n.1, p.1-5, Jan. 1968.

BOULANGER, J. C. Alguns componentes lingüísticos no ensino da terminologia. *Ciência da Informação*, Brasília, v.24, n.3, p.313-318, set./dez. 1995.

BRÄSCHER, M.; CAFÉ, L. Organização da Informação ou Organização do Conhecimento? In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 9, 2008, São Paulo, *Anais*. São Paulo: ANCIB, 2008. Disponível em: < <http://www.enancib2008.com.br> >. Acesso em: 20 fev. 2009.

BRITO, M. A. B. **Análise do processo de leitura de futuros bibliotecários para elaboração de resumos**. 2001. Trabalho de Conclusão de Curso (Graduação em Biblioteconomia) – Universidade Estadual Paulista, Marília. 2001.

BUCKLAND, M. Information as a thing. *JASIS*, v.42, n.5, p.351-360, 1991.

BUSH, V. As we may think. *Atlantic Monthly*, v. 176, n. 1, p. 101-108, July 1945.

CABRÉ, M. T. **La terminología**: teoría, metodología, aplicaciones. Barcelona: Editorial Antártida/Empuries, 1993.

CABRÉ, M. T. La terminologia hoy: concepciones, tendencias y aplicaciones. *Ciência da Informação*, Brasília, v.24, n.3, p.289-298, set./dez. 1995.

CABRÉ, M. T. La Terminología hoy: replanteamiento o diversificación. *Organon*, Porto Alegre, v. 12, n. 26, p. 33-41, 1998.

CABRÉ, M. T. **La terminología**: representación y comunicación: elementos para una teoría de base comunicativa y otros artículos. Barcelona: Institut Universitari de Lingüística Aplicada, 1999.

CABRÉ, M. T. Terminologie et linguistique: la théorie des portes. *Terminologies Nouvelles*, n.21, p.10-15, 2000.

CABRÉ, M. T. Terminologie ou terminologies? Spécialité linguistique ou domaine interdisciplinaire? *Meta*, v.36, n.1, p.55-63, 1991.

CABRÉ, M. T. A Terminologia hoje: concepções, tendências e aplicações. In: KRIEGER, M. G.; ARAÚJO, L. (Orgs.) A Terminologia em Foco. **Cadernos de Tradução**, Instituto de Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre, v. 17, out-dez. 2004.

CAMPOS, A.T. Linguagens documentárias. *Revista de Biblioteconomia de Brasília* v. 14, n. 1, p. 85-88, jan./jun. 1986.

CAMPOS, M. L. A. **Linguagem documentária**. teorias que fundamentam sua elaboração. Niterói: EdUFF, 2001.

CAMPOS, M. L. A. **Em busca de princípios comuns na área de representação da informação**: uma comparação entre o método de classificação facetada, o método de tesouro-baseado-em-conceito e a teoria geral da terminologia. 1994. 196f. Dissertação

(Mestrado em Ciência da Informação) - Universidade Federal do Rio de Janeiro, Rio de Janeiro.

CAMPOS, M. L. A. Perspectivas para o estudo da área de representação da informação. **Ciência da Informação**, Brasília, v.25, n.2, 1995.

CAMPOS, M. L. A.; GOMES, H. E. Metodologia de elaboração de tesouro conceitual: a categorização como princípio norteador. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 11, n. 3, p.348-359, set./dez. 2006.

CAMPOS, M. L. A. et al. Estudo comparativo de softwares de construção de tesouros. **Perspectivas em Ciência da Informação**, Belo Horizonte, v.11 n.1, p. 68-81, jan./abr. 2006

CARNEIRO, M. V. Diretrizes para uma política de indexação. **Revista da Escola de Biblioteconomia da UFMG**, Belo Horizonte, v.14, n.2, p.221-241, set. 1985.

CAVALCANTI, M. C. **Interação leitor-texto**: aspectos de interação pragmática. Campinas: Ed. da UNICAMP, 1989.

CAVALCANTI, M. C.; ZANOTTO, M. S. Introspection in Applied Linguistics: meta-research on verbal protocols. In: BARBARA, L.; SCOTT, M. (Ed.). **Reflections on Language Learning**. Clevedon, Avon: Multilingual Matters, 1994. p.148-56.

CERÁVOLO, S. M. **Da palavra ao termo**: um caminho para compreender Museologia. 2004. Tese (Doutorado em Ciências da Comunicação: Ciência da Informação e Documentação) – Universidade de São Paulo, Escola de Comunicações e Artes, São Paulo.

CERVANTES, B. M. N. **Contribuição para a Terminologia do Processo de Inteligência Competitiva**: estudo teórico e metodológico. Marília, 2004. 183f. Dissertação (Mestrado em Ciência da Informação) Programa de Pós-Graduação em Ciência da Informação– Universidade Estadual Paulista, Marília.

CERVANTES, B. M. N.; FUJITA, M. S. L.; NARDI, M. I. A. Pesquisa terminológica para a elaboração de linguagem documentária. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 5., 2003, Belo Horizonte. **Informação, Conhecimento e Transdisciplinaridade**: anais eletrônicos. Belo Horizonte: UFMG, Escola de Ciência da Informação, 2003. 1 CD-ROM.

CERVANTES, B. M. N.; MORAES, J. B. E.; FUJITA, M. S. L. Terminologia em inteligência competitiva. In: VI ENCONTRO DE DIRETORES; V ENCONTRO DE DOCENTES DE BIBLIOTECONOMIA E CIÊNCIA DA INFORMAÇÃO DO MERCOSUL, EBCIN, 2002. Londrina: UEL, 2002.

CERVANTES, B. M. N.; FUJITA, M. S. L.; RUBI, M. P. Terminologias em política de indexação In: GARCÍA MARCO, F. J. (Ed.). **Avances y perspectivas en sistemas de información y documentación**, IBERSID, 2008. Zaragoza: Universidad de Zaragoza, 2008. (Número monográfico).

CERVANTES, B. M. N.; FUJITA, M. S. L.; MORAES, J. B. E. de. Terminología del Proceso de inteligencia Competitiva: estudio teórico y metodológico. In: X ENCUENTROS INTERNACIONALES SOBRE SISTEMAS DE INFORMACIÓN Y DOCUMENTACIÓN, IBERSID, 2005. Zaragoza. **Anais...** Zaragoza: Universidad de Zaragoza, 2005.

CERVANTES, B. M. N.; MORAES, J. B. E. de.; FUJITA, M. S. L. Pesquisa Terminológica Temática: o processo de identificação e confirmação de termos In: VI Simpósio em Filosofia

e Ciência, 2005, Marília. **Universidade e contemporaneidade**: produção do conhecimento e formação profissional - coletânea de textos. Marília: FFC/Marília/UNESP - Comissão Permanente de Publicações, 2005. v.1.

CERVANTES, B. M. N.; VALENTIM, M. L. P. Terminologia em Inteligência Competitiva Organizacional In: VALENTIM, M. L. P. (Org.) **Informação, conhecimento e inteligência organizacional**. 2. ed. Marília: FUNDEPE, 2007. (Cap. 12)

CERVO, A. L.; BERVIAN, P. A. **Metodologia científica**. 5. ed. São Paulo: Prentice-Hall, 2002. 242 p.

CHAUMIER, J. **Analisis y lenguajes documentales**. Barcelona: Mitre, 1986.

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

CINTRA, A. M. M. et al. **Para entender as linguagens documentárias**. 2.ed. São Paulo: Polis, 2002. (Coleção Palavra-Chave, 4).

COHEN, D. M. **O consumidor da informação documentária**: o usuário de sistemas documentários visto sob a lente da análise documentária. 1995. 145f. Dissertação (Mestrado em Ciência da Informação e Documentação) – Universidade de São Paulo, São Paulo.

CORRÊA, A. O. **A construção de tesouros na perspectiva da metodologia facetada**. Marília: Unesp, 1999. (Relatório Final de Pesquisa apresentado à FAPESP, orientado por Mariângela Spotti Lopes Fujita).

CROUCH, C. J. An approach to the automatic construction of global thesauri. **Information Processing & Management**, v.26, n.5, p.629-40, 1990.

CUNHA, I. M. R. (Comp.). **Análise documentária**: considerações teóricas e experimentações. São Paulo: FEBAB, 1989.

CURRÁS, E. **Tesouros**: manual de construcción y uso. Madrid: Kaher II, 1998.

CURRÁS, E. **Ontologias, taxonomia y tesouros**: manual de construcción y uso. 3. ed. atual. ampl. Gijón: TREA, 2005.

CUTTER, C. A. **Rules for a dictionary catalog**. 4. ed. Government Printing Office: Washington, 1904.

DAHLBERG, I. Knowledge organization: a new science? **Knowledge Organization**, Frankfurt, v.33, n.1, p. 11-19, 2006.

DAHLBERG, I. Knowledge organization: its scope and possibilities. In: **Knowledge Organization**, Frankfurt, v.20, n.4, p.211-222, 1993.

DAHLBERG, I. O futuro das linguagens de indexação. In: CONFERÊNCIA BRASILEIRA DE CLASSIFICAÇÃO BIBLIOGRÁFICA, 1976. **Anais...** Rio de Janeiro, IBICT/ABDF, 1979. v. 1, p. 323-334.

DAHLBERG, I. A referent-oriented, analytical concept theory of Interconcept. **International Classification**, Frankfurt, v. 5, n. 3, p. 122-151, 1978a.

DAHLBERG, I. Teoria da classificação, ontem e hoje. In: CONFERÊNCIA BRASILEIRA DE CLASSIFICAÇÃO BIBLIOGRÁFICA, 1976. **Anais...** Rio de Janeiro, IBICT/ABDF, 1979. v. 1, p. 352-370.

DAHLBERG, I. Teoria do conceito. **Ciência da Informação**, Rio de Janeiro, v.7, n.2, p.101-107, 1978b.

DAVENPORT, T. H. **Ecologia da informação**: porque só a tecnologia não basta para o sucesso da informação. São Paulo: Futura, 1998.

DEPECKER, L. Terminologie et standardisation. **Hieronymus Complutenses**, Madrid, n. 3, p. 85-88, ene.-jun. 1996. Disponível em: <  
[http://cvc.cervantes.es/lengua/hieronymus/pdf/03/03\\_085.pdf/](http://cvc.cervantes.es/lengua/hieronymus/pdf/03/03_085.pdf/) >. Acesso em: 23 jun. 2009

DEVADASON, F.J. On-line construction of alphabetic CLASSAURUS: a vocabulary control and indexing tool. **Information Processing & Management**, v.21, n.1, p.11-26, 1985.

DIAS, J. G. **A terminologia da geografia política**. 1999. Dissertação (Mestrado em Filologia e Língua Portuguesa) – Universidade de São Paulo, São Paulo.

DICIONÁRIO Eletrônico Houaiss da língua portuguesa. Rio de Janeiro: Objetiva, 2001. 1 CD-ROM.

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

DUBOIS, J. et al. **Dicionário de Lingüística**. 8.ed. São Paulo: Cultrix, 2001.

DUBUC, R. **Manual práctico de terminología**. 3. ed. corr. actual. Traducción de Ileana Cabrera. Providencia, Chile: RiL, 1999. 236 p. (Colección de Terminología y Traducción).

DUCROT, O.; TODOROV, T. **Dicionário enciclopédico das ciências da linguagem**. 2.ed. São Paulo: Perspectiva, 1988.

FAULSTICH, E. **Base metodológica para pesquisa em socioterminologia**: termo e variação. Brasília: UnB, 1995a.

FAULSTICH, E. A socioterminologia na comunicação científica e técnica. **Ciência e Cultura**. abr./ jun. 2006, v.58, n.2, p.27-31.

FAULSTICH, E. Metodologia para projeto terminológico. In: SIMPÓSIO LATINO-AMERICANO DE TERMINOLOGIA, 2., 1990, Brasília. **Anais...** Brasília: IBICT, 1992. p. 206-217.

FAULSTICH, E. Terminologia: disciplina da nova era. **Ciência da Informação**, Brasília, v.24, n.3, p. 275-365, set./dez., 1995b.

FÁVERO, L. L.; KOCH, I. G. V. **Lingüística textual**: uma introdução. 5.ed. São Paulo: Cortez, 2000.

FEDOR DE DIEGO, A. de. **Terminologia**: teoria y practica. Venezuela: União Latina, 1995.

FELBER, H. **Manuel de terminologie**. Paris: Unesco-Infoterm, 1984, 1987.



FELBER, H. Manuel de terminologie. Paris: Unesco-Infoterm, 1987. Apud FEDOR DE DIEGO, A. de. **Terminologia**: teoria y práctica. Venezuela: União Latina, 1995.

FERREIRA, A. B. de H. **Novo dicionário eletrônico Aurélio**. versão 5.0. São Paulo: Positivo, 2004.

FOSKETT, A. C. **A abordagem temática da informação**. Tradução de Antonio Agenor Briquet de Lemos. São Paulo: Polígono, 1973. 437 p.

FREITAS, H.; JANISSEK, R. **Análise léxica e análise de conteúdo**: técnicas complementares, seqüenciais e recorrentes para exploração de dados qualitativos. Proto Alegre: Sagra Luzzatto, 2000.

FUJINO, A.; JACOMINI, D. D. Produtos e serviços de informação na sociedade do conhecimento: da identificação ao uso. In: GIANNASI-KAIMEN, M. J.; CARELLI, A. E. (Org). **Recursos informacionais para compartilhamento da informação**: redeseenhando acesso, disponibilidade e uso. Rio de Janeiro: E-papers, 2007. 226 p.

FUJITA, M. S. L. Análise e síntese documentárias para compreensão de leitura de textos didáticos: uma proposta de aplicação do sistema de indexação PRECIS. **Informare**, v.5, n.1, p. 77-94, 1999.

FUJITA, M. S. L. Aspectos evolutivos das bibliotecas universitárias em ambiente digital na perspectiva da rede de bibliotecas da UNESP. **Informação & Sociedade**: estudos, João Pessoa, v.15, n. 2, p. 97-112, jul./dez. 2005.

FUJITA, M. S. L. **O contexto sócio-cognitivo do catalogador em bibliotecas universitárias**: perspectivas para uma política de tratamento da informação documentária. Marília: UNESP, 2006. 17 f. (Projeto de Pesquisa)

FUJITA, M. S. L. A Estrutura de categorias do tesauro: modelos de elaboração. **Cadernos do F.F.C**, Marília, v.7, n.12, p. 107-120, 1998.

FUJITA, M. S. L. **A leitura documentária do indexador**: aspectos cognitivos e lingüísticos influentes na formação do leitor profissional. 2003. 321f. Tese (Livre Docência em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília.

FUJITA, M. S. L. A leitura do indexador: estudo de observação. **Perspectivas em Ciência da Informação**, Belo Horizonte, v.4, n.1, p. 101-116, jan./jun. 1999.

FUJITA, M. S. L. **Linguagem documentária em Odontologia**: uma aplicação do sistema de indexação PRECIS. 1992. Tese (Doutorado em Ciências da Comunicação) – Escola de Comunicação e Artes, Universidade de São Paulo, São Paulo. 3v.

FUJITA, M. S. L. **Linguagens documentárias alfabéticas em análise documentária**: aspectos de estrutura e funcionalidade. Marília: [UNESP], 2005. (Material didático).

FUJITA, M. S. L. Política científica em Ciência da Informação da graduação à pós-graduação. In: GUIMARÃES, J. A. C.; FUJITA, M. S. L. (Orgs.). **Ensino e pesquisa em Biblioteconomia no Brasil**. Marília: Fundepe; São Paulo: Cultura Acadêmica, 2008b.

FUJITA, M. S. L. Organização e representação do conhecimento no Brasil: análise de aspectos conceituais e da produção científica do ENANCIB no período de 2005 a 2007. **Tendências da Pesquisa Brasileira em Ciência da Informação**, Brasília, v. 1, n. 1, 2008a.

Disponível em: <[inseer.ibict.br/ancib/index.php/tpbci/article/viewPDFInterstitial/4/13](http://inseer.ibict.br/ancib/index.php/tpbci/article/viewPDFInterstitial/4/13)>. Acesso em: 12 jan. 2008.

FUJITA, M. S. L. **Precis na língua portuguesa: teoria e prática de indexação**. Brasília: Ed. UnB, 1988.

FUJITA, M. S. L.; CERVANTES, B. M. N. Abordagem Cognitiva do Protocolo Verbal na Confirmação de Termos para a Construção de Linguagem Documentária em Inteligência Competitiva. In: VALENTIM, M. L. P. (Org.). **Métodos Qualitativos de Pesquisa em Ciência da Informação**. São Paulo: Polis, 2005. (cap. 1).

FUJITA, M. S. L.; CERVANTES, B. M. N. A abordagem cognitiva do Protocolo Verbal para a construção de linguagem documentária em inteligência competitiva. In: SIMPOSIO IBEROAMERICANO DE TERMINOLOGÍA (RITERM), 10. Montevideo, Tema central: **Terminología, Conocimientos, Sociedad y Poder**. 2006.

FUJITA, M. S. L.; NARDI, M. I. A.; FAGUNDES, S. A. A observação da leitura documentária por meio de Protocolo Verbal. In: RODRIGUES, G. M.; LOPES, I. L. (Org.). **Organização e representação do conhecimento na perspectiva da Ciência da Informação**. Brasília: Thesaurus, 2003. (Estudos Avançados em Ciência da Informação; v.2).

FUJITA, M. S. L.; NARDI, M. I. A.; SANTOS, S. A leitura em análise documentária. **Transinformação**, Campinas, v.10, n.3, p. 13-31, set./dez. 1998.

GARCIA GUTIERREZ, A. Elementos de lingüística en sistemas de información y documentación. **Revista Latina de Comunicación Social**, n.7, jul., 1998. Disponível em: <http://www.lazarillo.com/latina/a/66ant.htm>. Acesso em: 30 jan. 2004

GARCIA GUTIERREZ, A. **Estructura lingüística de la documentación: teoría y metodo**. Barcelona: Universidad de Murcia, 1990a.

GARCIA GUTIERREZ, A. Suficiencia estructural y tipología de la omisión en análisis documental. **Documentación de las Ciencias de la Información**, Madrid, n.13, p. 73-86, 1990b.

GARCIA GUTIERREZ, A.; LUCAS FERNÁNDEZ, R. **Documentación automatizada en los medios informativos**. Madri: Paraninfo, 1987.

GARDIN, J-C. Document analysis and linguistic theory. **Journal of Documentation**, London, v.29, n.2, p.137-168, Jun. 1973.

GARDIN, J-C. **Les analyses des discours**. Neuchatêl: Delchaux et Niestlé, 1974.

GARDIN, J-C. Éléments d'un modèle pour la description des lexiques documentaires, **BBF**, 1966, n. 05, p. 171-182. Disponível em: <<http://bbf.enssib.fr>> Acesso em: 23 jun 2008.

GAUDIN, F. **Pour une socioterminologie: des problèmes sémantiques aux pratiques institutionnelles**. Rouen: Université de Rouen, 1993.

GIL, A. C. **Como elaborar projetos de pesquisa**. 4. ed. São Paulo: Atlas, 2008. 175 p.

GIL, A. C. **Métodos e técnicas de pesquisa social**. São Paulo: Atlas, 1999.

GIL URDICIÁIN, B. **Manual de lenguajes documentales**. 2. ed. rev. e ampl. Madrid: Ed. NOESIS, 2004. 280 p.

GILREATH, C. T. Harmonization of terminology: an overview of principles. **International classification**, Frankfurt, v. 19, n. 3, p. 135-139, 1992.

GOMES, H. E. (Coord.). Elaboração de tesauro documentário. In: **Biblioteconomia, Informação & Tecnologia da Informação (BITI)**. Rio de Janeiro, [2004]. Disponível em: <http://www.conexaorio.com/bitit/index.htm> Acesso em: 20 ago. 2008.

GOMES, H. E. **Manual de elaboração de tesouros monolíngües**. Brasília: Programa Nacional de Bibliotecas das Instituições de Ensino Superior, 1990. 78 p.

GOMES, H. E. **Classificação, tesauro e terminologia**: fundamentos comuns. Disponível em: <<http://www.conexaorio.com/bitit/tertulua/tertulua.htm#7>>. Acesso em: 10 set. 2001.

GOMES, H. E.; CAMPOS, M. L. de A. Tesauro e normalização terminológica: o termo como base para intercâmbio de informações. **DataGramZero** - Revista de Ciência da Informação, v.5, n.6,dez. 2004, art.02. Disponível em: [www.dgz.org.br/dez04/Art\\_02.htm](http://www.dgz.org.br/dez04/Art_02.htm)>. Acesso em: 23 abr. 2007.

GONZALEZ DE GOMEZ, M. N. Dos estudos sociais da informação aos estudos do social desde o ponto de vista da Informação.. In: AQUINO, M.A (org.), **O Campo da Ciência da Informação – Gênese, conexões e Especificidades**. João Pessoa: Ed. Universitária, 2002. p. 25-48.

GONZALEZ DE GOMEZ, M. N. A representação do conhecimento e o conhecimento da representação: algumas questões epistemológicas. **Ciência da Informação**, Brasília, v.22, n.3, p.217-222, set./dez. 1993.

GRAND Dictionnaire de la psychologie. Paris: Larousse, 1991

GUIMARÃES, J. A. C. A análise documentária no âmbito do tratamento da informação: elementos históricos e conceituais. In: RODRIGUES, G.M.; LOPES, I.L. (Org.). **Organização e representação do conhecimento na perspectiva da Ciência da Informação**. Brasília: Thesaurus, 2003. p. 100-117. (Estudos Avançados em Ciência da Informação; v.2)

GUIMARÃES, J. A. C. Perspectivas de ensino e pesquisa em organização do conhecimento em cursos de Biblioteconomia: uma reflexão. In: CARRARA, K. (Org.). **Educação, Universidade e Pesquisa**. Marília: Unesp-Marília-Publicações; São Paulo: FAPESP, 2001.

GUIMARÃES, J. A. C. Análise documentária em jurisprudência: uma metodologia de indexação de acórdãos para instrumentos terminológicos. SIMPOSIO IBEROAMERI-CANO DE TERMINOLOGÍA – RITerm. 4. Buenos Aires, 1994. **Terminología y desarrollo**: actas.Disponível em: <http://www.riterm.net/actes/4simposio/chaves.htm>> Acesso em: 20 ago. 2008.

GUIMARÃES, J. A. C. **Condensação documentária em legislação e jurisprudência**: elementos instrumentais para a construção de ementas. 2000. Tese (Livre Docência em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2000.

GUIMARÃES, J. A. C. A recuperação temática da informação. **Revista Brasileira de Biblioteconomia e Documentação**, São Paulo, v.16, n.3/4, p.112-130, jan./dez. 1990.

- GUINCHAT, C.; MENO, M. As linguagens documentais. In: \_\_\_\_\_. **Introdução geral às ciências e técnicas da informação e documentação**. 2. ed. corr. aum. por Marie France Blanquet. Tradução de de Míriam Vieira da Cunha. Brasília: IBICT, 1994. p. 131-165.
- GUNTZER, V. et al. Automatic thesaurus construction by machine learning from retrieval sessions. **Information Processing and management**, v.25, n.3, p.265-73, 1989.
- HJELMSLEV, L. **Prolegômenos a uma teoria da linguagem**. Tradução de J. Teixeira Coelho Netto. 2. ed. São Paulo: Perspectiva, 2006. (Estudos; 43).
- HJØRLAND, B. J. Domain analysis in information science: eleven approaches – traditional as well as innovative. **Journal of Documentation**, London, v. 58, p.422-462, 2002
- HJØRLAND, B. J. Theory of knowledge organization and the feasibility of universal solutions. In: **International ISKO Conference**, 8.: London, 2004. London, ISKO, 2004
- HJØRLAND, B. J. What is Knowledge Organization (KO)? **Knowledge Organization: international journal devoted to concept theory, classification, indexing, and knowledge representation**, Frankfurt, v. 35, n. 2/3, p. 86-101, 2008.
- HJØRLAND, B; ALBRECHTSEN, H. Toward a New Horizon in Information Science: domain analysis. **Journal of the American Society for Information Science**, 1995, v.46, n.6, p.400-425.
- HODGE, G. **Systems of knowledge organization for digital libraries: beyond traditional authority files**. Washington: The Digital Library Federation, The Council on Library and Information Resources, 2000. 37 p. Disponível em: <<http://www.clir.org/pubs/abstract/pub91abst.html>>. Acesso em: 20 mar. 2008.
- HUDON, M. Multilingual thesaurus construction: integrating the views of different cultures in one gateway to knowledge and concepts. **Knowledge Organization**, Frankfurt, v. 24, n. 2, p. 84-91, 1997.
- HOSENFELD, C. A preliminary investigation of the reading strategies of successful and unsuccessful second language learners. **System**, v.5, n. 2, p.110-123, 1977.
- HUERTAS-LOPEZ, M. J. Thesaurus structure design: a conceptual approach for improved iteration. **Journal of Documentation**, v.53, n.2, p. 139-177, mar. 1997.
- IGLESIAS MATURANA. M. T. **Análisis y recuperación de información documental**. Santiago: I.P.S., 1992.
- INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA. **Diretrizes para elaboração de tesouros monolíngües**. Brasília: IBICT, 1984.
- INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA. **Tesouro de Ciência da Informação**: versão preliminar. Brasília: IBICT, 1989a.
- INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA. **Programa para elaboração de tesouros em microcomputador (TECER)**. Brasília: IBICT, 1989b.
- INTERNATIONAL SOCIETY FOR KNOWLEDGE ORGANIZATION. ISKO, 2008. **Mission**. Disponível em: <<http://www.isko.org/about.html>>. Acesso em: 12 abr. 2009.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 2788-1986:** Documentation - Guidelines to establishment and development of monolingual thesauri. Geneva : ISO, 1986. Disponível em: <http://www.collectionscanada.ca/iso/tc46sc9/standard/2788e.htm> Acesso em: 12 maio. 2006.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 704 - Terminology work – principles and methods** = Travail terminologique – principes et méthodes. 2.ed. Geneva : ISO, [2000, 2001].

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 1087 – Terminologia – Vocabulário**. Proposta de versão brasileira. [Rio de Janeiro: ABNT, 1996].

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 1087-1 - Terminology work – Vocabulary – p.t. 1: Theory and application**. [Geneva] : ISO, 1995, [2000].

IPHAN. Centro Nacional de Folclore e Cultura Popular. Tesouro de Folclore e Cultura Popular Brasileira. Rio de Janeiro, 2005. Disponível em: <<http://www.cnfcp.gov.br/tesouro/>> Acesso em: 17 fev. 2009

IZQUIERDO ARROYO, J.-M. El marco documental. In: \_\_\_\_\_. **La organización documental del conocimiento**. Madrid: Tecnidoc, 1995. v.1/1, p. 19-21.

JACOB, E.K., SHAW, D. Sociocognitive perspectives on representation. In: WILLIAMS, M. E. (Ed.). **Annual Review of Information Science and Technology**. Medford: Information Today for the American Society for Information Science, 1998. v.33, p. 131-185.

JANIK, Sophie. La mise à jour d'un thésaurus. **Documentaliste**, v.24, n.6, nov./dez., p.215-229, 1987.

JÁTIVA MIRALLES, M. V. Indicadores de calidad aplicables al análisis, evaluación y comparación de opacs. **El profesional de la información**, Barcelona, v. 13, n. 1, p. 28-44, ene/feb. 2004.

KING, D. W.; TENOPIR, C. A publicação de revistas eletrônicas: economia da produção, distribuição e uso. **Ciência da Informação**, Brasília, v.27, n.2, p.176-182, 1998.

KOBASHI, N. Y. Análise documentária: considerações sobre um modelo lógico-semântico. In: **ANÁLISE** documentária: considerações teóricas e experimentações. São Paulo: FEBAB. 1989.

KOBASHI, N. Y. Análise documentária e representação da informação. **Informare**, Rio de Janeiro, v.2, n.2, p.5-27, jul./dez. 1996.

KOBASHI, N. Y. **A elaboração de informações documentárias**: em busca de uma metodologia. 1994. Tese (Doutorado em Ciências da Comunicação) – Universidade de São Paulo, São Paulo.

KOBASHI, N. Y. Fundamentos semânticos e pragmáticos da construção de instrumentos de representação da informação. **DataGramaZero** - Revista de Ciência da Informação, Brasília, v.8, n.6, art. 01, dez. 2007.

KOCH, I. V. **A Inter-ação pela Linguagem**. 10 ed. São Paulo: Contexto, 2006.

KOCH, I. G. V. **A inter-ação pela linguagem**: linguagem e sociedade, a construção interativa dos sentidos no texto, estratégias dos jogos de linguagem. 4.ed. São Paulo: Contexto, 1998.

KRIEGER, M. G. Terminologia em contextos de integração: funcionalidade e fundamentos. **Organon**, Porto Alegre, v. 12, n. 26, 1998.

KRIEGER, M. G.; ARAÚJO, L. (Orgs.) A Terminologia em Foco. **Cadernos de Tradução**, Instituto de Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre, v. 17, out-dez. 2004.

KRIEGER, M. G. Terminologia revisitada. In: KRIEGER, M. G.; MACIEL, A. M. B. **Temas de terminologia**. Porto Alegre: Ed. Universidade/ UFRGS; São Paulo: Humanitas/USP, 2001.

KRIEGER, M. da G.; BEVILACQUA, C. R. **A pesquisa terminológica no Brasil**: uma contribuição para a consolidação da área. 2005. Disponível em: <[http://www.riterm.net/revista/n\\_1/krieger.pdf](http://www.riterm.net/revista/n_1/krieger.pdf)>. Acesso em: 08/01/2006.

KRIEGER, M. G.; FINATTO, M. J. B. **Introdução à terminologia**: teoria e prática. São Paulo: Contexto, 2004.

LANCASTER, F. W. **Construção e uso de tesouros**: curso condensado. Tradução de Antonio Agenor Briquet de Lemos. Brasília: IBICT, 1987. 106 p.

LANCASTER, F. W. **El control del vocabulario en la recuperación de información**. 2. ed. València: Universitat de València, 2002. (Educació. Materials, 12).

LANCASTER, F. W. **Indexação e resumos**: teoria e prática. Tradução de Antonio Agenor Briquet de Lemos. 2. ed. Brasília: Briquet de Lemos, 2004.

LANCASTER, F. W. **El control del vocabulario en la recuperación de información**. Tradução de Alejandro de la Cueva Martín. València: Universitat de València, 1995. (Educació. Materials, 12).

LANCASTER, F. W. **Vocabulary control for information retrieval**. Arlington, Virgínia : Information Resources, 1986.

LARA, M. L. G. de. Algumas contribuições da semiologia e da semiótica para a análise das linguagens documentárias. **Ciência da Informação**, Brasília, v.22, n.3, p.223-226, set./dez. 1993a.

LARA, M. L. G. de. Linguagens documentárias, instrumentos de mediação e comunicação. **Revista Brasileira de Biblioteconomia e Documentação**, São Paulo, v.26, n.1/2, p.72-80, jan./jun. 1993b.

LARA, M. L. G. de. **Linguística documentária**: seleção de conceitos. São Paulo, 2009. Tese (Livre-Docência em Análise Documentária) – Universidade de São Paulo, Escola de Comunicações e Artes, São Paulo.

LARA, M. L. G. de. **A representação documentária**: em jogo a significação. 1993c. Dissertação (Mestrado em Ciências da Comunicação) – Universidade de São Paulo, São Paulo.

LARA, M. L. G. de. **Representação e linguagens documentárias**: bases teórico-metodológicas. 1999. Tese (Doutorado em Ciências da Comunicação) – Universidade de São Paulo, Escola de Comunicações e Artes, São Paulo.

LARA, M. L. G. de. O unicórnio (o rinoceronte, o ornitorrinco...), a análise documentária e a linguagem documentária. **DataGramZero**, v.2, n.6. 2001. Disponível em: <[http://www.datagramazero.org.br/dez01/Art\\_03.htm](http://www.datagramazero.org.br/dez01/Art_03.htm)>. Acesso em 20 abr. 2003

LARA, M. L. G. de, A terminologia como instrumento para a construção de ferramentas semânticas. In: CONGRESSO BRASILEIRO DE BIBLIOTECONOMIA E DOCUMENTAÇÃO E CIÊNCIA DA INFORMAÇÃO, 20., 2002, Fortaleza. **Anais...** Fortaleza: UFC, 2002. 1 CD-ROM.

LARA, M. L. G. de. Termos e conceitos da área de comunicação e produção científica. In: POBLACION, Dianh; WITTER, Geraldina Porto; SILVA, José Fernando Modesto da Silva (Orgs.). **Comunicação & Produção Científica**: contexto, indicadores e avaliação. São Paulo: Angellara, 2006.

LARA, M. L. G. de. Linguagem documentária e Terminologia. **Transinformação**, Campinas, v. 16, n. 3, p. 231-240, set./dez. 2004. Disponível em: <<http://revistas.puc-campinas.edu.br/transinfo/>>. Acesso em: 20 ago. 2008..

LARA, M. L. G. de.; TÁLAMO, M. de F. G. M. Uma experiência na interface Linguística Documentária y Terminología. **DataGramZero**. Revista de Ciência da Informação, Rio de Janeiro, v.8, n.5, art. 01, out. 2007. Acesso em 20 maio. 2009.

LARA, M. L. G.; TALAMO, M. F. G. M. Linguística documentária e terminologia: experiência didática na interface das disciplinas. In: **VII Encontro Nacional de Pesquisa em Ciência da Informação**, 2006, Marília. A dimensão epistemológica da Ciência da Informação e suas interfaces teóricas, políticas e institucionais nos processos de produção, acesso e disseminação da informação. Marília: UNESP: ANCIB, 2006.

LARIVIERE, L. *Vers un produit unifié en Terminologie et en Documentation le thesaurus terminologique*. **Meta**, v. 34, n.3, p. 457-467, 1989.

LE COADIC, Y-L. **A ciência da Informação**; tradução de Marial Yeda F. S. Filgueiras Gomes. 2. ed. rev., atual. Brasília: Briquet de Lemos/Livros, 2004.

LIMA, G. A. B. Interfaces entre a ciência da informação e a ciência cognitiva. **Ciência da Informação**, Brasília, v.32, n.1, p.77-87, jan./abr. 2003.

LIMA, V. M. A. **Comunicação e representação documentária**. São Paulo: APB, 1999. (Ensaio APB, 62).

LIMA, V. M. A. **Da classificação do conhecimento científico aos sistemas de recuperação de informação: enunciação de codificação e enunciação de decodificação da informação documentária**. 2004. Tese (Doutorado em Ciências da Comunicação: Ciência da Informação e Documentação) – Universidade de São Paulo, Escola de Comunicações e Artes, São Paulo.

LIMA, V. M. A. *et al.* Estudos para implantação de ferramenta de apoio à gestão de linguagens documentárias: Vocabulário Controlado da USP. **Transinformação**, Campinas, v. 18, n. 1, p. 17-25, jan./abr., 2006.

- LIMA, V. M. A. **Terminologia, comunicação e representação documentária**. 1998. Dissertação (Mestrado em Ciências da Comunicação) – Universidade de São Paulo, Escola de Comunicações e Artes, São Paulo.
- LÓPEZ ALONSO, M. A. Los tesauros conceptuales como herramienta de precisión en los sistemas de organización científica. **Revista Interamericana de Bibliotecología**, Medellín, Colombia, v.22, n.1, p.21-36, ene./jun. 1999.
- LÓPEZ-HUERTAS, M.J. Gestión del conocimiento multidimensional en los sistemas de organización del conocimiento. In: RODRÍGUEZ BRAVO, B.; ALVITE DÍAZ, M.A. **La interdisciplinariedad y la transdisciplinariedad en la organización del conocimiento científico**. Actas del VIII Congreso ISKO-España. León: Universidad de León, 2007, p. 1-26.
- LOURENÇO, A. **A compilação de termos para elaboração de tesauros**: revisão de literatura para conceituação e procedimentos. 1996b. Trabalho de Conclusão de Curso (Graduação em Biblioteconomia) – Universidade de Marília, Marília.
- LOURENÇO, A. **O método de compilação dedutivo para elaboração de tesouro**. Marília: UNESP, 1996a. (Relatório parcial de pesquisa de iniciação científica).
- LYONS, J. **Lingua(gem) e lingüística**: uma introdução. Rio de Janeiro: LTC, 1987.
- LYONS, J. (Org.). **Novos horizontes em lingüística**. São Paulo: Cultrix: Ed. da Universidade de São Paulo, 1976.
- LYONS, J. **Semântica**. Lisboa: Presença: Martins Fontes, 1977. v.1
- MACEDO, N. D., DIAS, M. M. K. Subsídios para a caracterização da Biblioteca Universitária. **Revista Brasileira de Biblioteconomia e Documentação**, São Paulo, v.25, n.3/4, p.40-48, jul./dez. 1992.
- MACIEL, A. M. B. TERMISUL e terminótica. In: KRIEGER, M. da G.; MACIEL, A.M.B (Org.). **Temas de Terminologia**. Porto Alegre: Ed. Universidade/UFRGS; São Paulo: Humanitas/USP, 2001. p. 344-363.
- MANUEL AGUADO, P. Los sistemas expertos y la recuperación documental: ejemplos de aplicación. **Scire**: representación y organización del conocimiento. Universidad Zaragoza, v.1, n.2, p. 21-32, jul./dic. 1995.
- MARDER, J.V. The development of the British Education Thesaurus: a personal account. **Education Libraries Journal**, v.32, n.1, p.2-8, 1989.
- MARI, H. Dos fundamentos da significação à produção do sentido. **Perspectivas em Ciência da Informação**, Belo Horizonte, v.1, n.1, p.93-109, jan./jun. 1996.
- MARKEY, K. **The process of subject searching in the library catalog**: final report of the SARP: Subject Access Research Project. Dublin: OCLC, 1983.
- MARTELETO, R. M. Lugares dos signos e contextos de informação. In: SEMINÁRIO DE ESTUDOS DE INFORMAÇÃO, 1., 1997, Niterói. **Anais...** Niterói: EDUFF, 1997. p.21-24.
- MARTIN-BARBERO, J. **Dos meios às mediações**: comunicação, cultura e hegemonia. 5. Ed. Rio de Janeiro: UFRJ, 2008.



MARTINEZ, M. L. C. **Classificação de recursos trabalhistas**: uma análise crítica da Classificação Decimal de Direito com vistas a uma proposta metodológica de extensão. 2005. 144 f. Dissertação (Mestrado em Ciência da Informação) – Programa de Pós-Graduação em Ciência da Informação, Universidade Estadual Paulista, Marília, 2005.

MARTYN, J.; LANCASTER, F. W. **Investigative methods in library and information science**: an introduction. Arlington: Information resources Press, 1981. 251 p.

MATURANA, H. R.; VARELA, F. J. **A árvore do conhecimento**: as bases biológicas da compreensão humana; tradução de Humberto Mariotti e Lia Diskin. São Paulo: Palas Athena, 2001.

MAYA MONTALVO, L. A. **Proposta de procedimentos metodológicos para elaboração de tesouros-facetados**. 1986. 245f. Tese (Doutorado em Ciências da Comunicação) – ECA/USP, São Paulo.

MEADOW, C. T. **Text information retrieval systems**. San Diego: Academic Press, 1992.

MEDEIROS, M. B. Terminologia brasileira em Ciência da Informação: uma análise. In: SIMPOSIO LATINOAMERICANO DE TERMINOLOGÍA. 1., Caracas, 1988. **Actas...** Disponível em: <http://www.ritem.net/actes/1simposio/medeiros.htm>

MELO, L. B. Avaliação de sistemas de recuperação de informação: breve retrospectiva dos principais projectos. In: **Cadernos BAD**, Lisboa, n. 2, 1994, p. 29-46

MELLO, L. F. **Para um glossário bilíngüe (português-inglês) de termos da área de secretariado**. 2002. 233f. Dissertação (Mestrado em Letras) – CLCH/UUEL, Londrina.

MILLER, G. The psychology of communication. Harmondsworth: Penguin Books, 1966. Apud McGURRY, K. **O contexto dinâmico da informação**: uma análise introdutória. Brasília: Briquet de Lemos/Livros, 1999.

MILSTEAD, J. L. **ASIS Thesaurus of Information Science and Librarianship**. 2. ed. New Jersey: Information today, 1998.

MIRANDA, A. L. C. Globalización y sistemas de información: nuevos paradigmas y nuevos desafios. **Ciência da Informação**, Brasília, v.23, n.3, p.308-317, 1994.

MIRANDA, L.M.C.D. *et al.* Elaboração de tesouros utilizando-se o programa de elaboração de tesouros em microcomputador (TECER). **Rev. Bibliotec. de Brasília**, v.18, n.2, p.185-192, 1990.

MIRANDA, A.; FORCELLINI, F. As tecnologias de informação e comunicação e a (ex.) inclusão das bibliotecas universitárias. In: VERGUEIRO, W.; MIRANDA, A. (Orgs.). **Administração de unidades de informação**. Rio Grande : Ed. da FURG, 2007.

MOCHÓN BEZARES, G.; SORLI ROJO, Á. **Tesouro de Biblioteconomía y Documentación**. Madrid: Consejo Superior de Investigaciones Científicas, 2003.

MONTGOMERY, C. A. Linguistics and information science. **Journal of the American Society for Information Science**, v.23, p.195-219, May/Jun. 1972.

MORAES, J. B. E.; GUARIDO, M. D. M. Análise de categorias conceituais da CDD intermediada pela terminologia. In: García Marco, F. J. (Ed.). **Avances y perspectivas en**

**sistemas de información y documentación** - IBERSID, 2008. Zaragoza: Universidad de Zaragoza, 2008. p. 199-209.

MOREIRA, M. P.; MOURA, M. A. Construindo tesouros a partir de tesouros existentes: a experiência do TCI - Tesouro em Ciência da Informação. **DataGramZero** - Revista de Ciência da Informação, v.7, n.4, ago. 2006. Disponível em: <[www.datagramazero.org.br](http://www.datagramazero.org.br)>. Acesso em: 18 ago. 2008.

MOREIRA, M. P. **TCI: Tesouro em Ciência da Informação**. Belo Horizonte: UFMG: PUCMINAS, 2006. Disponível em: <<http://www.inf.pucminas.br/ci/tci/>>. Acesso em: 18 ago. 2008.

MORIN, E. **O método 3: o conhecimento do conhecimento**. Porto Alegre: Sulina, 1999.

MOTTA, D. F. **Método relacional como nova abordagem para a construção de tesouros**. Rio de Janeiro: SENAI, 1987. (Coleção Albano Franco; 12).

MOURA, M. A. et al. Linguagens de indexação em contextos cinematográficos: a experiência de elaboração do tesouro eletrônico do cinema brasileiro. **Perspectivas em Ciência da Informação**, Belo Horizonte, v.10, n.1, p. 54-69, jan./jun. 2005. Disponível em : <<http://www.eci.ufmg.br/pcionline/>>. Acesso em: 12 abr. 2008.

MULTYSYSTEMS. **MultiTes**. Disponível em: <<http://www.multites.com>> Acesso em: 18 out. 2006.

NAKAYAMA, H. **Terminologia aplicada à ciência da informação: da produção de vocabulário técnico-científico bilíngüe (japonês-português), na área do ensino da língua japonesa**. 1996. Tese (Doutorado em Lingüística) – Universidade de São Paulo, Faculdade de Filosofia, Letras e Ciências Humanas, São Paulo, 1996..

NARDI, M. I. A. **A metáfora e a prática de leitura como evento social: instrumentos do pensar a Biblioteconomia do futuro**. 1999. Tese (Doutorado em Lingüística Aplicada e Estudos da Linguagem) – Programa de Lingüística Aplicada e Estudos da Linguagem, Pontifícia Universidade Católica, São Paulo.

NARDI, M. I. A. **As expressões metafóricas na compreensão de texto escrito em língua estrangeira**. 1993. Dissertação (Mestrado em Lingüística Aplicada e Estudos da Linguagem) – Programa de Lingüística Aplicada e Estudos da Linguagem, Pontifícia Universidade Católica, São Paulo.

NAUMIS PEÑA, C. **Los tesouros documentales y su aplicación en la información impresa, digital y multimedia**. México: Universidad Nacional Autónoma de Mexico; Buenos Aires: Alfagrama, 2007. 288 p.

NAUMIS PEÑA, C. Análisis de la confluencia entre término y descriptor en la elaboración de tesouros. **Investigación Bibliotecológica: archivonomía, bibliotecología e información**. v.14, n. 29 jul./dic. 2000.

NAUMIS PEÑA, C. Dinámica en la construcción de lenguajes documentales y el acceso a la información. In: **La información en el inicio de la era electrónica**. México, 1998. (Organización del conocimiento y sistemas de información).

NOVELLINO, M. S. F. Instrumentos e metodologias de representação da informação. **Inf.**, Londrina, v.1, n.2, p.37-45, jul./dez. 1996.

- NOVELLINO, M. S. F. Organização do conhecimento: as contribuições da ciência da informação e da lingüística no contexto de uma ação comunicativa. In: SEMINÁRIO DE ESTUDOS DE INFORMAÇÃO, 1., 1997, Niterói. **Anais...** Niterói: EDUFF, 1997. p. 105-110.
- NUNES, M. C. A organização do conhecimento com fins documentários: aspectos teóricos da indexação. **InfoCiência**, São Luís, v.1, p.43-52, 1998.
- O'SULLIVAN, T. et al. Key concepts in communication. London: Merthuen, 1983. Apud KOBASHI, N. Y. **A elaboração de informações documentárias**: em busca de uma metodologia. 1994. 195f. Tese (Doutorado em Ciências da Comunicação) –Universidade de São Paulo, São Paulo.
- OLIVEIRA, Marlene de. Origens e evolução da Ciência da Informação. In: \_\_\_\_\_. (Coord.). **Ciência da Informação e Biblioteconomia**: novos conteúdos e espaços de atuação. Belo Horizonte: Editora UFMG, 2005. cap. 1. p. 9-28. (Coleção Didática).
- OTLET, P. **Traité de documentation**: le livre sur le livre, théorie et pratique. Bruxelles: Editiones Mundaneum, 1934.
- OTLET, P. **Documentos e documentação**: discurso pronunciado no Congresso de Documentação Universal, Paris, 1937. Rio de Janeiro: Imprensa Nacional, [200?]. Disponível em: <[www.conexaorio.com/bit/otlet](http://www.conexaorio.com/bit/otlet)>. Acesso em 14 out. 2009.
- PAIS, C. T. Monografias de semiótica e linguística. In: GREIMAS, A. J. **Semiótica do discurso científico. Da modalidade**: prefácio e tradução de Cidmar Teodoro Pais. São Paulo: Difel: SBPL, 1976. (Monografias de Semiótica e Linguística, n.1).
- PAVEL, S., NOLET, D. **Manual de terminologia**. Canadá: TERMIUM, 2002.
- PEIRCE, C. S. **Semiótica**. São Paulo: Perspectiva, 1977.
- PICHT, H. La terminologia como factor de desarrollo. **Revista Interamericana de Bibliotecologia**, Medellín, Colombia, v.19, n.1, p.7-30, ene./jun. 1996.
- PIEIDADE, M. R. **Introdução à teoria da classificação**. 2. ed. Rio de Janeiro: Interciência, 1983.
- PINHO, F. A. **Aspectos éticos em representação do conhecimento**: em busca do diálogo entre Antonio García Gutiérrez, Michèle Hudon e Clare Beghtol. Marília, 2006. Dissertação (Mestrado) – Universidade Estadual Paulista, Faculdade de Filosofia e Ciências, Marília, 2006.
- RAIMUNDO, E. M.; CERVANTES, B. M. N.; PEREIRA, L. M. R. Análise terminológica: termos e contextos. In.: V Selisigno e VI Simpósio de Leitura da UEL, 2006, Londrina. **Anais do V Selisigno e VI Simpósio de Leitura da UEL**. Londrina: UEL, 2006.
- RAYWARD, W. B. The origins of information science and the International Institute of Bibliography/International Federation for Information and Documentation (FID). **Journal of American Society for Informatin Science**, v. 48, n.4, p. 239-300, 1997.
- REGO, T. C. **Vygotsky**: uma perspectiva histórico-cultural da educação. 19. ed. Rio de Janeiro: Vozes, 2008.
- ROBREDO, J. **Da ciência da informação revisitada**: aos sistemas humanos de informação. Brasília: Thesaurus: SSRR Informações, 2003.

ROBREDO, J. **Documentação de hoje e de amanhã**: uma abordagem revisitada e contemporânea da Ciência da Informação e de suas aplicações biblioteconômicas, documentárias, arquivísticas e museológicas. 4. ed. rev. e ampl. Brasília: Edição de autor, 2005.

RODRÍGUEZ BRAVO, B. (Dir.). **Tesouro de Ciencias de la Documentación**. Autores, M<sup>a</sup> Luisa Alvite Díez, Ángela Díez Díez, Josefa Gallego Lorenzo... [et al.]. León: Universidad de León, 2004. Disponível em: < <http://eprints.rclis.org/archive/00003067/> > Acesso em: 17 nov. 2008.

RONDEAU, G. **Introduction à la terminologie**. 2.ed. Québec, Canadá: Gaëtan Morin, 1984. 238 p.

RUBI, M. P. **Política de indexação para construção de catálogos coletivos em bibliotecas universitárias**. 2008. 169 f. Tese (Doutorado em Ciência da Informação) – Programa de Pós-Graduação em Ciência da Informação, Universidade Estadual Paulista, Marília, 2008.

SAGER, J-C. A practical course in terminology processing. Amsterdam: John Benjamins, 1990. Apud MELLO, L. F. de. **Para um glossário bilíngüe (português-inglês) de termos da área de secretariado**. 2002. Dissertação (Mestrado em Letras) – Universidade Estadual de Londrina, Londrina.

SAGER, J-C. Prólogo: la Terminología, puente entre varios mundos. In: CABRÉ, M. T. **La Terminología**: teoría, metodología, aplicaciones. Traducción castellana de Carles Tebé. Barcelona: Ed. Antártica/Empúres, 1993. p. 11-17.

SANTOS BORBUJO, A. Terminología y socioterminología. In: REAL, E., JIMÉNEZ, D., PUJANTE, D.; CORTIJO, A. (Eds.), **Écrire, traduire et représenter la Fête**. Valencia: Universitat de València, 2001, p. 657-664,

SANTOS, L. C. dos; CERVANTES, B. M. N.; RAIMUNDO, E. M.; COSTA, G. C. da; MELLO, L. F. de. Pesquisa Terminológica Temática Trilíngüe (português-inglês-francês) de Termos do domínio da Biblioteconomia/Ciências da Informação. In: SIMPOSIO IBEROAMERICANO DE TERMINOLOGÍA (RITERM), 10. **Terminología, Conocimientos, Sociedad y Poder**. Montevideo, 2006.

SARACEVIC, T. Interdisciplinarity nature of Information Science. **Ciência da Informação**, Brasília, v.24, n.1, p.36-41, 1996.

SHANNON, C.; WEAVER, W. **A teoria matemática da comunicação**. São Paulo: DIFEL, 1975. Apud McGURRY, K. **O contexto dinâmico da informação**: uma análise introdutória. Brasília: Briquet de Lemos/Livros, 1999.

SILVA, J. M. P. **Linguagem documentária para mediação de recuperação de informações empresariais em sistemas de Data Warehouse**. 2002. Dissertação (Mestrado em Ciência da Informação) – Universidade Estadual Paulista (Campus de Marília) – Faculdade de Filosofia e Ciências, Marília.

SMIT, J. W. (Comp.). **Análise documentária**: a análise da síntese. Brasília: IBICT, 1987.

SMIT, J. W. Arquivologia/Biblioteconomia: interfaces das Ciências da Informação. **Inf. Inf.**, Londrina, v.8, n.1, jun./dez. 2003.

SMIT, J. W. A determinação do campo científico da Ciência da Informação: uma abordagem terminológica. **DataGramZero** – Revista de Ciência da Informação. Rio de Janeiro, v.5, n.1, fev., art. 03., 2004.

SMIT, J. W. A determinação do campo científico da Ciência da Informação: uma abordagem terminológica. In: Encontro Nacional de Pesquisa em Ciência da Informação, 5., 10-14 nov. 2003. **Anais...** Belo Horizonte, Escola de Ciência da Informação da UFMG, 2003.

SMIT, J. W. A pesquisa na área de Ciência da Informação. **Transinformação**, v.14, n.1, Jan./jun, 2002.

SMIT, J. W. Documentação e lingüística: inter-relação e campos de pesquisa. **Revista Brasileira de Biblioteconomia e Documentação**, São Paulo, v.11, n.1/2, p. 33-42, 1978.

SMIT, J. W. O documento audiovisual ou a proximidade entre as 3 Marias. **Revista Brasileira de Biblioteconomia e Documentação**, São Paulo, v.26, n.1/2, p. 81-85, 1993.

SMIT, J. W. ; BARRETO, A. de A. Ciência da Informação; base conceitual para a formação do profissional. In: VALENTIN, M. L. (Org.) **Formação do Profissional da Informação**. São Paulo: Pólis, 2002, cap. 1, p. 9-23

SMIT, J.; TÁLAMO, M. F. G. M.; KOBASHI, N. Y. A determinação do campo científico da Ciência da Informação: uma abordagem terminológica. **DataGramZero** – Revista de Ciência da Informação. Rio de Janeiro, v.5, n.1, fev., artigo 03. 2004.

SOERGEL, D. **Indexing languages and thesauri: construction and maintenance**. Los Angeles, Cal., Melville Publ., 1974. 632 p.

STRAIOTO, A. C. **A Análise em facetas como dimensão teórica e prática na organização do conhecimento**. 2001. 163f. Dissertação (Mestrado em Ciência da Informação) – Programa de Pós-Graduação em Ciência da Informação, Universidade Estadual Paulista, Marília.

SVENONIUS, E. **The intellectual foundation of information organization**. Cambridge: The MIT Press, 2000.

SVENONIOUS, E. Unanswered questions in the design of controlled vocabularies. **Journal of American Society for information Science**, v. 37, n. 5, p. 331-340, 1986.

TÁLAMO, M. de F. G. M. et al. Instrumentos de controle terminológico: limites e funções. In: SIMPÓSIO LATINO-AMERICANO DE TERMINOLOGIA, 2.; ENCONTRO BRASILEIRO DE TERMINOLOGIA TÉCNICO-CIENTÍFICA, 1., 1990, Brasília. **Anais**. Brasília: IBICT, 1990.

TÁLAMO, M. de F. G. M. Informação: organização e comunicação. In: SEMINÁRIO DE ESTUDOS DE INFORMAÇÃO, 1., 1997, Niterói. **Anais**. Niterói: EDUFF, 1997a. p.11-14.

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

TÁLAMO, M. de F. G. M. Curso de atualização: elaboração e uso do tesouro. In: \_\_\_\_\_. **Lingüística e análise Documentária**. [S.l.: s.n.], 1997c. (Apostila para uso didático).

TÁLAMO, M. de F. G. M. Terminologia e documentação. **TradTerm**: Revista do Centro Interdepartamental de Tradução e Terminologia, São Paulo, n.7, p.141-152, 2001.

TÁLAMO, M. de F. G. M. ; LARA, M. L. G. de. O campo da Lingüística Documentária. **Transinformação**, v. 18, p. 203-211. 2007. Disponível em: <<http://revistas.puc-campinas.edu.br/transinfo/viewarticle.php?id=182>>. Acesso em: 20 nov. 2008.

TÁLAMO, M. de F. G. M. ; LARA, M. L. G. de; KOBASHI, N. Y. Contribuição da terminologia para a elaboração de tesouros. **Ciência da Informação**, Brasília, v.21, n.3, p.197-200, set./dez. 1992.

TÁLAMO, M. de F. G. M; LENZI, L. A. F. Terminologia e documentação: a relação solidária das organizações do conhecimento e da informação no domínio da inovação tecnológica. **DataGramaZero**. Revista de Ciência da Informação, Rio de Janeiro, v.7, n.4, art. 03, ago. 2006.

TÁLAMO, M. de F. G. M. et al. Otlet, o criador de estruturas informacionais. In: CONGRESSO BRASILEIRO DE BIBLIOTECONOMIA E DOCUMENTAÇÃO E CIÊNCIA DA INFORMAÇÃO, 20., 2002, Fortaleza. **Anais**. Fortaleza: UFC, 2002. 1 CD-ROM.

TELIN, M. F.; ALMEIDA, G. M.; ALUISIO, S. M. **Extração manual e automática de terminologia**: comparando abordagens e critérios. [2003] Disponível em: [http://www.nilc.icmc.usp.br/til2003/oral/Teline\\_Almeida\\_Aluisio\\_37.pdf](http://www.nilc.icmc.usp.br/til2003/oral/Teline_Almeida_Aluisio_37.pdf)

TEMMERMANN, R. Teoria Sociocognitiva da Terminologia. KRIEGER, M. G.; ARAÚJO, L. (Orgs.) A Terminologia em Foco. **Cadernos de Tradução**, Instituto de Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre, v. 17, out-dez. 2004.

UNESCO. **Diretrizes para o estabelecimento e desenvolvimento de tesouros monolíngües**. [2ª ed. rev. por Derek Austin e Peter Dale]; Trad. de Bianca Amaro de Melo; rev. de Lígia Maria Café de Miranda. Brasília: IBICT; SENAI, 1993.

UNIVERSIDADE ESTADUAL PAULISTA. PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO. **Linhas de Pesquisa do PPGCI/UNESP**. Disponível em: <<http://www.marília.unesp.br/index.php?CodigoMenu=363&CodigoOpcao=366>> Acesso em: 26 dez. 2008.

VALENTIM, M. L. P. Construção de conhecimento científico. In: \_\_\_\_\_. (Org.). **Métodos qualitativos de pesquisa em ciência da informação**. São Paulo: Polis, 2005. (Cap.1).

VALENTIM, M. L. P. Processo de Inteligência Competitiva Organizacional In: \_\_\_\_\_. **Informação, conhecimento e inteligência organizacional**. 2. ed. Marília: FUNDEPE, 2007. (Cap. 1).

VALENTIM, M. L. P. Terminologia em Inteligência Competitiva Organizacional In: \_\_\_\_\_. **Informação, conhecimento e inteligência organizacional**. 2. ed. Marília: FUNDEPE, 2007. (Cap. 12).

VALENTIM, M. L. P. Terminologia em Inteligência Competitiva Organizacional In: \_\_\_\_\_. **Informação, conhecimento e inteligência organizacional**. Marília: FUNDEPE, 2006. (Cap. 12).

VAN DER LAAN, R. H. **Tesouro e Terminologia**: uma inter-relação lógica. Porto Alegre, 2002. 185 f. Tese (Doutorado em Letras) – Universidade Federal do Rio Grande do Sul. 2002.

VAN DJIK, T. A. **La ciência del texto**. Barcelona: Paidós, 1989.

VAN SLYPE, G. **Los lenguajes de indización**: concepción, construcción y utilización en los sistemas documentales. Traducción del francés: Pedro Hípola, Félix de Moya. Madrid: Fundación Germán Sanchez Ruipérez, 1991.

VICKERY, B. C. **Classificação e indexação nas ciências**. Tradução Maria Christina Girão Pirolla. Rio de Janeiro: BNG/Brasilart, 1980.

VICKERY, B. C. Ontologies. In: **Journal of Information Science**. v. 23, n. 4, p. 277-286, 1997.

VIZCAYA ALONSO, D. **Lenguajes documentários**. Rosario: Nuevo Parahadigma, 1997.

VOGEL, M. J. **A noção de estrutura lingüística e de processo de estruturação e sua influência no conceito e na elaboração de linguagens documentárias**. São Paulo, 2007. Dissertação (Mestrado em Ciência da Informação) – Universidade de São Paulo, Escola de Comunicações e Artes, São Paulo, 2007.

WANDERLEY, M. A. Linguagem documentária: acesso à informação, aspectos do problema. **Ciência da Informação**. Rio de Janeiro, v.2, n.2, p. 175-217, 1973.

WEBCHOIR. The Thesaurus Construction System (TCS). Disponível em: <<http://www.webchoir.com>>. Acesso em: 28 set. 2008.

WERSIG, G. & NEVELING, U. (Comp.). **Terminology of documentation**: a selection of 1200 basic terms published in English, French, German, Russian and Spanish. Paris: UNESCO, 1976. 274 p.

WÜSTER, E. **Introducción a la teoría general de la terminología y a la lexicografía terminológica**. Barcelona: Institut Universitari de Lingüística Aplicada, 1998 ; 2003.

WÜSTER, E. L'Étude scientifique générale de la Terminologie, zone frontalière entre la Linguistique, la Logique, l'Ontologie, l'Informatique et les Sciences des Choses. In: RONDEAU, G., FELBER, F. (Org.). **Textes choisis de Terminologie**. I. Fondements théoriques de la terminologie. Québec : GIRSTERM, 1981. p.106.

ZAMBEL, M. M. **Glossário de termos usuais em Biblioteconomia**. São Carlos, Fundação Theodoretto Souto, 1978.

ZEMGULIS, A. M. A normalização terminológica e o desenvolvimento tecnológico. In: SIMPÓSIO LATINO-AMERICANO DE TERMINOLOGIA, 2., 1990, Brasília. **Anais**. Brasília: CNPq; IBICT, 1992. p. 419-424.

ZINS, C. Conceptions of Information Science. **Journal of the American Society for information Science and technology**, v.58, n. 3, p.335-350, 2007.

**APÊNDICES**



## Apêndice A

ETAPAS DE CONSTRUÇÃO DE UM TESAURO DE ACORDO COM AS DIRETRIZES			
CATEGORIAS TEMÁTICAS	DIRETRIZES... IBICT (1984)	DIRETRIZES... UNESCO (1993)	DIRETRIZES... ANSI/NISO (2005)
1	<p><b>Trabalho preliminar:</b></p> <p>a) “Comunicação de intenção – quando um serviço decide a preparar um tesauro novo, essa intenção deve ser comunicada em periódicos apropriados (por ex., num periódico de documentação, de classificação)”;</p> <p>b) Não duplicação de trabalho – a verificação da existência de tesauro que cobre a mesma área de conhecimento ou parte dele, pode ser útil como ponto de partida. Todavia, orientam que “cada tesauro deve refletir as necessidades de uma comunidade específica de usuários” (IBICT, 1984, p. 52).</p>	<p><b>Métodos de compilação:</b></p> <p><b>a) Método dedutivo</b> - os termos são extraídos de documentos durante uma etapa preliminar de indexação. Não é feito controle do vocabulário, nem mesmo para determinar as relações entre os termos. A não ser que um número suficiente de termos tenha sido coletado. Nesse caso, todos os termos são revistos por um grupo de especialistas no assunto. Primeiro, devem ser identificados os termos que representam categorias genéricas, e os termos restantes devem ser relacionados nestas categorias conforme as suas relações lógicas. Deste modo, as categorias tendem a ser estabelecidas sobre uma base de termo genérico à específico. O controle do vocabulário deve ser efetuado no estágio em que as categorias são estabelecidas, observando os princípios descritos nas seções de 5 a 7, dessa Norma. A seção 5 trata do Controle de Vocabulário. As seções 6 e 7, abordam dos Termos de Indexação e Termos Compostos, respectivamente.</p> <p><b>b) Método indutivo</b> - novos termos são admitidos no tesauro tão logo apareçam nos documentos. O controle do vocabulário é feito desde a etapa inicial, e cada termo, à medida que é admitido, é designado como membro de uma ou mais categorias genéricas constituídas sobre uma base <i>ad hoc</i>. O tesauro é, portanto, organizado sobre uma base de termo específico a genérico. A elaboração do tesauro é, desde o princípio, uma operação contínua e, apesar de ser necessária a assistência de técnicos especialistas, eles não precisam fazer parte de um grupo editorial formal ou comitê;</p> <p><b>c) Combinação de Métodos</b> - “na prática, é</p>	<p><b>Evitar duplicação de vocabulários</b></p> <p>Antes de iniciar a construção de um vocabulário controlado (tesauro), deve-se averiguar se um vocabulário controlado existente cobre o mesmo domínio ou um domínio de conhecimento coincidente.</p>

		<p>possível empregar-se tanto o método dedutivo como indutivo em uma ou outra etapa da elaboração do tesouro". Por exemplo, um grupo editorial composto por indexadores e especialistas pode primeiramente estabelecer as categorias dos termos indutivamente e depois examiná-las pelo ponto de vista dedutivo. Ambas as técnicas são essencialmente empíricas e, desde o início, deve-se ter em mente que algumas decisões tomadas durante a etapa de compilação possam ter que ser revisadas à medida que se adquire maior experiência. A adição de novos termos a uma linguagem de indexação freqüentemente esclarece as decisões tomadas anteriormente quanto aos termos e suas inter-relações. Os compiladores devem revisar freqüentemente os termos e as hierarquias para assegurar que as relações entre os termos, a fatoração de termos compostos etc., seguem os princípios gerais que têm sido aplicados. Recomenda-se particularmente que se busque o auxílio de especialista na escolha dos termos sinônimos e formas variantes dos termos.</p>	
2	<p><b>Métodos de compilação:</b></p> <p><b>a) método dedutivo:</b> os termos são extraídos da literatura durante o estágio preliminar de indexação; entretanto, até que um número suficiente de termos tenha sido coletado, todos os termos são considerados por uma comissão, formada de preferência por indexadores e especialistas, que decidem primeiro quais termos representam as classes mais amplas, fixando os demais termos nessas classes e relacionando-os logicamente com base em suas conceituações;</p> <p><b>b) método indutivo:</b> novos termos são admitidos no tesouro logo que são encontrados na literatura, sendo cada termo designado como membro de uma ou mais classes, estabelecidas em bases <i>ad hoc</i> durante o ato de indexação. A elaboração do tesouro é considerada, desde o início, como uma operação contínua e, embora se busque assistência de especialistas sempre que for preciso, tais profissionais não são membros de uma comissão formal. Na oportunidade, ainda que tenham sido evidenciadas as duas abordagens descritas acima, nessa publicação aparece mais uma abordagem de</p>	<p><b>Registro dos termos</b></p> <p>Registro individual em ficha.</p>	<p><b>Determinar a estrutura e formatos de apresentação:</b></p> <p>A estrutura do vocabulário controlado/tesouro (estrutura genérica e específica, apresentação hierárquica e/ou gráfica) e formato de apresentação (impresso, online, ou disponível na web) deveriam ser decididos antes que os termos fossem coletados e considerados como candidatos para inclusão.</p>

	<p>compilação, denominada:</p> <p><b>c) método relacional*</b> - em notas de rodapé, consta a seguinte explicação: “*Dahlberg aponta um terceiro tipo – a abordagem relacional – que parte do aspecto formal, categorial” (IBICT, 1984, p. 52).</p> <p><b>d) Combinação de métodos.</b> Durante a sequência de etapas de construção do tesouro, empregam-se, de modo igual, os métodos dedutivo e indutivo. Por exemplo: as classes de termos que tenham sido estabelecidas de início de forma indutiva podem ser revistas mais tarde por uma comissão de indexadores e especialistas da área do tesouro; ambas as técnicas são empíricas. Nesse sentido, decisões tomadas nos primeiros estágios da compilação podem ser revistas, à medida que os indexadores ganham mais experiência. A inclusão de novos termos, a conferência dos relacionamentos entre os termos, a fatoração, as notas de aplicação, devem seguir uma linha lógica e consistente. A assessoria de especialistas na escolha entre termos sinônimos e outras formas variantes é recomendada.</p>		
3	<p><b>Registro de termos</b></p> <p>Registro individual em ficha.</p>	<p><b>Verificação dos termos</b></p> <p>Termos candidatos à inclusão no tesouro, consultar as fontes autorizadas como: dicionários e enciclopédias especializadas, tesouros e esquema de classificação, índice de periódicos especializados, tratados e manuais atualizados, e especialistas no assunto, principalmente aqueles que possuem algum conhecimento de indexação ou documentação.</p>	<p><b>Métodos de construção/compilação</b></p> <p><b>a) Abordagem de Comitê</b> - especialistas no domínio de assunto do vocabulário controlado fazem uma lista de termos chave na área e indicam as relações entre eles, com assistência de especialistas em criação de vocabulário controlado. Listas de termos podem ser retiradas de várias fontes ou submetidas a vários usuários/comunidades. Dois métodos principais para criar vocabulários controlados por Comitê têm sido utilizados: 1) <b>Top down</b> (Descendente) – os termos mais genéricos são identificados primeiro e então termos mais específicos são selecionados para atingir o nível desejado de especificidade. As estruturas hierárquicas e relações necessárias são estabelecidas à medida que o trabalho evolui; 2) <b>Bottom up</b> (Ascendente) – este caso ocorre frequentemente quando listas de termos forem derivadas de um <i>corpus</i> de objetos de conteúdo e devem então ser incorporadas em um vocabulário controlado. Como no caso acima, as estruturas e relações hierárquicas necessárias são criadas à medida que o trabalho avança, mas inicia-se com os termos que têm uma abrangência mais específica e segue para os termos mais genéricos. Se um novo vocabulário controlado/tesouro será construído, a abordagem “descendente” é preferida. No caso de um</p>

		<p>vocabulário controlado/tesauro já publicado, a abordagem “ascendente” é mais usada para o acréscimo de novos termos para cobrir novos conceitos;</p> <p><b>b) Abordagem Empírica</b> - a abordagem empírica evidencia dois métodos básicos: <b>i) Método Dedutivo</b> – termos são extraídos de objetos de conteúdo (por humanos ou computadores), opcionalmente durante um estágio preliminar de indexação, mas nenhuma tentativa é feita para controlar o vocabulário, nem para determinar relações entre termos, até que um número suficiente de termos tenha sido coletado. Todos os termos são revisados por um grupo de especialistas, preferencialmente, composto de especialistas em informação e no assunto tratado. Eles deveriam primeiro identificar termos que representam as classes mais amplas, e então alocar os termos remanescentes para aquelas classes segundo suas relações lógicas, a fim de que as hierarquias tendam a ser estabelecidas sobre uma base da mais genérica para a mais específica. O controle de vocabulário deveria ser desenvolvido no estágio onde hierarquias e outras relações são estabelecidas, seguindo os princípios descritos nas seções que tratam de Termos compostos; e de Relacionamentos entre os termos; <b>ii) Método Indutivo</b> - novos termos são selecionados para potencial inclusão no vocabulário controlado/tesauro como eles forem encontrados em objetos de conteúdo. O controle de vocabulário é aplicado desde o início. Se o vocabulário em construção tiver algum tipo de arranjo hierárquico, cada termo admitido é designado como um membro de uma ou mais classes genéricas que são construídas sobre uma base <i>ad hoc</i> em um estágio anterior. O vocabulário controlado é, portanto, estabelecido sobre uma base de termo do mais específico para o mais genérico. A construção de vocabulário controlado é considerada do início como uma operação contínua. Assistência de especialistas no assunto é altamente recomendada; estes especialistas podem servir como membros de uma comissão ou comitê editorial/formal;</p> <p><b>c) Combinação de Métodos</b> - na prática, é possível empregar mais de uma dessas abordagens em um estágio ou outro durante a construção de um vocabulário controlado. Por exemplo, hierarquias e outras relações entre termos que foram primeiro estabelecidos indutivamente poderiam mais tarde ser examinadas a partir de um ponto de vista dedutivo. Ambas as técnicas são essencialmente empíricas, e deveria ser aceito desde o início que algumas decisões com relação aos termos e suas inter-relações, que foram feitas durante os primeiros estágios de compilação, possam ser revisadas como uma experiência posterior. Os compiladores deveriam verificar os</p>
--	--	--

			<p>termos e hierarquias frequentemente para assegurar uma aplicação consistente de princípios em tais procedimentos como: o estabelecimento de inter-relações entre termos e a divisão de termos compostos;</p> <p><b>d) Assistência da Máquina</b> – assume-se de acordo com essa Norma que a construção de vocabulário controlado/tesauro requer decisões intelectuais. A assistência da máquina pode ser empregada, no entanto, para tarefas de identificação de termo tais como as seguintes: i) Identificação de termos candidatos – termos candidatos podem ser identificados automaticamente a partir de texto legível por máquina, (por exemplo: títulos e/ou resumos). O número de termos potenciais deveria ser reduzido pelo uso de palavras vazias ou uma lista vazia (palavras funcionais – artigos, conjunções, e preposições tais como “um/uma”, “e”, e “para” – além de outras palavras consideradas sem valor para a recuperação). Remover as palavras vazias da consideração é geralmente o primeiro passo para a indexação auxiliada pela máquina. Os termos remanescentes deveriam ser comparados com aqueles já registrados no vocabulário controlado. Termos que não foram registrados poderiam ser considerados como candidatos à inclusão; ii) Registro da frequência de designação (indicação) de termo – Em sistemas de indexação computadorizados, a frequência com a qual um termo ter sido usado em indexação, pode ser registrada automaticamente. Termos com resultados altos ou baixos podem ser considerados candidatos à modificação ou exclusão; iii) Registro de termos a partir das consultas do usuário – Termos encontrados em consultas de usuário que não combinam com um ou mais termos de entrada também podem ser considerados para inclusão, especialmente quando um dado termo ocorre em múltiplas consultas. Para assegurar privacidade, os usuários não deveriam ser identificados quando se rastreia e coleta consultas de usuário. Termos que ocorrem em linguagem natural e são formas comuns de descrever um conceito são incluídos no vocabulário controlado/tesauro como termos de entrada.</p>
<b>4</b>	<p><b>Verificação do termo</b></p> <p>Termos candidatos à inclusão no tesauro, consultar as fontes autorizadas como: dicionários e enciclopédias especializadas, tesouros e esquema de classificação, índice de periódicos especializados, tratados e manuais atualizados, e especialistas no assunto.</p>	<p><b>Especificidade</b></p> <p>Um serviço de indexação que trabalhe com documentos relacionados com mais de um campo de conhecimento, pode ser necessário desenvolver vários tesouros especializados relacionados e compatíveis a um tesauro geral [...] e seja produzido pelo mesmo serviço.</p>	<p><b>Registro de termos</b></p> <p>Registro individual em ficha.</p>
<b>5</b>	<p><b>Especificidade</b></p> <p>Se um serviço de informação compreende mais de</p>	<p><b>Inclusão e exclusão de termos</b></p> <p>Os termos e suas relações devem refletir a maneira</p>	<p><b>Verificação de termo</b></p> <p>Antes de um termo ser admitido em um vocabulário</p>

	uma área de conhecimento, pode tornar-se necessário desenvolver tesouros específicos para as áreas, ligados a um tesouro geral.	pela qual os especialistas da matéria utilizam a linguagem no campo coberto pelo tesouro.	controlado/ tesouro, ele deveria ser validado de acordo com regras propostas
<b>6</b>	<p><b>Admissão e exclusão de termos</b></p> <p>O tesouro deve ser submetido a uma revisão técnica após determinado período de uso e, a partir daí, a intervalos regulares, levando em conta mudanças no uso de termos dentro do campo de cobertura do tesouro.</p>	<p><b>Uso de equipamentos de processamento automático de dados</b></p> <p>a) a seleção de termos - a escolha da forma etc. requerem decisões intelectuais por parte dos indexadores.</p> <p>b) o equipamento e o suporte lógico disponíveis podem ocasionalmente limitar a reprodução de um tesouro a partir de arquivos conservados em máquina. Os indexadores devem ter acesso a equipamentos e programas capazes de satisfazer as necessidades especiais de um tesouro. Não deveria ser necessário adaptar o tesouro às limitações de uma máquina inadequada.</p>	<p><b>Termos candidatos</b></p> <p>São termos propostos que não passaram por todos os procedimentos de validação. Estes termos deveriam ser marcados por um símbolo especial ou frase no registro do termo. Assim que um termo candidato é aprovado como termo, o símbolo ou frase deve ser excluído. Em um sistema online no qual o vocabulário controlado está integrado a uma base de dados única, os termos candidatos, geralmente, não são mostrados ao usuário.</p>
<b>7</b>	<p><b>Uso de equipamento automático de processamento de dados</b></p> <p>a) a seleção de termos: 1) seleção automática de termos candidatos extraídos de títulos e resumos; 2) determinação automática (nos sistemas automatizados de indexação) da frequência com que os termos são usados na indexação e nas perguntas de busca;</p> <p>b) o uso do computador sofre restrições impostas por limitações do equipamento disponível ou do software: 1) número de caracteres permitido em um termo; 2) conjunto de caracteres, inclusive sinais de pontuação, disponíveis na impressora; 3) número de níveis de subordinação permitido em um tesouro sistemático.</p>	<p><b>Forma e conteúdo de um tesouro</b></p> <p>a) as seguintes partes devem ser claramente distinguidas: 1) página de rosto; 2) sumário; 3) introdução; 4) apresentação sistemática ou gráfica (quando for o caso); 5) seção alfabética;</p> <p>b) todos os tesouros devem conter uma introdução abrangente que mencione claramente: 1) o propósito do tesouro; 2) o(s) campo(s) temático(s), identificando separadamente as áreas centrais e relacionadas; 3) os significados de todas as convenções e abreviaturas; 4) o número total de termos, com subtópicos de termos preferidos e não-preferidos; 5) os critérios adotados para selecionar as formas preferidas dos termos de indexação e para estabelecer suas inter-relações; 6) as regras de registro empregadas, seguindo (e citando), quando possível, uma norma nacional ou internacional apropriada; 7) os significados de todos os sinais de pontuação utilizados de maneira não convencional; 8) um resumo da política de atualização e o nome e endereço do serviço responsável, ao qual devem ser enviados comentários e sugestões; 9) a data de inclusão do último termo. Sempre que possível estes pontos devem ser ilustrados com exemplos.</p>	<p><b>Níveis de especificidade</b></p> <p>A adição de termos específicos é normalmente restrita à área central do campo de assunto coberto por um vocabulário controlado/tesouro, porque a proliferação de tais termos em áreas periféricas é susceptível de conduzir a um vocabulário controlado que é difícil de gerenciar. Em uma organização que lida com objetos de conteúdo que cobrem mais de um domínio de conhecimento, pode ser necessário desenvolver um número de vocabulários controlados especializados, cada um, ligado a, e compatível com, um vocabulário controlado geral que tem um nível inferior de especificidade, produzido pela mesma organização.</p>

8	<b>Forma e conteúdo do tesouro</b>	<b>Outras questões editoriais</b>	<b>Termos não designados</b>
	<p>1) não é possível estabelecer um formato padrão de apresentação de tesouros impressos. Os seguintes elementos devem ser destacados: a) página de rosto; b) sumário; c) introdução; d) apresentação sistemática ou planigráfica (quando for possível); e e) parte alfabética;</p> <p>2) todos os tesouros devem conter uma introdução abrangente, que informe de maneira clara: a) finalidade; b) área do conhecimento que abrange, com as áreas de assunto central e periférica identificadas separadamente; c) significados de todas as convenções e abreviaturas; d) número total de termos; e outros.</p>	<p>a) notificação da intenção - em uma publicação profissional apropriada;</p> <p>b) não duplicação de esforços – consultar preliminarmente um ou mais tesouros em campos relacionados.</p> <p>Fazer uma revisão completa do tesouro depois de um período de uso, e depois, a intervalos regulares, levando em conta todas as mudanças no uso dos termos dentro do campo coberto pelo tesouro.</p>	<p>Quando hierarquias são estabelecidas em um vocabulário controlado/ tesouro, os termos que ainda não foram usados na indexação são, frequentemente, admitidos no vocabulário controlado com o fundamento de que eles são necessários para completar uma hierarquia (por exemplo, como termos genéricos), e que têm valor potencial como termos de indexação.</p>

**Quadro 6** – Síntese das Etapas de Construção de um Tesouro segundo as Diretrizes IBICT (1984); Diretrizes UNESCO (1993); e Diretrizes ANSI/NISO (2005).

## Apêndice B

DIRETRIZES	MÉTODOS DE COMPILAÇÃO DE TERMOS: TERMINOLOGIAS E DEFINIÇÕES			
<b>DIRETRIZES...</b> IBICT (1984)	<b>Método Dedutivo</b>  Os termos são extraídos da literatura durante o estágio preliminar de indexação; entretanto, até que um número suficiente de termos tenha sido coletado, todos os termos são considerados por uma comissão, formada de preferência por indexadores e especialistas, que decidem primeiro quais termos representam as classes mais amplas, fixando os demais termos nessas classes e relacionando-os logicamente com base em suas conceituações.	<b>Método Indutivo</b>  Novos termos são admitidos no tesouro logo que são encontrados na literatura, sendo cada termo designado como membro de uma ou mais classes, estabelecidas em bases <i>ad hoc</i> durante o ato de indexação. A elaboração do tesouro é considerada, desde o início, como uma operação contínua e, embora se busque assistência de especialistas sempre que for preciso, tais profissionais não são membros de uma comissão formal.	<b>Combinação de Métodos</b>  Durante a sequência de etapas de construção do tesouro, empregam-se, de modo igual, os métodos dedutivo e indutivo.  Por exemplo: as classes de termos que tenham sido estabelecidas de início de forma indutiva podem ser revistas mais tarde por uma comissão de indexadores e especialistas da área do tesouro; ambas as técnicas são empíricas. Nesse sentido, decisões tomadas nos primeiros estágios da compilação podem ser revistas, à medida que os indexadores ganham mais experiência. A inclusão de novos termos, a conferência dos relacionamentos entre os termos, a fatoração, as notas de aplicação, devem seguir uma linha lógica e consistente. A assessoria de especialistas na escolha entre termos sinônimos e outras formas variantes é recomendada.	<b>Método Relacional*</b>  *Dahlberg aponta outro tipo, a abordagem relacional – que parte do aspecto formal, categorial” (IBICT, 1984, p. 52).
<b>DIRETRIZES...</b> UNESCO (1993)	<b>Método Dedutivo</b>  <b>a) Método dedutivo</b> - os termos são extraídos de documentos durante uma etapa preliminar de indexação. Não é feito controle do vocabulário, nem mesmo para determinar as relações entre os termos. A não ser que um número suficiente de termos tenha sido coletado. Nesse caso, todos os termos são revistos por um grupo de especialistas no assunto. Primeiro, devem ser identificados os termos que representam categorias genéricas, e os termos	<b>Método Indutivo</b>  Novos termos são admitidos no tesouro logo apareçam nos documentos. O controle do vocabulário é feito desde a etapa inicial, e cada termo, à medida que é admitido, é designado como membro de uma ou mais categorias genéricas constituídas sobre uma base <i>ad hoc</i> . O tesouro é, portanto, organizado sobre uma base de termo específico a genérico. A elaboração do tesouro é, desde o princípio, uma operação contínua e, apesar de ser necessária a assistência de	<b>[Combinação de Métodos]</b>  “Na prática, é possível empregar-se tanto o método dedutivo como indutivo em uma ou outra etapa da elaboração do tesouro”. Por exemplo, um grupo editorial composto por indexadores e especialistas pode, primeiramente, estabelecer as categorias dos termos indutivamente e depois examiná-las pelo ponto de vista dedutivo. Ambas as técnicas são essencialmente empíricas e, desde	



	<p>restantes devem ser relacionados nestas categorias conforme as suas relações lógicas. Deste modo, as categorias tendem a ser estabelecidas sobre uma base de termo genérico à específico. O controle do vocabulário deve ser efetuado no estágio em que as categorias são estabelecidas, observando os princípios descritos nas seções de 5 a 7, dessa Norma. A seção 5 trata do Controle de Vocabulário. As seções 6 e 7, abordam dos Termos de Indexação e Termos Compostos, respectivamente.</p>	<p>técnicos especialistas, eles não precisam fazer parte de um grupo editorial formal ou comitê.</p>	<p>o início, deve-se ter em mente que algumas decisões tomadas durante a etapa de compilação possam ter que ser revisadas à medida que se adquire maior experiência. A adição de novos termos a uma linguagem de indexação freqüentemente esclarece as decisões tomadas anteriormente quanto aos termos e suas inter-relações. Os compiladores devem revisar freqüentemente os termos e as hierarquias para assegurar que as relações entre os termos, a faturação de termos compostos etc., seguem os princípios gerais que têm sido aplicados. Recomenda-se particularmente que se busque o auxílio de especialista na escolha dos termos sinônimos e formas variantes dos termos.</p>	
<p><b>DIRETRIZES... ANSI/NISO (2005)</b></p>	<p><b>a) Abordagem de Comitê</b></p> <p>- especialistas no domínio de assunto do vocabulário controlado fazem uma lista de termos chave na área e indicam as relações entre eles, com assistência de especialistas em criação de vocabulário controlado. Listas de termos podem ser retiradas de várias fontes ou submetidas a vários usuários/comunidades. Dois métodos principais para criar vocabulários controlados por Comitê têm sido utilizados:</p> <p>1) <b>Top Down</b> (Descendente) – os termos mais genéricos são identificados primeiro e então termos mais específicos são selecionados para atingir o nível desejado de especificidade. As estruturas hierárquicas e relações necessárias são estabelecidas à medida que o trabalho evolui;</p> <p>2) <b>Bottom up</b> (Ascendente) – este caso ocorre frequentemente quando listas de termos forem derivadas de um <i>corpus</i> de</p>	<p><b>b) Abordagem Empírica</b></p> <p>- a abordagem empírica evidencia dois métodos básicos:</p> <p><b>i) Método Dedutivo</b> – termos são extraídos de objetos de conteúdo (por humanos ou computadores), opcionalmente durante um estágio preliminar de indexação, mas nenhuma tentativa é feita para controlar o vocabulário, nem para determinar relações entre termos, até que um número suficiente de termos tenha sido coletado. Todos os termos são revisados por um grupo de especialistas, preferencialmente, composto de especialistas em informação e no assunto tratado. Eles deveriam primeiro identificar termos que representam as classes mais amplas, e então alocar os termos remanescentes para aquelas classes segundo suas relações lógicas, a fim de que as hierarquias tendam a ser estabelecidas sobre uma base da mais genérica para a mais específica. O controle de vocabulário deveria ser desenvolvido no</p>	<p><b>c) Combinação de Métodos</b></p> <p>- na prática, é possível empregar mais de uma dessas abordagens em um estágio ou outro durante a construção de um vocabulário controlado. Por exemplo, hierarquias e outras relações entre termos que foram primeiro estabelecidos indutivamente poderiam mais tarde ser examinadas a partir de um ponto de vista dedutivo. Ambas as técnicas são essencialmente empíricas, e deveria ser aceito desde o início que algumas decisões com relação aos termos e suas inter-relações, que foram feitas durante os primeiros estágios de compilação, possam ser revisadas como uma experiência posterior. Os compiladores deveriam verificar os termos e hierarquias frequentemente para assegurar uma aplicação consistente de princípios em tais procedimentos como: o estabelecimento de inter-relações entre termos e a divisão de termos</p>	<p><b>d) Assistência da Máquina</b></p> <p>– assume-se de acordo com essa Norma que a construção de vocabulário controlado/tesauro requer decisões intelectuais. A assistência da máquina pode ser empregada, no entanto, para tarefas de identificação de termos tais como as seguintes: i) Identificação de termos candidatos – termos candidatos podem ser identificados automaticamente a partir de texto legível por máquina, (por exemplo: títulos e/ou resumos). O número de termos potenciais deveria ser reduzido pelo uso de palavras vazias ou uma lista vazia (palavras funcionais – artigos, conjunções, e preposições tais como “um/uma”, “e”, e “para” – além de outras palavras consideradas sem valor para a recuperação). Remover as palavras vazias da consideração é geralmente o primeiro passo para a</p>

	<p>objetos de conteúdo e devem então ser incorporadas em um vocabulário controlado. Como no caso acima, as estruturas e relações hierárquicas necessárias são criadas à medida que o trabalho avança, mas inicia-se com os termos que têm uma abrangência mais específica e segue para os termos mais genéricos. Se um novo vocabulário controlado/tesauro será construído, a abordagem “descendente” é preferida. No caso de um vocabulário controlado/tesauro já publicado, a abordagem “ascendente” é mais usada para o acréscimo de novos termos para cobrir novos conceitos.</p>	<p>estágio onde hierarquias e outras relações são estabelecidas, seguindo os princípios descritos nas seções que tratam de Termos compostos; e de Relacionamentos entre os termos;</p> <p><b>i) Método Indutivo</b> - novos termos são selecionados para potencial inclusão no vocabulário controlado/tesauro como eles forem encontrados em objetos de conteúdo. O controle de vocabulário é aplicado desde o início. Se o vocabulário em construção tiver algum tipo de arranjo hierárquico, cada termo admitido é designado como um membro de uma ou mais classes genéricas que são construídas sobre uma base <i>ad hoc</i> em um estágio anterior. O vocabulário controlado é, portanto, estabelecido sobre uma base de termo do mais específico para o mais genérico. A construção de vocabulário controlado é considerada do início como uma operação contínua. Assistência de especialistas no assunto é altamente recomendada; estes especialistas podem servir como membros de uma comissão ou comitê editorial/formal.</p>	<p>compostos;</p>	<p>indexação auxiliada pela máquina. Os termos remanescentes deveriam ser comparados com aqueles já registrados no vocabulário controlado. Termos que não foram registrados poderiam ser considerados como candidatos à inclusão; <b>ii)</b> Registro da frequência de designação (indicação) de termo – Em sistemas de indexação computadorizados, a frequência com a qual um termo ter sido usado em indexação, pode ser registrada automaticamente. Termos com resultados altos ou baixos podem ser considerados candidatos à modificação ou exclusão; <b>iii)</b> Registro de termos a partir das consultas do usuário – Termos encontrados em consultas de usuário que não combinam com um ou mais termos de entrada também podem ser considerados para inclusão, especialmente quando um dado termo ocorre em múltiplas consultas. Para assegurar privacidade, os usuários não deveriam ser identificados quando se rastreia e coleta consultas de usuário. Termos que ocorrem em linguagem natural e são formas comuns de descrever um conceito são incluídos no vocabulário controlado/ tesauro como termos de entrada.</p>
--	---	---	-------------------	---

**Quadro 7 – Métodos de Compilação de termos: terminologias e definições.**