

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

CARLA TATIANE VIEIRA

QUALIDADE NA REPRESENTAÇÃO TEMÁTICA:
algumas considerações

Porto Alegre
2005

CARLA TATIANE VIEIRA

QUALIDADE NA REPRESENTAÇÃO TEMÁTICA:
algumas considerações

Monografia apresentada como requisito parcial para a obtenção de título de graduação em Biblioteconomia da FABICO/UFRGS.

Orientadora: Prof^a. Dr^a. Regina Helena van der Laan

Porto Alegre
2005

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Prof. Dr. José Carlos Ferraz Hennemann.

Vice-Reitor: Prof. Dr. Pedro Cezar Dutra Fonseca

FACULDADE DE BILIBLIOTECONOMIA E COMUNICAÇÃO

Diretora: Prof. Valdir Jose Morigi

Vice Diretor: Prof. Ricardo Schneiders da Silva

DEPARTAMENTO DE CIÊNCIAS DA INFORMAÇÃO

Chefe: Profª. Iara Conceição Bitencourt Neves

Chefe Substituta: Profª. Jussara Pereira Santos

Orientadora :

Profª. Drª. Regina Helena van der Laan

e-mail: rhvd@ufrgs.br

V685q Vieira, Carla Tatiane.

Qualidade na representação temática: algumas considerações./ Carla Tatiane da Silva Vieira; Orientadora Regina Helena Van der Laan . Porto Alegre: UFRGS, 2005.

61p.

1. Representação temática – Qualidade. I. Van der Laan, Regina Helena. II. Título.

CDU 025.4

DEPARTAMENTO DE CIÊNCIAS DA INFORMAÇÃO

Rua Ramiro Barcelos, 2705

CEP: 90035-007 – Porto Alegre – RS

Tel.: (051) 3316 5146

Fax: (051) 3316 5

E-mail: fabico@ufrgs.br

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

BANCA EXAMINADORA

Prof^a. Glória Isabel Sattamini Ferreira

Prof^a. Martha E. K. Bonotto

ORIENTADORA

Prof^a. Dra. Regina Helena van der Laan

Porto Alegre, ____ de _____ de 2005.

VIEIRA, Carla Tatiane. **Qualidade na Representação Temática: algumas considerações.** Porto Alegre: UFRGS, 2005.

RESUMO:

Relato de pesquisa qualitativa, que consiste em uma Revisão de Literatura sobre a Representação Temática no Processo de Indexação. Objetiva verificar o atual estado-da-arte na qualidade da representação temática, a fim de identificar as contribuições teóricas e atuais tendências, a partir de um levantamento na literatura periódica publicada na área, nos últimos cinco (5) anos. Por tratar-se de estudo teórico e não o resultado da observação direta de dados, seu corpus de pesquisa consistiu no levantamento de informações publicadas somente em artigos de periódicos eletrônicos e impressos publicados no Brasil. A partir das afirmações dos autores, delinear-se os pontos comuns encontrados na literatura acerca da qualidade da representação temática apontados por Lancaster e tratados com maior profundidade por outros autores. Aborda o caráter interdisciplinar do conhecimento humano e descreve a importância do conhecimento e utilização prática das contribuições teórico-metodológicas em que se pautam a terminologia e a lingüística textual para oferecer qualidade à representação temática. Destaca o aspecto problemático do fenômeno da explosão informacional que repercute na área de representação temática devido ao seu caráter mutante, e como o desenvolvimento de habilidades de leitura profissional e o domínio dos instrumentos de indexação qualificam a etapa de representação. Descreve os fatores intervenientes na qualidade da representação que estão ligados ao indexador, ao processo, ao vocabulário e ao documento. Investiga na literatura científica elementos que ofereçam à análise documentária subsídios teóricos consolidados para que se construa, a partir das reflexões dos autores, seu próprio referencial teórico e metodológico que leve em consideração os objetivos do centro de informação no atendimento das demandas sociais e individuais. Constata que a análise documentária carece ainda de muitas reflexões e considerações sobre a interferência das variáveis presentes no processo de indexação e como influem na qualidade da representação.

Palavras-chave: Representação temática. Análise Documentária.

VIEIRA, Carla Tatiane. **Qualidade na Representação Temática**: algumas considerações. Porto Alegre: UFRGS, 2005.

ABSTRACT:

Account of quality investigation, that treat in a Revision of Literature above Representation of subject in Processe of Indexing. Objective inspect the present state of the art in the quality representation of subject, the in order that identify in to contribute theory and present tendency, the beginning from survey in the literature periodical publish in the area, in the last five (5) year. By it is a question of the theoretical study and not result of the data direct observation, your body of the research treat of the publish information survey only in article electronics periodicals publish and print in the Brazil. It is a question of the authors, deliate the common points meet in the literature about of the representation of subject quality sharpen by Lancaster and treat with greater profundity by others authors. Board the interdisciplinary character of the human knowledge and describe the importance of the knowledge and using pratic of the contribution theoric-methodological in that guide in Therminology and Textual Linguistics for to offer quality at the representation of subject. Singularize the aspect problem of the phenomenon of the information explosion who reflect in the área of the representation of subject due to your mutation character, and how the development of the hability professional reading and the domain of the instrumento f indexing qualify the stage of the representation. Describe the aspects intervening in the quality of the representation of subject who be linked an the indexador, and the processe, an the vocabulary an the document. Search in the cientific literature element who offer an the analisys documentation theorics subsidy consolidate for who if construct the beginning from reflects of the authors if peculiar fature theoric referencial and metodologic who consider the objective of the information center in the answer of the socials and individuals demand. Verify who the analisys documentation to want very reflections and considering above the inference of the variable presents in the processe of the indexing and what induce in the quality of representation.

Key-words: Representation of Subject. Analisys Documentation.

LISTA DE FIGURAS

Ilustração 1 – Fatores que afetam a qualidade da indexação	15
Ilustração 2 – Conhecimento por parte do indexador.....	17
Ilustração 3 – Situações possíveis entre as variáveis leitor, texto e contexto	22
Ilustração 4 – Relações entre as variáveis leitor, texto e contexto.....	23
Ilustração 5 – Tensões entre paradigmas da linguagem natural e da linguagem documentária	50

SUMÁRIO

1 INTRODUÇÃO	10
2 METODOLOGIA	12
3 A REPRESENTAÇÃO TEMÁTICA E OS FATORES QUE AFETAM A QUALIDADE	14
3.1 Fatores Ligados ao Indexador	16
3.1.1 Conhecimento temático	16
3.1.2 Experiência	18
3.1.3 Concentração	19
3.1.4 Capacidade de síntese, análise e compreensão	19
3.1.5 Falta de coerência	23
3.2 Fatores Ligados ao Vocabulário	25
3.2.1 Especificidade/Sintaxe	25
3.2.2 Ambigüidade ou imprecisão	27
3.2.3 Disponibilidade de instrumentos auxiliares	28
3.2.4 Qualidade da estrutura e qualidade do vocabulário	28
3.3 Fatores Ligados ao Documento	30
3.3.1 Conteúdo temático	31
3.3.2 Complexidade	31
3.3.3 Língua e linguagem	33
3.3.4 Extensão	34
3.3.5 Apresentação e sumarização	35
3.4 Fatores Ligados ao Processo	35
3.4.1 Tipo de indexação	36
3.4.2 Política de indexação	36
3.4.3 Produtividade exigida	40
3.4.4 EXAUSTIVIDADE DA INDEXAÇÃO	4

4 QUALIDADE DA REPRESENTAÇÃO E OS SISTEMAS DE CLASSIFICAÇÃO	42
5 QUALIDADE DA REPRESENTAÇÃO E O CARÁTER MUTANTE DA INFORMAÇÃO	45
6 CONSIDERAÇÕES SOBRE A ANÁLISE DOCUMENTÁRIA	56
7 DISCUSSÃO	61
8 CONCLUSÃO	65
REFERÊNCIAS.....	67

1 INTRODUÇÃO

Este trabalho é fruto de algumas considerações e reflexões em torno da qualidade da representação temática, surgidas ao longo de estágios em algumas bibliotecas, que nos deram a percepção da existência de certa discrepância entre a teoria proposta e discutida pelos diferentes autores e a prática nas bibliotecas. A discussão sobre estas questões, levantadas em sala de aula, despertou-nos a curiosidade de saber o atual estado-da-arte na representação temática de nosso país, motivando-nos a propor este trabalho, com a intenção de contribuir, indicando tendências teóricas que objetivam qualificar o processo de indexação.

Como poderemos constatar adiante, esta discrepância está mais relacionada ao desconhecimento, por parte do indexador, das variáveis presentes no processo de indexação e das teorias para operacionalização e estabelecimento de critérios metodológicos de análise e representação de documentos no contexto profissional, e não à inexistência desses aparatos teórico-conceituais.

Neste contexto, iniciamos uma revisão de literatura para verificar o atual estado-da-arte na qualidade da representação temática, visando identificar as contribuições teóricas e atuais tendências, a partir de um levantamento na literatura periódica brasileira publicada na área, nos últimos cinco (5) anos.

As transformações decorrentes do fenômeno de explosão informacional e o vertiginoso crescimento da produção de conhecimento trouxeram novos e grandes desafios à representação temática, gerando dificuldades na escolha de termos de indexação, uma vez que muitos dos instrumentos de indexação não são suficientemente flexíveis para acompanhar essas mudanças. Se, por um lado, o

aumento da produção do conhecimento ocasionou uma maior agilidade no acesso da informação, a partir do desenvolvimento de novas tecnologias de busca/recuperação da informação, por outro, essa aparente “comodidade”, obriga os profissionais bibliotecários, a um estudo aprofundado dos instrumentos de indexação, sua organização e das teorias que o fundamentam.

Essa imensa quantidade de informações, produzidas e disponibilizadas pelas diferentes áreas do saber humano, dificulta a identificação dos limites de domínio do conhecimento a que pertence, ao mesmo tempo em que acarreta sérios problemas de inconsistência na indexação.

Nesta acepção, o conhecimento dos aspectos lingüísticos do texto e a compreensão dos mecanismos lógicos e semânticos envolvidos na análise temática vêm contribuindo para a qualidade da representação.

Os princípios em que se pauta a teoria da Terminologia vêm oferecendo à análise documentária contribuições teóricas e metodológicas significativas em conseqüência do progresso alcançado por outras áreas do conhecimento, e com as quais, a análise documentária mantém esse intercâmbio.

2 METODOLOGIA

Este trabalho é uma pesquisa qualitativa, que consiste em uma Revisão de Literatura sobre a Representação temática no Processo de Indexação.

Devido ao fato, de ser um estudo teórico e não o resultado da observação direta de dados, seu corpus de pesquisa consistiu no levantamento de informações publicadas somente em artigos de periódicos eletrônicos e impressos, no período de 2000 a 2005, publicados no Brasil.

As fontes de consulta utilizadas para a identificação dos artigos publicados e que constituíram o corpus de pesquisa, inicialmente, foram:

- a) LISA – periódico de resumos, contendo ampla cobertura de títulos de periódicos na área de biblioteconomia e ciência da informação, cerca de 550 títulos de periodicidade mensal, cobertura internacional, apresenta o resumo dos artigos;
- b) WEB OF SCIENCE – base de dados contendo referências de artigos publicados em todas as áreas do conhecimento, em diversas línguas;

Pelo fato, de não haverem sido encontradas, nas fontes de consulta acima mencionadas, um número suficiente de artigos que tratassem da qualidade da representação temática no período delimitado inicialmente por esta revisão, optamos por utilizar ferramentas de busca disponíveis na WEB (Google scholar) para localização de artigos eletrônicos.

As fontes para localização/obtenção de documentos foram utilizadas através do portal de periódicos do CAPES, o catálogo de bibliotecas da UFRGS, e o catálogo coletivo nacional do IBICT.

Foram elaboradas fichas de citações, a fim reproduzir as conclusões e afirmações dos autores, os problemas e as soluções propostas nos artigos, quando encontradas, juntamente com a referência bibliográfica e a localização do artigo. Na fase de análise e interpretação, a leitura do texto foi realizada, a fim de extrair as principais idéias dos autores.

Seguimos um roteiro para a elaboração do trabalho, no qual o assunto foi dividido em tópicos e subtópicos, conforme a necessidade apresentada para que a abordagem do assunto se tornasse mais clara.

As conclusões dos autores permitiram-nos delinear o corpus textual deste trabalho, a partir dos pontos comuns encontrados na literatura acerca da qualidade da representação temática apontados por Lancaster (2003)¹ e tratados com maior profundidade por outros autores.

Esta revisão se viu limitada pelo escasso número de artigos encontrados, que tornou precárias as condições para a elaboração deste trabalho.

Foi necessário recorrer, à análise global da temática para inferir, a partir dela, os problemas específicos da representação temática e, não fosse a opção de elaborá-lo embasado nos estudos de Lancaster sobre os fatores que influenciam a qualidade da indexação, não teria sido concluído, devido a dificuldade de identificar, na literatura periódica, dentro das circunstâncias espaço-temporais pré-estabelecidas no projeto deste trabalho, a menção de todos os fatores que influem na qualidade da representação temática, apontados por este autor. Este fato pode nos levar a supor que carecemos, ainda, de investigações mais aprofundadas sobre o tema.

¹ LANCASTER, F. W. **Indexação e resumos**: teoria e prática. Brasília: Briquet de Lemos, 2003. 360 p.

3 A REPRESENTAÇÃO TEMÁTICA E OS FATORES QUE AFETAM A QUALIDADE

Neste capítulo, serão abordados os fatores que podem afetar a qualidade da representação temática. Apesar de serem poucas as pesquisas realizadas sobre a qualidade da representação temática, far-se-á neste trabalho, a tentativa de identificação dos fatores que influenciam sua qualidade, tomando-se por base os estudos realizados por Lancaster (2003).

A representação temática é a descrição do conteúdo de um documento, mediante emprego de um ou vários termos de indexação selecionados a partir de um vocabulário controlado, que servem como pontos de acesso para recuperação e localização de um item, que visa, normalmente, a “[. . .] atender as necessidades dos usuários de determinado centro de informação ou de uma publicação específica.” (LANCASTER, 2003, p.7-9).

O mesmo autor apresenta como fatores interferentes no processo de indexação, os representados na figura abaixo. Os fatores sinalizados por asterisco, (*) foram modificados por terem sido considerados terminologicamente imprecisos.

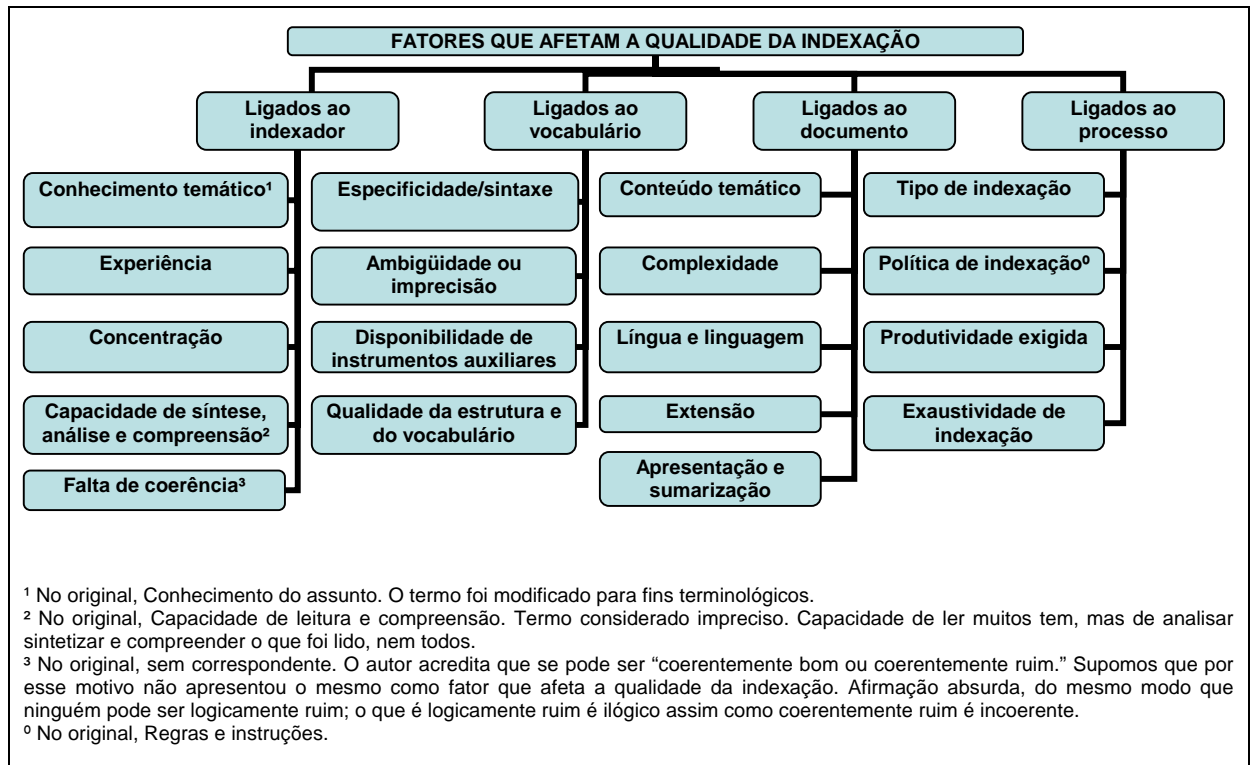


Ilustração 1 – Fatores que afetam a qualidade da indexação²

Para fins didáticos, optamos por analisar separadamente esses fatores, mesmo sabendo estarem ligados, a um ou mais dos outros fatores (ao indexador, ao vocabulário, ao documento e ao processo), pois são parte de um processo indissociável. Na tentativa, de não tornar repetitivas as explicações, preferimos enquadrá-los uma única vez em cada categoria, mesmo estando estes interconectados no processo de indexação.

² FONTE: LANCASTER, F. W. Qualidade da Indexação. In: _____. Indexação e resumos: teoria e prática. Brasília: Briquet de Lemos, 2003. cap. 6, p.89

3.1 Fatores Ligados ao Indexador

Os principais fatores ligados ao indexador são: o conhecimento temático, a experiência, concentração, capacidade de análise e síntese, compreensão e falta de coerência (percepção dos itens a serem indexados de forma diferente em épocas distintas).

3.1.1 Conhecimento temático

O processo de análise documentária, tal como ocorre no processo de conhecimento, no processo de representação e construção da informação documentária, conforme Lara (2002, p. 128), “[. . .] recorre a recortes de conteúdo, à analogia e a generalização.” A autora, tecendo um paralelo entre o processo de conhecimento e o processo de representação da informação documentária, demonstra que na representação, esta segmentação (divisão) de conteúdo é obtida através:

[. . .] de referências que já possuímos, formando agrupamentos em função de suas propriedades comuns, ou mais exatamente, das características que julgamos pertinentes para os nossos propósitos, já que nunca esgotamos as propriedades das coisas e fenômenos.

.....
 No processo documentário a representação de conteúdos se relaciona a propósitos e concepções definidas de antemão. [. . .] ou fazemos uso de classificações existentes (se elas forem consideradas ‘adequadas’ para acolher mais membros), ou ainda, segmentamos mais uma vez o conteúdo propondo uma nova hipótese de organização.

O conhecimento temático prévio, por parte do indexador, auxilia na etapa da representação, a partir do momento em que o mesmo recorre a conhecimentos que

já possui (recortes de conteúdo, generalização, analogia), permitindo a associação dos conhecimentos precedentes e do conhecimento dos interesses dos usuários, ao que uma informação proporciona de novidade, interrelacionando-os com outros documentos já indexados. Baseando-se nestas teorias, podemos entender o processo de representação, no contexto do indexador, como leitor profissional, dependente de alguns fatores que auxiliarão na qualidade na representação, os quais, inspiraram-nos a representação esquemática, conforme o quadro abaixo:

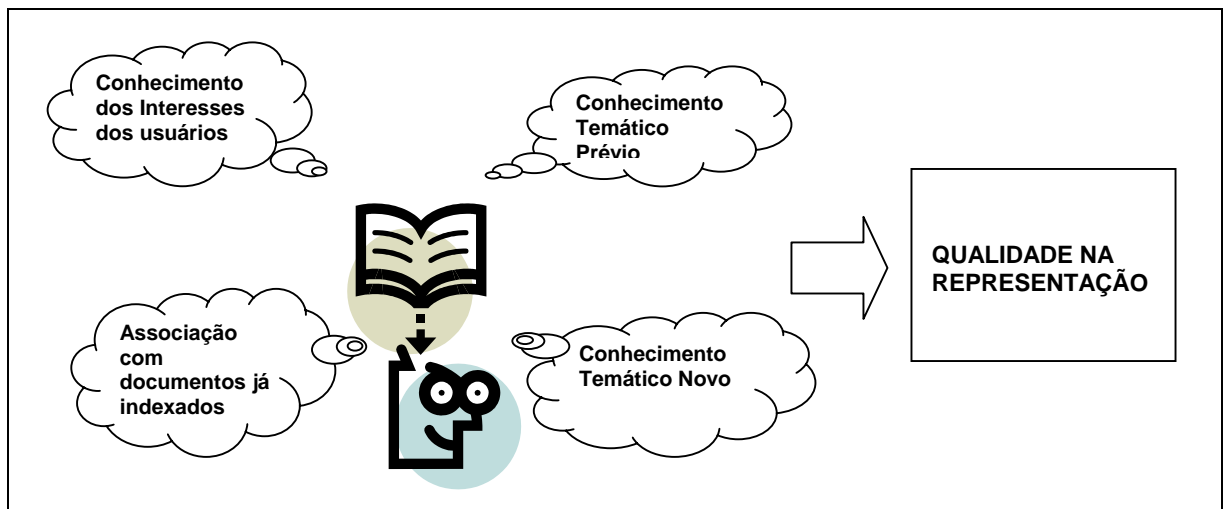


Ilustração 2 - Conhecimento por parte do indexador

Lancaster (2003, p. 88) acredita que tanto os profissionais muito especialistas como aqueles que possuem pouco conhecimento temático, podem comprometer a qualidade da indexação. Os primeiros, podem apresentar uma tendência a interpretar o conteúdo de um documento indo além do que autor quis afirmar ou censurando afirmações por não aceitá-las ou considerarem-se *experts* no assunto. Já os últimos, por sua falta de conhecimento, podem indexar excessivamente um documento, mas por outro motivo: a incapacidade de distinguir entre dois termos, podendo atribuir ambos, quando, o emprego de um só seria correto.

Muitos dos fatores de qualidade da indexação estão ligados à capacitação profissional do indexador, quer seja no conhecimento de teorias, ou no emprego de metodologias de análise e representação do conhecimento, que mais se aproximem da função social desempenhada pelo centro de informação. A capacitação profissional, quando guiada a partir das necessidades da comunidade usuária e dos objetivos institucionais pré-estabelecidos em manuais e políticas de indexação, tornam altamente convencional o sistema, revelando o que Nunes (2000, p. 107) denomina “saber lingüístico.” Quanto maior for este, menos arbitrária se torna a indexação e maior será a rede de relações entre as unidades (termos) de que o indexador poderá dispor para representar a informação, garantindo melhorias na obtenção do produto final, a informação solicitada.

3.1.3 Experiência

A experiência do indexador influi na qualidade da indexação, na medida em que possibilita ao indexador um bom domínio sobre o assunto, maior conhecimento dos interesses dos usuários, maior conscientização sobre seu próprio conhecimento, suas potencialidades, capacitação e aprimoramento de sua formação profissional para o atendimento das demandas de sua comunidade usuária.

Um estudo realizado por Fujita (2004), a fim de coletar dados a respeito da atuação profissional de indexadores, aqui entendidos como leitor, e sua formação, mediante revisão de literatura, entrevistas com indexadores e diagnóstico de sistemas de informação brasileiros, levou em conta cada uma das variáveis do processo de leitura documentária (leitor, texto e contexto), constatando que os

indexadores tinham como fontes de experiência em indexação, as adquiridas através de treinamentos, uso de manuais e pelo próprio esforço.

3.1.4 Concentração

A capacidade de concentração do profissional interfere na qualidade da representação. Por ser desgastante e exaustiva a execução do trabalho intelectual, na análise e compreensão dos conceitos expressos em um documento, tanto as condições ambientais desfavoráveis (ruídos, trânsito muito intenso de pessoas no local onde se executa a tarefa de indexação, etc.), como a própria capacidade de concentração do indexador, podem atuar negativamente nos resultados da indexação, mas não são as que mais interferem na qualidade da representação. (LANCASTER, 2003, p. 89).

3.1.5 Capacidade de síntese, análise e compreensão

O indexador deve preocupar-se em estudar metodologias de análise temática que, não só o auxiliarão a obter maior qualidade e precisão na indexação, como também criarão uma uniformidade de procedimentos, auxiliarão a complementar a política de indexação, e entender o funcionamento da terminologia empregada no sistema de informação adotado pelo centro de informação. (FUJITA, 2004; LANCASTER, 2003; LARA, 2001).

Fujita (2004, p.1) define a leitura profissional como “[. . .] um ato social porque existe um processo de comunicação e de interação entre o leitor (indexador) e o autor do texto, ambos com objetivos estabelecidos anteriormente dentro do contexto de cada um.” A autora analisa o processo de leitura profissional, dentro da perspectiva de abordagem de Giasson³ (1993) que vê a leitura como um processo interativo de três variáveis leitor (indexador), texto e contexto.

A leitura de cunho profissional efetua-se de forma diferente da leitura comum, sendo especialmente voltada aos objetivos da indexação, ou seja, representar conteúdos, através do uso de termos que serão recuperados pelos usuários do sistema de informação. Enquanto leitor, o indexador deve conhecer procedimentos e estratégias de análise de documentos. (FUJITA, 2004, p.2)

Ao ter consciência de como se efetua o processo de leitura para fins documentários, o indexador torna-se mais capacitado para compreender suas próprias dificuldades e todo o mecanismo de engendramento dos componentes textuais a partir de sua tipologia discursiva, sua estrutura textual e superestrutura. Por serem parte dos fatores ligados ao documento e que igualmente influenciam na qualidade da indexação, serão explicitados mais adiante, muito embora estejam interconectados aos fatores ligados ao indexador, uma vez que são parte de um processo maior, sendo portanto, indissociáveis.

Sendo assim, a capacidade de síntese, análise e compreensão da temática envolvida na análise de textos, vai depender fundamentalmente dos conhecimentos prévios do indexador sobre estas tipologias e estruturas textuais. A primeira está relacionada à classificação dos textos (científicos, técnicos, literários, etc.) e a

³ GIASSON, J. **A compreensão na leitura**. Lisboa: ASA, 1993. *apud* FUJITA, 2004, p.6-7.

segunda, à maneira como as idéias, os conceitos são organizados no texto que o autor tenciona transmitir. (GIASSON⁴, 1993, p.36 *apud* FUJITA, 2004, p.7)

A capacidade de leitura do indexador, analisada sob a perspectiva de Fujita (2004, p. 2) levou a autora a concluir que “[. . .] existe uma lacuna quanto ao conhecimento disponível sobre o processo de leitura e sua influência nos resultados de representação temática da informação.” A autora, combinando as variáveis interferentes no processo de indexação, analisa três situações e as dificuldades que podem gerar, afetando a qualidade da indexação, que são as seguintes:

- a) se o indexador não realiza uma leitura compreensiva do texto porque não tem domínio do assunto, tem pouco conhecimento profissional de metodologias de análise, mas domina a política do sistema de informação (manual e linguagem documentária) e conhece bem o usuário porque trabalhou no atendimento de buscas em bases de dados, ele terá dificuldades na leitura do texto mas se apoiará no seu conhecimento sobre o sistema e, provavelmente, terá um resultado que certamente será o da representação do texto por conceitos baseados na demanda e não no conteúdo;
- b) se o indexador realiza uma leitura compreensiva do texto porque tem bom domínio do assunto, tem conhecimento profissional de metodologias de análise, mas não domina a política do sistema de informação (manual e linguagem documentária) e não conhece o usuário porque nunca trabalhou no atendimento de buscas em bases de dados, ele terá dificuldades na leitura do texto porque não consegue decidir sobre o conceito mais significativo, tendo em vista a recuperação pelo usuário e provavelmente, terá um resultado que certamente será o das representação do texto por conceitos baseados apenas no seu conteúdo;
- c) se o indexador realiza uma leitura compreensiva do texto porque tem bom domínio do assunto, mas não tem conhecimento profissional de metodologias de análise e desconhece a estrutura textual do texto, apesar de dominar a política do sistema de informação (manual e linguagem documentária) e conhecer o usuário em suas demandas, ele terá mais dificuldades na leitura do texto porque não tem um procedimento sistematizado de abordagem do texto e terá uma demora maior para encontrar o conceito mais significativo ou se limitará a retirar seus conceitos do resumo ou do título, sem explorar a estrutura textual. (FUJITA , 2004, p. 1-27).

Estas não são as únicas situações que combinam as variáveis. Segundo a mesma autora, é preciso ter consciência, conhecimento e compreensão de todas as variáveis presentes no processo de indexação, que analisadas separadamente, irão

⁴ GIASSON, J. **A compreensão na leitura**. Lisboa: ASA, 1993. *apud* FUJITA, 2004, p. 7.

influenciar na qualidade da indexação, na mesma medida em que o processo de leitura influi no resultado da indexação.

Giasson⁵ (1993 *apud* FUJITA, 2004, p. 5) exemplifica três situações que dificultam a compreensão do texto, na ilustração 3 representamos graficamente as explicações do autor e reproduzimos seu modelo sobre as relações entre as variáveis leitor, texto e contexto na ilustração 4 da página seguinte:

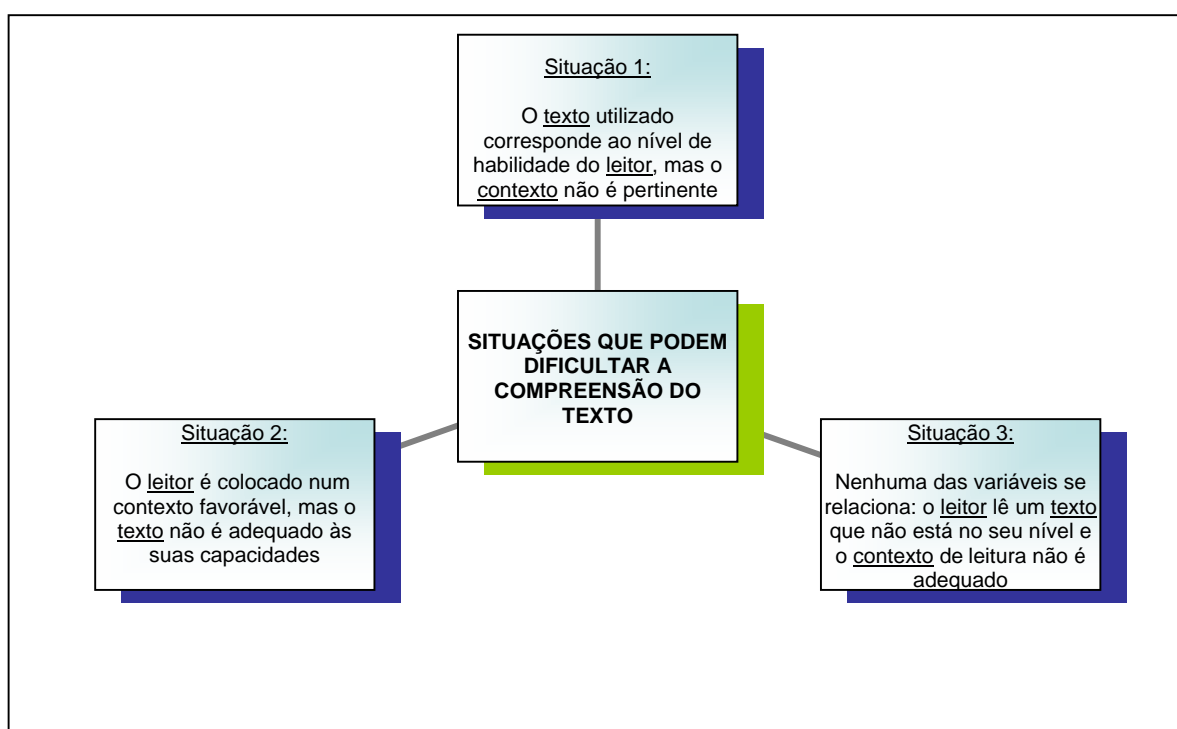


Ilustração 3 – situações de possíveis combinações entre as variáveis leitor, texto e contexto⁶

⁵ GIASSON, J. **A compreensão na leitura**. Lisboa: Asa, 1993. *apud* FUJITA, 2004, p. 5.

⁶ FONTE: FUJITA, Mariângela Spotti Lopes. A leitura documentária na perspectiva de suas variáveis: leitor-texto-contexto. **DataGramZero**: revista de ciência da informação, v.5, n.4, ago. 2004. Disponível em: <<http://www.dgz.org.br>>. Acesso em: 25 jul. 2005.

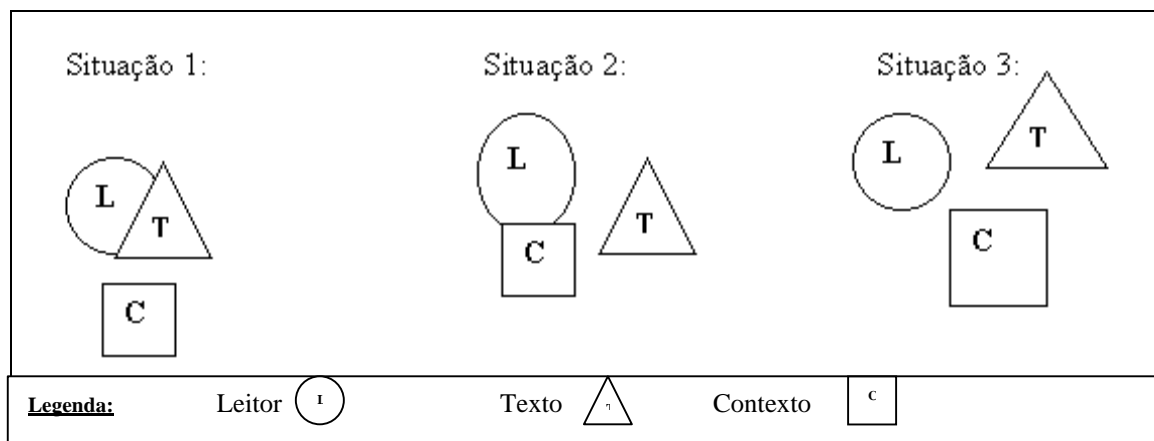


Ilustração 4 – Relações entre as variáveis leitor, texto e contexto⁷

Deste modelo, e conforme expõe Fujita (2004, p. 6), pode-se inferir que no processo de compreensão, “[. . .] a combinação entre as variáveis (leitor, texto e contexto) produzem a ‘situação ideal’ na análise de um documento.”

Isso quer dizer que as variáveis interferentes no processo de **leitura profissional** não estão somente ligadas ao leitor-indexador, e sim, do ponto de vista da análise documentária, a todo o processo de indexação determinando sua eficiência, a partir do conhecimento da rede relacional, estabelecida simultaneamente, entre as variáveis leitor-texto-contexto, a serem combinadas no momento da análise.

3.1.6 Falta de coerência

Um mesmo indexador, em momentos distintos, pode perceber de forma diferente o conteúdo real do documento.

⁷ FONTE: GIASSEN, J. **A compreensão na leitura**. Lisboa: Asa, 1993. *apud* FUJITA, 2004, p. 6.

A utilização de uma política de indexação e vocabulário controlado para nortear as atividades do indexador no momento da representação temática, é de máxima importância. A primeira orienta a indexação de acordo com as necessidades de informação dos usuários, enquanto que o segundo serve para estabelecer o modo em que se dará a representação temática de forma que haja maior coerência entre os termos indexados.

Pode-se determinar o nível de coerência alcançado na indexação mediante alguma forma de consenso entre os diferentes indexadores, medida com o emprego de um padrão, obtido através de “[. . .] um conjunto de termos estabelecido de comum acordo por indexadores altamente experientes.” (LANCASTER, 2003, p. 93)

Nem todos os autores concordam sobre a aplicabilidade destes indicadores de nível de coerência. Enquanto alguns acreditam que **não são praticáveis**, outros os classificam como **não confiáveis**.

A qualidade da representação temática compreende a análise da **temática envolvida em um documento** e a atribuição de termos específicos para representá-la, nas devidas áreas de conhecimento as quais pertence, devendo, portanto, recorrer à utilização de teorias que contemplem a análise desses conceitos dentro de sua perspectiva conceitual, a exemplo da Terminologia, e não sob perspectivas estatísticas, pois não encerram, em si mesmas, um princípio organizador do conhecimento, e sim um princípio aleatório baseado em **quantificadores de consenso entre indexadores**.

3.2 Fatores Ligados ao Vocabulário

Os fatores ligados ao vocabulário também interferem na qualidade da indexação. Os termos utilizados para representar conceitos nos vocabulários controlados nem sempre permitem representar o conteúdo em sua totalidade, sendo às vezes necessário o uso de outros vocabulários na fase de tradução para os termos da linguagem de indexação, o que significa transformar os conceitos selecionados, em termos ou símbolos autorizados para representá-los no sistema.

Van der Laan (2002, p. 27) define vocabulário controlado como “[. . .] um instrumento utilizado pelos indexadores, com a finalidade de representar o conteúdo temático dos documentos [. . .]” a partir de listas estruturadas de termos autorizados ou preferidos para controle de sinônimos, formas de preferência de grafia e casos de homônimos.

“Quando indexamos um documento por assunto, nós o colocamos em uma ou mais classes, e cada classe deve ter um número ou rótulo, que são chamados (sic) **termos indexadores.**” (NAVES, 2004, p.9). O conjunto desses termos forma uma linguagem de indexação.

3.2.1 Especificidade/Sintaxe

Segundo Naves (2004, p. 19), sintaxe “[. . .] é o conjunto de artifícios empregados para revelar as relações entre os conceitos e as regras para estabelecer os descritores e determinar a ordem em que devem ser citados.”

Essas relações podem ser de dois tipos: semânticas ou sintáticas. As relações semânticas (ou de significado) têm por objetivo, indicar os termos alternativos ou substitutivos de indexação e podem ser de 3 tipos: de equivalência, hierárquicas e associativas. As relações sintáticas originam-se da interseção entre duas ou mais classes de conceitos distintos.

Os termos utilizados para representar conceitos devem traduzir o mais perto possível, o conteúdo temático de um documento (especificidade). A generalidade ou especificidade deve ser analisada e a escolha dos termos deve encerrar os conceitos expressos **no documento**. (MOREIRO, 2002, p. 58)

No processo de indexação devemos indexar pelos conceitos específicos e não pelos genéricos, pois quanto maior o grau de especificidade, maior a precisão ou pertinência relacionada à atinência do documento (de que trata). Por precisão podemos entender a capacidade do sistema de evitar documentos inúteis numa pesquisa, que não sejam do interesse do usuário.

O sistema deve permitir a recuperação de itens que considerem as necessidades informacionais do usuário e que coincidam com aquilo que o autor oferece, para que a comunicação entre o usuário e o sistema seja satisfatória. (LANCASTER, 2003; VAN DER LAAN, 2002)

Para tanto, os vocabulários controlados, norteiam a atividade de indexação através do controle de sinônimos, de forma padronizada (com remissivas para todas as outras entradas), da diferenciação de homógrafos e da ligação entre termos que apresentam entre si uma relação mais estreita.

Embora o uso de vocabulários controlados possibilite uma representação mais coerente e mais próxima do item informacional, deve-se estar atento para não

comprometer a especificidade do texto em função da generalidade dos termos utilizados, representando fielmente seu conteúdo.

3.2.2 Ambigüidade ou imprecisão

A linguagem controlada tem por objetivo controlar as ambigüidades do léxico comum a partir de um conjunto controlado de termos que estabelecem relações hierárquicas e relações de vizinhança. (VAN DER LAAN, 2002, p.35)

Desta forma, um termo deve expressar conceitos determinados segundo suas relações com outros termos, de forma que estas unidades (termos) possam representar a temática de um documento. (CINTRA et al., 2002; LANCASTER, 2003)

Na linguagem controlada, o descritor tem por objetivo representar conceitos evitando ambigüidades. No caso de tesouros, a ambigüidade dos termos se reconhece quando necessitam de contexto adequado ou notas explicativas para seu emprego ou, ainda, interpretação correta, que orientem o indexador para o termo mais específico ou adequado para representar o conteúdo temático de um documento.

Em um vocabulário controlado, o uso de notas explicativas e de escopo, serve para diminuir ou neutralizar a multiplicidade de significações para que se possa, de maneira efetiva, obter uma interpretação adequada por parte do indexador, que corresponda ao sistema de conceitos de determinado domínio disciplinar do conhecimento.

3.2.3 Disponibilidade de instrumentos auxiliares

Os instrumentos auxiliares de indexação, como o próprio nome diz, servem de ferramentas ao indexador, na medida em que permitem ao mesmo utilizar-se de recursos extras para compreensão da temática abordada, como por exemplo, o uso de dicionários técnicos e glossários especializados.

3.2.4 Qualidade da estrutura e qualidade do vocabulário

O vocabulário controlado deve contar com uma estrutura suficientemente completa. No caso de tesouros, por exemplo, as abreviaturas TG(termo genérico)/TE (termo específico)/TR ou TA (termo associado) serve para guiar o indexador até o termo mais adequado para representar determinado tópico. (LANCASTER, 2003)

Quanto a algumas das funções da linguagem de indexação, estão as de substituir a grande variedade de expressões da linguagem natural por uma linguagem formal; estabelecer um ponto comum entre a linguagem do autor, a linguagem do indexador e a linguagem do usuário; estabelecer um vocabulário que mostre as relações entre termos e conceitos: relações semânticas (equivalência, hierarquia e associação) e relações sintáticas (coordenação de dois ou mais elementos para formar assuntos compostos). (NAVES, 2004)

Abaixo, serão mencionadas algumas especificações de critérios, que podem ajudar ao indexador a refletir sobre a qualidade dos tesouros. VAN DER LAAN (2002) recomenda que, antes da utilização indiscriminada de um tesouro, o

indexador atente para as seguintes informações, que poderão oferecer um panorama qualitativo do mesmo, que consiste da análise:

das informações gerais: este campo compreende a indicação do propósito do tesouro, definição do campo temático abrangido, indicação da equipe de trabalho, indicação de outros tesouros na área, indicação da norma de elaboração de tesouros, o tipo de pesquisa preliminar em bibliografias e periódicos especializadas, a sua sujeição e apreciação por órgãos e elementos especializados, a indicação de temas periféricos, ou seja, das áreas de especialidade complementares ao tema central e o grau de especificidade com que foram trabalhados, indicação da política de atualização do tesouro, indicação do número total de temas e subtotais dos descritores e não descritores, relação de data de inclusão do último termo de indexação. Todas as decisões tomadas na fase de planejamento e elaboração do tesouro. Essas informações introdutórias devem permitir uma clara compreensão da temática central e as periféricos.

das informações sobre a estrutura: compreende as explicações, definições e exemplificações sobre convenções e abreviatura a fim de evitar qualquer possibilidade de interpretação equivocada, a explicação de todos os sinais de pontuação utilizados, a explicação sobre notas a fim de prestar esclarecimentos sobre o significado do descritor. Já as notas de escopo ou qualificadores especificam os descritores para que não apresentem um sentido vago ou ambíguo.

das informações sobre os descritores: essas informações fornecem dados qualitativos sobre a elaboração dos tesouros e fornecerão subsídios para sua manutenção. Aqui devem apresentar-se a indicação do método de recolha dos candidatos a descritor sendo ou não eles recolhidos de fontes atualizadas e conceituadas, passando após pelo exame de especialistas da área.

Deverão também indicar das principais fontes de coleta dos candidatos a descritor e os critérios de determinação do descritor preferido seja pela frequência em que os termos foram utilizados ou mencionadas nas novas terminologias de cada área. Indica-se, se for o caso, a utilização de dicionários e glossários especializados e fontes de referência atualizadas para determinação dos mesmos ou os critérios de escolha a que ficaram a cargo dos especialistas de cada área.

Também deverão estar contidas informações que indiquem a forma de validação dos descritores, a indicação de especialistas, consultores ad hoc, para validação dos descritores que fornecem a garantia literária e as explicações referentes a às relações de equivalência entre o descritor preferido e os não-descritores.

das informações sobre a organização conceitual: explicação sobre as diferentes relações hierárquicas e como visualmente são reconhecidas através das siglas convencionadas para seu reconhecimento e das diferentes relações associativas.

Todas as redes relacionais devem estar devidamente especificadas, permitindo ao indexador a representação de conceitos dentro dos domínios de especialidade a partir das relações lógicas (genéricas, específicas, analíticas e oposicionais), ontológicas (partitivas de sucessão, material-produto) e de efeito (causalidade, instrumental, de descendência) presentes nas classes de conceitos

(cadeias e renques). Estas indicações e explicações oferecem forte indício sobre a qualidade do tesouro.

Deve-se sempre estar atento à qualidade do vocabulário. Quanto mais interdisciplinar, flexível e abrangente for a fundamentação teórica a que recorrem os especialistas na sua elaboração, tanto mais próximo estará de servir aos objetivos da indexação.

Lancaster (2003, p. 90) afirma que a especificidade do vocabulário torna mais minuciosos “os matizes de significado que permite expressar”, de modo que estes matizes dificultam o estabelecimento de “[. . .] diferenças entre termos muito afins” e o “emprego de termos de modo coerente”. Adverte também, sobre o emprego de elementos adicionais (subcabeçalhos ou indicadores de função) que “[. . .] aumentam a especificidade e complicam o trabalho de indexação.”

Nos tesouros terminológicos, existe uma grande preocupação, por parte dos especialistas, com a precisão dos termos, e a relação entre conceitos que esses termos representam. Por esse motivo podemos considerar mais pertinente sua utilização, uma vez que abarcam os aspectos multidisciplinares do conhecimento.

3.3 Fatores Ligados ao Documento

Os fatores ligados ao documento são bem mais complexos do que podem parecer inicialmente, pois envolvem não só as variáveis presentes na leitura, como também as tipologias discursivas, a estrutura textual, etc.

3.3.1 Conteúdo temático

A temática envolvida em um documento pode tornar mais difícil a compreensão do assunto. Lancaster (2003, p. 90) atribui isto, ao “[. . .] grau de ‘correspondência’ entre o conteúdo temático do documento e o conhecimento ou interesses do indexador.”

Esse fator, diz respeito à temática e aos conceitos tratados no texto, no qual o indexador pode demonstrar ou não, interesse pela temática envolvida. Portanto, o cuidado a tomar, é não deixar que seu desinteresse afete o trabalho de análise.

Se o indexador não possui conhecimentos prévios a respeito da temática envolvida em um documento, que sejam suficientes o bastante, para permitirem a compreensão do assunto, isto poderá refletir negativamente no trabalho de indexação.

3.3.2 Complexidade

Quanto mais habilidade e familiaridade acerca de tipologias e estruturas textuais possuir o leitor, mais facilidade terá na busca por compreensão. (Cintra⁸,1987; Kato⁹,1986; Kobashi¹⁰, 1994 *apud* FUJITA, 2004, p.6).

⁸ CINTRA, A. M. M. Estratégias de leitura em documentação. In: SMIT, J. W. (Coord.). **Análise documentária: a análise da síntese**. 2.ed. Brasília: IBICT, 1987. p. 29- 37. *apud* FUJITA, 2004, p.6.

⁹ KATO, M. A. **No mundo da escrita: uma perspectiva psicolinguística**. São Paulo: Ática, 1986. *apud* FUJITA, 2004, p.6.

¹⁰ KOBASHI, N. Y. **A elaboração das informações documentárias: em busca de uma metodologia**. 1994. 195f. Tese (Doutorado em Ciências da Comunicação) – Escola de Comunicação e Artes, Universidade de São Paulo, São Paulo, 1994. *apud* FUJITA, 2004, p.6.

Visto por esse ângulo, o conhecimento desse “código de comunicação” lingüístico, permite a decodificação do texto pelo leitor, por ser o texto um “produto lógico do pensamento” do autor a ser captado pelo leitor. (Giasson¹¹, 1993; Koch¹², 2002 *apud* FUJITA, 2004, p. 6-7).

Quanto maior for o conhecimento do indexador acerca das estruturas textuais no momento da análise documentária, mais próximo estará o indexador de extrair seu conteúdo temático. As estruturas estão associadas ao “[. . .] modo como as idéias são organizadas no texto [. . .]”, ou seja, ao conteúdo, ao tema e aos conceitos tratados. (FUJITA, 2004, p. 8)

Van Dijk¹³ (1992 *apud* FUJITA, 2004, p. 7-8) considera, que no momento da leitura, o leitor procura a informação importante. O autor as divide em duas categorias: informação **textualmente importante** e **contextualmente importante**. A primeira, é a considerada pelo autor, enquanto que a outra, é a considerada pelo leitor.

No mesmo sentido, Cavalcanti¹⁴ (1989 *apud* FUJITA), descreve o princípio de relevância na comunicação, que denomina de “saliência-autor/relevância-leitor”. A saliência-autor é entendida como as idéias que o autor salienta no texto, ou a tematicidade intrínseca. A relevância-leitor diz respeito às que são consideradas mais importantes pelo leitor no momento da leitura, ou tematicidade extrínseca.

Certamente, a identificação das estruturas contidas em um texto, amplifica a compreensão da leitura na análise temática, enquanto permite ao leitor-indexador,

¹¹ GIASSON, J. **A compreensão na leitura**. Lisboa: ASA, 1993. *apud* FUJITA, 2004, p.6-7.

¹² KOCH, I. G. V. **Desvendando os segredos do texto**. São Paulo: Cortez, 2002. p.16. *apud* FUJITA, 2004, p.6-7.

¹³ VAN DIJK, T. A. **La ciência del texto**: um enfoque interdisciplinário. Barcelona: Paidós, 1992. *apud* FUJITA, 2004, p. 7-8.

¹⁴ CAVALCANTI, M. C. **I-n-t-e-r-aç-ã-o Leitor-texto**: aspectos de interpretação pragmática. Campinas: UNICAMP, 1989. 271 p.

extrair a idéia principal do texto, levando em conta o gênero e os diferentes tipos de texto.

[. . .] para entender um texto os leitores devem ser conscientes de que devem buscar uma estrutura organizativa do texto, descobrir distintas chaves de estrutura e saber como modificar o marco de referência à medida que avançam, até que tenham construído uma representação do texto **de acordo com seus propósitos**. (CALFEE; CURLEY¹⁵, 1997 *apud* FUJITA, 2004, p. 8, grifo nosso)

Além das estrutura do texto, as superestruturas descritas por Fujita (2004, p.8) como “esquemas abstratos que estabelecem a ordem de um texto”, devem ser analisadas durante a leitura documentária.

A combinação da superestrutura organizativa de um texto, elaborada por seu autor, servem ao indexador, na medida em que permitem a este, inferir através da organização interna do texto e a assimilação de seu conteúdo temático, a informação retirada do texto. (FUJITA, 2004)

O conhecimento da estrutura textual permite ao indexador, criar estratégias que ofereçam maior facilidade na compreensão do conteúdo temático, na mesma proporção em que diminuem a dificuldade gerada por sua complexidade, acarretam em maior domínio por parte do indexador e permitem maior coerência e precisão na análise de assuntos.

3.3.3 Língua e linguagem

Uma das grandes dificuldades na análise temática gira em torno da linguagem utilizada por alguns especialistas, que dificulta a compreensão do texto.

¹⁵ CALFEE, R. C.; CURLEY, R. **Estructuras discursivas en las diferentes areas del conocimiento**. P. 53-80. In: RODRIGUES, Emma. LAGER, Elisabeth (comp.). La lectura. Santiago de cali: Editorial Universidad del Vale. 1997. 225p. *apud* FUJITA, 2004, p. 8.

Um autor pode usar a linguagem de forma mais clara que outros para expor a mesma idéia, o que simplifica o trabalho de indexação. (LANCASTER, 2003, p. 91)

Neste ponto, podemos considerar o problema de como a linguagem é utilizada pelo autor. Essas dificuldades, no momento da análise, são muitas vezes originárias da capacidade do autor em comunicar o conteúdo informacional de forma erudita ou rebuscada. Sem contar aqueles que têm a incrível aptidão em utilizar a linguagem de maneira tão articulada e tão inovadora, que o produto informacional acaba sendo o que podemos chamar de informação-parábola ou informação-charada, e além de tudo, ainda requerem do indexador alta capacidade de decifração. Além de tudo isso, um documento publicado em uma língua ou idioma que o indexador não domina gera grandes dificuldades no momento da representação temática.

3.3.4 Extensão

Lancaster (2003, p. 89) somente refere-se à extensão do documento, não explicitando o que quer dizer com isso. Podemos considerar que a extensão a que o autor se refere, seja o número de volumes de uma obra o que pode tornar cansativa e complicada a análise do mesmo, uma vez que o autor classificou-a como fator ligado ao documento.

3.3.5 Apresentação e sumário¹⁶

Por considerarmos o termo utilizado por Lancaster (2003, p.89) – “sumarização” - terminologicamente impreciso, optamos por utilizar o termo sumário. O autor também não explicita o que entende por “apresentação e sumarização”.

A partir disso, podemos inferir que o autor se refere a esses dois fatores, por proporcionarem, tanto o sumário como a apresentação do texto, um esboço do panorama temático encerrado no documento, dando uma idéia, ainda que muito diminuta, sobre o que trata o documento e algumas vezes, da perspectiva abordada pelo autor, os limites dos domínios de especialidade aos quais está circunscrita a temática quanto ao reconhecimento dos pontos de vista analisados sob a perspectiva disciplinar do autor.

Na apresentação, muitas vezes, é explicitada as justificativas e as circunstâncias de produção do texto, como por exemplo, a necessidade de mudanças paradigmáticas frente à sociedade, novas descobertas no campo da ciência em questão, soluções metodológicas ou teóricas na resolução de determinados problemas, entre inúmeros outros.

3.4 Ligados ao Processo

Os fatores interferentes na qualidade da indexação que estão ligados ao processo são: o tipo de indexação, as políticas de indexação, a produtividade exigida e a exaustividade da indexação.

¹⁶ No original, o termo utilizado por Lancaster foi sumarização.

Quanto mais claras e forem as diretrizes e os critérios adotados no centro de informação como as políticas de indexação, política de desenvolvimento de coleções, manuais de serviço, entre outros, maior qualidade conferem aos objetivos do centro de informação para o atendimento das demandas.

A falta de políticas claramente definidas interfere na indexação. Quando não estão devidamente especificadas e explicitadas, ficam a mercê de subjetividades administrativas e gerenciais fundamentadas no “achômetro” que, longe de atender às demandas sociais e individuais, desvinculam-se dos propósitos dos centros de informação em seu conjunto.

3.4.1 Tipo de indexação

Existem dois tipos de indexação: por extração e por atribuição. Na indexação por extração utilizam-se os termos que o autor emprega para expressar suas idéias, utilizando-se a linguagem natural e, na indexação por atribuição os termos são selecionados a partir de um vocabulário controlado que estabelece distinção entre homônimos, controle de sinônimos, relações de equivalência, etc.

3.4.2 Política de indexação

A utilização de uma política de indexação para nortear as atividades do indexador, no momento da representação temática, é de máxima importância,

servindo para estabelecer a forma em que se dará a representação temática para alcançar maior coerência entre os termos indexados.

A política de indexação é uma decisão de ordem administrativa condicionada à existência de condições como infra-estrutura, recursos humanos e objetivos institucionais, que conforme Fujita (2004, p.20), “[. . .] deverão oferecer a identificação de condutas e procedimentos de análise e recuperação de informação.”

Para alcançar a qualidade na indexação é essencial que em qualquer centro de informação se adote uma política de indexação que sirva como diretriz, no sentido de formalizar procedimentos que deverão ser observados na indexação.

Em acordo com a assertiva supra citada, Nunes (2004, p.57) a complementa ao afirmar que:

[. . .] com um pouco de reflexão, o bibliotecário saberá articular os argumentos apropriados para sustentar a decisão de se formalizar - e pôr em prática – uma política de indexação, compreendendo o estabelecimento de princípios e critérios que balizarão as decisões que o bibliotecário precisa tomar quando do cumprimento do trabalho rotineiro de indexação.

Longe de gerar dificuldades na indexação, a formalização desses procedimentos garante sua consistência. Isto não quer dizer, que seja desnecessária a permanente atualização de suas políticas de acordo com as mudanças organizacionais e a evolução do conhecimento humano. Segundo o mesmo autor, evitará “[. . .] que sofram demasiado a influência pessoal do bibliotecário.” (NUNES, 2004, p.57)

A cobertura de assuntos centrais e periféricos, o grau de profundidade da indexação dos diferentes domínios, a linguagem de indexação utilizada pelo sistema, também deverão ser dispostos na política de indexação.

Nunes (2004, p.58) ainda afirma que no caso da utilização de tesouros, por não serem gerais como os cabeçalhos de assuntos, “[. . .] será necessário recorrer a tantos tesouros quanto necessário (sic) forem para dar conta da totalidade de assuntos.”

A forma de atualização da linguagem adotada também deve estar fixada na unidade de informação pelo que o autor denomina de catálogo decisório. Este catálogo consiste no registro de todas as decisões tomadas no cotidiano da unidade de informação, sendo que as mesmas, de tempos em tempos, devem ser revisadas e atualizadas, para garantir o controle e a consistência das decisões relativas ao processo de indexação, pois atuam de forma direta sobre os níveis de exaustividade, especificidade, revocação e precisão do sistema.

Ponto em que concorda Fujita (2004, p. 21), ao destacar a imprescindibilidade na definição da política de indexação, do número de descritores a serem extraídos do documento (baseada no critério de exaustividade) e da indexação de conceitos selecionados, mediante o emprego de um termo, que seja o mais específico possível, na linguagem de indexação adotada (baseada no critério de especificidade).

Nunes (2004, p. 59) acrescenta que a revocação e a precisão comportam-se “inversamente condicionadas pela exaustividade e especificidade. [. . .] sendo que o estabelecimento de seus parâmetros considerados eficientes e eficazes para uma biblioteca resultará de sucessivos ensaios [. . .].”

Isto quer dizer, que quanto maior a exaustividade na indexação de documentos, maior o número de documentos recuperados (revocação) e menor a precisão, o contrário do que acontece com a especificidade, quanto maior esta, menor a revocação.

Rubi e Fujita (2003, p. 69) esclarecem que “[. . .] a indexação pode ser observada em dois momentos distintos dentro do sistema: na entrada – no tratamento temático da informação – e na saída – na busca e recuperação da informação.”

Nunes (2004, p. 58) confirma esta observação ao afirmar que:

[. . .] é conveniente jamais esquecer que a indexação de um documento é feita uma única vez enquanto que a recuperação da informação contida em um documento é feita inúmeras vezes, do que se conclui que o tempo ‘gasto’ na indexação, é de fato, economia de tempo para inúmeros usuários no futuro.

Lancaster (2004, p. 89) apresenta como uma das variáveis interferentes na qualidade da indexação os fatores ligados ao “processo”, no qual se inclui a política de indexação. Diferentemente de Fujita (2004, p. 22) que a inclui no contexto sociocognitivo do indexador que se refere:

[. . .] ao modelo mental do leitor em que estamos supondo existir o processo de análise de assunto, a linguagem documentária do sistema, a política de indexação do sistema e seu manual de serviço [. . .].

.....
O contexto sociocognitivo do indexador, circunscrito pelo conhecimento da situação comunicativa e de suas regras, está implícito no contexto de trabalho desenvolvido por indexadores em sistemas de informação: a política de indexação, as regras e procedimentos do manual de indexação, a linguagem documentária para representação e mediação da linguagem do usuário e os interesses de busca do usuário.

Do ponto de vista da autora, o contexto sociocognitivo serve para “[. . .] nortear a concepção de análise de assunto do indexador”. Portanto, não é só uma política de indexação mal elaborada que afeta a qualidade da indexação, mas também, a falta de conhecimento da “situação comunicativa” (leitor-texto-contexto) e suas regras (linguagem, política de indexação, manual de indexação, vocabulário, etc.). Estas mesmas regras, quando não estão bem formuladas geram dificuldades

para “corresponder às demandas do usuário” porque o indexador não as conhece ou não as compreende.

3.4.3 Produtividade exigida

Algumas instituições ou organizações criam o que podemos chamar de linha de montagem, exigindo que a “produção” alcance um número pré-estabelecido pelas unidades de informação.

Lancaster (2003, p. 91), a esse respeito, declara que “[. . .] se for exigido do indexador que ele dê conta de certos números de itens por dia, ele poderá sentir-se pressionado e isso levará a erros por descuido, especialmente se a instituição tiver uma expectativa excessiva de produção diária.”

Sendo assim, a indexação torna-se um processo de trabalho racionalizado e mecanizado, em que a qualidade está fadada ao fracasso. Qualidade demanda tempo, capacitação, habilidade, concentração, desenvolvimento de métodos que incluem simultaneamente, uma rede de relações empiricistas (demanda, conhecimento dos interesses do usuário, conhecimento temático, objetivos da instituição, etc) e teóricas.

3.4.4 Exaustividade da indexação

O nível de exaustividade é uma decisão de ordem administrativa, que deve estar claramente definido na política de indexação.

Lancaster (2003, p.23) define a exaustividade como sendo “[. . .] o emprego de termos em número suficiente para abranger o conteúdo temático do documento de modo bastante completo.”

Podemos complementar a assertiva acima, com a de VAN DER LAAN (2002, p. 25), que define a exaustividade como “[. . .] uma medida de extensão que se refere a identificação e posterior tradução para a linguagem de indexação utilizada pelo sistema de todos ou de um número significativo de tópicos temáticos abordados em um determinado item.” Em contraposição à seletividade, caso em que um item é indexado por tópicos principais considerados de maior relevância para a unidade de informação. Geralmente, recomenda-se que o número de descritores selecionados sejam tantos quantos forem necessários para atenderem o critério de exaustividade fixado na política de indexação, podendo variar de acordo com o tipo de biblioteca, o tipo de usuário, o tipo de documento, o objetivo da instituição, o tempo de resposta do sistema, o grau de revocação e precisão requeridos pelo sistema, etc.

Podemos evidenciar mais uma vez, a necessidade de uma boa política de indexação, dado que esta “[. . .] permeia todo o processo de indexação e determina os procedimentos a serem observados [. . .]” que qualificam o trabalho de indexação. (VAN DER LAAN, 2002, p. 27)

4 QUALIDADE DA REPRESENTAÇÃO E OS SISTEMAS DE CLASSIFICAÇÃO

Uma das dificuldades encontradas pelos profissionais da informação quanto à representação temática de documentos, está ligada ao fato de que os assuntos, nas tabelas ou esquemas de classificação, algumas vezes, não se adaptam aos assuntos (temática) contidos e identificados em documentos, criando desta forma, problemas na indexação dos mesmos. Por conseqüência, acabam afetando a qualidade da indexação, pois não mais acompanham e comportam o aspecto interdisciplinar do conhecimento.

Alguns exemplos podem ser mencionados, como os que estão relacionados às inúmeras modificações e sucessivas edições dos Sistemas de Classificação Bibliográficos Universais (CDD e CDU), que vem sofrendo progressivamente, um afastamento das funções e objetivos iniciais de segmentação (divisão) do conhecimento e, conseqüentemente, dos efeitos decorrentes da ausência de uma alteração radical nestes sistemas.

Também podemos citar os problemas decorrentes da utilização de listas de cabeçalhos de assuntos que não formam um princípio organizador uno, mesmo quando apresentadas em forma de tesouro. Por não possuírem estruturação de conteúdo, controle de sinônimos, distinção entre relações hierárquicas e associativas e garantia de que o sistema conceitual das disciplinas do conhecimento sejam representados, o quê resulta em um agrupamento de palavras, sem a determinação do sentido de cada termo.

Os critérios baseados no uso de parâmetros estatísticos ou sistemas de indexação automáticos, feitos com base na freqüência e ocorrência de termos, resultam em uma classificação arbitrária e, por esse motivo, não podem ser

considerados eficazes, principalmente na área das ciências humanas, visto não possuírem terminologia consolidada e alto grau de formalização terminológica.

Moreira, Alvarenga e Oliveira (2004), declaram que critérios baseados em parâmetros sintáticos e estatísticos, ainda estão muito distantes de conferir qualidade aos sistemas de indexação. Diferentemente dos critérios baseados em aspectos semânticos da linguagem que “[. . .] de modo geral ainda não são possíveis devido à complexidade de se obter a semântica de uma palavra, de uma unidade lexical, em um determinado contexto.”

Os problemas decorrentes da ausência de critérios explicitados em alguns tesouros, tais como os de inclusão de descritores e seus significados bem claros e sobre a organização conceitual da área, sinalizam que talvez os mesmos não tenham sido baseados em princípios teórico-metodológicos específicos.

O conhecimento por parte do indexador, dos problemas na utilização de tais instrumentos, serve para que o mesmo se utilize de instrumentos mais eficientes na representação, que sejam flexíveis o bastante para acompanhar o caráter mutante da informação e represente mais fielmente os aspectos interdisciplinares do conhecimento.

No entender de Moraes e Arcello (2000), todas essas dificuldades são conseqüências da influência das correntes filosóficas e dos pressupostos teóricos no contexto histórico em que tais esquemas ou sistemas de classificação foram produzidos.

De forma diferente, Ginez de Lara (2002, p. 133, grifo nosso), vê o processo de construção da informação documentária “[. . .] como resultado da segmentação de conteúdos feitos a partir de **hipóteses de organização**” realizadas através da

analogia e da generalização. Afirma ainda a autora, que a informação é sempre uma construção visto que:

As generalizações a que procedemos são um produto cultural, fruto de nossa experiência com a realidade.

Enquanto produtos culturais as classificações (e generalizações) partem necessariamente de hipóteses; podem existir concomitantemente distintas hipóteses, que privilegiam determinadas características em detrimento de outras. [. . .] qualquer proposta de classificação universal é datada, isto é, parcial, momentânea e centrada num ponto fixo de enunciação. Isso explica porque as classificações bibliográficas de natureza enciclopédica¹⁷ e quase exclusivamente hierárquicas oferecem dificuldade para dar conta de distintos aspectos não contemplados de início. Tais propostas classificatórias têm uma articulação dura entre o plano do conteúdo e o plano de expressão e funcionam exclusivamente a partir de operações de encaixe.

Moraes e Arcello (2000, p. 118) atribuem a crise dos sistemas de classificação bibliográfica às “[. . .] suas bases epistemológicas. Há que se aceitar a razão como dinâmica e histórica, transformada dia a dia, abandonando o determinismo universal.” Afirmam ainda que a solução:

[. . .] está em trabalhar com representação de forma interdisciplinar, ou seja com epistemologia, lingüística, ciência da informação, comunicação, informática e também com a estrutura das ciências e a organização do conhecimento como um todo. (MORAES; ARCELLO, p. 119)

A inflexibilidade, a insuficiência estrutural, a rigidez na representação do conhecimento e o contexto epistemológico em que os sistemas de classificação tradicionais foram elaborados, podem ser apontados como uma das causas da crise desses sistemas. Tendo o conhecimento múltiplas dimensões, estas devem ser analisadas e representadas com o uso de instrumentos que comportem sistemas conceituais, lingüísticos, semióticos, baseados em conceitos teórico-metodológicos de análise discursiva e lógico-semânticas representados *no documento*.

¹⁷ **N.A** O entendimento de enciclopédias neste contexto, refere-se à pretensão universalizante de abarcar, num só sistema, todo o universo do conhecimento. A enciclopédica entretanto, é um postulado semiótico: o que se consegue registrar é sempre parcial frente à infinita possibilidade de interpretação enciclopédica.

5 QUALIDADE DA REPRESENTAÇÃO E O CARÀTER MUTANTE DA INFORMAÇÃO

O conhecimento por parte do profissional da informação, dos aspectos lingüísticos do texto e a compreensão dos mecanismos lógicos e semânticos envolvidos na análise de um documento proporcionam maior qualidade à representação. Na medida em que o profissional da informação apropria-se das idéias do autor e as sintetiza, com a utilização de aparatos teóricos de áreas correlatas, para fins de indexação, passa a analisar as variáveis presentes na produção de um texto e qualifica o trabalho profissional.

A subjetividade por parte do indexador, não deve ser utilizada para justificar possíveis falhas na recuperação de informação. Deve-se recorrer a fundamentações teórico-metodológicas que possibilitem ao indexador, a análise de documentos com a menor intervenção possível de juízos de valor não descritos pelo autor, buscando sempre que possível, utilizar a objetividade científica em favor do estabelecimento de relações conceituais entre os registros do conhecimento, principal função da representação temática, possibilitando à ciência sempre novas descobertas.

Nunes (2002, p. 187) alerta a respeito da crença na univocidade da representação que:

A idéia de que seja possível precisar inequivocadamente o sentido de um termo significa aprisionar-lhe o sentido, impedindo que o mesmo continue tendo vida, isto é, venha a incorporar variações semânticas decorrentes do próprio processo de investigação e produção de novos conhecimentos. Produz-se assim um paradoxo, porquanto, com o objetivo de assegurar a comunicação entre os cientistas, constrói-se uma linguagem pretensamente unívoca, a qual, se atende (ainda que incompletamente) ao desiderato atual por homogeneidade de significado, por outro lado deixa de atender a necessidade futura de heterogeneidade, imprescindível ao progresso da ciência.

.....
Se há incerteza no conjunto do conhecimento científico estabelecido e, mais se a parcela dos resultados de pesquisa comunicados no interior da comunidade científica tipificados como irrelevantes ou incorretos é

absurdamente majoritária (estimada por alguns autores em 90% de tudo que se publica), é muita pretensão exigir-se ou esperar-se que as linguagens documentárias alcancem a univocidade para elas preconizada como condição de sua eficácia e eficiência, enquanto ferramentas intermediárias do processo de produção-consumo de informações científicas, vale dizer, da circulação do conhecimento científico no interior do sistema tecno-científico.

O ideal de univocidade na representação temática e conseqüentemente da informação, seu objeto de estudo, não pode ser analisado sob uma perspectiva unívoca sendo esta diametralmente oposta ao caráter dinâmico e multidisciplinar do conhecimento. Tal tentativa incorre em um contra senso, ou seja, não se pode representar unidades dinâmicas (conceitos), a partir de arranjos estáticos que desconsiderem as variações semânticas de termos consagrados pelo uso e que dentro do contexto cultural, histórico, ideológico, religioso ou sociológico carregam consigo um sentido próprio podendo ser satisfatórias somente em um **determinado ponto no tempo**.

Qualquer tentativa para solucionar o problema da representação temática, baseada em teorias que não contemplem o conhecimento como um todo, que não sofram atualizações contínuas a partir de um processo cooperativo entre cientistas e profissionais da informação, numa velocidade compatível à verificada no desenvolvimento da ciência, com linguagens que disponham de mecanismos de ajuste (supressões, acréscimos, atualizações que contemplem mudanças de paradigmas) ocorridas nos ramos do saber, jamais poderão ser consideradas satisfatórias na indexação.

Neste sentido Ziman¹⁸ (1996, *apud* NUNES, 2000, p. 189) sobre a efetividade da comunicação do conteúdo dos conhecimentos científicos e os demais conhecimentos acumulados pela humanidade explica:

¹⁸ ZIMAN, J. **O Conhecimento Confiável**: uma exploração dos fundamentos para a crença na ciência. Campinas: Papyrus, 1996. 252 p. *apud* NUNES, 2000, p. 189.

[. . .] que seu conteúdo é **consensível**. [. . .] cada mensagem não deve ser obscura ou ambígua a ponto de seu receptor ficar incapacitado de dar-lhe um consentimento sincero ou opor-lhe objeções bem fundamentadas. A meta da ciência, além disso, é alcançar o máximo de **consensualidade**. Idealmente, o corpo geral do conhecimento científico deve consistir de fatos e princípios firmemente estabelecidos e aceitos, sem dúvidas sérias, por uma maioria esmagadora de cientistas competentes e bem informados. Tal como veremos, convém fazer uma distinção entre uma mensagem **consensível** com **potencial** para vir a contribuir para um consenso, e uma afirmação **consensual** plenamente testada e objeto de um acordo universal. Pode-se dizer, com efeito, que a consensibilidade é uma condição necessária para qualquer comunicação científica, ao passo que apenas uma pequena proporção de todo o corpo da ciência é inegavelmente consensual em qualquer momento dado. (grifo do autor)

Cabe dizer que a univocidade pretendida na representação encontra grandes dificuldades, como no caso de teorias ou hipóteses que trabalham com as fronteiras ou limites entre as áreas do conhecimento, devido à utilização, por parte dos cientistas, de uma linguagem demasiadamente variável.

A esse respeito Nunes (2000, p. 190) afirma:

[. . .] a dificuldade surge no momento em que a univocidade pretendida é estabelecida convencionalmente, enquanto que o consenso é formado socialmente, isto é, ele existe, porém não é o resultado de qualquer procedimento formal, contrariamente ao que ocorre com o processo de formação de uma linguagem documentária. [. . .] o conceito não é estático; ao contrário, poderá ser substituído por um novo consenso - e de fato o será, tão logo deixe de ser útil, eficaz para o equacionamento dos problemas[. . .].

Só pode ser utópica a expectativa de alguns autores em utilizar uma linguagem documentária universal para representar uma linguagem variável, sem que as Linguagens Documentárias acompanhem a terminologia empregada na comunicação de conteúdos científicos estabelecidos e aceitos na comunidade científica.

Sendo seus membros sujeitos principais deste processo, que se abastecem da informação representada pela linguagem documentária e fornecem subsídios para sua elaboração, numa cadeia de retroalimentação, tanto na sua utilização como na sua representação, decorre disso, que é através da cooperação e interação entre

os mesmos, que se pode criar um processo efetivo, eficiente e eficaz de comunicação que contribua no progresso da ciência e no processo de produção e representação do conhecimento (autocorretivo como a própria ciência).

A linguagem científica (terminologia) deve acompanhar o avanço científico, ampliando ou limitando a área conceitual, considerando-se a evolução do significado dos termos na representação, que acompanhem as modificações, alterações, inter-relações ocorridas nas teorias e em hipóteses que trabalham com as fronteiras do conhecimento, contribuindo no processo de comunicação entre especialistas. Isto pode ser confirmado, tomando por empréstimo a afirmação abaixo:

A terminologia surge da necessidade de denominar os sistemas de conceitos das diferentes disciplinas, com o objetivo de permitir uma comunicação eficiente entre especialistas. Este objetivo, ao ser atingido, supõe a obtenção de outros não menos importantes: a elaboração de uma terminologia da Ciência da Informação, contemplando conceitos próprios e de empréstimo, permitirá reconhecê-la na sua autonomia. Na prática, significa estabelecer a linguagem da área.

.....

Em larga medida, portanto, o conhecimento e a compreensão de uma área de conhecimento, vinculam-se ao domínio da linguagem desta mesma área. O núcleo específico de uma linguagem de especialidade é seu vocabulário, que normalizado e organizado semântica e logicamente constitui a terminologia da área. (KOBASHI; SMIT; TÁLAMO, 2001)

À mesma conclusão chega Nunes (2000, p. 102) ao afirmar que as contribuições da terminologia à análise documentária podem advir "dos empréstimos teóricos e metodológicos [. . .] de controle da linguagem natural tal como empregada em contextos comunicacionais de especialidade."

Cabré¹⁹ (1995, *apud* NUNES, 2000, p. 100) declara que o objeto de estudo da terminologia está centrado nas unidades terminológicas da Linguagem Natural (termos) de um domínio disciplinar possuindo três abordagens de pesquisa: dos

¹⁹ CABRÉ, M. T. La terminología hoy: concepciones, tendencias y aplicaciones. **Ciência da Informação**, v. 24, n. 3, p. 289-298, 1995. *apud* NUNES, 2000, p. 100.

domínios, filosófica e lingüística. Enquanto que a primeira está centralizada em refletir um conceito a partir de suas relações com os conceitos de outros domínios de especialidade e, também, na relação entre o conceito-termo estabelecendo uma relação interdisciplinar, a segunda ocupa-se da classificação de conceitos em categorias e em estabelecer a relação entre terminologia e documentação. Já a terceira, está baseada na terminologia como subconjuntos do léxico de uma língua de especialidade (sublinguagens).

Portanto, o interesse da abordagem filosófica para a Análise Documentária, está na fundamentação teórica e metodológica que esta pode oferecer para o estabelecimento de conceitos e termos unificados, com a finalidade de normalizar o uso destes termos, em função do nível de desenvolvimento do conhecimento em qualquer um domínio de especialidade, quer seja na sua construção de uma linguagem documentária ou na sua manutenção.

Cabré (1995, *apud* NUNES, 2000, p. 49) considera que o objeto da terminologia e o da análise documentária, historicamente falando, são diametralmente opostos. Enquanto a primeira está mais vinculada ao consumidor da informação, a segunda encontra-se mais vinculada ao produtor, revelando o que denomina de tensão entre paradigmas. Acredita ser este o desafio da análise documentária, equilibrar as várias funções, aparentemente opostas, não só para atender aos objetivos da indexação, mas também para que se faça uma reflexão acerca de contribuições metodológicas levando em consideração essas tensões para emprestar maior consistência às fundamentações teórico-metodológicas das atividades documentárias. Estas fundamentações, tomadas por empréstimo dos princípios teórico-metodológicos da terminologia, ainda não romperam os liames do

tradicionalismo das antigas estruturas nas quais exercia sua atividade. (NUNES, 2000, p. 106)

Esquemáticamente podemos resumir essa tensão de paradigmas da linguagem natural e da linguagem documentária na seguinte ilustração:

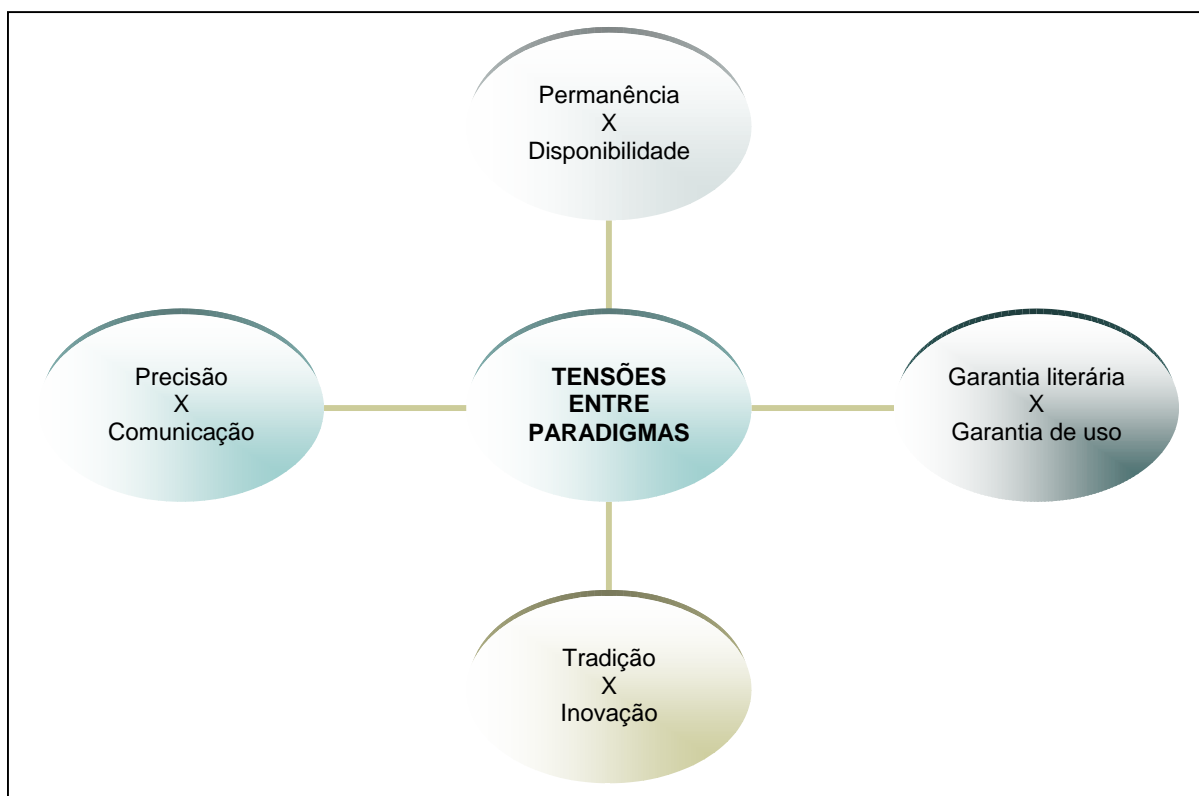


Ilustração 3 - Tensões entre paradigmas da linguagem natural e da linguagem documentária ²⁰

A permanência é descrita por Nunes (2000, p.103), como a “necessidade de preservar a informação em seu estado original”, enquanto que a disponibilidade “implica uma condição dinâmica, de mutação”, devido às variações decorrentes do nível de desenvolvimento do conhecimento, que quanto maior for este, maiores as variações em tempo mais curto.

²⁰ FONTE: NUNES, Cláudio Omar Iahnke. A relação da análise documentária com a terminologia. Biblos, Rio Grande, v.12, p. 99-113, jan./dez. 2000.

Tentar estabelecer um equilíbrio entre estas funções é um desafio para assegurar a eficácia do processo de comunicação entre o sistema de informação e o usuário.

Estas dificuldades podem ser minimizadas, e em parte solucionadas, com as contribuições que a terminologia vem oferecendo à análise documentária, a partir de seu corpus teórico-conceitual, na medida em que cria hipóteses de organização do conhecimento de acordo com o domínio de especialidade na qual faz parte, e que simultaneamente, impele a análise documentária a buscar suas próprias fundamentações teóricas, com a finalidade de organizar o conhecimento de acordo com as necessidades de seu consumidor, o usuário.

Podem até parecer antagônicas, as perspectivas do consumidor e do produtor da informação, mas não são, elas são complementares, correspondem a um ciclo.

O produtor pode ser, num momento subsequente, o consumidor e vice-versa. Cabe ao indexador, o desenvolvimento de habilidades, metodologias, capacitação e conhecimento dos instrumentos utilizados que melhor atendam a esse processo de comunicação.

Nunes (2000, p. 103) ainda reflete sobre as tensões existentes entre a garantia literária e a garantia de uso, explicando a diferença entre os tesouros e os sistemas de classificação, em que os primeiros explicitam:

[. . .] uma **segmentação**, baseada em hipóteses de organização do conhecimento. A estrutura dessas linguagens baseia-se na hipótese adotada para sua construção, o que é feito mediante o procedimento metodológico designado como **garantia literária**, cuja aplicabilidade implica um aporte fundamental das terminologias de especialidade. A garantia literária é complementada por um outro recurso metodológico, não menos importante, designado como **garantia de uso**, a qual, ao contrário do que sua designação possa sugerir não se limita a apenas "como o falante fala", mas a explicitar seu desempenho no contexto social em que o falante realiza o ato de fala, isto é, levando em consideração o conjunto de falantes de um domínio de especialidade ou de uma instituição em particular. (grifo do autor)

O ideal é que se possa estabelecer o equilíbrio entre estas duas garantias na elaboração da linguagem documentária. Enquanto que a garantia literária está mais ligada à tradição, à formalidade, a um modelo preponderante no âmbito científico, a garantia de uso, mais informal, oferece a possibilidade de inovação, de novos paradigmas como parte do processo revolucionário do qual faz parte a ciência e o conhecimento.

A tradição e a inovação são as outras duas tensões entre paradigmas. A tradição sendo rígida, reflete os próprios sistemas tradicionais de classificação bibliográfica, que como vimos até agora, não conferem qualidade ao processo de segmentação do conhecimento em seu nível atual. Por ser o conhecimento, dinâmico, mutável, interdisciplinar, etc, em cada domínio, com efeito, só pode ser segmentado sob a mesma perspectiva.

Enquanto que a informação documentária está circunscrita pela natureza da instituição da qual faz parte, por sua abrangência e competência informacional, na linguagem, o funcionamento do signo é uma questão individual, em que o sistema e o processo de comunicação da informação não estão internalizados pelo sujeito, mas, que em contrapartida, a circulação da informação documentária no interior de uma instituição:

[. . .]contribui para a redução das incertezas com que se defrontam os sujeitos, em decorrência das escolhas que realizam face à multiplicidade de opções comunicativas que se apresentam, imediatamente antes do momento em que as escolhas são feitas. Em lingüística, a redução de incerteza refere-se ao ato de que o ato da fala reduz a um conjunto de possibilidades antes que se dê escolha ao falante. (NUNES, 2000, p.106)

Muitos dos fatores de qualidade da indexação estão ligados à capacitação profissional do indexador, quer seja no conhecimento de teorias ou no emprego de metodologias de análise e representação do conhecimento que mais se aproximem da função social desempenhada pelo centro de informação.

“Dominar fragmentariamente o vocabulário” acarretará na queda de qualidade da indexação. “A terminologia realiza o controle da significação das palavras de um dado domínio de especialidade para viabilizar a comunicação e a transferência de informações” quer seja entre uma comunidade de estudiosos ou pesquisadores deste domínio, prevalecendo a opção pela precisão do significado atribuído aos termos de indexação. (NUNES, 2000, p. 107)

Esta precisão, em que se baseia a relevância no uso tesouros em sistemas de informação, de acordo com Moreira, Alvarenga, Oliveira (2004), está ligada ao fato de os mesmos permitem :

[. . .] determinar quais termos podem ser usados no sistema; determinar quais termos podem ser usados na busca para que esta tenha um resultado satisfatório; e permitir a introdução de novos termos em sua estrutura de termos e relações de modo a aproximar a linguagem do usuário à do sistema e realizar alterações de sentidos dos termos existentes.

Smit, Tálamo e Kobashi (2004) não são tão otimistas em relação às dificuldades a serem enfrentadas no âmbito da ciência da informação, pois afirmam estar baseada, em frágeis princípios teórico-metodológicos, como poderemos ver em sua assertiva:

[. . .] apresenta-se como um ponto de junção entre subjetividades que compartilham de uma prática comum. Não existe um acordo fundamental e uma linguagem de especialidade de natureza científica que comunique uma realidade integrada. A fragmentação conceitual, o recurso constante a terminologias de outras áreas, a ausência de projeto de consolidação acabam retardando o desenvolvimento da área.

As considerações feitas, até o presente momento, sobre o processo de produção do conhecimento e a interdisciplinaridade entre os campos do saber, podem ser facilmente observadas, em nossa própria área que, ao tomar por

empréstimo teorias de áreas afins, sinaliza que não construiu o seu próprio paradigma. Encontra-se ainda, ausente de fundamentações teóricas próprias que possam delimitar o domínio de especialidade ao qual pertence. É, pois, paradoxal trabalhar com limites do conhecimento, quando ainda não se conhece seus próprios limites, mas que vêm sendo solucionados em parte, através das contribuições que a terminologia vem oferecendo à área.

Segundo Kobashi, Smit e Tálamo (2001), a ausência de objeto teórico da chamada ciência da informação “[. . .] funciona como mero significante, o que torna difícil fazer a distinção entre o que lhe é próprio e o que lhe é acessório ou estranho [. . .]” quando:

Inadvertidamente, muitas vezes, utiliza-se o sinal que expressa o conceito, mas não o próprio conceito. O discurso torna-se vazio ou obscuro sem que o cientista social perceba que a sua linguagem pode dificultar a comunicação... Os conceitos metodológicos desprovidos de suas características limitar-se-ão a nomeações e classificações rituais de postura sem qualquer influência nas estratégias de investigação.

Os conceitos assimilados de outros domínios do conhecimento tornam-se vazios de significado, pois os esquemas conceituais e sistemas teóricos ao serem somente importados de outras áreas, não perdem sua identidade com o objeto teórico que lhe deu origem, não sofrem a devida reformulação a partir de seu novo objeto, que possui outras características, além daquelas que por definição, caracterizavam seu antigo objeto teórico. Ou seja, é como se importássemos rótulos e aviássemos receitas de xampu para fabricar sabonete. Podemos até utilizar o xampu como sabonete, mas ele não é o sabonete.

Neste sentido, a terminologia vem oferecendo contribuições na medida em que estabelece os limites conceituais de cada domínio sob o que Kobashi, Smit e Tálamo (2001) denominam como:

[. . .] pontos de vista específicos que orientam um modo próprio de explicar e interpretar a realidade. Na ausência das linguagens de especialidade, "sabemos o mundo" segundo o que nos dita a linguagem natural, inserindo-nos no senso comum. No entanto, não é apenas com o recurso à linguagem natural que se introduz a indeterminação conceitual. Saber o mundo através de um conjunto de termos sem consistência conceitual, provenientes, por exemplo, de diferentes áreas, na ausência de normalização, equivale a ter em mãos vários fragmentos que, se juntados, não fazem sentido ou o fazem à custa de muito esforço. Em larga medida, portanto, o conhecimento e a compreensão de uma área de conhecimento vincula-se ao domínio da linguagem desta mesma área. O núcleo específico de uma linguagem de especialidade é seu vocabulário, que normalizado e organizado semântica e logicamente constitui a terminologia da área.

6 CONSIDERAÇÕES SOBRE A ANÁLISE DOCUMENTÁRIA

A análise documentária, na visão de Fujita (2003, p. 61) é “uma combinação metodológica altamente estratégica entre o tratamento do conteúdo de documentos e sua recuperação por um usuário.” A autora sugere, que ao realizar a identificação de conceitos, o indexador o faça recorrendo a um “esquema de categorias existente na área coberta pelo documento”, por exemplo, fenômeno, processo, propriedade material, etc.

Lancaster (2003, p. 22) adverte, que o indexador não deve realizar a análise conceitual sob a influência do vocabulário, no momento exato em que faz a leitura para identificar e selecionar conceitos, por acreditar que isto poderá “impedir a compreensão do contexto do documento e sua adequada representação.”

Isto quer dizer que o indexador poderá deixar de extrair termos significativos que representam conceitos ou idéias, dentro da perspectiva ou visão de mundo do autor, porque o vocabulário não os contempla. Isto pode levar o indexador a realizar uma análise superficial de seu conteúdo, pois apenas ajusta as palavras do texto ao vocabulário e também acarreta a perda de compreensão do universo considerado, ou seja, o conteúdo do documento.

Também poderá fazer com que o indexador deixe de analisar a estrutura profunda do texto, por estar reduzindo as outras possibilidades de análise do texto, pois já representou anteriormente os conceitos extraídos do documento na fase de análise conceitual, o que deveria ser feito na fase de tradução. Na análise conceitual, as metodologias de representação informacional estão mais orientadas para o conteúdo do documento. Já na etapa de tradução, a representação está mais

orientada para a demanda, sendo, pois, complementares e não distintas, estando ligadas uma a outra por sua própria natureza, tanto a saliência-autor quanto a relevância-leitor. Desta forma, a indexação trata de traduzir e representar conteúdos informacionais para sua posterior disponibilização para o usuário. Assim serão atendidas às demandas do centro de informação, dentro da perspectiva das necessidades do usuário, que estão reduzidas em sua multiplicidade de opções comunicativas à natureza da instituição, sua missão, objetivos, etc.

Este ponto pode ser confirmado por Blair²¹; Hjörland²²; Soergel²³; Weinberg²⁴ (1990,1992,1985,1988 *apud* FUJITA 2003, p. 79) ao enfatizarem a função primária da indexação e recomendar que “[. . .] o indexador não deve direcionar a análise exclusivamente sobre o conteúdo, mas antecipar o impacto e o valor de um determinado documento para uso potencial”.

A distinção entre o conteúdo intrínseco e o extrínseco de um documento foi feita por Begthol²⁵ (1986 *apud* FUJITA 2003, p. 80) ao definir o conteúdo intrínseco de um documento (*aboutness*) como sendo independente:

[. . .] do uso temporal que um indivíduo possa fazer do mesmo em análise e que o faz possuir uma tematicidade relativamente permanente e um número variável de *meanings* (significados)”, podendo ser medido de acordo com o uso particular do documento tendo em vista os usuários. (grifo nosso)

Enquanto que *aboutness* refere-se ao “conteúdo relativamente permanente”

do documento, meanings diz respeito ao “significado compreendido” *pelo usuário*.

²¹ BLAIR, D. C. **Language and representation in information retrieval**. Amsterdam: Elsevier Science Publisher, 1990. *apud* FUJITA 2003, p. 79.

²² HJÖRLAND, B. The concept of subject in information science. **Journal of Documentation**, London, v. 48, n. 2, p.172-200, 1992. *apud* FUJITA 2003, p. 79.

²³ SOERGEL, D. **Organizing information**: principles of database and retrieval systems. New York: Academic Press, 1985. *apud* FUJITA 2003, p. 79.

²⁴ WEINBERG, B. H. Why indexing fails the researcher. **The Indexer**, London, v. 16, n. 1, p. 3-6, 1988. *apud* FUJITA 2003, p. 79

²⁵ BEGHTOL, C. Bibliographic classification theory and text linguistics: aboutness analysis, intertextuality and the cognitive act of classifying documents. **Jornal of Documentation**, London, v. 42, n. 2, p. 84-113, 1986. *apud* FUJITA 2003, p. 80)

Lancaster (2003, p.7-9) afirma que a indexação pode ser orientada para a consulta, ou “[. . .] o provável interesse das informações contidas em um documento, para um usuário ou grupo deles.”

Adverte ainda, quanto a indexação somente voltada para a consulta (demanda de usuários), a relatividade do atendimento das necessidades de informação dos usuários, uma vez que esta pode ser temporariamente satisfatória: “[. . .] aos usuários atuais e as necessidades de informação atuais”.

Há uma relação muito estreita entre um documento e uma necessidade de informação (atendência) ou de um documento e um enunciado de necessidade de informação (relevância) que orientam o indexador quanto ao nível de especificidade ou exaustividade a ser adotada.

Van der Laan (2000, p. 23) afirma que com o auxílio da Lingüística textual podemos ainda representar tematicamente um documento utilizando os sete padrões de textualidade de Beaugrand²⁶, observando:

1) coesão: indicam como os componentes ligam-se na produção de sentido.

“A coesão está ligada à dependência gramatical, ao sentido e ao uso [. . .]”, proporcionam a compreensão do texto.

2) coerência: como os conceitos e relações são mutuamente acessíveis e relevantes no texto, permitindo inferências e deduções não tão explícitas, mas que, segundo o contexto, podem ser assimiladas. “Uma ação pode ser causa para a subsequente; pode possibilitar ou ser a razão da ação.”

3) intencionalidade: o propósito, o objetivo, a intenção do autor. Uma mudança no plano discursivo afeta o entendimento e pode, conforme

²⁶ BEAUGRAND, Robert-Alain de. **Introduction to text Linguistics**. London: Longman, 1981. *apud* VAN DER LAAN, 2000

obseva VAN DER LAAN, levar o indexador a “[. . .] indexar um texto, equivocadamente, em um assunto não tratado no mesmo.”

- 4) **aceitabilidade:** este padrão envolve questões sociais, culturais e individuais. Na análise temática o indexador deve estar atento a essas questões de forma a evitar a censura temática, ou seja, representar o assunto de um documento de forma equivocada em proveito de suas convicções pessoais, sejam elas sociais, culturais, políticas, ideológicas ou religiosas. A análise deve ser feita levando-se em consideração as convicções *do autor*, a sua visão de mundo.
- 5) **informatividade:** ao indexar um documento o bibliotecário deve estar atento à obviedade e pouca informatividade que oferecido pelo autor “[. . .] uma informação pouco consistente pode gerar mais poluição no sistema do que resultar em um dado importante para o leitor.”
- 6) **situacionalidade:** “fatores que fazem um texto relevante para a situação de ocorrência.”
- 7) **intertextualidade:** relaciona-se muito de perto com o conhecimento prévio do leitor e sua capacidade de inter-relação com outros texto ou contextos vividos por ele.

Como podemos constatar até agora, são inúmeros os fatores que irão se refletir na qualidade da representação temática, portanto cabe a nós enquanto profissionais, aplicarmos o conhecimento teórico disponível em nossa área no sentido de qualificarmos o trabalho profissional e passarmos deste ponto em diante a buscar novas alternativas para novos problemas que ainda terão de ser enfrentados no âmbito biblioteconômico.

7 DISCUSSÃO

Pelo que podemos constatar, os problemas com que se depara a representação temática, referentes à qualidade, são mais de ordem individual do que global. Nosso estudo possibilitou perceber que o aparato teórico conceitual disponível no âmbito das ciências da informação não se encontra bem delineado ou, ainda, podemos inferir que a área não construiu o seu próprio paradigma. Não por ausência de um referencial que permita desenvolver o corpus teórico-metodológico de que a área necessita, mas tão somente pelo fato de que a grande maioria de profissionais da área não se utiliza, na prática, das contribuições das áreas correlatas de que se serve “na teoria.” Bastando que fique no mesmo discurso, da falta de fundamentações teóricas próprias que garantam a criação de metodologias que sejam eficazes, quando o que se vê é um tipo bem raro de laborofobia intelectual que pode levar o profissional a uma atrofia irreversível.

Não queremos dizer com isso, que não devemos buscar as fundamentações teóricas próprias, mas na impossibilidade que nos encontramos, neste momento, de criá-las por faltar o conhecimento de que ainda carece a área, procedente da falta de aplicação prática das teorias que se dispõem, pelo menos façamos o esforço de utilizar o que se dispõe. Como então poderemos obter avanços significativos na área, se nos falta quem acumule experiência prática para evidenciar sua inadequação ou impraticabilidade?

A representação temática, de maneira alguma, poderá ser reduzida de forma simplista, pois seu objeto, a informação, é demasiadamente complexo devendo ser considerada no próprio universo que lhe confere materialidade ou plasticidade: o universo cognitivo. Conhecer as variáveis interferentes no processo de indexação,

sejam elas ligadas ao indexador, ao texto ou ao contexto, possibilitará que se busquem soluções efetivas para minorar sua interferência.

Também é certo que o uso de metodologias adequadas, garante qualidade à representação temática. O conhecimento das estruturas e superestruturas ligadas ao texto conferem maior consistência a representação, na medida em que oferecem mais facetas a serem analisadas nas informações veiculadas pelo texto.

Evidentemente, a abordagem filosófica da terminologia oferece pragmatismo à segmentação baseada em hipóteses de organização do conhecimento e ao processo de informar.

Devemos refletir sobre a alteração dos currículos na área, que contemplem disciplinas que trabalhem com a teoria do discurso, tipologia das formas discursivas, suas condições de produção e várias outras disciplinas que irão remeter o indexador a repensar a leitura com fins profissionais e o tradicionalismo dos métodos de tratamento da informação.

Os avanços alcançados em vários domínios do conhecimento e que tem cada vez mais revelado o caráter interdisciplinar e mutante do mesmo, ainda não se encontram devidamente abalizados na esfera teórica.

Existe uma grande lacuna a ser preenchida no campo teórico, que é consequência da falta de reflexões acerca das práticas profissionais da área de organização e tratamento da informação, que estão sendo modificadas, em função da introdução das inovações tecnológicas, impondo novos e grandes desafios a serem transpostos pelos profissionais da informação. Essas inovações, se por um lado oferecem maior agilidade na disponibilização de informações a partir de sistemas informatizados, em contrapartida observa-se uma certa falta de reflexão

teórica, por parte de alguns profissionais, em relação ao aperfeiçoamento e atualização dos instrumentos de indexação.

Haja vista os interesses comuns da ciência da informação com os demais domínios do conhecimento, no que tange à interdisciplinaridade, às contribuições da semiótica e da semiologia, que vêm oferecendo os aportes necessários para a construção de linguagens de indexação. Seu objetivo é a melhora da performance profissional e da qualidade dos instrumentos de que se utiliza: que levem em consideração a profusão e o impacto das novas tecnologias disponíveis, bem como as necessidades e exigências dos usuários.

Um dos problemas na utilização de alguns instrumentos de indexação é a rigidez e a generalidade que fazem com que na fase de tradução dos conceitos veiculados pelos textos, perca-se a especificidade e a originalidade da informação, o que afeta, conseqüentemente, a sua recuperação.

A qualidade da representação, também fica afetada, pelo caráter simultaneamente hodierno e obsoleto da informação. Em função do princípio dinâmico e aceleradíssimo de produção do conhecimento, o vocabulário torna-se sempre defasado e, como efeito, esta defasagem reflete-se no processo de comunicação entre especialistas.

Os tesauros terminológicos, por sua própria estrutura e rede de relações entre conceitos, permitem que se efetue uma aproximação entre a linguagem do usuário e o sistema de informação. Não basta que os indexadores somente partilhem de práticas de indexação em comum sem que, no entanto, façam sobre a prática referida as devidas reflexões.

Torna-se necessário entender a informação para que possamos conferir certa plasticidade a esse elemento mutante, como alternativa para sua recuperação, acesso e disponibilidade.

Com isto, podemos nos perguntar, em que lugar se enquadram os “sistemas ou máquinas inteligentes” que resolverão todas as dificuldades. Essa crença sustentada por alguns bibliotecários, aos quais podemos denominar de tecnológico-visionários, beiram a puerilidade.

Essas posturas, longe de oferecerem soluções ao campo biblioteconômico, que ainda carece de fundamentações teórico-metodológicas para seu avanço, demonstram a imaturidade e o desconhecimento profissional diante do papel social desempenhado pelos profissionais da informação.

Na briga ferrenha instaurada entre o domínio do conhecimento intitulado Ciência da Informação, recentemente promovido de patente e dicotomizado de sua genitora, a rebaixada Biblioteconomia, atualmente operando no campo da **informação documentária**, podemos observar a tantas vezes criticada ingratidão do filho com mãe, que acreditamos, ter gerado esse sentimento de exclusão e rejeição social alimentado por alguns bibliotecários, por ainda não compreenderem seu papel social.

8 CONCLUSÃO

A qualidade da indexação depende mais diretamente da qualidade dos instrumentos que utiliza e da performance e habilidade dos indexadores, que só é obtida, quando os mesmos se propõem, constantemente, a recorrer à educação continuada. O conhecimento de teorias, que ofereçam subsídios à análise do texto, oferecidas pela lingüística textual, é uma valiosa ferramenta, por facilitar a leitura e a compreensão do texto.

Igualmente no contexto de trabalho, a explicitação e articulação das políticas de indexação, como diretriz ao trabalho profissional, quando em seus pormenores estão contempladas, permitem ao indexador, compreender e “incorporar” os objetivos do sistema de informação, a partir de suas regras, procedimentos e diretrizes. Acrescendo-se a isto, o conhecimento do processo de leitura documentária de sua parte, o auxiliará a mantê-lo centralizado, focado, em uma estratégia que facilite a representação do texto, levando-se em conta a tematicidade intrínseca e extrínseca do texto.

As áreas do conhecimento possuem peculiaridades que as caracterizam. Em função da riqueza de suas especificidades, o indexador deve possuir uma sólida formação profissional inter-relacionando a sua experiência como indexador e mediador na comunicação entre o sistema e as necessidades de seu usuário.

A teoria da terminologia vem prestando contribuições significativas na análise documentária. Os desafios e as limitações com que ela se depara atualmente, são só um pequeno vislumbre do longo caminho a ser percorrido para que se alcance, efetivamente, a tão almejada qualidade da representação. No entanto, o fato de conhecermos essas limitações, já possibilita que se articulem estratégias internas

nos serviços de informação para minorar o impacto resultante da produção acelerada do conhecimento, pelo fato de já, neste momento, termos presente a necessidade de incorporarmos as novas tecnologias ao trabalho bibliotecário.

A atividade biblioteconômica não se resume em tarefas mecanicistas, mas sim, no produto de grandiosos esforços para adquirir a fundamentação teórica-metodológica circunscrita ao domínio do qual a análise documentária faz parte, contribuindo, assim, na troca e assimilação de novos conhecimentos teóricos, sem entretanto, adquirir somente os rótulos de outros domínios, por acarretarem inconsistência na utilização de conceitos que não foram rearranjados ou readaptados ao âmbito biblioteconômico.

REFERÊNCIAS:

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 10520**: informação e documentação: citações em documentos: apresentação. Rio de Janeiro, 2002.

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

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 6024**: informação e documentação: numeração progressiva das seções de um documento escrito: apresentação. Rio de Janeiro, 2003.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 6027**: informação e documentação: sumário: apresentação. Rio de Janeiro, 2003.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 6028**: informação e documentação: resumo: apresentação. Rio de Janeiro, 2003.

CINTRA, A. M. M. et al. Para entender as linguagens documentárias. 2.ed. São Paulo: Polis, 2002. 92p.

DUMONT, Lígia Maria Moreira. Os múltiplos aspectos e interfaces da leitura. **DataGramaZero**: revista de Ciência da Informação, v.3, n.6, dez. 2002. Disponível em: <<http://www.dgz.org.br>>. Acesso em: 25 jul. 2005.

FUJITA, Mariângela Spotti Lopes. A leitura documentária na perspectiva de suas variáveis: leitor-texto-contexto. **DataGramaZero**: revista de ciência da informação, v.5, n.4, ago. 2004. Disponível em: <<http://www.dgz.org.br>>. Acesso em: 25 jul. 2005.

FUJITA, Mariângela Spotti Lopes. A identificação de conceitos no processo de Análise para indexação. **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v. 1, n. 1, p. 60-90, jul./dez. 2003. Disponível em: <<http://www.unesp.br/revistadigital>>. Acesso em: 01 ago. 2005.

KOBASHI, Nair Yumiko; SMIT, Johanna W.; TÁLAMO, Maria de Fátima G. M. A função da terminologia na construção do objeto da Ciência da Informação. **DataGramaZero**: revista de Ciência da Informação, v.2, n.2, abr.2001. Disponível em: <<http://www.dgz.org.br>>. Acesso em: 25 jul. 2005.

LANCASTER, F. W. **Indexação e resumos**: teoria e prática. Brasília: Briquet de Lemos, 2003. 360 p.

LARA, Marilda Lopes Ginez de. O processo de construção da informação documentária e o processo de conhecimento. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 7, n. 2, p. 127-139, jul./dez. 2002.

LARA, Marilda Lopes Ginez de. O unicórnio (o rinoceronte, o ornitorrinco ...), a análise documentária e a linguagem documentária. **DataGramaZero**: revista de ciência da informação, v.2, n.6, dez. 2001. Disponível em: <<http://www.dgz.org.br>> Acesso em: 19 jul. 2005.

MORAES, Alice Ferry de; ARCELLO, Etelvina Nunes. O conhecimento e sua representação. **Informação e Sociedade**, João Pessoa, v.10, n.2, p. 105-121, jul./dez. 2000.

MOREIRA, Alexandra; Alvarenga, Lídia; Oliveira, Alcione de Paiva. O nível do conhecimento e os instrumentos de representação: tesouros e ontologias **DataGramaZero**: revista de Ciência da Informação, v.5, n.6, dez. 2004. Disponível em: <<http://www.dgz.org.br>> Acesso em: 19 jul. 2005.

MOREIRO, José Antonio. Critérios e indicadores para evaluar la calidad del análisis documental de contenido. **Ciência da Informação**, Brasília, v.31, n.1, p. 53-60, jan./abr. 2002.

NAVES, Madalena Martins Lopes. Estudo de fatores interferentes no processo de análise de assunto. **Perspectivas em Ciência da Informação**, Belo Horizonte, v.5, n.2, p. 189-203, jul./dez. 2001.

NUNES, Cláudio Omar Iahnke. A relação da análise documentária com a terminologia. **Biblos**, Rio Grande, v.12, p. 99-113, jan./dez. 2000.

NUNES, Cláudio Omar Iahnke. Análise documentária: os limites da univocidade. **Biblos**, Rio Grande, v.14, p. 188-199, jan./dez. 2002.

NUNES, Cláudio Omar Iahnke. Algumas considerações acerca da ausência de políticas de indexação em bibliotecas brasileiras. **Biblos**, Rio Grande, v.16, p. 55-61, jan./dez. 2004.

RUBI, Milena Polsinelli; FUJITA, Mariângela Spotti Lopes. Elementos de política de indexação em manuais de indexação de sistemas de informação especializados. **Perspectivas em Ciência da Informação**, Belo Horizonte, v.8, n.1, p. 66-77, jan./jun. 2003.

SMIT, Johanna W.; TÁLAMO, Maria de Fátima G. M. A determinação do campo científico da Ciência da Informação: uma abordagem terminológica. **DataGramZero**: revista de Ciência da Informação, v.5, n.1, fev. 2004. Disponível em: <<http://www.dgz.org.br>>. Acesso em: 19 jul. 2005.

VAN DER LAAN, Regina Helena van der. Análise de assunto e lingüística textual: uma tentativa de aproximação. In: CONGRESSO BRASILEIRO DE BIBLIOTECONOMIA E DOCUMENTAÇÃO, 19., 2000, Porto Alegre. **Anais...** Porto Alegre: UFRGS, 2000. 1 CD-ROM.

VAN DER LAAN, Regina Helena van der. **Terminologia**: uma inter-relação. Tese (Doutorado) – Universidade Federal do Rio Grande do Sul. Instituto de Letras, Faculdade de Pós-Graduação em Letras. Porto Alegre, 2002. Orientadora: Prof^a. Dr^a. Maria da Graça Krieger.