

Leitura documentária para indexação: abordagens interdisciplinares
8 – A tematicidade textual na análise de assunto da indexação: aspectos
inerentes à atuação profissional do indexador

Roberta Cristina Dal'Evedove Tartarotti

SciELO Books / SciELO Livros / SciELO Libros

TARTAROTTI, R. C. D.E. A tematicidade textual na análise de assunto da indexação: aspectos inerentes à atuação profissional do indexador. In.: FUJITA, M. S. L., ALVES, R. C. V., and ALMEIDA, C. C., eds. *Modelos de leitura documentária para indexação: abordagens teóricas interdisciplinares e aplicações em diferentes tipos de documentos* [online]. Marília: Oficina Universitária; São Paulo: Cultura Acadêmica, 2020, pp. 195-216. ISBN: 978-65-8654-607-1. Available from: <http://books.scielo.org/id/96v3r>. <https://doi.org/10.36311/2020.978-65-86546-07-1.p195-216>.



All the contents of this work, except where otherwise noted, is licensed under a [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Todo o conteúdo deste trabalho, exceto quando houver ressalva, é publicado sob a licença [Creative Commons Atribuição 4.0](https://creativecommons.org/licenses/by/4.0/).

Todo el contenido de esta obra, excepto donde se indique lo contrario, está bajo licencia de la licencia [Creative Commons Reconocimiento 4.0](https://creativecommons.org/licenses/by/4.0/).

A TEMATICIDADE TEXTUAL NA ANÁLISE DE ASSUNTO DA INDEXAÇÃO: ASPECTOS INERENTES À ATUAÇÃO PROFISSIONAL DO INDEXADOR

Roberta Cristina Dal'Evedove TARTAROTTI

RESUMO: No domínio da Organização do Conhecimento (OC), a abordagem teórica da indexação objetiva a representação temática em sistemas de recuperação da informação. No entanto, o processo de indexação realizado pelo indexador reveste-se de uma complexidade característica notadamente em dois momentos distintos: a análise de assunto/identificação de conceitos (tematicidade intrínseca) e a seleção de conceitos (tematicidade extrínseca). Nesse sentido, o trabalho tem como objetivo trazer algumas considerações teóricas em torno da tematicidade textual na análise de assunto do processo de indexação, particularmente sobre a tematicidade extrínseca textual da etapa de seleção de conceitos. Neste viés, busca-se contribuir para o entendimento dos fatores interferentes no processo de indexação e da melhoria da atuação do bibliotecário indexador como um leitor profissional. Conclui-se que o estabelecimento de interfaces interdisciplinares da Organização do Conhecimento (OC) com outros campos científicos contribui não apenas para a resolução de problemas teórico-práticos em torno da representação e recuperação temáticas da informação como também para seu fortalecimento como campo científico.

PALAVRAS-CHAVE: Indexação. Análise de assunto. Tematicidade textual. Indexador.

ABSTRACT: In the Knowledge Organization (OC) domain, the indexing theoretical approach aims at thematic representation in information retrieval systems. However, the indexing process carried out by the indexer has a complexity characteristic, especially in two different moments: the subject analysis/concept identification (intrinsic thematicity) and the concept selection (extrinsic thematicity). Therefore, the work aims to bring some theoretical considerations around the textual theme in the subject analysis of the indexing process, particularly on the extrinsic textual theme of the concepts selection stage. In this bias, it was sought to contribute to the understanding of the interfering factors in the indexing process and the improvement of the indexer librarian's role as a professional reader. It is concluded that the establishment of interdisciplinary interfaces of the Knowledge Organization (OC) with other scientific fields contributes not only to the resolution of theoretical-practical problems around the representation and the information thematic retrieval but also to its strengthening as a scientific field.

KEYWORDS: Indexing. Subject analysis. Textual thematicity. Indexer.

1 INTRODUÇÃO

Diversos autores têm se debruçado nas questões teórico-metodológicas visando à consolidação da Organização do Conhecimento (OC) como disciplina científica e domínio de conhecimento. Por definição, o campo da Organização do Conhecimento (OC) é “[...] um sub-domínio-chave da Ciência da Informação, que é dedicado à ordem conceitual do conhecimento” (SMIRAGLIA, 2011, p. 1, tradução nossa) e “[...] um domínio inter e multidisciplinar indispensável para as atividades científicas, que também é praticada dentro da moderna Ciência da Informação” (BABIK, 2014, p. 328, tradução nossa).

Na Ciência da Informação, a leitura documental é pressuposto básico das atividades desempenhadas pelos profissionais que lidam com o tratamento temático da informação. O objetivo deste tratamento temático é possibilitar a representação por assunto de recursos informacionais e sua posterior recuperação pelos usuários nos sistemas de recuperação da informação: catálogos online, bases de dados ou mais atualmente em repositórios institucionais ou repositórios de dados de pesquisa no contexto de bibliotecas acadêmicas, dentre outros. Assim, uma das características imprescindíveis em um sistema de recuperação da informação é representar o conteúdo intelectual dos documentos para prover acesso aos mesmos baseando-se no seu **assunto**.

A análise de assunto, que se caracteriza pelo conjunto dessas subetapas, é considerada na literatura e na prática profissional como a mais importante de todo o processo de indexação, pois resultará na representação do assunto principal do recurso informacional, tornando-se a “operação base para todo o procedimento de recuperação de informações” (CESARINO; PINTO, 1980, p. 32). A importância da análise de assunto é reforçada por Fujita (2003, p. 78), para quem “[...] o assunto ou tematicidade do documento é o cerne principal e mais carente de esclarecimentos dentro dos estudos em análise documental”. Entretanto, as variáveis na indexação são muitas e pouco conhecidas, considerando-se as inúmeras possibilidades de significados e ideias contidos nos recursos informacionais em cada elemento no processo, além da presença da subjetividade, que depende particularmente do indexador (MAI, 1997a, p. 65).

A abordagem teórica da indexação, ao adaptar ou propor uma metodologia que vise uma análise de assunto adequada dos recursos informacionais deve admitir que esta ocorre na leitura documental. Neste sentido, abordagens interdisciplinares com outros campos científicos tal como a Linguística que têm o texto como objeto de estudo contribuem para um maior entendimento desta atividade (TARTAROTTI; BOCCATO, 2013, p. 43). Nesse viés, busca-se tecer algumas considerações teóricas em torno da tematicidade textual da análise de assunto da indexação, especialmente no que tange à **tematicidade extrínseca textual da etapa de seleção de conceitos**.

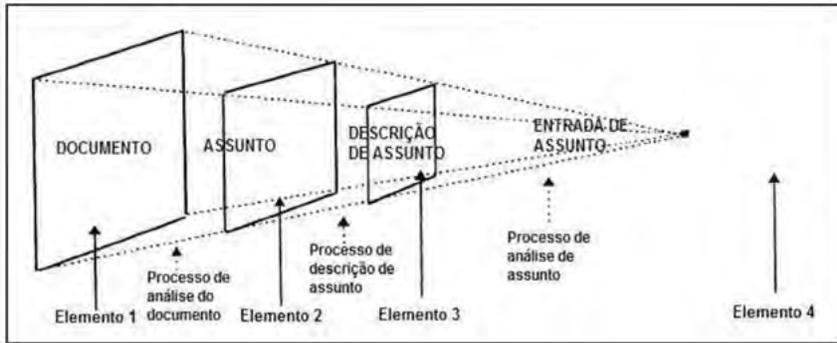
2 A ANÁLISE DE ASSUNTO PARA IDENTIFICAÇÃO E SELEÇÃO DE CONCEITOS NO PROCESSO DE INDEXAÇÃO

Conceitualmente, a indexação de assuntos caracteriza-se como um processo formado por subprocessos ou etapas que identificam o conteúdo de um recurso informacional por meio de uma metalinguagem construída – a linguagem documental –, visando a efetiva recuperação da informação (TARTAROTTI; BOCCATO, 2013, p. 40). Em termos teóricos, qualquer recurso informacional pode ser indexado/representado por meio de conceitos em busca da satisfação de necessidades informacionais específicas (GIL LEIVA, 2008, p. 64). Logo, a indexação vai além da representação temática da produção científica – foco dos produtos e serviços – podendo

ser aplicada, em um contexto mais amplo, a qualquer recurso informacional ou objeto passível de ser denominado documento.

Embora não consensuais entre os autores da literatura especializada no temário, para Mai (2000, p. 294) o processo de indexação pode ser desconstruído revelando três etapas: *processo de análise do documento*, *descrição do assunto* e *processo de análise de assunto* e quatro elementos: *documento*, *assunto*, *descrição do assunto* e *entrada do assunto*, assim representada geometricamente (Figura 1):

Figura 1. O processo de indexação de assunto



Fonte: Mai (1997a, p. 61) adaptado de Miksa (1983)²⁹.

Nesta **abordagem interpretativa** do processo de indexação, o autor considera a indexação como uma atividade intelectual que requer a interpretação do texto do recurso informacional por parte do indexador, indo além de uma atividade puramente técnica. Ao imprimirmos tais etapas e elementos no domínio dos modernos sistemas de recuperação da informação tais como bases de dados, catálogos online, repositórios institucionais ou repositórios de dados de pesquisa, na primeira etapa do processo de indexação é realizada a análise do recurso informacional digital ou, na terminologia do autor - objeto digital - visando à descrição temática, denominada de *processo de análise*. O primeiro elemento é o *objeto digital* que está sendo analisado. A segunda etapa é a formulação de uma frase de indexação ou descrição de assunto, denominada de *processo de descrição*, uma formulação mental ou escrita do assunto pelo indexador, tendo como

²⁹ MIKSA, F. *The subject in the dictionary catalog from Cutter to the present*. Chicago: American Library Association, 1983.

segundo elemento o *assunto* do documento ou do objeto digital, que pode estar presente apenas na mente do indexador. Já na terceira etapa ocorre a tradução da descrição de assunto em uma linguagem de indexação ou esquema de classificação, denominada de *processo de análise de assunto*. O terceiro elemento nesta etapa é a *descrição formal do assunto*, que pode ser escrita. Já o quarto elemento, denominado de *entrada de assunto*, é o produto da tradução da descrição formal do assunto em uma determinada linguagem de indexação do repositório institucional (MAI, 1997b, p. 55).

Cabe esclarecer que o autor propõe o uso do termo “análise do documento” para a primeira etapa do processo e “análise de assunto” para a última etapa do processo. Na literatura percebe-se a existência de mais de uma denominação para o termo *análise de assunto*, podendo ser denominada ainda de *análise temática* ou *análise conceitual*. Entretanto, entendemos que o termo **análise de assunto** é o que melhor representa o processo de análise do conteúdo temático do recurso informacional, considerando-se seu sólido estabelecimento e aceitação da terminologia na literatura científica do campo da Organização do Conhecimento (OC).

Conforme Fujita, Rubi e Boccato (2009, p. 24), as etapas do processo de indexação não precisam ser necessariamente realizadas de maneira sequencial, visto que o indexador profissional já familiarizado com o processo pode realizá-las simultaneamente e até de maneira automática, situação corroborada na prática em estudo realizado com catalogadores e indexadores por Tartarotti (2014, p. 198), para quem quanto mais familiarizado com a área científica especializada no qual atua, mais estas etapas são superpostas durante o processo de indexação pelo profissional. Além disso, de acordo com Mai (2000, p. 280), a ação tomada por indexadores com diferentes níveis de experiência pode não ser exatamente a mesma, mas as etapas e elementos do processo de indexação de assunto são considerados fundamentais para qualquer processo de indexação.

No processo de indexação, a **análise de assunto** é considerada a etapa mais complexa e mais importante da indexação, pois determina os resultados das estratégias de busca realizadas pelos usuários do sistema de recuperação da informação, no momento da recuperação da informação. Durante a análise de assunto, o indexador examina o documento, procurando compreender o texto por meio da **identificação** do assunto, com o objetivo posterior de **selecionar** os conceitos indexáveis. Chu

e O'brien (1993) consideram a análise de assunto como fase inicial do processo de indexação que decidirá sobre os principais tópicos do assunto de um documento, precedendo a fase de tradução desses tópicos de acordo com a linguagem documental adotada pelo sistema.

Segundo Taylor e Joudary (2009, p. 305), a “[...] análise de assunto é a parte do processo de criação de metadados que identifica e articula o assunto do recurso informacional que está sendo descrito”. O processo de análise de assunto abarca três subetapas; 1) análise conceitual para determinar de que trata o item; 2) descrição da tematicidade em uma declaração por escrito; e 3) utilização dessa declaração de tematicidade para atribuir termos de vocabulário controlado e/ou notações de classificação. Já para Raju e Raju (2006, p. 14), teóricos da abordagem teórica da catalogação de assunto, a análise de assunto consiste no “processo de analisar o assunto contido em um item antes da tradução desta análise conceitual em uma linguagem”. Em síntese, a análise de assunto implica em determinar a tematicidade do documento mediante a identificação e seleção dos conceitos que comporão o assunto ou temas principal e secundários.

No contexto da análise de assunto realizada pelo indexador, a leitura documental possui um fundamental, pois “[...] para o indexador, seu cotidiano é concentrado no ato da leitura, de forma a viabilizar o acesso à informação contida nos documentos aos usuários dos sistemas de informação” (NEVES; DIAS; PINHEIRO, 2006, p. 141). Conforme Fujita (2003, p. 94), é a leitura do documento que dá início à atividade de indexação, diferindo de uma leitura normal ao exigir procedimentos diferentes, “[...] ainda que os conhecimentos necessários para um bom entendimento de um texto sejam comuns a ambas”. Aqui, “[...] esse leitor-indexador tem objetivo definido: identificação e seleção de conceitos que representem o conteúdo do texto e que coincidam com as necessidades informacionais da comunidade usuária do sistema de informação” (LARA, 1993, p. 50).

Todavia, é notável a dificuldade apresentada pelo indexador ao ler o documento com o objetivo de identificar e selecionar os conceitos representativos do assunto do documento, considerando-se que, na maioria das vezes, não é um especialista sobre o assunto que indexa. Por outro lado, o indexador bibliotecário pode ser tornar um especialista na

área em que atua devido à experiência adquirida ao longo do tempo na prática na atividade de indexação, além de cursos especializados (FUJITA, 2003, p. 84).

Considerando-se que a construção de qualquer texto de um recurso informacional é baseada em um ou mais temas, é por meio da **leitura profissional** na representação temática da informação que ocorre a busca pela **tematicidade** documental, revelada por meio da **análise de assunto** realizada pelo indexador - leitor profissional que interage com o texto para cumprir o objetivo da indexação. Sendo realizada durante a leitura documental, a análise de assunto é subdividida em outras três etapas: **compreensão** do conteúdo do recurso informacional; **identificação dos conceitos** que representam este conteúdo; e **seleção dos conceitos** válidos para recuperação. Se a análise de assunto, como primeira etapa da indexação, é realizada através da leitura - uma leitura documental -, é no momento da leitura em que os conceitos tratados em um documento são **identificados** e **selecionados**, para posterior representação em termos de indexação. Nesse sentido,

[...] a **representação por conceitos** na **análise de assunto** para identificação e seleção de conceitos é diferente da **representação realizada na tradução** porque, em primeiro lugar, é realizada durante a análise de conteúdo do documento, ou seja, com o documento na íntegra e em segundo lugar, porque utiliza conceitos com os quais são representados os termos e assim identificada a tematicidade intrínseca do documento. Na tradução, de outro modo, a **representação** se dá com os **termos extraídos do documento analisado**, portanto, **fora do contexto documental** em que o autor desenvolveu o conteúdo significativo e utiliza uma linguagem *documental* que representa o vocabulário de áreas de especialidade com a finalidade de compatibilizar os termos identificados com os termos buscados pelo usuário (FUJITA, 2013, p. 50, grifo nosso).

Conforme apresenta Dal'Evedove (2002, p. 58), em uma abordagem ideal da prática da indexação, as etapas de leitura documental, análise de assunto/identificação de conceitos e seleção de conceitos precisam ocorrer durante a leitura, enquanto que a etapa de tradução dos conceitos que representam os conceitos em descritores da linguagem do

sistema deve ser feita após a leitura do documento. Esta visão permite que a análise seja realmente conceitual e compreensiva, pois a preservação do conteúdo do documento é uma garantia de relevância de recuperação, objetivo da eficiente indexação de conteúdo.

Referente às subetapas da análise de assunto, a **compreensão de leitura** é considerada uma condição necessária à leitura, ou seja, “[...] não existe leitura sem compreensão. Então, quando falamos em leitura para indexação, podemos dizer que o indexador necessita compreender o texto para identificar e selecionar conceitos, pois somente o fará a contento se houver compreensão” (FUJITA, 2003, p. 82). Sobre isso Farrow (1991, p. 151) adverte que o indexador compreende o texto basicamente do mesmo modo que um leitor fluente, porém, sob a influência de condições de tempo, objetivo definido, modelo a ser produzido e áreas temáticas definidas com estrutura textual padronizada dos recursos informacionais, podendo induzi-lo a um processo repetitivo e automático.

Na **identificação de conceitos**, o indexador, “após o exame do texto, passa a abordá-lo de uma forma mais lógica a fim de selecionar os conceitos que melhor representem seu conteúdo”. Esta etapa depende “[...] da tematicidade do texto e está atrelada à leitura do indexador e às suas concepções de análise de assunto adquiridas pela sua formação, objetivos e políticas de indexação” (FUJITA, 2003, p. 64-85). Deste modo, a identificação dos conceitos tratados num documento envolve como aspecto principal a compreensão do conteúdo do documento pelo leitor, de forma a representar e propiciar a seleção dos conceitos válidos para recuperação.

Já a etapa de **seleção de conceitos**, que ocorre após a identificação dos conceitos, está diretamente relacionada aos “objetivos para os quais as informações são indexadas”, sendo que “nem todos os conceitos identificados serão necessariamente selecionados” (FUJITA, 2003, p. 64).

3 A TEMATICIDADE TEXTUAL NA ANÁLISE DE ASSUNTO: DA TEMATICIDADE INTRÍNSECA À TEMATICIDADE EXTRÍNSECA

Como visto, a análise de assunto é subdividida nas etapas de **compreensão** do conteúdo do recurso informacional; **identificação dos conceitos** que representam este conteúdo; e **seleção dos conceitos**. Todo

texto possui uma **tematicidade intrínseca** – que podemos denominar de **tematicidade básica**, ou seja, um tema nuclear inerente ao texto e ligado à **produção** do mesmo, identificada na etapa de análise de assunto ou identificação de conceitos. De acordo com Todd (1992, p. 102), o nível de relação entre tematicidade e significado varia, pois depende “do uso que a pessoa pode encontrar da tematicidade do documento numa certa época, e o mesmo documento pode vir a ter diferentes significados para o mesmo leitor em diferentes épocas, entretanto o documento possui uma atinência fundamental”. Esta etapa se refere à **identificação de conceitos**.

Por outro lado, a **tematicidade extrínseca** – que podemos denominar de **tematicidade profunda** – está diretamente relacionada com o **contexto** de produção e de recepção do texto, com a necessidade informacional dos usuários do sistema de recuperação da informação e com a política de indexação da instituição, ligada à **recepção** do texto. Estes fatores, quando conhecidos pelos profissionais que executam a atividade de indexação, são decisivos para a adequação temática ao contexto e respectiva recuperação temática. Esta etapa se refere à seleção de conceitos.

Na tematicidade do documento, destaca-se o conteúdo relativamente **permanente do documento** (*aboutness*), o que nos leva à **tematicidade intrínseca textual**, enquanto que o **significado compreendido pelo usuário** (*meanings*) refere-se à **tematicidade extrínseca textual**. De acordo com Fujita (2003, p. 80), Begthol (1986)³⁰ fez distinção entre *aboutness* e *meanings*: “o *aboutness* é o conteúdo intrínseco do documento, que independe do uso temporal que um indivíduo possa fazer do mesmo em análise e que o faz possuir uma tematicidade relativamente permanente”; já o *meanings* (significados) pode ser medido de acordo com o uso particular do documento tendo em vista os usuários (FUJITA, 2003, p. 80).

Ainda para Fujita (2003, p. 79), não existe um consenso entre os autores em torno do termo *aboutness* em português, cunhado em 1969 por Fairthorne (1969). Enquanto para alguns autores significa “do que trata um texto”, outros utilizam “atinência” ou ainda “**tematicidade**”. Wellish (2000, p. 5) define *aboutness* como “[...] o conjunto de assuntos ou tópicos tratados em um documento, inclusive a intenção do autor e o

³⁰ BEGHTOL, C. Bibliographic classification theory and text linguistics: aboutness analysis, intertextuality and the cognitive act of classifying documents. *Journal of Documentation*, London, v. 42, n. 2, p. 84-113, 1986.

possível uso pelos leitores” Para Albrechtsen (1993, p. 220), tal conceito passou a ser pesquisado em substituição ao conceito de *subject* ou “tema”. Nesse sentido, conceitualmente podemos considerar a tematicidade como sinônimo de tema. De acordo com Fujita (2001, p. 61), o tema possui uma estrutura temática composta por conceitos ou categorias ou facetas. A identificação das categorias ou conceitos na estrutura textual do recurso informacional ocorre por meio da análise conceitual e o conjunto das categorias identificadas formulará o tema do documento. Segundo Guimarães (2009, p. 35), o tema “[...] é uma representação abstrata da estrutura global do significado de um texto”, compreendido ainda como um

[...] núcleo informativo fundamental ou elemento em torno do qual se estrutura a mensagem, sua identificação permite ao receptor considerar “entender” o texto, daí poderá partir para a elaboração de resumo do texto, ou para exercício de retextualização sob diversas formas (paráfrases, comentários, resenhas, resenhas, resenhas, etc.). O processo seguido pelo receptor é, de certo modo, oposto ao utilizado pelo emissor. Se este procede desenvolvendo o tema, o receptor deve, reduzindo as informações que lhe são transmitidas, limitar-se ao fundamental, até chegar a esse núcleo informativo (GUIMARÃES, 2009, p. 35).

De forma geral, enquanto o *aboutness* refere-se ao “conteúdo relativamente permanente do documento”, o *meanings* é entendido como “o significado compreendido pelo usuário” do sistema de recuperação da informação (FUJITA, 2003, p. 80). A literatura científica sinalizava que o interesse do indexador deve limitar-se ao *aboutness*, ou seja, à tematicidade textual, tendo em vista que o indexador lida com os conteúdos permanentes dos recursos informacionais (com as ideias propostas pelo autor), representadas através da linguagem textual.

Entretanto, os estudos mais contemporâneos defendem o estabelecimento de um novo paradigma no processo de análise de assunto, pois o *meanings*, ou seja, o contexto, também deve ser considerado pelo indexador durante a atividade de indexação, considerando-se que embora a tematicidade seja o conteúdo relevante do documento no contexto próprio do autor, algumas variáveis irão influenciar na determinação

desse conteúdo. Logo, a determinação do(s) assunto(s) do recurso informacional também está ligada a estes fatores. Segundo Cavalcanti (1989) a tematicidade intrínseca refere-se ao tema importante para o autor, enquanto que a tematicidade extrínseca refere-se ao tema importante do ponto de vista do leitor. Isso é denominado pela autora como “saliência autor – relevância leitor”.

Em síntese, no processo de representação da informação existem dois tipos de tematicidade: a **tematicidade intrínseca** (resultado da fase de **identificação de conceitos** na análise de assunto) e a **tematicidade extrínseca** (resultado da fase de **seleção de conceitos** na análise de assunto), conforme apresentado no Quadro 1:

Quadro 1 - A tematicidade textual na análise de assunto do processo de indexação

ETAPA DA INDEXAÇÃO	SUBETAPA DA INDEXAÇÃO	TIPO DE TEMATICIDADE
Leitura documental	-	-
Análise de assunto	Compreensão de leitura	-
	Identificação de conceitos	<i>Tematicidade intrínseca</i>
	Seleção de conceitos	<i>Tematicidade extrínseca</i>
Tradução dos conceitos	-	-

Fonte: Elaborado pela autora.

Considerando-se que a tematicidade intrínseca e a tematicidade extrínseca são aspectos norteadores do processo de indexação, entendemos que a prática ideal desta atividade realizada pelo indexador deve abarcar tanto o *aboutness* (contexto de produção) quanto o *meaning* (contexto de recepção) dos recursos informacionais.

4 A TEMATICIDADE EXTRÍNSECA DA ETAPA DE SELEÇÃO DE CONCEITOS NA ANÁLISE DE ASSUNTO

Na prática, a indexação é analisada sob três diferentes concepções teóricas ou perspectivas de análise que refletem na atuação profissional do indexador: indexação orientada para o documento (ênfase no documento), indexação orientada para o usuário (ênfase nos usuários) e indexação orientada para o domínio (engloba o **contexto**, o documento e os usuários) (GIL LEIVA, 2008, p. 62). De acordo com Fujita (2003, p. 72), a concepção de leitura orientada para o conteúdo deve orientar a identificação de conceitos e a concepção orientada para a demanda, a seleção de conceitos, sendo que a concepção do indexador está “diretamente vinculada com sua formação educacional (concepção orientada pelo conteúdo) e com a postura do sistema de informação (concepção orientada pela demanda) e não pelo fato de ele ser um leitor menos ou mais habilitado”. Por abarcar uma visão mais ampla, a abordagem da indexação centrada no domínio representa o ideal em termos de indexação, pois considera outros elementos além do documento ou o usuário.

Sob a ótica de Smiraglia (2009, p. 674), a interpretação de um recurso informacional envolve o contexto em que o recurso informacional é interpretado, visto que anteriormente ao processo de indexação houve a criação do recurso informacional por uma determinada **comunidade discursiva**, resultado de diversos **atos de interpretação** por parte de seus atores sociais. Uma vez concluído, publicado e analisado por meio do processo de indexação possibilitando que o recurso informacional seja acessível, o objetivo é que este seja recuperado e utilizado pelos usuários da informação, estejam ou não inseridos dentro da própria comunidade discursiva que o elaborou (MAI, 2000, p. 272; MAI, 2001, p. 604).

No processo de indexação, a seleção de conceitos corresponde à etapa da análise de assunto que resultará na extração da **tematicidade extrínseca textual** do recurso informacional em análise pelo profissional indexador. Logo, a etapa de seleção de conceitos é influenciada pela variável **contexto**, que refere-se à análise de assunto em contextos de informação especializados e todos os elementos influentes no momento da leitura profissional, isto é, “[...] todas as condições nas quais se encontra o leitor quando entra em contato com um texto.” (GIASSON, 1993, p. 40). Além disso, “[...] representa não apenas o **conhecimento prévio profissional** e

os **objetivos da atividade em questão**, mas o **ambiente** em que é realizada esta atividade profissional em leitura documental” (DAL’EVEDOVE, TARTAROTTI; FUJITA, 2015, p. 610).

Nessa perspectiva, para que o significado do conteúdo informacional de um documento seja o menos ambíguo possível, o mesmo deve ser analisado de acordo com o **contexto** no qual está inserido, uma vez que tal variável tende a agregar valor em determinados conceitos, os quais passam a assumir novos significados apenas mudando-se o ambiente de análise (PINTO MOLINA, 1993). Desse modo, “o conceito somente assume sentido quando levado em consideração o contexto de situação, em que o bibliotecário indexador deve ajustá-lo as suas decisões conforme a necessidade exigida pelo meio” (DAL’EVEDOVE; TARTAROTTI; FUJITA, 2015, p. 610).

Entretanto, como observa Olson (2007, p. 536), a maioria dos relacionamentos na **abordagem temática da informação** como os tesouros (produtos da concepção teórica da indexação), listas de cabeçalhos de assunto (produtos da concepção teórica da catalogação) ou esquemas de classificação (produtos da concepção teórica da classificação) está limitada às relações paradigmáticas, ou seja, são **relações intrínsecas**; não dependendo do contexto. Nesta **abordagem interpretativa** do processo de indexação, os principais problemas da representação de documentos estão relacionados ao significado e à linguagem, conseqüentemente que o **indexador** é o fator mais interferente do processo de indexação, pois os referentes de cada elemento no processo são entendimentos **subjetivos** e, portanto, que os resultados do processo são bastante incertos, e dependendo em grande parte de quem realiza a análise (MAI, 2000, p. 270; MAI, 1997a, p. 61).

Na opinião de Neves (2006, p. 43), o momento mais crucial do processo de indexação é a análise de assunto, pois reveste-se de uma subjetividade inerente à política de indexação, que determina prioridades a alguns assuntos. Para minimizar esta subjetividade na indexação, a formação e capacitação profissional desempenham um papel-chave. Entretanto, conforme alertam Neves, Dias e Pinheiro (2006, p. 151), “[...] apesar desse treinamento, em sua busca por termos para representar o assunto tratado no documento, o indexador baseia-se no discernimento próprio e na prática adquirida no exercício da profissão, o que, além

de tornar o processo de indexação **subjetivo**, questiona a eficácia da formação profissional”.

No que concerne aos aspectos cognitivos do indexador enquanto leitor, ao considerarmos que todo fato cultural, atividade ou prática social constituem-se como práticas significantes, ou seja, práticas de produção de linguagem e de sentido (SANTAELLA, 2007, p. 10), isto nos leva ao campo da Linguística (a ciência da linguagem verbal), mais especificamente à abordagem teórica da **Linguística Textual**, considerando-se a relevância do objeto texto como *materializador de sentidos* no processo de indexação no âmbito da Organização do Conhecimento (OC).

De acordo com Koch (2009, p. 154), a multiplicidade cultural, histórica e funcional que caracteriza a moderna evolução do conhecimento seria impossível sem a existência de textos. Para Discini (2007, p. 14), “[...] o texto é uma unidade de sentido, dada por recorrência daquilo que é dito e de um modo próprio dizer”.

Cavalcanti (1989, p. 48) considera a interação leitor-texto como um processo caracterizado por quatro principais aspectos: **ativo** (que envolve não compreensão, mas interpretação e exige competência comunicativa); **construtivo** (baseia-se em interpretação das expressões indexicais sinalizadas no texto); **interativo** (influenciado por variáveis de desempenho); e **reconstrutivo** (permite a introdução de vieses tanto das variáveis de desempenho como dos sistemas de valores do indexador). Neste contexto, a interação entre o indexador e o texto “[...] consiste numa amplitude de procedimentos coordenados e de operações perceptíveis, linguísticas e cognitivas, em que as características da estrutura textual fornecem subsídios ao conhecimento do bibliotecário indexador, aumentando a compreensão do assunto abordado” (DAL’EVEDOVE; TARTAROTTI; FUJITA, 2015, p. 609).

Nos estudos mais recentes sobre o texto e o discurso da Linguística Textual, recai uma ênfase na concepção de cognição humana de forma integrada, “[...] onde a linguagem, percepção, afeto, atenção, memória, estrutura cultural e outros componentes do sistema cognitivo encontram-se definitivamente inter-relacionados”, assim como uma retomada do interesse pelas tipologias e gêneros textuais (BENTES, 2011, p. 282). No âmbito da Ciência da Informação, a cognição é percebida em todas as atividades realizadas pelos profissionais da informação, tanto no

tratamento descritivo ou temático quanto na recuperação da informação. Nesse sentido, “[...] a cognição implica que cada ato de processamento da informação, seja ele perceptivo ou simbólico, é mediado por um sistema de categorias e conceitos os quais, para o mecanismo de processamento da informação, constituem um modelo de mundo” (NEVES, 2006, p. 42).

Entre as diversas concepções de texto que fundamentaram os estudos em Linguística Textual, destacam-se: **concepção gramatical**: texto como frase complexa ou signo linguístico mais alto na hierarquia do sistema linguístico; **concepção semiótica**: texto como signo complexo; **concepção semântica**: texto como expansão tematicamente centrada de macroestruturas; **concepção pragmática**: texto como ato de fala complexo; **concepção discursiva**: texto como discurso ‘congelado’, como produto acabado de uma ação discursiva; **concepção comunicativa**, texto como meio específico de realização da comunicação verbal; **concepção cognitivista**: texto como processo que mobiliza operações e processos cognitivos e, por fim, **concepção sociocognitiva-interacional**: texto como lugar de interação entre atores sociais e de construção interacional de sentidos (KOCH, 2009, p. 12). Na concepção sociocognitiva-interacional, “a compreensão deixa de ser vista como simples “captação” de uma representação mental ou como a decodificação de mensagem resultante de uma codificação de um emissor”, mas entendida como uma “atividade interativa altamente complexa de produção de sentidos”, realizada “com base nos elementos linguísticos presentes na superfície textual e na sua forma de organização, mas que requer a mobilização de um vasto conjunto de saberes (enciclopédia) e sua reconstrução no interior do evento comunicativo” (KOCH, 2009, p. 18).

Ao considerarmos o processo de indexação com base em Koch (2009, p. 19), o sentido de um texto de um determinado recurso informacional é construído na interação recurso informacional-indexador (texto-sujeito) com três elementos: o **autor (produtor)**, que procura viabilizar o seu “projeto de dizer”, recorrendo a diversas estratégias de organização textual e orientando o indexador profissional por meio de marcações textuais para a construção dos (possíveis) sentidos; o **recurso informacional (texto)**, organizado estrategicamente de dada forma, em decorrência das escolhas feitas pelo autor entre as inúmeras possibilidades de materialização oferecida pela língua e estabelecendo limites quanto

às leituras possíveis; e o **indexador (leitor)**, que, a partir do modo como o texto se encontra linguisticamente construído, das sinalizações que lhes oferece, bem como pela mobilização do contexto relevante à interpretação, vai proceder à construção dos sentidos a fim de cumprir o objetivo da indexação.

De acordo com Koch (2009, p. 14-15), a concepção de sujeito – ou mais especificamente, a de *sujeito da linguagem* (neste caso, de autor e indexador) varia de acordo com a concepção de língua adotada. Em uma concepção de **língua como estrutura**, caracteriza-se um autor/indexador “não consciente”, inerte, “assujeitado” pelo sistema, onde o comportamento individual é apenas um reflexo do comportamento coletivo, seja este linguístico ou social. O autor/indexador é inserido numa ideologia, numa instituição da qual é apenas porta-voz: é um discurso anterior que fala através dele. Ainda para a autora, se pensarmos na **língua como representação do pensamento**, percebe-se um autor/indexador individual, dono de sua vontade e de suas ações, social na medida em que se constrói em sociedade e com isto adquire a habilidade de interagir, que constrói uma representação mental e deseja que esta seja “captada” pelo interlocutor da maneira como foi mentalizada. Já na concepção de **língua como lugar de interação**, percebe-se o autor/indexador como ativo na produção do social e da interação. O autor/indexador reproduz o social na medida em que participa ativamente da definição da situação na qual se acha engajado, sendo um ator na atualização das imagens e das representações, sem as quais a comunicação não poderia existir. Desse modo, a concepção de língua como lugar de interação e do autor/indexador como um elemento ativo neste processo caracteriza o **contexto sociocognitivo**.

Conforme a autora, atualmente na Linguística Textual o **contexto** abrange não apenas o co-texto, como a situação de interação imediata, a situação mediata (entorno sociopolítico-cultural) e também o contexto sociocognitivo dos interlocutores (que integra os demais), mas também todos os tipos de conhecimentos arquivados na memória dos atores sociais (neste caso, indexadores e atores durante o processo de indexação). Estes conhecimentos são classificados como: o **conhecimento linguístico** propriamente dito; o **conhecimento enciclopédico** (declarativo ou episódico, frames e *scripts*); o **conhecimento da situação comunicativa** e

de suas ‘regras’ (situacionalidade); o **conhecimento superestrutural** (tipos textuais); o **conhecimento estilístico** (registros, variedades de língua e sua adequação às situações comunicativas); o **conhecimento sobre os variados gêneros** adequados as diversas práticas sociais; o **conhecimento de outros textos** que permeiam nossa cultura (intertextualidade) (KOCH, 2009, p. 24).

Esta interação ocorre com o uso de estratégias cognitivas (inferências, focalização, busca de relevância, etc.), estratégias sociointeracionais (preservação das faces, polidez, atenuação, atribuição de causas a (possíveis) mal-entendidos etc.) e estratégias textuais (formas de organização do texto com vistas à produção de determinados sentidos) (VILELA; KOCH, 2001, p. 24, p. 464). Assim, todos estes conhecimentos são importantes para o indexador realizar uma análise conceitual e compreensiva no **processo de indexação**. Percebe-se que a literatura da área tem se voltado principalmente para as estratégias cognitivas e metacognitivas do indexador. Do ponto de vista da Linguística Textual, o processo de indexação é um **processamento textual**, e, portanto, torna-se relevante mais estudos sobre estas outras estratégias de interação descritas por este campo.

Se um texto escrito não é recebido ativamente pelo indexador, pressupondo energia de processamento cognitivo por parte do indexador. Desse modo, a cognição será mais eficaz se completada pela **metacognição**, isto é, “pelos princípios que regem a desautomatização consciente das estratégias cognitivas ou o pensamento sobre nosso próprio pensamento”. Assim, “[...] um texto não tem sentido por si mesmo, mas graças à interação que se estabelece entre o conhecimento apresentado no texto e o conhecimento de mundo armazenado na memória do interlocutor” (GUIMARÃES, 2009, p. 17). De acordo com Vilela e Koch (2001, p. 559), enquanto os **modelos cognitivos** (frames, *scripts*, cenários, modelos episódicos, esquemas, modelos mentais) controlam o conteúdo semântico do texto, isto é, o que está sendo ou será dito, os **modelos de contexto** controlam a maneira como os interlocutores vão formular tais conteúdos em função do contexto em que a interação se realiza.

Somada às características inerentes à etapa de identificação de conceitos da análise de assunto que trazem uma complexidade própria à atividade de indexação por envolver aspectos cognitivos, a etapa de seleção

de conceitos é influenciada notadamente pelos elementos de **política de indexação** e pelo perfil dos **usuários**, e em uma abordagem mais ampla, pela política de informação presentes no contexto institucional em que o indexador atua. Embora o assunto seja a informação relevante abordada no recurso informacional, a seleção do assunto sofre a influência da política de indexação do sistema de informação à qual pertence, pois a instituição decidirá, dentre outros elementos, se o tema extraído do documento terá um nível mais específico ou mais genérico com base no perfil da comunidade usuária. Nesse cenário,

A atuação profissional no tratamento temático da informação deve ser regida por uma política própria e única, que considere a instituição como um todo, seus usuários, colaboradores, estrutura física e financeira, entre outros. Como um processo que vai muito além de um fazer técnico – mas sim, uma atividade altamente cognitiva e humana, sendo, portanto, passível de influências sociais - a qualidade do tratamento temático da informação no contexto de bibliotecas universitárias depende ainda da coleção e da comunidade usuária local. [...] São diversos os fatores que afetam o processo de indexação, já que não existe uma única maneira de indexar corretamente um documento (TARTAROTTI, 2014, p. 228).

Assim, ao refletir os fundamentos teórico-metodológicos do campo da Organização do Conhecimento e, ao mesmo tempo, as necessidades profissionais, o estabelecimento de uma política de indexação formalizada em um manual de indexação contribui com o processo de melhorias contínuas na atuação profissional do indexador, pois fornece as diretrizes de indexação ao profissional indexador em seu contexto de atuação.

5 CONSIDERAÇÕES FINAIS

Como considerações finais, verifica-se que os aportes teóricos da Linguística Textual são relevantes e podem contribuir, em **termos teóricos**, para o fortalecimento da abordagem teórica da indexação do campo da Organização do Conhecimento (OC) e em **termos práticos**, para a atuação do indexador como um leitor profissional, reforçando ainda mais

a relação interdisciplinar entre a Ciência da Informação e a Linguística, em torno da representação e materialização de sentidos.

A complexidade inerente ao processo de análise de assunto – realizada durante a leitura documental, ou seja, com o texto “em mãos” – é ampliada ao lançarmos o olhar para a etapa de **seleção de conceitos**, que ocorre após a leitura do documento, ou seja, não mais com o texto. Isto porque na seleção de conceitos são vários os fatores interferentes. Desse modo, a etapa de seleção de conceitos é tão ou ainda mais complexa que a etapa de identificação de conceitos, por conter ainda mais elementos e fatores interferentes no processo, **fatores extrínsecos** que independem da leitura documental e, portanto, do contexto onde o profissional atua. Reforça-se ainda a visão de Dal'Evedove, Tartarotti e Fujita (2015, p. 612) sobre a importância de considerarmos tanto a abordagem sociocognitiva como a abordagem sociocultural nos estudos em torno da indexação, mais especificamente na etapa de seleção de conceitos que abriga a tematicidade extrínseca textual da análise de assunto, tanto na formação quanto na capacitação do indexador.

Diante destas reflexões em torno da tematicidade extrínseca textual da análise de assunto para indexação e do texto como objeto de estudo da Linguística Textual, recomenda-se estudos comparativos em torno dos fatores interferentes na análise de assunto, especialmente sobre as convergências e divergências entre os fatores interferentes na etapa de identificação de conceitos e os fatores interferentes na etapa de seleção de conceitos, bem como a relação entre a seleção de conceitos e a recuperação da informação. Além disso, pesquisas interdisciplinares que versem sobre a abordagem teórica da análise do discurso, tal como a abordagem bakhtiniana, podem contribuir com aportes teórico-práticos da tematicidade intrínseca e da tematicidade extrínseca na representação temática da informação no processo de indexação

Finalmente, ao *incorporar novas, diferentes e futuras vozes* (McTAVISH, 2014, p. 330) nos estudos sobre a análise de assunto do processo de indexação, acredita-se que a representação temática da informação nesta abordagem discursiva e contextual como um modelo alternativo permita realmente a recuperação dos recursos informacionais nos sistemas de recuperação da informação, contribuindo para o entendimento

dos fatores interferentes no processo de indexação e da melhoria da atuação do bibliotecário indexador como um leitor profissional.

REFERÊNCIAS

- ALBRECHTSEN, H. Subject analysis and indexing: from automated indexing to domain analysis. *The indexer*, London, v. 18, n. 4, p. 219-223, Oct. 1993.
- BABIK, W. *et al.* ISKO and Knowledge Organization's 25th anniversary: the future of Knowledge Organization and ISKO Panel Discussion. Reported by Rebecca Green. *Knowledge Organization*, Wurzburg, v. 41, n. 4, p. 327-331, 2014.
- BENTES, A. C. Linguística textual. *In*: MUSSALIM, F.; BENTES, A. C. (org.). *Introdução à linguística: domínios e fronteiras*. 9. ed. São Paulo: Cortez, 2011. v.1.
- CAVALCANTI, M. do C. *Interação leitor-texto: aspectos de interpretação pragmática*. Campinas: Ed. da Unicamp, 1989.
- CESARINO, M. A. N., PINTO, M. C. M. F. Análise de assunto. *Revista de Biblioteconomia de Brasília*, Brasília, v. 8, n. 1, p. 32-43, jan./jun.1980.
- CHU, C. M.; O'BRIEN, A. Subject analysis: the critical first stage in indexing. *Journal of Information Science*, Cambridge, v. 19, n. 6, p. 439-454, 1993.
- DALEVEDOVE, P. R.; TARTAROTTI, R. C.D.; FUJITA, M. S. L. Reflexões acerca do desenvolvimento de metodologias para análise de assunto. *In*: GUIMARÃES, J. A. C.; DODEBEI, V. (org.). *Organização do conhecimento e diversidade cultural*. Marília: ISKO-Brasil; FUNDEPE, 2015. p. 607-614. (Série: Estudos Avançados em Organização e Representação do Conhecimento, v. 3).
- DALEVEDOVE, R. C. *Compreensão de leitura em análise de assunto para identificação e seleção de conceitos*. 2002. 154 f. Trabalho de Conclusão de Curso (Graduação em Biblioteconomia) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2002.
- DISCINI, N. *A comunicação nos textos*. São Paulo: Contexto, 2007.
- FAIRTHORNE, R. A. Content analysis, specification, and control. *Annual Review of Information Science and Technology*, Medford, v. 4, p. 73-109, 1969.
- FARROW, J. F. A cognitive process model of document indexing. *Journal of Documentation*, London, v. 47, n. 2, p. 149-66, 1991.
- FUJITA, M. S. L. A identificação de conceitos no processo de análise de assunto para indexação. *Revista Digital de Biblioteconomia e Ciência da Informação*, Campinas, v. 1, n. 1, p. 60-90, dez. 2003.
- FUJITA, M. S. L. *Leitura em análise documentária: uma contribuição à formação do indexador*. Marília: UNESP; CNPq, 2001. 190 p. (Relatório Parcial de Pesquisa).

- FUJITA, M. S. L. Representação documentária no processo de indexação com o modelo de leitura documentária para textos científicos e livros: uma abordagem cognitiva com protocolo verbal. *Ponto de Acesso*, Salvador, v. 7, n. 1, p. 42-66, abr. 2013.
- FUJITA, M. S. L.; RUBI, M. P.; BOCCATO, V. R. C. As diferentes perspectivas teóricas e metodológicas sobre indexação e catalogação de assuntos. In: FUJITA, M. S. L. (org.). *A indexação de livros: a percepção de catalogadores e usuários de bibliotecas universitárias*. São Paulo: Cultura Acadêmica, 2009. p. 19-42.
- GIASSON, J. *A compreensão na leitura*. Lisboa: Asa, 1993.
- GIL LEIVA, I. *Manual de indización: teoría y práctica*. Gijón: Trea, 2008.
- GUIMARÃES, E. *Texto, discurso e ensino*. São Paulo: Contexto, 2009.
- KOCH, I. G. V. *Desvendando os segredos do texto*. 6. ed. São Paulo: Cortez, 2009.
- LARA, M. L. G. *A representação documentária: em jogo a significação*. 1993. 133 f. Dissertação (Mestrado em Ciência da Comunicação) – Escola de Comunicação e Artes, Universidade de São Paulo, São Paulo, 1993.
- MAI, J. E. Deconstructing the indexing process. *Advances in Librarianship*, New York, v. 23, p. 269-298, 2000.
- MAI, J. E. Semiotics and indexing: an analysis of the subject indexing process. *Journal of Documentation*, London, v. 57, n. 5, p. 591-622, 2001.
- MAI, J. E. The concept of subject: on problems in indexing. *Knowledge Organization for Information Retrieval: Proceedings of the 6th International Study Conference on Classification Research*, v. 6, p. 60-67, 1997a.
- MAI, J. E. The concept of subject in a semiotic light. Digital Collections: Implications for Users, Funders, Developers and Maintainers. *Proceedings of the ASIS Annual Meeting*, v. 34, p. 54-64, 1997b.
- McTAVISH, J. ISKO and Knowledge Organization's 25th anniversary: the future of Knowledge Organization and ISKO Panel Discussion. Reported by Rebecca Green. *Knowledge Organization*, v. 41, n. 4, p. 327-331, 2014.
- NEVES, D. A. B. Ciência da Informação e cognição humana: uma abordagem do processamento da informação. *Ciência da Informação*, Brasília, v. 35, n. 1, p. 39-44, jan./abr. 2006.
- NEVES, D. A. B.; DIAS, E. W.; PINHEIRO, Â. M. V. Uso de estratégias metacognitivas na leitura do indexador. *Ciência da Informação*, Brasília, v. 35, n. 3, p. 141-152, set./dez. 2006.
- OLSON, H. How we construct subjects: a feminist analysis. *Library Trends*, Champaign, v. 56, n. 2, p. 509-541, 2007.

- PINTO MOLINA, M. *Análisis documental: fundamentos y procedimientos*. 2. ed. rev. aum. Madrid: EUDEMA, 1993.
- RAJU, J.; RAJU, R. *Descriptive and subject cataloguing: a workbook*. Oxford: Chandos Publishing, 2006.
- SANTAELLA, L. *O que é semiótica*. 13. ed. São Paulo: Brasiliense, 1995.
- SMIRAGLIA, R. P. Bibliocentrism, cultural warrant, and the ethics of resource description: a case study. *Cataloging & Classification Quarterly*, New York, v. 47, n. 7, p. 671-686, 2009.
- SMIRAGLIA, R. P. Domain coherence within Knowledge Organization: people, interacting theoretically, across geopolitical and cultural boundaries. In: ANNUAL CAIS/ACSI CONFERENCE, 39., 2011, Canada. *Proceedings* [...]. Canada: University of New Brunswick, 2011. p. 1-6.
- TARTAROTTI, R. C. D. *Atuação bibliotecária no tratamento temático da informação em unidades informacionais: um estudo comparativo qualitativo-quantitativo*. 2014. 277 f. Dissertação (Mestrado em Ciência, Tecnologia e Sociedade) – Universidade Federal de São Carlos, São Carlos, 2014.
- TARTAROTTI, R. C. D., BOCCATO, V. R. C. A abordagem teórica bakhtiniana e o processo de indexação: diálogos. *Versão Beta: sob o signo da palavra*, São Carlos, ano 11, v. 74, p. 33-47, 2013.
- TAYLOR, A. G.; JOUDAREY, D. N. *The organization of information*. 3. ed. Wesport: Libraries Unlimited, 2009.
- TODD, R. J. Academic indexing: what's it all about? *The Indexer*, London, v. 18, n. 2, p. 101-104, 1992.
- VILELA, M.; KOCH, I. G. V. *Gramática da Língua Portuguesa*. Coimbra: Livraria Almedina, 2001.
- WELLISH, H. H. *Glossary of terminology in abstracting, classification, indexing and thesaurus construction*. 2. ed. Medford: American Society of Indexers, 2000.