



UNIVERSIDADE ESTADUAL PAULISTA  
"JÚLIO DE MESQUITA FILHO"  
Campus de Marília



**CULTURA  
ACADÊMICA**  
*Editora*

## **Análise documentária e semântica discursiva:**

contributos para o tratamento temático de objetos informacionais

Raimunda Fernanda dos Santos

Dulce Amélia de Brito Neves

Maria Elizabeth Baltar Carneiro de Albuquerque

**Como citar:** SANTOS, R. F.; NEVES, D. A. B.; ALBUQUERQUE, M. E. B. C. Análise documentária e semântica discursiva: contributos para o tratamento temático de objetos informacionais. *In:* FUJITA, M. S. L.; ALVES, R. C. V.; ALMEIDA, C. C. (org.). **Modelos de leitura Documentária para Indexação:** abordagens teóricas interdisciplinares e aplicações em diferentes tipos de documentos. Marília: Oficina Universitária; São Paulo: Cultura Acadêmica, 2020. p. 117-138.  
DOI: <https://doi.org/10.36311/2020.978-65-86546-07-1.p117-138>



All the contents of this work, except where otherwise noted, is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported.

Todo o conteúdo deste trabalho, exceto quando houver ressalva, é publicado sob a licença Creative Commons Atribuição - Uso Não Comercial - Partilha nos Mesmos Termos 3.0 Não adaptada.

Todo el contenido de esta obra, excepto donde se indique lo contrario, está bajo licencia de la licencia Creative Commons Reconocimiento-NoComercial-CompartirIgual 3.0 Unported.

# 5

## **ANÁLISE DOCUMENTÁRIA E SEMÂNTICA DISCURSIVA: CONTRIBUTOS PARA O TRATAMENTO TEMÁTICO DE OBJETOS INFORMACIONAIS**

*Raimunda Fernanda dos SANTOS*

*Dulce Amélia de Brito NEVES*

*Maria Elizabeth Baltar Carneiro de ALBUQUERQUE*

**RESUMO:** Discute aspectos relativos aos contributos da Semântica Discursiva para a produção significativa de metadados semânticos no processo de análise documental de recursos informacionais. Destaca que os percursos temático e figurativo possuem qualidades metodológicas a serem aplicadas na leitura documental, com isso faz-se necessário que o indexador reconheça as superestruturas responsáveis pela caracterização do tipo do objeto informacional, considerando as suas especificidades. Na revisão de literatura são apontadas pesquisas na área de Organização e Representação da Informação e do Conhecimento que constata a pertinência e adequação do percurso figurativo e do percurso temático de Greimas para a análise documental de objetos informacionais. Conclui ressaltando que os contributos da Semântica Discursiva para a análise documental resultam em novo método para extrair os conteúdos de recursos informacionais, trazendo como diferencial a redução do nível de subjetividade, tornando a indexação mais próxima possível de atender os parâmetros conceituais dos documentos.

**PALAVRAS-CHAVE:** Análise Documentária. Semântica Discursiva. Indexação.

<https://doi.org/10.36311/2020.978-65-86546-07-1.p117-138>

**ABSTRACT:** Discusses aspects related to the Discursive Semantics contributions to the significant production of semantic metadata in the process of documentary analysis of information resources. Emphasizes that the thematic and figurative paths have methodological qualities to be applied in documentary reading, so it is necessary that the indexer recognizes the superstructures responsible for characterizing the type of informational object, considering their specificities. Through a literature review are pointed researches in the area of Organization and Representation of Information and Knowledge that find the relevance and adequacy of the figurative and Greimas thematic path for the documentary analysis of informational objects. Concludes by emphasizing that the contributions of Discursive Semantics to documentary analysis result in a new method for extracting indicators of information resources, bringing as a differential the reduction of the subjectivity level, making the indexing closer to meet the conceptual parameters of the documents.

**KEYWORD:** Documentary Analysis. Discursive semantics. Indexing.

## 1 INTRODUÇÃO

No decurso da história, a representação está presente nas práticas sociais, na cultura popular e nos diversos contextos de construção e transfiguração da sociedade. Neste limiar, o conhecimento pode abranger o que o sujeito pode discernir logicamente e o que é possível ser observado em uma determinada realidade por meio da linguagem. Assim, as relações entre a humanidade e o conhecimento evidencia a necessidade do homem em representar e organizar o conhecimento com vistas à perpetuação, preservação, recuperação da memória social.

Partindo desse pressuposto, a representação é vista como uma ação criativa, e socialmente construída, que faz uso da linguagem e se configura como objeto de análise crítica e científica de uma realidade. A linguagem, por sua vez, é compreendida como o meio pelo qual o significado é produzido e intercambiado e se constitui como elemento fundamental para os sentidos e para a cultura (HALL, 2016).

Esse aspecto multifacetado da representação e da linguagem desafia aqueles que se propõem a estudá-las no contexto da Ciência da Informação através de pesquisas teóricas e práticas sob o viés da Organização e Representação da Informação e do Conhecimento. A representação no contexto da Ciência da Informação, além de se constituir como uma atividade pragmática, se amplia como campo de estudos teóricos e aplicados

que objetivam responder aos problemas e efeitos da informação a fim de tornar os registros informacionais acessíveis à sociedade.

A Ciência da Informação consiste em um campo interdisciplinar que se preocupa, principalmente, em tornar a informação acessível para aqueles que dela necessitam, tendo em vista que os itens informacionais, independente do seu formato<sup>21</sup> (áudio, texto, imagem, vídeo, etc.), devem passar por um conjunto de tarefas de armazenamento, organização e representação para que seja possível a sua recuperação, acesso e uso em sistemas de recuperação da informação. Nesse cenário, os objetos informacionais devem passar por esse conjunto de tarefas.

A indexação, considerada por Chaumier (1988) como operação mais importante da análise documentária, visa descrever os conteúdos dos itens informacionais por meio de termos, também denominados de descritores, que se configuram como, ou se propõem a ser, pontos de acesso para fins de disponibilização das informações pertinentes às necessidades dos usuários. Desse modo, a indexação condiciona os resultados de uma estratégia de busca e a qualidade dessa atividade reflete na recuperação, no acesso e no uso da informação pela geração hodierna e futura.

Assim, todas as fases do tratamento temático da informação, como análise conceitual e tradução, fazem parte das técnicas realizadas pelos profissionais da informação, sendo influenciadas por fatores linguísticos, cognitivos, lógicos, ideológicos, contextuais e culturais, simultaneamente.

Com isso, verifica-se a imprescindibilidade do compromisso do indexador na execução da análise documentária, tendo em vista que essa operação consiste em uma atividade intelectual considerada como o limite dos aparelhos tecnológicos, uma vez que o computador, por exemplo, ainda não realiza o trabalho humano de abstrair, distinguir, compreender e assimilar o que está presente nos itens informacionais (SANTOS, 2016).

Considerando que a indexação é uma operação relacionada à semântica, faz-se necessário discutir aspectos relativos aos contributos da Semântica Discursiva para a produção significativa de metadados semânticos no processo de análise documentária de recursos informacionais.

---

<sup>21</sup> Nesta pesquisa, o conceito de formato está relacionado à configuração física, forma dos objetos informacionais, também denominados como materiais informacionais ou itens informacionais.

## 2 ANÁLISE DOCUMENTÁRIA E TRATAMENTO TEMÁTICO DA INFORMAÇÃO

**A análise documentária centra-se nos procedimentos voltados para a identificação e seleção de conceitos para fins de representação do conteúdo dos itens informacionais. A indexação, operação que integra a análise documentária, está metodologicamente associada a duas etapas intelectualmente distintas: a *análise de assunto* e a *tradução*.**

A *análise de assunto* diz respeito à extração de conceitos que representem o conteúdo de um material informacional. Seguindo essa linha de raciocínio, essa etapa se inicia por meio da leitura técnica através da análise das principais partes ou dos elementos representativos do material informacional em questão. Nesta etapa, faz-se necessário que o indexador conheça o tipo e estrutura desse item informacional (imagem, texto, áudio, vídeo, etc.) para que a leitura seja iniciada para fins específicos. Em seguida, é efetuada a extração dos conceitos<sup>22</sup> constituindo-se o segundo momento da análise de assunto. Essa atividade é necessária para a representação do conteúdo do material informacional em termos adequados. Desse modo, o assunto do objeto informacional é determinado com base nesses conceitos e no contexto em que o mesmo se encontra inserido. Por isso, é importante verificar a complexidade do assunto do material informacional analisado para que seja possível elencar termos que representem efetivamente o seu conteúdo.

Após a etapa de análise do assunto é realizada a *tradução* desses descritores em termos provenientes de uma linguagem de indexação, ou seja, em termos autorizados para representar o conteúdo do objeto informacional analisado. Nesse entendimento, a etapa de *tradução* consiste na representação desse material informacional em termos provenientes de instrumentos de representação da informação, denominados de linguagens de indexação ou linguagens documentárias, como listas de cabeçalhos de assunto, tesouros, ontologias e taxonomias, por exemplo.

Destarte, as linguagens documentárias são instrumentos de controle terminológico que podem ser empregados na etapa de tradução dos conceitos, visando facilitar a comunicação entre a linguagem dos usuários e a linguagem do sistema, mediante a padronização de termos para a descrição dos conteúdos dos itens informacionais. Esses instrumentos

---

<sup>22</sup> Conceitos são, portanto, definidos como **unidades do conhecimento** identificadas através de enunciados verdadeiros sobre um item de referência e representados por uma forma verbal (termo ou palavra) (MEDEIROS, 1986, p. 136, grifo do autor).

são frutos de trabalhos colaborativos entre a Terminologia, Teoria da Classificação, Linguística Documentária, Ciência da Computação, Análise Documentária, Ciências Cognitivas, por exemplo.

Lancaster (2004) ressalta que esses instrumentos consistem em listas de termos ou descritores controlados, construídas para fins de tratamento e recuperação da informação objetivando controlar sinônimos<sup>23</sup>, diferenciar homógrafos<sup>24</sup> e agregar termos cujos significados apresentem alguma relação mais estreita entre si. Diante disso, observa-se que esses instrumentos objetivam controlar as dispersões lexicais existentes nas linguagens naturais; orientar o indexador no que concerne à escolha dos termos para representar o assunto dos materiais informacionais e indicar os melhores termos para a indexação.

Motta (1987, p. 25), conceitua a linguagem documentária como:

Um sistema de vocabulário baseado em conceitos, incluindo termos preferidos (descritores), termos não preferidos (não descritores) e suas inter-relações, que se aplica a um determinado ramo do conhecimento e que se destina a controlar a terminologia utilizada para a indexação/recuperação de documentos.

Com a ampliação da quantidade de sistemas de recuperação da informação e o incremento de tecnologias aplicadas às atividades de representação, organização, recuperação, acesso e uso da informação, surgiram outros instrumentos complementares aos sistemas de classificação como as listas de cabeçalhos de assunto, taxonomias, tesauros, ontologias, mapas conceituais, por exemplo. Esses instrumentos têm sido denominados genericamente por *Knowledge Organization Systems* (KOS), Sistemas de Organização do Conhecimento (SOC), em português.

Zeng (2008) apresenta uma lista de Sistemas de Organização do Conhecimento por meio das suas tipologias, a saber:

✓ Listas de Termos

- Listas propriamente ditas (*pick lists*): conjunto de termos em ordem sequencial;

<sup>23</sup> Palavras que possuem significados semelhantes.

<sup>24</sup> Termos que possuem a mesma grafia, mas apresentam pronúncia e significados diferentes.

- Dicionários: listas de termos organizados alfabeticamente com as suas respectivas definições, apresentando significados variados de acordo com a aplicação desses termos;
- Glossários: listas de termos organizados alfabeticamente com as suas respectivas definições;
- Anéis de sinônimos: conjunto de descritores considerados como equivalentes para fins de representação e recuperação da informação.

✓ *Modelos semelhantes a metadados:*

- Arquivos de autoridade: listas de termos empregados para fins de controle das variações de nomes para uma autoridade;
- Diretórios: listas de nomes e informações de contato associadas;
- Gazetteers: dicionários geoespaciais de tipos e nomes de lugares.

✓ *Classificação e categorização:*

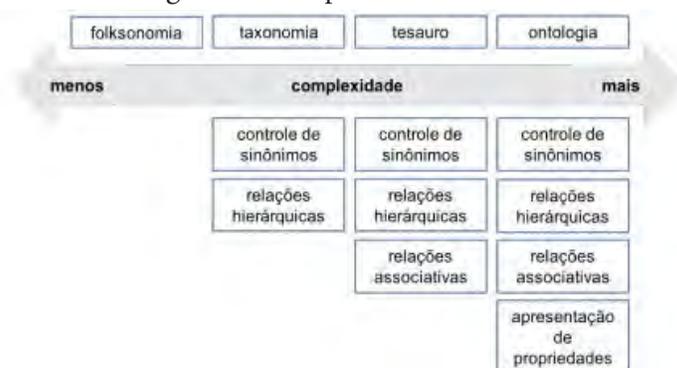
- Cabeçalhos de assunto: esquemas que apresentam um conjunto de descritores controlados para representar o assunto dos itens em uma coleção e dispõem um conjunto de diretrizes para combinar descritores em cabeçalhos compostos;
- Esquemas de categorização: esquemas de agrupamento “vagamente formados” (ZENG, 2008, p. 161);
- Taxonomias: separação de itens em grupos ordenados ou categorias, considerando características particulares;
- Esquemas de classificação: arranjos hierárquicos e facetados que apresentam notações alfabéticas ou numéricas visando representar assuntos genéricos.

✓ *Modelos de relações:*

- Tesouro: listas de termos autorizados de um domínio específico do conhecimento, constituída de base léxica composta por termos que apresentam relações hierárquicas, de equivalência e associativas entre eles.
- Redes semânticas: conjunto de descritores modelados como nós em uma rede de variados tipos de relações objetivando representar conceitos;
- Ontologias: modelos de conceitos específicos que representam relações complexas entre objetos, incluindo regras e axiomas ausentes em redes semânticas.

Considerando que as funções dos SOC diversificam conforme o grau de complexidade e o objetivo de sua aplicação, Moreira (2018) também representa em sua pesquisa a complexidade desses sistemas:

Figura 1- Complexidade dos SOC



Fonte: Moreira (2018, p. 102).

Não obstante, as vantagens das linguagens documentárias sobressaíam em relação à linguagem natural, em algumas situações a aplicação dessa última espécie de linguagem se faz apropriada mediante a aplicação dos métodos colaborativos de indexação. Lancaster (2004) afirma essa conveniência em ambientes que tratam materiais informacionais com conteúdos completos e altamente subjetivos que

viabilizam variadas interpretações - como é o caso das imagens - tendo em vista a sua capacidade de fomentar a exaustividade e especificidade da indexação e, conseqüentemente, recuperação da informação. Todavia, considera-se a importância de um tratamento linguístico mínimo dos descritores apresentados pelos usuários na atividade de representação do item informacional.

Em síntese, verifica-se que as etapas de *análise de assunto e tradução* são intelectualmente distintas, apesar de não serem diferenciadas de maneira clara nas produções científicas sobre a temática. Campos (1987) afirma que embora sejam operações distintas, elas se sucedem haja vista o caráter integrativo das mesmas, objetivando atender às necessidades de uma determinada comunidade de usuários.

Todavia, não se pode definir o assunto de um recurso informacional sem que sejam contemplados diversos aspectos, além do contexto no qual esse item informacional está inserido. Assim, corroboramos com Lucas (2000) quando a mesma revela que a leitura documentária não é resumida aos procedimentos e técnicas de tratamento dos itens informacionais, cujas etapas foram citadas anteriormente. Nas práticas de indexação devem ser consideradas as características do material informacional e os gestos de leitura desse sujeito enquanto leitor e usuário da informação, os quais são inerentes à sua formação discursiva configurada em processos de significação.

Entendendo isso, observa-se que a complexidade da indexação está associada ao fato de que a descrição do conteúdo intelectual, operação altamente subjetiva, pode ser realizada de maneira incorreta, tendo em vista a influência de fatores linguísticos, cognitivos, lógicos, ideológicos, contextuais, culturais, etc. Mesmo estando correta, essa atividade pode ser considerada insuficiente para demarcar congruentemente os descritores, pactuando com as necessidades dos usuários (BLAIR, 2003).

Fica claro, portanto, que a indexação é uma atividade de representação e que, por sua vez, está atrelada ao conceito de substituição, de “estar no lugar de outra coisa”. Para fins de realização dessa atividade, o indexador se coloca entre o objeto informacional a ser representado e o usuário final. Por esse motivo, essa representação não consiste na apresentação do registro do conhecimento propriamente dito.

Seguindo essa linha de raciocínio, Pato (2015, p. 73) atenta que:

[...] quem ou aquele que representa se apresenta no lugar de outrem. Ao se colocar no papel de alteridade, do autor, aquele que agrega polifonicamente ao seu texto os textos de vários outros sujeitos, o esforço do indexador será no sentido de traduzir, concorrer com as representações desse(s) outrem(ens) e alimentar um sistema de organização da informação.

Sobre esse prisma, a indexação consiste em uma prática de interpretação e, ao mesmo tempo de representação, a qual, por meio de ações seletivas, privilegia certos elementos acerca de objetos informacionais e oculta outros. Nesse cenário, o indexador exerce o seu papel crucial de mediação entre o item informacional e o usuário final, buscando apurar termos representativos mediante a pluralidade de sentidos existentes no registro do conhecimento. Isso ocorre pelo fato de que, ao interpretar o conteúdo intelectual deste, são acionadas variadas leituras e esse profissional é quem apresenta o que há de informação nesse registro, captando alguns aspectos e ignorando outros.

Concorda-se com o pensamento de Pato (2015, p. 73) quando o mesmo enfatiza a importância do indexador, considerando o fato de que esse profissional é coparticipante do processo de atribuição de significados dos itens informacionais, os quais se relacionam com o contexto social, cultural e econômico do qual emerge, pois, “a leitura de um texto [ou registro do conhecimento] depende da interação entre leitor, texto, autor e mundo socialmente estruturado”.

Essa assertiva também se aproxima das ideias de Lucas (2000, p. 49) quando essa autora reflete que a indexação “[...] encaminha o leitor por vários percursos, enriquece ou confunde os possíveis olhares que serão lançados à obra”, uma vez que o olhar leitor do profissional silencia ou expõe diferentes leituras.

Sobre esse prisma, os traços semânticos existentes no material informacional e que são selecionados pelo indexador revelam a visão de mundo desse profissional, por isso essa atividade é considerada difícil de ser ensinada, haja vista que não há um manual a ser seguido que inclua os aspectos subjetivos para o desenvolvimento da operação.

Em síntese, a indexação assume uma posição crucial e condicionante na recuperação da informação, ela “é um fator de auxílio na comunicação entre o sistema de recuperação da informação e o usuário que deseja satisfazer sua necessidade informacional” e deve ser encarada como a materialização de uma dinâmica de natureza ética, refletindo universos culturalmente distintos (FRANCELIN; PINHO, 2011, p. 61).

Ainda de acordo com esses autores, as competências específicas para os profissionais da informação, em especial aqueles que lidam com o tratamento temático da informação, não se esgotam no “como fazer”, mas encontram terreno fértil de indagação do “por que” fazer e no “para que” fazer. Principalmente porque é necessário considerar que tanto o produto quanto o processo de indexação devem estar imbuídos da preocupação com a sua utilidade e com a sua adequação para determinada comunidade ou grupo-alvo.

Nesse entremeio o sujeito cognoscente, ou seja o interagente, conforme sinaliza Capurro (2003, *online*), é essencial neste processo, pois o processo de tradução é algo que a final de contas, terá a sua razão de existir quando direcionado a seu público, pois, em essência se apresenta para organizar e se organiza para recuperar a informação seja qual for o suporte, a mídia, a ambiência e a maneira como a comunicação é empregada.

### **3 SEMÂNTICA DISCURSIVA E ANÁLISE DOCUMENTÁRIA**

A representação da informação consiste em uma prática de enunciação das propriedades de um objeto informacional (unidade de informação organizável) ou das relações desse objeto com outros que o identificam, através de operações que visam a sua descrição.

Conforme Fogl (1979, p. 21) a informação se configura como uma unidade de três elementos, são eles:

1. Conhecimento (conteúdo da informação);
2. Linguagem (instrumento de expressão de itens de informação);
3. Suporte (objetos materiais ou energia).

Para esse autor, a fonte de origem da informação é o conhecimento, portanto, ela pode ser analisada e representada segundo pontos de vistas semânticos (cognitivos), pragmáticos (reais), inferenciais (juízos de valor) e de acordo com o formato utilizado (textos, imagens, áudios, vídeos, etc.), levando em conta as propriedades do seu conteúdo e a sua função social.

Nesse processo, vê-se que o registro informacional, independente do meio, suporte ou ambiente, é o aspecto central da representação da informação, colocando tanto o produtor, quanto o indexador e o interagente (público-alvo), em uma situação de constante vigilância, pois os estoques oriundos desse processo desafia constante e cotidianamente o profissional da informação, uma vez que a relação dialógica do acesso e uso da informação é cada vez mais dependente do ato de representar. Por isso, Gaudêncio e Alburquerque (2014, p. 47), afirmam que “a representação da informação faz parte da vida do profissional da informação que lida diariamente com as questões documentárias e que busca recuperar com mais eficácia determinados conteúdos”.

A análise documentária diz respeito a uma operação associada à semântica que se concentra nas informações extraídas de um recurso informacional e dos referentes internos (cognição) do(s) sujeito(s) que o analisa. Nesse entendimento, independente da configuração desse recurso – seja de caráter visual, textual, sonoro, etc. ele pode ser lido e analisado. Para a realização dessa operação é importante que o sujeito tenha conhecimento prévio acerca do conteúdo presente no item informacional a ser indexado. Todavia, o conhecimento prévio do sujeito não deve ser fator condicionante para a execução dessa operação, uma vez que, para auxiliar na compreensão do documento a ser analisado podem ser extraídas informações extrínsecas em outras fontes de informação.

Sob esse viés, nesta pesquisa busca-se estudar as contribuições da semântica no nível discursivo, contemplando os procedimentos de tematização e figurativização como subsídios para leitura documentária de objetos informacionais.

A semântica, enquanto área do conhecimento, emergiu no final do século XX no intuito de responder as demandas da sociedade científica por meio do estudo do significado ou teoria da significação (FIORIN, 1999). Desse modo, o seu foco não está ligado apenas às palavras, mas

também às orações, gestos, símbolos, imagens e entre outros elementos de significação.

De acordo com Greimas (1966, p. 11) a significação consiste em elemento central das preocupações da semântica, uma vez que essa estuda os problemas relacionados ao significado de todas as coisas que são pensadas, produzidas e transmitidas.

Desse modo, a atividade de representação advém do processo de significação e esse, por sua vez, contribui para a atividade de organização - o que justifica a relevância e necessidade das investigações de natureza semântica nos processos de organização e tratamento da informação na Ciência da Informação.

A semântica atribui sentido aos signos, esclarecendo o que é apreendido no mundo das coisas, e se configura como um componente semiótico-discursivo (ALBUQUERQUE, 2011).

A semiótica, por sua vez, consiste no estudo da significação, a qual é compreendida como a relação dependente entre o conteúdo e a expressão de signos verbais e não-verbais (BATISTA, 2001). Greimas e Coutés (2008, p. 447) afirmam que a semiótica realiza uma “operação que, ao instaurar uma relação de pressuposição recíproca entre a forma e a expressão e a do conteúdo” e entre “o significante e o significado”, “[...] produz signos, resultando em uma semiose<sup>25</sup>.

Santaella (1993) infere que a semiótica estuda todos os tipos possíveis de ações sígnicas, ou seja, o centro da sua teorização são as representações de todas as linguagens, uma vez que estuda os meios pelos quais o homem se comunica, sendo eles verbais ou não (gestos, ideias, símbolos, imagens, etc.).

Para Albuquerque (2011), a semiótica de origem francesa, também conhecida como semiótica greimasiana, parte da concepção de signo criada por Hjelmslev (1975), a qual enriquece o conceito de signo delineado por Saussure (2004, p. 80) como “uma entidade psíquica” que se associa a dois elementos: o significante (conceito) e o significado (imagem acústica).

---

<sup>25</sup> De acordo com Batista (2001, p. 146), a semiose é o “[...] o processo de produção, acumulação e transformação da função semiótica”. Assim, a semiose se constrói e se manifesta ao longo do discurso e só atinge a sua completude no percurso sintagmático do discurso por inteiro.

Enquanto abordagem metodológica, a Semântica Discursiva foi idealizada por Algirdas Julien Greimas e busca atribuir sentido ao discurso com base no modelo teórico-metodológico denominado Percurso Gerativo da Significação (PGS) ou Percurso Gerativo do Discurso (PGD), cujo foco é o estudo da produção e interpretação de sentidos. Embora a noção de signo seja considerada, Greimas não parte do signo para estruturar e propor a sua metodologia, uma vez que são levadas em consideração as denominadas “figuras”.

Para Fiorin (1999, p. 38) “[...] a noção de percurso gerativo de sentido constitui um simulacro metodológico para explicar o processo de entendimento, em que o leitor precisa fazer abstrações, a partir da superfície do texto, para poder entendê-lo [...]”. Portanto, esse modelo teórico-metodológico representa um processo que parte do nível simples ao mais complexo, são eles: *profundo (ou fundamental), narrativo e discursivo*.

Cada nível supracitado contém um elemento sintático e um elemento semântico, considerando que a sintaxe busca estudar questões relativas à estrutura do vocabulário e a semântica busca focar “[...] nas regras que presidem às relações entre os vocábulos, as construções das orações e as relações interacionais.” (ANTÔNIO, 2008, p. 12).

Seguindo essa linha de raciocínio, aqui são estudadas as contribuições da semântica de nível discursivo, no intuito de considerar os procedimentos de tematização e figurativização como subsídios para a construção da metodologia de indexação de imagens, em especial de xilogravuras.

A Semântica Discursiva integra duas formas básicas de discursos: os discursos concretos e os discursos abstratos. Eles não são antagônicos entre si, pois coexistem e estão interligados apresentando um caráter lógico que percorre do concreto (figuras) para o abstrato (temas). Desse modo, o discurso concreto construído com figuras é denominado figurativo. E o discurso abstrato edificado por temas é definido como temático.

Para Fiorin e Savioli (2003, p. 89) os textos figurativos “[...] produzem um efeito de realidade, e por isso representam o mundo, criam uma imagem do mundo, com seus seres, seus acontecimentos.”. Os textos temáticos, por sua vez, “explicam as coisas do mundo, ordenam-nas,

classificam-nas, interpretam-nas, estabelecem relações e dependências entre elas, fazem comentários sobre suas propriedades”.

Nesse prisma, os textos figurativos dirigem-se a algum aspecto da realidade, criando imagens da mesma, bem como dos seres, das coisas, das ações, acontecimentos e qualidades associados a ela. Para Antônio (2008, p. 66) “[...] quando fala-se mundo real devemos pensar também no mundo construído.”.

Tal como as figuras, os temas “[...] também se encadeiam em percursos, isto é, em conjuntos organizados. São os percursos temáticos. Para apreender o tema geral, é preciso perceber esse encadeamento dos temas e depreender a unidade subjacente à diversidade.” (FIORIN; SAVIOLI, 2003, p. 101).

Assim, os temas expressam elementos abstratos por meio de uma natureza conceitual representada por classes e categorizações, conforme explica Barbalho (2006, p. 88):

A tematização expressa elementos abstratos buscando explicar a realidade e representar o mundo através de um investimento conceptual. Os temas organizam, categorizam e ordenam a realidade significante de modo a permitir sua interpretação. Figuras e temas são para manter a coerência interna do texto, necessitam seguir um percurso ou encadeamento lógico de modo a gerar sentido. As figuras devem ser vistas através do conjunto por elas composto e não isoladamente.

Albuquerque (2011, p. 173, grifo da autora), por sua vez, apresenta as relações entre os níveis temático e figurativo da Semântica Discursiva:

A *semântica discursiva* tem como componente a *tematização* – elementos abstratos presentes no texto – e a *figurativização* – elementos concretos presentes no texto – que dão concretude ao tema. As figuras do texto formam uma rede, uma trama que, para entendê-las, é necessário conhecer o primeiro nível temático assim como o nível figurativo em que palavras e expressões apresentam traços comuns de significação, que podem ser agrupados.

Com isso, verifica-se que a tematização e a figurativização são componentes da Semântica Discursiva e consistem em procedimentos semânticos da discursivização. No Quadro a seguir são apresentadas as principais características dos textos figurativos e temáticos:

Quadro 1 - Principais características dos textos figurativos e temáticos

TEXTOS FIGURATIVOS	TEXTOS TEMÁTICOS
*Apresentam função descritiva ou representativa	*Elementos abstratos no texto cuja função é explicar a realidade
*Representam o mundo	*Classificam e ordenam a realidade
*Criam imagem do mundo	*Estabelecem relações e dependências
*Criam imagem dos seres	*Apresentam função predicativa ou interpretativa
*Criam os acontecimentos do mundo	*Apresentam comentários sobre as propriedades do mundo
*Referem-se ao concreto presente no texto	*Apresentam temas genéricos (classes) que contempla assuntos principais
*São palavras ou expressões correspondentes ao mundo natural	*Apresentam coerência ao texto principal
*Constroem simulacro da realidade	

Fonte: Adaptado de Antônio (2008).

A relação entre as figuras apresentadas para a construção do sentido e, conseqüentemente, a identificação do tema atinente a elas é denominada por Albuquerque (2011) de “encadeamento de figuras”. Portanto, a sequência dessas figuras resultante do seu agrupamento produz uma rede relacional, denominada no sistema semiótico como *percurso figurativo*, e expressa os temas subentendidos no texto - tornando viáveis as possibilidades significativas em um dado contexto.

Depreende-se, portanto, que os estudos que possuem maior conexão com a Semântica Discursiva e que contribuem significativamente para a Ciência da Informação são aqueles relacionados à Organização e Representação da Informação e do Conhecimento, especificamente sobre a análise documentária. Esses estudos, em conjunto, podem

colaborar para o processo de análise e síntese do sentido das informações contidas em textos verbais e não verbais, considerando que os sujeitos possuem mentes interpretadoras que executam processos semióticos, interpretam e ressignificam diferentes visões de mundo apresentadas por estruturas conceituais.

Diante do exposto, as metodologias de análise documentária, ao considerarem a Semântica Discursiva, devem contemplar a discursivização semântica, privilegiando os processos de figurativização e tematização. Nesse limiar, a análise dos objetos informacionais deve estar pautada em níveis descritivos (figuras) e interpretativos (temas). As figuras consideram os aspectos denotativos, ou seja, aquilo que o item informacional (imagem, texto, áudio, vídeo, etc.) mostra de forma explícita (elementos concretos); e os temas levam em conta os aspectos conotativos (elementos abstratos), ou seja, aquilo que pode ser simulado pelo objeto informacional – associando-se ao processo de tematização.

Os temas não necessariamente precisam refletir o contexto ou significado “real” do material informacional, mas o que esse pode (res) significar na concepção de quem a analisa. Eles são extraídos de classes ou categorizações, podendo ser provenientes de instrumentos de controle terminológicos como tesouros, ontologias, taxonomias, etc. ou até mesmo da linguagem natural como as *tags* (etiquetas) atribuídas pelos usuários em um dado sistema colaborativo.

Pesquisas na Ciência da Informação, em especial na área de Organização e Representação da Informação e do Conhecimento, evidenciam a pertinência e adequação do percurso figurativo e do percurso temático de Greimas, onde ocorrem os processos de tematização e figurativização, para a análise documentária de objetos informacionais como textos e imagens, apontando diretrizes e metodologias para a indexação desses recursos. Tal aspecto contribui para a interdisciplinaridade das pesquisas sobre a área supracitada e a linguística.

Albuquerque (2011) em sua pesquisa constata a eficácia da Semântica Discursiva para a indexação textos, estabelecendo como recorte os folhetos de cordel, a partir da identificação de figuras e temas que subsidiam a indexação dessa espécie documental. Além disso, essa pesquisadora propôs temas e 27 classes temáticas para a organização de acervos que possuem folhetos de cordel. De outro modo disposto,

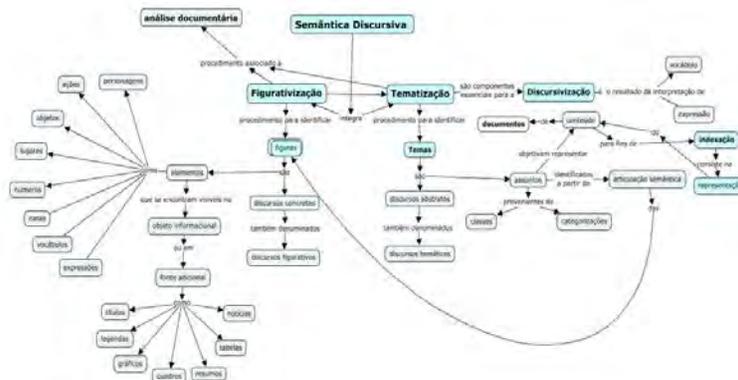
Gandier e Pinho (2018) ressaltam a importância da Semântica Discursiva para a análise documental de textos ficcionais, especificamente crônicas, constatando que esse método pode ser satisfatoriamente aplicado a essa tipologia textual.

De maneira complementar, Santos (2019) em sua pesquisa doutoral constata as contribuições da Semântica de nível discursivo para a indexação de imagens, estabelecendo como recorte as xilogravuras de cordel. Essa autora enfatiza que os métodos e técnicas de indexação de imagens existentes na literatura pautam-se, principalmente, na análise descritiva de fotografias em diferentes níveis de detalhamento, não contemplando elementos intrínsecos de imagens específicas como as xilogravuras, por exemplo. **Diante disso, surge a necessidade de encontrar alternativas que auxiliem o trabalho do profissional da informação na prática de indexação desses recursos imagéticos, fornecendo caminhos** para que os temas desses objetos informacionais atendam os parâmetros conceituais das imagens, reduzindo o nível de dispersão interpretativa do profissional indexador - aspecto inerente à operação.

Diante do exposto, percebe-se que os percursos temático e figurativo possuem qualidades metodológicas a serem aplicadas na leitura documentária. Para tanto, faz-se necessário que o indexador reconheça as superestruturas, ou seja, as estruturas globais responsáveis, independentemente do conteúdo, pela caracterização do tipo do objeto informacional, considerando as suas especificidades. Tal processo contribui para que esse profissional reconheça quais os pontos de acesso no objeto informacional que facilitam a identificação das figuras e dos temas para fins de indexação, podendo utilizar fontes adicionais como subsídios para o reconhecimento do(s) tema(s) a ser(em) utilizado(s) na indexação.

A Figura a seguir apresenta de maneira sistematizada os contributos da Semântica Discursiva, através dos procedimentos de figurativização e tematização, para a análise documentária a partir dos aspectos discutidos anteriormente.

Figura 2 - Contributos da Semântica Discursiva para a análise documentária



Fonte: Elaborado pelas autoras com base nas ideias de Santos (2019).

Conforme elucidado anteriormente, a Semântica Discursiva integra os discursos figurativos (concretos) e os discursos abstratos (temáticos). Aqui os discursos figurativos são tratados como figuras e correspondem aos elementos que se encontram visíveis ou descritos no item informacional, como por exemplo: personagens, ações, objetos, lugares, números, vocábulos, expressões, etc. Seguindo essa lógica, o procedimento de análise e identificação dessas figuras no objeto informacional é denominado de figurativização.

Já os discursos abstratos correspondem aos temas, os quais são identificados a partir da confluência ou da imbricação das figuras identificadas em um dado recurso informacional, no intuito de representar o seu conteúdo por meio de conceitos oriundos de classes e categorizações. Tal procedimento é denominado de tematização.

As figuras e os temas também podem ser identificados e extraídos em fontes adicionais relacionadas ao recurso informacional, como por exemplo: títulos, legendas, gráficos, quadros, resumos, tabelas, sinopses, etc. Nesse caso, deve-se considerar a relação semântica entre as figuras extraídas dessas fontes e o material informacional analisado.

Portanto, a figurativização e a tematização são procedimentos que pertencem à Semântica Discursiva e são componentes essenciais ao processo discursivização - resultado da interpretação de um vocábulo

ou expressão com base em tais procedimentos. Por exemplo, o vocábulo “*tortura*” figurativiza o tema “*sofrimento*”, que significa sensação consciente ou inconsciente de dor, mal-estar ou infelicidade. Portanto, “*tortura*” é uma figura associada ao tema “*sofrimento*”, cujo significado apontado anteriormente corresponde à discursivização.

Desse modo, considera-se também a importância de o indexador conhecer as características do sistema, bem com as funções e usos do recurso a ser indexado no contexto da unidade de informação. Esses aspectos também podem ser aplicados e/ou aperfeiçoados para contribuir no processo de análise documentária de diferentes tipologias de recursos (imagens, textos, áudios, vídeos, etc.).

Portanto, observa-se que a Ciência da Informação não pode se isentar dos estudos sobre as tendências contemporâneas de representação da informação que visam contribuir para o desenvolvimento da competência dos profissionais da informação, pesquisadores e usuários que lidam com as práticas de representação temática de recursos informacionais.

#### **4 CONSIDERAÇÕES FINAIS**

A análise documentária e a linguística integram ferramentas que podem fornecer subsídios teóricos e práticos para a correta interpretação do conteúdo de itens informacionais, objetivando compreender os sentidos e significados dos assuntos explorados nos mesmos para fins de indexação, recuperação, acesso de uso da informação.

Sujeitos que fazem parte de determinadas culturas podem não conseguir (ou apresentar dificuldades para) reconhecer, assimilar e interpretar os personagens, elementos e ações retratadas em objetos informacionais pelo fato de não possuírem conhecimentos adequados para entendê-los. Isso se deve ao fato de que os indivíduos atribuem sentido às coisas pela maneira como as utilizam e associam em suas práticas cotidianas. Por isso a importância de considerar as novas metodologias que contribuem para a análise documentária, no intuito de contribuir para a produção significativa de metadados semânticos.

Com base nessa perspectiva, a análise documentária deve estar pautada, primeiramente, no objeto informacional e na articulação

semântica do mesmo com o conteúdo das fontes adicionais relacionadas ao recurso (como o título, legendas, resumos, sinopses, quadros, tabelas, *tags* atribuídas pelos usuários em um dado sistema, etc.), contemplando os procedimentos de figurativização e tematização da Semântica Discursiva. Portanto, esta proposta está diretamente associada à indexação, não incluindo procedimentos voltados para a representação descritiva desses recursos informacionais.

Diante do exposto, acredita-se que os contributos da Semântica Discursiva para a análise documentária resultam em uma nova maneira de extrair os conteúdos de recursos informacionais, trazendo como diferencial a redução do nível de subjetividade, tornando a indexação mais próxima possível de atender os parâmetros conceituais dos documentos. Para tanto, a criação de uma política de indexação em consonância com as especificidades e singularidades dos recursos informacionais a serem tratados, consiste em um fator importante para a execução de um trabalho eficiente em bancos de dados e sistemas de recuperação da informação.

## REFERÊNCIAS

- ALBUQUERQUE, M. E. B. C. *Literatura popular de Cordel: dos ciclos temáticos à classificação bibliográfica*. 2011. Tese (Doutorado em Letras) – Universidade Federal da Paraíba, João Pessoa, 2011. Disponível em: <https://repositorio.ufpb.br/jspui/handle/tede/6183>. Acesso em: 26 abr. 2019.
- ANTONIO, D. M. *O Percurso Gerativo de Sentido aplicado à análise documental de textos narrativos de ficção: perspectivas de utilização em bibliotecas universitárias*. 2008. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2008. Disponível em: [http://bdtd.ibict.br/vufind/Record/UNSP\\_af8c4a251e7e06516d941efe1b9e8c22](http://bdtd.ibict.br/vufind/Record/UNSP_af8c4a251e7e06516d941efe1b9e8c22). Acesso em: 26 abr. 2019.
- BARBALHO, C. R. S. Fazer semiótico: subsídios para exame do espaço concreto. *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação*, Florianópolis, 2. n. esp., p. 79-96, 2006.
- BATISTA, M. F. B. de M. O discurso semiótico. In: ALVES, E.; CHRISTIANO, M. E. (org.). *Linguagem em foco*. João Pessoa: Editora Universitária; Ideia, 2001. p. 120-139.
- BLAIR, D. C. Information retrieval and the philosophy of language. *Annual Review of Information Science Technology*, Medford, v. 37, p. 3-50, 2003.

- CAMPOS, A. T. A indexação. *Revista de Biblioteconomia de Brasília*, Brasília, v. 15, n. 1, p. 69-72, jan./jun. 1987. Disponível em: <http://www.brapci.ufpr.br/documento.php?dd0=0000003009&dd1=33e85>. Acesso em: 26 abr. 2019.
- CAPURRO, R. Epistemologia e ciência da informação. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 5., 2003, Belo Horizonte. *Anais [...]*. Belo Horizonte: ANCIB, 2003. Disponível em: <https://bit.ly/2qQTUZZ>. Acesso em: 26 abr. 2019.
- CHAUMIER, J. Indexação: conceito, etapas, instrumentos. *Revista Brasileira de Biblioteconomia e Documentação*, São Paulo, v. 21, n. 1/2, p. 63-79, jan./jun. 1988.
- FIORIN, J. L. *Elementos de análise do discurso*. São Paulo: Contexto, 1999.
- FIORIN, J.; SAVIOLI, F. *Para entender o texto*. São Paulo: Ática, 2003.
- FOGL, J. Relations of the concepts 'information' and 'knowledge'. *International Fórum on Information and Documentation*, The Hague, v. 4, n. 1, p. 21-24, 1979.
- FRANCELIN, M. M.; PINHO, F. A. *Conceitos na Organização do Conhecimento*. Recife: Editora Universitária da UFPE, 2011.
- GANDIER, A. M.; PINHO, F. A. A importância da semântica discursiva para a análise documental: um estudo em texto ficcional. *Brazilian Journal of Information Science*, Marília, v. 12, n. 2, p. 13-23, 2018. Disponível em: <http://www2.marilia.unesp.br/revistas/index.php/bjis/article/view/7897>. Acesso em: 26 abr. 2019.
- GAUDÊNCIO, S. M.; ALBUQUERQUE, M. E. B. C. Diálogos teóricos em representação da informação. In: ALVES, E. C. et al. (org.). *Práticas e abordagens contemporâneas em ciência da informação*. João Pessoa: Ed. da UFPB, 2014. p. 57-78.
- GREIMAS, A. J. Condições de uma semântica científica. In: GREIMAS, A. J. *Semântica estrutural*. Tradução Haquira Osakabe. 2. ed. São Paulo: Cultrix; USP, 1966. p. 28-53.
- GREIMAS, A. J.; COURTÉS, J. *Dicionário de semiótica*. São Paulo: Contexto, 2008.
- HALL, S. *Cultura e Representação*. Tradução Daniel Miranda e Wiliam Oliveira. Rio de Janeiro: PUC-Rio; Apicuri, 2016.
- HJELMSLEV, L. *Prolegômenos a uma teoria da linguagem*. São Paulo: Perspectiva, 1975.
- LANCASTER, F. W. *Indexação e resumos: teoria e prática*. 2. ed. Brasília: Briquet de Lemos, 2004.
- LUCAS, C. R. *Leitura e interpretação em Biblioteconomia*. Campinas: Ed. da Unicamp, 2000.
- MEDEIROS, M. B. B. M. Terminologia Brasileira em Ciência da Informação: uma análise. *Ciência da Informação*, Brasília, v. 15, n. 2, p. 135-142, jul./dez. 1986.

Disponível em: <http://ridi.ibict.br/bitstream/123456789/283/1/BRASCHERCI1986.pdf>. Acesso em: 26 abr. 2019.

MOREIRA, W. *Sistemas de Organização do Conhecimento: aspectos teóricos, conceituais e metodológicos*. 2018. Tese (Livre-Docência em Sistemas de Organização do Conhecimento) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2018. Disponível em: [http://repositorio.unb.br/bitstream/10482/7465/1/2010\\_ElianaCarlan.pdf](http://repositorio.unb.br/bitstream/10482/7465/1/2010_ElianaCarlan.pdf). Acesso em: 26 abr. 2019.

MOTTA, D. F. *Método Relacional como nova abordagem para a construção de tesaurus*. Rio de Janeiro: SENAI/DN/DPEA, 1987.

PATO, P. R. G. *Imagens: polissemia versus indexação e recuperação da informação*. 2015. Tese (Doutorado em Ciência da Informação) – Faculdade de Ciência da Informação, Universidade de Brasília, Brasília, DF, 2015. Disponível em: <http://repositorio.unb.br/handle/10482/19050>. Acesso em: 26 abr. 2019.

SANTAELLA, L. *A percepção*. São Paulo: Experimento, 1993.

SANTOS, R. F. *Modelos colaborativos de indexação social e a sua aplicabilidade na Base de Dados Referencial de Artigos de Periódicos em Ciência da Informação (BRAPCI)*. 2016. Dissertação (Mestrado em Ciência da Informação) - Universidade Federal de Pernambuco, Recife, 2016. Disponível em: <https://repositorio.ufpe.br/handle/123456789/17218>. Acesso em: 26 abr. 2019.

SANTOS, R. F. *Indexação de xilogravuras em versos: a representação entre o real e o imaginário coletivo*. 2019. Tese (Doutorado em Ciência da Informação) - Universidade Federal da Paraíba, João Pessoa, 2019.

SAUSSURE, F. de. *Course in general linguistics*. Paris: Payot, 2004.

ZENG, M. L. Knowledge Organization Systems (KOS). *Knowledge Organization*, Baden-Baden, v. 35, n. 2/3, p. 160-182, 2008.