



# TENDÊNCIAS ATUAIS E PERSPETIVAS FUTURAS EM ORGANIZAÇÃO DO CONHECIMENTO

ATAS DO III CONGRESSO ISKO ESPANHA-PORTUGAL  
XIII CONGRESSO ISKO ESPANHA

*Universidade de Coimbra, 23 e 24 de novembro de 2017*

Com a coordenação de

---

Maria da Graça Simões, Maria Manuel Borges

TÍTULO

Tendências Atuais e Perspetivas Futuras em Organização do Conhecimento: atas do III Congresso ISKO Espanha e Portugal - XIII Congresso ISKO Espanha

COORDENADORES

Maria da Graça Simões  
Maria Manuel Borges

EDIÇÃO

Universidade de Coimbra. Centro de Estudos Interdisciplinares do Século XX - CEIS20

ISBN

978-972-8627-75-1

ACESSO

<https://purl.org/sci/atas/isko2017>

COPYRIGHT

Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

OBRA PUBLICADA COM O APOIO DE



FLUC FACULDADE DE LETRAS  
UNIVERSIDADE DE COIMBRA



CEIS 20  
CENTRO DE ESTUDOS  
INTERDISCIPLINARES  
DO SÉCULO XX  
UNIVERSIDADE DE COIMBRA

**FCT**  
Fundação para a Ciência e a Tecnologia  
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

PROJETO UID/HIS/00460/2013



## TEORIAS DA LINGUAGEM NOS ESTUDOS DE INDEXAÇÃO: UMA ANÁLISE A PARTIR DOS PERIÓDICOS *JOURNAL OF DOCUMENTATION* E *KNOWLEDGE ORGANIZATION*

Raquel Luise Pret<sup>1</sup>, Rosa Inês de Novais Cordeiro<sup>2</sup>

<sup>1</sup>Professora Assistente do Departamento de Ciência da Informação da Universidade Federal Fluminense (UFF/Brasil), [pretraquel@vm.uff.br](mailto:pretraquel@vm.uff.br)

<sup>2</sup>Professora Associada do Departamento de Ciência da Informação da Universidade Federal Fluminense (UFF/Brasil), [rosacordeiro@vm.uff.br](mailto:rosacordeiro@vm.uff.br)

**RESUMO** Este artigo é resultado de um estudo analítico sobre a influência das teorias linguísticas, em especial da Semântica, nas abordagens sobre indexação. A intenção é mostrar como o Realismo e o Mentalismo enquanto paradigmas que influenciam a compreensão da linguagem no campo da Ciência da Informação. No Realismo, a linguagem expressa o mundo real. O sentido é aquilo que descreve a realidade. É possível identificar conceitos, leis, lógicas universais porque as coisas do mundo são criadas a partir da mesma substância. Dessa forma, a vocação da linguagem é declarar, informar sobre elas. Já no Mentalismo, de inspiração aristotélica, o universal é a mente. A faculdade racional subordina a linguagem a comunicar o pensamento. Estes dois paradigmas têm grande influência nas clássicas teorias linguísticas como a Semântica Formal de Frege e a Gramática Gerativa de Chomsky. A Lógica Representacional da Semântica fregeana prioriza as relações formais, ou seja, os aspectos objetivos do significado. A notação valida sentenças a partir de critérios que estabelecem o que é verdadeiro ou falso, eliminando as dúvidas e os mal entendidos da linguagem natural. De forma semelhante, a teoria da Gramática Gerativa de Chomsky compartilha do pressuposto aristotélico que a linguagem é usada como ferramenta para dizer o que pensamos. O aparato combinatório e interpretativo da gramática consiste em um conjunto de princípios e regras, inconscientes, mas que podem ser analisados precisamente porque, de algum modo, eles são análogos a operações matemáticas. Os estudos de indexação publicados pelos *Journal of Documentation* e *Knowledge Organization* apresentam em suas abordagens a influência dos paradigmas Realista e Mentalista na compreensão da linguagem que impacta diretamente na representação temática dos documentos.

**PALAVRAS-CHAVE** *Indexação; Realismo; Mentalismo; Journal of Documentation; Knowledge Organization*

**ABSTRACT** This paper is the result of an analysis on the influence of linguistic theories, especially the Semantic field, on indexing approaches. The intention is to show how the Realism and the Mentalism as paradigms influence the understanding about language in Information Science. In Realism, language expresses the real world. Meaning is what describes reality. It is possible to identify universal concepts, laws, logics because the things of the world are created from the same essence. So, the vocation of language is to declare, to inform about them. On the other hand, for The Mentalism of Aristotelian inspiration, the universal is the mind. The rational faculty subordinates language to thought. - both paradigms have great influence on classic Linguistic theories like Frege's -Formal Semantics and Chomsky's Generative Grammar. The Representational Logic in Frege's Semantics prioritizes formal relations, that is, the objective aspects of meaning. Chomsky's Generative Grammar theory shares the Aristotelian assumption that language is used as a tool to say what we think. The combinatorial and interpretative apparatus of a grammar consists of an unconscious set of principles and rules, that can be precisely analyzed - because, -, they are, somehow,

analogous to mathematical operations. The approaches of the indexing studies published by the *Journal of Documentation* and by the *Knowledge Organization* journal demonstrate the influence of the Realism and the Mentalism paradigms on the understanding of the language, which impacts directly in the subject analysis of records.

KEYWORDS *Indexing; Realism; Mentalism; Journal of Documentation; Knowledge Organization*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

## INTRODUÇÃO

Este artigo é produto de um estudo analítico sobre a influência das teorias linguísticas, em especial no campo da Semântica, nas abordagens sobre indexação. A intenção é mostrar como o Realismo e o Mentalismo influenciam os autores da área da indexação a pensar este processo.

De acordo com Silva e Fujita (2004, p. 136), o conceito de indexação surgiu a partir da elaboração de índices num período em que não estava tão demarcada a separação entre representação descritiva e representação temática. Para Lancaster (2003, p. 1), os processos de catalogação descritiva identificam autores, fontes, títulos e outros elementos bibliográficos. Já a indexação é um processo que identifica o assunto de que trata um documento, ou seja, representa o seu conteúdo.

A indexação consiste em uma assembleia de referências para itens de informação. Os itens podem ser o assunto tratado por todos os documentos ou fatos gravados nos documentos, mas dessa massa de referências o pesquisador deve ser capaz de selecionar dessas referências um assunto específico (Vickery, 1950, p. 140, tradução nossa)

A necessidade de estabelecer o assunto (*subject*) ou a tematicidade (*aboutness*) de um documento está diretamente relacionada à questão da linguagem: qual é o seu sentido? Instintivamente, respondemos esta questão a partir de nossa concepção filosófica, mesmo sem perceber quais os princípios que governam a nossa decisão. Percebemos que nas abordagens sobre indexação há dois paradigmas da filosofia da linguagem frequentes: o **realismo**, aquele que identifica parcelas da realidade, e o **mentalismo**, aquele que partilha eventos mentais entre falantes e ouvintes.

No Realismo, a linguagem expressa o mundo real. O sentido é aquilo que descreve a realidade e as suas coisas. Essa diáde é associada ao pensamento de Platão sobre as palavras e sua função. Em *Crátilo*, Platão apresenta sua compreensão sobre como o mundo é formado. Para ele, há um mundo ideal onde a essência de todas as coisas reside. Esta essência transcende à experiência humana. É possível identificar conceitos, leis, lógicas universais porque as coisas do mundo são criadas a partir da mesma substância e o que muda é a maneira em que a essência é configurada no mundo sensível onde vivemos. A linguagem funciona como ferramenta para descrever as coisas do mundo.

O Mentalismo, de inspiração aristotélica, partilha do pensamento platônico sobre a essência universal das coisas. Contudo, neste paradigma, a linguagem não diz o real, mas diz o que pensamos sobre o real. Para Aristóteles, não há divisão entre o mundo ideal e o mundo sensível. É possível tocar o universal sobre as coisas porque as almas humanas vêm da mesma substância. O universal é a mente. Então, a função da linguagem é dizer o que a mente pensa. A faculdade racional ocupa o lugar central no desenvolvimento do pensamento e a linguagem um lugar secundário, subordinado à mente. Estes dois

paradigmas - Realismo e Mentalismo - têm grande influência nas clássicas teorias linguísticas como a Semântica Formal de Frege e a Gramática Gerativa de Chomsky.

Frege, filósofo matemático, herdeiro da tradição aristotélica, procurou resolver problemas da Lógica criando um sistema de representação formal. Seu trabalho objetiva o controle do significado nas equações matemáticas. A partir desta perspectiva, ele escreve *Sentido e Referência* (1976), um livro sobre a necessidade de controlar os sentidos na linguagem, removendo suas imperfeições. A Lógica Representacional da Semântica de Frege prioriza as relações formais, ou seja, os aspectos objetivos da linguagem. Os sentidos são modos de apresentação das referências, a existência de objetos referenciais independem de seu observador. Em Frege (1976), as relações entre sentenças e seus elementos constitutivos poderiam ser melhor compreendidas por uma linguagem lógica. A notação valida sentenças a partir de critérios que estabelecem o que é verdadeiro ou falso, eliminando as dúvidas e os mal entendidos da linguagem natural. "É a busca pela verdade, qualquer que ela seja, que dirige o sentido à referência" (Frege, 1976, p. 69). O sentido é objetivo porque "a humanidade possui um tesouro comum de pensamentos" (Frege, 1976, p. 65). Nesta perspectiva, é importante refrear a "euforia da linguagem" para que os sistemas de representação possam ser eficientemente comunicativos. A padronização sempre opera a partir de relações binomiais entre palavras, conceitos, termos, sentenças e expressões como: certo e errado, verdadeiro e falso, válido e inválido.

Assim como na teoria da Semântica Formal de Frege, a teoria da Gramática Gerativa de Chomsky compartilha do pressuposto aristotélico que a linguagem é usada como ferramenta para dizer o que pensamos. Segundo Chomsky (1968), o aparato combinatório e interpretativo da gramática consiste em um conjunto de princípios e regras, inconscientes, mas que podem ser analisados precisamente porque, de algum modo, eles são análogos a operações matemáticas. De acordo com a teoria da Gramática Gerativa, há uma gramática comum a todas as línguas, a mente. A Gramática Gerativa fornece um esquema geral: regras e princípios universais governam o comportamento da linguagem e ajuda a gerá-la. "A diferença das línguas está no léxico" (Chomsky, 1968, p.46). Esta Gramática Universal, segundo Chomsky, é um componente biológico humano. Se "a língua é a manifestação mais ampla da nossa vida mental" (Chierchia, 2003, p. 24), é necessário compreender os padrões mentais e reproduzi-los na linguagem. É possível fazer a correspondência entre palavras de qualquer língua porque possuímos um aparelho cognitivo parecido, a formulação dos pensamentos é semelhante entre os indivíduos. Chomsky afirma que as estruturas linguísticas variam pouco de um país para outro. A maior diversidade está nas possibilidades de combinação entre as estruturas. Nesta concepção, nós possuímos um "algoritmo inconsciente" capaz de combinar estruturas finitas da linguagem, criando sequências infinitas de palavras, sentenças, expressões e textos. A maneira com que amalgamamos os significados das palavras depende exclusivamente da maneira com que as combinamos sintaticamente.

Os estudos sobre indexação foram influenciados por esta forma de pensar a linguagem: palavras declaram conceitos universais sobre o mundo. Os elementos, características e atributos transformarão os conceitos em bem sucedidos se suas combinações chegarem próximo à descrição das coisas e do pensamento. Há uma objetividade essencialista no mundo captada pela mente e representada pela linguagem. Nesta visão instrumentalista, a principal função da linguagem é informar sobre as coisas.

## METODOLOGIA

Desse modo, procuramos mapear os estudos de indexação em dois importantes periódicos, reconhecidos internacionalmente no campo da Ciência da Informação - o *Journal of Documentation* e o *Knowledge Organization*.

O *Journal of Documentation* é o periódico mais antigo na área da Ciência da Informação, publicado desde 1945, consolidou-se por sua frequência em divulgar teorias, filosofias, conceitos, modelos sobre documentação, organização do conhecimento, indexação, classificação e outros estudos no mesmo campo. Este periódico possui 73 volumes e 329 números. Já o *Knowledge Organization*, periódico oficial da International Society for Knowledge Organization, publica pesquisas sobre teorias, fundamentos filosóficos do conhecimento, operações práticas associadas à indexação e à classificação, além de problemas de terminologia. Fundado em 1976, possui 44 volumes e mais de 161 números.

Assim, delimitamos como escopo para a análise dessa pesquisa às publicações referentes aos últimos 10 anos nos dois periódicos, entre 2006 e 2016. No *Journal of Documentation* (JD) analisamos todos os números a partir do volume 62, totalizando 63 números, nos quais encontramos em seus títulos e resumos 323 artigos abordando a temática da indexação (*indexing*), 22 sobre análise temática (*subject analysis*) e 63 sobre estudos de linguagem (*language*). Números consideráveis que apontam a preocupação do campo com questões referentes à indexação.

Já os artigos do *Knowledge Organization* (KO) analisados possuem diferentes características relacionadas às teorias linguísticas. O escopo desta análise também utilizou como marco os últimos dez anos de publicações (2006-2016), ou seja, todos os artigos a partir do volume 33. Portanto, analisamos 53 publicações, dentre as quais encontramos entre títulos e resumos 53 artigos sobre indexação (*indexing*), 5 artigos sobre análise temática (*subject analysis*) e 2 artigos sobre estudos de linguagem (*language*). Entendemos que a grande diferença nos números de artigos referentes a indexação e estudos de linguagem, se compararmos aos números do *Journal of Documentation*, está associada à própria questão filosófica que, de certa forma, serve de diretriz ao KO.

## ANÁLISE DOS RESULTADOS

Tradicionalmente, o *Journal of Documentation* publicou importantes estudos sobre indexação como os de Vickery, Farradane, Coates, Langridge, Faithorne e Foskett entre os anos de 1950 e 1970. O Classification Research Group, formado na Inglaterra em 1952 por A. J. Wells e B. C. Vickery a partir das discussões acerca da classificação levantadas na *Royal Society Scientific Information Conference* de 1948, teve seus primeiros boletins publicados neste periódico. A proposta inicial do CRG era reunir bibliotecários interessados em discutir os sistemas de classificação de suas instituições, sobretudo a aplicabilidade de princípios classificatórios. O grupo sentia a necessidade de compartilhar suas experiências na organização sistemática de documentos por assunto, divulgando práticas e aplicações teóricas bem sucedidas. Os debates registrados em seus boletins foram importantes referenciais para os estudos de indexação. Essas teorias sobre indexação possuem em comum o fato de pensar formas de sistematização da linguagem na entrada dos documentos em um sistema de recuperação da informação. Neste período, os computadores estavam sendo inseridos nos sistemas de recuperação da informação (SRIs). Portanto, os estudos da indexação visavam buscar maneiras de padronização da linguagem, realizando a sua tradução para uma linguagem artificial que fosse compreensível para os serviços de

informação automatizados, facilitando assim a recuperação dos documentos de tais bases de dados. Essa busca aproximou os estudos acerca da indexação dos estudos de linguagem.

A construção de linguagens artificiais, controladas e padronizadas possui forte influência da lógica encontrada na semântica formal de inspiração fregeana. A premissa semelhante à lógica quantificadora de predicados e sentidos de Frege é a do controle da linguagem pela indexação a fim de realizar a correspondência entre assunto requerido e conteúdo do documento. Assim, os primeiros estudos publicados no *Journal of Documentation* associando a indexação aos serviços de informação como abordados por Vickery (1953), Saracevic & Rees (1967) e Foskett (1975) vão se aproximar dos estudos linguísticos para realizar suas proposições. Esses estudos possuem em comum a premissa de entender a indexação como uma operação que deve reconhecer os assuntos (*subjects*) mais significantes de um documento para auxiliar a recuperação da informação. Mas como precisar o que é mais significativo? Tais teorias apontam para o uso dos princípios da terminologia que procuram controlar as significações das palavras (Cabré, 1993). A proposta é estabelecer uma tradução da linguagem natural a um sistema lógico tentando criar um vocabulário em que cada palavra tenha apenas um significado, esses estudos procuram estabelecer leis para controlar as variações na indexação, expressando a relação entre o tamanho do vocabulário e a especificidade com que um conceito pode ser descrito (Gardin, 1973, p. 143).

A linguagem representacional deste modelo matemático-terminológico de indexação aproxima-se com a semântica formal proposta por Frege (1976), marcada pela influência do paradigma realista da linguagem. Tanto nessa teoria da indexação, quanto na semântica formal de Frege (1976), a linguagem é compreendida como representação da realidade. O sentido do documento estaria no conteúdo expresso no texto.

A influência deste pensamento pode ser apreendida nos artigos produzidos nos últimos anos acerca de indexação automática. Nos 36 artigos encontrados sobre indexação automática como o trabalho de Tudhope *et al* (2006), *Query expansion via conceptual distance in thesaurus indexed collections*, o estudo de Zhang e Zeng (2014), *A new similarity measure for subject hierarchical structures; Classifications and concepts: towards an elementary theory of knowledge interaction*, de Smiraglia e Heuvel (2013); *Untangling search task complexity and difficulty in the context of interactive information retrieval studies*, de Wildemuth, Freund, Toms (2014) são alguns exemplos da influência do paradigma realista na abordagem acerca da indexação e da recuperação da informação. Esses artigos apresentam uma concepção de linguagem semelhante à semântica formal. Tais estudos apresentam um princípio em comum que é procurar estabelecer possibilidades - seja na entrada dos documentos, seja na sua saída dos sistemas - para que o computador realize a melhor conexão entre os documentos armazenados nas bases de dados e a pergunta do pesquisador. Os estudos enfatizam o controle terminológico como uma condição primordial na recuperação da informação. Outrossim, a construção de modelizações, leis, padronizações, para o controle de termos, conceitos e palavras pela indexação automática filia-se à semântica formal ao procurar *traduzir* a linguagem para uma representação lógica, tentando evitar o que Frege chamava de *imperfeições* da linguagem. Problemas também identificados na indexação como a sinonímia, a polissemia e os sentidos fora do contexto eram questões que a semântica formal procurava eliminar a partir do controle da linguagem pela notação lógica e o cálculo de predicados.

Já os estudos mais recentes sobre indexação automática que propõem a utilização de métodos a partir da padronização por sintagmas como o artigo de Ibekwe-Sanjuan (2006), *Constructing and maintaining*

*knowledge organization tools*, de Rafferty (2011), *Knowledge Representation in the Social Semantic Web* são marcados pela influência da gramática gerativista chomskyana. Segundo Chomsky, a linguagem é uma faculdade mental, um equipamento inato e universal, ilimitado e criativo (é um repertório de unidades e princípios de combinação que oferecem ilimitadas possibilidades de expressão. Se "a linguagem é a manifestação mais ampla de nossa vida mental" (Chierchia, 2003, p. 24), é necessário que compreendamos as estruturas gramaticais procurando padronizá-las, pois a forma de pensar é semelhante.

Sendo assim, a tradução seria a simples operação de substituir a palavra em determinada língua por a sua correspondente, o mesmo se daria para as sentenças e até mesmo para os textos. O sentido não se altera porque ele é estabelecido pela mente e não pela linguagem. De acordo com Chierchia (2003, p. 38), possuímos um "algoritmo inconsciente" capaz de combinar as regras e padrões, conjunto finito e matricial, de maneiras infinitas que se desdobram em variadas sentenças, textos e línguas.

A Gramática Gerativa tem sido amplamente utilizada como suporte teórico dos sistemas de indexação automática. Alguns métodos como a indexação por extração automática sintática, a indexação por atribuição semântica, a identificação de palavras *in full context* são exemplos de operações que buscam estabelecer padrões de conceitos, de relações hierárquicas entre as palavras, termos e sintagmas, prevendo inúmeras possibilidades de combinação entre elas.

Esta abordagem da indexação a partir da gramática gerativista e, por conseguinte, do paradigma mentalista, facilitou a interface entre o indexador e o computador, possibilitando o aparecimento de diversas bases de dados com graus elevados de precisão e revocação na recuperação dos documentos.

O *Knowledge Organization* fundamenta-se no conceito de organização e representação do conhecimento, diferentemente do *Journal of Documentation* que privilegia o tratamento temático da informação e o comportamento de uso. Dessa forma, operações como a indexação estão inseridas nas discussões, teorias e problematizações da própria área da Organização do Conhecimento.

Dahlberg, ao analisar a Organização do Conhecimento enquanto campo científico em 1993, quando da mudança de nome do periódico de *International Classification* para *Knowledge Organization*, estabelece que a organização do conhecimento é um domínio da biblioteconomia e da ciência da informação comprometido em organizar, ordenar e dispor a informação (Dahlberg, 1993, p. 211). Este pensamento está em consonância com a sua teoria do conceito que estabelece as formas em que os indivíduos organizam seu modo de pensar. De acordo com a autora, os conceitos "identificam as características, os atributos das coisas designadas" (Dahlberg, 1978, p. 102). Segundo esta teoria, o processo de análise dos conceitos somente é possível "coletando-se os enunciados verdadeiros que sobre determinado objeto se podem formular" (Dahlberg, 1978, p. 102). Assim como nos pressupostos platônicos, Dahlberg indica que nos conceitos há elementos que constituem enunciados verdadeiros sobre as coisas, passíveis de serem identificados e classificados de acordo com suas características naturais. Nesta concepção, existe uma autonomia simbólica dos enunciados que compõem os textos. Os elementos, as características e os atributos vão compor os conceitos, a tessitura bem sucedida dessas combinações é que vão indicar se a representação do real será bem ou mal sucedida. A eficácia dos conceitos está relacionada diretamente à fidelidade de sua descrição do real. Encontramos, portanto, a influência do pensamento realista de tradição platônica na definição de *conceito* estabelecida por Dahlberg.



Percebemos que a tendência majoritária dos artigos publicados nos últimos dez anos no *Knowledge Organization* é de uma abordagem semelhante acerca do processo de indexação e organização do conhecimento. *Semantic relations in knowledge organization systems*, de Brascher (2014), faz alusão ao trabalho de Dahlberg sobre a teoria do conceito na construção da categoria *organização do conhecimento*. O artigo enfatiza a necessidade do mapeamento das estruturas e funções semânticas e sintáticas dos KOS a fim de facilitar a identificação e, por conseguinte, a representação dos conceitos. Esta operação visa a padronização da linguagem artificial, facilitando a recuperação da informação. Os são compreendidos como unidades universais e representativas do pensamento do indivíduo, tal como na gramática gerativista chomskiana. Nesta perspectiva, a linguagem representa sistemas de conceitos elaborados e organizados pelo pensamento, produto da cognição característica comum a todos os indivíduos.

Portanto, a representação do conhecimento é construída a partir de conceitos articulados e relacionados entre si, pois ela é produto da mente. A língua apenas expressa o conteúdo mental cuja estrutura sintática e semântica reproduzem o modo de pensar dos indivíduos em seus contextos sociais, culturais, geográficos e históricos. De maneira semelhante ao gerativismo chomskyano, percebemos nesta abordagem a influência dos paradigmas realista e mentalista, por tratar os conceitos como entidades abstratas que designam as coisas no mundo e por considerar a linguagem como expressão do pensamento influenciado por seu meio. Encontramos uma abordagem semelhante acerca da indexação no artigos *Using a semantic analysis tool to generate subject access points*, de Zeng *et al* (2014); *Method for selecting specialized terms from a general language*, de Anguiano Peña e Naumis Peña (2015); *A survey of the coverage and methodologics of schemes and vocabularies use to describe information resources*, de Hider (2015); além do artigo *Terminology web services*, de Binding e Tudhope (2010).

As publicações analisadas vão procurar tratar de questões voltadas à organização e representação do conhecimento a partir da compreensão de que existem conceitos universais que expressam o pensamento tal como vimos em Dahlberg (1978). Portanto, os artigos vão procurar refletir sobre formas de classificar, enquadrar, estabelecer hierarquias e nivelamentos do conhecimento. Há uma predominância dos estudos sobre ontologias - encontramos 56 artigos abordando esta temática nos últimos 10 anos dentre eles *Formal Ontology and the Foundation of Knowledge Organization*, de Herre (2013), *Organizing Acts and Objects*, de Budd (2014) e *Boundary Objects*, de Marchese e Smiraglia (2013). Ademais, as linguagens documentárias e taxonomias também destacam-se como temas de muitos artigos no KO. São 30 artigos abordando a questão da taxonomia na organização do conhecimento como *Towards a Taxonomy of KOS*, de Souza, Tudhope e Almeida (2012), *Concepts and Terms in the Faceted Classification*, de Broughton (2010), *Naming and Reframing*, de Gross (2015). Todos os artigos citados possuem em comum a filosofia de que o conhecimento é uma entidade autônoma com estruturas passíveis de serem classificadas, codificadas e organizadas de maneira a promover o seu acesso. Esses métodos possuem uma grande proximidade do paradigma mentalista. Os sistemas de organização do conhecimento (KOS), assim como a mente humana, possuem padrões limitados, inclusive de estruturas linguísticas tal como os sintagmas, no entanto podem ser utilizados de infinitas maneiras, de acordo com as combinações feitas de forma criativa. Os algoritmos tão utilizados na Matemática e na Ciência da Computação, e aplicados à recuperação da informação, são exemplos dessas possibilidades. A sequência finita de regras, instruções bem definidas e não ambíguas, cada uma das quais devendo ser executadas mecânica ou eletronicamente em um intervalo de tempo finito e com uma quantidade de esforço finita que podem gerar infinitas combinações. Assim como a mente é capaz de realizar conexões entre os pensamentos, os computadores por sua programações são

capazes de realizar, em muitos casos, a recuperação de documentos úteis ao usuário, eliminando o ruído e o silêncio pelo controle e padronização da linguagem natural.

## CONSIDERAÇÕES FINAIS

Aproximamo-nos de nosso ponto de chegada, que talvez também seja outro ponto de partida, no sentido de suscitar questionamentos que não se encerram aqui. Nosso percurso procurou refletir sobre como o realismo e o mentalismo, enquanto filosofias da linguagem, que influenciam os estudos acerca da indexação. Escolhemos como campo empírico para tal análise os periódicos *Journal of Documentation* e o *Knowledge Organization*, mediante exame dos títulos e resumos dos artigos, que são fragmentos dos textos na íntegra, é sabemos dessa delimitação na observação dos artigos na pesquisa. Com o objetivo de refinar e apurar a pesquisa, procuramos delimitar nossa investigação usando como marco temporal as publicações dos últimos dez anos de cada jornal, analisando os artigos que problematizaram o processo da indexação. Dessa forma, procuramos compreender como complexas teorias linguísticas tal como o gerativismo chomskyano e a semântica formal de inspiração fregeana encontram-se presentes nas abordagens dos artigos analisados, sobretudo aqueles voltados à indexação e à organização do conhecimento. Percebemos o quanto a tradição da filosofia socrática influencia até hoje a nossa forma de compreender a linguagem. A concepção da linguagem como um instrumento comunicacional, de base universal, cuja função básica é designar os nomes das coisas no mundo é algo presente na forma em que abordamos a indexação dos documentos em nossa contemporaneidade.

Compreendemos que o pensamento aristotélico da língua como sistema de representação da vida mental influenciou inclusive a construção de diversos sistemas de recuperação da informação.

Destarte, desfecha-se este artigo, ressaltando-se que este foi apenas um esforço experimental, resultante de uma pesquisa em andamento para a conclusão de uma tese de doutoramento, que procurou mapear como algumas filosofias da linguagem influenciam as abordagens e estudos sobre indexação e organização do conhecimento. Nossa intenção foi apresentar algumas interações entre elas, de forma panorâmica, procurando não tratá-las como blocos teóricos indivisíveis.

## REFERÊNCIAS BIBLIOGRÁFICAS

Anguiano Peña, G.; Naumis Peña, C. (2015). Method for selecting specialized terms from a general language corpus. *Knowledge Organization*, 42(3), 164-175.

Aristotle. (1990). *Metaphysics*. Madrid: Editorial Gredos.

Binding, C.; Tudhope, D. (2010). Terminology web services. *Knowledge Organization*, 37(4), 287-298.

Borko, H. (1977). Toward a theory of indexing. *Information Processing & Management*, 13 May: 355-365.

Brascher, M. (2014). Semantic relations in knowledge organization systems. *Knowledge Organization*, 41(2), 175-180.

Broughton, V. (2010). Concepts and terms in the faceted classification: the case of UDC. *Knowledge Organization*, 37(4), 270-279.

Budd, J. Organizing acts and objects: metaphysical foundations. *Knowledge Organization*. 41(6), 419-428.

Cabré, M. T. (1993). *La terminología: teoría, metodología y aplicaciones*. Barcelona: Ed. Antárdida/Empúries.

Chierchia, G. (2003). *Semântica*. Campinas: Ed. Unicamp.

Chomsky, N. (1968). *Language and mind*. New York: Harcourt, Brace & World, Inc.

Dahlberg (1993). Knowledge organization: its scope and possibilities. *Knowledge Organization*, 20( 4), 211-222.

\_\_\_\_\_ (1978). Teoria do conceito. *Ciência da Informação*, 7(2), 101-107.

Farrow, J. (1995). All in the mind. *The indexer*, 19(4) Octutober: 243-247.

Foskett, D. J. (1975). *Classification and indexing in the social sciences*. London: LexisNexis Butterworths Co.

Frege, G. (1976). *Sense and Reference*. São Paulo: Cultrix/USP.

Gardin, J. C. (1973). Document analysis and linguistic theory. *Journal of Documentation*, 29(2), 137-168.

Plato (2008). *Republic*. Project Gutenberg.

\_\_\_\_\_ (2011). *Cratylus*. Cambridge: Cambridge University Press.

Farrow, J.F (1991). A cognitive process model of document indexing, *Journal of documentation*, 47(2), 149-166.

Foskett, D.J. (1970). Classification and indexing in the social sciences. *Journal of Documentation*, 22(3), 90-101.

Gross, T. (2015). Naming and reframing: a taxonomy of attacks on knowledge organization. *Journal of Documentation*, 42(5), 263-268.

Harviainen, J. T.; Hamari, J. (2015). Seek, share, or withhold: information trading in MMORPGs, *Journal of Documentation*, 71(6), 1119-1134.

Hemmig, W. S. (2008). The information-seeking behavior of visual artists: a literature review. *Journal of Documentation*, 64(3), 343-362.

Herre, H. (2013). Formal ontology and the foundation of knowledge organization. *Knowledge Organization*, 40(5), 332-339.

Hider, P. (2015). A Survey of the coverage and methodologies of schemas and vocabularies used to describe information resources. *Knowledge Organization*, 42(3), 154-163.

- Hjorland, Birger (2017). Subject (of documents). *Knowledge Organization*, 44(1), 55-64.
- Howarth, L. C. (2010). Clare Beghtol: exploring new approaches to the organization of knowledge. *Knowledge Organization*, 37(2), 95-100.
- Ibekwe-Sanjuan, F (2006). Constructing and maintaining knowledge organization tools: a symbolic approach. *Journal of Documentation*, 62(2).
- Kwaśnik, B. H. (2010). Semantic warrant: a pivotal concept for our field. *Knowledge Organization*, 37(2), 106-110.
- Marchese, C.; Smiraglia, R. P. (2013). Boundary objects: CWA, and HR firm, and emergent vocabulary. *Knowledge Organization*, 40(4), 254-259.
- Martínez-Ávila, D.; San Segundo, R. (2013). Reader-interest classification: concept and terminology historical overview. *Knowledge Organization*. 40(2), 102-114.
- Rafferty, P. (2011). Knowledge representation in the social semantic web. *Journal of Documentation*, 67(5), 896-899.
- Robson, A.; Robinson, L (2015). The information seeking and communication model: a study of its practical application in healthcare. *Journal of Documentation*, 71(5), 1043-1069.
- Sabbar, C.; Xie, I (2016). Language in the information-seeking context: A study of US scholars using non-English sources. *Journal of Documentation*, 72(1). 103-126.
- Satja, M. P.; Madalli, D. M.; Dutta, B. (2014). Modes of growth of subjects. *Knowledge Organization*, 41(3), 195-204.
- Saracevic, T.; Rees, A. M. (1967). Towards the identification and control of variables in information retrieval experimentation. *Journal of Documentation*, 23(1), 7-19.
- Smiraglia, R. P.; Heuvel, C. Classifications and concepts: towards an elementary theory of knowledge interaction. *Journal of Documentation*, 69(3), 360-383.
- Souza, R. R.; Tudhope, D.; Almeida, M. B. (2012). Towards a Taxonomy of KOS: Dimensions for Classifying Knowledge Organization Systems. *Knowledge Organization*, 39(3), 179-192.
- Tudhope, D.; Binding, C.; Blocks, D.; Cunliffe, D (2006). Query expansion via conceptual distance in thesaurus indexed collections, *Journal of Documentation*, 62(4), 509-533.
- Vállez, M.; Pedraza-Jiménez, R.; Codina, L.; Blanco, S.; Rovira, C (2015). Updating controlled vocabularies by analysing query logs. *Journal of Documentation*, 39(7), 870-884.
- Vickery, B. C. (1953). Systematic subject indexing. *Journal of Documentation*, 9(1), 48-57.
- Wildemuth, B.; Freund, L.; Toms, E. G (2014). Untangling search task complexity and difficulty in the context of interactive information retrieval studies. *Journal of Documentation*, 70(6), 1118-1140.
- Zeng, M. L., Gracy, K. F., Žumer, M. (2014). Using a semantic analysis tool to generate subject access points: a study using Panofsky's theory and two research samples. *Knowledge Organization*. 41(6), 440-451.

Zhang, J.; Zeng, M. L (2014). A new similarity measure for subject hierarchical structures. *Journal of Documentation*, 70(3), 364-391.