

Sistemas de recuperação da informação: uma discussão a partir de parâmetros enunciativos¹

Information retrieval systems: a discussion based on enunciative parameters

Vera Maria Araujo Pigozzi de ARAUJO²

Resumo

A temática deste artigo são os Sistemas de Recuperação da Informação na interface com a teoria da enunciação. O objetivo é discutir a relação que se estabelece entre os sistemas de informação e os usuários. Com o propósito de identificar como se dá essa relação, examinam-se os modelos de comunicação documentária vistos em revisão de literatura, e identificam-se os parâmetros sob os quais são construídos. Por questionar a validade desses modelos, analisam-se as características dessa relação. Identifica-se, no aparelho formal da enunciação formulado por Benveniste, a possibilidade de se contemplarem as peculiaridades que a caracterizam, considerando os seguintes parâmetros: categorias de pessoa (eu-tu), espaço (aqui) e tempo (agora). Conclui com uma proposta alternativa aos modelos apresentados, considerando bibliotecários, usuários e sistemas como interlocutores, e sugerindo ser uma relação de interação e não de comunicação. Esta proposta é fundamentada nas noções de subjetividade, intersubjetividade, categorias de pessoa, de espaço e de tempo.

Palavras-chave: Aparelho formal da enunciação. Comunicação documentária. Sistemas de recuperação da informação. Teoria da enunciação.

Abstract

The theme of this article is the interface of Information Retrieval Systems with the theory of enunciation. The aim of this study was to discuss the relationship established between information systems and users. In order to identify how this relationship works, it examines the documentary communication models seen in a literature review, and identifies the parameters they are built on. By questioning the validity of these models, it analyzes the characteristics of this relationship. In the formal apparatus of enunciation proposed by Benveniste, it identifies an opportunity to contemplate the peculiarities that characterize the relationship, considering the following parameters: the categories of person, space and time. It concludes with an alternative proposal to the models presented, considering librarians, users and systems as interlocutors, and suggests that it is a relationship of interaction and not one of communication. This proposal is based on notions of subjectivity, intersubjectivity, categories of person, space and time.

Keywords: Formal apparatus of enunciation. Documentary communication. Information retrieval systems. Theory of enunciation.

Introdução

Este artigo trata dos Sistemas de Recuperação da Informação (SRI), e examina-se nele a comunicação documentária, a partir dos referenciais de Araujo (1995), Lima (1998), Lara (2002), Ortega e Lara (2010) e

Biscalchin e Boccato (2011). Por questionar a validade desses modelos, que identificam sistema e usuário como emissor e receptor desse processo de comunicação, examinam-se os SRI, apresentando sucinta-

¹ Artigo elaborado a partir da pesquisa de doutorado em curso, desenvolvido por V.M.A.P. ARAUJO.

² Doutoranda, Universidade Federal do Rio Grande do Sul, Programa de Pós-Graduação em Letras. Av. Bento Gonçalves, 9500, 90040-060, Porto Alegre, RS, Brasil. E-mail: <verapigozzi@hotmail.com>.

Recebido em 29/10/2011, reapresentado em 21/5/2012 e aceito para publicação em 5/6/2012.

mente sua conceituação e função sob a visão de Cendón (2005) e Souza (2006).

No exame da relação dos SRI com os usuários, identificam-se diversas peculiaridades que não estão contempladas nos modelos apresentados. Na busca de um referencial mais satisfatório, percebe-se a possibilidade de desenvolver uma proposta alternativa a esses modelos. É no aparelho formal da enunciação que essa possibilidade é visualizada.

Para discutir a relação que se estabelece entre os SRI e os usuários, o artigo apresenta uma breve revisão de literatura sobre comunicação documentária e sistemas de recuperação da informação. A partir da análise dos modelos de comunicação e dos sistemas na sua relação com os usuários, oferece-se, ainda, uma proposta fundamentada no aparelho formal da enunciação formulado por Benveniste.

Comunicação documentária

Comunicação documentária, conforme Lima (1998), é o processo de comunicação entre o acervo de um sistema de informação e o seu usuário. Ela é alcançada quando um sistema codifica o conteúdo informacional dos documentos, objetivando sua recuperação pelo usuário. A linguagem documentária é o código utilizado nesse processo.

Os modelos de comunicação documentária têm origem na teoria da informação proposta por Shannon e Weaver (1949), que entende que o processo de comunicação se dá a partir de um sistema (fonte), que influencia outro sistema (receptor), transmitindo signos por intermédio de um canal. O canal é o meio que permite a

comunicação e que viabiliza a transferência da informação. Essa teoria, ainda que tenha recebido alterações no decorrer do tempo, é a que usualmente fundamenta o processo de comunicação documentária.

Lara (2002) entende que as representações documentárias são utilizadas para transmitir informação. No processo de transferência da informação, é possível identificar: 1) o emissor (instituição que formula o sistema documentário); 2) o canal (veículo pelo qual são veiculadas as mensagens documentárias e que seriam os catálogos eletrônicos³); 3) o código intermediário (a linguagem documentária usada na indexação e na recuperação da informação e que deve funcionar, simultaneamente, como um sistema de comunicação e de significação); 4) o destinatário (usuário final do sistema). Pelo modelo de comunicação documentária assim proposto, a relação entre SRI e usuários é caracterizada de forma bem simplificada (Quadro 1).

Conforme Ortega e Lara (2010, p.9):

O usuário é o ator que dispara uma situação de comunicação, pois é o ato da recepção que determina o estabelecimento ou não de uma comunicação. [...] para que a comunicação documentária ocorra é necessário que a informação documentária - produto que é transmitido nesse processo - seja portadora de um significado tal que o usuário possa ser capaz de processá-la mentalmente, a fim de obter a informação desejada. A comunicação documentária é então um processo que exige, de um lado, a informação documentária e, de outro, o usuário que se apropria desta informação. Os sistemas documentários lidam com atos de comunicação materialmente explícitos, isto é, o documento, sua representação, a pergunta do usuário, a ação comunicativa do profissional de informação etc.

Quadro 1. Modelo de comunicação documentária.

Elementos da comunicação na teoria da informação	Representação dos elementos da comunicação na comunicação documentária
Emissor	Instituição que formula o sistema documentário
Canal	Catálogo eletrônico
Código intermediário	Linguagem documentária
Destinatário	Usuário final do Sistema

Fonte: Elaborado pela autora.

³ Os catálogos eletrônicos são reconhecidos como SRI por integrarem diferentes atividades: seleção, desenvolvimento de coleções, processamento técnico, busca e recuperação da informação.

Ainda, para que esse processo de comunicação ocorra, é necessário que: 1) a informação documentária - produto que é transmitido nesse processo - seja portadora de um significado, e 2) o usuário - ator que dispara uma situação de comunicação - seja capaz de processar a informação documentária para obter a informação desejada.

Tendo como foco a insatisfação do usuário na relação com os SRI, aborda-se, a seguir, a relação que ocorre entre ambos. Dentre os autores que questionam e procuram respostas para essa insatisfação presente nos processos de busca e recuperação da informação, Araujo (1995, p.1), cuja tese discute essa questão, faz a seguinte afirmação:

Por que vêm falhando os sistemas de informação? E não há dúvidas de que vêm falhando: uma rápida análise da literatura da área evidencia uma série de estudos, projetos etc., visando corrigir, ao menos parcialmente, suas falhas e proporcionar 'maior satisfação aos usuários'. Observa-se, inclusive, que tais estudos incluem 'modernas abordagens' ao planejamento, gestão e, principalmente, *marketing* de seus serviços e produtos, buscando a inserção dos sistemas de informação em uma sociedade pós-industrial, pós-moderna, conhecimento-intensiva.

Em resposta à pergunta *por que vêm falhando os sistemas de informação* e à consequente insatisfação do usuário, Biscalchin e Boccato (2011), em um estudo sobre o uso de linguagem documentária em catálogos coletivos de bibliotecas universitárias, apontaram três aspectos principais: 1) a incompatibilidade existente entre a linguagem de busca e a adotada pelo sistema; 2) a ausência de remissivas; e 3) as inconsistências sintático-semânticas verificadas na linguagem. A pesquisa revelou que a grande maioria dos usuários desconhece a linguagem usada pelo sistema. Acrescenta-se a isso o desconhecimento dos recursos de pesquisa que o catálogo eletrônico oferece. Lima (1998) observa, ainda, que a presença de várias linguagens documentárias em um único SRI prejudica a transferência da informação e, desse modo, a comunicação entre o sistema e o usuário.

Em busca de uma maior satisfação do usuário, são feitas algumas considerações sobre o fato de o bibliotecário, como profissional da informação, ter sob sua responsabilidade o gerenciamento dos SRI. Ele precisa estar atento tanto ao modo de expressar da maneira mais

fiel possível a representação do conteúdo temático de um documento, como ao conhecimento das estratégias de busca oferecidas pelo sistema.

Neste artigo, o aspecto a ser considerado é o da representação fiel do conteúdo temático do documento. Para Lima (1998), a compatibilização das linguagens documentárias com a terminologia oferece condições propícias a uma comunicação documentária mais eficaz.

Sistemas de recuperação da informação

Os sistemas de informação tratam da representação, do armazenamento, da organização e da localização dos itens de informação. Para organizar e comunicar a informação, eles utilizam linguagens documentárias, que estabelecem um importante elo entre os SRI e os usuários. Conforme Souza (2006), as funções de um SRI são as seguintes: 1) representação das informações contidas nos documentos e expressas pelos processos de indexação e descrição dos documentos; 2) armazenamento e gestão física e ou lógica desses documentos e de suas representações; e 3) recuperação das informações ali contidas e dos próprios documentos armazenados no sistema.

Das vantagens que Cendón (2005, p.62) apresenta a favor dos sistemas no formato eletrônico, há a considerar:

[...] as facilidades que os SRI automatizados oferecem para busca de informação [...] maior número de pontos de acesso que os SRI não-automatizados [...]. Além disso, permitem realizar pesquisas mais complexas, em que vários conceitos necessitam ser relacionados pois pode-se combinar grande número de termos de busca [...], de maneiras que não seriam possíveis nos SRI impressos.

Com a intenção de apresentar uma proposta alternativa que caracterize a relação estabelecida entre SRI e usuários, desenvolve-se a seguir uma reflexão sobre as características apresentadas por esses sistemas. Pela identificação de tais características, verificou-se que os modelos de comunicação até então apresentados não configuram essa relação.

Essa reflexão parte do entendimento de que a função principal de um SRI é dispor de informações contidas nos documentos indexados, a partir de uma descrição sintética, objetiva e representativa de seu con-

teúdo formal e temático. Um SRI integra diferentes processos: seleção, aquisição, indexação e, ainda, busca e recuperação das informações.

Na análise da relação que se estabelece entre os SRI e os usuários, observa-se que ela apresenta peculiaridades que não são explicadas nos modelos de comunicação documentária vistos em revisão de literatura. O que se verifica é que esses sistemas apresentam características bem diferenciadas e muito mais complexas do que aquelas que fundamentam esses modelos, como descrito a seguir:

1) São sistemas em que os locutores (bibliotecário, usuário e SRI) são de natureza heterogênea e não estão presentes quando o alocutário busca a informação;

2) Os códigos utilizados e representados pelas linguagens documentárias são reveladores de diferentes sistemas semânticos; uma mesma Unidade de Informação (cabeçalho de assunto ou descritor) não cobre a mesma noção de um vocabulário para outro, e a gramática das linguagens documentárias é diferente de um sistema para o outro;

3) O processo de indexação faz uso de mais de uma linguagem: linguagem especializada (linguagem do documento) e vocabulário controlado (lista de cabeçalho de assuntos ou tesouro);

4) O processo de busca e recuperação da informação oferece pesquisas por palavras, em catálogos eletrônicos diferenciados quanto à estrutura e estratégias de pesquisa;

5) Os processos que viabilizam a recuperação da informação e a produção do conhecimento se manifestam em lugares e momentos múltiplos;

6) No processo de indexação, bibliotecários e usuários se posicionam como interlocutores, mas no processo de busca e recuperação da informação, usuários e sistemas são os interlocutores;

7) As marcas da subjetividade e intersubjetividade estão presentes e são relevantes nos processos de indexação e de busca e de recuperação da informação;

8) O documento processado é a fonte da informação em ambos os processos.

Visando ao uso da informação e à produção do conhecimento, propõe-se uma aproximação entre os

processos anteriormente mencionados. Entende-se que a relação que ocorre entre SRI e usuários não se limita ao processo de busca e recuperação da informação. A construção de um sistema que se propõe a estabelecer uma relação mais satisfatória com o usuário precisa levar em consideração aquela que ocorre entre o bibliotecário e o usuário durante o processo de indexação. Mas ainda fica a pergunta: que tipo de relação é esta? Trata-se mesmo de uma relação de comunicação? Busca-se esta resposta na teoria de Benveniste.

Teoria da enunciação de Benveniste

Visando compreender o funcionamento das linguagens documentárias, e a relação SRI e usuário, recorre-se à teoria da enunciação. É sob a ótica da semântica, e não da semiótica, que se desenvolve essa reflexão. O semântico toma a seu encargo o conjunto de referentes, e é produzido pelo discurso e se identifica com a enunciação.

Do texto *A forma e o sentido na linguagem* (Benveniste, 1989), apresenta-se a ideia de que a forma é o arranjo formal dos elementos linguísticos ao nível linguístico relevante, enquanto o sentido é o conjunto de procedimentos de comunicação identicamente compreendidos por um grupo de locutores. Para o linguista, a significação é da natureza da linguagem.

Com a noção de semântica, Benveniste entende a língua em emprego e em ação. A frase é a expressão semântica por excelência; com ela nos comunicamos e nos deslocamos para a produção do discurso, para aquilo que o locutor quer dizer. O sentido da frase é sua ideia, que é sempre única, e o das palavras, seu emprego. O sentido da frase se constitui no momento da enunciação, ou seja, quando o locutor exprime sua ideia. Sendo a palavra a unidade semântica, ela mesma guarda suas referências e significações próprias. Entretanto, quando há o agenciamento de palavras para a construção da frase, a referência passa a ser o próprio ato de enunciação; dele decorre o sentido da frase. A noção de referência é essencial no processo de construção do sentido. Sobre as noções de sentido e de referência, o linguista propõe que se estabeleça uma distinção entre elas.

Se o 'sentido' da frase é a ideia que ela exprime, a 'referência' da frase é o estado de coisas que a

provoca, a situação de discurso ou de fato a que ela se reporta e que nós não podemos jamais prever ou fixar [...]. A frase é então cada vez um acontecimento diferente; ela não existe senão no instante em que é proferida e se apaga neste instante [...] (Benveniste, 1989, p.231).

Subjetividade é um conceito fundamental na sua teoria; ela procede da relação 'eu e tu', ou seja, 'pessoa subjetiva e pessoa não-subjetiva'. Mas as noções de intersubjetividade, locutor, referência e categorias de pessoa, espaço e tempo também são importantes. Quanto ao tempo da enunciação, segundo o linguista, ele é sempre presente, uma vez que corresponde ao momento em que 'eu' se diz 'eu'.

Em *A natureza dos pronomes*, Benveniste (1988) aborda a categoria de pessoa. A noção de pessoa está presente apenas nas duas primeiras pessoas, estando ausente na terceira pessoa. O 'ele' é diferente de 'eu e tu' pela sua função e pela sua natureza, e se caracteriza por não remeter a ele mesmo, mas a uma situação objetiva.

No texto *Da subjetividade na linguagem* (Benveniste, 1988), a subjetividade é entendida como a capacidade do locutor para se propor como sujeito. Ao inserir a noção de referência no contexto da comunicação, o linguista questiona a função da linguagem como ferramenta da comunicação. Fundamentado no argumento de que opor o homem à linguagem é colocá-lo em oposição a sua própria natureza, ele nega o aspecto instrumental da linguagem e afirma que a língua não é um instrumento de comunicação, pois isso vai contra o princípio segundo o qual a linguagem está na natureza do homem.

Dentre as teorias revisadas a fim de se formular uma proposta alternativa aos modelos de comunicação documentária, identificados em revisão de literatura, foi na teoria da enunciação de Benveniste, em especial no aparelho formal da enunciação, que foram identificados vários elementos capazes de contemplar as peculiaridades da relação que ocorre entre SRI e usuários.

A opção por essa teoria justifica-se pelas seguintes razões: 1) ela entende que é pela linguagem que o homem se inter-relaciona, recebe e transmite uma mensagem; 2) olha para o estudo da língua do ponto de vista do sentido; 3) permite considerar peculiaridades próprias do tipo de relação que se estabelece entre os SRI e os usuários, e que não são visíveis nos modelos de comunicação apresen-

tados na literatura documentária; 4) caracteriza-se por ser construída a partir de noções como linguagem, subjetividade, intersubjetividade, referência e contexto, que são características presentes no referencial teórico que fundamenta a proposta apresentada.

No tópico a seguir, discute-se o aparelho formal da enunciação com a intenção de propor um novo referencial para a relação que se estabelece entre os SRI e os usuários.

Aparelho formal da enunciação e a relação SRI e usuários

O aparelho formal da enunciação proposto por Benveniste pode ser entendido como um dispositivo que permite que as línguas possam ser enunciadas; ele nada mais é do que a marcação da subjetividade na estrutura da língua. Segundo Benveniste (1989, p.82), "A enunciação é esse colocar em funcionamento a língua por um ato individual de utilização", pois, conforme Flores e Teixeira (2005, p.37), "[...] cada vez que o locutor se apropria do aparelho formal da enunciação - e por ele se apropria da língua toda - produz um uso novo e como tal irrepitível. Essa irrepitibilidade deve-se ao fato de que jamais tempo, espaço e pessoa - categorias fundamentais em enunciação - podem ser perenizados no uso da língua".

Pela compreensão das noções advindas do aparelho formal da enunciação - subjetividade, intersubjetividade, referência e contexto - e pelas categorias fundamentais em enunciação - pessoa (eu e tu), espaço (aqui) e tempo (agora) -, tornou-se possível perceber as peculiaridades e a complexidade da relação que se estabelece entre SRI e usuários, assim como elaborar e construir outro referencial sobre essa relação, que contempla os diversos aspectos que a caracterizam. A partir da configuração registrada (Quadro 2), são feitas algumas considerações.

Ao se inserirem as noções de subjetividade e de intersubjetividade, reconhece-se uma situação de interação entre os diferentes interlocutores envolvidos. A noção de referência aparece como definidora de sentido no nível semântico, sendo a referência ao mundo parte integrante do ato de enunciação: aqui entra a noção de contexto. É em relação ao contexto de situação que o sentido de uma palavra é obtido; fora do contexto, nem

Quadro 2. Aparelho formal da enunciação e a relação SRI e usuários.

Aparelho formal da enunciação	Relações SRI e usuários	
	<i>Primeiro momento</i>	<i>Segundo momento</i>
Categoria de pessoa subjetiva (eu)	Bibliotecário (eu)	Usuário (eu)
Categoria de pessoa não-subjetiva (tu)	Usuário (tu)	SRI (tu)
Categoria de não-pessoa (ele)	⇒ Linguagem especializada (ele)	Linguagem especializada (ele)
Categoria de espaço (aqui)	Vocabulário controlado (aqui)	Catálogo eletrônico (aqui)
Categoria de tempo (agora)	Processo de indexação (agora)	Processo de busca e recuperação da informação (agora)

Fonte: Elaborado pela autora.

sempre é possível compreender o sentido que resulta da junção das palavras.

Compreendendo que não se trata de uma situação de comunicação, mas de interação, essa proposta introduz a noção de interlocutores, identificando como interlocutores não apenas o sistema e o usuário, mas também o bibliotecário. São identificadas duas situações de interação, que ocorrem em espaço e tempo diferenciados: 1) uma entre bibliotecários e usuários, interagindo durante o processo de indexação, e 2) outra entre usuários e catálogos eletrônicos, também reconhecidos como SRI, na qual a ação de interlocução acontece durante o processo de busca e recuperação da informação.

Ressalta-se, ainda, que os termos e expressões que identificam o conteúdo informacional de um documento são localizados nos vocabulários controlados e nos catálogos eletrônicos. Os vocabulários controlados são as ferramentas com as quais os bibliotecários reconhecem a terminologia que identifica o conteúdo informacional de um documento. Os catálogos eletrônicos são as ferramentas utilizadas pelos usuários para recuperar a informação que desejam. Nas duas situações, caracterizadas pelos processos de indexação e de busca e recuperação da informação, a linguagem especializada é a fonte de informação.

Assim, entende-se que a relação estabelecida entre os SRI e os usuários não se limita ao processo de busca e recuperação da informação. É preciso que seja considerada também a relação de interação que ocorre durante o processo de indexação entre bibliotecários e usuários. A partir desses pressupostos, conclui-se que essa relação não pode ser vista do modo como tem sido apresentada em modelos de comunicação documentária,

ou seja, a partir de um emissor, canal, código intermediário e destinatário.

Considerações Finais

Finaliza-se este artigo afirmando que a teoria da enunciação apresenta vantagens teóricas e metodológicas importantes aos estudos desenvolvidos em linguagens documentárias. Esse referencial conduz o bibliotecário a novas reflexões e posicionamentos diante das atividades documentárias, seja na execução das atividades de indexação, seja na elaboração de vocabulários controlados, seja no estabelecimento de uma política de indexação para os sistemas.

O questionamento sobre a relação que se estabelece entre os SRI e os usuários é o seguinte: é essa realmente uma situação de comunicação? Essa discussão, feita a partir de parâmetros enunciativos, possivelmente levará a uma mudança de paradigma no âmbito documentário. Por ora, entende-se que, entre usuários e SRI, estabelece-se uma ação de interação, mas não de comunicação documentária, e como tal não pode ser descrita pelos elementos que usualmente a representam: emissor, canal, código intermediário e usuário.

A insatisfação do usuário na consulta aos catálogos eletrônicos leva à conclusão de que os SRI não estão cumprindo sua função de transferência da informação. Hoje, uma das maiores dificuldades enfrentadas pelo usuário de um sistema não é exatamente a recuperação de documentos, mas a recuperação de informações relevantes para seus propósitos de pesquisa, nos documentos recuperados. Essa realidade leva à seguinte pergunta: o que se espera de um SRI? Seu verdadeiro papel será o da transferência da informação?

Referências

- ARAUJO, V.M.R.H. Sistemas de informação: nova abordagem teórico conceitual. *Ciência da Informação*, v.24, n.1, p.54-76, 1995.
- BENVENISTE, E. *Problemas de linguística geral I*. 2.ed. Campinas: Pontes, 1988. p.277-293.
- BENVENISTE, E. *Problemas de linguística geral II*. Campinas: Pontes, 1989.
- BISCALCHIN, R.; BOCCATO, V.R.C. Estudo do uso de linguagem documentária em catálogos coletivos de bibliotecas universitárias: uma abordagem qualitativa-sociocognitiva pela perspectiva do usuário. In: CONGRESSO BRASILEIRO DE BIBLIOTECONOMIA, DOCUMENTAÇÃO E CIÊNCIA DA INFORMAÇÃO, 24., 2011, Maceió. *Anais eletrônicos...* Maceió: CBBD, 2011. Disponível em: <<http://febab.org.br/congressos/index.php/cbbd/xxiv/paper/view/39/508>>. Acesso em: 23 mar. 2012.
- CENDÓN, B.V. Sistemas e redes de informação. In: OLIVEIRA, M. (Coord.). *Ciência da informação e biblioteconomia: novos conteúdos e espaços de atuação*. Belo Horizonte: UFMG, 2005. p.45-75.
- FLORES, V.N.; TEIXEIRA, M. *Introdução à lingüística da enunciação*. São Paulo: Contexto, 2005.
- LARA, M.L.G. O processo de construção da informação documentária e o processo de conhecimento. *Perspectivas em Ciência da Informação*, v.7, n.2, p.127-139, 2002.
- LIMA, V.M.A. *Terminologia, comunicação e representação documentária*. 1998. Dissertação (Mestrado em Ciências da Informação) - Escola de Comunicação e Artes, Universidade de São Paulo, São Paulo, 1998. Disponível em: <<http://www.teses.usp.br/teses>>. Acesso em: 20 mar. 2012.
- ORTEGA, C.D.; LARA, M.L.G. A noção de estrutura e os registros de informação dos sistemas documentários. *TransInformação*, v.22, n.1, p.7-17, 2010. Disponível em: <<http://revistas.puc-campinas.edu.br/transinfo/viewissue.php?id=24#Artigos>>. Acesso em: 19 maio 2011.
- SHANNON, C.E.; WEAVER, W. *The mathematical theory of communication*. Illinois: The University of Illinois, 1949.
- SOUZA, R.R. Sistemas de recuperação de informações e mecanismos de busca na web: panorama atual e tendências. *Perspectivas em Ciência da Informação*, v.11, n.2, p.161-173, 2006.

