

Modelos de leitura documentária para indexação de diferentes tipos de documentos  
13 – Modelo de leitura documentária para indexação de artigos de jornal

Silvana Aparecida Fagundes

SciELO Books / SciELO Livros / SciELO Libros

FAGUNDES, S. A. Modelo de leitura documentária para indexação de artigos de jornal. In.: FUJITA, M. S. L., ALVES, R. C. V., and ALMEIDA, C. C., eds. *Modelos de leitura documentária para indexação: abordagens teóricas interdisciplinares e aplicações em diferentes tipos de documentos* [online]. Marília: Oficina Universitária; São Paulo: Cultura Acadêmica, 2020, pp. 323-345. ISBN: 978-65-8654-607-1. Available from: <http://books.scielo.org/id/96v3r>.  
<https://doi.org/10.36311/2020.978-65-86546-07-1.p323-346>.



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

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

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

## MODELO DE LEITURA DOCUMENTÁRIA PARA INDEXAÇÃO DE ARTIGOS DE JORNAL

*Silvana Aparecida FAGUNDES*

**RESUMO:** A leitura documentária consiste na identificação e extração de conceitos. A leitura documentária de jornais tem que ser rápida, devido à quantidade de material que necessita ser tratado. Propõe-se a elaboração de um modelo de leitura para indexação de artigos de jornais mediante revisão de literatura, estudo de caso com o Arquivo do jornal *O Estado de S. Paulo* e no Departamento de Documentação da Editora Abril (DEDOC) e observação da leitura dos indexadores com aplicação de Protocolo Verbal. Os resultados do teste do modelo indicam que o grupo de ação obteve melhor desempenho na análise porque selecionou termos que possibilitaram elaborar o enunciado de assunto, contendo a macroestrutura global do texto. Concluiu-se que o modelo de leitura proposto possibilita procedimentos que direcionam o indexador experiente ou inexperiente à uma compreensão durante a leitura documentária sem apelação ao “bom senso”, permitindo assim realizar uma representação do conteúdo do texto.

**PALAVRAS-CHAVE:** Indexação de jornal. Leitura documentária. Estrutura textual.

**ABSTRACT:** The documental reading consist of the identification and extraction of concepts. The documental reading of newspapers must be fast, due to the quantity of materials that needs to be treated. It is proposed the elaboration of a reading model to the indexing of newspaper articles according to literature review, case study in Archive of the journal *O Estado de São Paulo* and, Department of Documentation of Editora

Abril (DEDOC) and observation study of indexers reading with application of the Verbal Protocol. The results of the model test indicate that the action group had better performance in the analysis because they select terms that could elaborate the subject, containing the global macrostructure of the text. It is concluded that proposed reading model facilitates procedures that direct the experienced or unexperienced indexer to a comprehension during the documental reading without making use of the “common sense”, this way allowing them to achieve a representation of the text content.

**KEYWORDS:** Newspaper indexing. Documental reading. Textual structure.

## 1 INTRODUÇÃO

A Análise Documentária é uma área da Biblioteconomia e Documentação que engloba também a atividade de indexação dos documentos por meio do processo de exame do documento, ou seja, da leitura documentária.

A leitura documentária é a fase inicial da operação de análise documentária que se constitui basicamente das fases: leitura do texto e, neste momento, também ocorre a identificação e seleção de conceitos que representem o conteúdo do texto, e representação desses conceitos selecionados através de uma linguagem documentária. O produto dos procedimentos de Análise Documentária são os índices de assuntos e resumos.

Na leitura documentária ocorrem os mesmos processos mentais presentes na leitura normal. E Kato (1986) distingue dois tipos de estratégias utilizados durante a leitura:

- estratégias metacognitivas: conscientes, usadas frente a um determinado problema;
- estratégias cognitivas: automáticas, subconscientes, usadas na leitura fluida sem obstáculos para o leitor.

Destacamos que na leitura para fins de indexação a Norma ISO (5963, 1985) orienta a proceder o exame do documento buscando informação precisa em curto período de tempo, analisando: título; resumo; lista de conteúdos; introdução, frases que iniciam capítulos e parágrafos,

e conclusão; ilustrações, diagramas, tabela; palavras ou grupos de palavras que se encontram destacadas.

A Análise Documentária deixou de restringir o corpus documental com o qual trabalha, que na maioria das vezes é o texto escrito geralmente em suporte papel, e passou a englobar outros documentos em diferentes suportes.

Evidenciando que nosso objetivo de pesquisa é a observação da indexação de artigos de jornal, enfatizaremos a estrutura textual do texto jornalístico.

De acordo com Van Dijk (1983, p. 37), a superestrutura do texto jornalístico na cultura ocidental é composta por categorias que exercem uma relação de subordinação e superordenação entre si, como se observa abaixo:

- 1- Sumário/Introdução
  - 1.1 Cabeçalho
  - 1.2 Lead
- 2. Episódio(s)
  - 2.1 Informações prévias
    - 2.1.2 Antecedentes
    - 2.1.2.3 Eventos Presentes
    - 2.1.3.1 Explicação
      - 2.1.3.1.1 Contexto
      - 2.1.3.1.2 Experiências passadas
    - 2.1.3.2 Eventos principais
  - 2.2 Consequências/ Reações
- 3. Comentários
  - 3.1 Expectativas
  - 3.2 Avaliação (Tradução nossa)

Para o linguista Van Dijk (1983), no texto jornalístico é convencional a apresentação de um resumo do acontecimento que foi abordado. Esse resumo pode ser expresso pelas letras grandes separadas do resto do texto, resumo ou introdução denominado de “Lead”,

componente inicial da estrutura textual do artigo de jornal. Com essa determinação é possível chamar a atenção do leitor que decidirá se o texto lhe interessa ou não.

E na opinião de Silva (2001), o LEAD constitui a abertura e arquitetura do texto e que nas notícias e nos textos de abertura, no 1º e 2º cadernos, o LEAD deverá respeitar formalmente as regras clássicas contendo: O quê (o que aconteceu, está ou vai acontecer); Quem (os agentes da ação); Quando (dia da semana e do mês, horas); Onde (o local do acontecimento); Como (as circunstâncias) e Porque (os motivos e as razões).

Após conhecer a área de Análise Documentária, surgiu o interesse em investigar a atividade de leitura dos indexadores, uma vez que, na literatura sobre o assunto, são raras as publicações sobre a atividade de leitura documentária. A nosso ver, essa deveria ser caracterizada como a atividade primordial do processo de indexação, porque sendo a primeira fase da análise do documento, o bom ou o mau resultado desta refletirá na boa, ou na má qualidade das outras atividades (elaboração de resumos e índices).

Com o conhecimento da escassez de pesquisas sobre leitura documentária, optamos então por investigar a atividade de leitura dos indexadores de jornais porque essa modalidade de leitura, segundo a literatura da área de Análise Documentária, deve ser realizada rapidamente já que “ao indexador raramente é dado o luxo de ler um texto do começo ao fim” (LANCASTER, 1993, p. 20). Considerando a grande quantidade de documentos que o indexador tem para tratar, além das outras atividades que, muitas vezes, tem para desenvolver dentro do centro de informação no qual trabalha e a grande quantidade de matérias jornalísticas publicadas diariamente que precisam ser indexadas, podemos supor que, no caso da indexação de jornais, a leitura documentária deverá ser mais rápida ainda. E também podemos supor que a observação dessa leitura rápida proporcionará a outros indexadores o conhecimento dos métodos utilizados pelos indexadores experientes e suas vantagens.

Neste contexto, percebemos a necessidade de observar a leitura dos indexadores de jornais para sabermos como agiam esses profissionais diante da realização da tarefa de indexar, procurando observar se utilizavam determinadas estratégias e quais eram elas e se tais estratégias eram consideradas compatíveis com o que indica a literatura da área. E com

os dados obtidos, comparados com informações relevantes recomendadas pela revisão de literatura sobre a temática, elaboramos um modelo de leitura para indexação de jornais.

## **2 OBSERVAÇÃO DA LEITURA DO INDEXADOR DE ARTIGO DE JORNAL: PROCEDIMENTOS METODOLÓGICOS**

Ressaltamos que decidimos observar a leitura dos indexadores de artigos de jornais no contexto dos centros de informações especializados que possuíam acervo organizado e desenvolviam a atividade de indexação. Nessa pesquisa, a ambiência da observação da prática de indexação de jornais ocorreu primeiramente no contexto do Arquivo do jornal *O Estado de S. Paulo* e foi realizada durante o nosso curso de graduação, como trabalho de conclusão de curso, em 1997. Durante o triênio de 1999-2001, continuamos com a pesquisa no curso de mestrado em Ciência da Informação na Universidade Estadual Paulista “Júlio de Mesquita Filho” (UNESP) – Campus de Marília, tendo como ambiência de observação o Departamento de Documentação da Editora Abril (DEDOC). O DEDOC foi escolhido visando obter subsídios capazes de complementar a pesquisa anterior, enriquecendo-a com a observação da indexação em duas ambiências: arquivo do jornal *O Estado de S. Paulo*, realizado durante o curso de graduação em 1997, e DEDOC da Editora Abril, para elaborarmos o modelo de leitura para indexação de jornais.

Para atingir os objetivos propostos, a pesquisa optou por uma metodologia de coleta e análise de dados cujos processos estão descritos na seqüência.

Dentre os processos metodológicos adotados, primeiramente, foi realizado um levantamento bibliográfico em fontes de informação na área, onde se pesquisou a literatura relevante publicada sobre: leitura documentária, estrutura do texto jornalístico, identificação de conceitos e metodologias para avaliação de política de indexação e linguagens documentárias.

Com o levantamento bibliográfico, identificamos bibliografias relevantes que compõem os pressupostos teóricos desta pesquisa, em que se destacam: a metodologia para indexação manual de jornais mencionada por Ahmad (1991); tipologias e estruturas textuais de textos trabalhados

pela Análise Documentária citada por Kobashi (1994), a classificação das concepções de análise de assunto proposta por Albretchtsen (1993), a estrutura do texto jornalístico elaborada por Van Dijk (1983) analisada e comentada por Amaro (1991) e ainda as considerações de Silva (2001) para essa estrutura textual, atividades metacognitivas presentes na leitura consciente estabelecidas por Brown (1980), Kato (1986), Norma ISO (5963, 1985), elementos de política de indexação estabelecidos por Carneiro (1985), avaliação da estrutura externa de tesouros elaborada por Corrêa (1998), metodologia de avaliação da estrutura interna de linguagens documentárias elaborada por Ginez de Lara (1993), metodologia de análise de assunto do sistema de indexação PRECIS comentada por Fujita (1989) e modelo de leitura para indexação de artigo de periódico científico também elaborado por Fujita (2000-2002).

Posteriormente, para a observação da leitura dos indexadores foram necessários estudos de casos nos centros de informação especializados (Arquivo do jornal *O Estado de S. Paulo* e DEDOC) que desenvolviam a indexação de jornais. Nessa fase de estudos de casos, foi realizada consulta à documentação e conversa com os responsáveis pelos setores, possibilitando uma caracterização da infraestrutura física e funcional dos centros englobando: histórico breve e área física, o organograma da instituição, recursos humanos, recursos materiais (mobiliário, equipamento e acervo), usuários e serviços prestados. E também foi elaborada entrevista para a caracterização dos indexadores dos sistemas de informações, com a qual foram obtidos dados referentes à formação educacional, experiência profissional anterior e no serviço de indexação, outras atividades realizadas na instituição, dificuldades na atividade de indexação e procedimentos de identificação de conceitos e fontes de experiência de indexação.

Identificamos a política de indexação dos centros de informações como uma das variáveis que influenciam a leitura do indexador. Percebemos que a política de indexação desses centros influencia na escolha da linguagem de indexação e, conseqüentemente, no resultado da leitura do indexador.

Quando realizamos os estudos de casos analisamos, também, a política de indexação do Arquivo do jornal *O Estado de S. Paulo* e do DEDOC, por meio da verificação da ocorrência de alguns elementos estabelecidos por Carneiro (1985) para implantação de uma política de indexação.

E como a escolha da linguagem documentária adotada pelo centro de informação sofre influência da política de indexação, e esta determina a comunidade usuária que a instituição servirá, fez-se necessário avaliar a estrutura externa das linguagens documentárias, seguindo a metodologia elaborada por Corrêa (1998), e também avaliar a estrutura interna somente da linguagem do DEDOC, seguindo a metodologia de Ginez de Lara (1993), porque a linguagem do Arquivo do jornal *O Estado de S. Paulo* não apresentou parâmetros funcionais para se avaliar a sua estrutura interna.

A observação da leitura documentária dos indexadores de jornal, no Arquivo do jornal *O Estado de S. Paulo* e no DEDOC, foi realizada com a aplicação do Protocolo Verbal, um método de observação introspectivo, muito utilizado na área de Linguística Aplicada para investigar as ações do indivíduo durante a realização da leitura, de modo a gravar a exteriorização de seus processos mentais em busca da compreensão. Esse método é denominado “Pensar Alto”.

A utilização da técnica introspectiva denominada “Pensar Alto”, ou seja, Protocolo Verbal, possibilita que os pensamentos dos sujeitos (indexadores) sejam gravados e transcritos literalmente, produzindo protocolos verbais. Os protocolos são definidos como relatos verbais dos processos mentais conscientes do sujeito. Em outras palavras, eles se referem ao “Pensar Alto” do sujeito enquanto realiza uma tarefa de qualquer natureza. (CAVALCANTI, 1989; CAVALCANTI; ZANOTTO, 1994).

Esse método de observação ultrapassou os limites da área de Psicologia Cognitiva e passou a ser utilizado na Linguística Aplicada, na pesquisa sobre leitura por meio dos trabalhos de Hosenfeld (1977) referentes à leitura em língua estrangeira, e do trabalho de Olshavsky (1976-1977) sobre leitura em língua materna.

No Brasil, particularmente no Programa de Pós-Graduação em Linguística Aplicada ao Ensino de Línguas da PUC/São Paulo (LAEL), protocolos têm sido usados como instrumentos de pesquisa em dissertações como a de Nardi (1993), cujos parâmetros metodológicos serviram de base para o desenvolvimento do Projeto Integrado “Leitura em Análise Documentária”, pertencente ao Grupo de Pesquisa da Universidade Estadual Paulista “Júlio de Mesquita Filho” (UNESP) - Campus de Marília, coordenado por Fujita (2000-2002).



É inédita no Brasil a observação do processo de leitura documentária para indexação com uso do Protocolo Verbal. Destaca-se que a pesquisa realizada por Fujita no Projeto Integrado “Leitura em Análise Documentária”, com indexadores da Sub-Rede de Informações em Ciências da Saúde Oral, provavelmente é pioneira na utilização deste método.

Na literatura internacional o uso do Protocolo Verbal para observação do processo de indexação foi relatado por Gotoh (1983) em artigo que discute os problemas do comportamento de processamento da informação na atividade de indexação humana. Nesse trabalho, o autor realizou um experimento com dois sujeitos indexadores utilizando a técnica de Protocolo Verbal.

A técnica de Protocolo Verbal é, no momento, segundo Nardi (1993), o único instrumento de coleta disponível que possibilita observar processos do leitor durante a compreensão de um texto. Destaca-se que a referida pesquisadora utilizou a técnica em sua pesquisa de mestrado (NARDI, 1993), e obteve êxito nos resultados, o que justifica a nossa escolha por esta técnica.

Nesta pesquisa, com os indexadores de jornais, os procedimentos da coleta de dados em ambas as instituições (Arquivo do jornal *O Estado de S. Paulo* e DEDOC) foram:

## **1- PROCEDIMENTOS ANTERIORES À COLETA DE DADOS:**

- **Seleção do Texto-Base**

Para a seleção do texto-base escolheram-se textos inéditos, ou seja, desconhecidos e ainda não indexados por nenhum dos indexadores.

Quanto à temática os textos selecionados eram referentes aos Cadernos, ou Editorias dos Jornais com os quais cada indexador trabalhava, variando-se, assim, a temática de cada texto.

Destaca-se que para a observação da leitura dos indexadores do Arquivo do jornal *O Estado de S. Paulo* e do DEDOC foram selecionados dez textos, sendo que cada indexador (caracterizados como sujeitos 1,2,3,4 e 5) indexou dois textos.

- **Seleção dos Sujeitos**

A seleção dos sujeitos considerou como critério o tempo de permanência no sistema de informação e na atividade de indexação acrescentando ainda “habilidade em indexação” refletida na quantidade de documentos indexados. Foram selecionados cinco indexadores que foram classificados como sujeitos 1 e 2 (pertencentes ao Arquivo do jornal *O Estado de S. Paulo*) e sujeitos 3, 4 e 5 (pertencentes ao DEDOC).

- **Conversa informal com os sujeitos**

Foi realizada uma conversa informal com cada um dos sujeitos, durante a qual foram expostos os objetivos da pesquisa, destacando sua relevância para o desenvolvimento da área de Análise Documentária. Nessa conversa, ressaltou-se que a identidade dos sujeitos manter-se-ia oculta, visando assim deixá-lo mais à vontade e não comprometer os dados coletados. Foi solicitado que cada sujeito realizasse a leitura do texto-base naturalmente como o faz no seu dia-a-dia tendo como objetivo a identificação e seleção de conceitos para a indexação.

- **Familiarização com a tarefa do “Think Aloud” (Pensar Alto)**

Antes da aplicação do Protocolo Verbal, foi realizada uma familiarização com a tarefa de “Pensar Alto” por meio da leitura do texto “Instruções aos Sujeitos” elaborado por Nardi (1993), tendo o propósito de descontrair e, ao mesmo tempo, apresentar procedimentos que pudessem auxiliar os sujeitos no desempenho da tarefa.

## **2- PROCEDIMENTOS DURANTE A COLETA DE DADOS:**

- **Gravação do “Pensar Alto” durante a leitura do texto-base**

Anteriormente à gravação, foi entregue aos sujeitos o texto-base, solicitando-os a “Pensar Alto” durante toda a leitura e exteriorizar seus processos mentais, procurando esquecer a presença da pesquisadora que somente controlava o gravador e lembrava que era preciso “Pensar Alto”. A aplicação do Protocolo Verbal foi realizada, com cada um dos indexadores,

bem como a conversa informal, a familiarização com a realização da tarefa e a leitura do texto “Instruções aos Sujeitos”.

### 3- PROCEDIMENTOS APÓS A COLETA DE DADOS

- **Transcrições literais das gravações da leitura dos sujeitos**

As transcrições foram feitas de maneira a destacar a compreensão dos sujeitos, suas dúvidas, equívocos, identificação e seleção de termos. Para melhor visualização dos processos adotados pelos sujeitos, destacamos cada um deles com um tipo de notação específica, como descrito em Fagundes (2001, p. 132):

Letra normal: vocalização das palavras exatas do texto durante a leitura

*Itálico*: comentários do leitor

**Negrito**: termos identificados

..PS: pausa

( ) : comentários da pesquisadora

⇐ : indexador volta a trechos do texto

Destacamos um trecho da análise da observação da atividade de leitura, de um sujeito dos estudos de caso fazendo uso das notações adotadas na transcrição (grifadas abaixo):

Bailarina acusa o marido ator de espancamento e dá queixa na polícia. *Bailarina acusa marido e ator registra queixa de agressão*. Briga de casal teria acontecido durante a festa de aniversário do filho. Ana Carolina Torres. (A pesquisadora observou que nesse momento o indexador seleciona o nome do autor e digita na ficha de indexação, onde o mesmo é obrigatório). *Então essa aqui como envolve pessoas públicas quer dizer bailarina, artista o casal trabalhou em novelas coisa do tipo então a gente marca o que aborda a matéria*. A atriz e bailarina Luciana Bessa, de 21 anos, registrou queixa contra o marido, o ator Ademir Zanyor, de 26, na 32ª DP (Jacarepaguá), no domingo passado. Ela o acusa de tê-la espancado durante a festa de aniversário de dois anos de Uriel, filho único do casal. ...PS.. E

*you vai pro nome da bailarina* (Nesse momento o indexador atribui como assunto da matéria o nome da bailarina) BESSA, Luciana e o ator chama ZANYOR, Danilo. (Posteriormente o indexador comenta) *Entramos pelo nome do agressor e da vítima então* aqui lemos ⇐ bailarina acusa marido ator ..PS.. *Então, aqui não precisa colocar bailarina, ator porque o título já está definindo isso e quando o título define mais ou menos o que está na matéria não tem necessidade de colocar muitos termos.*

Durante a observação da leitura dos indexadores experientes, estabelecemos um parâmetro de análise que considerou os aspectos destacados no QUADRO 1 abaixo:

Quadro 1 - Identificação da Natureza Metacognitiva Na Leitura Documentária.

ASPECTOS METACOGNITIVOS DE LEITURA ESTABELECIDOS POR BROWN REFORMULADAS POR KATO	RECOMENDAÇÕES DA NORMA ISO 5963 ADAPTADAS AO TEXTO JORNALÍSTICO	OUTROS ASPECTOS DO COMPORTAMENTO DO INDEXADOR DURANTE A LEITURA
<ul style="list-style-type: none"> <li>· Explicitação dos objetivos da leitura e/ ou manutenção dos objetivos na mente;</li> </ul>	<ul style="list-style-type: none"> <li>· Análise do documento com domínio da estrutura textual, considerando partes do texto é uma recomendação da Norma ISO. Verificaremos a exploração da estrutura do texto jornalístico proposta por Van Dijk (1983); analisada e comentada por AMARO (1991) e SILVA (2001)</li> </ul>	<ul style="list-style-type: none"> <li>· Associação com Linguagem;</li> </ul>
Identificação de aspectos importantes da mensagem;	<ul style="list-style-type: none"> <li>· Identificação de conceitos (abordagem sistemática mediante questionamento)</li> </ul>	<ul style="list-style-type: none"> <li>· Seleção de conceitos (Termos ou descritores)</li> </ul>

Alocamento de atenção a áreas importantes;	· Seleção de conceitos	
* Exploração da estrutura textual		
· Monitoramento do comportamento através de:		
* Engajamento em revisão e auto-indagação para ver se o objetivo está sendo atingido;		
*Tomada de ações corretivas quando são detectadas falhas na compreensão		
<b>*Recobramento de atenção quando a mente se distrai ou faz digressões</b>		

Fonte: FAGUNDES (2001)

### 3 RESULTADOS DA OBSERVAÇÃO DA LEITURA DOS INDEXADORES DE ARTIGO DE JORNAL

A análise das transcrições dos textos permitiu identificar que os indexadores experientes, durante a indexação, seguem uma sistemática de identificação de conceitos, utilizando seis tipos de comportamento:

- Realizam o destaque de palavras do texto por meio de argolamento ou digitação (esse argolamento não foi expresso verbalmente, a pesquisadora observou e anotou);
- Realizam associações com a linguagem após reconhecimento de termos;
- Fazem questionamento durante a leitura do texto quanto à escolha dos descritores;
- Identificam que há maneiras diferentes de ler;
- Evidenciam conhecimento da estrutura do texto jornalístico;
- Usam conhecimento prévio do assunto e recorrem à memória de longo prazo.

É relevante mencionar que com os resultados obtidos durante a observação da leitura dos indexadores experientes e também com os dados dos estudos de revisão da literatura conseguimos obter subsídios para elaboração do modelo. Assim, foi possível identificar as respostas às nossas suposições, referindo-nos, primeiramente, ao fator tempo disponível para ler, em que supúnhamos que a leitura documentária de jornais deveria ser rápida devido à grande quantidade de material publicado. A análise dos protocolos evidenciou que a leitura dos indexadores de jornais é rápida, e que um dos fatores que proporciona esta rapidez é o conhecimento da estrutura do texto jornalístico. Tal estrutura está composta de tal forma que, geralmente, nos primeiros parágrafos encontramos um resumo com respostas aos questionamentos - Quem? O que? Como? (Para que?) Onde? Porque? (Para que?) Quando? - que permitem ao leitor compreender o texto e identificar o tema abordado antes de ler o texto detalhadamente. Entretanto, os indexadores não explicitaram, durante a gravação da indexação dos textos selecionados, conhecerem o questionamento proposto pela estrutura do texto, mas demonstraram indícios de conhecer a estrutura textual, porque concentraram a seleção de conceitos nos itens “*Lead*” e “Episódio” componentes da estrutura textual da notícia, que se localizam no começo do texto jornalístico.

Além disso, o que proporciona rapidez na leitura dos indexadores de jornais é o conhecimento prévio sobre o tema a indexar, porque percebemos que o leitor não identifica e seleciona o tema do texto se não tiver compreendido o texto. E para que ocorra compreensão, é preciso que o conhecimento prévio (adquirido com indexação de textos anteriores relacionados ao tema presente, ou por meio de leitura, etc.) referente ao tema interaja, no momento da leitura, com as informações do texto. Percebemos, então, que a memória do indexador é um fator primordial para a acumulação de conhecimento prévio sobre o assunto com o qual trabalha, porque as matérias são publicadas por vários dias, então algo novo vai se introduzir nas matérias novas, mas o tema abordado, já conhecido, provavelmente será o mesmo.

Uma outra suposição era que os indexadores de jornais experientes deveriam utilizar determinadas estratégias para realizar essa leitura rápida, e interessava-nos descobrir quais eram essas estratégias. A análise dos protocolos tornou possível o estabelecimento da sistemática

de identificação de conceitos, que os indexadores utilizaram durante a leitura, recorrendo a diferentes estratégias como exploração da estrutura textual, considerando, na maioria dos textos, a leitura do “Lead”, uso do conhecimento prévio, entre outras. Entretanto, notou-se que após seleção de conceitos, realizaram pouca associação com Linguagem. (Linguagem Documentária do Centro de Informação.)

Evidencia-se que questionávamos, também, se as estratégias utilizadas eram compatíveis com o que recomendava a literatura. Os resultados da observação da leitura dos indexadores nos comprovaram que sim, porque eles utilizam exploração da estrutura textual, uso do conhecimento prévio, entre outras estratégias.

Sendo assim, confirmamos nossas suposições, evidenciando que os indexadores experientes de jornais são estratégicos, rápidos e dominam a estrutura textual para identificação e seleção de conceitos. Com a confirmação das hipóteses, acima, atingimos, primeiramente, um dos nossos objetivos que era a identificação de estratégias de leitura para indexação de jornal. E isto conjuntamente com dados relevantes encontrados na revisão de literatura, nos possibilitou alcançar um segundo objetivo que era elaborar um modelo de leitura para indexação de artigos de jornal, considerando os resultados da observação de leitura dos indexadores e recomendações da literatura.

#### **4 LEITURA DOCUMENTÁRIA PARA INDEXAÇÃO DE ARTIGOS DE JORNAIS: PROPOSIÇÃO E AVALIAÇÃO DE MODELO**

Com a análise dos resultados da revisão de literatura sobre a temática e a observação prática de indexação de jornal, elaboramos o modelo de leitura para indexação de artigos de jornais, visualizado abaixo, que acompanha um roteiro de utilização:

##### **ROTEIRO DE UTILIZAÇÃO DO MODELO**

- 1 - Inicie a leitura do texto buscando prever por meio de interpretação da ilustração (fotografia) ou da leitura do título, o assunto que o texto abordará;
- 2 - Posteriormente, comece a ler o texto utilizando o seu conhecimento prévio sobre o assunto, direcionando a sua leitura para os cinco

primeiros parágrafos, sem realizar leitura linear, buscando encontrar respostas aos questionamentos propostos pelo modelo de leitura descrito mais especificamente no quadro abaixo (alertamos que nem todas as respostas das questões estarão presentes num texto);

3 - Encontre no texto as possíveis respostas aos questionamentos: O que?; Quem?; Quando?; Onde?; Como?; (ou Para que?); Por que?;

4 - Identifique as respostas considerando, nesse momento, o conteúdo do texto, visando compreender do que o texto trata, qual o seu assunto;

5 - Selecione, dentre as respostas encontradas, somente os conceitos que interessem à comunidade usuária que fará uso de determinado texto;

6 - Após selecionar o conceito que representa o conteúdo do texto e que interessa aos usuários, compare os conceitos selecionados com a Linguagem do sistema (no caso DEDOC);

7 - Por fim, elabore um resumo considerando o questionamento do Lead organizando as respostas obtidas em ordem lógica para as questões: O que?; Quem?; Quando?; Onde?; Como?; (ou Para que?); Por que?; (ou Para que?).

Quadro 2 - Modelo de leitura para indexação de jornais: identificação de conceitos por questionamento evidenciando a estrutura textual jornalística.

<p>CONCEITO PARA ANÁLISE DE ASSUNTO DO SISTEMA (PRECIS)</p>	<p>PARTE DA ESTRUTURA TEXTUAL (VAN DIJK, 1983) ONDE PODEMOS ENCONTRAR O QUESTIONAMENTO DO LEAD E OS ELEMENTOS DO SISTEMA PRECIS</p>	<p>QUESTIONAMENTO AO (LEAD) DA ESTRUTURA DO TEXTO JORNALÍSTICO SEGUNDO VAN DIJK (1983); SILVA (2001)</p>	<p>LEMBRE-SE QUE PARA ATINGIR O OBJETIVO DA LEITURA DOCUMENTÁRIA O INDEXADOR DEVE:</p>
<p>AÇÃO</p>	<p>Sumário/Introdução (Cabeçalho e Lead)</p>	<p>O que (o que aconteceu, está ou vai acontecer)</p>	<p>Identificar conceitos considerando o conteúdo do documento</p>



AGENTE	Sumário/Introdução (Lead)	Quem (os agentes da ação)	Selecione os conceitos considerando as necessidades informacionais da comunidade usuária
MÉTODOS DO AGENTE	Sumário/Introdução (Lead) ou Episódio (s)	Como, ou Para quê (as circunstâncias)	Padronize os conceitos com termos de uma linguagem de indexação. E se preciso crie novos termos.
LOCAL OU AMBIÊNCIA	Sumário/Introdução (Cabeçalho) Sumário/introdução (cabeçalho e Lead)	Quando (dia da semana e do mês, horas)  Onde (o local do acontecimento)	
CAUSA E EFEITO	Sumário/Introdução (Lead) ou Episódio (s)	Por que, ou Para que (os motivos e as razões)	

Fonte: FAGUNDES (2001).

Após elaboração do modelo, realizamos uma aplicação com um corpus de textos selecionados para testar a viabilidade de seu funcionamento. Selecionamos no DEDOC três indexadores considerados inexperientes, porque realizavam a atividade de indexação esporadicamente, para participarem da fase de avaliação do modelo de leitura, visando verificar o seu funcionamento na atividade de indexação, considerando o contexto dos centros de informação especializados que tratam a notícia e que demandam rapidez no tratamento dessa tipologia documentária.

A fase de avaliação do modelo, na qual participaram os três indexadores caracterizados como: Clara, Ana e Maria, seguiu

os procedimentos adotados para observar a leitura dos indexadores experientes. Ressalta-se que, diferentemente da primeira fase, não houve uma preocupação em selecionar textos de temática específica porque seria utilizado por sujeitos (indexadores) que indexam textos diversos, mas nos preocupamos em selecionar dois textos, caracterizados como (A e B), que apresentavam mais completamente a estrutura completa do texto jornalístico.

O texto A foi utilizado pelos três sujeitos Ana, Maria e Clara, mas somente o sujeito Clara o indexou fazendo uso do modelo de leitura (porque já conhecia o modelo e sabia utilizá-lo) e o texto B foi utilizado para a pesquisadora ensinar ao indexador Clara sobre o funcionamento do modelo, numa fase anterior à indexação do Texto A. Um outro procedimento diferente ocorrido na fase de avaliação foi a gravação do depoimento do sujeito Clara quanto à utilização do modelo, a fim de que, se preciso, o mesmo pudesse ser aperfeiçoado.

Destaca-se que na fase de avaliação do modelo de leitura com o segundo grupo de indexadores do DEDOC, (indexadores inexperientes), os dados levantados do protocolo do indexador que num primeiro momento foi orientado sobre o funcionamento do modelo (caracterizado como sujeito Clara) para que num segundo momento tivesse condições de utilizar individualmente o modelo, demonstraram que o questionamento do Lead presente no modelo foi um dos fatores que, conjuntamente com o conhecimento prévio mínimo do assunto, possibilitaram ao indexador realizar a indexação do texto selecionado. Esse fato foi evidenciado durante a gravação do Protocolo Verbal nos momentos em que o sujeito indaga que no dia-a-dia deixaria de atribuir termos (assuntos) para representar o texto por se tratar de um texto difícil de compreender, pois era um texto da área médica.

Na discussão dos resultados quanto à avaliação do modelo de leitura proposto, notou-se que fazendo uso das orientações do modelo, os conceitos selecionados pelo indexador que foi orientado para utilizar o modelo, possibilitaram uma representação mais fidedigna do tema global comunicado pelo texto, do que os conceitos selecionados pelos dois indexadores que não fizeram uso e desconheciam o modelo, porque aumentou as possibilidades de recuperação do texto, selecionando mais termos.

Para uma melhor visualização dos dados obtidos na fase de avaliação do modelo elaboramos os QUADROS 3 e 4:

Quadro 3 - Indexação do texto A não fazendo uso do modelo de leitura proposto, realizada pelo sujeito Ana.

CONCEITOS (PRECIS)	PARTE DA ESTRUTURA TEXTUAL (VAN DIJK)	CONCEITO (LEAD)	TERMOS IDENTIFICADOS	LINGUAGEM DEDOC
AÇÃO	Lead	O que (o que aconteceu, está ou vai acontecer)	Experiência Científica	Experiência Científica
AGENTE	Cabeçalho e Lead	Quem (os agentes da acção)	Instituto de Pesquisa Scripps	
MÉTODOS DO AGENTE	Lead	Como (as circunstâncias)		
LOCAL OU AMBIÊNCIA	Lead	Onde (o local do acontecimento)		
CAUSA E EFEITO	Lead	Por que (os motivos e as razões)		

**Enunciado de Assunto: O Instituto de Pesquisa Scripps realizou uma experiência científica.**

Fonte: FAGUNDES (2001).

Percebemos que o sujeito Ana, com os conceitos que atribuiu ao texto, conseguiu responder o que aconteceu de uma forma bem genérica, e quem o realizou, sem conseguir descrever a informação que o texto comunica de forma global.

Quadro 4 - Indexação do texto A não fazendo uso do modelo de leitura proposto, realizada pelo sujeito Maria.

CONCEITOS (PRECIS)	PARTE DA ESTRUTURA TEXTUAL (VAN DIJK)	CONCEITO (LEAD)	TERMOS IDENTIFICADOS	LINGUAGEM DEDOC
AÇÃO	Lead	O que (o que aconteceu, está ou vai acontecer)		Experiência Científica
AGENTE	Cabeçalho e Lead	Quem (os agentes da acção)		
MÉTODOS DO AGENTE	Lead	Como (as circunstâncias)	Medicamento-antibiótico	
LOCAL OU AMBIÊNCIA	LeadLead	Onde (o local do acontecimento)		
CAUSA E EFEITO		Por que (os motivos e as razões)	Infecção hospitalar	

**Enunciado de Assunto: Vai se desenvolver um medicamento-antibiótico para se combater infecção hospitalar.**

Fonte: FAGUNDES (2001).

Notamos que o sujeito Maria, sem usar o modelo de leitura proposto, consegue, muito vagamente, com os termos que selecionou, responder o que se causou e como se pretende combater o problema.

Quadro 5 - Indexação do texto A fazendo uso do modelo de leitura proposto, realizada pelo sujeito Clara.

CONCEITOS (PRECIS)	PARTE DA ESTRUTURA TEXTUAL (VAN DIJK)	CONCEITO (LEAD)	TERMOS IDENTIFICADOS	LINGUAGEM DEDOC
AÇÃO	Lead	O que (o que aconteceu, está ou vai acontecer)	Bactéria	Bactéria
AGENTE	Cabeçalho e Lead	Quem (os agentes da acção)	Instituto de Pesquisa Scripps	
MÉTODOS DO AGENTE	Lead	Como (as circunstâncias)	Molécula	Molécula
LOCAL OU AMBIÊNCIA	LeadLead	Onde (o local do acontecimento)	La Jolla/ Califórnia	
CAUSA E EFEITO		Por que (os motivos e as razões)	Infecção hospitalar	Infecção-hospitalar

**Enunciado de Assunto: O Instituto de Pesquisa Scripps localizado em La Jolla na Califórnia criou uma nova classe de moléculas para combater bactéria que causa infecção hospitalar.**

Fonte: FAGUNDES (2001).

## 5 CONSIDERAÇÕES FINAIS

Com a utilização do modelo de leitura proposto, os conceitos selecionados pelo sujeito Clara, provavelmente abrangeram todas as possibilidades de termos para recuperação do texto e possibilitou organizar o enunciado de assunto de maneira lógica, resultando em uma representação mais fidedigna do tema global vinculado pelo texto.

Pretendemos na continuidade da pesquisa aprimorar o modelo de leitura proposto realizando uma investigação teórica mais abrangente

a respeito da estrutura do texto jornalístico, visando identificar em sua organização elementos linguísticos de apresentação do tema que compõem as respostas ao questionamento proposto pelo “Lead” e presentes no modelo, bem como, uma hierarquização desses elementos, diferenciando os relevantes dos secundários no relato de uma notícia.

Percebe-se que identificar o tema é uma tarefa difícil de realizar, como também de ensinar, pois em muitos casos este se encontra no texto de maneira complexa e pouco explícita.

Logo, identificar e representar o tema no documento não é uma tarefa fácil, e até o momento é realizada de maneira mais eficaz pelo indexador humano e essa é temática carente de investigação na literatura da área.

Concordamos com Campos (1987) que a área de indexação deve apresentar instrumentos para análise e interpretação do documento e que esses venham fornecer ao indexador possibilidades de compreender coerentemente o texto analisando o seu conteúdo e considerando as necessidades informacionais da comunidade usuária.

O modelo de leitura que elaboramos é provavelmente original e a possibilidade de sua utilização na indexação de jornais é relevante para a área de Análise Documentária, como para os centros de informação especializados que tratam a notícia.

## **REFERÊNCIAS**

- AHMAD, N. Newspaper indexing: an international overview. *The Indexer*, Liverpool, v. 17, n. 4, p. 257-266, 1991.
- ALBRETCHTSEN, H. Subject analysis and indexing: from automated indexing to domain analysis. *The Indexer*, Liverpool, v. 8, n.4, p. 219-224, 1993.
- AMARO, R. K. F. *Contribuição da análise do discurso para a análise documentária: o caso da documentação jornalística*. 1991. Dissertação (Mestrado em Ciências da Comunicação) - Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 1991.
- BROWN, A. L. Metacognitive developmental and reading. In: SPIRO, R. J. *et al.* (org.). *Theoretical issues in reading comprehension*. Hillsdale: Lawrence Erlbaum, 1980. cap. 19.

- CAMPOS, A. T. A indexação. *Revista de Biblioteconomia de Brasília*, Brasília, v. 15, n. 1, p. 69-72, 1987.
- CARNEIRO, M. V. Diretrizes para uma política de indexação. *Revista da Escola de Biblioteconomia: UFMG*, Belo Horizonte, v. 14, n.2, p. 221-241, 1985.
- CAVALCANTI, M. C. *Interação leitor-texto: aspectos de interação pragmática*. Campinas: UNICAMP, 1989.
- CAVALCANTI, M. C.; ZANOTTO, M. S. Introspection in applied linguistics: meta-research on verbal protocols. In: SCOTT, B. (ed.) *Reflections on language learning*. Cleverdon: Multilingual Matters, 1994. p. 148-156.
- CORRÊA, A. O. *A construção de tesouros na perspectiva da metodologia facetada*. 1998. 117 p. Relatório (Pesquisa) - UNESP, FAPESP, Marília, 1998.
- FAGUNDES, S. A. *Leitura em análise documentária de artigos de jornal*. 1997. Trabalho de Conclusão de Curso (Bacharelado em Biblioteconomia e Documentação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista “Júlio de Mesquita Filho” Paulista, Marília, 1997.
- FAGUNDES, S. A. *Leitura em análise documentária de artigos de jornal*. 2001. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista “Júlio de Mesquita Filho”, Marília, 2001.
- FUJITA, M. S. L. *Leitura em análise documentária: uma contribuição à formação do indexador*. 185 p. Relatório Parcial (Projeto Integrado de Pesquisa) - UNESP, CNPq, Marília, 2000-2002.
- FUJITA, M. S. L. *Precis na língua portuguesa: teoria e prática de indexação*. Brasília: UNB, 1989.
- GINEZ DE LARA, M. L. *A representação documentária: em jogo a significação*. 1993. Dissertação (Mestrado em Ciências da Comunicação) – Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 1993.
- GOTOH, T. Cognitive structure in human indexing process. *Library and Information Science*, Bingley, n. 21, p. 209-226, 1983.
- HOSENFELD, C. A preliminary investigation of the reading strategies of successful and unsuccessful second language learners. *System*, Oxford, v. 5, p. 110-123, 1977.
- INTERNATIONAL ORGANIZATION FOR STANDARDIZATION.  
Documentation: methods for examining documents, determining their subjects, and selecting indexing terms. Suíça: ISO. 5 p. 9 (ISO 5963-1985).
- KATO, M. *No mundo da escrita: uma perspectiva psicolinguística*. São Paulo: Ática, 1986.

- KOBASHI, N. Y. *A elaboração de informação documentária: em busca de uma metodologia*. 1994. Tese (Doutorado em Ciências da Comunicação) – Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 1994.
- LANCASTER, F.W. *Indexação e resumos: teoria e prática*. Tradução: Antônio Agenor Briquet de Lemos. Brasília: Briquet de Lemos, 1993.
- NARDI, M. I. A. *As expressões metafóricas na compreensão de texto escrito em língua estrangeira*. 1993. Dissertação (Mestrado em Linguística Aplicada ao Ensino de Línguas) – Pontifícia Universidade Católica, São Paulo, 1993.
- OLSHAVSKY, J. E. Reading as problem solving: an investigation of strategies. *Reading research Quartely*, Hoboken, v. 12, n. 4, p. 654-674, 1976-1977.
- SILVA, V. J. *Regras de construção*. Disponível em: [http://www.publico.pt/nos/livro\\_estilo/paginas/1e12\\_regras.htm](http://www.publico.pt/nos/livro_estilo/paginas/1e12_regras.htm). Acesso em: 27 mar. 2001.
- VAN DIJK, T. A. Discourse analysis: its development and application to the structure of news. *Journal of Documentation*, Bingley, v. 33, n. 2, p. 20-23, 1983.