

SILVANA APARECIDA FAGUNDES

“LEITURA EM ANÁLISE DOCUMENTÁRIA DE ARTIGOS DE JORNAIS”

Mestrado em Ciência da Informação

UNESP

MARÍLIA - SÃO PAULO

2001

SILVANA APARECIDA FAGUNDES

“LEITURA EM ANÁLISE DOCUMENTÁRIA DE ARTIGOS DE JORNAIS”

*Dissertação de Mestrado
apresentado ao Curso de
Pós-Graduação em Ciência
da Informação.*

Orientadora: Prof. Dr. Mariângela Spotti Lopes Fujita

*Marília
São Paulo - 2001*

Fagundes, Silvana Aparecida.

Leitura em Análise Documentária de artigos de jornais / Silvana Aparecida Fagundes. -- Marília, 2001.

261 p. ; 30 cm.

Dissertação (mestrado) - - Universidade Estadual Paulista. Faculdade de Filosofia e Ciências de Marília, 2001.

Bibliografia: p. 239-249

1. Leitura Documentária 2. Indexação de jornais. I. Autor. II. Título.

CDD20 – 025.35

BANCA EXAMINADORA

Prof. Dr. Mariângela Spotti Lopes – Orientadora

Prof. Dr. Maria Isabel Asperti Nardi

Prof. Dr. Jose Moreiro Gonzalez

Prof. Livre Docente José Augusto Chaves Guimarães

Prof. Dr. Maria Lúcia Balestreiro

AGRADECIMENTOS

Primeiramente a DEUS que me ajudou na conclusão de mais essa etapa na minha vida.

À minha FAMÍLIA, especialmente a minha mãe a quem devo a vida e a vontade de viver.

À minha orientadora Professora Doutora Mariângela Spotti Lopes Fujita a quem devo o ensinamento de se viver a vida acadêmica , tudo aquilo que eu sou ou espero ser eu devo ao anjo que foi minha orientadora .

A professora Doutora Maria Isabel Asperti Nardi que aceitou o convite de ter participado da banca de qualificação desse trabalho e que muito ajudou com esclarecimentos de “expert” no assunto e que vem agora participar também da banca de defesa dessa pesquisa.

Ao Professor Livre Docente José Augusto Chaves Guimarães por ter aceito participar da banca de qualificação dessa pesquisa e que muito contribuiu com seus conhecimentos.

Ao professor Doutor José Antonio Moreiro González que gentilmente aceitou participar da banca de defesa dessa pesquisa e que muito poderá contribuir com seus conhecimentos.

A Professora Doutora Maria Lúcia ---- que agradavelmente aceitou o convite para participar da banca de defesa dessa pesquisa.

Aos meus amigos especialmente a: Maria, Maura, Daniela, Ana Cláudia, Aline, Lúcia, Marisa, Júlio, Patrícia, Vinícius,

Denise, Ana Beatriz, Ivete, Ivone, Sandra, Raquel, Fernando, Juliana, Yoshi e os mais recentes: Malú, Tatiana, Maria Vitória e Joana aos quais considero como se fossem antigos e aqueles que por ventura esqueci de agradecer, me perdoem.

Aos funcionários da Biblioteca, Seção de Pós-Graduação e Escritório de Pesquisa que sempre gentilmente prestaram seus serviços.

À Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) que contribuiu e financiou o desenvolvimento da pesquisa.

RESUMO:

“Leitura em Análise Documentária de Artigos de Jornais”.

A leitura documentária consiste na identificação e extração de conceitos que irão representar o conteúdo dos documentos. Difere da leitura normal porque ocorre em condições específicas e contextos diferentes. Considerando que a leitura documentária de jornais tem que ser rápida, devido à quantidade de material que necessita ser tratado, interessa investigar os procedimentos para realizar essa leitura rápida. 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 São Paulo” e no Departamento de Documentação da Editora Abril Cultural (DEDOC) e estudo de observação da leitura dos indexadores com aplicação de Protocolo Verbal. Os resultados de informações relevantes encontradas na revisão de literatura e os dados dos estudos de casos conjuntamente comparados e acoplados com a análise da leitura dos indexadores experientes, possibilitou atingir os objetivos de identificar estratégias de leitura para jornais e fornecer subsídios teórico-práticos para a análise documentária de artigos de jornal e, principalmente, elaborar o modelo de leitura documentária para indexar jornais. O funcionamento do modelo foi verificado a partir de um teste com outros dois grupos de indexadores inexperientes do DEDOC. Esse teste consistiu na observação da leitura com Protocolo Verbal de dois grupos de indexadores: um grupo foi orientado para uso do modelo de leitura proposto (grupo de ação); enquanto o outro grupo indexou o mesmo texto sem orientação do modelo de leitura (grupo de controle). Os resultados obtidos antes da elaboração do modelo de leitura revela que os indexadores experientes são estratégicos e rápidos. O que proporciona essa leitura rápida é o domínio da estrutura textual do texto jornalístico e conhecimento prévio dos indexadores. 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 proposto, contendo a macroestrutura global do texto que representou mais adequadamente o conteúdo do texto em uma frase que organizou, de forma coerente e lógica, os conceitos extraídos. 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:

“Newspaper Article Reading in Documental Analysis”

The documental reading consists of the identification and extraction of concepts that will represent the content of the documents. It differs from the normal reading because it occurs in specific conditions and different contexts. Considering that the documental reading of newspapers must be fast, due to the quantity of materials that needs to be treated, it is interesting to investigate the procedures to achieve this fast reading. It is proposed the elaboration of a reading model to the indexing of newspaper articles according to literature review, case study in the Archive of the journal “O Estado de São Paulo” and, Department of Documentation of Editora Abril Cultural (DEDOC) and observation study of indexers’ reading applying Verbal Protocol. The results of relevant information found in the literature review and the data of the case studies all compared and together with the analysis of reading of experient indexers made us achieve the objectives of identifying reading strategies to newspapers and supply theoretical and practical subsidies to the documental analysis of newspaper articles and, mainly, to elaborate the model of documental reading to index newspapers. The performance of the model was verified by a test with two other groups of unexperienced indexers from DEDOC. This test consisted in the observation of reading through Verbal Protocol with two groups of indexers: one group was oriented to use the proposed reading model (action group); while the other group indexed the same text without orientation of the reading model (control group). The obtained results before the elaboration of the reading model revealed that the experienced indexers are strategic and fast. What makes this reading fast is the knowledge of textual structure of the newspaper text and previous knowledge of the indexers. The results of the model test indicate that the action group had better performance in the analysis because they selected terms that could elaborate the subject, containing the global macrostructure of the text, representing more adequately the content of the text in a sentence that organized, in a logical and coherent way, the extracted concepts. It is concluded that the 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.

Key-words: journal indexing; documental reading; textual structure

SUMÁRIO

INTRODUÇÃO	1
Capítulo 1: As Concepções de Leitura para a Leitura Documentária	11
1.1 Leitura Documentária	32
Capítulo 2: A Estrutura Textual na Leitura do Indexador	38
2.1 A Estrutura Textual de Artigo de Jornal no Processo de Leitura Documentária	45
2.2 A Diplomática e o Texto Jornalístico: Análise da Tipologia Documentária	58
Capítulo 3: Identificação de Conceitos na Indexação	64
3.1 Indexação de Documentos	85
3.2 Indexação de Artigos de Jornal	95
Capítulo 4: Metodologia	100
Capítulo 5: Leitura Documentária para Indexação de Texto Jornalístico em Centros Referenciais Especializados: Estudos de Casos	104
5.1 Arquivo do jornal “O Estado de São Paulo”	105
5.2 Departamento de Documentação da Editora Abril Cultural (DEDOC)	108
5.3 Avaliação de Linguagens Documentárias: Estudos de Casos	112
Capítulo 6: Modelo de Leitura para Indexação de Artigos de Jornais: subsídios da observação leitura documentária do indexador	129
6.1 Observação de Estratégias do Indexador na Leitura Documentária: resultados e discussão	132
6.2 Modelo de Leitura para Artigo de Jornal	196
Capítulo 7: Avaliação do Modelo de Leitura para Indexação de Jornais: metodologia e resultados	203
Capítulo 8: Considerações Finais	230
9 Referências Bibliográficas	241
APÊNDICES:	

Apêndice A: Grade Diplomática do Texto Jornalístico	253
Apêndice B : Organização e Funcionamento do Arquivo	254
Apêndice C: Caracterização dos Indexadores do Arquivo	255
Apêndice D: Organização e Funcionamento do DEDOC	256
Apêndice E: Caracterização dos Indexadores do DEDOC	257
Apêndice F: Transcrição da Leitura dos Indexadores Experientes do Arquivo e DEDOC	258
Apêndice G: Transcrição da Leitura dos Indexadores Inexperientes do DEDOC	259
ANEXOS:	
Anexo A: Exemplos da Linguagem do Arquivo	260
Anexo B: Cópia do Texto-Base utilizados pelos Indexadores Experientes	261
Anexo C: Cópia do Texto Instruções aos Sujeitos	262
Anexo D: Fichas de Indexação do DEDOC	263
Anexo E: Cópia do Texto-Base utilizados pelos Indexadores Inexperientes	264

SUMÁRIO DE QUADROS

Quadro 1: Classificação dos Conectores	22
Quadro 2: Identificação da Natureza Metacognitiva na Leitura Documentária	133
Quadro 3: Análise das Operações Desenvolvida pelos Sujeitos 1 e 2	152
Quadro 4: Análise da Seqüência das Operações Desenvolvida pelos Sujeitos 3 e 4	153
Quadro 5: Análise da Seqüência das Operações Desenvolvida pelo Sujeito 5	154
Quadro 6: Seqüência da Estrutura Textual Seguida pelos Sujeitos 1, 2, 3, 4 e 5	158
Quadro 7: Identificação de Conceitos e Partes do Texto Sujeito 1	161
Quadro 8: Identificação de Conceitos e Partes do Texto Sujeito 2	162
Quadro 9: Identificação de Conceitos e Partes do Texto Sujeito 3	163
Quadro 10: Identificação de Conceitos e Partes do Texto Sujeito 4	164
Quadro 11: Identificação de Conceitos e Partes do Texto Sujeito 5	165
Quadro 12: Termos Identificados e Seleccionados pelos Sujeitos 1 e 2	173
Quadro 13: Termos Identificados e Seleccionados pelos Sujeitos 3 e 4	174
Quadro 14: Termos Identificados e Seleccionados pelo Sujeito 5	175
Quadro 15: Termos Identificados e Seleccionados Comparados com Linguagem pelos Sujeitos 1, 2 e 3	177
Quadro 16: Termos Identificados e Seleccionados Comparados com Linguagem pelos Sujeitos 4 e 5	178
Quadro 17: Demonstração da Aplicação do modelo de leitura para texto Científico (FUJITA, 2000)	198
Quadro 18: Modelo de Leitura para Texto Jornalístico	201
Quadro 19: Aplicação do Modelo de Leitura para Texto Jornalístico	202
Quadro 20: Análise das Operações Desenvolvida pelos Sujeitos Ana, Maria e Clara	212
Quadro 21: Seqüência da Estrutura Textual Seguida pelos Sujeitos Ana, Maria e Clara	215

Quadro 22: Identificação de Conceitos e Partes do Texto – Sujeito Ana	217
Quadro 23: Identificação de Conceitos e Partes do Texto – Sujeito Maria	217
Quadro 24: Identificação de Conceitos e Partes do Texto – Sujeito Clara	218
Quadro 25: Termos Identificados e Seleccionados pelo Sujeito Ana	221
Quadro 26: Termos Identificados e Seleccionados pelo Sujeito Maria	222
Quadro 27: Termos Identificados e Seleccionados pelo Sujeito Clara	222
Quadro 28: Termos Identificados e Seleccionados Comparados com Linguagem pelos Sujeitos Ana, Maria e Clara	224
Quadro 29: Indexação do Texto A pelo Sujeito Ana sem fazer uso do Modelo de leitura	226
Quadro 30: Indexação do Texto A pelo Sujeito Maria sem fazer uso do Modelo de leitura	227
Quadro 31: Indexação do Texto A pelo Sujeito Clara fazendo uso do Modelo de leitura	228
Quadro 32: Modelo Definitivo de Leitura para Indexação de Texto Jornalístico	236

1- INTRODUÇÃO

A Análise Documentária é uma das áreas da Biblioteconomia e Documentação que engloba tanto o tratamento temático dos documentos por meio de atividades como a classificação bibliográfica e a indexação, como também se encarrega do tratamento descritivo dos documentos incluindo aqui a atividade de catalogação dos mesmos.

Evidenciam-se que as atividades de classificação bibliográfica, indexação e catalogação desenvolvidas pela Análise Documentária têm como resultado final tornar os documentos organizados de tal maneira que possam ser recuperados por uma determinada comunidade usuária, preferivelmente em um curto espaço de tempo, e que venham, certamente, satisfazer os interesses informacionais dessa comunidade.

A área de Análise Documentária, com as pesquisas desenvolvidas pelo Grupo de Pesquisa *Tratamento Temático da Informação* do curso de Graduação em Biblioteconomia da Universidade Estadual Paulista – Campus de Marília, vem desenvolvendo experimentações para adaptar, ou elaborar metodologias que possam tornar a atividade de indexação mais ágil e eficiente e, para tanto, tem pesquisado a elaboração de modelos de leitura para indexação de textos científicos, resumos e artigos de jornal. E ainda investiga a interpretação da metáfora na leitura do indexador, e elaborou uma metodologia caracterizada como método diplomático também desenvolvido por um integrante do grupo para auxiliar na leitura do indexador por meio da evidência de algumas partes da estrutura dos documentos, independente da tipologia documental e suporte que pertence.

Com as investigações iniciadas pelo Grupo de Pesquisa *Tratamento Temático da Informação* do Curso de Graduação em Biblioteconomia e Documentação da Universidade Estadual Paulista, a área de 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 como: rótulo de shampoo, plantas arquitetônicas, receitas culinárias, entre outros.

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

Com o conhecimento da escassez de pesquisas sobre leitura documentária, talvez até por ser vista como um assunto não relevante, optamos, então, por investigar a atividade de leitura dos indexadores de artigos de jornal porque se a leitura documentária, segundo a escassa literatura deve ser realizada rapidamente já que *ao indexador raramente é dado o luxo de ler um texto do começo ao fim* dada a grande quantidade de documentos que tem para tratar, além das outras atividades, que muitas vezes, tem que desenvolver dentro do sistema de informação no qual trabalha. Sendo assim, supúnhamos que, no caso da indexação de jornal, a leitura documentária deveria ser mais rápida ainda, devido à grande quantidade de material (matérias jornalísticas) publicada diariamente e considerando essa problemática, supúnhamos também, que esses indexadores deveriam fazer uso de algumas estratégias para realizar uma leitura tão rápida.

Decidimos observar a leitura dos indexadores de artigos de jornais no contexto de sistemas de informações especializados que possuíam acervo organizado e que desenvolvessem a atividade de indexação. Nessa pesquisa, a ambiência da observação da prática de indexação de jornais ocorreu no contexto do Arquivo do Jornal “O Estado de São Paulo” ressaltando que essa foi realizada já durante o nosso curso de graduação como trabalho de conclusão de curso em 1997 e, agora em 2000-2001, durante o mestrado, no Departamento de Documentação da Editora Abril Cultural, visando comparar os dados obtidos na pesquisa anterior, na graduação, para confirmar com uma segunda análise da observação de leitura nossas suposições e, também, para termos suficientes subsídios com a observação da indexação em duas ambiências (arquivo do jornal “O Estado de São Paulo e DEDOC da editora Abril) a fim de alcançarmos os objetivos desta pesquisa de mestrado: 1- proporcionar aos indexadores, subsídios metodológicos para a análise de assunto, de textos com estrutura diversificada; 2- identificar estratégias de leitura para artigos de jornal e 3- fornecer subsídios teórico-metodológicos que contribuam para a elaboração de um modelo de leitura documentária de artigos de jornal.

Para atingir os objetivos propostos, a pesquisa optou por uma metodologia que se compôs de cinco procedimentos seguidos de quatro metodologias de coleta e análise dos dados. Primeiramente, foi realizado o procedimento de revisão de literatura cuja metodologia foi consulta em fontes bibliográficas na área (revistas especializadas, bases de dados e Internet) sobre: leitura, leitura documentária, estrutura textual de artigos de jornal, identificação de conceitos, metodologias para avaliação de linguagens documentárias, e para caracterizar a política de indexação dos sistemas de informações envolvidos na

pesquisa. Os dados relevantes encontrados na literatura foram trabalhados nos capítulos referentes aos pressupostos teóricos e englobam os capítulos 1, 2 e 3.

Portanto, evidenciamos agora as informações relevantes presentes nos pressupostos teóricos mencionando que, no capítulo 1, referente a concepções de leitura, enfocamos que a leitura global é uma atividade criativa que exige do leitor atitudes conscientes frente a tarefa de ler, que os tipos de leitura que se devem utilizar dependem do objetivo da leitura, que esse objetivo geralmente é encontrar a informação que considera importante, e que os processos de leitura seguidos pelo leitor, enquanto realiza a leitura, estão descritos nos modelos de leitura existentes na literatura também citados nesse capítulo da pesquisa.

Destacamos que os processos de leitura, assim como o conhecimento permanente do leitor estão armazenados na memória de longo-termo em formas de esquemas que permitem uma recuperação rápida de algum esquema ativado durante a leitura.

Elucidamos ainda que existem modelos de compreensão de leitura que mencionam que a compreensão do texto envolve três variáveis: leitor, texto e contexto e também os princípios comunicacionais presentes num ato de comunicação entre um emissor para um receptor, e que o não-reconhecimento dessas variáveis influi na compreensão de leitura. E ressaltamos, ainda, que durante a leitura ocorre, dependendo de algumas variáveis, a transferência e aquisição de cultura por meio do texto, vendo nessa situação a leitura como um evento social de transmissão de conhecimento entre indivíduos da mesma ou de diferentes sociedades. Posteriormente, comentamos as estratégias de leitura, identificando que existem estratégias metacognitivas (conscientes) e cognitivas (inconscientes). E que o conhecimento de estratégias pelo leitor, durante a leitura, ajuda-o a buscar a compreensão do texto, e que o conhecimento de estrutura textual é uma estratégia de leitura. E ainda, nesse capítulo, relatamos a atividade de leitura documentária identificando que a mesma difere da leitura global, por ter objetivo definido e estar sob limite de tempo exigindo, portanto, outros procedimentos previstos em normas metodológicas para exame de documentos da área de Análise Documentária e por autores que pesquisam sobre leitura documentária. Damos ênfase também às estratégias de leitura evidenciando, nesse momento, que, a cada novo texto, o indexador poderá utilizar estratégias de leitura diferentes das usadas em textos anteriores e que, por isso, não existem estratégias fixas podendo o uso dessas depender da finalidade de leitura, do grau de conhecimento do indexador sobre o assunto do texto que está lendo, entre outras variáveis.

Com as informações do capítulo 1 foi possível conhecer mais sobre a atividade de leitura e o processos envolvidos nesta e, atingir o objetivo 2- *identificar estratégias de leitura para artigos de jornal*.

Posteriormente, ainda comentando os dados da revisão de literatura no capítulo 2 sobre os pressupostos teóricos, tratamos as estruturas textuais encontradas com a revisão de literatura evidenciando que essas exercem fortes influências na compreensão do leitor, destacamos como estão estruturadas e quais os conceitos pertencentes a cada parte das estruturas dos textos. Mencionamos as estruturas textuais utilizadas na Análise Documentária, identificamos uma estrutura de organização para o texto científico e a estrutura textual de artigos de jornal, como também evidenciamos o que é a notícia jornalística, quais as formas de narrativas jornalísticas, citamos a proposta de um modelo de metodologia de ensino da redação jornalística encontrado na literatura e, realizamos a aplicação do método diplomático, citado anteriormente, com textos jornalísticos evidenciando as partes mais significativas dessa estrutura textual para os fins de indexação.

No capítulo 3, último dos pressupostos teóricos, discorremos sobre a identificação de conceitos na atividade de indexação, ocorrida durante a leitura documentária do documento, destacando as recomendações de como identificar conceitos elaboradas pelas normas metodológicas da área de Análise Documentária, pelo sistema de indexação Preci e uma metodologia específica para a atividade de indexação de textos científico encontrada na literatura, bem como as observações de pesquisadores (autores) sobre identificação do tema durante a leitura, como o modelo de identificação da macroestrutura temática do texto proposto por lingüistas que pesquisam a compreensão na leitura e a metodologia de análise de assunto por meio de adaptação do modelo de identificação da macroestrutura temática proposta pelos lingüistas conjuntamente com o uso da análise proposicional (PA) em determinado texto. Ainda trabalhamos, brevemente, nesse capítulo, elementos que estão envolvidos na indexação, como a política de indexação do sistema e mencionamos, nesse momento, que caracterizamos a política de indexação das instituições envolvidas na pesquisa, utilizando a metodologia para estabelecimento de uma política de indexação de uma autora conceituada encontrada na revisão de literatura; citamos também de forma breve a relevância da linguagem do sistema de informação, destacando as metodologias de avaliação de linguagens de indexação consideradas, a nosso ver, como as mais relevantes na literatura, e destacamos que fizemos uso destas para avaliar as linguagens

documentárias das instituições envolvidas na pesquisa; evidenciamos ainda, nesse capítulo, as concepções dos sistemas de informação quanto à identificação de conceitos encontradas na literatura referindo-nos às concepções orientadas para o conteúdo do documento, para a demanda tendo em vista os usuários do sistema, e a simplista que considera os assuntos como entidades absolutas.

Com as informações obtidas nos capítulos 2 e 3 foi possível atingir o objetivo 1- *proporcionar aos indexadores, subsídios metodológicos para a análise de assunto de textos com estrutura diversificada* e mencionar também que levantamos, nesse capítulo, os procedimentos para indexação de jornal baseando-nos em um texto levantado na literatura que aborda a indexação manual de jornais e, com isso, atingimos mais especificadamente o objetivo 2- *identificar estratégias de leitura para artigos de jornal*.

No capítulo 4, discorremos sobre as metodologias utilizadas na pesquisa como: revisão de literatura, o segundo procedimento referente à caracterização da ambiências da pesquisa por meio de consulta a documentos e entrevistas com responsáveis pelo setor nas instituições, o terceiro procedimento sobre a avaliação das linguagens de indexação das instituições fazendo uso de metodologias de avaliação de linguagens encontradas na literatura e o quarto procedimento que diz respeito à caracterização da política de indexação das instituições por meio de uma metodologia encontrada na literatura para esse fim; e comentamos mais especificamente a metodologia denominada “Protocolo Verbal”, ou “Pensar Alto” surgida na Psicologia Cognitiva e adaptada para a leitura, que foi utilizada aqui para realizar o quinto procedimento que foi observar a atividade de leitura dos indexadores do DEDOC e do Arquivo do jornal “O Estado de São Paulo”, porque possibilita posterior análise de todos os pensamentos ocorridos durante a gravação da leitura em áudio. Destacamos ainda os procedimentos que adotamos para aplicar o Protocolo Verbal nas instituições observadas como: textos inéditos para os indexadores lerem, e gravação da leitura e, na seqüência, foi realizada a transcrição literal da leitura dos textos adotando sinais que possibilitassem destacar os processos mentais utilizados pelos leitores.

No capítulo 5, descrevemos o segundo procedimento da pesquisa que foi a caracterização das instituições Departamento de Documentação da Editora Abril Cultural e Arquivo do jornal “O Estado de São Paulo” onde se observou a prática da atividade de indexação de jornais.

Evidenciamos a organização, ou caracterização da instituição por meio de metodologia como consulta a documentação, entrevista com responsáveis pelo setor e indexadores. Destacamos, nessa etapa, histórico breve, infra-estrutura (em apêndices), caracterização dos indexadores destacando formação educacional, fontes de experiência de indexação, dificuldades, procedimentos de identificação de conceitos, nível de especificidade e exaustividade exigido no sistema, outras atividades que realizam na instituição. Posteriormente, comparamos os dados obtidos nas duas instituições.

Em seguida, ainda no capítulo 5, realizamos a avaliação das linguagens de indexação das instituições porque estas muitas vezes marcam influência no resultado da leitura dos indexadores e também para conhecer as características dessa linguagem documentária presente nos sistemas de informação jornalístico que difere-se de instituições bibliotecas que já conhecemos até então pela vasta publicações que a literatura da área de Análise Documentária menciona. Portanto as avaliações basearam-se numa metodologia da área visando análise formal da estrutura das linguagens e, como se fez necessário, utilizamos outra metodologia de avaliação de linguagens objetivando analisar a estrutura interna das linguagens de indexação ambas mencionadas no capítulo 4 de metodologias adotadas na pesquisa. Entretanto, destacamos que a metodologia de análise da estrutura interna da linguagem foi possível aplicar somente à linguagem do DEDOC. Posteriormente, apresentamos os resultados da avaliação das linguagens.

No capítulo 6, apresentamos a análise dos dados coletados com a observação da leitura dos indexadores experientes, fazendo uso da metodologia *Protocolo Verbal*. Ressaltamos que iniciamos a análise dos protocolos verbais, que antes foram gravados e transcritos, buscando identificar na leitura os aspectos metacognitivos estabelecidos como relevantes, quando discorremos sobre o assunto na revisão de literatura, e ainda os associados aos da Norma ISO (5963 de 1985) e a exploração da estrutura textual do texto jornalístico citado nos pressupostos teóricos, e ainda o uso da linguagem no momento da leitura e associação com linguagem após seleção de termos, ambos foram demonstrados em quadros, comentados e comparados.

A análise dos protocolos evidenciou que a leitura dos indexadores de jornais é rápida e que um dos fatores que proporcionam 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 quê? Como? (Para quê?) Onde? Por quê? (Para quê?) Quando? que permite ao

indexador compreender o texto e identificar o tema abordado antes de ler o texto detalhadamente. Os indexadores experientes evidenciaram conhecer o questionamento proposto pela estrutura do texto e apresentaram conhecimento da estrutura textual, porque concentraram a seleção de conceitos nos itens “Lead e Episódio”. Além disso, o que proporciona rapidez na leitura dos indexadores de jornais é o conhecimento prévio do 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 conhecimento prévio (adquirido com indexação de textos anteriores referente ao tema presente, por meio de leitura, etc...) referente ao tema e que o mesmo interaja no momento da leitura, com as informações do texto. Percebemos, então, que a memória do indexador é um fator primordial para o uso 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 provavelmente é o mesmo.

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”, o questionamento para verificar compreensão, o uso do conhecimento prévio, etc... Entretanto, notou-se que após seleção de conceitos poucos realizam associação com linguagem. Nós perguntávamos 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, etc...

Em seguida, nesse capítulo, mencionamos os dados relevantes levantados com a revisão de literatura e os estudos de casos, e dividimos este tópico em parte conceitual que envolve as considerações teóricas e práticas enfocadas na literatura pertinente à Análise Documentária evidenciando a atividade de leitura dos indexadores de jornais por meio da observação de identificação e seleção de conceitos, e da exploração da estrutura textual. E a outra é a parte diagnóstica na qual se realiza uma avaliação comparada dos dados coletados a partir da observação da leitura dos indexadores dos dois estudos de casos, o arquivo do jornal “O Estado de São Paulo” e o Departamento de Documentação da Editora Abril. Portanto, tratamos nesse capítulo com a divisão dos dois tópicos (parte conceitual e diagnóstica) de levantar as informações relevantes obtidas no estudo de revisão de literatura realizado nos capítulos 1, 2, e 3 buscando identificar as informações da literatura

nos resultados da análise do diagnóstico incluindo desde a observação da leitura com protocolo verbal, avaliação das linguagens e organização, e funcionamento das instituições observadas com o intuito de comentá-los e compará-los, evidenciando resultados obtidos para propor a elaboração do modelo de leitura.

Com comparação e análise de dados da revisão de literatura e da observação prática da leitura dos indexadores de jornal confirmamos nossas suposições, evidenciando que os indexadores de jornais são estratégicos, rápidos e dominam a estrutura textual para a identificação e seleção de conceitos. Com a confirmação das hipóteses acima, atingimos primeiramente dois objetivos que são: identificar estratégias de leitura para artigos de jornal, e proporcionar aos indexadores subsídios metodológicos para a análise de assunto de textos com estrutura diversificada. Isso nos possibilita alcançar o terceiro objetivo que é elaborar um modelo experimental de leitura para artigos de jornal, considerando os resultados da observação de leitura dos indexadores experientes e recomendações da literatura, possibilitando assim que outros indexadores de jornais (que deduzimos ter pelo menos conhecimento prévio mínimo do assunto) conheçam e utilizem esse modelo para suprir deficiências referentes à localização dos conceitos durante a indexação.

Ainda no capítulo 6, atingimos o terceiro objetivo que é 3- *fornecer subsídios teórico-metodológicos que contribuam para a elaboração de um modelo de leitura documentária de artigos de jornal* quando propomos o modelo experimental para leitura de artigos de jornais nos baseando no modelo encontrado na literatura direcionado à leitura de textos científicos e nos dados relevantes focalizados aqui nos capítulos 1, 2, 3, 4, 5 e 6. Destacamos, então, que o modelo para indexação de jornal proposto abordou a concepção de análise de assunto orientada para o conteúdo do documento; o questionamento de análise de assunto do sistema de indexação Precís; a estrutura textual de artigos de jornais evidenciando os elementos componentes do “Lead” da estrutura textual e as partes mais relevantes dessa estrutura textual identificadas quando se aplicou o método diplomático desenvolvido por integrante do grupo de pesquisa *Tratamento Temático da Informação* nessa tipologia documentária (artigo de jornal) e, baseou-se ainda no modelo de leitura para indexação de texto científico elaborado também por integrante do nosso grupo de pesquisa *Tratamento Temático da Informação*. É relevante evidenciar que o modelo foi demonstrado no final desse capítulo 6 por meio de um exemplo de indexação de um texto de jornal que realizamos.

No capítulo 7, após elaboração do modelo de leitura, realizou-se a avaliação do modelo proposto utilizando novamente a técnica de observação de leitura Protocolo Verbal ou “Pensar Alto” com três indexadores inexperientes com prováveis dificuldades em indexação, do DEDOC (Departamento de Documentação da Editora Abril Cultural), visando verificar se o modelo de leitura é eficiente, ou seja, se auxilia os indexadores que apresentam dificuldades durante a indexação na identificação de conceitos representantes do conteúdo de determinado documento porque esse é o objetivo maior da pesquisa.

Esclarecemos que o modelo de leitura proposto foi avaliado da seguinte maneira: o funcionamento do modelo foi ensinado a um indexador que compõe o grupo de indexadores do DEDOC que realiza a atividade de indexação esporadicamente porque realiza outras atividades na instituição e também porque não possui experiência na realização da tarefa e provavelmente apresenta dificuldades para identificar conceitos nos textos. No mesmo momento, aplicamos também a técnica de Protocolo Verbal sem fazer uso do modelo de leitura proposto para observarmos a leitura de um par, ou seja, dois indexadores do DEDOC que também realiza a atividade de indexação esporadicamente. Essa etapa foi realizada para que no final pudéssemos comparar os resultados atingidos pelo par e pelo outro o indexador e verificar a provável eficácia do modelo proposto. Os resultados da avaliação do modelo, comparando o desempenho da leitura do par de indexadores e do outro indexador observados, nesse momento, consideraram também os mesmos parâmetros de análise dos dados utilizados anteriormente na análise dos dados dos protocolos do indexadores experientes descritos aqui no capítulo 6. Destacamos que os resultados obtidos demonstraram que o modelo de leitura para indexação de jornal possibilita, tanto ao indexador com anos de prática na atividade de indexação, como a um indexador iniciante considerando que ambos apresentam dificuldades para identificar conceitos, eficácia na realização da tarefa de indexação por meio de uma leitura rápida do texto jornalístico o que vem a ser necessário devido a quantidade de artigos a indexar.

No capítulo 8, destacamos que a elaboração do modelo de leitura para indexação de jornal, conciliando as recomendações da revisão de literatura e a prática de indexação de jornal, que era o objetivo maior da pesquisa, foi alcançado com êxito porque demonstrou eficácia na realização da tarefa de identificação de conceitos durante a leitura para indexação de artigo de jornal. E também, na conclusão da pesquisa, foi indicado ao DEDOC que avalie a sua política de indexação procurando verificar o grau de satisfação dos seus usuários com o sistema de recuperação do acervo em funcionamento na

instituição porque os dados levantados baseando-se em depoimentos de indexadores revelam que as diversas possibilidades de recuperação de documentos oferecidas pelo sistema de informação é o ideal, e que os usuários as estão utilizando satisfatoriamente mas é preciso investigar, com quem faz uso das mesmas, se, na prática de utilização do sistema de recuperação do acervo da instituição, estas diversas possibilidades de acesso ao acervo são eficazes, precisas e lhes deixam realmente satisfeitos.

Por fim, concluímos a pesquisa, neste capítulo, mencionando que o modelo de leitura demonstrou ser eficiente na identificação de conceitos durante a indexação. Entretanto, destacamos que o modelo considerou, na sua elaboração, a exploração e representação do conteúdo do documento, não se atendo a questão de considerar os interesses informacionais da comunidade usuária à qual o documento se destina. Por isso, foi incluído, no roteiro de utilização do modelo de leitura proposto, um item alertando o indexador que pense, no momento de selecionar o conceito, se o assunto atribuído ao documento será de interesse da comunidade usuária porque organizar para recuperar o que se deseja é o objetivo da indexação.

CAPÍTULO 1: AS CONCEPÇÕES DE LEITURA PARA A LEITURA DOCUMENTÁRIA

A atividade de leitura, com o passar dos séculos, não mais se restringe ao texto escrito em suporte de papel, mas engloba também outros diferentes suportes impressos, ou eletrônicos: disquetes, CD-ROM, fita K-7, Internet, micro-fichas, etc.... Como consequência, o processo de aprendizagem também se modificou. Pesquisadores confirmam que os computadores estão sendo utilizados para auxiliar no processo de aprendizagem da leitura, complementando-o, sem pretender jamais substituir os professores.

Entretanto, limitaremos este capítulo referente à atividade de leitura global, estratégias de leitura e leitura documentária ao suporte tradicional, mas pode ser que as estratégias descritas neste trabalho venham a se adaptar ao meio eletrônico, porém enquanto não verificamos essa adaptação recomendamos sua aplicação ao suporte tradicional. (papel)

Com o intuito de verificar algumas concepções de leitura e de estratégias de leitura, enfatizaremos primeiramente o que é a atividade de leitura e quais as suas características; consequentemente abordaremos o que o leitor procura na leitura, quais as formas e os tipos. Relataremos em seqüência que a leitura é classificada como uma atividade interativa envolvendo as três variáveis texto, leitor e contexto envolvidas no processo de compreensão .

Em seguida, verificaremos o papel da memória na leitura destacando como está organizado o conhecimento na memória, quais os processos de leitura, os principais modelos de leitura e as estratégias de leitura a que o leitor deve recorrer para compreender o texto, bem como a utilização do conhecimento de estrutura textual como estratégia de leitura e, finalizamos, caracterizando a leitura documentária realizada por um indexador.

A leitura é uma atividade que exige do leitor olhar atento sobre as páginas do texto e o que for retido nesse momento da chamada “olhadela” por todo o texto é simultaneamente relacionado com os conhecimentos que o leitor possui armazenados na memória.

Portanto, definimos a leitura como uma atividade interativa que envolve “*pensamento e criatividade*” em busca da compreensão do texto.

Segundo Smith (1989), a leitura apresenta as seguintes características: objetiva, seletiva, antecipatória e flexível. (SMITH, 1989)

A leitura é objetiva porque o leitor tem o objetivo de compreender o texto, e a compreensão pode envolver tanto a apreensão do significado do que o texto aborda de forma global, como também a identificação do assunto principal do texto; é seletiva porque na busca do assunto principal, o leitor deixa de lado as informações que não lhe interessam saltando, assim, os trechos considerados irrelevantes para ele; é antecipatória porque o leitor, com base no seu conhecimento anterior sobre o assunto tratado no texto, antecipa o que o autor abordará nos parágrafos seguintes sobre determinado assunto; é flexível porque pode ser controlada pelo leitor; este tem o poder de determinar o momento de parar, ou de voltar a trechos já lidos.

Mas o que o leitor procura na leitura? Para Van Dijk (1978) o leitor nela procura a informação importante, e essa pode variar de um leitor para outro.

Dessa forma, podemos considerar que existem duas categorias de informação importantes:

1. Informação textualmente importante: porque o autor a apresenta como importante;
2. Informação contextualmente importante: porque o leitor a considera como tal devido a sua intenção de leitura.

A visão de Cavalcanti (1989), considerando a leitura como um processo comunicativo que envolve leitor e texto, distingue saliência textual de relevância-leitor. Evidenciando a saliência textual, podemos mencionar que essa caracteriza-se como as idéias que o autor destaca no texto (as mais relevantes para ele); já com relação à relevância-leitor, nós mencionamos as idéias que o leitor seleciona como relevantes para interagir com o texto durante a leitura.

Para Nardi (1999, p.24), interpretando a visão de Cavalcanti (op. Cit) sobre a atividade de leitura ressalta-se que

“nem sempre o que o autor salienta no texto é o que o leitor considera importante. A relevância não se restringe ao texto, mas emerge da relação entre os aspectos salientes do texto e o conhecimento prévio do leitor.”

Elucidando o que venha a ser informação importante destacamos que pode ser tanto a idéia principal do texto que foi explorada com mais ênfase pelo autor dentro do contexto de determinado assunto narrado, como também, pode ser o assunto geral do texto, ou seja, o que o texto trata de forma global. (GIASSON, 1993).

Entretanto, a idéia principal e o assunto global do texto serão realmente informação importante desde que satisfaça os anseios de informação de determinado leitor.

Acrescentamos aqui que, dependendo da informação importante que o leitor procura para atingir um objetivo fim, ele poderá utilizar diferentes tipos de leitura. Segundo Medeiros (1997) os tipos de leitura são:

- Skimming: o leitor procura captar a tendência geral do texto, fazendo uma leitura superficial de títulos, subtítulos e de alguns parágrafos;
- Do significado: procura obter uma visão geral do texto, fazendo uma leitura rápida, sem se deter, nem retomar parágrafos anteriores já lidos;
- De estudo: engloba ler, reler, anotar e resumir;
- Crítica: envolve reflexão, avaliação, comparação com o que se leu anteriormente;
- Scanning: caracteriza-se como aquela pela qual se procura certo tópico de uma obra.

As formas de leitura que o leitor utiliza pode influenciar o desempenho de compreensão do leitor frente ao texto lido.

De acordo com Kato (1986) temos às seguintes formas de leitura:

- Leitura em voz alta: realizada quando nos deparamos com trechos do texto de difícil compreensão;
- leitura silenciosa: considerada como resposta-estímulo que é associada a um significado;
- leitura subvocalizada: considerada como lenta, pois obtém dupla decodificação: letras em sons e sons em significados, impede impor rapidez à leitura;

Ressaltamos também que existem estudos na área de Lingüística e outras áreas que trabalham com o texto, para identificação de estrutura textual e das tipologias dos documentos, isto é, as partes que compõem o texto e as informações pertencentes a cada parte.

Por exemplo, a tipologia do resumo é composta das partes: introdução, objetivo, justificativa, metodologia, resultado e conclusão.

Os resultados das pesquisas afirmam que nos textos existem partes principais que são mais significativas, e que se o leitor conhecer cada parte dessas estruturas e utilizá-las, durante a leitura, encontrará a informação importante que procura realizando uma leitura mais rápida.

Enfocamos até o momento o que é a leitura e suas características, o que o leitor nela procura, suas formas, seus tipos, existentes de leitura e concluímos que a leitura é uma atividade interativa e portanto comunicativa exigindo uma espécie de cooperação entre os interlocutores. E segundo Grice (1975), citada por Ginez de Lara (1993, p.46-47), esse princípio de cooperação reporta-se a regras dispostas sob quatro categorias:

- 1- Quantidade: relativa ao volume de informações que deve ser fornecida pelo locutor ao seu ouvinte;*
- 2-Qualidade: refere-se à veracidade das informações;*
- 3-Relação: diz respeito ao propósito da informação;*
- 4-Modalidade: refere-se à forma pela qual se apresenta o dito, incluindo clareza, ausência de obscuridade, ausência de ambiguidade, brevidade e método.*

Na opinião de Ginez de Lara (1993), a eficiência na comunicação incluindo aqui a leitura por estabelecer uma atividade de comunicação, está relacionada à observação ou não dos interlocutores (autor e leitor) das máximas conversacionais de Grice (Op. Cit) descritas acima. E concordamos com isso porque, o autor, quando se propõe a redigir um texto e se preocupa em fornecer o máximo de informações para que o seu leitor alvo compreenda as suas idéias originais, e tem o cuidado de remeter, no texto, às fontes fidedignas sobre as quais formulou os seus argumentos, e discorre de forma coerente, sem contradições ou informações implícitas que possam causar dupla interpretação, com certeza ele atingirá o seu propósito de transmitir informação. E quando o leitor começa a observar a presença ou não dessas máximas de Grice no texto poderá por exemplo: avaliar se o texto realmente se destina ao seu perfil enquanto leitor? Se será necessário recorrer a outras fontes de informação para buscar compreender aquelas informações que o texto omite por não prever um outro perfil de leitor diferente do idealizado pelo autor?.

Encontramos também na revisão da literatura sobre leitura um consenso entre os pesquisadores que a leitura é interativa porque se concretiza por meio de um modelo de

compreensão que envolve intrinsecamente três variáveis: o texto, o leitor e o contexto. (GIASSON, 1993).

A primeira variável a ser abordada deve ser o leitor, variável classificada como a mais complexa do modelo de compreensão. O ponto de partida é que o leitor, durante a leitura, em busca da compreensão, utiliza seus conhecimentos sobre a língua na qual se encontra o texto e, obviamente, conhecimento sobre o mundo.

O conhecimento de mundo é o ingrediente mais importante; diz respeito aos conhecimentos adquiridos anteriormente e que se encontram armazenados na memória. Dentre esses conhecimentos, por exemplo, pode existir conhecimento sobre o assunto abordado no texto que se está lendo. Tanto o conhecimento sobre a língua como o conhecimento de mundo referem-se às “*estruturas cognitivas*” do leitor, ou seja, às características que o leitor possui e que podem variar de um leitor para outro. Pode ocorrer que os conhecimentos referentes à língua e o conhecimento de mundo, utilizados pelo leitor para compreender um texto, não sirvam para a compreensão de um outro texto, ficando o leitor obrigado a recorrer a novos conhecimentos.

O leitor, durante a atividade de leitura, está, também, sujeito a transferir suas estruturas afetivas, isto é, suas atitudes frente ao texto a ser lido e seus interesses individuais pelo texto, influenciando, assim, o processo de compreensão.

Dentre as estruturas afetivas que o leitor possui estabelecidos por Giasson (Op. Cit) destacamos:

- atração que o leitor possui pelo texto;
- ou seu sentimento de indiferença, ou repulsa pela sua leitura;
- a capacidade de se arriscar durante a leitura em busca da compreensão;
- a sua auto-imagem em geral;
- a sua auto-imagem como leitor;
- o medo do insucesso na leitura e seus interesses específicos.

Particularmente, com relação aos seus interesses específicos, estes se desenvolvem fora do contexto de leitura e influenciam muito na compreensão do leitor, pois, em se tratando de textos de temas específicos, o leitor pode sentir uma enorme atração pelo texto se o tema abordado fizer parte dos seus interesses específicos e, conseqüentemente, realizará uma leitura interativa relacionando os seus conhecimentos específicos sobre o tema com o conhecimento que o texto traz. Mas, o leitor pode, também, sentir repulsa se o tema abordado no texto não fizer parte dos seus interesses

específicos, nesse caso, a leitura se tornará “*massante*”, “*cansativa*” e com uma grande possibilidade de não se atingir a compreensão, uma vez que o leitor provavelmente não possui conhecimentos anteriores sobre o tema do texto. (GIASSON, 1993)

Com relação à variável leitor e, após destacarmos que na leitura utilizam-se estruturas cognitivas e afetivas, ressaltamos que também se utilizam processos de leitura, isto é, “*habilidades necessárias para se abordar o texto ao desenrolar das atividades cognitivas durante a leitura*”. A importância dos processos de leitura será melhor evidenciada nas seções seguintes. (GIASSON, 1993).

A outra variável, o texto, do modelo de compreensão de leitura proposto por Giasson (Op. Cit), tem a função de transmitir a idéia do autor. Portanto, o autor é representado pelo texto. Nele, o autor aborda informações principais e secundárias e utiliza dicas tipográficas como: grifo, sublinhado, negrito, notas, tabelas, etc..., para evidenciar algo que deseja destacar no conteúdo do texto.

Todos os textos apresentam uma estrutura textual com as partes informacionais que os compõem organizadas numa sequência lógica diferenciada de uma tipologia textual para outra.

Geralmente, o autor determina o estilo de estrutura textual a ser utilizada para representar as suas idéias, com base nas suas intenções, tais como: “*agir sobre as emoções do leitor*”, o autor apresenta um texto que mexe com as emoções do leitor como: riso, choro, nostalgia, etc...; “*agir sobre os comportamentos do leitor*”, o autor apresenta um texto que com instruções que podem ser uma receita culinária ou regras de jogo; “*agir sobre os conhecimentos do leitor*”, o autor apresenta um texto que permite ao leitor comparar os seus conhecimentos anteriores com o conteúdo do texto, e com isso adquirir conhecimentos novos. As intenções do autor também estão direcionadas para o perfil do leitor que pretende atingir. (MARSHALL, 1984).

O autor, quando se propõe a redigir um texto, deve procurar usar vocabulário conhecido pela comunidade usuária, e quando usar palavras novas, esclarecer o significado dessas no texto. O autor ainda deve ficar atento aos períodos textuais com muitas orações subordinadas, compridas, e para o contrário, ou seja, uso somente de orações simples, sem termos de ligação com excessiva simplificação, que pode ser um obstáculo à compreensão. (CALFEE & CURLEY, 1997).

No modelo de compreensão de leitura elaborado por Giasson (Op. Cit) que envolve leitor, texto e contexto, o contexto refere-se a todas as condições próprias do

leitor, bem como as condições que lhe são impostas pelo meio em que vive. Podemos classificar o contexto em: psicológico, social e físico.

No contexto psicológico estão as condições próprias do leitor: grau de interesse pelo texto, sua motivação e intenção de leitura. Destacamos que as condições próprias do leitor que compõem o contexto psicológico podem influenciar na compreensão do texto, pois se não tiver intenção de leitura não saberá que informações deverá reter do conteúdo do texto para atingir a compreensão, e alcançar o objetivo da leitura. Além disso, se o leitor não estiver motivado para ler um texto em um determinado momento porque está, por exemplo, preocupado, ou com problemas particulares, isso, com certeza, afetará o desempenho do leitor na leitura do texto.

O contexto social remete para todas as formas de interação que podem ocorrer durante a leitura e que afetam diretamente o leitor. Dentre as formas de interação, destacamos como exemplo, o caso em que o leitor é obrigado a ler em voz alta perante um grupo e, posteriormente, é chamado a expor a sua compreensão sobre o que leu. Certamente o leitor compreenderá menos o texto, se comparado com a situação em que é colocado a fazer uma leitura silenciosa individual sobre o mesmo texto, e expor a sua compreensão, posteriormente. Observa-se aqui que as formas de interação podem ser referentes a situações de interação do leitor com seus pares, ou seja, outras pessoas que juntamente com o leitor formam um grupo, como por exemplo, grupo de alunos, entre aluno leitor e o professor, etc....

O contexto físico compreende todas as condições materiais em que se desenvolve a leitura como: temperatura ambiente não pode ser nem muito alta nem muito baixa e sim moderada, ou, adequada, não afetando assim as condições físicas do leitor e possibilitando-lhe a realizar a leitura num lugar agradável; qualidade de reprodução dos textos: o texto a ser lido deve ser legível possuindo assim, letras em tamanho apropriadas, uniformes, a tinta usada no texto deve ser de determinada cor que não perturbe nem canse a visão do leitor; ruído, refere-se ao barulho próximo ou no local em que está se realizando a leitura o que dificulta a concentração do leitor; a ausência de barulho é essencial para que o leitor formule a sua compreensão sobre o texto, não correndo, assim, o risco de dispersar seus pensamentos durante a sua leitura.

Elucidamos também com relação à leitura que, recentemente, existem autores como Bloome (1983), Spradley (1980), Nardi (1999), entre outros, que têm visões da atividade de leitura como evento social porque defendem que durante a leitura os

conhecimentos culturais adquiridos, seja com leituras anteriormente realizadas, conhecimento de mundo (nossa vida familiar e na sociedade a que pertencemos), são transmitidos tanto pelo autor que redigiu o texto para o leitor, como do leitor que está lendo o texto e confrontando-o com seu conhecimento cultural que adquiriu, e que a partir de então pode aceitar, refutar, ter outra interpretação da informação transmitida pelo texto e transmiti-la a outros indivíduos do grupo a que pertence, e todo esse ciclo de transmissão e aquisição de cultura, por meio da leitura, deixará registradas as idéias do autor que estão submetidas à interpretação do leitor e que vêm evidenciar os costumes, isto é, as idéias de uma sociedade contextualizando-o numa determinada época.

Nardi (1999, p. 38) evidencia essa nossa explanação quando comenta

“a meu ver, as possibilidades de interação social e cultural que se abrem numa leitura, até mesmo na que envolve um único participante e um texto, são inúmeras e imprevisíveis: o leitor pode interagir não só com o autor do texto, como também com outros leitores que tenham tornado explícitas suas interpretações anteriores do mesmo texto, com outros autores de outros textos que de alguma forma se relacionem ao texto sendo lido etc... construindo assim “mapas” inéditos de “navegação”.

Após termos evidenciado a leitura como atividade interativa que engloba um modelo de compreensão esquematizado pelas três variáveis leitor, texto e contexto, é necessário considerar que durante a leitura a memória do leitor desempenha uma função primordial na busca da compreensão do texto, porque armazena, de forma organizada, grande parte dos conhecimentos que obtemos, e quando necessário podemos consultá-la e recuperar os conhecimentos que desejamos.

Portanto, é na memória que armazenamos o nosso conhecimento sobre a língua escrita, adquirido com a alfabetização em língua materna, conhecimentos sobre outros idiomas, nosso conhecimento de mundo que engloba conhecimentos sobre os tipos de leitura, suas formas, seus processos, seus modelos, enfim conhecimentos que nos auxiliam na compreensão do texto.

Segundo Kato (1986, 48 -52), a memória se divide em: memória de curto termo, memória de médio termo e memória de longo termo. As duas primeiras armazenam informações temporárias que, por meio da memorização, são transmitidas à memória de longo termo. A de longo termo é uma terceira divisão da memória onde se armazenam

informações permanentes e que podem ser recuperadas a qualquer momento. Tanto o leitor como o autor obtêm, nessa memória, o conhecimento lingüístico e as regras gramaticais, o léxico, bem como as instruções para o uso de ambos, e tantos outros conhecimentos.

Os conhecimentos contidos na memória de longo termo são passíveis de recuperação porque estão armazenados de forma organizada. A essa organização dos conhecimentos na memória denominamos de “esquemas” que são os *“conceitos ou conhecimentos genéricos referentes a situações, acontecimentos, seqüências de acontecimentos, ações e seqüências de ações”*, que adquirimos com nossa vivência diária, por meio de leitura, conversa informal ou formal, etc... (RUMELHART, 1997).

Os esquemas possuem componentes denominados variáveis como, por exemplo; na situação compra temos: o vendedor, o objeto a ser negociado, o comprador e a forma de pagamento; portanto, essas são as variáveis da situação compra e, como são variáveis essenciais para que ocorra o desenvolvimento da situação, são denominadas de superesquemas e, nesse caso, são também variáveis fixas, pois se faltasse alguma delas a situação compra não ocorreria. (FUJITA, 1999)

Os esquemas apresentam algumas situações em que as variáveis denominadas superesquemas (essenciais) tanto quanto as variáveis subesquemas (secundárias) variam de indivíduo para indivíduo, devido às experiências próprias que cada indivíduo possui de determinada situação. Por exemplo: na situação “estudar”, alguns dizem que estudar engloba ler, reler, anotar e resumir, e outros somente ler; isto ocorre porque os indivíduos possuem esquemas diferentes para determinar a situação “estudo”, independentemente de ter o mesmo objetivo de se obter uma nota escolar. Isso ocorre porque os indivíduos em situações passadas fizeram uso de determinadas variáveis e alcançaram assim o objetivo que buscavam, e por isso defendem essas variáveis como sendo precisas ou essenciais.

Além de superesquemas, os esquemas apresentam também subesquemas que são variáveis de peso inferior, isto é, a ausência delas não afeta em nada o desenvolvimento de uma situação como por exemplo: na situação “compra”, citada acima, os subesquemas seriam: o comprador ser do sexo feminino, com olhos verdes, etc...

Mas, importante é destacar que, quando recorremos aos esquemas que estão armazenados na memória, podemos acioná-los pelos superesquemas, e quando isto ocorre, dizemos que ocorreu um processamento “Top-Down”, ou seja, do todo para as partes, ou ainda, partiu-se das variáveis superiores denominadas assim porque são essenciais, para as

inferiores ou secundárias que são classificadas assim por não serem essenciais, para o desenvolvimento de uma determinada situação. Quando o leitor aciona os subesquemas dizemos que houve um processamento “Bottom-Up”, isto é, das partes para o todo, ou seja, das variáveis secundárias para as variáveis essenciais.

Os esquemas são muito importantes na leitura pois, se o leitor possui esquemas que lhe permitam compreender a relação entre as variáveis de determinada situação enfocada pelo autor no texto, bem como o efeito que causam no conteúdo do texto, alcançará uma compreensão coerente do significado do texto. Entretanto, quando o leitor não possui esquemas para descobrir essas variáveis e suas relações, o texto parecerá desarticulado e incompreensível. (RUMELHART, 1997)

Conforme Rumelhart (1997, p.47 - 48) o leitor, muitas vezes, não compreende uma passagem ou um trecho do texto. Existem três razões implícitas na teoria dos esquemas para explicar porque isso acontece:

1- “*é possível que os leitores não tenham os esquemas apropriados, nesse caso, simplesmente não podem compreender o conceito comunicado.*”

2- “*leitores podem ter os esquemas apropriados, mas as informações previstas pelo autor podem ser insuficientes para ativá-los*”. Nesse caso, os leitores entenderiam pouco o texto, mas com informações adicionais apropriadas podem chegar a compreendê-lo;

3- os leitores podem encontrar uma “*interpretação consistente do texto*”, mas não encontram o que o autor pretendia, nesse caso os leitores compreendem o texto, mas interpretam mal o autor.

Nessa exposição de Rumelhart percebemos que o autor quando redige um texto, na maioria das vezes, inclui informações referentes aos seus esquemas próprios de conhecimentos de determinadas situações e, em muitos casos, conhecidos somente por indivíduos que tenham conhecimentos semelhantes por pertencerem a mesma classe profissional ou por terem vivenciado situações semelhantes. Porém, quando o leitor não possui esses esquemas apresenta dificuldade para compreender o texto. De outra forma, o leitor possui os esquemas, mas as informações fornecidas pelo autor são insuficientes para que ele atinja a compreensão; deve-se, então, procurar obter informações complementares através de outras leituras sobre o assunto, consulta a obras de referência, conversa com especialista no assunto, etc...

Como mencionamos anteriormente o que o leitor procura no texto é a informação que considera importante para atingir um objetivo que deseja, e para isso utiliza determinado tipo de leitura. Entretanto, para utilizar esse ou aquele tipo de leitura, o leitor

deve optar pelo melhor processo ou por um conjunto de processos de leitura que o possibilite atingir a compreensão que almeja.

Irwin (1986) considera cinco categorias de processos de leitura que se dividem em componentes:

1. Microprocessos: englobam o reconhecimento de palavras, leitura de grupos de palavras e micro – seleção, e são usados para explorar a informação contida numa frase.

a) reconhecimento de palavras:

Na leitura, muitas vezes, o conhecimento do significado de determinada palavra pode ser importante para que o leitor compreenda uma frase ou conjunto de frases que envolvam determinada palavra que seja relevante para desvendar o significado do texto ou de um tópico do texto.

b) leitura de grupos de palavras:

No que se refere à leitura de grupos de palavras, o leitor deve saber agrupar as palavras em unidades significativas mesmo que as compreenda individualmente, pois, durante a leitura, as informações que extraímos são retidas na memória de curto prazo até serem tratadas e transferidas para a memória de longo prazo. A memória de curto prazo somente retém quatro ou cinco elementos de cada vez para serem tratados em unidades de sentido e transportados para a memória de longo prazo; se forem enviados à memória de curto prazo novas informações antes das anteriormente enviadas serem tratadas, as últimas informações se perderão. Por isso, o leitor, durante a leitura, deve agrupar as palavras porque a informação será tratada mais rapidamente na memória de curto prazo, e a procura por significação será facilitada.

c) micro-seleção:

O leitor sempre está atualizando o seu vocabulário a cada nova leitura, com isso, tem facilidade em reconhecer as palavras no texto, podendo decidir que informação considera importante numa frase e, posteriormente, retê-la na memória de curto prazo. Esse processo deverá repetir-se em todas as frases do texto, procurando sempre integrar a informação nova à antiga. A informação retida é considerada importante tanto na frase como no contexto do conjunto de frases que compõem o texto.

2- Processos de integração: tem a função de efetuar ligações entre as proposições ou as frases utilizando referentes, conectores e inferências baseadas em esquemas, com o objetivo de assegurar a coesão do texto. Os processos de integração são metacognitivos, ou seja, são processos conscientes do leitor. O leitor deve conhecer esses elementos de ligação, uma vez que o autor nunca descreve uma situação no seus mínimos detalhes.

a) referentes

Os referentes compreendem marcas explícitas entre as frases e podem ser uma palavra ou expressão utilizada para substituir outra, como: pronome, sinônimo, termo genérico ou, em muitos casos serem subentendidos.

b) conectores

São palavras utilizadas na ligação de duas proposições ou duas frases.

Os conectores se classificam segundo Irwin (1986) em:

Quadro1: Classificação dos Conectores

Conjunção: e, também	Contraste: contrariamente a
Disjunção: ou	Oposição: apesar de, embora
Exclusão: excepto, com excepção de	Concessão: se bem que
Tempo: antes, logo que	Consequência: de maneira a , a tal ponto que
Lugar: diante de, por cima de	Fim: para, a fim de
Causa: porque, por causa de	Condição: se, a menos que
	Modo: como

No texto, podem aparecer de forma explícita ou implícita e, muitas vezes, serem omitidos, estarem muito afastados ou serem imprecisos.

Para compreender os conectores implícitos ou afastados no texto, o leitor pode se apoiar nas inferências baseadas em esquemas para inferir o significado desses conectores e atingir a compreensão do texto. Esse tipo de inferência é denominada de “pragmática”, por ser baseada no conhecimento do leitor. O leitor pode, também, inferir determinado

significado com base em dicas ou pistas fornecidas pelo texto, denominadas “inferências lógicas”. Quando o conhecimento do leitor, comum a um certo grupo de leitores, é utilizado para inferir significados durante a leitura, denominamos “inferências criativas”. (CUNNINGHAM, 1987).

3- Macroprocessos: orientam-se para a compreensão global do texto e compreendem a identificação das idéias principais, resumo e utilização da estrutura do texto.

a) identificação das idéias principais

Durante a leitura, o leitor procura identificar a idéia principal que pode ser uma mensagem do autor, ponto de vista principal ou idéia central do texto. Entretanto, a determinação da idéia principal depende do objetivo da leitura do leitor, isto é, do que ele procura na leitura do texto. Esse objetivo poderá ser a informação textualmente importante ou a informação contextualmente importante.

A idéia principal deve variar de acordo com a estrutura textual como, por exemplo: num texto narrativo a idéia principal pode ser um acontecimento ou a sua interpretação, já num texto informativo pode ser uma regra, um conceito, ou uma generalização. Quando a idéia principal aparece implícita, o leitor deve inferi-la com base nas informações fornecidas pelo texto e nos seus conhecimentos anteriores sobre o assunto.

b) resumo

No decorrer da leitura, o leitor extrai informações que lhe permitam construir a representação global do texto, isto é, o resumo do texto, descartando assim as informações redundantes ou secundárias.

E quando é possível fazer um resumo do texto, o leitor procura recontar mentalmente a si próprio ou a outro o que o texto aborda no seu conjunto, pois o reconto é uma forma de o leitor avaliar a sua compreensão e reestruturar o texto.

c) utilização da estrutura do texto

O conhecimento de estruturas textuais possibilitará ao leitor identificar a parte do texto que traz a idéia principal ajudando, assim, a compreender de forma global o texto e a realizar uma leitura mais rápida, pois já conhece as partes que tem a explorar e os conceitos pertencentes a cada parte. O leitor poderá explorar a estrutura textual por meio

do questionamento, perguntando, por exemplo, na exploração de um texto científico: qual o problema que o texto trata? Qual o objetivo? etc....

4- Processos de elaboração: permitem ao leitor ir além do texto, efetuar inferências não realizadas pelo autor. Os processos de elaboração são: fazer previsões, construir imagens mentais, resposta afetiva (reagir emotivamente), ligação com os conhecimentos para integrar informações antigas (conhecimento de mundo) às abordadas no texto, raciocínio para compreender o texto como um todo, e ser consciente do objetivo que pretende alcançar na leitura.

a) previsões

O leitor elabora previsões sobre o que poderá encontrar no texto apoiando-se na leitura do título ou nas subdivisões do capítulo e baseando-se em seus conhecimentos de mundo, que englobam a sua experiência em leitura e situações vividas no seu dia-a-dia para, por meio do raciocínio, compreender o que o autor quis dizer em cada parte do texto, e de forma geral.

As previsões em leitura permitem ao leitor autocorrigir um pensamento, uma vez que as previsões não confirmadas no decorrer do texto podem ser corrigidas ou modificadas e, conseqüentemente, reformuladas.

Portanto, a previsão tem a função de aumentar a motivação e envolvimento do leitor com o texto efetuando uma melhora na compreensão.

b) construir imagens mentais

As imagens mentais em leitura estão relacionadas à dimensão visual e exige do leitor capacidade de reunir grupos de palavras enquanto lê para facilitar o envio das informações à memória e, conseqüentemente, obter um significado da imagem (informações) enviada.

c) resposta afetiva

As respostas afetivas dizem respeito ao envolvimento do leitor com o texto. O leitor que se envolve emotivamente na leitura “tem mais possibilidade de compreender e de reter as informações do texto”. (IRWIN, 1986).

Ressalta-se, porém, que as respostas afetivas nem sempre são apropriadas a todos os tipos de textos, pois existem textos em que o autor “usa” os sentimentos do leitor para

persuadi-lo e, nesse caso, o leitor deve avaliar o texto com base no seu raciocínio sobre o texto.

d) raciocínio: ligado à leitura crítica

Portanto, na leitura, os processos de elaboração permitem que o leitor utilize o raciocínio para distinguir, no texto, opiniões de fatos, evitando assim que manipulem a sua compreensão do texto. Isso pode ser evitado, também, por meio da verificação da fonte onde prevê a informação abordada no texto. O leitor utiliza, também, o raciocínio para estabelecer relações entre o texto e os seus conhecimentos pessoais necessários à compreensão.

e) ligação com os conhecimentos

Os macroprocessos e os processos de elaboração são considerados processos metacognitivos, em que o leitor realiza consciente de que busca um determinado fim.

5- Processos metacognitivos: geram a compreensão e permitem ao leitor adaptar-se ao texto e à situação. Os processos metacognitivos são: identificação da perda da compreensão e reparação da perda de compreensão, ou “monitoração da compreensão”. (BROWN, 1980).

Os processos metacognitivos exigem que, durante a leitura, o leitor interogue se está compreendendo o texto e, quando percebe que algo lhe passou despercebido, ocasionando falhas na compreensão que o impossibilitam de atingir o seu objetivo de leitura, volte a trechos do texto já lidos e realize uma releitura para tentar compreender o texto. Portanto, os processos metacognitivos são realizados conscientemente pelo leitor e este deve ser estratégico na resolução das falhas de compreensão para saber escolher qual atividade (estratégia) melhor o ajudará a resolver tal problema.

Conforme Brown (1980, p. 354), os processos metacognitivos exigem que o leitor auto-avalie a sua compreensão por meio de quatro processos de gestão da compreensão:

- 1- saber quando compreendemos e quando não compreendemos;
- 2- saber o que compreendemos e o que não compreendemos;
- 3- saber de que precisamos para compreender;
- 4- saber que podemos fazer alguma coisa quando não compreendemos.

Para sintetizar o enfoque dado aos processos de leitura, as autoras Kato (1986) e Kleiman (1989) ressaltam que quando lemos devemos realizar os seguintes processos:

- procurar obter uma visão do texto como um todo, buscando através do correr dos olhos o que nos interessa;
- procurar incorporar rapidez a leitura saltando trechos do texto que são irrelevantes ou previsíveis;
- partir da compreensão dos termos mais genéricos para os mais específicos;
- formular hipóteses baseados no nosso conhecimento sobre o assunto que serão, através do processo de leitura, confirmadas ou refutadas;
- utilizar nosso conhecimento de mundo ou conhecimento global baseado na nossa experiência de vida ou com base no próprio texto para imprimirmos sentido ao texto;
- procurar identificar o tipo de estrutura textual;
- ter um objetivo na leitura, pois melhora a capacidade de processamento e memória;
- realizar leitura seletiva: busca de informações através de tabelas, resumos, primeiro e último período do parágrafo, título e subtítulo quando o título coincide com o tópico, mas leva à falha na compreensão quando não existe tal coincidência;
- identificar a hierarquia das informações no texto, entre informações principais e periféricas.

A sequência de processos que o leitor utiliza durante a leitura gera os mais variados modelos de leitura. De acordo com Galvez & Pinto (s.d.), os modelos de leitura constituem-se nos resultados dos processos de leitura. Encontramos na revisão da literatura os Modelos Seriais, Modelos em Paralelo, Modelos Ascendentes e Descendentes, Modelo Interativo, Modelo de Concepção Modular e Modelo de Concepção Global.

Os Modelos Seriais obedecem a uma estrutura hierárquica e a leitura discorrerá da seguinte forma: o leitor começará a ler partindo de reconhecimento de rasgos visuais, reconhecimento de letras e palavras, reconhecimento de orações, interpretação semântica e representação do significado do texto.

Como se pode perceber, nesse modelo, o leitor começa a ler obedecendo a uma ordem crescente iniciando-se com os microprocessos, depois utiliza os processos de integração indo até aos macroprocessos para obter o significado do texto. A revisão da

literatura demonstrou que Forster (1979), Garrett (1975) e Mitchell (1952) trabalharam com esse modelo.

Nos modelos em paralelos, o leitor utiliza, durante a leitura, todos os processos necessários para a sua compreensão, desde os microprocessos até os processos metacognitivos, não obedecendo a uma ordem seqüencial crescente e nem a uma ordem decrescente. Trabalharam com os modelos em paralelos Marslen-Wilson (1975), Dell & Reich (1981) e Stemberger (1985).

Com o modelo ascendente, conhecido como processo “bottom-up”, o processamento ocorre de baixo para cima, iniciando com a visão ortográfica do texto, depois o reconhecimento de palavras, de frases, conectores que ligam as frases e os parágrafos, obedecendo, portanto, a uma ordem crescente, ou seja, partindo de processos inferiores até chegar, a processos mais elevados que exigem um nível maior de conhecimento. Gough (1972) e Laberge & Samuels (1974) trabalharam com o modelo ascendente.

De outro modo, no modelo descendente, conhecido como “top-down”, o processo ocorre de cima para baixo, isto é, o leitor utiliza, na leitura, seus conhecimentos de mundo, inferências, formulação de hipóteses ou seja, conhecimentos referentes aos processos que exigem maior esforço mental, para depois utilizar processos inferiores como: reconhecimento de letras, palavras, etc... que exigem esforço cognitivo menor, para com isso obter a compreensão do texto. Portanto, a seqüência de processos seguida pelo leitor obedece a uma ordem decrescente partindo de processos metacognitivos (conscientes) até chegar aos microprocessos. Foi Goodman (1971) o teórico que trabalhou com o modelo descendente.

Com o modelo interativo, o fluxo da informação pode ir de cima para baixo (top-down) e de baixo para cima (bottom-up), utilizando todos os processos necessários para atingir a compreensão sem obedecer a uma ordem de seqüência crescente nem decrescente, mostrando-se semelhante ao modelo em paralelo. Rumelhart (1977), De Vegga (1990) trabalharam os modelos interativos.

No modelo de concepção modular, o processo seguido pelo leitor vai em ordem crescente partindo dos microprocessos, isto é, reconhecimento dos rasgos visuais para as letras, destas para as sílabas, depois para o léxico e assim sucessivamente. Nesse modelo, o

leitor lê em função do texto, exigindo dele um escasso esforço cognitivo. O teórico que trabalhou esse modelo de concepção modular foi Fodor (1983).

No modelo de concepção global, a leitura consiste num sistema global e unitário, onde intervêm todos os atributos da mente humana: percepção, atenção, memória e pensamento. É, portanto, um modelo complexo porque o leitor realiza uma leitura consciente do objetivo que pretende alcançar e dos processos que está utilizando para atingir esse objetivo. Anderson (1983) trabalhou esse modelo.

Destacamos que o uso de determinado(s) processo (s) de leitura resultará num determinado modelo de leitura que vem a representar a seqüência de processos que o leitor utilizou durante a leitura. Portanto, o leitor pode escolher melhor os processos que lhe serão úteis, o que depende do conhecimento prévio do leitor sobre a língua em que se encontra o texto, da consciência do seu objetivo de leitura, da escolha do momento certo para usar determinado processo e, acima de tudo, dos “esquemas” que o leitor possui armazenados em sua memória para que possa recuperá-los quando necessário e utilizá-los.

Até aqui, verificamos os seguintes componentes envolvidos na atividade de leitura que funcionam como auxiliares para se atingir a compreensão: a informação que o leitor procura, isto é, o objetivo de leitura; o papel da memória na leitura e sua organização privilegiada que beneficia a recuperação do conhecimento; os tipos de leitura que refletem as informações que o leitor procura no texto; os processos de leitura, ou seja, conhecimentos tanto sobre a língua em que o texto se encontra como o conhecimento de mundo do leitor, que geram os modelos de leitura, uma seqüência dos processos utilizados pelo leitores durante a leitura. Esperamos que este texto tenha contribuído para resgatar a importância desses componentes envolvidos na compreensão da leitura, uma atividade complexa que não pode ser tratada isoladamente.

Destacamos, na seqüência, as atividades que o leitor pode realizar para atingir a complementação desse (s) ou daquele (s) processo (s) de leitura em busca da compreensão do texto. Essas atividades são denominadas de estratégias de leitura.

Na busca pela compreensão do texto, o leitor faz uso de atividades que lhe permitem compreender o texto, sendo essas atividades denominadas de estratégias e “skills”.

Conforme Nardi (1993, p. 20), alguns autores fazem distinção entre estratégias e habilidades automáticas (skills):

- Skills: seria uma estratégia que teria sido adquirida em algum momento da aprendizagem e se tornado automática (raramente “recuperada” pelo leitor proficiente);
- Estratégia: seria o uso consciente de uma “skill” ou de uma nova “tática”, em momentos de solução de problemas.

Vários autores expressam sua opinião sobre o que se caracteriza como estratégia:

Para Faerch & Kasper (1987), estratégias são atividades utilizadas na resolução de um problema que pode estar impedindo o alcance da compreensão do texto.

Na opinião de Brown (1980, p. 465), “estratégias são qualquer controle deliberado e planejado de atividades que levam à compreensão.”

Para Oxford (1989), citada por Nardi (1993), as estratégias de leitura são ações utilizadas para um objetivo/fim, que podem ser observáveis, ensináveis e flexíveis, pois podem ser controladas pelo leitor, ou seja, o leitor pode selecionar e combinar a (s) estratégia (s) que quiser.

Entretanto, Kato (1986) distingue dois tipos de estratégias que utilizamos durante a leitura:

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

Na opinião de Kato (Op. Cit) as estratégias cognitivas são regidas por dois princípios básicos:

- Princípio da Canonicidade (ou da ordem natural sintática e semântica), cujo conhecimento permite ao leitor predizer, por exemplo, a categoria gramatical de uma palavra desconhecida, facilitando a inferência de seu significado;
- Princípio da Coerência: busca da coerência no texto, um dos princípios de Grice (1975).

A busca pela coerência ocorre em três níveis: global, local e temático. Buscar a coerência global significa buscar as intenções do autor com a produção do texto em questão. Buscar a coerência local relaciona-se com tentar fazer sentido de trechos

específicos (por exemplo: predizendo o que foi dito até o trecho em questão, ou estabelecer um paralelo, um contraste...). A coerência temática relaciona-se ao uso da informação-tópico do texto para fazer sentido de trechos. Na busca da coerência temática, manter em mente o título do texto é interessante.

Para melhor caracterizar o uso de estratégias metacognitivas, Brown (1980, 354) identifica aspectos que o leitor utiliza durante a leitura:

- *Explicitação dos objetivos da leitura;*
- *Identificação de aspectos importantes da mensagem;*
- *Alocamento de atenção a áreas importantes;*
- *Monitoração do comportamento para ver se está ocorrendo compreensão;*
- *Engajamento em revisão e auto-indagação para ver se o objetivo está sendo atingido;*
- *Tomada de ações corretivas quando são detectadas falhas na compreensão;*
- *Recobramento de atenção quando a mente se distrai ou faz digressões.*

Evidenciamos que a capacidade para usar estratégias metacognitivas depende segundo (Leffa, 1996, p.64) de alguns fatores como:

- Desenvolve-se com a idade cronológica do leitor;
- Correlaciona-se com a proficiência em leitura (leitores fluentes têm mais consciência de seus comportamentos de leitura, possuem mais flexibilidade para ajustar a leitura a objetivos específicos);
- O comportamento metacognitivo melhora com a instrução (o treinamento específico das habilidades metacognitivas tem feito o aluno responder de modo mais eficaz a mensagens ambíguas do falante)

Oxford (Op. Cit) afirma que as estratégias não podem ser prontamente observáveis, pois somente é possível a observação de ações comportamentais do leitor durante a leitura como por exemplo: o virar de páginas, ou a procura de uma palavra no dicionário. Entretanto, as ações mentais como associações e deduções, durante a leitura, não podem ser vistas.

Portanto, o fato de as estratégias não serem observáveis durante a leitura causa contradição entre os pesquisadores sobre o grau de consciência das mesmas. Pois alguns autores como Brown (1980) acreditam que as ações são intencionalmente selecionadas,

enquanto outros, como Faerch e Kasper (1987) referem-se a ações potencialmente conscientes.

Mesmo que as estratégias não possam ser observáveis durante a leitura, podem ser verbalizadas através da aplicação do Protocolo Verbal durante a leitura, pois esse método permite que o indivíduo expresse o seu pensamento enquanto lê em voz alta e o pesquisador grava em audio para posterior análise. O Protocolo Verbal foi utilizado por: Nardi (1993), Fujita (1998 e 1999), Fagundes (1997) e Rubi (1999) em leitura para fins documentários.

De acordo com Cunningham (1987), uma estratégia importante é também a inferência, na atividade de adivinhar que informações o texto irá nos trazer com base no nosso conhecimento de mundo sobre o assunto tratado no texto (inferências pragmáticas), ou através das informações contidas no próprio texto (inferências lógicas), e ainda por meio de inferências criativas restritas a certos grupos de leitores que possuem conhecimento sobre um assunto qualquer, pois “quanto mais conhecimento sobre um assunto possuir um leitor, mais possibilidades terá de fazer inferências criativas”.

Em síntese, o leitor poderá recorrer, durante a leitura a estratégias que o auxiliam na busca pela compreensão, como:

- Ler marcas tipográficas como: palavras grifadas, sublinhadas, tabelas, notas de rodapé e subdivisões do texto;
- Formular hipóteses com base no seu conhecimento de mundo referente ao assunto tratado no texto;
- Consultar o dicionário para procurar significados de palavras desconhecidas que são relevantes para a compreensão do texto;
- Fazer inferências sobre o que o texto abordará na sua sequência;
- Utilizar seu conhecimento sobre a língua na qual o texto se encontra bem como os processos de leitura que citamos no capítulo sobre concepções de leitura;
- Fazer uso da intertextualidade: comparar as informações obtidas num determinado texto com as informações de outros textos que leu referente a um determinado assunto;
- Recorrer à memória de longo prazo para verificar se possui esquemas referentes a determinados assuntos ou situações;

- Utilizar conhecimento prévio sobre o assunto para prever as informações que serão abordadas na sequência do texto.

Na sequência, abordaremos as características da atividade de leitura documentária estudada pela área de Análise Documentária nessa abordagem poderá se verificar que trata-se de um tipo de leitura que difere-se da leitura global porque tem objetivo definido e exige outros procedimentos, mas ressalta-se que em ambas a compreensão é fundamental para atingir determinado objetivo, quer seja a compreensão visando adquirir informação, realizar uma análise crítica e, ou, ainda indexar, entre outros.

1.1 LEITURA DOCUMENTÁRIA

Como relatamos acima, a leitura documentária difere da leitura global exigindo, portanto, outros procedimentos.

Destacamos que a leitura documentária é fase inicial da operação de análise documentária que se constitui basicamente das fases: leitura do texto e, nesse 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 assunto e resumos.

A compreensão do texto pelo indexador, segundo Farrow (1991, p. 151), acontece da mesma maneira que os leitores fluentes compreendem. Mas com quatro aspectos diferentes, como:

- 1- *Indexador tem a pressão do tempo para indexar os documentos, o que faz com que leia o documento rapidamente ao invés de ler normalmente, ou extensivamente;*
- 2- *Muitos indexadores compreendem o texto apenas com o propósito de classificar, indexar ou resumir documento;*
- 3- *A compreensão de texto é seguida diretamente pela produção de um resumo, conjunto de índices, ou classificação;*
- 4- *Muitos indexadores trabalham dentro de um âmbito estreito de tipos de textos e campos de assunto e o elemento repetitivo consequentemente no trabalho deles conduz ao processamento automático (...)* (Tradução Livre)

Durante a leitura documentária, o leitor, denominado de indexador, pode utilizar o seu conhecimento prévio sobre estrutura textual e obter conhecimento do assunto

avancando na sua compreensão, e proporcionando também velocidade à leitura, não precisando ler palavra por palavra, letra por letra.

Como parâmetro auxiliar na leitura documentária, a Norma ISO (5963,1985) recomenda o exame do documento como metodologia para a leitura do documento.

Nesse exame do documento, “trata a operação como busca objetiva de informação precisa em pouco 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.

Devemos destacar, porém, que essa metodologia para leitura do documento da Norma ISO, é direcionada para o texto técnico científico, apesar da Norma ISO não mencionar isso.

O processamento do texto pelo leitor normal e pelo indexador, conforme Farrow (1991), pode ser de dois níveis: perceptivo e conceitual - no nível perceptivo, os dois tipos de leitores buscam a compreensão por meio da procura de dicas como: *palavras longas, palavras em itálico*, utilizando seu *conhecimento de estrutura textual do texto* para fixar atenção aos trechos relevantes. E ainda em nível perceptivo, encontramos as *dicas verbais* que são: *frequência da palavra*, ou seja, a observação da ocorrência repetitiva de uma palavra ou frase é um fator importante para se considerar, para fins de indexação; *redes semânticas: a interação entre o processo conceitual e perceptivo* é tanta que, uma vez registrada uma palavra, o indexador vai ficar atento à ocorrência de palavras semanticamente relacionadas; e *características estruturais*: os resumidores têm na cabeça um conjunto de frases que apontam as características estruturais de um texto, isto é, estrutura textual e seus componentes. E quanto ao processamento, em nível conceitual, ele engloba por exemplo: *seleção de uma frase do texto para ler completamente ou superficialmente* e pode utilizar, ou não a informação contida na frase para produzir uma *representação coerente do texto todo*. E os leitores que realizam a leitura rapidamente, denominados por alguns pesquisadores como *estratégicos*, ou *proficientes* utilizam, na

leitura, habilidades, ou estratégias mais conceituais do que perceptivas; como exemplo, podemos citar a habilidade de fazer inferências que é possível quando já se *conhece o assunto e o material*, ou documento com que trabalha, isto é, a estrutura do texto. Portanto, destacam-se aqui a necessidade dos indexadores terem conhecimento específico dos assuntos dos documentos a indexar e da estrutura textual desses documentos.

Segundo Farrow (1995, p. 243), os indexadores mantêm armazenados na memória uma gama de conceitos que engloba: *o assunto a ser indexado, a estrutura dos textos a serem escaneados, os sistemas que estão usando, os usuários destes sistemas e um background de conhecimento geral.*

A aquisição da memória habilidosa está baseada em três princípios: *uso de associações significantes para codificar o conhecimento na memória de longo termo, uso de dicas de recuperação e a importância da prática em aumentar a velocidade de operações de codificar e recuperar.* Portanto, o conhecimento conceitual de um indexador leva tempo para se desenvolver. (Farrow, 1995)

Na leitura documentária, o modelo de leitura (seqüência dos processos de leitura) seguido pelo indexador, obedece tanto a movimentos “botton-up”, lendo na dependência estrita do contexto, e extraindo linearmente dos símbolos impressos o significado, como ao movimento “top-down”, em que há maior dependência do conhecimento prévio do leitor, ao fazer generalizações e predições a partir de “esquemas” que tem armazenados em sua memória (processamento conceitual).

Ao caminhar do todo para as partes (“top-down”), vai confirmando, ou não, suas hipóteses sobre o texto e, nesse movimento, o leitor formula hipóteses decisivas para a compreensão, além de fortalecer a própria organização do conhecimento individual.

Enfatizando a comparação entre os modelos de leitura, citados anteriormente, destacamos que o modelo interativo é o que melhor representa a atuação do indexador. Pois, quando o assunto do texto é desconhecido, ou que tenha informações insuficientes, o indexador procederá a leitura do documento partindo dos níveis mais inferiores aos superiores, realizando assim, uma leitura lenta do tipo ascendente ou “botton-up”; quando o texto for familiar, realizará uma leitura descendente ou “top-down” antecipando informação e dirigindo, se possível, a uma representação do conteúdo global do texto. O modelo de concepção modular também é aceito para fins de leitura documentária porque

possui os processos de leitura que o indexador realiza durante a indexação do documento. (Galvez & Pinto, s.d).

Para atingir a compreensão, o indexador utiliza os processos de leitura sugeridos pela Norma ISO (5963 de 1985) e, para concluir esses processos, apoia-se segundo (CINTRA, 1987) nas seguintes estratégias:

- durante a leitura não esquecer sua finalidade e compreensão;
- verificar as partes relevantes do texto;
- dar ênfase ao conteúdo principal, especialmente sobre o tema;
- realizar a divisão do texto com base na sequência das informações principais;
- utilizar-se dos “esquemas”, ou seja, conhecimentos armazenados na memória para buscar a compreensão e com isso adquirir novos conhecimentos;
- utilizar conhecimento prévio sobre o assunto do texto o que resulta na compreensão e imprime maior velocidade à leitura;
- fazer uso das estruturas textuais, pois para captar com mais facilidade as idéias centrais do texto;
- deve utilizar estratégias cognitivas e metacognitivas

Conforme Farrown (1995), o leitor indexador deve ainda realizar a análise, ou escaneamento do texto na busca da informação da seguinte forma:

- Estratégias de nível perceptual
 - observar palavras que apresentam mais letras que as que apresentam menos letras;
 - observar palavras não comuns;
 - observar os efeitos visuais do texto (ilustrações, quadro, palavra em itálico, etc..)
 - observar definições que aparecem no texto.
- Estratégias de nível conceitual
 - indexadores podem aplicar seu conhecimento de estrutura textual para procurar áreas de objetivo significante para o texto;
 - em relação às palavras, a semântica é uma dica importante;

- se uma declaração importante é exemplificada, o leitor pode decidir lê-la cuidadosamente;
- quando o leitor lê uma informação que está relacionada ao principal, pode selecionar proposições chaves para ajudar a formar macro-estruturas coerentes da nova informação processada;
- pode fazer inferências para ajudar a ligar as proposições.

De acordo com Galvez & Pinto (s.d), o analista deve usar durante a leitura documentária a estratégia de inferências e se o

“leitor tiver uma ampla formação cultural e conhecimentos específicos na área temática dos documentos a tratar ou analisar, esses conhecimentos o ajudarão a inferir melhor e completar os “vazios temáticos”, que podem ser uma informação que se encontra implícita que pode ser imprescindível para uma correta interpretação do texto. “ (Bernier , 1969 citado por Galvez & Pinto, s.d)

Segundo as autoras (Galvez & Pinto, s.d) acima citadas, as funções das inferências são:

- unir o tema ou tópico aos temas parciais ou subtópicos completando as informações implícitas. São mecanismos de coesão.
- eliminar a ambigüidade contribuindo para que as orações ou proposições e as seqüências de proposições não tenham distintas interpretações;
- ajudar o estabelecimento do contexto lingüístico quando as circunstâncias material, situacional ou as coordenadas espaço-tempo do texto permanecerem ausentes.

Para utilizar a estratégia de inferência é preciso ter conhecimento das estruturas textuais dos diversos textos, pois isto permite “ predizer e antecipar a organização da informação e seu objetivo.”

Segundo Cintra (1987), no processo de leitura podem-se utilizar várias estratégias como já vimos, mas a seleção dessas depende:

- da finalidade da leitura
- da experiência do leitor
- da sua maturidade frente à tarefa de ler
- do tipo de texto que está lendo

- da atenção mais concentrada em partes do texto
- do grau de novidade do texto e até mesmo da motivação para ler

Portanto, evidenciamos que as estratégias a serem utilizadas pelo indexador são selecionadas de acordo com o tipo de texto, levando em consideração o grau de conhecimento do leitor em relação ao texto, do objetivo da leitura, da experiência do leitor com determinado tipo de texto, do objetivo de tornar os conceitos selecionados pertinentes às formas de buscas dos usuários, do tempo disponível para leitura, barulho no local onde ela se realiza, interesse pelo assunto do texto, entre outras variáveis.

Ressaltamos que com o desenvolvimento das idéias enfocadas no capítulo 1, alcançamos um dos objetivos da pesquisa que é *identificar estratégias de leitura para artigos de jornal*, ou seja, as estratégias levantadas nessa revisão de literatura sobre leitura são provavelmente na sua maioria adaptáveis para se realizar a leitura documentária, visando indexar jornais porque são utilizadas numa atividade de leitura que almeja a compreensão do leitor frente ao texto que está lendo e na leitura documentária também ocorre o processo de compreensão.

Destacamos que no próximo capítulo exploraremos a relevância do conhecimento e utilização das estruturas textuais dos textos para se compreender a idéia abordada num determinado texto, isto é, trabalharemos a estrutura textual como uma estratégia de leitura a ser utilizada visando atingir a compreensão sobre o que o texto aborda, ou seja, qual o seu tema. E mencionaremos mais especificadamente a estrutura textual do texto jornalístico por ser o objetivo de estudo dessa pesquisa.

CAPÍTULO 2: A ESTRUTURA TEXTUAL NA LEITURA DO INDEXADOR

O presente capítulo está dividido em dois tópicos: o primeiro aborda de maneira superficial a relevância de variadas estruturas textuais¹ na leitura do indexador e o segundo tópico dedica-se a explorar mais o conhecimento da estrutura textual do texto jornalístico, chegando a realizar uma experimentação que evidencia as partes mais significativas de conteúdo dessa tipologia documentária.

Primeiramente, destacamos que o conhecimento sobre as diferentes tipologias documentárias e suas estruturas textuais possibilita ao leitor conhecer, ou reconhecer, as idéias principais e secundárias tratadas no texto. Isso é possível porque conhecendo a estrutura do texto, o leitor sabe os componentes dele e a que parte pertence determinado componente, podendo assim realizar uma leitura mais direcionada explorando as partes do texto que apresentam as idéias que pretende obter e, assim, atingir o seu objetivo de leitura. Evidenciando que nosso objetivo de pesquisa é a observação da indexação de artigos de jornal, enfatizaremos a estrutura textual do texto jornalístico.

Ressaltamos que cada tipologia documentária teoricamente apresenta uma estrutura textual, isto é, uma superestrutura, desenvolvida com base nos diferentes tipos de construção dos textos, estabelecida pelos autores que optam por esta ou aquela construção que melhor expresse suas idéias.

Portanto, podemos perceber que não existe uma teoria geral sobre superestrutura do texto, pois está atrelada ao estilo de escrita do autor. Entretanto, o que se tem evidenciado com as pesquisas referentes à linguagem humana é uma teoria, até então, convencional da estrutura textual referente aos textos narrativos e argumentativos que têm possibilitado caracterizar esses textos. (Van Dijk, 1992).

Para Kobashi (1994, p. 112-113), a origem dos estudos referentes à superestrutura textual remonta aos primeiros estudos sistemáticos sobre os “gêneros textuais” e as “partes do discurso” desenvolvidos pelo filósofo Aristóteles. Encontramos reforço para essa argumentação na obra “Arte Retórica e Arte Poética” de Aristóteles traduzida por Antonio Pinto de Carvalho (1959,p.13) menciona que:

¹ Mencionamos estrutura textual nos referindo a superestrutura global do texto que é diferenciada considerando determinada tipologia documentária que por sua vez reflete os objetivos do autor no momento em que se dispõem a redigir determinado texto a uma estabelecida comunidade de leitor.

“a Arte Retórica conclui com uma série de considerações relativas às diferentes partes do discurso: exórdio, narração, demonstração, peroração; nem sequer fica excluída a facécia.”

Evidencia-se que o filósofo Aristóteles provavelmente foi o primeiro a distinguir a natureza e o papel de cada uma das partes do discurso por defender que *“é pelo discurso que persuadimos, sempre que demonstramos a verdade ou o que parece ser a verdade, de acordo com o que sobre cada assunto, é suscetível de persuadir.”*

Elucidando agora o que venham a ser as estruturas textuais, ou superestruturas, assim denominadas pelos lingüistas Van Dijk & Kintsch (1983), esclarecemos que uma superestrutura é “um tipo de forma do texto”.

E. para Moreira González (1993, p. 54),

“ a superestrutura é a ordem do texto, ou seja, a estrutura esquemática global, isto é, a distribuição do texto em partes com função diferenciada cuja a missão é organizar formalmente o texto.” (Tradução Livre)

O que autor (escritor) de modo geral procura evidenciar na superestrutura de um texto é o tema, denominado também de “macroestrutura textual”, que corresponde ao conteúdo do texto. (Van Dijk, 1992). Destacamos então que, dependendo da função específica do texto, este apresentará uma forma que pode, por exemplo, ser: uma narração, uma argumentação ou outra forma de texto e isso constituirá, uma determinada estrutura textual.

Moreiro Gonzáles (Op. Cit, 54) define a macroestrutura como “um super conceito para outros conceitos que vinculam uma seqüência de proposições.” (Tradução Livre)

Podemos perceber que a superestrutura textual é fundamental para a compreensão da leitura e que, tanto o leitor quanto o autor devem conhecer o formato de texto para buscar a compreensão do mesmo, identificando assim o tipo de informação que deverá encontrar, ou, no caso do autor que quer expor, se optar por esta ou aquela estrutura textual. (KATO, 1986).

Cintra (1987, p.31), baseando-se em pesquisas que evidenciam a facilidade de leitura para leitores com conhecimentos sobre estruturas textuais, considera que “o leitor que domina as superestruturas textuais, capta com mais facilidade as idéias centrais do texto, pois tem como parâmetro a identificação dos constituintes básicos” do texto.

Para compreender o texto, os leitores devem: ser conscientes de que devem buscar uma estrutura organizativa do texto, descobrir as distintas chaves da estrutura, e saber como modificar o marco de referência a medida que avançam, até que tenham construído uma representação do texto de acordo com seus propósitos. (CALFEE & CURLEY, 1997).

O lingüista Van Dijk (1992) considera que se o leitor pretende obter, durante a leitura a identificação e compreensão do tema do texto (macroestrutura textual), pode fazer uso das macrorregras: 1- Omitir toda a informação de pouca importância e não essencial a apreensão do tema; 2- selecionar somente a proposição essencial deixando de lado outras que podem ser consideradas como consequência óbvia; 3- generalizar, construir uma proposição geral sobre um conceito que se encontra omitido no texto e que são importantes para o entendimento do tema e, por fim 4- construir um conceito geral sem omitir ou selecionar, ou seja, agrupar todos os aspectos detalhados e formar um conceito geral.

Durante a aplicação dessas macrorregras, o leitor deve ter consciência que o resultado final deve ser equivalente ao significado do texto porque foram derivadas do texto, portanto devem-se evitar opiniões individuais, atitudes e ideologias por parte do leitor.

Como se percebe o uso das macrorregras está ligado ao conhecimento prévio do leitor sobre o assunto abordado no texto e somente o leitor que obtém esse conhecimento é capaz de aplicar corretamente as macrorregras, pois conhecimentos distintos resultarão em interpretações globais distintas do mesmo discurso. Por exemplo, se o tema do texto é "M" e determinado leitor quis dizer em sua aplicação das macrorregras, que quase compreendeu que era "M" mas não disse explicitamente, então não atingiu a compreensão do tema.

Mas é importante destacar que, para o leitor-indexador, o tema global do texto pode não interessar aos usuários e caso interesse somente uma vertente desse tema, o indexador deve indexar o tópico que interessa aos usuários conjuntamente com o tema global do texto, ou documento.

Como mencionamos nos itens anteriores onde falamos sobre a atividade de leitura global e documentária, o leitor durante a leitura pode utilizar o seu conhecimento sobre estrutura textual para localizar o tema abordado num determinado texto, com leitura rápida, por não precisar ler palavra por palavra, letra por letra, porque conhece as partes do

texto e as informações mais significativas pertencentes a cada parte. Com isso, atingirá o seu objetivo de leitura que é a identificação e compreensão do tema tratado no texto, mas a não obtenção de conhecimento científico ou elaboração de uma interpretação crítica sobre o tema.

O texto pode classificar-se em narrativo, descritivo, argumentativo e discurso científico, dentre outros. Podemos considerar que os textos narrativos são formas básicas muito importantes da comunicação textual, os argumentativos são mais utilizados em filosofia e teoria lógica, mas ambos servem de base para o texto científico (VAN DIJK, 1992).

Segundo Van Dijk (1992, p.160), a estrutura do texto argumentativo é composta de categorias que dependem do tipo de argumentação como:

ARGUMENTAÇÃO, que se divide em JUSTIFICATIVA e CONCLUSÃO;

JUSTIFICATIVA que se divide em MARCO e CIRCUNSTÂNCIA;

CIRCUNSTÂNCIA que se divide em PONTOS DE PARTIDA E FATOS;

PONTOS DE PARTIDA que se divide em LEGITIMIDADE e REFORÇO.

A estrutura do texto narrativo é composta, segundo Van Dijk (1992, p.156) tal como segue

NARRAÇÃO, composta de HISTÓRIA e *MORAL da HISTÓRIA*

HISTÓRIA se divide em TRAMA e EVOLUÇÃO

TRAMA se divide em EPISÓDIO

EPISÓDIO, se divide MARCO e o SUCESSO

SUCESSO se divide em COMPLICAÇÃO e RESOLUÇÃO

O texto científico se diferencia do texto argumentativo e narrativo por algumas categorias como: a CONCLUSÃO, e a JUSTIFICATIVA e ainda incluem-se as categorias PROBLEMA e SOLUÇÃO.

No caso de um texto científico que relata um informe experimental para publicar seus resultados, podemos perceber a seguinte estrutura textual: observação para encontrar uma explicação; para cada circunstância formulará hipótese, e dessa derivam previsões que podem ser comprovadas por TEST (teste); temos ainda as pessoas envolvidas no

ensaio, a estruturação do experimento, diferentes condições de investigações e a execução do experimento, resultado, conclusão (conclusões) e solução.

Na opinião dos autores Léon & Garcia Madruga (1991, citado por GALVEZ & PINTO s.d., p. 21) os textos se classificam em descritivos, narrativos e expositivos. Podemos caracterizar como textos expositivos: ensaios; artigos científicos e artigos de jornal.

Destaca-se no textos expositivos que seu conteúdo é organizado em três níveis obedecendo uma hierarquia entre as idéias: macro-nível onde se encontra a idéia principal; nível intermediário onde se encontram as idéias subordinadas e micro-nível onde se localizam os detalhes. (MEYER, 1977, citado por GALVEZ & PINTO , s.d., p. 21) .

Percebemos que os textos expositivos e argumentativos são muito parecidos, e algumas vezes parecem se complementarem para originar o texto científico.

O texto expositivo na opinião de Selinker (1976, citado por GALVEZ & PINTO, s.d.) possui seu conteúdo organizado seguindo uma hierarquia e as funções retóricas se encontram classificadas em uma série de níveis: detalhar um experimento, fazer uma recomendação, formular um problema, definição, descrição e causalidade.

No caso específico do texto científico, as autoras Galvez & Pinto (s.d., p.23) caracterizam-no a partir de:

- *Estrutura textual que contém: objetivo, metodologia, resultados e conclusões;*
- *“ Um estilo cuidadoso e altamente formalizado;*
- *A objetividade do conteúdo está de acordo com a realidade científica;*
- *O emprego de uma linguagem científica diferente da linguagem habitual;*
- *Prioridade ao explícito, ao conhecido, à informação acumulada durante o desenvolvimento da humanidade e à própria essência como criação humana.”*

Ressaltamos ainda, com relação ao texto científico, que esse pressupõe grande quantidade de informação, e a pressuposição é um importante fator a considerar no exercício das tarefas analíticas realizadas pelo indexador (GALVEZ & PINTO, s.d.).

Segundo as autoras acima citadas, na estrutura textual das várias tipologias expostas percebe-se que a informação principal se retém e se evoca com maior facilidade do que a informação secundária. Destaca-se também que cada grupo de textos corresponde

a distintos esforços cognitivos por parte do receptor, dentre esses os textos expositivos exigem ou requerem a contribuição de mais recursos cognitivos que os descritivos ou narrativos.

Com base nas suas observações, Kato (1986) apresenta os tipos de estrutura textual e o texto que corresponde a cada estrutura:

- Forma Piramidal: começa com a idéia principal e geral, para depois ir dando detalhes e particularidades, é a *estrutura do texto jornalístico*.
- Forma argumentativa: começa com uma introdução, definição do problema, solução proposta, argumentos a favor e contra, refutação de argumentos contrários e conclusão- nessa estrutura encontramos os *argumentos científicos e legais*.
- Formato de Eliminação: monta um argumento eliminando, sucessivamente, várias propostas até chegar a sua ou descreve vários acontecimentos para justificá-los, no conjunto, somente no fim, é a *estrutura do texto científico*.
- Formato de uma Narrativa: privilegiado por estruturar histórias, é usado para descrever o que foi pensado e feito em uma sequência temporal, como uma narrativa típica, é a *estrutura dos textos de histórias*.

Kobashi (1994, p. 112-119) afirma que, se o indexador conhecer e utilizar a exploração das estruturas textuais dos documentos com os quais trabalha na indexação, identificará e selecionará o tema que considera relevante em determinado documento realizando uma leitura documentária rápida; isso lhe favorece devido a grande quantidade de documentos que tem para indexar e a inexistência de disponibilidade de tempo para realizar uma leitura integral de cada documento. A autora ainda determina os tipos de textos usados na Análise Documentária e suas estruturas:

1- Texto científico

Constituído por:

- Tema: assunto que se irá desenvolver
- Problemas: dificuldade que se quer solucionar cientificamente
- Hipótese: proposição que se antecipa à comprovação de uma realidade, porque propõe-se, através dela, uma resposta a um problema que poderá ser comprovado ou refutado pelas observações a que os fatos serão submetidos.

- Metodologia: procedimentos e operações que possibilitem a observação racional e controlada dos fatos, de modo a permitir a interpretação e a explicação adequada do fenômeno observado.

- Resultado e conclusão:

Resultado: implica na aceitação ou não das hipóteses formuladas ou na reformulação das mesmas;

Conclusão: comentário final, em que se discutem as possibilidades de aplicação e de utilização dos resultados, isto é, a incorporação ou não destes últimos a um sistema teórico.

2- ARGUMENTATIVO OU DISSERTATIVO

- Tese: apresentação de um ponto de vista;
- Argumentos: são as provas apresentadas para qualificar positivamente o ponto de vista do autor e, simultaneamente, desqualificar um ponto de vista inicialmente apresentado;
- Conclusão: confirma o ponto de vista inicialmente apresentado;

3- EXPOSITIVO

- Problema: dificuldade que se pretende resolver, verificar cientificamente,
- Causas: causas e conseqüências do problema;
- Solução: resposta para a solução do problema.

Para Tálamo (1994), o discurso científico possui elementos tópicos que são básicos no procedimento científico e que permitem ao leitor que conhece esses elementos caminhar do não-saber para o saber, através da individualização de cada componente do discurso, e esses componentes são: problema, hipótese, argumentação e conclusão.

Tálamo (1994, p. 16-18) propõe uma estrutura de organização para o discurso científico que abrange:

- **Pressuposição:** *é um conjunto de informações sobre as quais se baseia o texto e que não é motivo de discussão ou contestação. É, na maioria dos casos, um tipo implícito que permeia todo o discurso, é o dado, a partir do qual se propõe um problema, uma hipótese, enfim uma investigação;*
- **Problema e o tema:** *o problema é a questão que, tendo em vista uma teoria particular, não obteve uma solução adequada. É uma representação conceitual das dificuldades com as quais se defronta o pesquisador e que serão resolvidas sucessivamente por conjecturas e testes. O problema, normalmente, é formulado de maneira bastante específica, pois se relaciona a um tema, este sim, uma proposição abrangente.*
- **As Conjecturas:** *definem-se como a resposta hipotética dada a um problema, portanto o motivo da investigação, a qual deve, em princípio, validá-las ou não.*
- **Os Testes:** *nesta etapa, os diversos argumentos favoráveis ou não, para a aceitação das conjecturas, defrontam-se agora. A estrutura polêmica do discurso é considerada justamente pela batalha travada entre os argumentos favoráveis e contrários à resposta hipotética. É só a partir dos testes que se pode falar em validação de conjecturas e coerência do texto.*

Como demonstramos no decorrer deste trabalho, a estrutura textual facilita a apreensão do tema do texto auxiliando, assim, na compreensão da leitura, sendo um fator a ser considerado quando se pretende realizar uma leitura rápida de determinado documento, que é caso da indexação de documentos realizados por indexadores em centros de informação especializados ou não. Podemos, portanto, considerar o uso de estrutura textual durante a leitura como uma estratégia metacognitiva e evidenciamos essa afirmação com a citação na íntegra exibida abaixo que menciona que, segundo Sanchez, Orrantia & Rosales (1992) citados por Galvez & Pinto (s.d, 62), o uso da estrutura textual é considerado uma estratégia de leitura a medida que permite:

1- detectar la organización interna de los textos, 2- utilizar esse patrón como instrumento para asimilar la información del texto, 3- ordenar en esse patrón o armazón da información que sera derivado del texto, 4- y si llega el caso, usar esse esquema organizador pra planejar memoria.

2.1 A ESTRUTURA TEXTUAL DE ARTIGOS DE JORNAL NO PROCESSO DE LEITURA DOCUMENTÁRIA

Com intuito de verificar como é a organização da estrutura textual do artigo de jornal, primeiramente enfatizaremos o que é o texto jornalístico e quais os processos para a sua construção e, na sequência do capítulo, aplicaremos o método diplomático no texto

jornalístico visando evidenciar as partes relevantes da estrutura textual dessa tipologia documentária.

De acordo com Kientz (1973, p.86), o texto jornalístico se origina com a transformação de um acontecimento ou fato social em notícia. Esse deve passar pelos seguintes processos para que ocorra essa transformação:

a- Filtragem em função da originalidade: Processo de seleção dos acontecimentos ou fatos sociais que se deseja noticiar podendo constituir este um assunto invulgar, surpreendente, comum, ou incomum. Realiza, também, uma seleção dos itens ou elementos que se quer discorrer dentro de determinado acontecimento. Segundo Marcondes Filho (1985), a transformação do fato em notícia depende dos interesses dos jornais e de suas relações com representantes de grupos ou setores políticos, econômicos ou sociais;

b-Tratamento da informação em função da inteligibilidade: Condensação do texto através do uso de palavras-chave e aumento no texto de explicações, repetições, explicitações com o objetivo de fornecer uma melhor compreensão do texto pelos leitores;

c-Tratamento da informação em função do grau de implicação: A retenção da notícia está relacionada com o distanciamento psicológico e geográfico do leitor em relação a ela, pois quanto mais distante estiver a notícia mais dificuldade se terá de retê-la;

d-Tratamento da informação em função da profundidade psicológica: As informações retidas pelos leitores são aquelas que atingem o inconsciente dos mesmos. E para se obter isso, é preciso que a informação receba uma coloração emocional e uma dramatização da mensagem.

Segundo Charaudeau (1983), após ter identificado a notícia entre os fatos ou acontecimentos, o jornalista e o jornal devem discorrer sobre os fatos obedecendo os seguintes contratos:

- Contrato de autenticidade: o fato deve ser narrado como acontece no real, ou seja, sem distorções.

- **Contrato de Seriedade:** tanto o jornalista como o jornal têm responsabilidade em relação à objetividade da notícia e com a sua transmissão, pois a informação deve ser divulgada e não se manter secreta.

Segundo Benitez (1983, p. 104), após ter se identificado a notícia, deve-se discorrer sobre os seus elementos constitutivos, determinados itens que se deseja discorrer dentro de determinado assunto, podendo para isso fazer uso da seguinte estrutura narrativa:

- a- **Desenvolvimento dos fatos em ordem cronológica:** os fatos narrados obedecem a uma ordem cronológica onde os itens mais importantes são narrados no final do texto; é caracterizado, por isso, como estilo antigo do jornalismo e ocorre em textos literários;
- b- **Sistema misto de narração:** composto por uma síntese do fato no início do texto, seguida de um relato detalhado sobre o fato e, no final do texto, apresenta uma justificativa da síntese do fato exposto no início do texto. Essa estrutura é usada em textos de suspense, em alguns tipos de notícia, e na maioria das vezes está atrelada ao Jornalismo Sensacionalista;
- c- **Desenvolvimento da notícia pela ordem de importância dos seus elementos:** apresenta uma síntese da notícia no início do texto e deixa as informações secundárias para o final do texto; essa técnica é denominada de Pirâmide Invertida e constitui o estilo de narrativa mais usado pela imprensa ocidental.

O texto jornalístico conforme Amaro (1991) pode ser caracterizado quanto ao tipo de mensagem que se deseja transmitir ao leitor, através dos seguintes tipos de discursos:

- a- **Discurso Relatado:** apresenta citação direta ou indireta e descrição de atividade argumentativa em causa. O jornalista obedece ao contrato de autenticidade e ocorre também um distanciamento entre o fato e o jornalista;
- b- **Discurso Alusivo:** Produz um enunciado que pode ser compreendido de diversas maneiras por membros de uma mesma comunidade, levando-se em consideração o conhecimento do leitor sobre determinado enunciado. Esse tipo de discurso é sutil, pois provoca no inconsciente do leitor um apelo à convivência com a interpretação definida pelo jornal sobre determinado enunciado, podendo o mesmo veicular a ideologia desejada;

c- Atualização existencial: Faz-se uso do tempo verbal no presente induzindo o leitor a interpretar o momento da leitura como se fosse o mesmo do acontecimento narrado. Essa estratégia é utilizada mais nos títulos dos textos jornalísticos;

d- Contrato do insólito: Consiste em narrar o fato como algo que está acontecendo na realidade e ao mesmo tempo fixa o fato como algo irracional, de ficção. Esse tipo de discurso é utilizado para os fatos de grande impacto como: sequestros, assaltos, atentados, guerra, acidentes, catástrofes, etc.....

O jornalista deve também na redação do acontecimento fazer uso de sentenças curtas, simples, com estrutura padrão (substantivo-verbo) pois podem ser mais fácil de compreender do que se optar por utilizar sentenças longas complexas. E quanto à escolha dos itens lexicais usados para determinar participantes do evento descrito orienta-se que tenha muita atenção pois os significados de determinado item (s) lexical (is) escolhido (s) expressam as crenças, opiniões, ideologias específicas do jornal ou do jornalista. Mas somente a escolha desse ou daquele item lexical não é suficiente, deve-se também observar a conotação denotativa da palavra, por isso, se necessário, fornecer sempre maiores detalhes que expliquem melhor o evento que está narrando. (VAN DIJK, 1983).

Van Dijk (1983) também alerta para o cuidado que se deve ter com as palavras consideradas neutras como por exemplo a palavra *simpatizante*, pois estas podem receber conotação negativa em determinados contextos. E na apreensão do tema do texto (macroestrutura), o conceito geral negativo de determinado participante é formado mesmo que nem todas as denominações sejam realmente negativas, uma vez que é o conceito geral que é o mais prontamente recuperado na memória.

O jornalista deve obter os conhecimentos necessários sobre a produção do texto jornalístico no curso de graduação em Comunicação Social para que se torne um profissional competente no mercado de trabalho. Defendendo essa afirmação, Pedrosa (1996) analisou dez cursos de graduação em Comunicação Social no Brasil e propôs uma metodologia para o ensino da redação jornalística.

Conforme Pedrosa (1996, p.210), o aluno-futuro jornalista deve ser “*agente da sua própria aprendizagem*” e o professor de redação jornalística pode conseguir com que o aluno aprenda redação oferecendo a ele as seguintes formulações de ação pedagógica que o encaminhará num melhor desempenho na hora de redigir o evento:

- a- conhecer a Língua Portuguesa, sua ortografia, morfologia, semântica, sintaxe e estilística;
- b- conhecer e exercitar a qualidade e a técnica do texto jornalístico;
- c- conhecer a rotina da produção jornalística;
- d- captar (por entrevista, reportagem, pesquisa e observação) e relacionar informações;
- e- experimentar variações na redação do texto, isto é, saber utilizar os recursos expressivos da Língua;
- f- distinguir e redigir os diversos gêneros jornalísticos;
- g- realizar entrevistas e conhecer as técnicas de pesquisa e de investigação nas Ciências Sociais e Humanas aplicadas ao Jornalismo;
- h- realizar pesquisa com fontes bibliográficas e com materiais não-convencionais;
- i- exercitar a técnica (e a ética) de edição;
- j- elaborar pauta;
- k- titular matéria;
- l- compreender e praticar a técnica da verificação (checagem permanente de fontes e de informações);
- m- contextualizar os acontecimentos, isto é, produzir reportagens interpretativas e investigativas) e ensaios jornalísticos;
- n- planejar e executar atividades de laboratório de agência de notícias;
- o- conhecer e praticar a técnica de copidescagem e de revisão de texto;
- p- planejar, executar e avaliar atividades de jornais, revistas, rádios e tevês laboratórios;
- q- redigir textos para jornais diários e para revistas semanais de informação;
- r- redigir para editoriais e cadernos especializados;
- s- redigir notícias, entrevistas e reportagens; editoriais, artigos de fundo e críticas.

Ainda para Pedroso (1996, p. 211), o aluno-futuro jornalista, ainda na graduação, deve ser direcionado pelo professor a avaliar os procedimentos técnicos da atividade jornalística. Sendo necessário para isso:

- a- refletir sobre questões fundamentais do exercício da atividade jornalística, tais como:
 direito à informação e direito à opinião; interesse público e interesse do público; verdade e privacidade; distorção deliberada e distorção inadvertida; manipulação técnica e manipulação ideológica; concentração e abuso de poder; precisão da notícia e verificação permanente, monopólio e impunidade; sociedade indefesa e peso do marketing; denunciismo e sensacionalismo e ética profissional;
- h- desenvolver a capacidade de analisar e avaliar o significado de acontecimentos da atualidade;

- c- conhecer os princípios básicos da teoria de texto e de redação, tais como: a organização de idéias ajuda a estruturar o texto; clareza é a primeira condição do texto;
- d- conhecer os fundamentos da Teoria do Jornalismo, tais como: a transformação do acontecimento em notícia; a seleção da notícia; o caráter de mercadoria da informação; os condicionantes econômicos, políticos, culturais e ideológicos do jornalismo, etc;
- e- conhecer os princípios fundamentais da boa apuração e cobertura dos acontecimentos.

Após ter sido desenvolvido com os alunos as atividades pedagógicas descritas acima, e também, refletindo sobre os procedimentos técnicos da atividades jornalística, o professor pode utilizar com os alunos a metodologia da redação jornalística chamada de “método da roda-crítica”, elaborada por Pedroso (1996), que consiste basicamente de: cada aluno elaborar uma pauta e posteriormente um texto jornalístico, apresentá-lo à turma que o avaliará e, posteriormente, o professor avaliará a matéria objetivando que todos os alunos tenham suas matérias avaliadas pela turma e pelo professor. Pedroso (1996, p.212) sugere que o professor faça uso do seguinte conteúdo programático na disciplina de ensino da redação jornalística:

- a- o processo de captação e de apuração da redação jornalística: discussão e elaboração de pauta; pesquisa, entrevista e verificação das informações; b- o relato da informação (a opinião implícita e a opinião explícita) e a organização do texto (titulação, hierarquização, contextualização e edição); c- a redação como resultado de um processo de seleção (ideologia profissional), ordenação, hierarquização, nomeação e estilo; d- o texto jornalístico como resultado do trabalho de apuração (pesquisa, entrevista, observação e verificação), investigação e interpretação dos acontecimentos.

O que nos chamou atenção no trabalho de Pedroso (1996, p.214) foram os parâmetros estabelecidos para a avaliação da matéria pelo professor (descritos abaixo), pois a estrutura textual do artigo de jornal é evidenciada e o “lead” é descrito como o elemento primeiro e, posteriormente, o desenvolvimento e a conclusão como veremos a seguir quando explorarmos a estrutura textual do texto jornalístico.

- 1- qualidade de pauta: abordagem, angulação e enfoque; 2- qualidade da apuração: a pesquisa e o contraponto; a observação e a verificação; 3- qualidade do texto: atualidade (gancho, novidade); correção da linguagem; correção gramatical; correção da sintaxe; adequação às normas técnicas da produção de originais; precisão dos dados e das informações; 4- estruturação do texto: clareza, concisão e coerência; unidade central (lead), desenvolvimento e conclusão; construção dos parágrafos (encadeamento, interrelação, sequência lógica e coerência). (Grifo nosso)

Mencionamos acima algumas considerações de Pedroso (1996) para evidenciar a preocupação de transmitir ao jornalista conhecimentos que lhe permitam coletar dados e redigir uma redação coerente, pois que, não podemos deixar de citar, o jornalista raramente é testemunha direta do acontecimento por não estar presente no local exato e na hora exata em que ocorreu um acontecimento que precisa narrar, tendo esse que se basear em discursos de outras pessoas, tais como: relato de testemunhas oculares.

Com isso, a produção de notícias é uma forma de processamento do discurso, uma vez que o jornalista redige o discurso das testemunhas oculares, dos entrevistados, as declarações de autoridades, entre outros. (VAN DIJK, 1983)

Enfocaremos agora a estrutura textual do artigo de jornal desenvolvida por Van Dijk (1983) para o texto jornalístico produzido na cultura ocidental.

Ressaltamos essa estrutura aqui, porque acreditamos que conhecendo a estrutura textual é possível localizar as informações necessárias para a extração dos termos no processo de indexação do jornal, sendo possível realizar de forma rápida o tratamento da enorme quantidade de informações diária de um jornal. Destacamos que, baseados nessa superestrutura, compararemos os resultados obtidos na coleta de dados da observação da indexação de jornais verificando a exploração dessa estrutura textual pelos indexadores na ambiência de nosso local de pesquisa descrito em trechos seguintes.

De acordo com Van Dijk (1983, p. 37), essa superestrutura 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 Conseqüências/ Reações

2.2.1 Eventos

2.2.2 Atos da fala

3. Comentários

3.1 Expectativas

3.2 Avaliação (Tradução Livre)

De acordo com Van Dijk (1983) as definições para as categorias dessa superestrutura são as seguinte:

1- Sumário/Introdução, “no qual são introduzidos os eventos principais , os participantes e a localização dos eventos principais,

1.1 Cabeçalho (título, subtítulo, olho) fazem parte do sumário/introdução;

1.2 Lead “é uma síntese do fato apresentada normalmente no primeiro parágrafo do texto que contém, com o menor número possível de palavras, as respostas lógicas às perguntas que o leitor se faz em função de qualquer fato: Quem? Que? Quando? Onde? Por que? Como?

2. Episódio(s) “compõe-se de duas outras categorias: os eventos e as conseqüências. Na categoria dos eventos temos: eventos presentes que é o núcleo da estória da notícia e os Antecedentes ou Informações Prévias que fornecem informações sobre os fatos que precedem o evento presente.

2.1 Informações prévias ”fornecem informações sobre os fatos que precedem o evento presente.

2.1.2 Antecedentes:

2.1.3 Eventos presentes englobam a categoria dos eventos principais que descrevem os fatos principais em foco.

2.1.3.1 Explicação; “apresenta algum contexto específico dos eventos presentes, bem como a experiência passada”.

2.1.3.1.1 Contexto

2.1.3.1.2 Experiências passadas: “fornecem informações históricas, cultural ou política dos eventos presentes, seus participantes, países envolvidos ou quaisquer outras informações dessa natureza.”

2.1.3.2 Eventos principais: Pontos relevantes identificados no desenrolar da matéria

2.2 Conseqüências/ Reações

2.2.1 Eventos

2.2.2 Atos da fala: quando a personagem indaga algo.

3. Comentários

3.1 Expectativas: possíveis respostas para um determinado problema

3.2 Avaliação: resultados obtidos com base no texto

Para o lingüista Van Dijk, no texto jornalístico é convencional apresentar 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 descrita acima. Com essa determinação é possível chamar a atenção do leitor que decidirá se o texto lhe interessa ou não.

O autor ainda destaca, que se o leitor obtiver informação sobre um acontecimento que vem sendo publicado no jornal, poderá na leitura das matérias novas sobre o acontecimento realizar uma leitura dos componentes iniciais do texto e compreenderá superficialmente do que o texto trata. Isso é possível porque os componentes primeiros da estrutura do jornal, letras escritas geralmente em maiúsculo e em negrito que, juntamente com o *Lead*, têm a função de expor ao leitor um breve resumo sobre o tema que o texto abordará em sua seqüência, funcionando assim, como *macroestrutura global do texto*, deixando os detalhes ou especificidades para a seqüência da estrutura textual. (VAN DIJK, 1983).

Mas, isso não quer dizer que nas categorias seguintes da estrutura textual as informações descritas no CABEÇALHO e LEAD não venham a aparecer, pelo contrário, isso pode ocorrer pois a função do CABEÇALHO e LEAD é transmitir ao leitor o tema que o texto abordará para que esse possa decidir se lerá ou não o texto. Portanto, as informações descritas ou mencionadas no LEAD poderão ser descritas mais profundamente nas outras categorias da estrutura textual.

Evidenciamos então que dados de uma informação podem aparecer mais de uma vez na estrutura textual do artigo de jornal, funcionando o LEAD bem escrito como um “representante” fiel do texto.

Elucidamos a afirmação acima com uma citação que diz:

(...) o título e o cabeçalho são elementos informativos importantes que sinalizam qual será a provável macroestrutura do texto. (...) (VAN DIJK, 1992, p.77, tradução livre)

Percebemos então que a macroestrutura, isto é, o tema no artigo de jornal, é bem sinalizado na estrutura textual, podendo ser facilmente percebido pelo leitor por apresentar-se em posição de destaque na disposição do todo, apresentar mudança nos tipos, ou seja, letras maiúsculas e geralmente em negrito ou itálico e que chamam a atenção mencionando aqui o componente CABEÇALHO da estrutura textual. Acrescentamos também o componente LEAD como integrante da macroestrutura, ressaltando que esse não apresenta mudança de tipos de letras, pois está presente no corpo do texto no (s) primeiro (s) parágrafo (s) do jornal.

Destacamos que os componentes CABEÇALHO e LEAD são elementos do nível mais alto da estrutura textual porque

“chamam a atenção do leitor para o artigo; permitem que se decida ler ou não o resto do discurso; apresentam o tema principal mesmo sem leitura mais aprofundada; ativam o conhecimento relevante da memória o que é necessário para se compreender o resto do texto e finalmente, eles formam a macroestrutura que servirá como a importante sugestão estratégica no controle da compreensão local do texto subsequente.” (VAN DIJK, 1983) (Tradução Livre)

Alertamos que o leitor somente pode procurar compreender um tema considerando o elemento TÍTULO que compõe o CABEÇALHO, se esse não se constituir

Os componentes ANTECEDENTES fornecem as informações sobre os fatos que precedem os verdadeiros acontecimentos, e que podem ou não terem sido relatados antes no LEAD e no EPISÓDIO pois, lembramos aqui, que no jornal geralmente uma matéria pode se estender no decorrer de novas publicações sobre a matéria.

Quanto ao componente ACONTECIMENTOS verdadeiros é a parte central da história da notícia e descreve os principais (novos) acontecimentos agora em foco.

Já o componente ACONTECIMENTOS CONSEQUENTES ou ATOS de FALA, que são tipicamente afirmações, declarações de participantes ou observadores politicamente importantes, podem ser colocadas antes e as “informações prévias” ou antecedentes colocadas mais tarde no texto.

Ressaltamos que as categorias que compõem a estrutura textual do jornal são teóricas, porque num discurso real de notícia: algumas delas não precisam ocorrer explicitamente, principalmente, quando se supõe que os leitores já dispõem dessa informação ou porque a informação é considerada irrelevante. Nesse caso, o leitor pode compreender parcialmente o texto porque dispõe da informação necessária para isso, ou então, será “conduzido” a compreender aquilo que o jornal/jornalista quer que ele compreenda se não obtiver a informação necessária para efetuar a compreensão.

Portanto, podemos dizer que o CABEÇALHO, o LEAD e o EPISÓDIO são os elementos mais importantes da estrutura textual de jornal. Os outros componentes da estrutura são responsáveis por fornecer os detalhes do acontecimento, deixando os menos importantes para o fim com a consequência prática que estes podem, se necessário, serem cortados pelo editor no momento da revisão da matéria e, dependendo da importância da informação, cortada no texto, prejudicando a compreensão do leitor.

A compreensão do discurso jornalístico pelo leitor ocorre da seguinte forma:

“ o leitor ativa para a leitura a memória de longo-termo (onde estão armazenados conhecimentos e crenças) responsável pela aquisição do conhecimento sobre acontecimentos políticos. O leitor terá também ativado e armazenado no sistema de controle o conhecimento geral importante sobre o jornal e seus pontos de vista dominantes sobre questões políticas (ideológica) e possivelmente terá expectativas concretas sobre o tipo de informação semântica. A memória episódica (a parte da memória de longo termo que armazena todas as experiências que aparecerão) produzirá pela estratégia de “lembrança” as informações sobre os acontecimentos que ela adquiriu nos últimos dias. Finalmente, o leitor de notícias, terá ativado conhecimento gerais sobre os assim chamados “esquemas” do discurso da notícia das novas matérias publicadas no jornal.” (VAN DIJK, 1983, p. 30) (Tradução livre)

Para Van Dijk (1983), a busca pela compreensão pode ser direcionada se observarmos os itens da estrutura da superfície, ou, estrutura superficial, como: número de página; posição na página, tamanho, tipos de impressão usados, números de colunas, presença de fotografias. O autor destaca, ainda, que os sinais de macroestrutura no nível intertextual são: localização, tamanho, etc... pois sinalizam qual de muitos fatos ou acontecimentos é relativamente o mais importante.

Mesmo tendo elaborado a estrutura textual para o texto jornalístico, Van Dijk (1983, p.40) faz algumas recomendações sobre a necessidade de realizar um trabalho empírico mais aprofundado para testar a relevância deste esquema para notícia. Para o autor, deve ser, também, empiricamente investigado se esse esquema de notícia tem alguma “realidade” cognitiva, ou seja, se os leitores, de fato, organizam o discurso da notícia de acordo com essas categorias e se eles as esperam ou sentem falta delas quando não estão presentes.

Van Dijk (1983) considera que a pesquisa empírica sobre leitura, leitura demorada e leitura rápida de jornais, revelará as implicações relevantes destas variáveis estruturais presentes na compreensão lembrança (memória).

Entretanto, podemos mencionar aqui que realizamos em nossa pesquisa de graduação, em 1997, a investigação sobre leitura documentária na indexação de jornais, e na análise dos resultados identificamos o uso dessa estrutura textual proposta por Van Dijk (1983). Com os resultados obtidos, estamos desenvolvendo, em sequência, nossa pesquisa de pós-graduação com a proposta de nova observação da atividade de leitura para a indexação de jornais a fim de elaborar um modelo de leitura para a indexação de jornal, explorando, entre outras sistemáticas de identificação de conceitos, o uso da estrutura textual de jornal.

Com os resultados da pesquisa de 1997, chegamos a conclusão de que algumas partes da estrutura textual eram suficientes para o indexador selecionar o assunto do texto jornalístico.

Em contato com os resultados das experimentações realizadas por Guimarães (1998) no campo de tratamento temático, verificamos subsídios teórico metodológicos da Diplomática para a análise de documentos, visando analisar as mais variadas tipologias documentárias como: rótulo de shampoo, (Straiotto, 1997); receitas culinárias (Bueno, 1998) entre outras, com o objetivo de identificar os procedimentos documentários que o Tratamento Temático e a Diplomática podem oferecer a essas tipologias documentárias.

Antes de mencionarmos a metodologia de Guimarães enfatizaremos o que é a Diplomática e o que se entende por documento diplomático dentro da Documentação.

2.2 A DIPLOMÁTICA E O TEXTO JORNALÍSTICO: ANÁLISE DA TIPOLOGIA DOCUMENTÁRIA

Segundo Guimarães (1998), a Diplomática surgiu no século XVII na Europa, com a função de estudar os documentos escritos e verificar a sua autenticidade. Assim, a Diplomática² trabalha com corpus de texto restrito que é o documento diplomático.

De acordo com Tessier (1952), citado por Guimarães (1998), a Diplomática na análise dos documentos trabalha com caracteres externos e internos, sendo esses:

■ Caracteres Externos: referem-se ao tipo de suporte (exemplo: atestado; certidão, etc...), e estão voltados para garantir a autenticidade dos documentos, destacando-se os seguintes itens:

- meio físico: (original versus cópia)
- formato (exemplo: atestado versus certidão, a certidão tem valor de autenticidade maior que o atestado;
- signos: são carimbos, selos, timbre, chancelas.

■ Caracteres Internos: estão voltados para o conteúdo do documento preocupando-se em garantir a coerência do documento e sua função, destacando-se os seguintes itens:

- língua: (de conteúdo)
- teor documental: estrutura do texto da informação

Por sua vez, o teor documental divide-se em texto e protocolo:

- texto: constitui o corpo do documento e expressa o fato documentado (divide-se em: preâmbulo, notificação, exposição, sanção, etc...) e são informações que precedem o texto.
- protocolo inicial: divide-se em inovação simbólica e verbal (exemplo: intitulação, direção e saudação)
- protocolo final: são informações posteriores ao texto dividindo-se em: subscrição, data (tópica e cronológica) e precação. Pode-se dizer que os Protocolos são conjuntos de regras formais não variáveis (ou que demoram para mudar). Devendo-se ressaltar, também, que

² "Dentre das ciências documentárias a Diplomática é a atividade que se ocupa de descrição e da explicação dos atos escritos; seu campo de aplicação são os documentos gerados na área pública, neles estabelecendo a forma que lhes conferirá validade legal" (BELLOTO, 1991. p.30).

esses elementos não aparecem mais ou menos estáveis em todos os documentos e que muitos nem são mais utilizados.

A Diplomática realiza a análise do documento diplomático com o objetivo de verificar a sua autenticidade, e para isso, procura estabelecer formas para conferir a validade legal ou autenticidade dos documentos diplomáticos, com base no estabelecimento dos Protocolos.

Guimarães (1998), com o conhecimento dos Protocolos, usados pela Diplomática para “*visualizar áreas de conteúdos temáticos específicos em textos da área jurídica*”, indica para uma eficaz realização das atividades de Tratamento Temático, que o bibliotecário/indexador, procure identificar nos documentos sua fórmula diplomática, ou seja, a estrutura textual baseando-se em itens que denominou como método diplomático.

No método diplomático, Guimarães (1998) indica quatro aspectos importantes:

- ① cada documento nasce com uma função;
- ② cada documento, para cumprir sua função original, possui uma estrutura (fórmula);
- ③ documentos de mesma função têm a mesma estrutura (tipologia documentária), sendo que essas tipologias podem ser divididas em espécies;
- ④ documentos podem ter outros usos além de sua função original. (exemplo: um bilhete de passagem que tem como função original o embarque no ônibus para determinado destino, mas pode ser usada para comprovar um valor gasto).

Com a descoberta da fórmula ou estrutura textual do documento, podem-se identificar os elementos diplomáticos em ordem de prioridade (informações principais e secundárias), destacando que elementos complementares não influenciam na função do documento.

Por ser aplicado a distintos tipos de documentos, o método diplomático é conveniente ao bibliotecário que trabalha com documentos em diferentes suportes e nos mais diversificados centros de informações. Essas tipologias documentárias devem passar por tratamento temático com objetivo de priorizar suas informações principais, proporcionando, assim, acesso mais rápido à informação desejada pelo usuário e, conseqüentemente, abrindo campos de trabalho para os bibliotecários nos centros de informações que armazenam diferentes tipologias documentárias.

Com o conhecimento do método diplomático para distintos tipos de documentos, procurou-se observar sua aplicação para a leitura de jornal, visando identificar se os procedimentos são úteis para elaboração de modelo de estratégias de leitura documentária.

Em Fagundes (1997), em pesquisa referente ao processo de leitura em artigos de jornal, entrevistou dois indexadores do Arquivo do jornal “ O Estado de São Paulo”, um jornalista e um bibliotecário, supondo que:

- a leitura documentária de jornais deve ser rápida devido à grande quantidade de material publicado a ser indexado;
- a nossa observação da rapidez da leitura proporcionaria a outros indexadores o conhecimento dos métodos utilizados e suas vantagens.

Para observação dos processos e estratégias de compreensão dos indivíduos durante a leitura, optamos pela técnica denominada Protocolo Verbal, no qual o indivíduo que realiza a leitura, interpreta e exterioriza verbalmente sua compreensão.

Ao investigarmos os procedimentos de leitura utilizados por indexadores de artigos de jornal, observamos que, na leitura dos artigos, os indexadores concentraram-se mais no início do texto, especificamente no item “Lead”, porque as informações contidas nesse item são suficientes. Constatamos também, que a estrutura narrativa presente nesses textos é a estrutura denominada “Pirâmide Invertida,” muito utilizada pela imprensa ocidental. Constitui-se do desenvolvimento da notícia pela ordem de importância dos seus elementos, isto é, o relato começa com a síntese da notícia deixando os detalhes secundários para o final. (BENITEZ, 1983, p.104)

Considerando-se esses resultados e o conhecimento da estrutura textual do texto jornalístico, verificou-se o uso do método diplomático nas atividades de tratamento temático.

A explicação do método diplomático foi iniciada com a análise dos quatro aspectos indicados por Guimarães (1998) e as contribuições que oferecem para a elaboração do modelo de leitura documentária de artigos de jornal.

① cada documento nasce com uma função:

O artigo de jornal se origina da transformação de um fato em notícia, cuja função é informar ou comunicar ao leitor a notícia sobre a cena do real, obedecendo a contratos de fala de autenticidade e de seriedade. (CHARAUDEAU, 1983, p. 101)

② cada documento, para cumprir sua função original, possui uma estrutura fórmula

Conforme resultados anteriores notamos que as partes da estrutura textual mais exploradas são: Cabeçalho, “Lead” e Episódio (s). Segundo Amaro (1991, p. 33), que também concorda com Van Dijk (1983), o “Lead” corresponde a uma síntese do texto referente ao primeiro parágrafo, consistindo num relato altamente condensado que contém, no menor número possível de palavras, as respostas lógicas às perguntas que o leitor se faz

em função de qualquer fato: Quem? Quê? Quando? Onde? Por quê? Como? . Portanto, de acordo com a autora, o “Lead permite a identificação de pessoas, lugares e fatos, de forma mais clara e rápida para a compreensão da informação”.

Com base na sistemática de identificação de conceito utilizada pelos indexadores, observou-se que, na maioria dos textos selecionados, é possível realizar a leitura documentária e extrair os conceitos lendo somente o primeiro parágrafo. Entretanto, em determinado texto, a leitura estendeu-se até o item Episódio (s) da estrutura textual. Com isso, ressalta-se que a leitura somente do Lead, em alguns casos, pode não fornecer as informações ou conceitos que o leitor indexador precisará para representar o assunto dos documentos.

De acordo com a aplicação do método diplomático e levando em consideração as partes da estrutura textual exploradas pelos indexadores, identificamos os seguintes elementos:

❖ Elementos principais

- Cabeçalho
- Lead
- Episódio (s)

❖ Elementos Secundários

- Comentários
- Expectativas
- Avaliação

Os elementos principais identificados na estrutura textual de artigos de jornal deram origem a uma grade diplomática que é um quadro-resumo onde se indica a função do documento e as partes mais importantes da sua estrutura textual que devem ser evidenciadas para fins documentários para que o documento não perca a sua função definida para determinado fim, que geralmente é a recuperação da informação relevante para objetivos específicos. No caso, a grade diplomática evidencia as partes mais importantes da estrutura textual de artigos de jornais para fins de indexação e a função original do artigo de jornal que é comunicar, ou, informar uma notícia. (representada no APÊNDICE A).

A aplicação do método diplomático confirma que as partes Lead e Episódio (s) são principais. Essas partes da estrutura textual são principais porque, com a exploração

dessas partes do texto, de fato, o indexador encontrará os conceitos que compõem as formas de buscas dos usuários como se vê a seguir:

Usuários (do arquivo do jornal “O Estado de São Paulo”) e tipos de buscas que realizam:

- usuários da redação (jornalistas), com o objetivo de acrescentar novos dados ao assunto de sua matéria, realizam buscas por, autor (Colaborador), data e coluna de publicação.
- Outros usuários

Entre esses encontram-se estudantes (do ginásio e universitários); assessoria de imprensa, órgão público, assistentes do jornal e curiosos. Realizam buscas por assunto, período e título.

③ documentos de mesma função têm a mesma estrutura (tipologia documentária); essas tipologias podem ser divididas em espécies.

Com a proposta de desenvolver o modelo de leitura documentária para artigos de jornal, a ambiência da pesquisa foi o Departamento de Documentação da Editora Abril Cultural (DEDOC). Sendo assim, aplicamos em alguns jornais indexados nessa instituição o método diplomático mencionado acima e confirmamos novamente a ocorrência desses elementos principais da estrutura textual de artigos de jornal no DEDOC e este será um fator que nos direcionará no desenvolvimento do modelo de leitura proposto.

④ documentos podem ter outros usos além de sua função original

Com a realização da atividade de leitura documentária, os jornais são indexados com objetivo de formarem índices de assunto, autor e título, de constituírem uma hemeroteca, ou seja, uma biblioteca de jornais, pastas de recortes com os textos selecionados como importantes.

Consideramos que a aplicação do método diplomático com texto jornalístico veio confirmar a conclusão da pesquisa da Graduação na qual observamos que a exploração da estrutura textual dos itens Lead e Episódio (s) é suficiente para que o indexador encontre os conceitos que formam as estratégias de buscas dos usuários. E como notou-se novamente a ocorrência desse fato na análise da observação da leitura dos indexadores do Departamento de Documentação da Editora Abril, a ser focalizada no capítulo 8, a seguir, destacamos assim, que o modelo de leitura que pretendemos propor deve direcionar-se para a exploração desses itens durante a leitura, visando uma identificação de conceitos

eficientes e uma recuperação da informação compatível com as formas de buscas da comunidade usuária.

Considerando o uso do conhecimento de estrutura textual como estratégia de leitura destacamos a afirmação de Kato (1986) que não existe um pacote pronto e acabado de estratégias para se realizar o ato de ler, pois diante de uma nova leitura o leitor vai buscar sempre novas estratégias para a compreensão, uma vez que o tipo de estratégia que se vai utilizar na leitura depende:

- do grau de maturidade do sujeito;
- do nível de complexidade do texto;
- do grau de conhecimento prévio do assunto tratado;
- do estilo individual do leitor, entre outras variáveis;
- da eficácia de uma determinada estratégia que depende do objetivo de uma determinada leitura.

Portanto o leitor durante a leitura deve explorar o texto apoiando-se nos conhecimento e habilidades que possui ou que pode vir a criar na busca pela compreensão.

É relevante mencionar que com o levantamento das informações que compõem o capítulo 2 atingimos parcialmente mais objetivo dessa pesquisa que é fornecer aos indexadores o *conhecimento de estruturas diversificadas de textos*.

No próximo capítulo, que se encontra organizado em três tópicos, enfocaremos a etapa de identificação de conceitos ocorrida na atividade de leitura documentária durante a indexação, evidenciando que se deve considerar a exploração da estrutura textual, ressaltamos ainda os procedimentos recomendados pelas metodologias da área de Análise Documentária para se identificar conceitos, mencionamos a atividade de indexação de artigos de jornal segundo a literatura e também comentamos outras variáveis que influem direta, ou indiretamente na realização da atividade de indexação como: a política de indexação e a linguagem documentária.

CAPÍTULO 3: IDENTIFICAÇÃO DE CONCEITOS NA INDEXAÇÃO

A identificação de conceitos, como mencionamos anteriormente, acontece durante a leitura documentária do documento. Portanto, trataremos de especificar aqui que a operação de identificar conceitos é um processo da atividade de indexação. A indexação constitui-se em uma forma de tratamento do tema dos documentos pertencente à área de Análise Documentária.

Elucidando melhor essas divisões, denominamos que a Análise Documentária trabalha o conteúdo dos documentos com o objetivo de promover a circulação de informações e de conhecimento por meio da disseminação do conteúdo dos documentos que trata.

O tratamento documentário estabelecido pela Análise Documentária engloba: a classificação do documento dentro de uma das áreas do conhecimento humano, sistematizadas nos códigos de classificação da área de Biblioteconomia e Documentação como Classificação Decimal Universal (CDU) e Classification Decimal Dewey (CDD), em que o documento é atribuída uma notação numérica que permite também localizar esse documento entre outros num determinado ambiente; síntese das idéias tratadas nos documentos por meio da atividade de indexação iniciada primeiramente pela leitura do documento que gerará um resumo do documento com conteúdo informacional equivalente ao documento original e, também, índice de assuntos com os conceitos pertinentes ao documento.

É importante mencionar que o tratamento documentário considera as necessidades dos usuários de um determinado sistema de informação, então podemos dizer que cada sistema de informação desenvolve uma política de tratamento dos seus documentos visando suprir as necessidades informacionais dos seus usuários, e isso será atingido se a política do sistema for estabelecida de tal forma que recupere *“o livro certo, para o usuário certo no momento no certo.”*

Destacamos que os documentos com os quais a Análise Documentária trabalha, e conseqüentemente a indexação, são diversos e já foram classificados por vários pesquisadores da área. Destacamos as tipologias documentárias classificadas por Pinto Molina & Gálvez (1996, p.65) sob três critérios:

1- pelo suporte físico: textuais, gráficos, iconográficos, sonoros, audiovisuais e legíveis por máquina;

- 2- pelo conteúdo informativo: documentos primários, secundários e terciários;
- 3- pela divulgação: publicados, inéditos e restritos.

A atividade de indexação de documentos, segundo Fujita (1999), engloba três fases: a análise realizada através da leitura do documento, durante a qual é feita a leitura e seleção de conceitos; a síntese com a elaboração de resumos; a identificação e seleção de conceitos em termos de uma linguagem documentária para a elaboração de índices, considerando os conhecimentos especializados dos usuários e suas necessidades informacionais que devem estar representadas numa Linguagem, cujo objetivo é padronizar as informações documentárias conforme as terminologias usadas nas diferentes áreas do conhecimento humano.

O processo de indexação, portanto, inicia-se no momento da leitura do documento e é realizado pelo sujeito-leitor caracterizado pelo indexador que pode ser um bibliotecário-indexador (um indivíduo com formação superior no curso de Biblioteconomia e Documentação) ou, um indexador-especialista (indivíduo com formação superior em determinada área do conhecimento a qual pertencem os documentos).

Com relação à formação especializada, indexador-bibliotecário deve ter conhecimento ou noção geral de como está estruturado o conhecimento na área em que trabalha para que possa diferenciar no documento as idéias relevantes das periféricas, que cumprem somente o papel de auxiliar na compreensão do documento ou, de esclarecer trechos do documento que permitam duplo julgamento. Outro fator que auxiliará na busca pela compreensão é o conhecimento da estrutura textual do documento que permitirá identificar partes relevantes e secundárias num determinado documento, bem como os conceitos pertencentes a cada parte. Esses conhecimentos, por parte do indexador-bibliotecário, podem ser adquiridos por meio de leituras sobre o assunto, conversa com especialistas da área em caso de dúvida, observação das formas de buscas dos usuários, e na sua experiência diária em indexação, tornando-se, assim, um indexador-bibliotecário especialista. (KOBASHI, 1994).

Como discurremos no capítulo anterior sobre leitura documentária, o indexador não realiza a leitura do documento na íntegra para fins de indexação, devido ao fato de que isso demandaria tempo e *“ao indexador raramente é dado o luxo de poder ler um documento atentamente do começo ao fim”*, pois sempre deparará com grande quantidade de material para indexar não tendo portanto tempo disponível para realizar tal atividade. (LANCASTER, 1993, p.20).

Este leitor-indexador tem a função de realizar um exame do documento textual, gráfico, iconográfico, sonoro, audiovisual e legível por máquina e tantos outros que possam vir a surgir, com o objetivo de selecionar a (s) idéia (s) consideradas relevantes para os interesses informacionais da comunidade usuária que utilizará determinado documento.

Sobre o momento de identificação de conceitos durante a leitura documentária enfatizamos a opinião de Kobashi (1994), segundo a qual a idéia relevante de um documento se constitui no tema do documento, ou seja, o foco principal. Sendo, portanto, o *“elemento em torno do qual se estrutura a mensagem”*, e caracterizando-se como o núcleo informativo do documento.

Entretanto, o autor, na intenção de discorrer sobre determinado tema, recorre a conhecimentos anteriores ou introduz outros conceitos, algumas idéias que considera essenciais para ilustrar determinados fatos, esses itens exercem determinadas funções no texto como, por exemplo, auxiliar a compreender o texto ou garantir a credibilidade de uma afirmação mencionada. Mas o leitor-indexador não pode deixar se enganar e tratar esses itens como tema. (KOBASHI, 1994).

Portanto, é oportuno destacar a observação de Lancaster (1993), que na indexação de documentos, o indexador não tem a responsabilidade de decidir somente sobre os assuntos tratados nos documentos, mas também de se questionar sobre a real importância desses assuntos selecionados para os usuários ou determinado grupo de usuários de um sistema de informação.

Percebemos, até então, que o indexador durante a leitura documentária busca compreender o documento por meio da identificação e seleção das idéias relevantes tratadas no documentos, confrontada com as necessidades informacionais dos reais usuários desse documento e, para isso, realiza procedimentos.

Para as pesquisadoras Kobashi (1994) e Tálamo (1987) a idéia relevante (s) ou o objetivo principal do documento é o tema do documento. Baseando-se nessa convicção, Tálamo (Op. Cit) definiu a estrutura temática do documento, ou seja, os componentes básicos do tema de um documento sendo esses: Quem? (ser); O que ?(tema); Como? (modo); Onde? (lugar) e Quando? (tempo).

Portanto, segundo Kobashi (Op. Cit), a atividade de leitura documentária incluindo a identificação e seleção do tema pode ser realizada pelo indexador por meio da procura dos mecanismos de respostas desses componentes estruturados por Tálamo.

Lancaster (1993) indica que o indexador, para uma melhor identificação e seleção das idéias relevantes do documento durante a leitura, deve elaborar questões sobre o documento: *De que trata? Por que foi incorporado a nosso acervo? Quais de seus aspectos serão de interesse para nossos usuários?*. Refletindo sobre essas questões, podemos dizer que a resposta ao primeiro questionamento sobre *o que o documento trata?* se refere à identificação das idéias relevantes que podem ser obtidas por meio do questionamento estabelecido por Tálamo (Op. Cit). Quanto à relevância do tema, essa é medida com base no nível de interesse da comunidade usuária pelo conteúdo do documento indexado e o grau desse nível determinará as respostas aos outros questionamentos como: *Por que foi incorporado a nosso acervo? e Quais de seus aspectos serão de interesse para nossos usuários?*.

Uma outra metodologia para identificação de conceitos é a análise conceitual do sistema de indexação PRECIS (Preserved Context Index System) pesquisado por Fujita (1989). A análise conceitual do sistema de indexação PRECIS foi estruturada em torno de "*conceitos universais*", recomendando-se para a etapa de identificação de conceitos uma análise conceitual baseada na interrogação do texto. O indexador, durante a leitura, procura identificar: a ação; o objeto da ação, o agente da ação e local onde ocorreu a ação.

O sistema PRECIS atingiu seu auge na indexação durante a década dos anos oitenta e na opinião de Fujita (1999, p.60),

"importantes considerações teóricas podem ser realizadas em torno da metodologia de indexação do PRECIS, porém, o fator tempo é uma das restrições do trabalho do indexador e a operacionalização extremamente sofisticada do PRECIS certamente inviabilizou a adaptação de indexadores no sistema."

Como estamos tratando os momentos de identificação e seleção das idéias relevantes na indexação de documentos, não poderíamos deixar de mencionar que muitos pesquisadores como Kobashi (1994), Van Dijk & Kintsch (1978), Farrow (1991), Cintra (1987), Fujita (1999), entre outros, recomendam que o leitor identifique a estrutura textual da tipologia documentária com a qual trabalha, para que possa ter conhecimento das partes mais significativas de determinada estrutura textual, onde o autor trabalha as idéias relevantes do documento e assim, localize-as mais rapidamente, resultando numa leitura documentária mais rápida do documento. Considerando que o indexador trabalha com limite de tempo devido à grande quantidade de material que precisa indexar, o

conhecimento e a utilização da estrutura textual na indexação é fator favorável para o objetivo documentário de identificar as idéias relevantes dos documentos.

Existe um modelo de processamento de texto elaborado pelos lingüistas Van Dijk & Kintsch (1978), citado por Farrow (1991), que estabelece as estruturas semânticas gerais do texto em diferentes níveis hierárquicos por meio de várias operações de redução semântica de informação. Essas operações visam obter a macroestrutura global do texto, ou a parcial de uma passagem do texto, e ambas constituem o tema (s) do texto.

A opinião de Farrow (1991, p. 245) sobre esse modelo é que “(...) *além de ser um elemento essencial na compreensão geral do texto, as macroestruturas fornecem ao indexador a base para decidir que partes do texto indexar (...).*” (Tradução livre).

O modelo de processamento de texto de Van Dijk & Kintsch (Op. Cit) estabelece a macroestrutura do texto por meio da redução das informações que podem ser obtidas com a aplicação de quatro regras denominadas de macrorregras cuja função é governar a redução das informações contidas num texto. Mas, há também uma regra geral que estabelece as condições nas quais a redução não deve ocorrer.

As macrorregras que possibilitam a compreensão geral do texto com a identificação do seu tema (macroestrutura) e, segundo Farrow (1996, p.245), ao indexador, explorar a estrutura textual do texto são:

- 1- *Supressão (deletion) fraca: a informação é simplesmente deixada de fora. Somente informação acidental deve ser suprimida, isto é, informação que pode ser omitida sem mudar o significado ou influenciar a interpretação das sentenças subsequentes do discurso;*
- 2- *Supressão forte: a supressão de informação localmente relevante. Aqui, a informação que é suprimida especifica associações normais ou esperadas;*
- 3- *Zero: nenhuma redução ocorre. Todas as informações são consideradas relevantes, e são admitidas diretamente na macroestrutura*.*
- 4- *Generalização: em que vários objetos ou propriedades da mesma classe superordenados são referidos, globalmente com o nome da classe superordenada*
- 5- *Construção, ou a combinação ou integração de informação que denota propriedades essenciais, causas, componentes, conseqüências, etc, de um fato de mais alto nível. Isto constitui “informação implícita”.*

Percebemos, portanto, que durante a indexação dos documentos, o momento de identificação e seleção das idéias relevantes depende do conhecimento do indexador da

* Essa categoria foi estabelecida por Van Dijk & Kintsch (op. Cit) no modelo de 1983.

terminologia de determinada área de assunto que pretende indexar. Esse conhecimento o indexador obtém por ser especialista na área ou simplesmente por ser um bibliotecário que adquiriu conhecimentos tanto por meio de leitura das obras publicadas na área ou por experiência profissional, realizando a atividade de indexação. Com isso, será capaz de identificar o que é essencial e secundário no conteúdo de um documento. Também utilizará seu conhecimento de estrutura textual para identificar as idéias relevantes, mais rapidamente, e assim conhecer quais as partes do documento que necessitam uma maior atenção e quais os elementos pertencentes a essas partes. A identificação e seleção de conceitos das idéias relevantes dependerão, inclusive do conhecimento das necessidades de informação da comunidade usuária para que se possa estabelecer uma concordância com as formas de buscas dos usuários.

Quanto ao uso das macrorregras propostas pelo modelo de Van Dijk & Kintsch (1978) na indexação, a opinião de Farrow (1996, p. 245; 247) é que, embora a teoria pareça ser realmente compatível com as atividades realizadas na indexação, ainda não se pode afirmar que sejam totalmente viáveis porque não foram testadas pelos indexadores. Para usá-las, o leitor-indexador teria que realizar a leitura em ritmo normal lendo de 200 a 300 palavras por minuto. Isto é viável para os indexadores que indexam livros fazendo leitura mais profunda da obra, mas para indexadores que indexam outras tipologias documentárias não é viável, porque o indexador habilidoso pode realizar a leitura documentária com fins de compreensão geral do texto, atingindo até 600 palavras por minutos porque não lê letra por letra, palavra por palavra, e salta trechos cuja informação é conhecida. O uso das macrorregras exige um conhecimento prévio sobre o assunto porque a informação apresentada no texto está sendo constantemente comparada com o conhecimento armazenado na memória de longo termo, e geralmente nos sistemas de informações, com exceção dos centros de informações especializados, o indexador-bibliotecário não é especialista (formação superior em área especializada, ou conhecimento especializado adquirido com a prática de indexação).

Uma outra metodologia para se identificar os conceitos que representam o tema tratado no documento é enfocada por Farrow (1996, p. 7) que menciona o uso da análise proposicional (PA) derivada da lógica formal. A análise proposicional consiste em *descrever um texto sentença por sentença e oração por oração para identificar as proposições. Considerando que uma proposição contém um predicado (um conceito de relação, muito comumente um verbo) e um ou mais argumentos (frequentemente dois: o*

sujeito e o objeto do verbo) o autor acima sugere aplicar conjuntamente PA e macrorregras elaboradas por Van Dijk & Kintsch (1983) para identificar macroestrutura do texto, isto é o tema.

Segundo Farrow (1996), o objetivo da PA é fazer uma representação das proposições, a mais simples possível, embora mantendo o significado subjacente delas. A metodologia da PA mais comumente usada para atingir esse objetivo consistem:

- 1- Identificação do verbo e sua transformação no infinitivo presente ativo;
- 2- Nomes (substantivos) são reduzidos à forma singular e os artigos são eliminados.

Conforme Farrow (1996, p.13), os lingüistas Van Dijk & Kintsch (1983) concluíram que o modelo acima citado de 1978 era super-rígido *“dependendo basicamente da leitura completa do texto a ser indexado”* e propuseram algumas modificações, surgindo o modelo de 1983 formado por uma

“construção mais versátil que tenta integrar velhas e novas informações podendo-se dizer que é orientado para o processo onde as macrorregras são complementadas pelas macroestratégias que são flexíveis e heurísticas, por meio das quais um leitor pode ser capaz de supor, ou adivinhar o tópico de uma passagem de um texto antes de chegar ao fim do parágrafo, ou capítulo, ou livro”. (Tradução Livre).

Portanto, Farrow (1996) destaca que a identificação das macroestruturas (tema) do texto consiste na aplicação de uma gama de macroestratégias contextuais e textuais. Sendo que as textuais envolvem *“conhecimento de estrutura textual e dicas perceptivas e estão relacionadas com a atividade de indexação”*. E as macroestratégias contextuais *“são concernentes ao processamento conceitual”*.

Farrow indica para os indexadores os principais grupos de estratégias contextuais:

- *Situacionais: os perfis dos usuários para os quais o índice de assunto é dirigido, ou pretendido;*
- *Profissionais: a aplicação de padrões profissionais como a ISO 5963 de 1985, BS3700 de 1988;*
- *Externas: restrições quanto ao tamanho, conteúdo, arranjo, etc do índice impostas pelos editores, pelo sistema de indexação usado, o tempo disponível para completar a tarefa, etc.*

Para Farrow (1996), o que forma o suporte das estratégias contextuais é o conhecimento de mundo do indexador, composto de habilidade de leitura e escrita, conhecimento contextual geral e conhecimento contextual específico.

A análise proposicional (PA) e as macrorregras, conforme Farrow (1996), mesmo com o auxílio das macroestratégias, ainda são inviáveis para os fins de indexação, porque as macroestratégias quando aplicadas a um texto por dois leitores diferentes resultam numa variedade de reduções do mesmo texto devido ao conhecimento de mundo e julgamento particular de cada leitor, e a PA demanda um trabalho intensivo sendo melhor aplicada às passagens curtas de texto ou fala, se prestando tipicamente para analisar o conteúdo de uma investigação, ou para medir a densidade de informação de um texto.

A possibilidade de aplicação de ambas na indexação de *um texto longo seria interminavelmente tedioso, ou chato*. (Farrow, 1996). Portanto, a PA como as macrorregras são um instrumento moroso e ineficiente para reduzir um texto a uma afirmação principal, ou tema, e se pretender gerar entradas de índices de assunto para o documento reduzido serão necessárias outras estratégias diferentes. (FARROW, 1996)

É importante ainda mencionar que de acordo com Ginez de Lara (1993), as variáveis que influem no momento da identificação e seleção de idéias relevantes, além do conhecimento prévio do assunto pelo indexador é a política de indexação da instituição na qual o indexador trabalha, pois a instituição estabelece a comunidade usuária que pretende servir, determinando, assim, o perfil desses usuários. Por exemplo, se a instituição é considerada especialista em determinada área do conhecimento humano, certamente a sua comunidade usuária será constituída de usuários específicos, e isso exigirá do indexador que realize a indexação dos documentos considerando a especificidade da sua comunidade usuária. Deverá selecionar, assim, as idéias do documento que realmente correspondam às tais especialidades dos usuários e, posteriormente, deverá traduzir adequadamente para uma Linguagem documentária e, quando não possível, deverá estabelecer, ao menos, um controle das idéias selecionadas, de modo a garantir a padronização dos termos de indexação. Dessa forma, poderá fazer com que o grupo de usuários tenha as suas necessidades ou interesses informacionais garantidos na boa qualidade e na eficiência da indexação dos documentos que, sem dúvida, são refletidas nos resultados das consultas realizadas pelo usuário ao acervo documentário de determinada instituição por meio de mecanismos de buscas informacionais, índices de assuntos em bases de dados. A boa

qualidade da indexação somente é garantida se o usuário obtém o documento que realmente satisfaz a sua necessidade ou interesse informacional.

Um outro caso seria a instituição que visa atender a uma comunidade usuária com conhecimentos gerais, ou seja, que engloba indivíduos das mais variadas áreas do conhecimento humano, como uma biblioteca pública. Nesse caso, o indexador deverá direcionar a indexação dos documentos para a seleção das idéias, considerando um nível mais genérico para a definição das idéias tratadas no documento, ou seja, realizar uma indexação exaustiva em que se pretende selecionar a idéia global do texto considerada relevante para garantir os interesses informacionais dos indivíduos pertencentes a cada uma das áreas de conhecimento existente na instituição, como se fossem uma única categoria de usuários. Posteriormente, o indexador traduzirá para uma linguagem documentária as idéias selecionadas. Quando não é possível a tradução, mantêm-se as idéias selecionadas e estabelece-se uma listagem controlada com os termos adotados, para os usuários as recuperarem e satisfazerem seus anseios por informações.

A determinação do perfil do usuário que a instituição quer beneficiar direcionará a seleção das idéias do documento e exigirá do indexador, no caso da instituição especializada, que selecione na indexação as idéias específicas tratadas no documento correlacionadas com as especialidades de área na qual o usuário se insere. No caso de uma instituição onde os usuários pertencem às mais variadas áreas do conhecimento humano, o indexador deverá indexar o documento selecionando as idéias que o documento trata considerando o seu conteúdo global, porque nessa situação, é mais difícil conhecer as necessidades informacionais específicas de todos os grupos de usuários, para que as idéias selecionadas venham interessar a qualquer uma das classes de usuários, pois o indexador, dessa forma, garante que o documento não seja recuperado somente por um grupo de usuário.

Percebemos que a identificação de conceitos está vinculada a objetivos institucionais variados que produzirão pontos de vistas diferentes para se identificar os conceitos nos documentos porque almejam resultados distintos. Dentro desse contexto, podemos considerar a existência de diferentes concepções de análise, que em determinadas situações de leitura (considerando que alguns textos são mais difíceis que outros) *afetam o desempenho do indexador enquanto leitor*. E Albretchtsen (1993, p.220) classifica os diferentes pontos de vistas em três tipos de concepções:

- concepção simplista: considera os assuntos como entidades objetivas absolutas, que podem derivar de uma abstração lingüística do documento ou de somas usando métodos estatísticos de indexação. De acordo com essa concepção a indexação pode ser totalmente automatizada;
- concepção orientada para o conteúdo: envolve uma interpretação do conteúdo do documento que vai além dos limites da estrutura superficial léxica e gramatical. A análise de assunto do conteúdo de documentos envolve identificação de tópicos ou assuntos que não estão explicitamente colocados na estrutura textual superficial do documento, mas que são facilmente percebidos por um indexador humano. Envolve, portanto, uma abstração indireta do documento.
- Concepção orientada pela demanda: considera o assunto como instrumento para transferência de conhecimento, portanto, direcionado para uma finalidade pragmática de informação e conhecimento. Conforme esta concepção, documentos são criados para comunicação do conhecimento, e assuntos devem, portanto, ser ajustados para funcionar como instrumentos de mediação e transmissão desse conhecimento para qualquer pessoa interessada. Dessa forma, quando o indexador analisa um documento não concentra-se em representar ou resumir a informação explícita ou implícita, mas questiona-se: como eu poderia tornar esse conteúdo ou parte dele, visível para o usuário potencial? Que termos deverei utilizar para levar esse conhecimento até o leitor interessado?

De acordo com Naves (1996), as duas últimas concepções, orientada pelo conteúdo e orientada pela demanda, se complementam. E ainda a concepção orientada para a demanda é realizada após a análise porque o objetivo explícito é a tradução dos conceitos extraídos do documento para os termos de uma linguagem de indexação tendo em vista os interesses dos usuários porque como pudemos mencionar na citação acima: “(...) assuntos devem, portanto, ser ajustados para funcionar como instrumentos de mediação e transmissão desse conhecimento para qualquer pessoa interessada (...)”, ou seja, a identificação de conceitos deve ser feita de tal forma que os seus produtos (resumos e índices de assunto) venham a funcionar como “elo” entre o documento e o usuário. Sendo assim, as observações de Fujita (2001) sobre a leitura de indexadores em centros referenciais especializados orientam que a identificação deve ser feita considerando o conteúdo do documento e que, nesse mesmo instante, a seleção de conceitos também deve ser realizada entretanto considerando as necessidades informacionais dos usuários, podemos dizer que são operações que se complementam e ocorrem num mesmo momento durante a leitura. (Grifo nosso)

Dentro dessa classificação de concepções podemos destacar que as visões de Farrow (1996), Tálamo (1987), Kobashi (1994), Fujita (1998), Van Dijk & Kintsch (1983) e Cintra (1987) quanto à identificação de conceitos está direcionada para a concepção orientada para o conteúdo porque prevê análise do texto por meio de questionamento e exploração da estrutura textual. E enfocando a concepção orientada para a demanda, destacamos novamente Farrow (Op. Cit) e Kobashi (1994).

Portanto, considerando as diferentes concepções sobre identificação de conceitos, o estabelecimento de uma política de indexação se faz necessário dentro de qualquer sistema de informação, e encontramos na revisão da literatura as diretrizes estabelecidas por Carneiro (1985) para se implementar uma política de indexação e seus elementos, e procuraremos identificar estes elementos durante a nossa pesquisa de campo no DEDOC da editora Abril e no arquivo do jornal "O Estado de São Paulo" descritos nos capítulos seguintes para verificar qual é a política de indexação dessas instituições.

Segundo Carneiro (Op. Cit), os requisitos para o estabelecimento de uma política de indexação devem abranger: 1- a identificação da organização a qual estará vinculado o sistema de indexação, porque isto permitirá conhecer os objetivos da instituição, e sobre estes, determinar o tipo de serviço que será implantado e 2- a identificação da clientela a que se destina o sistema, os recursos humanos, materiais e financeiros. Destacamos que a identificação do cliente possibilitará uma adequada política de seleção e aquisição de documentos para a formação do acervo, bem como a área de assunto de maior interesse para a instituição, e com base nos objetivos da instituição e no perfil de seus usuários, se estabelecerá, na medida do possível, o percentual necessário quanto a recursos financeiros, bem como os recursos humanos e materiais.

Com relação aos elementos de uma política de indexação devem-se considerar:

- a *cobertura de assunto* abrangida pelo sistema de informação que com certeza será baseada no estudo do usuário potencial do sistema diferenciando assuntos centrais e periféricos;
- o estabelecimento de uma política de *seleção e aquisição dos documentos-fonte* que deve considerar os interesses dos usuários para poder determinar a extensão da cobertura de assunto do sistema e a qualidade do documentos nessas áreas de assunto;
- como se realizará o *processo de indexação considerando níveis de exaustividade e especificidade* que serão determinados com base nos objetivos do sistema de informação que estão relacionados com a clientela a qual se destina e, conseqüentemente, se determinará se a indexação considerará os assuntos em sua conceituação geral (exaustividade) ou específica (especificidade);
- a *escolha da linguagem* nesse momento deve observar a clientela do sistema, o processo de indexação (exaustividade e especificidade).

Ainda com relação às variáveis que influem na identificação e seleção de conceitos, Ginez de Lara (1993) refere-se ainda às pertencentes ao próprio indexador: sua ideologia, sua visão de mundo e sua formação pessoal, que podem estar norteadando a seleção e descarte de idéias na leitura do documento.

Não podemos deixar de mencionar que pode acontecer de o indexador deixar de selecionar determinado tema por possuir preconceitos individuais ou por obter uma visão de mundo diferente do autor em relação a esse tema e isso, com certeza, afetará a qualidade da indexação. Na medida que o indexador deixa de considerar a relevância desse tema para os usuários faz com que o documento indexado talvez nunca seja recuperado, e se vier a ser, talvez seja descartado, principalmente, quando o usuário se prende somente às pistas de conteúdo oferecidas pelo indexador.

Entretanto a clareza dos objetivos da instituição e das necessidades dos usuários tornam menor a influência de fatores pessoais na indexação dos documentos.

É importante também destacar que existem variáveis pertencentes ao próprio documento que afetam a seleção do tema pelo indexador como: *idioma no qual o documento foi escrito; a capacidade do autor de discorrer com clareza as suas idéias; a legibilidade do texto.*

Elucidando essas variáveis, ressaltamos que, muitas vezes, o idioma no qual o documento se apresenta pode ser uma barreira para indexador se o mesmo não tiver conhecimento desse idioma. Quanto à produção do texto escrito, o indexador pode encontrar no documento idéias implícitas que foram consideradas pelo autor como subentendidas, mas pode acontecer de o autor expressar “alho” como subentendido e o indexador interpretar como “bugalho” causando assim falha, tanto na compreensão do texto por parte do indexador, como também na seleção do tema do documento, se o subentendido não compreendido for um “elo” de compreensão entre as idéias tratadas no texto pelo autor. Com relação à legibilidade do documento destacamos que textos com tipo impresso muito reduzido, pouco espaçamento entre as linhas do texto, a cor da tinta na escrita do texto, etc podem dificultar a leitura do indexador, causando-lhe cansaço na realização da leitura e, conseqüentemente, provocando sua desatenção, fazendo com que, muitas vezes, as idéias importantes do documento lhe passem despercebidas. (GINEZ DE LARA, 1993)

* Lancaster (1993) destaca que devem-se considerar também outras variáveis pertencentes ao documento que podem influenciar na seleção das idéias do documento, quando admitimos que

“ existem assuntos que são mais difíceis que outros; que o documento que aborda teoria sobre determinado tema é mais difícil que outro que exemplifica em seu conteúdo uma questão prática; título do documento tanto pode ser fiel ao conteúdo do documento como enganador; e a existência ou não de resumo que reflita o conteúdo do documento.”

A identificação e seleção de idéias num documento pelo indexador dependerá portanto do conhecimento individual do indexador da área de conhecimento que o documento aborda para que possa saber identificar precisamente o tema do documento, separando, assim, informações essenciais das acessórias. Destacamos ainda que a determinação do tema não implica na quantidade de informações que o documento apresenta sobre ele, pois no processo de indexação o que vai determinar a escolha das idéias relevantes é a política de indexação da instituição traçada conforme a comunidade usuária que pretende servir, o grau de conhecimento do indexador quanto aos interesses dos usuários pelo tema a ser indexado, e a quantidade de informações registradas em publicações da área sobre o conceito. Isso não quer dizer que temas recentes não deverão ser indexados, mas é preciso que o indexador defina a real importância de estar selecionando um determinado conceito para o tema recente pois, com o passar do tempo, essa escolha pelo conceito selecionado será justificada se vierem a ser publicados outros documentos sobre o tema cuja seleção de conceitos direciona-se novamente ao conceito anteriormente selecionado. Por isso o indexador deve, durante a seleção das idéias recentes do documento, tomar cuidado para não selecionar conceitos que ainda não foram sedimentados pela área do conhecimento ou correrá o risco, de serem considerados como modismo de uma época. (LANCASTER, 1993).

A tradução dos conceitos selecionados durante a leitura para uma Linguagem Documentária é a próxima etapa a ser analisada. As Linguagens Documentárias podem ser ou não vocabulários controlados de determinada área do conhecimento que apresentam estruturada a terminologia usada na área, ou seja, a “arquitetura conceitual” da área em questão. Para a determinação dos termos que deverão compor as linguagens, os profissionais responsáveis pela sua elaboração consultam fontes de conhecimento na área como: linguagem de especialidade (enciclopédias especializadas, glossários); Linguagem de uso corrente e a Linguagem natural (a forma escrita em que encontramos os documentos

nas publicações bibliográficas da área). A elaboração dessas linguagens é de responsabilidade de uma equipe interdisciplinar composta por profissionais especialistas na área e profissionais bibliotecários.

Quanto à sua função, as Linguagens Documentárias pretendem “*normalizar os conceitos da área, controlar seu uso e viabilizar a interface documentação-usuário*” porque a definição dos conceitos que compõem a linguagem documentária, deve considerar as necessidades informacionais da comunidade usuária pertencente a determinada instituição, e também o linguajar regional dessa comunidade no momento da tradução do conteúdo selecionado do documento para uma linguagem de indexação. (KOBASHI, 1994, p.65). Para atingir essa função, a Linguagem Documentária é formada de vocabulário e sintaxe.

Ao definirmos vocabulário, podemos dizer que se refere aos conceitos responsáveis por representar o conteúdo selecionado dos documentos. A sintaxe é um conjunto de regras utilizadas para estabelecer a combinação entre os termos selecionados de determinado documento na indexação.

Destacamos que as Linguagens Documentárias podem ser classificadas em duas categorias: hierárquicas representadas pelos Sistemas de Classificação Bibliográfica como por exemplo a Classification Decimal Dewey (CDD) e a Classificação Decimal Universal (CDU); alfabéticas, representadas pelos Tesouros e Listas de cabeçalhos de Assuntos exemplificando aqui a Library of Congress Subject Heading e vários tesouros especializados.

Os Sistemas de Classificação Bibliográfica foram os primeiros instrumentos de indexação elaborados; sua existência é datada do século XIX. Cumprem a dupla função: de classificar os documentos em determinada área do conhecimento e possibilitar a recuperação do documento. O documento, ao ser classificado, recebe uma notação ou símbolo constituído de números, letras, ou sinais que permitem a sua localização no acervo.

Já os Tesouros são constituídos de descritores que representam os conceitos de determinada área e podem ou não permitir relacionamentos semânticos, hierárquicos e associativos entre os termos, permitindo, assim, que se represente mais adequadamente o conteúdo do documento. Destacamos que a estrutura de organização dos Tesouros evidencia primeiramente a conceituação atribuída a determinado termo, e posteriormente,

estabelece as relações existentes entre ambos. Os tesouros são caracterizados como sistemas pós-coordenados.

As listas de Cabeçalhos de Assunto possuem descritores de determinada área organizados alfabeticamente, não obedecendo a uma ordem hierárquica. Os Cabeçalhos de Assunto consideram em sua estrutura termos sinônimos, quase sinônimos e homógrafos, e estabelecem regras tais como: forma correta de entrada do descritor, uso de singular/plural, uso de abreviaturas e de termos traduzidos. São diferentes dos Sistemas de Classificação Bibliográfica e dos Tesouros porque são caracterizados como sistema pré-coordenado.

Percebemos então, que a linguagem auxilia o indexador quanto à conceituação das idéias que quer representar, pois uma Linguagem Documentária deve apresentar: o vocabulário de determinada área do conhecimento, as relações hierárquicas, associativas e de equivalência existente entre os conceitos da área e quando necessário for, permitir que se combine conceitos para representar adequadamente o conteúdo do documento. Apresentando essa estrutura, a linguagem ajuda o indexador a se familiarizar com a estruturação do conhecimento da área, uma vez que, a maioria desses profissionais não é caracterizada como especialistas, isto é, indexadores com formação superior na área em que realizam a indexação, ou indexadores que possuem conhecimento especializado na área adquirido com a experiência diária de indexação, ou cursos.

Na revisão de literatura sobre linguagem documentária encontramos a pesquisa de Pinto (1985, 173-177) onde se afirma que existem dois tipos de relações entre os termos em uma linguagem documentária sendo:

1- *Relações Sintáticas: originam-se da necessidade de se poder recuperar a intersecção entre duas ou mais classes de conceitos distintos. Constituem-se da associação de termos para representar assuntos compostos, como por exemplo: "fundição do alumínio" e não "fundição" e "alumínio."*

2- *Relações Semânticas: devem ser controladas ou mostradas nas Linguagens de indexação com o objetivo de se indicarem os termos alternativos ou substitutivos de indexação e busca. E podem ser de três tipos:*

1- *Equivalência: fornece termos substitutivos de indexação e busca, através de:*

- *Sinônimos e antônimos*

Ex: diminutivos x aumentativos

- *Quase sinônimos*

Ex: datiloscopia x impressão digital

- *Grafias diferentes*

Ex: contato x contacto

- *Abreviaturas e acrônimos*

Ex: ONU – Organização das Nações Unidas

- *Traduções*

Ex: recall x recuperação

- *Ordem de citação*

Ex: automação de bibliotecas x bibliotecas, automação de

Ressaltamos aqui que a Linguagem Documentária indica a escolha de um único termo para representar o conceito, pois, a garantia de recuperação da linguagem está na precisão do vocabulário.

2- **Hierárquicas:** permitem a busca por termos alternativos. Expressam o posicionamento de um conceito em relação a termos mais amplos ou mais específicos.

São as relações de subordinação entre os termos como:

Ex: relação todo-parte

Livro x folha

Ex: relação gênero-espécie

Gato x mamífero

3- **Afinitivas ou Associativas:** revelam termos indexadores que desencadeiam uma ampliação da estratégia de busca. Mostra termos afins, sem expressar uma noção de hierarquia. As relações afinitivas possibilitam a busca de termos alternativos em relação à:

- *Coordenação*

Ex: portas x janelas

- *Genética*

Ex: pai x filho

- *Concorrente*

Ex: educação x ensino

- *Causa e efeito*

Ex: ensino x aprendizagem

- *Instrumental*

Ex: medicina preventiva x vacinação

- *Material*

Ex: esquadrias x alumínio

- *Similaridade de processo*
Ex: classificação x indexação

Destacamos aqui que o indexador no momento da leitura documentária (constatando-se que realiza primeiramente identificação e posteriormente a seleção entretanto num único momento) deve consultar a linguagem somente após a seleção das idéias do documento, porque pode ocorrer do conceito primeiramente identificado e constatado na linguagem não refletir adequadamente o conteúdo do texto, deixando o indexador de representar o conteúdo abordado no texto seja por meio de combinação de termos presentes na linguagem, ou de criação de novos termos. Considerando a especialidade da área do conhecimento à qual a linguagem se destina, e dependendo da quantidade de conceitos que a linguagem não permite representar, o indexador poderá deduzir se a linguagem está desatualizada e, se constatado, o indexador deve avisar as instituições responsáveis para atualizar tais linguagens a fim que as mesmas venham a se tornar mais adequadas e eficientes.

Percebemos que é preciso fazer uma avaliação da Linguagem documentária para constatar a eficácia na representação ou tradução dos termos selecionados na indexação. Na opinião de Cesarino & Pinto (1978, p. 272), a necessidade de controlar o vocabulário de uma determinada área é decorrente de alguns fatores:

- a- *Fatores humanos: diferenças de cultura, de experiência dos autores e domínio de terminologia entre indexadores, autores e usuários;*
- b- *Fatores referentes à própria Linguagem natural: vários termos podem designar o mesmo conceito (sinônimos); um mesmo termo pode designar conceitos diferentes (homógrafos); o sentido dos termos se altera pela posição que ocupa na frase (sintaxe);*
- c- *Fatores hierárquicos: um conceito implica em outros mais amplos e/ou mais restritos.*

Portanto, a avaliação da linguagem documentária que auxilia o indexador no momento de tradução dos termos é necessária, porque uma linguagem eficaz permite traduzir os termos selecionados sem perder a idéia principal tratada no texto original.

A preocupação em avaliar as linguagens documentárias sempre existiu; destacamos que na revisão de literatura encontra-se o trabalho retrospectivo de Piedade (1976) que relata as pesquisas mais conhecidas sobre avaliação de linguagens

documentárias como: o teste da ASTIA, os projetos Cranfield I e II, o teste de Alberystwyth, as experiências de James Carlies e Rober Fetter.

Mais recentemente, Lancaster (1995) propôs uma metodologia para avaliação de linguagens documentárias.

No Brasil, destacamos, entre outros, o estudo de Ginez de Lara (1993).

Evidenciamos também que no grupo de pesquisa “Análise Documentária”, da UNESP – Marília coordenado por Fujita existe a linha de pesquisa de avaliação de linguagem em que se realizam estudos, como as pesquisas já concluídas de iniciação científica de Corrêa (1998) e de Cessel (1998), ambas patrocinadas pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP).

Abordaremos brevemente aqui os critérios da metodologia de avaliação de linguagens propostos por Ginez de Lara (1993) e por Corrêa (1998), destacando que ambas, juntamente com a metodologia de Lancaster (1995), serão enfocadas aqui por termos conhecimento de pesquisas que as aplicaram e obtiveram resultados considerados relevantes para a avaliação das linguagens. É o caso da pesquisa de Cessel (1998). São, portanto, metodologias que devem ser mencionadas, uma vez que pretendemos evidenciar a importância da avaliação das linguagens o que pode ser percebido pelas metodologias recentes elaboradas.

Somando a isso, na pesquisa em que avaliamos as linguagens documentárias do arquivo do jornal “O Estado de São Paulo” e do Departamento de Documentação da Editora Abril, objetivando mencionar brevemente a influência da linguagem na indexação, descrita na sequência, destacamos que optamos por utilizar as metodologias de Ginez de Lara (1993) e de Corrêa (1998). Entretanto, descartamos a metodologia de Lancaster (1995) por ser direcionada a tesauros tidos como bons, ou seja, tesauros que possibilitam qualidade na representação dos termos que por sua vez produz eficácia na recuperação dos documentos indexados. Portanto, não se obterá êxito nos resultados se aplicarmos essa metodologia numa listagem alfabética controlada de termos que é o caso do arquivo do jornal “O Estado de São Paulo” e do Departamento de Documentação da Editora Abril.

Primeiramente, ressaltamos que o pesquisador Lancaster (1995), que se destaca por pesquisar o desempenho de sistemas de recuperação da informação (como, por exemplo, a avaliação que realizou do sistema MEDLARS) afirma que os métodos

utilizados para se avaliar uma linguagem de indexação podem ser de dois tipos: objetivo e subjetivo. Na utilização do método subjetivo a avaliação é baseada apenas nas opiniões dos usuários, enquanto que no método objetivo são adotados critérios que permitem a coleta de dados para análise e diagnóstico.

Destacamos que o método objetivo foi proposto por Lancaster (1995) baseando-se nas pesquisas de Van Dijk, M. (1976) e de Kochen & Tagliacozzo (1968), e que o autor manteve, no momento da elaboração do método objetivo, alguns dos critérios estabelecidos pelos pesquisadores precursores dos critérios. Enfocaremos somente o método objetivo por termos conhecimento da pesquisa de Cessel (1998) que o utilizou e obteve êxito na avaliação das linguagens avaliadas. No método objetivo de Lancaster (1995, p.174), estão inseridos seis critérios para se avaliar uma linguagem, sendo:

1. *A razão de equivalência, que é a razão entre o número de não descritores e descritores, realmente constitui uma medida da riqueza do vocabulário de entrada. Este valor deve ser maior que 1, quer dizer, que haja mais termos de entrada que descritores;*
2. *A razão de reciprocidade, que é o grau em que as relações de TG, TE e TR têm reenvios recíprocos;*
3. *Definição, representada pela equação $(b-a)/b$, onde a é o número de descritores que possivelmente são ambíguos porque não tem notas de aplicação, qualificadores, ou relações hierárquicas que os definam, e b é o número total de descritores do vocabulário;*
4. *Flexibilidade, que é a proporção de palavras nos descritores compostos que aparecem no vocabulário como descritores ou não descritores. Se recomenda um valor de 0,6 ou maior;*
5. *Nível de pré-coordenação, que é o número médio de palavras por descritor. Para os tesouros ingleses, espanhóis e franceses recomenda-se um valor entre 1,5/2,0 e de 1,1/1,2 para os tesouros alemães;*
6. *O tamanho dos grupos de termos (quer dizer, os grupos que formam a lista sistemática do tesouro). Recomenda-se de 30 a 40 termos por grupo. (Tradução Livre).*

Na opinião do próprio Lancaster (1995, p. 174) ressaltam-se algumas observações sobre a aplicação do método objetivo o que vem justificar a nossa atitude para não utilizar o método objetivo em nossa pesquisa cuja justificativa por descartar a aplicação de tal metodologia exibimos abaixo com apresentação de citação na íntegra:

"algumas de estas medidas son ingeniosas, pero otras resultan triviales (razón de reciprocidad) e incluso arbitrarias (la flexibilidad). Los valores recomendados

son bastante arbitrarios. Se establecieron a partir de algunos tesauros considerados como "buenos" y a los que se aplicaron las distintas medidas.
(Grifo nosso)

Desde luego, es imposible evaluar un tesauro si no se hace en condiciones de uso real". (...) (Grifo nosso).

O que gostaríamos de elucidar com a citação acima é que o uso do método objetivo foi descartado de nossa pesquisa na avaliação das linguagens do DEDOC e do arquivo do jornal "O Estado de São Paulo" por ser o mesmo direcionado especificamente para tesauros que são linguagens documentárias organizadas e logicamente estruturadas enquanto que as linguagens a serem avaliadas se constituem de uma listagem alfabética de termos que ressaltamos ser bem organizada mas que está longe de apresentar as características de um tesouro.

Ainda sobre a metodologia proposta por Lancaster (1995) citada acima, segundo Cessel (1998) que a utilizou na avaliação das linguagens DeCS e Inis Thesaurus, evidenciam-se por exemplo, as dificuldades para calcular o nível de pré-coordenação das linguagens. Destacando estas dificuldades evidenciou-se que o método recomenda que para estabelecer o nível de pré-coordenação, deve-se calcular o número de palavras por descritor, entretanto, não fornece maiores detalhes para informar o que podemos considerar como palavras; será que os advérbios podem ser considerados como palavras? e, até mesmo, se determinadas palavras podem ser consideradas como uma única palavra como por exemplo a palavra: EXAMES PRÉ-NUPCIAIS, será que PRÉ- NUPCIAIS deve ser contada como uma única palavra?; a mesma questão aparece também com a palavra inglesa LI-DRIFTEDN JUNCTION DETECTORS onde nos questionamos se podemos considerar LI-DRIFTEDN como palavra única, e ainda será que podemos separar as palavras separadas com hifen? .

Ressaltamos que todas essas dúvidas e outras referentes à aplicação do método deverão ser refletidas e melhores interpretadas provavelmente nas pesquisas referentes à metodologia de avaliações de linguagens que comportem a utilização de tal método, e isto não nos cabe aqui porque as linguagens que avaliamos sendo a do arquivo do jornal "O Estado de São Paulo" e a do Departamento de Documentação da Editora Abril não apresentam as características do tipo de linguagem ao qual o método se destina , isto é tesouro.

Quanto à metodologia de avaliação de linguagens desenvolvida por Ginez de Lara (1993, p. 85-101), que utilizamos em nossa pesquisa consiste da avaliação da estrutura interna das linguagens a partir de cinco critérios preestabelecidos:

1. *Categorias de assunto nas quais o descritor se insere;*
2. *Análise dos termos genéricos;*
3. *Análises dos termos relacionados;*
4. *Lógica de subordinação dos termos específicos, e*
5. *Consistência da relação de equivalência entre os termos.*

Na avaliação da linguagem documentária, é necessário observar também a estrutura externa da linguagem, e destacamos aqui a metodologia de análise formal proposta por Corrêa (1998) para avaliar a estrutura externa de uma linguagem que também utilizamos em nossa avaliação das linguagens do arquivo do jornal e do DEDOC citadas anteriormente. Os parâmetros definidos por Corrêa (Op. Cit) são:

- 1- *Aspectos Gerenciais*
 - a) *Compilação dos termos*
 - b) *Número de termos*
 - c) *Uso de software e/ou programa de computador*
 - d) *Instrução de uso/informações complementares*
 - e) *Utilização de normas para controle de vocabulário*
 - f) *Coleta temática*
- 2- *Estrutura da linguagem*
 - 1- *Apresentação da linguagem*
 - 2- *Termos de indexação*
 - a) *Termos equivalentes em outros idiomas;*
 - b) *Formas dos termos*
 - *substantivos*
 - *frases substantivadas*
 - *advérbios*
 - *verbos*
 - *escolha de formas singular e plural*
 - *homógrafos (ou homônimos)*
 - *siglas*
 - 3- *Relações semânticas e lógicas*
 - a) *relação hierárquica*
 - b) *relação associativa*

Com as breves descrições de alguns critérios para avaliações de linguagens mencionadas acima, esperamos ter destacado que a avaliação de linguagem faz-se necessária e destacamos algumas dentre muitas metodologias que podem ser utilizadas.

Evidenciamos que a metodologia proposta por Ginez de Lara (Op. Cit), e os parâmetros estabelecidos por Corrêa (1998) foram ambos utilizados na pesquisa de Cessel (1998) cujos resultados obtidos demonstraram êxito nas avaliações das linguagens, foi o que justificou a nossa escolha por utilizar essas metodologias em nossa pesquisa para a avaliação das Linguagens do Departamento de documentação da Editora Abril Cultural e do arquivo do jornal "O Estado de São Paulo" descritas adiante.

É oportuno evidenciar que, para cumprir as suas funções de recuperação e satisfação de informação dos usuários, as informações documentárias (índices e resumos) devem apresentar, conforme Kobashi (1994, 103), as propriedades estabelecidas por Grice (1975) já mencionadas no capítulo de leitura:

a) Concisão (princípio de quantidade) o resumo e o índice devem fornecer informação suficiente, não mais do que o necessário. Este aspecto é fundamental para a economia do sistema porque facilita a estocagem e a recuperação de informações. Mas não se deve esquecer da necessidade de fidelidade em conteúdo ao documento original;

b) Pertinência (princípio de qualidade e de relação) o resumo e o índice devem representar o conteúdo do documento tão fielmente quanto possível. Desse modo, não se pode integrar a eles informações que não estejam no texto original. Destacamos que o indexador deve indexar o documento que tem em mão e não que queria que o autor tivesse escrito Lancaster, 1993);

c) Precisão e objetividade (princípios de modo): o resumo e o índice não devem comportar ambigüidade, devendo ser formulados em termos precisos; no caso do resumo : apresentar as informações em ordem adequada e não ser objeto de duplo julgamento, e oferecer a possibilidade do usuário escolher se deseja consultar o documento original.

3.1 INDEXAÇÃO DE DOCUMENTOS

Abordaremos, nesta seção, a indexação de documentos conforme as recomendações das normas metodológicas e dos pesquisadores da área de Análise Documentária evidenciando mais especificadamente os procedimentos para indexar textos jornalísticos.

A atividade de indexação, como vimos, anteriormente, engloba leitura consciente do documento para identificação e seleção do assunto e tradução desses assuntos conforme os termos de uma linguagem de indexação.

Na revisão de literatura, encontramos como primeiro passo dado para se estabelecer uma metodologia de análise do documento para fins de indexação, os Princípios de Indexação estabelecidos pelo UNISIST em 1981 que definem a atividade de indexação como: (...)” *a ação de descrever e identificar um documento de acordo com seu assunto*” (...).

Com relação à operacionalização da indexação, os Princípios do UNISIST indicam dois estágios:

- 1- *determinação do assunto: estabelecimento dos conceitos tratados num documento;*
- 2- *representação de conceitos por termos de uma linguagem de indexação.*

Destacamos que o primeiro estágio segundo o UNISIST é subdividido em outros três estágios: compreensão do conteúdo do documento; identificação dos conceitos que representam esse conteúdo e seleção dos conceitos válidos para recuperação. Após mencionar que o estágio de determinação do assunto está subdividido em outros três estágios, o UNISIST destaca que, na prática, os três estágios se superpõem.

O Projeto Integrado “Leitura em Análise Documentária” da UNESP-Marília enfocou a observação da atividade de leitura dos indexadores por meio da metodologia de Protocolo Verbal. O desenvolvimento do Projeto Integrado possibilitou as pesquisas de Fujita (1998) na Sub-Rede Nacional de Informação em Ciência da Saúde Oral com a indexação do texto de artigos científicos em Odontologia, a pesquisa de Rubi (1999) no Centro de Energia Nuclear (CIN) com artigos científicos em energia nuclear, de Pires (1999) na CENAGRI (Coordenação Geral de Documentação em Agricultura), bem como a pesquisa de Iwashita (1999) que investigou a elaboração de resumos no Centro de Energia Nuclear (CIN) e de Fagundes (1997) que investigou a indexação de jornais no arquivo “O Estado de São Paulo” e ambas pesquisas mencionadas os resultados atingidos demonstraram que os estágios de identificação e seleção de conceitos ocorrem no mesmo instante. Ressalta-se que nessas pesquisas os indexadores identificam considerando o conteúdo e selecionam priorizando as necessidades informacionais dos usuários que visa como resultado da leitura documentária representar o conteúdo do texto e os interesses dos usuários.

Com relação ainda ao estágio de compreensão do conteúdo do documento, o UNISIST primeiramente identifica duas categorias de documentos: os documentos gráficos e não-gráficos. Posteriormente, estabelece que, para fins de indexação, a compreensão do conteúdo dos documentos não-gráficos (livros; monografias; jornais; periódicos; relatórios; teses) pode ser obtida por meio da leitura de partes importantes do texto como: *título; introdução e as primeiras frases de capítulos e parágrafos; ilustrações, tabelas, diagrama e suas explicações; conclusão e palavras ou grupos de palavras sublinhadas ou impressas com tipo diferente*. Entretanto, adverte que os primeiros itens do texto apresentam, geralmente, as intenções do autor, enquanto que as partes finais comunicam o alcance dessas idéias. Portanto, não recomenda a indexação somente com a leitura do título e/ou resumo, porque os títulos podem ser enganosos e ambos título e resumo podem ser inadequados. Assim, somente lendo essas partes textuais o indexador corre o risco de negligenciar informações relevantes do texto.

Ainda quanto aos documentos gráficos, o UNISIST adverte que o ideal seria uma leitura extensiva do documento mas em indexação isso é impraticável pois os indexadores sempre estão sufocados com o limite de tempo despendido para a indexação dos documentos, uma vez que há uma grande quantidade de documentos à indexar.

Uma outra categoria de documento estabelecida pelo UNISIST são os documentos não-gráficos que englobam: documentos audio-visuais, visuais ou sonoros. Nesse caso, o UNISIST (1981, p. 87) destaca que “(...) *pode não ser possível ao indexador um exame minucioso, profundo, desse tipo de material, de maneira que muitas vezes ao contrário dos procedimentos utilizados na indexação dos documentos gráficos nessa espécie de documento a indexação a partir do título ou da sinopse torna-se uma necessidade.*”

Com a compreensão do conteúdo, passaremos ao estágio de identificação de conceitos, o segundo estágio dentro da fase de determinação do assunto. Nessa fase, o UNISIST recomenda que o indexador aborde o documento de forma mais lógica a fim de selecionar os conceitos que melhor representem seu conteúdo. Para isso sugere que o indexador identifique por meio de questionamento algumas categorias existentes na área coberta pelo documento, como por exemplo: *o fenômeno; o processo; as propriedades; as operações; o material; o equipamento, etc...*

Após a fase de identificação de conceitos por meio do questionamento, faz-se necessária a seleção dos conceitos. Conforme o UNISIST, esta deve ser traçada tendo em vista os objetivos para os quais as informações são indexadas, que variam de instituição

para instituição. E, às vezes, para atingir um determinado objetivo, nem todos os conceitos identificados serão necessariamente selecionados.

Com a seleção de conceitos, o UNISIST alerta para a questão da exaustividade e especificidade durante a seleção de conceitos na indexação. Com relação à exaustividade, alerta que o indexador deve *“identificar todos os conceitos de um documento que têm um valor potencial para os usuários de um sistema de informação”* ou seja, o indexador que trabalha com usuários especialistas e não-especialistas da área, deve selecionar os conceitos que interessem aos dois perfis de usuários e não indexar o documento somente sobre o seu enfoque genérico, ou, somente sobre o específico.

Destaca-se também que o indexador não pode se deter em limite de descritores, mas selecionar todos os conceitos pertencentes ao documento que interessam aos usuários do sistema e, muitas vezes, incluir entre os conceitos indexados aqueles que foram abordados implicitamente no documento, mas que servem para conceitualizar uma determinada idéia no contexto apropriado.

Quanto à compatibilidade da linguagem documentária, no momento de tradução dos conceitos, o indexador deve ter consciência de que, se necessário, deve *“modificar tanto os instrumentos de indexação quanto os procedimentos com base no “feedback” de questões,”* pois as linguagens, em muitos casos, não permitem especificar adequadamente o conceito abordado no documento.

Um outro fator que influencia a seleção de conceitos é a questão da especificidade; o UNISIST ressalta que é praticamente uma regra identificar exatamente o conceito específico tratado nos documentos. E que os conceitos genéricos devem ser selecionados dependendo dos objetivos do sistema de informações como por exemplo as necessidades informacionais dos usuários que podem ser por informações mais genéricas, ou, mais específicas, ou, em muitos casos, até ambas (genérica e específica).

É importante mencionar que um conceito pode ter seu nível de especificidade elevado no texto por influência da importância dada pelo autor para determinado conceito. Em casos em que o indexador concluir que uma idéia não foi completamente desenvolvida ou foi, mas de forma superficial pelo autor, deve proceder a indexação num nível mais genérico do conceito.

Para o momento de tradução dos conceitos em uma linguagem documentária: tesouro, lista de cabeçalho de assunto e sistemas de classificação, o UNISIST (1981, p.90)

recomenda que o indexador obtenha um conhecimento profundo sobre essas linguagens para garantir “a organização de conceitos de uma forma útil e acessível” e que, ao mesmo tempo, esteja familiarizado com os procedimentos e regras desses instrumentos, não esquecendo, porém, que podem limitar a prática de seleção de conceitos. É o caso, por exemplo, dos sistemas de classificações que abordam os conceitos em nível genérico impedindo o indexador de indexar um aspecto específico do conceito. Ocorrendo isto o “(...) indexador deve usar descritores mais genéricos, e os novos conceitos ficam como candidatos para uma nova edição”. (UNISIST, 1981, p. 91).

Acreditamos que essa orientação do UNISIST deixa transparecer uma imagem do indexador “preso” a imposições da linguagem documentária e inconsciente da sua missão de representar o conteúdo documental, conforme as necessidades dos usuários que determinarão se um conceito abordado no documento será mais preciso se for indexado de forma genérica, ou específica, ou ambas.

Ao mesmo tempo, essa orientação chega a ser contraditória quando o UNISIST (1981, p.92-93, destacadas abaixo) descreve que o controle de qualidade da indexação depende das qualificações do indexador e da qualidade das linguagens documentárias. Posteriormente, menciona que quanto às qualidades do indexador, este deve ser de preferência um especialista na área em que trabalha. Com isso concordamos, também, destacando ainda que essa especialidade pode ser obtida com a realização de curso superior na área do conhecimento que indexa documentos, ou obtida por anos de prática na atividade de indexação em um sistema especializado tornando-se, assim, um indexador especialista na área. Para o UNISIST, o indexador deve compreender os termos encontrados nos documentos, as regras e os procedimentos do sistema de indexação em que trabalha. O controle da indexação será atingido mais eficientemente se o indexador mantiver contato com os usuários, pois “juntos podem determinar se alguns descritores produzem falsas recuperações ou criam problemas no estágio de saída.”

Ressalta quanto a linguagem que

“a qualidade de indexação depende também de certas propriedades dos métodos e procedimentos de indexação. É essencial que um índice seja flexível bastante para receber novos desenvolvimentos da terminologia, como também atender a novas necessidades dos usuários: isto é, deve permitir atualização freqüente.”

Posteriormente, baseada nos princípios do UNISIST, surge a Norma ISO 5963 (1985) que recomenda ao indexador, durante a leitura, examinar o documento considerando as partes mais relevante do documento, alertando assim que a leitura do documento todo é impraticável, devendo o indexador proceder ao exame do documento a partir da leitura do título, resumo, se houver, sumário, introdução, as frases e parágrafos de abertura de capítulos, as conclusões, ilustrações, gráficos, tabelas e respectivas legendas, palavras ou grupos de palavras que apareçam sublinhados ou impressos com tipos diferentes.

Porém, a Norma ISO não especifica que documentos apresentam essas características deixando subentender que a metodologia se aplica a todo e qualquer tipo de documento que se deseja indexar.

A Norma alerta, na página dois, que não é recomendável indexar somente com a leitura do título, pois podem ser enganosos e ambíguos, e nem somente com base no resumo, pois esse pode ser inadequado, e ambos não são uma fonte confiável para informação que o indexador requer.

Para a Norma ISO, tanto a análise como transcrição, ou tradução de conceitos devem ser realizadas com o auxílio de linguagens documentárias como tesauros e esquemas de classificação. Mas, na opinião de Lancaster (1993, p.22) "... a transcrição, é claro, não se realiza sem esses instrumentos, mas a análise independe completamente deles."

Em 1992, a ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS (ABNT) publicou a tradução da Norma ISO 5963 citada acima, sob n. 12676 intitulada "Métodos para análise de documentos- determinação de seus assuntos e seleção de termos de indexação."

O objetivo dessa Norma elaborada pela ABNT (n. 12676 de 1992) é

"(...) fixar as condições exigíveis para a prática normalizada do exame de documentos da determinação de seus assuntos e da seleção de termos de indexação (...). Destina-se, portanto, aos estágios preliminares de indexação (...)".

Na norma acima citada encontramos indicações de que esta destina-se a sistemas de indexação nos quais os assuntos dos documentos são expressos de forma resumida, e os conceitos são registrados através dos termos de uma linguagem documentária. A norma

direcionou sua aplicação para serviços de indexação independentes e serviços de indexação em rede.

A atividade de indexação, conforme a norma da ABNT, é entendida como *“ato de identificar e descrever o conteúdo de um documento com termos representativos dos seus assuntos (...)”* e consiste de três fases ou estágios:

- 1- Exame do documento e estabelecimento do assunto de seu conteúdo;
- 2- Identificação dos conceitos presentes no assunto;
- 3- Tradução desses conceitos nos termos de uma linguagem de indexação.

No primeiro estágio denominado de exame do documento, a norma se contradiz quando considera ideal a leitura total do documento e, conseqüentemente, aponta a impraticabilidade operacional de uma leitura extensiva do documento, recomendando ao indexador que analise cuidadosamente algumas partes do documento para identificar conceitos como: título e subtítulo, resumo, se houver, sumário, introdução, ilustrações, diagramas, tabelas, palavras em destaque (sublinhadas, impressas em tipos diferentes, etc) e referências bibliográficas.

Em nota de rodapé, a norma adverte que não se recomenda indexar um documento por qualquer uma das partes textuais descritas acima isoladamente.

Entretanto, a norma não especifica em qual tipologia documentária poderemos encontrar essas partes do texto mencionadas para análise de conteúdo, ou se todos os documentos indistintamente possuem essas partes. Mas, se analisarmos as estruturas textuais dos documentos a indexar, perceberemos que o documento científico é o que melhor apresenta essas partes textuais mencionadas pela norma.

Posteriormente ao estágio de exame do documento analisando partes da estrutura textual, a Norma da ABNT indica o estágio de identificação de conceitos como *“(...) uma abordagem sistemática para identificar aqueles conceitos que são essenciais na descrição do assunto”*.

Para orientar o indexador a identificar os conceitos relevantes no documento, a norma propõe uma abordagem sistemática de identificação de conceitos que vai além do “esquema de categorias existente na área coberta pelo documento” proposto pelo “Princípios de indexação do UNISIST” (1981,p.2) porque recomenda um questionamento através de questões preparadas para encontrar os conceitos relevantes, sendo essas:

- a)- O documento possui em seu contexto um objeto sob efeito de uma atividade?
- b)- O assunto contém um conceito ativo (por exemplo, uma ação, uma operação, um processo, etc.)?
- c)- O objeto é influenciado pela atividade identificada?
- d)- O documento possui um agente que praticou esta ação?
- e)- Este agente refere-se a modos específicos para realizar a ação (por exemplo, instrumentos especiais; técnicas ou métodos)?
- f)- Todos estes fatores são considerados no contexto de um lugar específico ou ambiente?
- g)- São identificadas algumas variáveis dependentes ou independentes?
- h)- O assunto foi considerado de um ponto de vista, normalmente não associado com o campo de estudo (por exemplo, um estudo sociológico ou religioso)?

Percebemos que a identificação de conceitos conforme a norma ABNT é realizada após a leitura do documento por meio das respostas ao questionamento acima adotado, e comprovamos isso quando no início do item de identificação de conceitos a norma menciona que “(...) *após examinar o documento (...) o indexador deve abordá-lo sistematicamente*” entendendo-se que a identificação de conceitos independe da leitura, por entender que o estágio de identificação de conceitos é algo separado do estágio de exame do documento. A norma deixa de mencionar se a leitura seria melhor realizada se utilizasse a abordagem sistemática do questionamento e se esse questionamento pode ser considerado como uma estratégia de leitura.

Entretanto, com os resultados das pesquisas de observação da prática de indexação realizada pelos pesquisadores do Projeto Integrado “Leitura em Análise Documentária” já mencionados anteriormente, pudemos perceber que a identificação de conceitos ocorre no momento da leitura, ou exame do documento. Além disso, a norma não menciona se a leitura seria melhor realizada usando-se a abordagem sistemática de questionamento. Por outro lado, concordamos com a norma que o questionamento ajuda o indexador a selecionar melhor os conceitos. Mas a norma falha quando não explicita quais questões seriam mais indicadas para cada parte do texto destacada para se examinar.

Em relação à seleção dos conceitos, a norma da ABNT diferencia dois momentos. Primeiramente, os conceitos selecionados durante o exame do documento, destacando que o indexador não precisa representar como termos de indexação todos os conceitos identificados, mas selecioná-los conforme os objetivos de uso dos termos.

Quanto à exaustividade e a especificidade, destaca como propósitos que podem decidir a escolha dos conceitos ressaltando, porém, que estão vinculados e dependentes de duas variáveis influentes na identificação e seleção de conceitos pelo indexador: o sistema documentário e o usuário desse sistema.

Um segundo momento da seleção de conceitos abordado pela norma diz respeito à “seleção de termos de indexação” para a tradução dos conceitos em termos de uma linguagem de indexação. Nessa fase, a norma recomenda procedimentos para a verificação de descritores controlados, bem como a preparação de uma listagem daqueles termos para os quais não houve no Tesauro uma exata representação de assuntos encontrados no documento.

A norma conclui seus procedimentos de indexação chamando atenção para o item controle de qualidade, mencionando a necessidade de qualidade e consistência em indexação, indicando os fatores que afetam a qualidade e a consistência: *a imparcialidade do indexador, o conhecimento do indexador sobre o campo coberto pelos documentos a serem indexados, as vantagens do contato direto com o usuário e a hospitalidade de linguagens de indexação para novos termos.*

Encontramos, também, na literatura o sistema automático de indexação PRECIS, já citado anteriormente, que estabelece uma metodologia para a identificação de conceitos durante a atividade de indexação por meio de questionamento ao texto, visando identificar:

- *QUE ACONTECEU? (AÇÃO)*
- *A QUE OU A QUEM ISTO ACONTECEU? (OBJETO DA AÇÃO - SISTEMA CHAVE)*
- *QUE OU QUEM FEZ ISTO? (AGENTE DA AÇÃO)*
- *ONDE ACONTECEU? (LOCAL) (FUJITA, 1989, p. 33-39)*

Encontramos, também, na revisão de literatura sobre metodologia de indexação a proposta de Kobashi (1994) que abrange a indexação de textos técnico-científicos com base na estrutura temática elaborada por Tálamo (1987), ou seja, identificar o tema do documento por meio das respostas ao questionamento: Quem? (ser), O que?, Quando?, Onde? e Como? descrita anteriormente. Ressaltamos aqui que na metodologia de Kobashi (Op. Cit) não se fez uso da categoria Quem?, (ser) caracterizada pelo traço humano, porque, aparentemente, essa categoria não é identificada em textos de natureza científica, sendo utilizada no texto narrativo e na notícia jornalística.

A metodologia proposta por Kobashi (1994, p.119-120) aborda os seguintes procedimentos:

1º Operação: Identificação do tema do texto

O indexador no momento da leitura do documento deve procurar identificar o objetivo do texto.

O parâmetro dessa leitura é a localização das categorias da estrutura temática descritas acima. Ressalta a autora que a categoria O quê? é essencial e as categorias Quando?, Onde? e Como? são acessórias, isto é, podem ou não ser encontradas no texto, sem que se possa determinar qualquer tipo de ordem de precedência entre elas no texto.

2º Operação: Elaboração do enunciado temático

Nessa fase o indexador deve selecionar e combinar as categorias temáticas pertinentes.

O parâmetro da seleção das categorias é a política do sistema de indexação.

3º Operação: Representação do tema

Nesse momento o indexador realiza a conversão do enunciado temático para a linguagem do sistema.

O parâmetro estabelecido para a conversão inclui a relação entre linguagem de indexação e a política do sistema.

Para Lancaster (1993), o indexador durante o exame do documento deve ler o título, resumo, resumo do autor denominado summary, as conclusões, os títulos das seções e legendas e das ilustrações em tabelas, e também olhar o restante do texto para confirmar se as informações extraídas nas partes consideradas relevantes apresentam uma imagem exata do que trata o texto, pois os termos atribuídos precisam refletir o documento todo.

Destacamos que as metodologias de indexação acima observadas segundo as concepções de identificação de conceitos estabelecidas por Albrechtsen (1993), dentre estas a Norma ISO 5963, ABNT (12676, 1992) e Princípios de indexação do UNISIST estão direcionadas para a análise de conteúdo do documento e para as necessidades dos usuários dos sistemas de informação. Enquanto a metodologia de Kobashi (1994) e do sistema PRECIS direcionam-se somente para o conteúdo do documento.

Ressaltamos que a identificação de conceitos na atividade de indexação deve direcionar-se tanto para a concepção orientada para o conteúdo (propõe análise do conteúdo do documento) quanto a orientada para a demanda (considera na tradução dos conceitos os interesses informacionais da clientela), pois ambas são fundamentais porque juntas visam uma recuperação da informação eficiente pelos usuários.

Percebemos nesse relato das metodologias existentes para se indexar documentos, que a fase inicial da indexação é o exame do documento realizado durante a leitura documentária, responsável muitas vezes pelo desempenho das outras atividades posteriores, porém, é pouco enfocado pelas metodologias de indexação descritas acima, o que nos faz recorrer, enquanto pesquisadores da área de Análise Documentária, aos sistemas de informações especializados para observar como o exame do documento é realizado na prática de indexação, visando estabelecer metodologias de análise de conteúdo das tipologias documentárias pesquisadas.

3.2 INDEXAÇÃO DE ARTIGOS DE JORNAL

A indexação é um processo decorrente da Análise Documentária, área da Biblioteconomia que realiza o tratamento do conteúdo temático do documento. Inicia-se pela leitura documentária com o objetivo de identificar conceitos relevantes considerando-se que essa relevância é determinada com base nos interesses dos usuários de determinada instituição. Após a identificação, os conceitos selecionados, podem ou não serem “traduzidos” em termos de indexação, mas independente disso deverão construir os produtos da indexação: os índices e resumos. Os índices de assuntos servirão de pontos de acesso para a consulta de usuários ao acervo e aos resumos informativos dos documentos que contêm as mesmas informações dos documentos originais indexados, servindo muitas vezes como substituto do documento original.

É importante relatar que a indexação trabalha com diversas tipologias documentárias como: livro, revista, jornal, CD-ROM, disquete entre outras.

Evidenciamos que muito pouco se tem pesquisado com relação a análise do conteúdo dos documentos, pois a literatura na área se detém em pesquisas referentes à representação do conteúdo do documento. E a problemática é mais extrema com relação ao nosso objeto de pesquisa, o artigo de jornal, que é descartado por não se conhecer uma

metodologia que permita tratar adequadamente esse documento e também por ser, na maioria das vezes, um documento com publicação diária onde o acúmulo de exemplares é sempre um problema por não se ter espaço físico para armazená-lo.

Na literatura sobre a metodologia de análise do conteúdo do documento, encontramos a Norma ISO (5963 de 1985) traduzida pela ABNT sob número (12676, 1992) e recomendada pelo UNISIST no seu relato sobre Princípios de Indexação (1981), além da metodologia de análise de assunto do sistema PRECIS de indexação (1989). Entretanto, ambas não especificam a que tipo de documento destina-se a metodologia de identificação de conceitos, ficando subentendido que essas metodologias se aplicam a toda e qualquer espécie de documento, ou, somente a documentos que apresentam as características descritas na metodologia de análise.

Como evidenciamos aqui a indexação de artigo de jornal, podemos mencionar primeiramente que, com relação ao armazenamento, uma solução é microfilmar o documento, ficando ainda sem saber como tratá-lo documentariamente. O que encontramos na literatura são pesquisas referentes à década de 1980 que visam o tratamento de conteúdo dos jornais locais por meio do computador, ou seja, indexação eletrônica, com o objetivo de produzir índices de jornal, via computador, por meio da análise dos títulos dos artigos. Dentre essas, podemos citar algumas pesquisas experimentais relatadas por Beare (1989) no texto *Projetos e resultados de indexação de jornal local* em que destacamos : *o projeto de indexação do The GLasglow Herald Index em 1984; The Scotman Index; The Index Meadley*; índices irlandeses como *The Country Down Spectator; The Moune Obrewer e Rerences to Itens of Country Down Interest* entre outras pesquisas experimentais.

Os resultados demonstram que a indexação eletrônica de jornal produz índices inconsistentes, não se encontrando, até então, um programa que execute a indexação de jornal de forma eficiente. Mas, não nos deteremos a relatar a indexação eletrônica, porque o nosso objetivo de pesquisa é elaborar um modelo de leitura para indexação de jornal que possa ser executado pelo indexador humano no contexto manual (jornal em suporte de papel). Portanto, realizaremos um esboço de como a literatura recomenda a indexação manual de jornal.

Como relatamos acima, a indexação manual de jornal é pouco pesquisada porque as instituições não possuem espaço físico para armazenar essa tipologia documentária, que se multiplica constantemente e a solução, então, seria microfilmar esse documento, mas isso demanda recursos financeiros que, na maioria das instituições, é sempre muito escasso. E mesmo as instituições privilegiadas que dispõem de recursos financeiros em

abundância o que se tem argumentado é que o jornal é um documento com vários artigos, e o tempo que o profissional responsável pela atividade de indexação dispõe para realizar a leitura documentária do jornal sempre se encontra “sufocado” entre tantas outras tipologias documentárias que necessitam de tratamento na instituição. Portanto a elaboração do modelo de leitura documentária para artigos de jornal possibilitará a realização de uma leitura rápida desse documento, por meio da exploração de sua estrutura textual, e com isso almejamos solucionar, em parte, o problema de tempo disponível para leitura de artigos de jornal, além de incentivar os profissionais a indexarem também esse documento.

O texto de Ahmad (1991) é um dos únicos que trata especificamente da indexação manual de jornal, sendo a nossa opção para compararmos os subsídios teóricos com os dados coletados nas instituições onde pesquisamos a atividade prática de indexação manual de jornais.

Ahmad (1991, 257) começa definindo a atividade de leitura na indexação de jornais como temos visto anteriormente, afirmando

a ocorrência de momentos simultâneos entre a leitura, identificação e seleção de conceitos pois para ela: *“indexar jornal implica na leitura hábil e rápida das notícias enquanto seleciona conceitos indexáveis. É o principal dever dos indexadores é tornar os conteúdos das notícias acessíveis pela indexação da abordagem dos assuntos. Eles devem juntar artigos informativos relacionados pela atribuição de descritores adequados (...)”* (Tradução Livre, com grifo nosso).

Enfocaremos agora os procedimentos manuais para indexação de jornais.

Primeiramente, temos observado que a indexação de jornal, geralmente nas instituições especializadas, é realizada em equipes de trabalho, em que cada indexador é responsável por determinada área de assunto específico que lhe foi atribuída.

Posteriormente, inicia-se a indexação do documento visando identificar conceitos por meio da leitura do documento. E para Ahmad (1991, p.259) não existe na fase de leitura do documento, um conjunto de critérios “fixos” pelos quais se possa analisar cada informação da mesma maneira (na leitura de um novo texto, na maioria das vezes, temos de utilizar ou criar novas estratégias de leitura distintas das usadas em textos anteriores para atingirmos a compreensão), então, os indexadores, para analisar a notícia jornalística, podem se direcionar pelos seguintes procedimentos de indexação de jornais identificados:

- 1- *compreender a linguagem e a terminologia do jornal;*
- 2- *compreender a mensagem geral transmitida pelo artigo/reportagem;*
- 3- *examinar cuidadosamente o título bem como o primeiro parágrafo e definir com exatidão, sublinhando as palavras significantes que representam os conteúdos do assunto;*
- 4- *correr os olhos por todo o texto e identificar os conceitos indexáveis; interpretar as palavras e os termos;*
- 5- *analisar as palavras e os conceitos e traduzi-los em linguagem de indexação;*
- 6- *casar os conceitos e as palavras com os descritores de um tesouro de notícias padrão;*
- 7- *criar/inventar novos descritores se eles não forem encontrados no tesouro de notícias;*
- 8- *identificar todos os nomes pessoais; nomes de corporações (empresas) importantes, etc (Tradução Livre).*

Evidenciamos que, no artigo de jornal, os conceitos podem expressar idéias, relações entre ocorrências, objetos ou sujeitos identificáveis. O autor de determinado texto, por exemplo, pode ser descritor do tema de um determinado artigo de jornal, quando esse for muito conhecido, podendo o indexador indexá-lo como autor e como assunto. A mesma situação poderá ocorrer com os sujeitos envolvidos na notícia e o título.

Nas resenhas de livro publicadas em jornal descreve-se o título, autor ou revisor. E considerando que *os títulos nem sempre são precisos, o indexador pode enriquecer os índices de assunto ou entradas temáticas com palavras explicativas entre colchetes; pode traduzir ou modificar a ortografia; quanto às abreviações, podem ser padronizadas e títulos vagos podem ser substituídos por outros mais significativos.* (AHMAD, 1991).

Para Ahmad (1991), a análise da notícia exige do indexador uma avaliação crítica dos seguintes aspectos: títulos, subtítulos, parágrafos mais importantes, termos isolados, expressões e afirmações que formam as notícias, a identificação dos temas centrais, subtemas, aspectos não-relacionados e relações sintagmáticas potenciais de sujeitos. Refletindo sobre esses aspectos, podemos perceber que a autora menciona implicitamente o conhecimento da estrutura textual do artigo do jornal para que o indexador saiba diferenciar as informações pertencentes a cada parte da estrutura do texto e identifique mais rapidamente as informações e os parágrafos mais importantes, uma vez que não tem tempo de ler todos os artigos do jornal. E o indexador precisará ter também uma boa memória porque nos jornais as matérias são contínuas durante semanas, e ocorrerão situações em que será necessário lembrar de um conceito já atribuído anteriormente,

deverá buscar adquirir conhecimento prévio diversificado, ou pelo menos sobre o assunto que indexa, porque no jornal abordam-se vários assuntos referentes a áreas do conhecimento humano, e desenvolver uma flexibilidade para elaborar estratégias de leitura que visam compreender o texto, pois a cada texto poderá ser exigido novas estratégias de leitura. Mas se o indexador não possuir e nem procurar adquirir esses conhecimentos e habilidades, a tarefa de ler e identificar o tema do artigo de jornal se tornará mais difícil e haverá perda de tempo com a atividade não se tornando viável para os fins de indexação.

Ressaltamos que os cabeçalhos do jornal (título) devem ser considerados pelo indexador como as menores unidades passíveis de referência, pois muitos títulos são enganosos ou se constituem metáforas como ocorre normalmente nos cadernos referentes a área de Economia. (AHMAD, 1991). Mas isso acontece porque, como descrevemos no capítulo 1, onde abordamos a estrutura textual de jornal, os títulos geralmente são elaborados para chamar a atenção dos leitores.

Na indexação de jornal, o indexador após selecionar os conceitos pode e deve fazer uso de uma linguagem controlada de indexação para representar o conceito conforme está estruturado na linguagem do jornal. Tal linguagem obedece geralmente à terminologia adotada pela área de jornalismo, assim estará garantindo, consistência na indexação, evitando a possibilidade de entradas diferentes de índices porque os conceitos estarão padronizados, e se não for possível contar com o auxílio de um vocabulário controlado, o indexador deve evitar indexar as notícias sob descritores mais amplos, ou genéricos. (AHMAD, 1991).

Mencionamos que segundo Ahmad (Op. Cit) existem algumas instituições que possuem linguagens documentárias para padronizar a indexação de jornal. Encontramos o New York Times Thesaurus; Canada News Index; The Times Index entre outros que geralmente são desenvolvidos no contexto das instituições que indexam jornal para uso próprio.

Como se pode perceber, a prática manual de indexação de artigos de jornal é pouco exercida nas instituições e, conseqüentemente, escassa na revisão de literatura sobre o assunto.

Então ressaltaremos aqui na seqüência alguns procedimentos de indexação de jornal identificados em nossa pesquisa de conclusão de curso de graduação em Biblioteconomia na UNESP-Marília, realizada em 1997, referente à observação de leitura

dos indexadores, durante a indexação de artigos de jornais no contexto do jornal “O Estado de São Paulo”. E procuraremos identificar a seguir os procedimentos estabelecidos por Ahmad (1991) descritos acima, nessa pesquisa já concluída.

.. Ressaltamos que com o desenvolvimento dos capítulo 2 e 3 alcançamos o objetivo da pesquisa de *proporcionar aos indexadores, subsídios metodológicos para a análise de assunto de textos com estrutura diversificada* e atingimos novamente mais especificadamente com o texto de Ahmad (1993) um outro objetivo antes alcançado com as informações trabalhadas no capítulo 1 que é de: *identificar estratégias de leitura para artigos de jornal*.

No próximo capítulo, mencionaremos as metodologias e os procedimentos utilizados para o desenvolvimento dessa pesquisa sobre a leitura documentária na indexação de jornal visando propor um modelo de leitura para esse determinado fim.

CAPÍTULO 4: METODOLOGIA

Para o desenvolvimento da pesquisa objetivando elaborar um modelo de leitura para indexação de artigos de jornais foi necessário adotar uma metodologia de coleta e análise de dados que se compôs dos seguintes processos descritos na sequência desse capítulo.

Primeiramente, foi feito um levantamento bibliográfico em fontes bibliográficas de informação na área, abrangendo os dez últimos anos onde se pesquisou a literatura relevante publicada sobre: leitura, leitura documentária, estrutura textual, estrutura textual do texto jornalístico e identificação de conceitos e metodologias para avaliação de política de indexação e linguagens documentárias.

Com o levantamento bibliográfico, encontraram-se bibliografias relevantes enfocadas nas páginas anteriores referentes aos pressupostos teóricos onde se pode destacar: a metodologia para indexação manual de jornais de Ahmad (1993); tipologias e estruturas textuais de textos trabalhados pela Análise Documentária citada por Kobashi (1994); a classificação das concepções de análise de assunto proposta por Albrecthsen (1993); a estrutura textual do texto jornalístico elaborada por Van Dijk (1983) e as considerações de Silva (2001) para essa estrutura textual; atividades metacognitivas presentes na leitura consciente estabelecidas por Brown (1980), elementos de política de indexação estabelecidos por Carneiro (1985), avaliação da estrutura externa de tesauros

101

elaborada por Corrêa (1998); metodologia de avaliação da estrutura interna de linguagens documentárias elaborada por Ginez de Lara (1993); metodologia de assunto do sistema de indexação Precís traduzido para língua portuguesa por Fujita (1989).

Posteriormente, para a observação da leitura dos indexadores, foram necessários estudos de casos em centro referencial especializado que desenvolvesse a indexação de jornais possuindo, portanto, acervo organizado. Os locais escolhidos foram: Arquivo do jornal "O Estado de São Paulo" e Departamento de Documentação da Editora Abril Cultural (DEDOC). Nessa fase de estudos de casos, foi realizada por meio de consulta à documentação e conversa com os responsáveis pelo setor, uma caracterização da infraestrutura física e funcional dos centros englobando: histórico breve e área física, o organograma da instituição, recursos humanos, recursos materiais (mobiliário, equipamentos e acervo), usuários e serviços prestados. E também foi elaborada uma entrevista para a caracterização dos indexadores dos centros na qual se obteve dados referentes a: formação educacional, experiência profissional anterior e no serviço de indexação, outras atividades que realiza no centro, dificuldades na atividade de indexação e procedimentos de identificação de conceitos e fontes de experiência de indexação.

Evidenciando que o objetivo da pesquisa é a observação da leitura dos indexadores de jornal para elaborar, com os resultados obtidos, um modelo de leitura para indexação de jornais e considerando a leitura documentária a operação primordial da atividade de indexação, porque a partir dela dependerá o bom ou mau resultado das outras atividades (elaboração de resumo e formação de índices de assunto), não poderíamos deixar de mencionar e querer observar, entre outras, as variáveis que influenciam na leitura do indexador como: a política de indexação do centro referencial, que, também, por sua vez, é refletida na escolha da linguagem de indexação do centro a qual também exerce influência muitas vezes, no resultado da leitura do indexador.

Portanto, os estudos de casos analisaram, também, a política de indexação do Arquivo do jornal "O Estado de São Paulo" e do Centro de Documentação da Editora Abril Cultural (DEDOC), por meio da verificação da ocorrência de alguns elementos estabelecidos por Carneiro (1985) para implantação de uma política de indexação. E como, por sua vez, a escolha da linguagem documentária do centro sofre influência da política de indexação que determina a comunidade usuária que a instituição servirá, desenvolveu-se uma avaliação da estrutura externa das linguagens documentárias seguindo a metodologia elaborada por Corrêa (1998), e avaliação da estrutura interna somente da linguagem do

102

DEDOC seguindo a metodologia de Ginez de Lara (1993), porque a linguagem do Arquivo não apresentou parâmetros funcionais para se avaliar a sua estrutura interna.

Por fim, a metodologia adotada nessa pesquisa, para a observação de estratégias da leitura documentária de jornal, no contexto do arquivo do jornal "O Estado de São Paulo" e no Departamento de Documentação da Editora Abril, foi 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, o "pensar alto" de sua compreensão.

Antes do surgimento do "Protocolo Verbal", a maioria dos pesquisadores, até a década de setenta, utilizava em suas pesquisas métodos para coleta dos dados em que se ressaltavam as respostas dos indivíduos.

Entretanto, a partir dessa década, a Psicologia Cognitiva passou a observar que seria fundamental investigar os comportamentos dos indivíduos durante a realização de tarefas a eles solicitadas. Dentre esses comportamentos, na tarefa de ler, destacam-se os movimentos dos olhos sobre o texto, o virar de páginas e os processos mentais que levam à compreensão. (Cavalcanti, 1989)

Para investigar essas ações do indivíduo durante a realização de uma tarefa, fez-se necessário desenvolver técnicas introspectivas que foram classificadas por Radford & Burton (1974), citados por Cavalcanti (1989, p. 138) em três grupos sendo:

1. Auto – observação (introspecção propriamente dita): o sujeito observado relata seus próprios eventos mentais;
2. Auto-relato ou Autopercepção (retrospecção): o sujeito conta sua experiência ao pesquisador/analista;
3. Pensar Alto (protocolos verbais ou análise de protocolo): o sujeito pensa em voz alta enquanto realiza uma tarefa.

Evidenciamos que, na utilização da técnica introspectiva denominada "Pensar Alto", os pensamentos do sujeito são gravados e transcritos literalmente, produzindo protocolos verbais. Protocolos são geralmente 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 & Zanotto, 1994).

103

Esse método de observação ultrapassou os limites da Psicologia Cognitiva e passou a ser utilizado na Lingüística Aplicada surgindo na pesquisa sobre leitura através 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 Lingüística 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", coordenado por Fujita, 1998.

É inédito, no Brasil, a observação do processo de leitura documentária para indexação com uso do Protocolo Verbal, realizada no Projeto Integrado com indexadores da Sub-Rede de Informação em Ciências da Saúde Oral. 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 no processo de indexação humana. Nesse trabalho, o autor realizou um experimento com dois sujeitos indexadores utilizando a técnica de Protocolo Verbal.

Mas a técnica do "Protocolo Verbal" é também criticada por alguns pesquisadores que argumentam que os dados coletados não são totalmente confiáveis porque os indivíduos observados podem distorcer os seus pensamentos buscando causar uma "boa impressão para não parecerem ignorantes"; como também podem processar outros pensamentos desvinculados do solicitado na tarefa.

Para Cavalcanti (1989), este problema pode ser evitado desde que haja empatia entre pesquisador e observado, ou seja, o sujeito tem que se sentir à vontade com o pesquisador, como também compreender e estar familiarizado com a tarefa, o que pode ser obtido por meio de instruções aos sujeitos e mediante sessão de familiarização com a realização da tarefa.

Ericsson & Simon (1987), que também defendem a utilização do "Protocolo Verbal", rebatem as críticas apoiados na teoria do processamento da informação explicando que o problema levantado em torno do processamento de outros pensamentos durante a aplicação do Protocolo Verbal não tem fundamento porque para eles: "a informação é armazenada em várias memórias, com diferentes capacidades e características de acesso" sendo que:

- na memória de curto prazo (STM) - ocorre duração intermediária de retenção e acesso rápido à informação;
- enquanto que na memória de longo prazo (LTM) - ocorre armazenamento e duração relativamente permanente sendo lenta a recuperação da informação”.

Sendo assim, Ericsson & Simon (Op. Cit, citado por Rubi, 2000, p. 35)

“ prevêm que a informação recém-apreendida pelo processador central é mantida na memória de curto prazo por algum tempo e é diretamente acessível para processamento subsequente, enquanto que na memória de longo prazo a informação precisa ser recuperada antes de ser relatada. Dessa forma, as informações coletadas em Relatos Verbais são as recém-apreendidas, diretamente acessíveis para processamento subsequente”.

Portanto, apesar de ainda controvertida, essa técnica é, segundo Nardi (1993, p. 81), o único instrumento de coleta, no momento, disponível, que possibilita observar processos do leitor durante a compreensão de um texto, o que justifica a nossa escolha. Além disso, a questão da confiabilidade dos dados provenientes de técnicas introspectivas é um problema de outras técnicas também.

Na sequência, no capítulo 5, enfocamos, por acreditar ser relevante para os objetivos da pesquisa, a caracterização dos sistemas de informações especializados no caso Arquivo do jornal “O Estado de São Paulo” e Departamento de Documentação da Editora Abril Cultural (DEDOC) onde se desenvolveu a observação da prática de leitura dos indexadores de jornais; caracterizamos a política de indexação dos sistemas e perfil dos indexadores e, realizamos, ainda, a avaliação das linguagens de indexação em uso nessas instituições.

CAPÍTULO 5: LEITURA DOCUMENTÁRIA PARA INDEXAÇÃO DE TEXTO JORNALÍSTICO EM CENTROS REFERENCIAIS ESPECIALIZADOS: estudos de casos

Este capítulo está dividido em três partes: na primeira e na segunda mencionamos a caracterização das instituições onde ocorreu a observação da leitura do indexador destacando a atividade de indexação, política de indexação, escolha da linguagem documentária, e na terceira parte, enfocamos a avaliação das linguagens por meio da definição da estrutura interna e externa de tais linguagens.

Abordaremos primeiramente aqui a caracterização da ambiência considerando o Arquivo do jornal “O Estado de São Paulo” onde realizamos a observação de leitura dos indexadores de jornais. Procuramos evidenciar infra-estrutura, histórico (APÊNDICE B),

caracterização dos indexadores, atividade de indexação, política de indexação e avaliação das linguagens documentárias.

Ressaltaremos no arquivo do jornal “O Estado de São Paulo” a caracterização dos seus profissionais indexadores, e consequentemente a política de indexação da instituição, por meio da identificação dos elementos e requisitos estabelecidos por Carneiro (1985), citados na revisão de literatura, como também a descrição da atividade de indexação e a escolha da linguagem.

5.1 Arquivo do jornal “O Estado de São Paulo”

A identificação e caracterização dos indexadores encontram-se demonstradas em formulários (APÊNDICE C). Os dados dos formulários foram coletados por meio de solicitação de informações aos indexadores individualmente abordando: *formação educacional, experiência profissional anterior no serviço de indexação, outras atividades que realiza no Arquivo, dificuldades na atividade de indexação, procedimentos de identificação de conceitos, fontes de experiência de indexação e nível de exaustividade e especificidade exigido na indexação.*

Evidenciando brevemente os resultados dos formulários dos indexadores do Arquivo do jornal “O Estado de São Paulo” destacamos que a maioria dos indexadores são jornalistas, adquiriram prática na atividade de indexação com a realização diária da tarefa, ambos indagam não possuir dificuldades para indexar e quanto ao grau de especificidade afirmam selecionar conceitos relevantes no documento.

a) A atividade de indexação no Arquivo

Os dados coletados sobre o processo de indexação no jornal foram obtidos por meio de solicitação de informações ao funcionário responsável.

Constatou-se em 1997 que o processo de indexação era feito, manualmente. Iniciava-se com a divisão por editoria, ou seja, divisão dos cadernos que compõem o jornal entre os quatro indexadores, que indexavam o *Jornal da Tarde* e o jornal *O Estado de São Paulo*.

Nesse trabalho, os indexadores realizavam uma leitura do artigo do jornal, procurando retirar o assunto que representasse o conteúdo tratado no artigo, para posterior

tradução para os termos estabelecidos pela “Linguagem do Jornal” (APÊNDICE D) considerada na época a única Linguagem Documentária do Sistema. Essa era composta por diversos assuntos que *geralmente aparecem nos jornais*, sendo uma mera lista de assuntos elaborada por um antigo funcionário do jornal.

Para extrair o assunto os indexadores faziam a si mesmos a seguinte pergunta: Do que trata esse artigo?. Identificando o assunto ou mais de um assunto no texto, que fossem essenciais, o indexador indicava os assuntos com uma caneta vermelha, no próprio texto, na parte que compõe o cabeçalho, como também o nome do autor do artigo. Com a identificação e destaque dos assuntos extraídos pelos indexadores, os digitadores preenchiam o catálogo automatizado que trazia os seguintes campos (assunto, autor do artigo, caderno em que foi publicado, página, data). Percebemos aqui que os indexadores do arquivo, conforme nos relatou Ahmad (1991) no capítulo anterior, buscam compreender a mensagem, ou seja, o conteúdo do texto.

Ressaltamos que podem ocorrer casos em que o assunto continua a ser publicado no dia seguinte, e nessa continuação é possível que o enfoque ou título dado ao assunto seja mais aceito que o primeiro, então, o indexador deve mudar a seleção do termo menos conhecido para o mais conhecido. Um exemplo disso ocorreu com uma matéria que recebeu um título e depois mudou para outro título: “Caso da Rua Cuba” pelo qual ficou mais conhecida; então os indexadores cancelaram o título dado primeiramente a matéria e indexaram, a partir de então, como: Caso da Rua Cuba. Podemos perceber aqui, que no arquivo, os indexadores criam novos conceitos visando a recuperação pelos usuários, o que demonstra que na instituição a identificação de conceitos além de ser direcionada para a análise do conteúdo do documento é voltada também para a demanda não se prendendo tanto nos termos da linguagem de indexação. Podemos demonstrar isto com o exemplo acima no qual foi mantido e, provavelmente, o título da matéria se tornou um termo de indexação porque assim a matéria seria mais facilmente recuperada do que se optasse somente por termos da linguagem do sistema.

b) Política de Indexação do Arquivo

Os dados quanto a política de indexação do O arquivo do jornal “O Estado de São Paulo” foi obtida por meio de solicitação de informação. Evidencia-se que no Arquivo essa

107

visita ocorreu durante a nossa pesquisa de graduação em 1997 e naquela ocasião não se possuía, um manual de política de indexação descrito em algum suporte informacional.

Mas, percebemos com a observação da leitura dos indexadores que, no Arquivo, os usuários alvo são os jornalistas e que a seleção de termos considera o conteúdo da matéria e os interesses dos usuários. E o que garante qualidade na indexação quanto à representação do conteúdo e recuperação é a experiência dos indexadores do Arquivo na realização da atividade.

c) Escolha da Linguagem de Indexação do Arquivo do Jornal “O Estado de São Paulo”

As informações sobre a linguagem do jornal foi obtida por meio de solicitação de informações ao funcionário responsável.

Destaca-se que a Linguagem do Arquivo foi elaborada por um antigo funcionário do jornal e se constitui numa listagem alfabética de termos.

A organização da linguagem do jornal não possui uma estrutura organizada; constitui-se de uma listagem de termos em ordem alfabética mantendo meio que subentendido que são termos de relação.

No Arquivo, durante nossa pesquisa, pudemos perceber que a linguagem estava desatualizada porque a maioria dos termos selecionados pelos indexadores dela não constava. Nesse caso, os indexadores tinham autonomia para criarem termos de indexação. Portanto, naquele momento, pudemos perceber que a linguagem já não satisfazia perfeitamente a tradução dos conceitos para termos de indexação, atribuindo-se aos indexadores toda a eficiência da recuperação da informação, porque estes tinham conhecimento das necessidades dos usuários e responsabilidade de listar os termos criados para manter controle do termos de indexação.

Destacamos que falaremos mais sobre essa Linguagem, posteriormente, quando tratarmos da avaliação das linguagens documentárias.

Enfocaremos agora, na sequência, a caracterização do segundo centro referencial onde se desenvolveu a observação da leitura dos indexadores de jornal no caso, o Departamento de Documentação da Editora Abril Cultural por meio de histórico breve, infra-estrutura física e funcional (APÊNDICE D) identificação e caracterização dos indexadores, identificação da política de indexação e escolha da Linguagem do DEDOC.

5.2 Departamento de Documentação da Editora Abril

Relataremos agora a caracterização dos indexadores que realizam a indexação no DEDOC por meio de formulários (Apêndice E) assim como no Arquivo no DEDOC os dados dos formulários foram coletados por meio de solicitação de informações aos indexadores individualmente procurando evidenciar: *formação educacional, experiência profissional anterior no serviço de indexação, outras atividades que realiza no Arquivo, dificuldades na atividade de indexação, procedimentos de identificação de conceitos, fontes de experiência de indexação e nível de exaustividade e especificidade exigido na indexação.*

Os resultados obtidos demonstraram que os indexadores do DEDOC possuem formação superior diversificada, a indexação é dividida por áreas de assunto entre os indexadores, a maioria diz ter adquirido prática na indexação, por meio da realização da tarefa diária, e afirmam não encontrar dificuldades na realização da tarefa de indexar e busca selecionar conceitos pertinentes ao conteúdo dos documentos.

a) A Atividade de Indexação do DEDOC*

A descrição do processo de indexação foi obtida por meio de solicitação de informação ao funcionário responsável. Constatou-se que o processo de indexação no DEDOC está automatizado. E os indexadores realizam a indexação do jornal eletrônico acompanhado com os jornais em papel porque, muitas vezes, os jornais on-line (DEDOC Eletrônico) não apresentam muitas matérias, ou partes das matérias presentes nos jornais em papel e que são relevantes para os usuários. Nesse caso, o indexador deve indicar a matéria, ou trecho que está faltando para que essas matérias dos jornais tradicionais possam ser escaneadas e posteriormente indexadas.

O processo de indexação no DEDOC, inicia-se com o acesso ao jornal on-line. Posteriormente, os indexadores começam uma seleção de artigos nos jornais a indexar durante o momento de indexação, porque eles dividiram, entre si, as áreas de assuntos que cada um indexará. Então, um deles começa a indexar um jornal selecionando aquelas matérias que serão incluídas no seu rol de assuntos, deixando de lado as matérias que

* O DEDOC* realiza indexação de diversos jornais. E inclusive os jornais utilizados nessa pesquisa durante a aplicação do Protocolo Verbal descritos a seguir são indexados no DEDOC, assim como os textos selecionados nesses jornais para a observação da leitura dos indexadores.

pertencem a outros indexadores. As áreas de assuntos dos indexadores são: geral, esporte, economia, política nacional e internacional, artes e literatura. A seleção das matérias que serão indexadas é realizada pelos indexadores que optam por aquelas que consideram relevantes dentro da área, objetivando, para tanto, os interesses informacionais dos usuários internos. Geralmente, os indexadores selecionam matérias que apresentam fato novo, que surpreende, matérias que apresentam resumo dos acontecimentos descritos em matérias anteriores relacionadas com a atual. Os indexadores selecionam os assuntos específicos tratados na matéria; após a seleção dos assuntos, devem preencher os campos de uma ficha de indexação (apresentada a seguir). A tradução dos conceitos para uma linguagem de indexação ocorre da seguinte maneira o indexador seleciona um conceito, digita-o na ficha de indexação; se o termo não for o termo controlado, o sistema exibe na tela do computador uma mensagem orientando o indexador a entrar pelo termo controlado que está indicando, e também orienta as maneiras corretas de entrada dos termos no sistema.

Percebemos que a identificação de conceitos do DEDOC, segundo as menções de Albretchtsen (1993), está direcionada para a concepção de demanda e de conteúdo, porque seleciona conceitos que reflitam o conteúdo das matérias e que satisfaçam às necessidades dos usuários.

Ficha de Indexação do DEDOC

Reg:

Título: *Título da matéria*

Autor: *Autor da matéria*

Public: *Jornal que publicou* **Data:** *dia de publicação* ≤ < > ≥ **Visualizar**

Seção: *Caderno do jornal* **Especial:** *Encontrar mais rápido*

OBS: *Encontrar assunto em determinada pasta*

Pasta:

URL/DIR: *Endereço Eletrônico do DEDOC*

Texto: *Seleção do texto* : *indexar* *atribuir termos ao texto selecionado*

Incluir: *Gravar o texto na URL* **Alterar:** **Excluir** **Salvar:**

Voltar : *ao jornal* **Pesquisar :** *Linguagem* **Base:** **Sair**

Palavras-Chave: *Atribuir os Termos selecionados*

b) Política de Indexação do DEDOC

O Departamento de Documentação (DEDOC) da editora Abril Cultural não possui uma política de indexação estabelecida em manual de indexação. Sua única recomendação é que os indexadores, além da indexação, realizem também a atividade de atendimento à pesquisa dos usuários, para que conheçam os conceitos referentes às necessidades informacionais dos usuários, para que possam durante a indexação efetuar uma identificação de conceitos direcionada a satisfazer às necessidades informacionais dos usuários refletidas no conteúdo específico da matéria.

Destacamos que este procedimento pode ser identificado na observação de leitura dos indexadores porque quando selecionavam termos sempre consideravam o conteúdo da matéria e mencionam as vezes as formas de recuperação do conteúdo pelos usuários.

c) Escolha da Linguagem de Indexação do DEDOC

As informações sobre a linguagem do DEDOC também foi obtida por meio de solicitação de informações. Evidenciou-se que a Linguagem do DEDOC encontra-se em formato eletrônico., e foi elaborada na própria instituição baseando-se em catálogo de fichas anteriormente existentes.

A estratégia de busca é feita por título, autor, assuntos, jornal, data, pasta de assunto e pode utilizar-se do operador lógico: AND para cruzar mais de uma estratégia de busca.

Sendo que AND (interseção) é o operador usado para relacionar os termos de uma pesquisa. Em uma estratégia de busca, utilizando-se o AND, pode-se recuperar matérias que apresentam os conceitos especificados na busca em um mesmo documento.

A recuperação de uma busca é demonstrada por meio de uma listagem das matérias nas quais atribuem os termos pesquisados, abrangendo os campos: título da matéria, autor, pasta em que foi indexada, jornal que a publicou, data, páginas.

Na Linguagem DEDOC estabeleceram-se algumas formas de organização dos termos como:

- a) os descritores autorizados e seus sinônimos (sendo os termos sinônimos identificados como termos proibidos)

Exemplo: tangerina (termo autorizado) sendo os sinônimos deste todos proibidos de serem representados pela linguagem como no caso (bergamota e os outros sinônimos – termos proibidos porque o autorizado é tangerina).

- b) modificadores (indicam aspectos que diferenciam conceitos escritos iguais mas pertencentes a um contexto específico por meio do sinal /

Ex: BALANÇO/CONTABILIDADE

BALANÇO/BRINQUEDO

- c) categorias às quais os descritores pertencem;
- d) Notas explicativas
- e) Descritores relacionados

A linguagem do DEDOC está estruturada em três categorias:

1- Categoria Descritores: encontramos nessa categoria os termos de indexação atribuídos aos documentos, ou seja, assuntos;

2- Categoria Personalidades: menciona pessoas famosas, públicas, desconhecidas que se transformaram em notícia;

4- Categoria Identificadores: encontramos empresas, Estados Brasileiros.

Na Linguagem, cada categoria se subdivide em outras subcategorias mantendo somente a ordem de termos relacionados. Todo descritor apresenta uma identificação da categoria a que pertence podendo ainda ser identificado como termos relacionados; sinônimos (termos proibidos); controlados e, se preciso, notas explicativas e observações.

Em síntese, destacaremos alguns aspectos das duas instituições DEDOC e Arquivo durante a caracterização da ambiência de pesquisa.

Os indexadores no Arquivo, em sua maioria, são jornalistas que possuem, ou deveriam possuir conhecimentos na área de Jornalismo e, portanto, da arquitetura de conceitos dessa área e definição de suas terminologias.

No DEDOC, encontramos profissionais-jornalistas e com outras formações e, pela entrevista, dizem não possuir dificuldades na indexação.

Em ambas as instituições a identificação e seleção dos conceitos durante a indexação estão direcionadas para os assuntos específicos tratados nos documentos e, ao mesmo tempo, percebe-se que demonstram estar também refletidos nas necessidades informacionais dos usuários. Sendo assim, nas instituições, a política de indexação para identificação de conceitos enquadra-se nas concepções de análise de assunto orientada para o conteúdo e orientada para a demanda (privilegia os jornalistas como usuários) mencionadas por Albrecthsen (1993, citado na p.59).

5.3 Avaliação de Linguagens Documentárias: estudos de casos

Como o nosso objetivo é observar as atividades de indexação de jornais, visando propor um modelo de leitura documentária para indexação de jornais, faz-se necessário avaliar as linguagens dos sistemas de informação envolvidos na pesquisa porque a linguagem tem influência na indexação à medida que estabelece, ou não, compatibilidade dos termos selecionados como representantes do conteúdo do texto o que norteia muitas vezes a seleção, ou não do termo pelo indexador. É relevante também conhecer um pouco as linguagens adotadas nos sistemas de informações que tratam a notícia jornalística e as suas influências e relevância nesse contexto diferente da biblioteca.

As linguagens documentárias do jornal “O Estado de São Paulo”, de que obtemos uma cópia da versão de 1997, e a linguagem do Departamento de Documentação da Editora Abril Cultural, que visitamos recentemente, encontram-se aqui avaliadas e comparadas conforme as metodologias de avaliação de linguagens estabelecidas por Ginez de Lara (1993) e Corrêa (1998), ambas descritas anteriormente, no capítulo 3, referente à indexação.

A metodologia de avaliação das linguagens documentárias do DEDOC e do jornal foi dividida em duas fases: análise formal da estrutura e avaliação da estrutura semântica das linguagens.

A análise da estrutura formal baseou-se nos parâmetros definidos por Corrêa (1998, p.77-87) para o conhecimento da estrutura formal de tesauros facetados, os quais foram utilizados para o conhecimento da estrutura das linguagens documentárias do DEDOC e do Jornal, com exceção da parte seis, específica para tesauros facetados.

Os parâmetros estabelecidos por Corrêa (op. Cit) foram mencionados no capítulo de indexação.

Com relação à avaliação da estrutura semântica das linguagens optou-se pela metodologia de Ginez de Lara (1993) que propõe análise por meio de análise de exemplos das linguagens a serem avaliadas a partir de cinco critérios mencionados no capítulo de indexação anteriormente.

- **Análise Formal da Estrutura da Linguagem Documentária do DEDOC conforme metodologia de Corrêa (1998) destacando:**

- A- Descrição da Estrutura da Linguagem documentária do DEDOC

- 1- Aspectos Gerenciais

- a) Compilação dos termos

Para a compilação dos termos partiu-se de arquivo de fichário interno. Apresenta poucos termos em inglês, utilizados somente quando o descritor em inglês não possui equivalente em português, quando já existem descritores em português mas são pouco difundidos ou há preferência em se usar o descritor em inglês:

Ex: SOFTWARE

Usaram neologismo quando não havia termo equivalente em português e se necessitava suprir a deficiência.

Ex: HARDWARE

Tradução;

b) Números de Termos

A linguagem do DEDOC apresenta 87.384 termos.

c) Uso de Software e/ou programa de computador

A linguagem do DEDOC encontra-se armazenada em um software denominado de ORACLE Sistema ISIS.

d) Instrução de uso/informações complementares

A linguagem apresenta em forma de ficha notas explicativas para alguns termos das categorias (Personalidade, Identificador e Descritor) informando se o termo é autorizado, relacionado, proibido, a que categoria pertence, observações, data de inclusão no sistema e data de alteração do termo.

e) utilização de normas para controle de vocabulário

A linguagem apresenta o termo denominado de modificador utilizado para diferenciar conceitos cujas escritas são iguais e podem representar mais de um significado, então, procura explicar o conceito daquele conceito.

É norma entrar por nome de associações sempre em caixa alta, isto é, letra maiúscula e conceitos com artigos sempre no final como: BECO, O/CASA DE ESPETÁCULO.

f) Cobertura Temática

São abrangidos assuntos diversificados porque o jornalismo trabalha com um rol de informações ampla classificados dentro das três categorias: Descritor (assuntos), Identificador (Empresas, Instituições e Corporações) e Personalidade (Pessoas famosas ou anônimas).

2 Estrutura da linguagem do DEDOC

a) Lista Alfabética

Os conceitos estão dispostos em ordem alfabética listados nas três categorias Descritor, Personalidade e Identificador, mantendo em cada categoria a ordem de relação entre os conceitos.

Ex: Medicina

Medicina veterinária

Medicina Homeopática e, assim, sucessivamente.

b) Lista Permutada

A linguagem não apresenta nenhuma outra forma de lista além da alfabética das categorias Descritor, Personalidade e Identificador. Cada palavra que compõe o conceito constitui uma entrada única, exceto artigos, conjunções, preposições, numerais. Pode-se localizar um descritor mesmo quando se conhece apenas uma palavra do mesmo.

Ex: Aberração

Aberração Genética

c) Lista hierárquica

Os descritores estão organizados nas três categorias somente por ordem de relação, associação não apresentando ordem hierárquica.

3 Apresentação da linguagem

3.1 Entrada

Os assuntos estão distribuídos em três categorias:

- 1- Categoria Descritor: encontramos aqui os assuntos indexados das mais variadas áreas do conhecimento;
- 2- Categoria Personalidade: localizamos nomes de pessoas famosas, importantes, desconhecida que virou notícia;
- 3- Categoria Identificador: refere-se a nomes de empresas, corporações, instituições.

3 Apresentação da linguagem

3.1 Entrada

a) Entrada da Lista Alfabética

Os conceitos apresentam-se listados alfabeticamente em ordem crescente dentro da categoria a que pertence e somente na categoria Descritor encontramos uma ordem de relação entre eles.

b) Entrada da Lista Permutada

Não apresenta nenhuma outra lista além da alfabética. Mas, durante a pesquisa, recupera termos ou conceitos controlados. E quando o termo não é um descritor, é classificado como termo proibido que envia para o termo descritor.

Ex: Termo proibido: BERGAMOTA Termo controlado: TANGERINA

c) Entrada da lista Hierárquica

Não possui ordem hierárquica entre os termos.

3.2 Componentes da Entrada

A linguagem apresenta os termos listados alfabeticamente dentro de cada categoria. E para saber como estão organizados os termos dentro de determinada categoria, deve-se selecionar o termo na listagem alfabética da categoria que deseja pesquisar e a linguagem recuperará em forma de uma ficha catalográfica como o termo se encontra classificado dentro de sua categoria.

A Linguagem apresenta, na tela de consulta aos conceitos, os componentes:

- Termo (descriptor controlado ,ou proibido)
- termo autorizado (apresentado quando se pesquisa por um termo sinônimo considerado como proibido)
- categoria a que pertence o descriptor (destaca a categoria podendo ser: Descriptor, Identificador e Personalidade).
- Menciona se é um termo controlado, ou proibido
- Nota de escopo (esclarece informações sobre o termo como, por exemplo, na categoria *personalidade* podemos mencionar a observação de que a pessoa é um cantor, pertence a determinada banda)
- Observações: introduz observações referentes ao termo como, por exemplo, na categoria Personalidade, dia de aniversário da pessoa)
- Data de criação do termo
- Data de alteração do termo
- Termos relacionados
- Termos proibidos (introduz termos sinônimos considerados proibido dentro da categoria descriptor, como também, termos com entradas erradas nas categorias Personalidade e Identificador.
- Quantidade de registros que possuem determinado termo pesquisado

Ex: TANGERINA

TR FRUTA CÍTRICA

TERMOS PROIBIDOS

BERGAMOTA

MEXERICA

MORGOTE

TANGERINA/FRUTA/

Os componentes: nota de escopo, observações, termos relacionados e termos proibidos não são fixos, podem aparecer ou não.

b) Entrada da Lista Permutada

Na linguagem, não foi encontrada nenhuma outra lista além da alfabética. Entretanto, quando o termo é proibido, remete-se para o termo controlado:

Ex: Termo Proibido COITO

Termo autorizado RELAÇÃO SEXUAL

d) Entrada lista Hierárquica

A linguagem não possui lista hierárquica.

4 Termos de Indexação

a) termos equivalentes em outros idiomas

Na lista alfabética, todos os termos são apresentados em português com exceção de termos em que o equivalente em português não é usado.

Ex: SOFTWARE

b) Forma dos Termos

- substantivos

Os descritores em sua maioria são substantivos.

Ex: ABDOME

- frases substantivadas

A linguagem não apresenta muitas frases substantivadas.

Ex: ABRIGO ANTINUCLEAR Ex: CONJUNTURA ECONÔMICA

- advérbio

Observou-se pouco uso de advérbio, encontramos estes funcionando como modificadores de conceitos.

Ex: AMAZONAS/ESTADO

- escolha de formas singular e plural

Percebeu-se que a maioria dos conceitos se encontram no singular.

Ex: RELAÇÃO SEXUAL

Mas existem algumas exceções que entram pelo plural.

Ex: RELAÇÕES EXTERIORES

Ex: RELAÇÕES COMUNITÁRIAS

- verbos

Há usos de verbo somente representado na forma de substantivos.

Ex: TRADUÇÃO (PROCESSO) INDICAÇÃO (PROCESSO)

- homógrafos (ou homônimos)

Os homógrafos são acompanhados de um modificador para evitar ambigüidade.

Ex: ANDRADE, JORGE/ESCRITOR

Ex: ANDRADE, JORGE/FUTEBOL

- siglas

As siglas são adotadas para termos nas categorias Descritor e Identificador.

Ex: DNA Termo proibido: ÁCIDO DESOXIRRIBONUCLEÍCO

Ex: OPAS Termo proibido: ORGANIZAÇÃO PAN-AMERICANA DE SAÚDE

5 Relações Semânticas e Lógicas

a) relação hierárquica

As relações de hierarquia mostradas através das relações TG e TE não estão presentes na linguagem do DEDOC.

b) relação associativa

Os termos associados são mostrados na tela de entrada e, ao mesmo tempo também, de recuperação de pesquisa, através da expressão TR encontrados nas três categorias.

Ex: RELAÇÃO SEXUAL

Ex: SANDY/CANTORA

TR FANTASIA SEXUAL

TR CHITÃOZINHO E XORORÔ/DUPLA SERTANEJA

TR MASTURBAÇÃO

JUNIOR/CANTOR/

TR ORGASMO

TR RELACIONAMENTO SEXUAL

TR REPRODUÇÃO HUMANA

TR SEXO

c) relação de equivalência

Os termos sinônimos são mostrados na tela modelo de ficha catalográfica usada para consulta da organização dos termos através da expressão Termo Proibido.

Ex: SANDY/CANTORA

Termos proibidos

LIMA, SANDY

LIMA, SANDY LEAH

SANDY/CANTORA MIRIM/

- **Avaliação da estrutura semântica da linguagem documentária do DEDOC aplicando os critérios para análise de exemplos com base na metodologia de Ginez de Lara (1993) considerou:**

- **categorias de assunto nas quais o descritor se insere**

A linguagem do DEDOC na tela de entrada após efetuada alguma busca permite identificar a categoria a que o termo pertence. Para comprovar isso realizaram-se buscas na Linguagem cujos exemplos abaixo evidenciam os resultados:

- 1) BERGAMOTA – Termo Proibido pertencente à categoria Descritor cujo termo autorizado é TANGERINA;
- 2) RELAÇÃO SEXUAL – Termo Controlado pertencente à categoria Descritor;
- 3) KENNEDY, Caroline – Termo Controlado pertencente à categoria Personalidades;
- 4) ABADIA, Maria de Lourdes – Termo controlado pertencente à categoria Personalidades;
- 5) AAA/ASSOCIAÇÃO ANTROPOLÓGICA AMERICANA – Termo Controlado pertencente à Categoria Identificador;
- 6) VALE DO RIO DOCE, CIA – Termo Controlado pertencente à Categoria identificador.

Destacamos, como se observa acima, que, entrando na tela de pesquisa e digitando o termo, se recupera a categoria a que o termo pertence.

- **análise dos termos genéricos**

A Linguagem do DEDOC não possui uma organização hierárquica entre os termos, não possuindo portanto termos genéricos.

- **análise dos termos relacionados**

Efetuuou-se busca na Linguagem para recuperar termos e seus termos relacionados, e analisar se ocorre reenvios recíprocos entre os termos.

- 1) VALE DO RIO DOCE, CIA

TR EMPRESA- VALE DO RIO DOCE DALBRAS

TR VALE DO RIO DOCE/EMPRESA
TR VALE DO RIO DOCE MINERAÇÃO

EMPRESA- VALE DO RIO DOCE DALBRAS
TR VALE DO RIO DOCE, CIA

VALE DO RIO DOCE/EMPRESA
TR VALE DO RIO DOCE, CIA

VALE DO RIO DOCE MINERAÇÃO
TR VALE DO RIO DOCE MINERAÇÃO, CIA
VALE DO RIO DOCE, CIA

- lógica de subordinação dos termos específicos

A Linguagem do DEDOC não possui termos específicos, não sendo possível verificar esse critério.

- consistência da relação de equivalência entre os termos

1) BERGAMOTA Termo Proibido	Termo Controlado TANGERINA
Termo Controlado TANGERINA	Termo Proibido BERGAMOTA
2) MEXERICA Termo Proibido	Termo Controlado TANGERINA
Termo Controlado TANGERINA	Termo Proibido MEXERICA
3) MORGOTE Termo Proibido	Termo Controlado TANGERINA
Termo Controlado TANGERINA	Termo Proibido MORGOTE
4) TANGERINA/FRUTA/ Termo Proibido	Termo Controlado TANGERINA
Termo Controlado TANGERINA	Termo Proibido TANGERINA/FRUTA

5) COITO Termo Proibido
Termo Controlado RELAÇÃO SEXUAL

Termo Controlado RELAÇÃO SEXUAL
Termo Proibido COITO

7) LIMA, SANDY Termo Proibido
SANDY/CANTORA
Termo Controlado SANDY/CANTORA

Termo Controlado
Termo Proibido LIMA, SANDY

8) LIMA, SANDY LEAH Termo Proibido
Termo Controlado SANDY/CANTORA

Termo Controlado SANDY/CANTORA
Termo Proibido LIMA, SANDY LEAH

9) SANDY/CANTORAMIRIM/ Termo Proibido
Termo Controlado SANDY/CANTORA
Termo Controlado SANDY/CANTORA
Termo Proibido SANDY/CANTORA MIRIM/

10) KENNEDY, Carol Termo Proibido
Termo Controlado KENNEDY, Caroline

Termo Controlado KENNEDY, Caroline
Termo Proibido KENNEDY, Carol

11) KENNEDY SCHLOSSBERG, CAROLINE Termo Proibido
Termo Controlado KENNEDY, Caroline

Termo Controlado KENNEDY, Caroline
Termo Proibido KENNEDY SCHLOSSBERG, CAROLINE

11) SCHLOSSBERG, CAROL Termo Proibido
Termo Controlado KENNEDY, Caroline
Termo Controlado KENNEDY, Caroline
Termo Proibido SCHLOSSBERG, CAROL

Com a observação dos exemplos acima, podemos afirmar que a Linguagem do DEDOC mantém uma consistência referente à relação de equivalência entre os termos.

- Explicação: o termo controlado é o termo autorizado para representar um assunto quando se tem vários termos sinônimos do assunto.

- **Análise Formal da Estrutura da Linguagem Documentária do Arquivo do jornal “O Estado de São Paulo” conforme metodologia de Corrêa (1998) destacando:**

I Aspectos Gerenciais

a) Compilação dos termos

Os termos foram criados com base no vocabulário controlado do jornal “NEW YORK TIMES” objetivando abranger todos os assuntos tratados na matéria jornalística. E a Linguagem foi elaborada por um antigo funcionário jornalista na instituição.

b) Números de termos

A linguagem do Arquivo em 1997 obtinha 947 descritores autorizados e 108 descritores proibidos. (sinônimos)

c) uso de Software e/ou Programa de Computador

A linguagem apresentava-se em suporte papel.

d) Instruções de uso/informações complementares

Encontram-se raramente informações sobre a forma de entrada de alguns descritores da linguagem e nenhuma menção de como usá-la, ressaltando porém que a mesma se encontra com falta de páginas.

e) Utilização de normas para controle de vocabulário

A linguagem apresenta remissivas **VER** indicada por V. que permitem proibir a entrada de termos sinônimos, entretanto, existem exemplos que V. indica termos relacionados e também as formas corretas de entrada para alguns termos.

f) Cobertura temática

Os temas cobertos pela Linguagem são diversos porque o Jornalismo trabalha com assuntos das diversas área do conhecimento humano.

Lista Alfabética

Os descritores estão dispostos em ordem alfabética de maneira não muito organizada distinta portanto dos padrões dos vocabulários controlados podemos destacar como exemplo que os descritores relacionados não trazerem a sigla TR como pode ser observado abaixo:

Ex: Acidente

- Acidente aéreo

- Brasil

- Acidente doméstico
- Acidente ecológico
- Acidente espacial
- Acidente ferroviário
- Acidente fluvial
- Acidente marítimo
- Acidente nuclear

c) Lista Permutada

A linguagem do Arquivo do jornal “O Estado de São Paulo” não fornece nenhum outro tipo de lista a não ser alfabética de descritores.

d) Lista Hierárquica

A linguagem do Arquivo do jornal “O Estado de São Paulo” não estabelece hierarquias entre os termos.

1 Apresentação da linguagem

Apresenta os descritores que deduzimos serem controlados e os relacionados seguidamente sem utilizar a sigla TR, e os termos sinônimos são indicados pela expressão VER entre parêntese ao lado do termo controlado, porém, como em alguns casos, com algumas falhas como mencionamos acima

a) Entrada

Descritor (v. indica-se o sinônimo a procurar)

Termos relacionados (indicados abaixo dos termos provavelmente controlados)

Ex: Acidente aéreo

Ex: Agricultura

Acidente aéreo – Brasil (*Provável TR*)

- Financiamento (v. Crédito rural)

b) Entrada da Lista Permutada

A Linguagem do “O Estado de São Paulo” não possui outra lista além da alfabética.

A) Entrada da Lista Hierárquica

Constatou-se que não apresenta hierarquia entre os termos.

1.2 Componentes da Entrada

A Linguagem avaliada possui: termos, termos relacionados e sinônimos.

A única lista existente é a alfabética onde os sinônimos estão indicados pela expressão VER representada por V., os descritores encontram-se sublinhados por um traço e os termos relacionados aparecem abaixo dos descritores sendo antecidos por hífen (-).

Ex: Agricultura

- Países
- Continentes
- Estados, Territórios e regiões e São Paulo
- Financiamento (V. Crédito rural)
- Preço

Ex: Arma

Armamento (v.t. Indústria bélica)

- Países
- Continentes
- qualificadores permitidos

Na lista alfabética o descritor é acompanhado dos qualificadores permitidos na indexação.

Ex: Automóvel Blindado

Automóvel a Álcool (V. Motor a Alcool)

- Categorias às quais o descritor pertence

A linguagem do “O Estado de São Paulo” constitui-se de uma listagem alfabética de termos não os classificando em nenhuma categoria.

- Notas explicativas

A Linguagem não apresenta notas explicativas referentes à organização dos termos.

- Notas de coordenação

A linguagem apresenta algumas notas referentes à entrada de termos.

Ex: Ensino primário e Ensino Secundário (usar sem subtítulo)

- notas de uso de qualificadores

Encontramos na linguagem somente indicações de que determinado termo possui um qualificador pois ao lado do termo encontra-se a expressão VER representada por V. seguida do qualificador indicado.

Ex: • Bebê

- Tráfico (V. Adoção)
- Troca

- Notas Gerais

A linguagem, com exceção dos tipos de notas que explicitamos acima, não apresenta nenhum outro tipo de nota.

2 Termos de indexação

- a) termos equivalentes em outros idiomas

Ex: Cinema Faroeste – Cinema western

- b) Forma dos Termos

- Substantivos

Os descritores em sua maioria são substantivo.

Ex: Boxe Ex: Boate

- frases substantivadas

Verificou-se o uso de frases substantivadas.

Ex: Indústria automobilística

- advérbios

Notou-se pouca presença de advérbios.

Ex: Educação para o trânsito

- adjetivos

Identificou-se a presença de adjetivos.

Ex: Administração Pública Municipal Ex: Armas Nucleares

- verbos

Há uso de verbos representados na forma de substantivos.

Ex: Apuração Ex: Exportação Ex: Habitação

- Homógrafos (ou homônimos)

Não encontramos a presença de termos homógrafos. Somente um termo com o mesmo significado acompanhado de um termo que podemos mencionar como modificador que distingue sua utilização.

Ex: Balão (esporte) ex: Balão (junino)

- siglas

Não foram encontradas siglas na linguagem.

- escolha de formas singular e plural

Na Linguagem os descritores em sua maioria encontram-se no singular mas observou-se também o uso de descritores no plural.

Ex: Exonerações Ex: Entorpecente

5 Relações Semânticas e lógicas

a) relação hierárquica

A linguagem não apresenta relações hierárquicas entre os termos.

b) relação associativa

As relações associativas existem mas não estão representadas por TR como também não possibilitam em sua maioria reenvios recíprocos.

Ex: Ecologia (v.t. Educação Ambiental, Crime Ecológico)

Remete para o seguinte exemplo: Educação

Ex: Crime Ecológico

Educação Ambiental

Educação artística

Educação Física

Educação para o trânsito

c) relação de equivalência

Os termos sinônimos na Linguagem são indicados ao lado do lado do termo proibido através da expressão VER representada por V. como:

Ex: 1- Boate (v. Casa Noturna – nome) 2- Ex: Arma Química (V. Guerra Química)

Recupera-se no primeiro exemplo Casa Noturna.

- Nome

Entretanto, para o exemplo de Arma Química não se identifica na Linguagem o termo Guerra Química.

Destacamos que agora deveríamos avaliar a aplicação dos critérios estabelecidos pela metodologia de Ginez de Lara (1993) na linguagem do Arquivo mas não foi possível porque a linguagem não possui uma organização lógica que nos permita identificar os componentes desses critérios como demonstra o anexo A.

Sintetizamos aqui os resultados obtidos na avaliação das Linguagens acima evidenciando que os parâmetros utilizados para uma análise formal da estrutura da linguagem DEDOC definidos por Corrêa (1998) revelaram-se eficazes para esta pesquisa. A Linguagem pode ser explicitada de um modo objetivo destacando-se os seus aspectos principais, ou seja, mais relevantes.

Foi constatado que ambas as Linguagens não apresentam uma hierarquia entre os termos que permita identificar termos genéricos e específicos, o que para fins de consistência na identificação de conceitos é uma barreira.

Na Linguagem do DEDOC, com relação à estrutura dos componentes da linguagem, a evidência do termo identificador é um fator que possibilita diminuir a ambigüidade dos termos sinônimos. E a estrutura da tela da linguagem de indexação utilizada para pesquisa se apresenta bem completa.

Com relação ainda à estrutura formal da linguagem do DEDOC, esta se apresenta em três categorias: Descritor (assuntos), Personalidades (pessoas) e Identificador (empresas, instituições e corporações) para o vocabulário; ela possibilita englobar o

conteúdo das matérias jornalísticas que tratam de assuntos variados e são as mesmas propostas por Ahmad (1991) na revisão da literatura sobre indexação de jornais.

Quanto à análise da estrutura formal da Linguagem do arquivo do jornal “O Estado de São Paulo”, os resultados demonstraram que a linguagem é uma listagem de assuntos ordenados em ordem alfabética, porém não de uma forma lógica, pois se encontram na Linguagem termos ambíguos que remetem para termo ambíguo equivalente, e não se identifica tal termo autorizado equivalente do primeiro termo ambíguo na linguagem, e essa mesma problemática acontece com os termos provavelmente relacionados que não possuem reenvios. Ainda existem casos em que a expressão VER significa remeter para termos equivalentes, em outros momentos, para termos relacionados. Quanto à expressão VER TAMBÉM, em alguns casos é usada para remeter aos termos relacionados, em outros, ela exerce a função de termos diferenciadores de sinônimos.

Devido à problemática encontrada na organização da Linguagem do Arquivo, não foi possível aplicar a metodologia de Ginez de Lara (1993) para a análise da estrutura semântica da Linguagem do Arquivo. (ANEXO D: Linguagem do Arquivo)

Com relação à aplicação da metodologia de Ginez de Lara (1993) à Linguagem do DEDOC, houve êxitos em sua aplicação porque os resultados obtidos permitiram identificar categorias da linguagem, como também outros dados importantes sobre os termos. Na análise dos termos relacionados, observou-se que os termos têm reenvios recíprocos. E a razão de equivalência mostrou-se consistente porque não permitiu duplo sentido de termos, além de recuperação pelo termo autorizado porque indica que o termo pesquisado é um termo proibido e que o autorizado é o indicado.

Na avaliação das linguagens, quanto à análise formal das estruturas propostas por Corrêa (1998), a linguagem do DEDOC apresenta uma estrutura organizada enquanto que na Linguagem do arquivo notaram-se falhas.

Em síntese, destacamos, com os resultados obtidos, que a linguagem do DEDOC, ao contrário do Arquivo mostra-se provavelmente como uma linguagem mais adequada para uma boa indexação.

E ainda evidenciaremos, na sequência, os resultados referentes à qualidade das linguagens, que foram obtidos no momento da análise da leitura dos indexadores fazendo uso do Protocolo Verbal quando se evidenciou a compatibilidade dos termos identificados e selecionados com as linguagens dos sistemas e, uso da linguagem pelos indexadores durante a leitura evidenciando se os indexadores realizavam associações com a linguagem.

Com o desenvolvimento do capítulo 5, tivemos oportunidade de conhecer um pouco mais sobre os sistemas de informação especializados que realizam a atividade de indexação de jornal e, que possuem acervo organizado, além de identificar um dado relevante que demonstrou que outros profissionais não-bibliotecários estão realizando, também, de forma eficiente, a atividade de indexação de documentos, embora, até onde sabemos, é disciplina pertencente à formação do indexador no curso de Biblioteconomia e Documentação.

No próximo capítulo, dividido em vários tópicos, realizamos a análise e discussão dos dados que foram coletados durante a gravação da leitura dos indexadores experientes do DEDOC e do Arquivo do jornal “O Estado de São Paulo”, considerando dados relevantes levantados nos capítulos de revisão de literatura. Posteriormente, no mesmo capítulo, fizemos um resgate das informações importantes levantadas na literatura e na discussão dos dados dos protocolos analisados, comparando os resultados obtidos para proposição de um modelo teórico de leitura para indexação de jornais.

CAPÍTULO 6: MODELO DE LEITURA PARA INDEXAÇÃO DE ARTIGOS DE JORNAIS: subsídios da observação da leitura do indexador

Neste capítulo, primeiramente mencionamos a metodologia adotada para aplicação do Protocolo Verbal. E na sequência analisamos e discutimos os resultados da observação da leitura dos indexadores experientes do Arquivo e do DEDOC. E posteriormente, elaboramos o modelo de leitura para indexação de jornais considerando os resultados obtidos com a observação da leitura e dados relevantes colhidos com a revisão de literatura descritas nos capítulos 1, 2 e 3.

• Metodologia de Aplicação do Protocolo Verbal no Arquivo do jornal “O Estado de São Paulo” e no Centro de Documentação da Editora Abril Cultural (DEDOC) com indexadores experientes visando subsídios para elaboração do modelo de leitura

A aplicação do Protocolo Verbal para a observação da atividade de leitura dos indexadores no contexto do arquivo do jornal “O Estado de São Paulo” e no Departamento de Documentação da Editora Abril obedeceu aos seguintes procedimentos:

1- Procedimentos anteriores à aplicação do Protocolo Verbal:

a) Seleção do Texto-Base

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

Temática do texto: os selecionados eram referentes aos cadernos dos jornais com os quais cada indexador trabalhava, variando-se, assim, a temática de cada texto.

Assim, para a observação da leitura dos indexadores do arquivo e do centro de documentação foram selecionados 10 textos (ANEXO B) sendo que cada indexador (caracterizados como sujeito 1, 2, 3, 4, e 5) indexou dois textos, sendo eles:

- Jornal “O Estado de São Paulo” - Caderno 2: “Lobão volta ao rock da melhor qualidade”- Texto 1; “Seminário prova que roupa também é documento” - Texto 2, indexados pelo indexador caracterizado como sujeito -1;
- Jornal “O Estado de São Paulo” - Caderno Cidades : “Execução de quadrilha afasta delegado no MA” - Texto 3; “Assaltante do BB de Campinas é preso”- Texto 4, que foram indexados pelo indexador caracterizado como Sujeito - 2.
- Jornal Folha de São Paulo - Seção Brasil: “Conflito entre PMs deixa cinco feridos” – Texto 5; “ Garotinho ajuda pedetista em São Gonçalo” – Texto 6 pertencente ao Jornal “O Estado de São Paulo”- Caderno Eleições 2000, que foram indexados pelo sujeito –3;
- Jornal “O Globo” – Seção RIO: “Agora é a vez do Cristo Redentor”- Texto 7; “Bailarina acusa o marido ator de espancamento e dá queixa na polícia” – Texto 8, que foram indexados pelo sujeito- 4;
- Jornal Folha de São Paulo – Caderno Geral – Cinema: “ Eu, Tu, Eles” é o candidato do Brasil ao Oscar” – Texto 9; “ Beck diz que agora vem” pertencente ao Caderno – Ilustrada, indexados pelo sujeito 5.

b) 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. Sendo assim foram selecionados cinco indexadores que foram classificados como sujeitos 1, 2, (pertencente ao Arquivo) 3, 4 e 5 (pertencente ao DEDOC).

c) **Conversa informal com os sujeitos** : foi realizada uma conversa informal com cada um dos sujeitos em que mencionaram-se 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-los 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.

d) **Familiarização com a tarefa do “Think Aloud” (“Pensar Alto”)** através do texto “Instruções aos sujeitos”. Antes da aplicação do protocolo verbal, foi realizada uma familiarização com a tarefa de “Pensar Alto” por meio da leitura do texto que contém “Instruções aos Sujeitos” (ANEXO C) fundamentado em NARDI (1993), tendo o propósito de descontrair e, ao mesmo tempo, apresentar procedimentos que podem auxiliar os sujeitos no desempenho da tarefa.

2- procedimentos durante a aplicação do protocolo verbal

e) **Gravação do “Pensar Alto” durante a leitura do texto – base**: Antes de iniciar a gravação foi entregue aos sujeitos o texto-base lembrando que é preciso “Pensar Alto” durante toda a leitura e exteriorizar seus processos mentais, procurando esquecer a presença da pesquisadora que estava presente somente para controlar o gravador e lembrar que é preciso “Pensar Alto”. Destacamos que a aplicação do protocolo verbal foi realizada, individualmente, com cada um dos indexadores, bem como a conversa informal, familiarização com a realização da tarefa e a leitura do texto “Instruções aos sujeitos”.

3- Procedimentos após a aplicação do Protocolo Verbal

f) **Transcrições literais das gravações da leitura dos sujeitos (APÊNDICE F)**: as transcrições foram feitas de maneira a destacar a compreensão dos sujeitos, suas dúvidas, equívocos, identificação e seleção de termos. E, para melhor visualização dos processos adotados pelos sujeitos, destacamos cada um deles com um tipo de notação específica como:

[REDACTED]

Itálico: comentários do leitor

[REDACTED]

..PS.. : pausa

[REDACTED]

...Bailarina acusa o marido e ator registra queixa de agressão. Briga de casal teria acontecido durante a festa de aniversário do filho. Ana Carolina Torres (...). 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 de anos, registrou queixa contra o marido, o ator Ademir Zanyor, de 26, na 32ª DP (jacarépagua), no Domingo passado. Ela o acusa de tê-la espancado durante a festa de aniversário de 2 anos de Uriel, filho único do casal. ..PS .. E você vai pro nome da bailarina (Nesse momento o indexador atribui como assunto da matéria o nome da bailarina) ... e o ator chama ... (Posteriormente o indexador comenta:). Entramos pelo nomes do agressor e da vítima então aqui lemos ... 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

Para uma melhor apresentação dos dados, o presente capítulo foi dividido em dois itens: primeiramente, a análise dos dados coletados com o Protocolo verbal buscando evidenciar na leitura dos indexadores, os aspectos da natureza metacognitiva elaborada por Brown (1980) e reformuladas por Kato (1986); num segundo momento, destacamos a discussão dos resultados obtidos com a análise dos dados coletados com o Protocolo Verbal.

- A evidência de estratégias metacognitivas listadas por Brown (1980) e reformuladas por Kato (1987);
- A natureza normativa de certas estratégias: algumas recomendações da Norma ISO (5963 de 1985) direcionando-as ao texto jornalístico;

- A evidência de estratégias metacognitivas listadas por Brown (1980) e reformuladas por Kato (1987);
- A natureza normativa de certas estratégias: algumas recomendações da Norma ISO (5963 de 1985) direcionando-as ao texto jornalístico;

- Estratégias do comportamento do indexador durante a leitura: evidenciaremos os momentos em que o indexador identifica, ressaltando que realiza essa atividade considerando o conteúdo do documento e, no mesmo momento, seleciona o conceito mantendo em mente as necessidades informacionais dos usuários. São atividades (identificação e seleção de conceitos) simultâneas. E justificamos nossa opção em observar esse comportamento dos indexadores porque nos dados coletados com o Protocolo Verbal por Fujita (2001), para observação da leitura dos indexadores da Sub-Rede Nacional em Ciências da Saúde Oral e do Centro de Energia Nuclear, desvendou-se a ocorrência desse comportamento dos indexadores na indexação de artigos científicos e resumos.

Os parâmetros se encontram organizados no quadro abaixo:

Quadro 2: Identificação da natureza metacognitiva na leitura documentária

Aspectos metacognitivos de leitura estabelecidos por BROWN KATO	Recomendações da Norma ISSO 5963 adaptadas ao texto jornalístico	Outros aspectos do comportamento do indexador durante a leitura
▪ Explicitação dos objetivos da leitura e/ou manutenção dos objetivos na mente; ▪ Identificação de aspectos importantes da mensagem; ▪ Alocamento de atenção a áreas importantes: → Exploração da estrutura textual ▪ Monitoração 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; → Recobrimento de atenção Quando a mente se distrai ou faz digressões.	• Análise do documento com domínio da estrutura textual, considerando partes do texto é uma recomendação da Norma. Verificaremos a exploração da estrutura textual do texto jornalístico proposta por VAN DIJK (1983) na leitura dos indexadores; • Identificação de conceitos (abordagem sistemática mediante questionamento) • Seleção de conceitos	• Associação com linguagem; • Seleção de conceitos (Termos ou descritores).

▪ **Análise dos Dados Coletados com o Protocolo Verbal**

Evidenciamos aqui a análise dos dados coletados com o Protocolo Verbal durante a leitura dos indexadores visando identificar aspectos metacognitivos (conscientes).

Análise dos dados do Sujeito 1 – Indexador do arquivo do jornal “O Estado de São Paulo” coletados através da Técnica do Protocolo Verbal, durante a leitura dos textos 1 e 2.

1. Explicitação dos objetivos da leitura

Texto - 1

Quase no final da leitura o sujeito evidenciou estar mantendo em mente o objetivo da leitura, ou seja, como o artigo pode ser recuperado.

... Quando a gente indexa é discos assim música, quando a pessoa que é sem que está sendo indexada não é famosa no caso o Lobão vai se fácil procurar porque o jornalista quando vier recuperar a matéria ele vai procurar direto por Lobão....

Texto - 2

Durante toda a leitura do texto 2, o sujeito não explicitou, em nenhum momento, os objetivos para os quais a estava realizando, talvez devido ao fato de já ter citado no texto anterior que o objetivo era facilitar o acesso à informação para a comunidade usuária (jornalistas).

2. Identificação de (e alocamento de atenção a) aspectos importantes da mensagem fazendo uso da estrutura textual.

Texto - 1

Logo no início da leitura o sujeito identifica aspectos que considera importantes no texto e faz predição do que poderia encontrar sobre o assunto do artigo.

Ex.: *Então a primeira página do Caderno tá hoje, tá tratando do Lobão é o cantor. Então a primeira coisa que eu faço vou direto no colaborador para não esquecer que existe o colaborador...*

sempre vou no fim da matéria pra ver si indica a página se continuidade de página porque o Caderno 2 sempre tem essa característica quase sempre vem numa página e continua na outra no caso de hoje continua na página 9. Então eu vou na página 9 trago perto da página 1 porque o assunto provavelmente é uma continuidade e é mesmo tá vendo, Lobão

Texto - 2

O sujeito aloca atenção ao final do texto que considerou importante, identifica as personagens da matéria como aspectos importantes da mensagem, ou seja, possíveis descritores de recuperação da informação quando o assunto tratado é moda .

Tá é um projeto de moda eu vou entrar no Joel Birman e na Cristina Franco que são os dois que tão promovendo essa conferência, daí pra frente vai volta a falar deles a primeira apresentação pública começou a dar os primeiros frutos a dupla foi convidada tá não tem mais o que ficha ali.

3. Monitoração do comportamento para ver se está ocorrendo compreensão: Engajamento em revisão e auto-indagação para ver se o objetivo está sendo atingido (a)/ Tomada de ações corretivas quando são detectadas falhas na compreensão (b).

Texto - 1

O sujeito deixou claro ter conhecimento prévio da estrutura textual e da forma em que informações são apresentadas no artigo. Tem consciência de que já apreendeu o assunto, de que seu objetivo foi atingido:

(a) *Olha eu não preciso mais ler que eu já sei que o resto da matéria tá falando dele, não me interessa mais nada, entendeu eu já paro de ler.*

(b) Não evidenciou porque não cometeu equívocos a serem corrigidos.

Texto - 2

O sujeito confirmou no meio do texto que o assunto tratado é moda e deixou claro ter conhecimento das informações que seriam apresentadas na continuidade do artigo.

(a) *Tá é um projeto de moda eu vou entrar no Joel Birman e na Cristina Franco que são os dois que tão promovendo essa conferência, daí pra frente vai volta a falar deles a primeira apresentação pública começou a dar os primeiros frutos a dupla foi convidada tá não tem mais o que fichar ali.*

(b) O sujeito em nenhum momento realizou voltas a trechos do texto para poder compreendê-los melhor, por não ter encontrado problema na compreensão.

4. Recobrimento de atenção quando a mente se distrai ou faz digressões.

Texto - 1

Não evidenciou esse tipo de comportamento porque não encontrou problema de compreensão.

Texto 2

Não evidenciou porque não cometeu equívocos a serem corrigidos.

Análise dos dados do Sujeito 2 - Indexador do arquivo do jornal “O Estado de São Paulo” coletados através da técnica do Protocolo Verbal durante a leitura dos textos 3 e 4.

1. Explicitação dos objetivos da leitura

Texto - 3

Durante a leitura não explicitou em nenhum momento os objetivos para os quais estava realizando a leitura.

Texto - 4

O sujeito não explicitou em nenhum momento a finalidade da leitura.

2. Identificação de (e alocamento de atenção a) aspectos importantes da mensagem

Texto - 3

No final do texto, através de uma leitura silenciosa, identificou aspectos importantes do texto.

... eu preciso saber se vou colocar na Polícia do Maranhão que é um delegado ou só no crime, mas pelo que eu tô entendendo aqui a Polícia que matou essa quadrilha certo não foi briga de quadrilhas então a Polícia aqui no caso é importante tá eu facho faço dois descritores Criminalidade Maranhão e a Polícia Maranhão, tá.

Texto - 4

O sujeito evidenciou ter identificado aspectos importantes da mensagem.

... Oh! não é o a o assalto que eu estava pensando não foi outro assalto foi no início da semana eu achei que era um outro assalto, tá vendo como que é a gente então a gente sempre tem que fazer a leitura dinâmica sempre tem que fazer viu como como é que é tá, então aqui no caso eu posso tirar o Banco do Brasil que não era um caso que já vinha acontecendo e é sempre assim, se é se é a matéria for enfocar é der continuidade é assaltar o Banco do Brasil eu facho o Banco do Brasil, não é aqui o caso não é importante Banco do Brasil, Banco do Brasil eu ficharia se alguma resolução do Banco, medidas econômicas, entendeu.

Durante a leitura, o sujeito identificou aspectos importantes da mensagem que causaram mudança na sua compreensão.

...Oh! não é o a o assalto que eu estava pensando não foi outro assalto foi no início da semana eu achei que era um outro assalto

3. Monitoração do comportamento para ver se está ocorrendo compreensão: Engajamento em revisão e auto-indagação para ver se o objetivo está sendo atingido (a)/ Tomada de ações corretivas quando são detectadas falhas na compreensão (b)

Texto - 3

Sujeito 2: Texto 1: O sujeito explicitou sua monitoração do comportamento procurando verificar sua compreensão na escolha dos descritores do texto.

... Aqui eu tô lendo mais atentamente ...

O sujeito, durante a leitura, se indaga quanto à escolha dos descritores que irão representar o texto.

(a) *...então eu preciso saber se vou colocar na Polícia do Maranhão que é um delegado ou só no crime, mas pelo que eu tô entendendo aqui a Polícia que matou essa quadrilha certo não foi briga de quadrilhas então a Polícia aqui no caso é importante tá eu ficho faço dois descritores Criminalidade Maranhão e a Polícia Maranhão, tá.*

OBS: Quanto ao item (b), o sujeito durante a leitura não realizou voltas a trechos do texto que denotassem tomadas de ação corretivas, para poder compreendê-los melhor.

Texto - 4

O sujeito, com base somente no título do artigo, explicitou ter conhecimento do assunto do artigo utilizando-se da estratégia de “feedback” à memória, ou seja, usou seu conhecimento prévio de artigos anteriores.

(a) *... BB é o Banco do Brasil tá, isso aqui também é um caso que aconteceu faz tempo e a gente tem o título Banco Assalto tá, no caso aqui a gente fichou também Banco do Brasil tá porque foi muito importante esse crime tá ele roubou milhões e milhões lá no...*

Durante a leitura do texto o sujeito faz uma revisão na compreensão que havia inferido somente pelo título do artigo e toma decisões sobre como agir para corrigir seu equívoco:

(a) *... Oh! não é o a o assalto que eu estava pensando não foi outro assalto foi no início da semana eu achei que era um outro assalto (a), tá vendo como que é a gente então a gente sempre tem que fazer a leitura dinâmica sempre tem que fazer viu como é que é tá, então aqui no caso eu posso tirar o Banco do Brasil(b) que não era um caso que já vinha acontecendo e é sempre assim, se é se é a matéria for enfocar é der continuidade é assaltar o Banco do Brasil eu ficho o Banco do Brasil, não é aqui o caso não é importante Banco do Brasil, Banco do Brasil eu ficharia se alguma resolução do Banco, medidas econômicas, entendeu*

(b) Não realizou voltas a trechos do texto.

4. Recobrimento de atenção quando a mente se distrai ou faz digressões

Texto - 3

Não evidenciou.

Texto 4

O sujeito durante a leitura recobra sua atenção para desfazer a compreensão que havia obtido somente através do título.

...Oh! não é o a o assalto que eu estava pensando não foi outro assalto foi no início da semana eu achei que era um outro assalto...

Análise dos dados do Sujeito 3 - Indexador do Centro de Documentação da Editora Abril coletados através da técnica do Protocolo Verbal durante a leitura dos textos 5 e 6.

1. Explicitação dos objetivos da leitura;

Texto 5

O indexador não evidenciou em nenhum momento o objetivo da leitura.

Texto 6

... como é o Garotinho uma figura importante eu vou por como termo o nome dele Garotinho ...

Percebemos, com a análise da frase acima, que o indexador deixa implícito que o objetivo da leitura é a recuperação pelos usuários do sistema, à medida que assume que o termo GAROTINHO é conhecido e, pode servir de ponto de acesso em uma tentativa de recuperar a matéria.

... Mas eu tô vendo que são coisas irrelevantes então eu não vou por nome do Garotinho ...

Posteriormente, em um segundo momento, pudemos observar que o objetivo da leitura além de direcionada para a recuperação objetiva também o era para o conteúdo enfocado na matéria, porque o indexador após ter identificado o termo GAROTINHO no início da leitura, não o seleciona como termo de indexação no final da leitura do texto por não ser pertinente ao conteúdo abordado na matéria.

2. Identificação de (e alocamento de atenção a) aspectos importantes da mensagem.

Texto 5

Logo no início da leitura, o indexador identifica aspectos que considera importantes para se compreender o conteúdo do texto.

Ex: . A greve dos policiais de fato teve um conflito é mais o conflito é dentro da greve então eu vou colocar greve policia vou colocar como termo policia militar

Texto 6

... e vou por nome da cidade que no nome da pasta não vai constar só consta Rio de Janeiro-Eleição Municipal então eu vou por São Gonçalo barra Rio de Janeiro São Gonçalo/Rio de Janeiro.

O indexador possui conhecimento da linguagem do sistema e, percebendo que a mesma não lhe permite especificar melhor o local onde acontece o fato da matéria, ele

inclui alterações na representação do termo local. Podemos perceber com isso que a linguagem é flexível porque permite ao indexador fazer, quando necessário, alterações quanto a representações de termos, e observamos ainda que a política de indexação do sistema nos demonstra que o indexador tem liberdade no momento de indexar as matérias.

**3. Monitoração do comportamento para ver se está ocorrendo compreensão:
Engajamento em revisão e auto-indagação para ver se o objetivo está sendo atingido
(a)/ Tomada de ações corretivas quando são detectadas falhas na compreensão (b).**

Texto 5

- (a) O indexador realizou auto-indagação para esboçar a sua compreensão sobre o que o texto estava abordando para, posteriormente, determinar mais um dos termos que realmente representaria o assunto enfatizado pelo autor, e que são relevantes para a recuperação do texto.

(a) ... *É protesto porque eles falam do protesto e o lugar onde está sendo feita a greve que é Recife é só Recife então Recife-Pernambuco ...*

- (b) Não evidenciou porque não cometeu equívocos a serem corrigidos.

Texto 6

(a) ... *como é o Garotinho uma figura importante eu vou por como termo o nome dele Garotinho ...*

O indexador, em um primeiro momento da leitura do texto, visando atingir um dos objetivos da leitura, no caso, a recuperação da matéria pelos usuários, indaga se é interessante ter o nome do governador como um dos termos de indexação, porque esse termo apresenta muitas possibilidades de buscas pelos usuários.

(b) ... Mas eu tô vendo que são coisas irrelevantes então eu não vou por nome do Garotinho então vai ficar vou por PDT porque o candidato é do PDT então vou por PDT barra partido PDT/Partido

Em um segundo momento, o indexador deixa claro que seu objetivo de indexação de extrair termos visando a recuperação da matéria pelos usuários está relacionado com o objetivo de representar também o conteúdo abordado na matéria. Percebemos, assim, que o indexador tem preocupação de realizar o que o Lancaster (1993) alerta que é a representação (termos de indexação ou resumo) fidedigna do conteúdo do documento.

4. Recobrimento de atenção quando a mente se distrai ou faz digressões;

Texto 5

Não evidenciou porque não cometeu equívocos a serem corrigidos.

Texto 6

... Mas eu tô vendo que são coisas irrelevantes então eu não vou por nome do Garotinho ...

O indexador nesse momento corrige o equívoco cometido durante o início da leitura onde visava selecionar o termo GAROTINHO por acreditar que o mesmo seria passível de recuperação e representante do conteúdo da matéria.

Análise dos dados do Sujeito 4 - Indexador do Centro de Documentação da Editora Abril coletados através da técnica do Protocolo Verbal durante a leitura dos textos 7 e 8.

1. Explicitação dos objetivos da leitura;

Texto 7

... se alguém quiser saber alguma coisa sobre o que aconteceu à época do Luis Conde então eu vou colocar e eles recuperam a matéria a troca da gestão do Pão de Açúcar ...

O indexador explicita que indexa objetivando a recuperação da matéria pelos usuários.

Texto 8

Durante a leitura do texto 8, o indexador deixa transparecer que o objetivo da leitura é a recuperação da matéria pelos usuários. Podemos perceber que, ao mesmo tempo, o indexador demonstra conhecer o funcionamento do sistema de busca da instituição.

(...) « Entao, 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 e também porque nosso sistema trabalha por palavra se voce colocar O Globo Jacarépagua todas as matérias que trouxerem a palavra Jacarépagua do jornal O Globo ele vai te jogar na tela.

2. Identificação de (e alocamento de atenção a) aspectos importantes da mensagem.

Texto 7

... esta materia é uma matéria rotineira é não teria grande importância se não fosse um monumento cartão postal do Rio de Janeiro como um monumento é quando se fala em Rio de Janeiro se fala em Pão de Açúcar se fala em Cristo Redentor essas coisas todas.

... então estão querendo dar uma nova estrutura a essas coisas que são coisas tradicionais na cidade do Rio de Janeiro é o Pão de Açúcar, o Corcovado, o bonde de Santa Tereza, coisas relativas à cultura da cidade ...

O indexador utiliza de seu conhecimento de mundo para verificar a importância de determinados assuntos dentro do contexto da matéria, com base no nível de importância desses assuntos num contexto extra-texto que vem a considerar esses itens como integrantes da cultura da cidade do Rio de Janeiro.

... então Cristo Redentor é um monumento e tá no título não há necessidade de colocar porque o sistema de busca realiza busca por qualquer palavra dentro da matéria inclusive do título ...

Analisando esse trecho da transcrição podemos destacar que o indexador tem conhecimento do funcionamento do sistema de busca da instituição, e utiliza esse conhecimento para definir os termos de indexação que deverão ser selecionados.

Texto 8

(...) Entao, 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...

O indexador demonstra ter conhecimento de que termos se deve selecionar quando as personagens envolvidas são pessoas, deixando subentendido que seleciona os nomes das pessoas e o assunto que envolve essas personagens.

3. Monitoração do comportamento para ver se está ocorrendo compreensão: Engajamento em revisão e auto-indagação para ver se o objetivo está sendo atingido (a)/ Tomada de ações corretivas quando são detectadas falhas na compreensão (b).

Texto 7

(a) (...)Então eles estão tentando tirar a concessão do bondinho e passar pra uma outra companhia e assim com o Corcovado também o Cristo Redentor também então aqui eu pego coisas normais eu ponho bonde ele vai como uma referencia

(a) se alguém quiser saber alguma coisa sobre o que aconteceu à época do Luis Conde então eu vou colocar e eles recuperam a matéria ...

(a) Essa aqui é uma matéria bastante simples não tem muito o que falar sobre ela então a pasta que ela vai ficar é Monumento-Brasil como então Cristo Redentor é um monumento e tá no título não há necessidade de colocar porque o sistema de busca realiza busca por qualquer palavra dentro da matéria inclusive do título então eu coloco o termo referente à gestão em que está sendo trocada sendo o termo CONDE, Luis Paulo e Cristo Redentor já está dentro da matéria e Monumento ...

(a) ... Tá vendo como eu falei eles falam dessas coisas ...

O indexador, com a leitura do trecho citado acima, reafirma a informação prévia mencionada em trechos anteriores onde destaca que os assuntos Pão de Açúcar, Corcovado são pontos turísticos primordiais da cidade do Rio de Janeiro assim como o são os termos que representaram a matéria.

(b) Não cometeu equívocos.

Texto 8

(a) 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 e também porque nosso sistema trabalha por palavra se você colocar O Globo Jacarépagua todas as matérias que trouxerem a palavra Jacarépagua do jornal O Globo ele vai te jogar na tela.

(a) ... e não coloco agressão como assunto porque o nome da pasta já informa que houve uma agressão contra mulher.

O indexador deixa transparecer que tem preocupação de registrar como assunto representante do conteúdo dessa matéria o termo agressão contra a mulher, e indaga se atingiu esse objetivo durante a indexação no momento em que optou por enviar a matéria para a pasta de assunto denominada agressão contra a mulher, porque os usuários podem recuperar a matéria além dos termos de indexação, também pelo nome da pasta.

(b) Não cometeu equívocos a serem corrigidos.

2. Recobrimento de atenção quando a mente se distrai ou faz digressões;

Texto 7

Não evidenciou porque não cometeu equívocos a serem corrigidos.

Texto 8

Não evidenciou porque não cometeu equívocos a serem corrigidos.

Análise dos dados do Sujeito 5 - Indexador do Centro de Documentação da Editora Abril coletados através da técnica do Protocolo Verbal durante a leitura dos textos 9 e 10.

1. Explicitação dos objetivos da leitura;

Texto 9

... eu coloco como termo para na hora que você for procurar dentro da matéria procura colocar termos que possam ter procura no caso aqui ele (autor) e possibilidades de busca pelo título.

... pode colocar atores que são casos de buscas como a Regina Casé como termo porque as vezes as pessoas podem lembrar dela e não lembrar o nome do filme aquele filme que ela fez e tal, (nesse momento o indexador inclui o nome da atriz como assunto do filme) o outro é o Lima Duarte ...

(...) porque na busca você pode cruzar os termos porque se você colocar somente indicação vai recuperar indicação até de efeito piolho mas sempre você vai acabar cruzando porque você vai falar há os filmes brasileiros que foram indicados ao Oscar digita indicação Oscar vem esse daqui e vai vim aqueles anteriores, há mais é aquele do Andrucha daí você coloca Andrucha né aí Andrucha, indicação já mata tá.

Percebemos, com os trechos destacados acima, que o indexador deixa claro que indexa objetivando selecionar assuntos que pertençam ao conteúdo da matéria e que sejam também passíveis de recuperação pelos usuários.

Texto 10

... Tem que pensar como o pesquisador vai procurar aquilo? Como facilitar a procura para o pesquisador porque para você tá com a matéria na mão tá fácil, tá vendo a matéria tal mas você pensa daqui um mês como você procuraria essa matéria? Então tem que se procurar facilitar com a atribuição de termos para o cara para facilitar a recuperação dele. Então o lance é esse. Porque para você tá fácil tá claro mas é aquela coisa do telefone sem fio onde o último lá pode estar pensando outra coisa. Então você tem que imaginar como o pesquisador vai pesquisar determinada matéria e procurar facilitar a busca.

Observamos, pelo trecho acima, que o indexador deixa claro que indexa objetivando representar o conteúdo da matéria e a sua recuperação pelos usuários.

2. Identificação de (e alocamento de atenção a) aspectos importantes da mensagem.

Texto 9

... Essa matéria esta no caderno cotidiano mas retranca para a ilustrada. Então quando eu for indexar a ilustrada essa matéria é a primeira que eu vou pegar. (Nesse momento o indexador pega o caderno ilustrada para localizar a matéria e comenta:). É mais a matéria saiu somente no cotidiano mesmo, mas no jornal eletrônico ele está na seção da ilustrada que é área dela tá.

Podemos perceber que o indexador tem conhecimento da estrutura textual do jornal (organização e classificação de assuntos dentro de determinados cadernos) porque, com base somente na leitura do texto, menciona que a matéria pertence ao caderno Ilustrada e que no atual caderno que é o Cotidiano deveria somente dar um pequeno enfoque à matéria e que a matéria integral estaria na Ilustrada que é o caderno que agrupa matérias referentes à arte, cultura, cinema.

... e tem um outro cara que é legal colocar como ator o terceiro marido mas nós não temos o nome dele aqui na matéria então eu não vou colocar.

O indexador recorre à memória de longo prazo em busca de informações que adquiriu (por meio de indexações anteriores referentes ao filme, ou quando assistiu o filme entre outras situações), que lhe possibilitarão lembrar o nome da terceira personagem da matéria por acreditar que este seja importante para a recuperação da matéria.

... E aí depois você coloca os termos tá por exemplo os termos do filme, ou peças de teatro vão para o diretor, autor isso é para a pasta autor – diretor tá ..

(Você coloca somente indicação não fica muito geral?) Não porque na busca você pode cruzar os termos porque se você colocar somente indicação vai recuperar indicação até de efeito piolho mas sempre você vai acabar cruzando porque você vai falar há os filmes brasileiros que foram indicados ao Oscar digita indicação Oscar vem esse daqui e vai vim

aqueles anteriores, há mais é aquele do Andrucha daí você coloca Andrucha né aí Andrucha, indicação já mata tá.

O indexador deixa claro nos trechos acima que tem conhecimento dos assuntos relevantes na área (Arte, Cinema) em que trabalha, como também das formas de buscas, demonstrando que visa imaginar como os usuários poderiam recuperar determinada matéria, e conhecimento prévio sobre o que indexar (termos), e do sistema de recuperação.

Texto 10

.. Aqui por exemplo nos campos termos é feito o controle de vocabulário dos dois (termos e nomes das pastas) então aq invés deu colocar o Beck eu vou para o vocabulário que vêm todos os Beck que tem.

... (pesquisadora: e no caso você não entra pelo nome dele como assunto porque pode ser recuperado pelo nome da pasta?). Pode se você der Back no campo de pesquisa já vem essa matéria e ainda você pode recuperar por pasta, palavra-chave todas as palavras do texto, por publicação (jornal).

Percebemos com os trechos acima que o indexador demonstra obter conhecimento das formas de buscas, ou recuperação, do sistema de informação e da necessidade de estabelecer um controle do vocabulário.

3. Monitoração do comportamento para ver se está ocorrendo compreensão: Engajamento em revisão e auto-indagação para ver se o objetivo está sendo atingido (a)/ Tomada de ações corretivas quando são detectadas falhas na compreensão (b).

Texto 9

(a) .. Então coloquei como termo de busca a Regina Casé, nome do filme, o Lima Duarte (ator). I é só.

O indexador indaga quais os termos que selecionou da matéria e posteriormente conclui que esses termos são suficientes.

(b) ... Olha aqui Regina Casé foi à sede da Conspiração comemorar a indicação.

“Esse filme merece representar o Brasil porque não traz só denúncias nem nos

mostra cantando e dançando o tempo todo.” (...) *vou ter que incluir termos* A indicação do Oscar *há 40 não 12 matérias sobre o filme desde seu lançamento mas o que marca essa matéria é a indicação pro Oscar. Porque poderia ser uma crítica, lançamento mas o que está marcando aqui essa matéria é a indicação então para facilitar a busca eu vou colocar termos agora como prêmio Oscar ...*

Nesse momento, o indexador percebe que estava faltando um termo muito relevante para representar o conteúdo da matéria que era uma situação nova quando comparada às matérias anteriores publicadas sobre o filme “Eu, Tu e Eles”, e que esse fator deveria ser indicado nos termos de indexação, isto é, que o filme agora havia sido indicado para concorrer ao Prêmio Oscar.

Texto 10

(a) .. *O que chama atenção nessa matéria é que ele vai vir no Rock in Rio 3 é uma entrevista dele é sobre um show então coloco como termos entrevista e show é genérico (nesse momento o indexador realiza uma “olhadela” pelo texto procurando encontrar um termo mais específico que represente a matéria) mas é só esses mesmos.*

(a) ... (pesquisadora: e no caso você não entra pelo nome dele como assunto porque pode ser recuperado pelo nome da pasta?). *Pode se você der Back no campo de pesquisa já vem essa matéria e ainda você pode recuperar por pasta, palavra-chave todas as palavras do texto, por publicação (jornal)...*

Percebemos que, nesse momento, o indexador indaga sobre o que a matéria trata e se interroga sobre os termos que selecionou para representar o conteúdo da matéria concluindo que os termos que selecionou são os únicos representantes do conteúdo da matéria mesmo.

Observamos também que o indexador tem conhecimento de como funciona o sistema de busca e demonstra que considera esse fator no momento da indexação.

4. Recobrimento de atenção quando a mente se distrai ou faz digressões;

Texto 9

... Olha aqui Regina Casé foi à sede da Conspiração comemorar a indicação. “Esse filme merece representar o Brasil porque não traz só denúncias nem nos mostra cantando e dançando o tempo todo.” (...) vou ter que incluir termos A indicação do Oscar há 40 não 12 matérias sobre o filme desde seu lançamento mas o que marca essa matéria é a indicação pro Oscar. Porque poderia ser uma crítica, lançamento mas o que está marcando aqui essa matéria é a indicação então para facilitar a busca eu vou colocar termos agora como Prêmio Oscar ...

O indexador recobrou sua atenção quando percebeu que estava cometendo o equívoco de indexar a matéria como se apresentasse a mesma situação (mesmo conteúdo, ou seja, comentários sobre o filme referentes às suas personagens, a mensagem que transmitia) das anteriores já publicadas sobre o assunto.

Texto 10

O indexador não evidenciou recobrimento de atenção porque não cometeu equívocos que necessitassem serem corrigidos.

▪ **Discussão dos Resultados da Análise dos Dados Coletados com Protocolo Verbal**

Discutiremos aqui os resultados da observação da leitura do indexador destacando novamente que, nesse estudo que utilizou a técnica do Protocolo Verbal, foi possível verificarmos a compreensão dos indexadores do arquivo do jornal “O Estado de São Paulo” (caracterizados como sujeitos 1 e 2), e os do Centro de Documentação da Editora Abril (caracterizados como sujeitos 3, 4 e 5), durante a atividade de indexação de jornais porque utilizamos um parâmetro de análise que nos permitiu identificar os aspectos metacognitivos listados por Brown (1980) reformuladas por Kato (1986), as recomendações de identificação de conceitos da Norma ISO 5963 adaptadas ao texto jornalístico e a exploração da estrutura textual de texto jornalístico proposto por Van Dijk (1983). Os aspectos abaixo, que compõem o parâmetro, são os mais representativos para a leitura com fins documentários:

- a) sequência de operações (aspectos) da leitura dos sujeitos;
- b) manter em mente o seu objetivo de representar o texto para futura recuperação

- c) exploração de seu conhecimento de estruturas textuais
- d) identificação de aspectos importantes da mensagem para seleção de conceitos;
- e) domínio da linguagem do sistema como conhecimento prévio.

Os quadros-resumos apresentados, a seguir, procuram evidenciar os aspectos notados na leitura.

a) Sequência de operações (aspectos) da leitura dos sujeitos

A verificação de procedimentos da leitura dos indexadores de jornais abordou primeiramente, a sequência de operações, sendo que essas operações são os aspectos da natureza metacognitiva da leitura, listados por Brown (1980) e reformulados por Kato (1986). Para isso, sintetizamos a sequência de operações obtidas a partir da análise dos dados coletados. Os quadros 3, 4, 5 e 6 a seguir, demonstram a sequência de operações realizadas pelos 5 sujeitos (indexadores) durante a leitura, criando também, um parâmetro de comparação entre a sequência da leitura de cada indexador, buscando verificar indícios da presença de um padrão de leitura.

Destacamos que, na leitura dos indexadores, foi identificada a existência de dez operações simultâneas sendo:

- Explicitação dos objetivos da leitura e/ou manutenção dos objetivos na mente;
- Identificação de aspectos importantes da mensagem;
- Alocamento de atenção a áreas importantes;
- Exploração da estrutura textual;
- Monitoração do comportamento para ver se está ocorrendo compreensão;
- Engajamento em revisão e auto-indagação para ver se o objetivo está sendo atingido;
- Tomadas de ações corretivas quando são detectadas falhas na compreensão;
- Recobrimento de atenção quando a mente se distrai ou faz digressões;
- Associação com Linguagem;
- Seleção de conceitos para indexação.

Quadro 3: Análise das operações (atividades) desenvolvidas pelos sujeitos 1 e 2 (indexadores do Arquivo do jornal “O Estado de São Paulo”) durante a leitura dos textos:

Sujeito 1 – Textos 1 e 2	
Texto 1	Texto 2
Identificação de aspectos importantes da mensagem	Alocamento de atenção a áreas importantes (título)
Alocamento de atenção a áreas importantes	Alocamento de atenção a áreas importantes (colaborador)
Exploração da estrutura textual	Identificação de aspectos importantes da mensagem
Seleção de conceitos para indexação	Seleção de conceitos para indexação
Sujeito 2 – Textos 3 e 4	
Texto 3	Texto 4
Alocamento de atenção à áreas importantes	Uso equivocado do conhecimento prévio
Identificação de aspectos importantes da mensagem	Associações com linguagem
Seleção de conceitos para a indexação	Monitoração do comportamento para ver se está ocorrendo compreensão
	Seleção de conceitos para a indexação

Quadro 4: Análise das operações (atividades) desenvolvidas pelos sujeitos 3 e 4 (indexadores do Centro de Documentação da Editora Abril) durante a leitura dos textos:

Sujeito 3 – Textos 5 e 6	
Texto 5	Texto 6
Alocamento de atenção á áreas importantes (seleção de itens para preencher a ficha de indexação)	Alocamento de atenção a áreas importantes (seleção de itens para preencher ficha de indexação)
- Identificação de aspectos importantes da mensagem	Identificação de aspectos importantes da mensagem
Seleção de conceitos para indexação	Explicitação dos objetivos da leitura
Engajamento em revisão e auto-indagação para ver se o objetivo está sendo atingido	Engajamento em revisão e auto-indagação para ver se o objetivo está sendo atingido
Seleção de conceitos para indexação	Tomadas de ações corretivas quando são detectadas falhas na compreensão
	Recobrimento de atenção quando a mente se distrai ou faz digressões
	Seleção de conceitos para indexação
Sujeito 4 – Textos 7 e 8	
Texto 7	Texto 8
Alocamento de atenção a áreas importantes	Alocamento de atenção a áreas importantes
Identificação de aspectos importantes da mensagem	Identificação de aspectos importantes da mensagem
Engajamento em revisão e auto-indagação para ver se o objetivo está sendo atingido	Seleção de conceitos para indexação
Seleção de conceitos para indexação	Engajamento em revisão e auto-indagação para ver se o objetivo está sendo atingido
Explicitação dos objetivos da leitura	Exploração da estrutura textual
Identificação de aspectos importantes da mensagem	Associação com Linguagem
Engajamento em revisão e auto-indagação para ver se o objetivo está sendo atingido	Explicitação dos objetivos da leitura
Seleção de conceitos para indexação	Seleção de conceitos para indexação
Associação com Linguagem	

Quadro 5: Análise das operações (atividades) desenvolvidas pelo sujeito 5 (indexadores do centro de Documentação da editora Abril) durante a leitura dos textos:

Sujeito 5 – Textos 9 e 10	
Texto 9	Texto 10
Identificação de aspectos importantes da mensagem	Identificação de aspectos importantes da mensagem
Alocamento de atenção a áreas importantes (seleção de itens para preencher ficha de indexação)	Alocamento de atenção a áreas importantes ((seleção de itens para preencher ficha de indexação)
Identificação de aspectos importantes da mensagem	Identificação de aspectos importantes da mensagem
Explicitação dos objetivos da leitura	Associação com Linguagem
Seleção de conceitos para indexação	Seleção de conceitos para indexação
Engajamento em revisão e auto-indagação para ver se o objetivo está sendo atingido	Engajamento em revisão e auto-indagação para ver se o objetivo está sendo atingido
Tomadas de ações corretivas quando são detectadas falhas na compreensão	Seleção de conceitos para indexação
Recobrimento de atenção Quando a mente se distrai ou faz digressões	Engajamento em revisão e auto-indagação para ver se o objetivo está sendo atingido
Seleção de conceitos para indexação	Associação com Linguagem
Associação com Linguagem	

Analizando os quadros de sequência das operações notamos que:

- Na leitura, os sujeitos utilizam algumas das estratégias metacognitivas (conscientes) listadas por BROWN (1980);
- Na leitura realizada pelo sujeito 1, dos textos 1 e 2, foram identificadas duas estratégias metacognitivas listadas por Brown como: Identificação de aspectos importantes da mensagem/ Alocamento de atenção a áreas importantes;
- O sujeito 2, na leitura do texto 3, utiliza as mesmas estratégias metacognitivas utilizadas pelo sujeito 1 nos textos 1 e 2. Essas estratégias (Identificação de aspectos importantes da mensagem/Alocamento de atenção a áreas importantes) possibilitam que os sujeitos, durante, a leitura identifiquem e selecionem o que consideram importante nos textos, ou

seja, separem informações principais das periféricas para poder efetuar as extrações de conceitos;

- deve-se ainda destacar que o sujeito 2, na leitura do texto 4, com base somente na leitura do título do texto, usa seu conhecimento prévio e se recorda que já indexou o assunto do texto que irá ler; então, parte para a opção de utilizar os mesmos conceitos utilizados na indexação feita anteriormente e se recorda também que os termos que foram extraídos na indexação anterior constam da Linguagem documentária do jornal (o que constatamos como Associação com Linguagem). Mas, ao prosseguir na leitura do texto, verifica que cometeu um equívoco (realiza monitoração do comportamento para ver se está ocorrendo compreensão), pois não se tratava do texto que conhecia e que já havia indexado. Com isso concluímos, de acordo com a Norma ISO 5963, que somente a leitura do título do texto não é suficiente para a indexação do texto;
- Evidencia-se também que o sujeito 2, na leitura do texto 4, realiza mais estratégias de monitoração do comportamento enquanto que, no texto 1, faz uma leitura mais fluida, sem obstáculos;
- Com relação à monitoração do comportamento, o sujeito 2, durante a leitura do texto 4, identificou maneiras diferentes de leitura, referindo-se à prática de leitura dinâmica do texto antes de identificar termos em partes, que reconhecia serem propícias. Podemos inferir que o sujeito 2 estava preocupado em não fazer o que Farrow (1991) denomina de “rastreamento” rápido do texto, porque o indexador volta ao início e explora cada parte da estrutura textual certificando-se, assim, que os termos identificados são representativos do conteúdo e, por isso, selecionados;
- Tanto o sujeito 1 como o sujeito 2 (indexadores do arquivo do jornal “O Estado de São Paulo”), durante a leitura da maioria dos textos selecionados, não explicitou os objetivos para os quais estava realizando a leitura, ressaltando-se que o sujeito 1, somente após o término da seleção de conceitos, explicita o objetivo da leitura percebido no final da transcrição do texto 1.
- Foi identificado que o sujeito 3, na leitura do texto 5, utilizou duas das estratégias metacognitivas listadas por Brown como: Identificação de aspectos importantes da mensagem/ Alocação de atenção a áreas importantes e Engajamento em revisão e auto-indagação para ver se o objetivo está sendo atingido;

- Na leitura do texto 6, o sujeito 3, do centro de documentação, necessitou utilizar um número maior de estratégias metacognitivas, o que demonstra que encontrou algum obstáculo, ou dificuldade para atingir a compreensão do texto. Sendo as estratégias utilizadas: identificação de aspectos importantes da mensagem, alocamento de atenção a áreas importantes, engajamento em revisão e auto-indagação para ver se o objetivo está sendo atingido, tomadas de ações corretivas quando são detectadas falhas na compreensão, recobrimento de atenção quando a mente se distrai ou faz digressões e explicitação dos objetivos da leitura;
- Destaca-se que o sujeito 4, na indexação dos textos 7 e 8, utiliza estratégias metacognitivas sendo essas: identificação de aspectos importantes da mensagem, alocamento de atenção a áreas importantes, explicitação dos objetivos da leitura e engajamento em revisão e auto-indagação para ver se o objetivo está sendo atingido;
- Durante a indexação dos textos 9 e 10, realizada pelo sujeito 5, percebemos que no decorrer da leitura, o indexador utilizou estratégias metacognitivas (conscientes) listadas por Brown (1980) como: identificação de aspectos importantes da mensagem, alocamento de atenção a áreas importantes, engajamento em revisão e auto-indagação para ver se o objetivo está sendo atingido, tomadas de ações corretivas quando são detectadas falhas na compreensão, explicitação dos objetivos da leitura e Recobrimento de atenção quando a mente se distrai ou faz digressões;
- Quando o sujeito 1, pertencente ao arquivo do jornal “O Estado de São Paulo”, comenta o objetivo da leitura no final da entrevista sendo esse a representação de conceitos visando a recuperação do texto pelos usuários, identificamos que a concepção de análise de assunto no arquivo visa o conteúdo e os interesses informacionais dos usuários;
- E os sujeitos 3, 4 e 5, indexadores do Centro de Documentação da Editora Abril, expressam o objetivo da sua leitura, destacando que indexam visando a representação de conceitos que possibilitem a recuperação do texto; também percebemos que no DEDOC se indexa explorando o conteúdo e o interesse dos usuários pelo conteúdo atribuído ao documento.
- Destacamos que tanto o arquivo como o DEDOC possuem uma concepção de indexação que visa representar o conteúdo do documento, como também a recuperação deste conteúdo pelos usuários, o que Albrecthsen (1993) classifica de concepção e análise de

assunto orientada para o conteúdo e concepção orientada para a demanda e que, segundo Fujita (2000), para os fins documentários é o ideal.

b) Manter em mente o seu objetivo de representar o texto para futura recuperação

Percebemos que os sujeitos 1, 3, 4, e 5 explicitaram manter em mente o objetivo da leitura na realização da tarefa, enquanto que o sujeito 2 não o fez, nem o mencionou. Podemos destacar que a indexação de jornais nos centros especializados (arquivo do jornal “O Estado de São Paulo” e Departamento de Documentação da Editora Abril) está direcionada para a representação fidedigna do conteúdo do texto como também para a qualidade na recuperação.

c) Exploração de seu conhecimento de estrutura textual

O conhecimento e a exploração da estrutura textual pelos indexadores são considerados muito importantes para a leitura documentária, pois proporcionam ao indexador domínio das partes que compõem o texto e do contexto de cada parte, para reconhecimento dos conceitos pertinentes que se pretende representar. A análise dos dados coletados indicou a exploração da estrutura textual seguida pelos 5 sujeitos (indexadores), destacada nos quadros seguintes.

Quadro 6: Seqüência da estrutura textual seguida pelos sujeitos 1, 2, 3, 4 e 5 nos textos:

Sujeito 1	
TEXTO 1	TEXTO 2
Cabeçalho Lead Sujeito 2	Cabeçalho Lead
Texto 3	Texto 4
Cabeçalho Lead Episódio (Eventos) Sujeito 3	Cabeçalho Lead
TEXTO 5	Texto 6
Cabeçalho Lead Sujeito 4	Cabeçalho Lead Episódio (Eventos)
Texto 6	Texto 7
Cabeçalho Lead Sujeito 5	Cabeçalho Lead
Texto 9	Texto 10
Cabeçalho Lead Episódio (s) Atos da fala, Quando a personagem indaga algo	Cabeçalho Lead

Os quadros de “Seqüência da estrutura textual seguida pelos sujeitos” demonstram que os sujeitos exploraram durante a leitura dos textos mais os itens “Cabeçalho” e “Lead”, sendo que estes itens da estrutura textual de jornais, na maioria dos textos, são significativos para a identificação de conceitos. Entretanto, percebe-se que o sujeito 2 (durante a leitura do texto 3) e o sujeito 3 (na leitura do texto 6) tiveram de estender a

leitura até o item “Episódios (Eventos)”, e o sujeito 5 (no texto 9) identificou conceitos pertinentes ao texto no item “Episódio (s) (Atos da fala)”. Isto nos leva a destacar que, em alguns textos, somente a leitura do “Cabeçalho” e “Lead” não é suficiente. E que na elaboração do modelo de leitura para a indexação, objetivo desta pesquisa, este é um fator que deve ser considerado.

Os dados, evidenciando a estrutura textual seguida pelos sujeitos, nos permitem, ainda, observar, com a análise dos quadros-resumos acima que:

- o sujeito 1 explorou apenas dois itens da estrutura textual do texto jornalístico (“Cabeçalho e Lead”) no texto 1 e utiliza a mesma sequência no texto 2;
- destacando a sequência da estrutura textual seguida pelo sujeito 2 no texto 1, podemos observar que ele precisou avançar um pouco mais na sequência de estrutura textual do texto jornalístico na busca por informações, chegando até ao item denominado Episódio. Portanto, podemos ressaltar que a sequência da estrutura textual seguida depende da complexidade do texto analisado, do conhecimento que o indexador possui sobre as diversas formas de narrativa e, das informações que interessam para a comunidade usuária;
- No texto 3, o sujeito 2 utiliza a mesma sequência usada pelo sujeito 1 (nos textos 1 e 2). Com isso, podemos ressaltar que a sequência de estrutura predominante nesses textos é: “Cabeçalho e Lead”. Portanto, as informações se concentram mais nesses itens, que se localizam no começo do texto, mais exatamente no (s) primeiro (s) parágrafo (s), o que torna a leitura, de certa forma, mais rápida quando as informações selecionadas nesses itens são pertinentes para a indexação;
- a análise do quadro-resumo enfatizando a sequência da estrutura textual seguida pelo sujeito 3 (no texto 5) demonstra que o indexador explorou apenas dois itens da estrutura textual do texto jornalístico, “Cabeçalho e Lead”;
- já na leitura do texto 6, o sujeito 3 precisou avançar um pouco mais na sequência da estrutura textual do texto jornalístico na busca por informações pertinentes para representar os conceitos, chegando até ao item denominado “Episódio (Eventos)”.

- O sujeito 4 também explorou (nos textos 7 e 8) os itens da estrutura textual “Cabeçalho e Lead” durante a indexação, sendo estes relevantes na extração dos assuntos que representavam o conteúdo dos textos;
- Durante a indexação do texto 9, o sujeito 5 explorou também os itens “Cabeçalho e Lead”, mas não leu o “Lead” até o fim, deixando de encontrar a informação que tanto desejava referente aos nomes das outras personagens envolvidas no filme, como também a informação de que se tratava da indicação do filme ao Oscar, o que veio a identificar no decorrer da leitura no item “Episódio (s) (atos da fala)”, e a exploração desse item lhe possibilitou que compreendesse que se tratava de uma indicação, e reparou a falha na compreensão que ocorreu por não ter lido todo o “Lead”;

É relevante destacar que, com os resultados obtidos na análise da exploração da estrutura textual, identificaram-se indícios de que os indexadores têm conhecimento da importância da exploração da estrutura textual durante a leitura, e que a utilizam como estratégia, pois sabem o momento de interromper a leitura afirmando que a continuidade do texto não lhe interessa (Sujeito 1 -- texto 2) e sabem também quando é preciso ler a continuidade do texto (Sujeito 2 – texto 4).

Ressaltamos ainda que as partes mais exploradas da estrutura textual foram o “Cabeçalho” e o “Lead” e, considerando que esses itens localizam-se geralmente no (s) primeiro (s) parágrafo (s) do texto jornalístico, supõe-se que isso seja um fator determinante para que ocorra uma leitura mais rápida. Ressalta-se que, em alguns textos, chegou-se até os itens “Episódios (Eventos)” e “Episódios (Atos da fala)” da estrutura textual de jornal proposta por Van Dijk (1983).

É relevante mencionar que o “Lead”, item mais explorado da estrutura textual e significativo em conteúdo para a identificação de conceitos durante a indexação, localiza-se, quase sempre, no primeiro parágrafo do texto jornalístico. E quando os indexadores necessitaram estender a leitura além do Lead, chegaram até o item Episódio (atos da fala), sendo provavelmente a leitura até este item suficiente, porque além dele as informações abordadas relatam as crenças do jornalista ou do jornal. Nos jornais utilizados “O Estado de São Paulo”, “Folha de São Paulo” e o “O Globo”, a leitura não ultrapassou o quinto parágrafo do texto, fator que, acreditamos, contribuirá na elaboração do modelo de leitura para indexação de jornais.

d) identificação de aspectos importantes da mensagem para identificação de conceitos

A identificação e seleção de conceitos é, tal como a exploração da estrutura textual, uma importante estratégia de leitura. Em nosso entendimento, é a estratégia vital para atingir o objetivo de representação para recuperação, pois é com esta estratégia que os conceitos são identificados.

Com essa perspectiva em mente, três aspectos da estratégia de identificação serão abordados e verificados na análise dos dados da entrevista:

d1) em que parte da estrutura textual foram identificados os conceitos;

d2) qual a sistemática utilizada para a identificação de conceitos e

d3) identificação e seleção de termos

Para uma melhor visualização dos resultados abordando o procedimento de identificação de conceitos e partes da estrutura textual, foram construídos os quadros-resumos evidenciados abaixo:

d1) Parte da Estrutura Textual em que foram Identificados os Conceitos

Quadro7: Identificação de conceitos e Partes do texto - Sujeito 1

Sujeito 1 - Texto 1	
Trechos em que o sujeito identificou os termos	Parte do Texto
<i>“ Lobão volta ao rock da melhor qualidade então a primeira página do caderno tá hoje tá tratando do Lobão é o cantor”</i> (Pesquisador observa que, nesse momento, o sujeito argola o título para preencher ficha de indexação com dados descritivos da matéria)	Cabeçalho (título)
<i>... “então vai tratar do Lobão, Lobão a gente entra em Lobão entre parêntese cantor” ...</i>	Lead
Sujeito 1 – Texto 2	
Trechos em que o sujeito identificou os termos	Parte do Texto
Longe das passarelas iluminadas a moda é um assunto que desperta o interesse de profissionais de áreas diversas. O que atrai essas pessoas que talvez não dominem as técnicas de modelagem nem sejam, hábeis costureiras, mas com certeza podem explicar por que a silhueta feminina muda tanto ao longo dos anos – é o lado documental da moda. Sujeito: <i>bom eu percebo desse primeiro parágrafo que o assunto abordado é moda então eu já vou indexar em moda”</i> .	Lead

Quadro 8: Identificação de conceitos e partes do texto - SUJEITO 2

Sujeito 2 – Texto 3	
Trechos em que o sujeito identificou os termos	Parte do Texto
... “o crime ocorreu na Quinta-feira. A quadrilha era composta por quatro bandidos .. Sujeito: <i>pelo que eu tô entendendo aqui a policia que matou essa quadrilha certo, não foi briga de quadrilhas então a policia aqui no caso é importante tá, eu ficho dois descritores criminalidade Maranhão e a Polícia Maranhão, tá.</i>	Episódio (Eventos)
Sujeito 2 – Texto 4	
Trechos em que o sujeito identificou os termos	Parte do Texto
... “a outra matéria Assaltante do BB de Campinas é preso. (Pesquisador: nesse momento o sujeito argolou o título para preencher ficha de indexação com dados descritivos da matéria) ... BB é o Banco do Brasil tá isso aqui também é um caso que aconteceu faz tempo e a gente tem o título Banco Assalto, tá. (Pesquisador: nesse momento o sujeito extrai esse Termo).	Cabeçalho (título)

Quadro 9: Identificação de conceitos e Partes do texto – Sujeito 3

Sujeito 3 – Texto 5	
Trechos em que o sujeito identificou os Termos	Parte do Texto
<i>... A greve dos policiais de fato teve um conflito é mais o conflito é dentro da greve então eu vou colocar greve policia vou colocar como Termo policia militar ...</i>	Lead
<i>... É protesto porque eles falam do protesto aqui e o lugar onde está sendo feita a greve que é Recife é só recife então Recife-Pernambuco já cobri está tá coberta.</i>	Lead
Sujeito 3 – Texto 6	
Trechos em que o sujeito identificou os Termos	Parte do Texto
<i>É eleição municipal São Gonçalo é onde? no Rio, então é a eleição municipal Rio de Janeiro então vai pra esta pasta como é o Garotinho uma figura importante eu vou por como termo o nome dele Garotinho e vou por nome da cidade que no nome da pasta não vai constar só consta Rio de Janeiro-Eleição Municipal então eu vou por São Gonçalo barra Rio de Janeiro São Gonçalo/Rio de Janeiro.</i>	Lead
<i>Ele vai inaugurar obras talvez seja alguma coisa pro governo do estado. Também está prevista para esta semana uma caminhada do governador com o candidato pelos principais bairros da cidade, segundo assessores de Ezequiel. Mas eu tô vendo que são coisas irrelevantes então eu não vou por nome do Garotinho então vai ficar vou por PDT porque o candidato é do PDT então vou por PDT barra partido PDT/Partido</i>	Episódio (s)

Quadro 10: Identificação de conceitos e partes do texto - Sujeito 4

Sujeito 4 – Texto 7	
Trechos em que o sujeito identificou os termos	Parte do Texto
<i>Então eles estão tentando tirar a concessão do bondinho e passar pra uma outra companhia e assim com o Corcovado também o Cristo Redentor também então aqui eu pego coisas normais eu ponho bonde ele vai como uma referência</i>	Lead
<i>Se alguém quiser saber alguma coisa sobre o que aconteceu a época do Luis Conde então eu vou colocar e eles recuperam a matéria a troca da gestão do Pão de Açúcar fala também eles querendo trocar as antigas concessões</i>	Lead
<i>Título então eu coloco o Termo referente a gestão em que está sendo trocada sendo o Termo CONDE, Luis Paulo e Cristo Redentor já está dentro da matéria e Monumento</i>	Lead
Sujeito 4 – Texto 8	
Trechos em que o sujeito identificou os Termos	Parte do Texto
<i>(...) E você 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 nomes do agressor e da vítima</i>	Lead
<i>Aberta a pasta eu incluo os assuntos: agressor e vítima e não coloco agressão como assunto porque o nome da pasta já informa que houve uma agressão contra mulher.</i>	Lead

Quadro 11: Identificação de conceitos e partes do texto - Sujeito 5

Sujeito 5 – Texto 9	
Trechos em que o sujeito identificou os Termos	Parte do Texto
<i>(...) pode colocar atores que são casos de buscas como a Regina Casé como Termo porque as vezes as pessoas podem lembrar dela e não lembrar o nome do filme aquele filme que ela fez e tal, (nesse momento o indexador inclui o nome da atriz como assunto do filme) o outro é o Lima Duarte</i>	Lead
<i>Vou Ter que incluir termos A indicação do Oscar há 40 não 12 matérias sobre o filme desde seu lançamento mas o que marca essa matéria é a indicação pro Oscar. Porque poderia ser uma crítica, lançamento mas o que está marcando aqui essa matéria é a indicação então para facilitar a busca eu vou colocar termos agora como prêmio Oscar</i>	Episódio (s) (Atos da fala)
<i>Tá, agora o outro termo que eu vou colocar é a indicação porque não é a premiação</i>	Episódio (s) (Atos da fala)
Sujeito 5 – Texto 10	
Trechos em que o sujeito identificou os termos	Parte do Texto
<i>... O que chama atenção nessa matéria é que ele vai vim no Rock in Rio 3 é uma entrevista dele é sobre um show então coloco como termos entrevista e show ...</i>	Lead

Mediante a análise dos quadros referentes à identificação de conceitos e partes da estrutura textual do texto jornalístico, verificou-se que:

- o sujeito 1, durante a indexação dos textos 1 e 2, inicia a leitura pelo item “Cabeçalho” e utiliza-o em dois momentos e, na sequência, conclui a leitura fixando-se no primeiro parágrafo que é considerado o “Lead”;
- Nota-se que o sujeito 2 inicia a leitura também com a visualização das informações do item “Cabeçalho” e, na sequência da leitura, encontra informações que lhe interessam somente no item “Episódio (Eventos)”, desconsiderando, na indexação deste texto, o “Lead”, por não apresentar conceitos pertinentes ao conteúdo do texto;
- Já na leitura do texto 4, o sujeito 2 explora o item “Cabeçalho”- título” em um único momento e identifica dois conceitos. Na sequência da leitura, confirma um conceito

identificado e exclui outro conceito por notar com a leitura do “Lead” que um dos conceitos não representa o conteúdo da matéria e conclui a atividade de leitura;

- Observa-se, novamente, a utilização dos itens (“Cabeçalho e Lead”) pois o sujeito 3 (no texto 5) também inicia a leitura pelo item “Cabeçalho” e, na sequência, conclui a leitura fixando-se no primeiro parágrafo que é denominado de “Lead”;

- Entretanto, percebemos que o sujeito 3 no texto 6 inicia a leitura no “Cabeçalho” com a extração dos itens descritivos que preenchem a ficha de indexação e prossegue a leitura em busca de conceitos explorando o item “Lead”. Na sequência da leitura, chegou até o item Episódio (s) (Eventos) e, nesse item, encontra informações que lhe permitem corrigir a falha e excluir um termo que não era o foco principal do texto;

- Destacamos que o sujeito 4 iniciou a indexação dos textos 7 e 8 com a seleção dos componentes necessários para preencher a ficha de indexação e, durante a leitura, explorou também o “Lead” na indexação do texto, encontrando, nesse item da estrutura textual, os conceitos relevantes para se representar o conteúdo dos textos;

- Durante a leitura dos textos 9 e 10 o sujeito 5 identificou conceitos no item “Lead”. E na indexação do texto 10, além de explorar o “Lead”, o sujeito estendeu a leitura até o item “Episódio (s) (Atos da fala)” da estrutura textual, e identificou um conceito que lhe permitiu representar mais especificadamente o conteúdo do texto;

- Destacamos que os sujeitos concentraram a identificação de conceitos durante a leitura da maioria dos textos no item “Lead” sendo este significativo em conteúdo. Percebe-se que, em alguns textos, a leitura estendeu-se até o item “Episódio (Eventos)” e “Episódio (Atos da fala)”, o que nos leva a concluir que na elaboração do modelo de leitura proposto, devemos considerar, provavelmente, a exploração da estrutura textual de jornais para fins de indexação até estes itens, percebidos, até o momento, como limite na utilização da estrutura do texto na leitura dos sujeitos.

d2) Sistemática para identificação de conceitos

A sistemática de identificação foi observada utilizando as transcrições dos textos e a análise dos quadros, (7,8,9,10). Observou-se que os sujeitos evidenciaram identificação de conceitos utilizando seis tipos de comportamentos: destaque de palavras

do texto por meio de argolamento, ou digitação; associações com linguagem após reconhecimento de termos; questionamento durante a leitura quanto à escolha dos termos; conhecimento da estrutura textual; identificação de maneiras diferentes de ler e conhecimento prévio, e recorre à memória a longo termo.

- **destaque de palavras do texto** (“argolamento”, ou digitação). Essa atividade não foi expressa verbalmente, a pesquisadora observou e anotou.

Como exemplo, temos a seguinte transcrição:

Bailarina acusa o marido ator de espancamento e dá queixa na polícia. Bailarina acusa o 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). (Sujeito 4 – Texto 8)

- **fazendo associações com Linguagem após reconhecimento de termos:**

Exemplos:

Sujeito 2 – texto 3: “ BB é o Banco do Brasil tá isso aqui também é um caso que aconteceu faz tempo e a gente (Pesquisador: Linguagem Possui) tem o título Banco Assalto, tá...”

Sujeito 4 - texto 7: ... e quando o título define mais ou menos o que está na matéria não tem necessidade de colocar muitos termos e também porque nosso sistema trabalha por palavra se você colocar O Globo Jacarépagua todas as matérias que trouxeram a palavra Jacarépagua do jornal O Globo ele vai te jogar na tela.

Sujeito 5 – texto 9: (Você coloca somente indicação não fica muito geral?) Não porque na busca você pode cruzar os termos porque, se você colocar somente indicação vai recuperar indicação até de efeito piolho, mas sempre você vai acabar cruzando, porque você vai falar há os filmes brasileiros que foram indicados ao Oscar digita indicação Oscar vem esse daqui e vai vim aqueles anteriores, há mais é aquele do Andrucha daí você coloca Andrucha né aí Andrucha, indicação já mata, tá.

Sujeito 5 – texto 10 realiza associação com linguagem: Aqui, por exemplo, nos campos termos, é feito o controle de vocabulário dos dois (termos(assuntos) e nomes das pastas), então ao invés d'eu colocar o Beck eu vou para o vocabulário que vem todos os Beck que tem ...

- Realiza questionamento durante a Leitura do texto quanto às escolhas dos descritores:

Exemplos:

Sujeito 2 – texto 3: ...*"eu preciso saber se vou colocar na Policia Maranhão que é um delegado ou só no Crime..."*

Sujeito 3 – texto 5 realiza essa atividade como: ... *A greve dos policiais, de fato, teve um conflito, é mais o conflito, é dentro da greve, então, eu vou colocar greve policia, vou colocar como termo policia militar*

Sujeito 3 – Texto 6 realiza essa atividade em: Mas eu tô vendo que são coisas irrelevantes, então eu não vou por nome do Garotinho, então vai ficar, vou por PDT porque o candidato é do PDT, então vou por PDT barra partido PDT/Partido

Sujeito 4 – texto 7 realizou a atividade como exemplo: *se alguém quiser saber alguma coisa sobre o que aconteceu à época do Luis Conde então eu vou colocar, e eles recuperam a matéria*

Sujeito 4 – texto 8 durante a leitura se questionou como exemplo: 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

Sujeito 5 – texto realiza questionamento como exemplo: ...PS.. vou ter que incluir termos A indicação do Oscar há 40, não 12 matérias sobre o filme desde seu lançamento, mas o que marca essa matéria é a indicação pro Oscar. Porque poderia ser uma crítica, lançamento mas o que está marcando aqui essa matéria é a indicação, então, para facilitar a busca, eu vou colocar termos agora como prêmio Oscar

Sujeito 5 – texto 10 se questionou durante a leitura quanto à escolha dos termos como exemplo: O que chama atenção nessa matéria é que ele vai vim no Rock in Rio 3 é uma entrevista dele é sobre um show, então coloco, como termos, entrevista e show é genérico

(nesse momento o indexador realiza uma “olhadela” pelo texto procurando encontrar um termo mais específico que represente a matéria) *mas é só esses mesmos*

- **Identifica que há maneiras diferentes de ler**

Sujeito 2 – texto 3 somente com a leitura do título identificou conceito que não representava o conteúdo do texto e percebe ser um erro indexar considerando somente o título, e admite ser necessário ler partes do texto, rapidamente, em busca da compreensão para a identificação de conceitos, como exemplo, destacamos: “... *Oh, não é assalto que eu estava pensando, não, foi outro assalto, foi no início da semana, eu achei que era outro assalto, tá vendo como que é, a gente, então a gente sempre tem que fazer a leitura dinâmica...*”(Sujeito 2 - Texto 3)

- **Conhecimento de estrutura textual**

Exemplos:

Sujeito 1 – texto 1 demonstra conhecer a estrutura textual do texto em dois momentos, como exemplo: “... *Olha eu não preciso mais ler que eu sei quando o resto da matéria vai tá falando dele....*”

Sujeito 5 – texto 9 evidencia conhecer a estrutura textual do caderno que indexa como exemplo: *Essa matéria está no caderno cotidiano mas retranca para a ilustrada. Então, quando eu for indexar a ilustrada, essa matéria é a primeira que eu vou pegar*

Sujeito 4 – texto 8 demonstra conhecer o elemento-título da estrutura textual que define o assunto que será abordado no texto como exemplo: “... *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*

- **Usa conhecimento prévio e recorre à memória de longo prazo**

Exemplos:

Sujeito 2 – texto 4” *BB é o Banco do Brasil tá, isso aqui também é um caso que aconteceu faz tempo e a gente tem o título Banco Assalto tá, no caso aqui a gente fichou também Banco do Brasil tá, porque foi muito importante esse crime tá, ele roubou milhões e milhões lá....*”

Sujeito 4 – texto 7 ... *então estão querendo dar uma nova estrutura a essas coisas que são coisas tradicionais na cidade do Rio de Janeiro, é o Pão de Açúcar, o Corcovado, o bonde de Santa Tereza coisas relativas à cultura da cidade.*

Sujeito 5 – texto 9 utiliza conhecimento prévio e recorre à memória durante a leitura em dois momentos, como exemplo: ... *e tem um outro cara que é legal colocar como ator o terceiro marido, mas nós não temos o nome dele aqui na matéria, então eu não vou colocar.* (Nesse momento o indexador recorre à memória tentando recordar informações que viveu ou por ter assistido ao filme, ou de outros textos já indexados, anteriormente, sobre o filme, e que lhe dão a certeza que existe um outro nome importante a indexar)

Com a demonstração da sistemática de identificação de conceitos acima, verificou-se que:

- Sujeito 1 na leitura do texto 1 faz destaque de palavras do texto (“argolamento”) e usa conhecimento de estrutura textual;
- e na leitura do texto 2 o sujeito 1 novamente faz destaque do texto (“argolamento”);
- Sujeito 2 no texto 3 realiza destaque de palavras do texto (“argolamento”); faz associações com Linguagem após reconhecimento de termos; realiza questionamento durante a leitura do texto quanto às escolhas dos descritores e identifica que há maneiras diferentes de se ler;
- e na leitura do texto 4, o sujeito 2 realiza destaque de palavras do título “Cabeçalho”; usa o seu conhecimento prévio e recorre à memória a longo prazo;

- O sujeito 3, no texto 5, durante a seleção dos termos de indexação, realiza duas vezes monitoração do comportamento para ver se está ocorrendo compreensão, e faz questionamento durante a leitura do texto quanto às escolhas dos descritores em dois momentos;
- O Sujeito 3, lendo o texto 6, durante a seleção dos termos de indexação, realiza também duas vezes monitoração do comportamento para ver se está ocorrendo compreensão; faz questionamento durante a leitura do texto quanto às escolhas dos descritores em dois momentos;
- Sujeito 4 texto 7: durante a seleção dos termos de indexação, realiza três vezes monitoração do comportamento para ver se está ocorrendo compreensão; faz questionamento durante a leitura do texto quanto às escolhas dos descritores em três momentos; faz também associações com Linguagem após reconhecimento de termos e usa conhecimento prévio, e recorre à memória à longo prazo;
- Sujeito 4 no texto 8: durante a seleção dos termos de indexação, realiza duas vezes monitoração do comportamento para ver se está ocorrendo compreensão; faz questionamento durante a leitura do texto quanto às escolhas dos descritores em quatro momentos, e usa conhecimento de estrutura textual;
- Sujeito 5 no texto 9: realiza associações com Linguagem após reconhecimento de termos; se questiona durante a leitura do texto quanto às escolhas dos descritores quatro vezes; durante a seleção dos termos de indexação, realiza monitoração do comportamento, três vezes, para ver se está ocorrendo compreensão; usa conhecimento prévio e recorre a memória à longo prazo em dois momentos, e utiliza conhecimento de estrutura textual;
- Sujeito 5 no texto 10: faz, duas vezes, questionamento durante a leitura do texto quanto às escolhas dos descritores; durante a seleção dos termos de indexação realiza uma única monitoração do comportamento para ver se está ocorrendo compreensão, e faz associações com Linguagem após reconhecimento de termos, em dois momentos.

Constatamos que os sujeitos realizam a identificação de conceitos utilizando variadas estratégias de leitura, conscientemente, e obedecendo, provavelmente, a uma sistemática constante. Pode-se observar que a indexação inicia-se com o preenchimento da ficha de indexação onde são selecionados título da matéria, jornal, caderno, seção, data, página e autor, explorando nesse momento o item “Cabeçalho” da estrutura textual de jornal, em busca desses elementos descritivos do texto. É importante destacar que o sujeito

2 (no texto 4) identifica o conceito *BANCO/ASSALTO* com a leitura do título e seleciona este conceito; entretanto, os títulos, muitas vezes, podem ser enganosos. Posteriormente, após ler o título, o indexador de jornal inicia a leitura do texto propriamente dito (explorando o “Lead”, “Episódios (Atos da fala)” e, nesse segundo momento, o sujeito recorre à memória, e utiliza seu conhecimento prévio sobre o assunto, na busca pela compreensão que é representada pela seleção de termos pertinentes ao texto. O sujeito confirma a pertinência do termo, se questionando sobre a escolha desse termo para representar o assunto da matéria e os interesses informacionais da comunidade usuária, estratégia de leitura que é definida como um dos aspectos de monitoração do comportamento para ver se está ocorrendo compreensão. Na etapa seguinte, confirmada a pertinência do termo, o sujeito verifica se a Linguagem de indexação possui este termo, ou se oferece preferência para outro termo-sinônimo, e caso a Linguagem não possua o termo, o indexador pode incluir o termo novo na Linguagem e finalizar a indexação. Entretanto, no DEDOC, a indexação não termina com a seleção dos termos porque posteriormente à atividade de indexação, um funcionário da instituição, responsável por estabelecer o controle da Linguagem, verifica os termos atribuídos aos textos e, se julgar necessário, exclui os termos considerados como não representantes do conteúdo do texto ou que podem ser recuperados pelo título, nome da pasta, ou por qualquer palavra do texto. Com isso, efetua uma alteração na indexação e esta é registrada no sistema, evidenciando o dia em que foi feita a alteração. Destacamos que as fichas de indexação dos textos selecionados, para observação da leitura dos indexadores do DEDOC, se encontram no anexo 8, para demonstrar a ocorrência ou não de tal alteração dos termos indexados.

d3) identificação e seleção de termos

Apresentamos, abaixo, os quadros referentes aos termos identificados e selecionados esclarecendo que, durante a leitura, o indexador realiza em um único momento, primeiramente, a identificação, e depois a seleção dos termos, sendo esta última confrontada com a representação do conteúdo do texto e os interesses informacionais dos usuários. Como exemplo, destacamos um trecho de uma transcrição:

O que chama atenção nessa matéria é que ele vai vim no Rock in Rio 3 é uma entrevista dele é sobre um show então coloco como termos entrevista e show é genérico (nesse momento o indexador realiza uma “olhadela” pelo texto procurando encontrar um termo mais específico que represente a matéria) mas é só esses mesmos. (pesquisadora: e

no caso você não entra pelo nome dele como assunto porque pode ser recuperado pelo nome da pasta?). Pode se você der Beck no campo de pesquisa já vem essa matéria, e ainda você pode recuperar por pasta, palavra-chave todas as palavras do texto, por publicação (jornal). Tem que pensar como o pesquisador vai procurar aquilo? Como facilitar a procura para o pesquisador porque para você tá com a matéria na mão tá fácil, tá vendo a matéria tal mas você pensa daqui um mês como você procuraria essa matéria? Então tem que se procurar facilitar com a atribuição de termos para o cara para facilitar a recuperação dele. Então o lance é esse. Porque para você tá fácil tá claro mas é aquela coisa do telefone sem fio, onde o último lá pode estar pensando outra coisa. Então você tem que imaginar como o pesquisador vai pesquisar determinada matéria e procurar facilitar a busca. Sujeito 5 – texto 10

Quadro 12: Termos identificados e selecionados - Sujeitos 1 e 2 nos textos:

Sujeito 1 – Texto 1	
Termos Identificados	Termos Selecionados
Lobão (cantor)	Lobão (cantor)
Sujeito 1- Texto 2	
Termos Identificados	Termos Selecionados
Moda	Moda
Joel Birman	Joel Birman
Cristina Franco	Cristina Franco
Sujeito 2 – Texto 3	
Termos Identificados	Termos Selecionados
Polícia Maranhão	Polícia Maranhão
Criminalidade Maranhão	Criminalidade Maranhão
Sujeito 2 – Texto 4	
Termos Identificados	Termos Selecionados
Banco Assalto	Banco Assalto
Banco do Brasil	

Quadro 13: Termos identificados e selecionados – Sujeitos 3 e 4 nos textos:

Sujeito 3 – Texto 5	
Termos Identificados	Termos Selecionados
Greve Policia Polícia Militar Recife-Pernambuco	Greve Polícia Polícia Militar Recife-Pernambuco
Sujeito 3 – Texto 6	
Termos Identificados	Termos Selecionados
Garotinho São Gonçalo/Rio de Janeiro PDT/Partido	São Gonçalo/Rio de Janeiro PDT/Partido
Sujeito 4 – Texto 7	
Termos Identificados	Termos Selecionados
Bonde Luis Conde CONDE, Luis Paulo Monumento	Bonde CONDE, Luis Paulo Monumento
Sujeito 4 – Texto 8	
Termos Identificados	Termos Selecionados
BESSA, Luciana ZANYOR, Danilo Agressor Vítima	BESSA, Luciana ZANYOR, Danilo Agressor Vítima

Quadro 14: Termos identificados e selecionados – Sujeito 5 nos textos:

Sujeito 5 – Texto 9	
Termos Identificados	Termos Selecionados
Regina Casé	CASÉ, Regina
Lima Duarte	LIMA DUARTE/ATOR
Prêmio Oscar	Prêmio Oscar
Indicação	Indicação
Sujeito 5 – Texto 10	
Termos Identificados	Termos Selecionados
Entrevista	Entrevista
Show	Show

A análise dos quadros referentes aos termos identificados e selecionados nos possibilita destacar que:

- sujeito 1: tanto no texto 1 como no texto 2, seleciona todos os termos que foram identificados, mostrando a eficácia da leitura, uma vez que o produto dessa atividade é a seleção de conceitos;
- sujeito 2: no texto 3, também selecionou todos os termos que identificou, o que demonstra a eficácia da leitura realizada pelo indexador.
- destaca-se ainda que o sujeito 2, no texto 4, identificou dois termos somente com a leitura do título e, na sequência da leitura, selecionou apenas dois termos, confirmando assim, que a busca da compreensão, ou da extração de termos, com base somente no título, não é suficiente. (Norma ISO 5963);
- novamente percebemos a eficácia da leitura desses indexadores de jornais pois o sujeito 3, no texto 5, seleciona todos os termos que foram identificados;
- já na leitura do texto 6, o sujeito 3 identificou o termo Garotinho, realizando uma previsão baseando-se nos seus conhecimentos prévios que diziam que este seria um dos assuntos do texto e, com a sequência da leitura, esse termo não foi selecionado por se constatar que não abordava o tema principal do texto;

- na leitura do texto 7, o sujeito 4 selecionou todos os termos identificados e realizou uma mudança na entrada para um dos termos;
- percebemos que o sujeito 4, no texto 8, selecionou todos os termos que identificou somente com a leitura do “Lead”;
- observa-se sujeito 5, na indexação do texto 9, selecionou todos os termos que indexou; entretanto, as transcrições nos deixam destacar que o indexador almejava incluir outros termos que considerava importante no conteúdo da matéria, como os nomes das outras personagens do filme, mas como não leu o “LEAD” até o fim, não obteve a informação que desejava e, tentando apelar para conhecimento anterior (conhecimento prévio) sobre o assunto, também não recuperou a informação por não se lembrar dos nomes;
- já na indexação do texto 10, o sujeito 5 selecionou todos os termos que identificou somente com a leitura do “LEAD”.

e) Compatibilidade dos termos selecionados com a Linguagem do Sistema

A análise da compatibilidade dos termos com a Linguagem de indexação, considerando os quadros-resumo acima, foi realizada procurando evidenciar se a Linguagem possuía os termos selecionados pelos sujeitos durante a leitura.

Para melhor analisar a compatibilidade dos termos com as Linguagens, foram elaborados os quadros abaixo:

Quadro 15: Termos identificados e selecionados pelos sujeitos 1, 2 e 3 nos textos comparados com a Linguagem do Jornal

Sujeito 1 – Texto 1	
Termos Identificados	Manual do Jornal

Lobão (cantor)

Não tem

Sujeito 1 – Texto 2	
Termos Identificados	Manual do Jornal

Moda

Não tem

Joel Birman

Não tem

Cristina Franco

Não tem

Sujeito 2 – Texto 3	
Termos Identificados	Manual do Jornal

Polícia Maranhão

Tem

Criminalidade Maranhão

Tem

Sujeito 2 – Texto 4	
Termos Identificados	Manual do Jornal

Banco Assalto

Tem

Banco do Brasil

Tem

Sujeito 3 – Texto 5	
Termos Identificados	Linguagem DEDOC

Greve Polícia

Tem

Polícia Militar

Tem

Recife-Pernambuco

Tem

Sujeito 3 – Texto 6	
Termos Identificados	Linguagem DEDOC

Garotinho

Tem

São Gonçalo/Rio de Janeiro

Não tem

PDT/Partido

Tem

Quadro 16: Termos identificados e selecionados pelos sujeitos 4 e 5 nos textos comparados com a linguagem do Jornal

Quadro 4 – Texto 7	
Termos Identificados	Linguagem DEDOC
Bonde	Tem
Luis Conde	Não tem
CONDE, Luis Paulo	Tem
Monumento	Tem

Quadro 4 – Texto 8	
Termos Identificados	Linguagem DEDOC
BESSA, Luciana	Tem
ZANYOR, Danilo	Tem
Agressor	Tem
Vítima	Tem

Quadro 5 – Texto 9	
Termos Identificados	Linguagem DEDOC
Regina Casé	Não Tem
CASE, Regina	Tem
Lima Duarte	Não tem
LIMA DUARTE/ATOR	Tem
Prêmio Oscar	Tem
Indicação	Tem

Quadro 5 – Texto 10	
Termos Identificados	Linguagem DEDOC
Entrevista	Tem
Show	Tem

Com a análise dos quadros de compatibilidade dos termos identificados e selecionados com a linguagem de indexação, verificou-se que:

- os termos selecionados pelo Sujeito 1 (nos textos 1 e 2) não constam na Linguagem do Jornal, destacando, assim, que a linguagem se encontra desatualizada em relação a alguns assuntos. Devemos ressaltar, porém, que os termos (nome do colaborador e o título) devem ser extraídos, pois são termos obrigatórios para preencher a ficha de indexação, segundo conversa informal com os indexadores do jornal;
- ressaltamos que a linguagem do jornal talvez não tenha os termos selecionados para os textos 1 e 2, indexados pelo Sujeito 1, por ser tratar de uma linguagem antiga, e porque a Editoria (Caderno 2), que trata desses assuntos, pode ser mais atual que a linguagem;
- No texto 3, o Sujeito 2 identificou três termos e a Linguagem do Jornal possuía somente um termo;
- destacamos que o sujeito 2, no texto 4, selecionou um termo e a linguagem do Jornal o possuía;
- Com a ausência de alguns dos termos selecionados na Linguagem do Jornal que visava representar a maioria dos textos, podemos destacar que os indexadores possuem muita experiência no assunto tratado e conhecem a comunidade usuária, conseguindo, assim, extrair os termos que, com certeza, irão satisfazer as buscas por informações de sua comunidade, já que a Linguagem não lhe disponibiliza a representação padronizada (orientação de uma linguagem de indexação) de tal conceito;
- Já todos os termos selecionados pelo sujeito 3 (no texto 5) a linguagem do DEDOC possuía;
- E na indexação do texto 6, realizada pelo sujeito 3, todos os termos selecionados constavam da Linguagem, destacando-se que o indexador achou conveniente, também, especificar melhor o local onde estava ocorrendo o fato, e incluiu o termo referente à localização do fato pois a Linguagem possibilita alterações;
- Com relação à seleção de termos para representar o texto 7, o sujeito 4 somente não encontrou na Linguagem um termo, como havia identificado, por se tratar de nomes pessoais e a Linguagem lhe ofereceu a entrada permitida para o nome, sendo sobrenome seguido do nome;
- na indexação do texto 8, o sujeito 4 encontrou na Linguagem todos os termos que selecionou;

- o sujeito 5, na seleção de termos para representar o texto 9, não encontrou na Linguagem termos pessoais, da maneira como se fala, pelo fato de a Linguagem apresentar esses termos de forma diferente, ou seja, sobrenome seguido do nome.
- já na indexação do texto 10, o sujeito 5 encontrou na linguagem todos os termos identificados durante a indexação.

Percebemos, com o resultado da compatibilidade dos termos com as Linguagens dos sistemas, que a linguagem do arquivo está desatualizada quanto aos termos extraídos pelo sujeito 1, não auxiliando totalmente o indexador no momento da representação, ficando a eficácia da tarefa por conta da experiência em indexação (termos selecionados e interesses dos usuários) que o indexador possui. Enquanto que na Linguagem do DEDOC, todos os termos selecionados são passíveis de representação, e podemos observar que a Linguagem é eficiente na representação do conteúdo dos documentos.

Evidenciamos, primeiramente, a comparação entre os procedimentos observados, durante a indexação, dos indexadores do Arquivo do jornal “O Estado de São Paulo” e no Departamento de Documentação da Editora Abril, e os procedimentos recomendados, pela revisão de literatura sobre indexação de jornal, evidenciados no capítulo 3 por Ahmad (1991).

Destacamos que os procedimentos de indexação dos indexadores, na prática, visando identificar conceitos foram:

- destaque de palavras do texto (“argolamento”, ou digitação). Sendo que o Lead foi o componente mais explorado da estrutura textual para a identificação de conceitos;

Ressaltamos que esses procedimentos são também estabelecidos por Ahmad (1991) no item 3, quando menciona o exame do título e do primeiro parágrafo, que é o componente Lead dentro da estrutura textual de artigos de jornais para identificar conceitos.

Percebemos também, que no item 3, a autora, acima citada, anteriormente recomenda que após leitura do título e lead deve-se definir sublinhando ou, escrevendo as palavras significativas que representam os conteúdos do assunto; (Grifo nosso)

- associações com Linguagem após reconhecimento de termos;

Quanto a este critério, Ahmad (1991) se refere, no item 5, em análise das palavras e dos conceitos, e tradução para uma linguagem de indexação.

Os indexadores realizam associação com Linguagem e os indexadores do DEDOC a realizam mais porque no Arquivo somente o sujeito 2, texto 3, deixa claro que a Linguagem possui o termo selecionado, no caso *BANCO/ASSALTO*.

- Realização de questionamento durante a Leitura do texto quanto às escolhas dos descritores;

Percebemos este procedimento durante a leitura dos sujeitos, indexadores do DEDOC, e do sujeito 2, no texto 3, pertencente ao Arquivo.

- Identificação que há maneiras diferentes de ler;

A indexação de jornal, na teoria, menciona uma leitura em que o indexador realiza o correr dos olhos por todo o texto buscando identificar os conceitos indexáveis, interpretar palavras e termos.

Destacamos que o sujeito 2, no texto 4, fez isto porque após identificar conceitos no título, prossegue a leitura pela sequência do texto objetivando identificar, confirmar, e como foi o caso, excluir conceito identificado.

- Uso de conhecimento da estrutura textual;

Ahmad (1991) recomenda também a exploração do conhecimento de estrutura textual à medida que orienta para o exame cuidadoso do título e do primeiro parágrafo.

Os indexadores concentraram a leitura dos textos no Lead que, na maioria das vezes, se fixa no primeiro parágrafo ou entre os cinco primeiros parágrafos, como foi percebido em nossas observações da leitura dos indexadores.

- Uso do conhecimento prévio e apelo à memória de longo prazo

A pesquisadora Ahmad (1991) refere-se à memória e ao conhecimento prévio, mencionando que a análise da notícia exigirá do indexador uma boa memória porque nos jornais as matérias podem ser publicadas, continuamente, durante semanas, conhecimento

sobre o assunto para adotar novas estratégias de leitura, se necessário, e também conhecimento prévio referente à estrutura textual dos textos que necessita indexar.

Os indexadores sujeitos 5, no texto 7, sujeito 1, texto 1 e o sujeito 5, texto 9, demonstraram ter conhecimento anterior sobre o assunto do texto que indexavam.

Mas somente o sujeito 5, no texto 9, deixa claro estar consultando a memória para recuperar a informação que possui sobre o texto.

Percebemos que os procedimentos identificados na atividade prática de indexação, no contexto do arquivo do jornal “O Estado de São Paulo” e no Departamento de Documentação da Editora Abril, foram evidenciados quando comparados com os procedimentos estabelecidos por Ahmad (1991) para teoria de indexação de jornal. E que o procedimento, questionamento, durante a leitura do texto, quanto às escolhas dos descritores foi realizado, em maior número, pelos indexadores do DEDOC.

A análise dos resultados da observação da leitura dos indexadores e a comparação entre teoria e a prática de indexação dos jornais pesquisados nos permitem destacar que, de modo geral, os indexadores observados são profissionais experientes, com boa formação educacional e possuem anos de prática na realização da atividade de leitura de jornais sendo portanto estratégicos e rápidos na identificação e seleção de conceitos.

As estratégias utilizadas são selecionadas de acordo com o temática do texto, levando em consideração o grau de conhecimento do leitor em relação ao assunto tratado no texto, o conhecimento da estrutura textual do texto, o objetivo da leitura, a experiência do leitor com determinado tipo de texto, entre outras variáveis. E com relação aos indexadores, destacamos que usam estratégias metacognitivas e o conhecimento de estrutura textual.

Apesar do conhecimento da estrutura textual do texto não ter sido explicitamente mencionado pelos indexadores do DEDOC, a leitura se concentrou também no “LEAD” e, quando necessário, atingiu o item “Episódio (Atos da fala) em busca de conceitos que possibilitassem representar o conteúdo do texto.

Ressalta-se, também, que, comparando o desempenho dos sujeitos entrevistados durante a leitura dos textos – base, percebe-se que a estratégia mais relevante utilizada por esses indexadores é o conhecimento prévio do assunto, e que os indexadores dominam a Linguagem Documentária utilizada, pois, após extração dos termos, ressaltam se a linguagem Documentária possui determinado termo ou não.

Quanto à Linguagem do arquivo do jornal “O Estado de São Paulo”, podemos dizer que é uma listagem alfabética de termos, que se encontra obsoleta em relação a alguns assuntos, pois na Análise do Protocolo Verbal, nem todos os conceitos selecionados existiam na Linguagem.

Com relação à Linguagem do DEDOC, ressalta-se que todos os termos identificados pelos indexadores constavam da Linguagem, o que indica a eficácia da Linguagem do sistema em representar os assuntos indexados. É importante destacar que a linguagem do DEDOC também auxilia o indexador durante a seleção dos termos, alertando quando o termo é controlado ou proibido, em se tratando de termos sinônimos ou com entradas diferentes, além de possibilitar, ao indexador, a inclusão de termos que considera relevantes durante a indexação, e que não constam da Linguagem.

Ressaltamos também que os indexadores, durante a leitura, fazem uso de “esquemas”, ou seja, utilizam o conhecimento prévio que têm armazenado na memória pois, o indexador denominado Sujeito 2, durante a leitura do texto 4, recorre à memória de longo prazo, para recordar se já havia indexado determinado assunto e quais termos foram extraídos, porque no jornal uma mesma matéria jornalística pode se prolongar no decorrer dos dias da semana. Caso a matéria tenha sido indexada, o indexador somente confirma os descritores anteriormente extraídos e, se necessário, incorpora a estes, novos descritores.

Quanto ao uso de estratégias de leitura, podemos destacar indícios do conhecimento da estrutura textual que os indexadores possuem, pois sabem o momento de interromper a leitura, uma vez que, no término, afirmam que não precisam mais ler pois não lhes interessa mais nada no texto, diferenciando informações principais das secundárias, e até que item da estrutura textual lhe interessa ler, na indexação.

Os sujeitos (indexadores) utilizam no processo de leitura, da maioria dos textos analisados, a mesma seqüência de exploração da estrutura textual para identificar conceitos, sendo a parte mais explorada o “Lead”.

Dentre as estratégias metacognitivas listadas por Brown, os indexadores, na maioria dos textos, explicitam o objetivo da leitura.

Quanto à sistemática de identificação de conceitos, podemos dizer que os indexadores realizam:

- realizam o destaque de palavras do texto através 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 às escolhas dos descritores:

- identificam que há maneiras diferentes de se ler;
- evidenciam conhecimento de estrutura textual;
- usam conhecimento prévio e recorrem à memória de longo prazo;

Com a observação da leitura dos indexadores e estudos de revisão da literatura, destacamos as suposições que formulamos quando decidimos investigar a respeito da atividade de leitura desses indexadores, referindo-nos, primeiramente, ao fator tempo disponível para ler, em que supúnhamos que a leitura documentária de jornais deveria ser rápida devido a 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 proporcionam 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 quê? Como? (Para quê?) Onde? Por quê? (Para quê?) Quando?, e que permite ao indexador compreender o texto e identificar o tema abordado antes de ler o texto detalhadamente. Os indexadores evidenciaram conhecer o questionamento proposto pela estrutura do texto e apresentaram conhecimento da estrutura textual, porque concentraram a seleção de conceitos nos itens “Lead e Episódio”. 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 referentes ao tema presente, por meio de leitura, etc...), referente ao tema que interaja no momento da leitura, com as informações do texto. Percebemos, então, que a memória do indexador é um fator primordial para a aquisição 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, provavelmente, é o mesmo.

Uma outra suposição era que os indexadores de jornais 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”, o questionamento para verificar compreensão, o uso do conhecimento prévio, etc... Entretanto, notou-se que após seleção de conceitos, realizam pouca associação com a Linguagem . Nos perguntávamos, 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, etc...

Sendo assim, confirmamos nossas suposições, evidenciando que os indexadores de jornais são estratégicos, rápidos e dominam a estrutura textual para a identificação e seleção de conceitos. Com a confirmação das hipóteses, acima, atingimos, primeiramente, um dos nossos objetivos que é a identificação de estratégias de leitura para artigos de jornal. E isto nos possibilita alcançar um segundo objetivo, que é elaborar um modelo experimental de leitura para artigos de jornal, considerando os resultados da observação de leitura dos indexadores e recomendações da literatura, possibilitando, assim, que outros indexadores de jornais (que deduzimos Ter, pelo menos, conhecimento prévio mínimo do assunto) conheçam e utilizem esse modelo para suprir deficiências referentes à localização dos conceitos, durante a indexação.

Portanto, propomos, no capítulo seguinte, a primeira versão do modelo experimental de leitura para indexadores de jornais, atingindo, assim, em parte, um objetivo de nossa pesquisa que é identificar as estratégias de leitura utilizadas pelos indexadores experientes visando compará-las com as recomendações da revisão de literatura, e desenvolver um modelo de leitura que auxilie os indexadores que apresentam dificuldades na identificação de conceitos.

Esclarecemos que o modelo de leitura a ser proposto será ensinado a um indexador integrante dos indexadores do DEDOC (caracterizado como sujeito - Clara) que realizam a atividade de indexação esporadicamente, porque não possuem experiência na realização da tarefa, e apresentam dificuldades para identificar conceitos nos textos, por serem inexperientes na realização da tarefa, ficam incumbidos de auxiliarem os indexadores experientes, esporadicamente, no momento de acúmulo de serviço e, no

restante do tempo, realizam digitação de pastas de recortes para a implantação do DEDOC Eletrônico que se constitui de artigos de jornais on-line e indexados. Posteriormente, observaremos o processo de indexação dos sujeitos do grupo. E, no mesmo momento, em que aplicaremos o uso do modelo, observaremos, também, a leitura de um par de indexadores do DEDOC (caracterizado como sujeitos – Ana e Maria) que também realiza a atividade de indexação, esporadicamente, mas destacamos que este grupo não utilizará o modelo para que, no final, possamos comparar os resultados atingidos pelos dois grupos, e verificar a eficácia do modelo proposto objetivando assim, possibilitar, tanto ao indexador com anos de prática na atividade de indexação, como a um indexador iniciante em que ambos apresentam dificuldades para identificar conceitos, eficácia na realização da tarefa de indexação por meio de uma leitura rápida do texto jornalístico, o que vem a ser necessário devido a quantidade de artigos a indexar.

Portanto, o objetivo de nossa pesquisa é propor um modelo de estratégias de leitura para indexação de artigos de jornais. Para tanto, foram realizados estudo teórico sobre o tema e estudo prático envolvendo análise da observação da prática de indexação de jornais com estudos de caso no arquivo do jornal “O Estado de São Paulo” (datado de 1997) e Departamento de Documentação da Editora Abril (datado de 2000). Foram realizados ainda, identificação da política de indexação das instituições, avaliação das linguagens documentárias, caracterização dos indexadores, e diagnóstico da infraestrutura física e funcional das instituições.

A análise de todos os capítulos desta dissertação, descrita a seguir, mostrou os resultados relevantes obtidos nos estudos teóricos e nas observações da atividade de leitura observando características especiais, dados relevantes de resultados que se relacionam com outros, observações teóricas e práticas de outros autores e, nossa análise. Os resultados desta análise propiciaram identificar critérios para a elaboração do modelo experimental de estratégia de leitura para indexadores de jornais.

Para melhor visualização, a análise do relatório foi dividida em duas partes: parte conceitual e parte diagnóstica. A parte conceitual envolve as considerações teóricas e práticas enfocadas na literatura pertinente à análise documentária, evidenciando a atividade de leitura dos indexadores de jornais por meio da observação de identificação e seleção de conceitos, e da exploração da estrutura textual. Enquanto que a parte diagnóstica ressalta a avaliação comparada dos dados coletados a partir da observação da leitura dos indexadores

dos dois estudos de caso, o arquivo do jornal “O Estado de São Paulo” e o Departamento de Documentação da Editora Abril.

a) PARTE CONCEITUAL

- a leitura documentária é uma atividade objetiva onde os conhecimentos do leitor são confrontados com o conteúdo do texto na busca da compreensão visando a seleção de conceitos;
- este conhecimento prévio do indexador, que fica armazenado na memória de longo termo referente ao assunto do texto que tem que indexar, o auxilia na compreensão;
- um outro conhecimento prévio é o de estrutura textual que auxilia o indexador na leitura à medida que ele conhece as partes do texto, consegue identificar o conceito pertencente a cada parte, e em que partes está a informação relevante que procura, realizando, assim, uma leitura mais rápida, o que para fins documentários é ideal devido a grande quantidade de material a tratar;
- na leitura do indexador, a seleção eficiente de conceitos considera as necessidades informacionais dos usuários, a capacidade de compreensão do indexador, política de indexação do sistema e a qualidade e flexibilidade da Linguagem na representação dos termos;
- a eficiência da seleção de conceitos pertinentes, considerados como representantes do conteúdo do texto que, por outro lado, devem refletir os interesses informacionais dos usuários, relaciona-se com o conhecimento prévio do indexador sobre o assunto do texto;
- a política de indexação influencia na seleção de conceitos porque estabelece a que categoria de usuários quer servir; e quando se escolhe uma classe em especial, opta-se pela seleção de conceitos mais específicos e; o contrário, sendo vários grupos de usuários de áreas diferentes, a seleção se direcionará para conceitos mais genéricos que possam ser recuperados pelos vários usuários;
- a qualidade da linguagem de indexação no momento da seleção de conceitos pode ser estabelecida pelo total de termos selecionados e compatíveis de representação pela Linguagem, por propiciar controle de vocabulário e inclusão de termos. Sendo assim podemos destacar que a linguagem do DEDOC é eficiente na indexação dos documentos porque se encaixa nesses itens;

- a leitura envolve compreensão e, sendo a primeira atividade da análise documentária, dela dependerá a qualidade das outras atividades;
- portanto, interessa confirmar que a identificação e seleção de termos são realizadas durante a leitura, e quais estratégias são utilizadas;
- somente com as observações da leitura será possível melhorar o desempenho de indexadores e verificar o uso de metodologias de indexação e, se preciso, adaptar e criar novas metodologias;
- na literatura encontramos como metodologia de indexação a Norma ISO 5693 que destaca que a leitura do documento é um exame de suas partes; percebemos que a norma aconselha a utilização da estrutura textual durante a leitura documentária;
- entretanto, para a Norma, a identificação de conceitos não está vinculada à leitura porque a Norma recomenda que o indexador identifique o conceito por meio de um questionamento que estabeleceu, e que deve ser realizado após a leitura do texto;
- E a Norma ainda ressalta que a seleção de conceitos é realizada após a identificação de conceitos;
- Mas entendemos, pela observação da leitura dos indexadores, que as orientações da Norma, na prática, não ocorrem, porque os resultados das observações destacam que a identificação e a seleção de conceitos presentes no conteúdo do documento são realizadas durante, e através da leitura;
- ressaltamos ainda que, mesmo que o indexador siga a metodologia de identificação e seleção de conceitos proposta pela Norma, somente conseguirá identificar e selecionar conceitos pertinentes, se tiver compreendido o texto;
- acreditamos que o uso do vocabulário controlado, durante a leitura, para a identificação e seleção de conceitos, impedirá a compreensão do contexto do documento porque o indexador pode ficar preso às formas de representação oferecidas pela Linguagem, e não selecionar o assunto considerando as informações abordadas no texto, e que são de interesses dos usuários;
- encontramos na literatura três concepções de análise de assunto: concepção simplista, concepção orientada para o conteúdo, e concepção orientada para demanda (Albrecthsen 1993, p.220);
- e segundo Fujita (2000) podemos considerar que os "Princípios do Unisist" e a Norma ISO contêm as duas concepções, a orientada para o conteúdo, porque propõe a abordagem sistemática para a identificação de conceitos, e a orientada para a demanda, porque orienta a seleção dos conceitos conforme objetivos de uso dos termos;

- O conteúdo do documento estará melhor representado, se a identificação e seleção de conceitos forem realizadas dentro das concepções orientadas para o conteúdo que seja refletida na demanda;
- A compreensão da leitura exige que o indexador tenha conhecimento prévio sobre o assunto, e que durante a leitura utilize este conhecimento de maneira consciente, comparando as informações presentes no texto com o conhecimento que possui armazenado na memória sobre o assunto do texto;
- Além do conhecimento sobre o assunto o indexador deve conhecer, como mencionamos acima, a organização do texto, ou seja, sua estrutura textual no caso do texto jornalístico; ele deve perceber que as informações contidas no (s) primeiro (s) parágrafo (s) apresentam uma síntese do assunto abordado na sequência do texto, e essa parte da estrutura textual é definida como “Lead” que contém respostas ao questionamento “quem?; quando?; onde?; como?, ou Para quê? Por quê?, ou Para quê?” (VAN DIJK, 1983; SILVA, 2001);
- Assim como o “Lead”, o título e subtítulo da matéria jornalística, denominados como “Cabeçalho” na estrutura textual, também condensam o assunto principal do texto; entretanto, ressalta-se que os textos que apresentam metáforas, como exemplo, os cadernos de Economia, quando analisados, somente com a leitura do título, podem distorcer a compreensão do conteúdo geral do texto;
- Percebemos que existem partes ou itens da estrutura textual que são mais significativos em conteúdo do que outros, e que o conhecimento deste diferencial possibilitará uma leitura mais rápida;
- A leitura documentária de jornais exige que o indexador identifique e selecione conceitos, e realize também a extração dos elementos descritivos do texto identificando autor, título, nome do jornal, dia e data de publicação, caderno que serão pontos de acesso para se recuperar o texto; (AHMAD, 1991)
- Na indexação, o indexador deve compreender a linguagem e a terminologia do jornal, utilizar para representar os conceitos um tesouro de notícias padrão, mas deve estar consciente de que, se necessário, deve “criar/inventar novos descritores quando eles não forem encontrados no tesouro de notícias”; (AHMAD, 1991)
- As observações de leitura dos indexadores (DEDOC e do arquivo do jornal “O Estado de São Paulo”) revelaram o uso de algumas estratégias metacognitivas durante toda a leitura, o que nos leva a considerar que a leitura, para fins documentários, exige que o

leitor seja metacognitivo (consciente) porque tem um objetivo na leitura que é a seleção de conceitos;

- “ a extração de informação documentária para produzir resumos e índices pode ser eficiente se o indexador conhecer a superestrutura textual, ou seja, como o texto está organizado” (FUJITA, 1999);
- os linguistas Van Dijk & Kintsch (1983) consideram que o leitor que domina essa superestrutura, torna-se capaz de monitorar a leitura, integrando as informações exploradas no texto àquilo que é central;
- e a estrutura do texto jornalístico foi elaborada por Van Dijk (1983);
- É importante destacar que, com os resultados das observações da leitura dos indexadores de jornal, foi possível identificar que na indexação, em busca da compreensão, o indexador deve explorar primeiramente o “Cabeçalho e o Lead” por apresentar um resumo da matéria, e quando estes não forem suficientes, a leitura do indexador se estenderá, no máximo, até o item “Episódios (Atos da fala da personagem)” porque, segundo o autor, os outros elementos da estrutura textual de jornais que são “Comentários” “,são a expressão típica, assinalada como tal, das crenças do jornalista, ou do jornal”

b) PARTE DIAGNÓSTICA

Destacamos aqui os pontos relevantes dos estudos de observação da leitura documentária com indexadores de jornais.

- Arquivo do jornal “O Estado de São Paulo”
 - A observação da leitura dos indexadores do Arquivo, devemos destacar, que foi realizada no ano de 1997 durante nossa pesquisa de curso de graduação. Naquele ano, o arquivo não apresentava nenhum manual de indexação e realizava a indexação dos jornais publicados pelo Jornal como também de outros jornais;
- Departamento de Documentação da editora Abril
 - No departamento de documentação da Editora Abril, a observação da atividade de leitura foi realizada no ano de 2000, mas também foi constatado que a instituição não possuía nenhum manual de indexação e que indexava vários jornais;

Foi constatado, por meio de conversa informal com os indexadores e pela análise dos dados das entrevistas que as instituições, (Arquivo do jornal “O Estado de São Paulo” e Departamento de Documentação da Editora Abril), apesar de não possuírem um manual de indexação, orientam os indexadores a realizar a leitura considerando o conteúdo das matérias e as necessidades informacionais dos jornalistas (o que determina a seleção de conceitos específicos para as necessidades informacionais destes), que são os principais usuários; as instituições “exigem” que a leitura seja rápida. Portanto, podemos afirmar que, nas instituições, as concepções de análise do documento são direcionadas para o conteúdo e para a demanda. É importante destacar que o DEDOC aconselha os indexadores a trabalharem no setor de atendimento à pesquisa para que conheçam os assuntos de interesse dos usuários e suas formas de recuperar os documentos.

Destacamos, ainda, que a política de indexação dos documentos no DEDOC inclui avaliação dos termos indexados, porque, passado um certo tempo, os conceitos antes atribuídos aos documentos podem sofrer alterações, incluindo, tanto inclusão de mais termos como exclusão dos anteriormente indexados, como exemplo: demonstramos que as fichas de indexação preenchidas durante nossa observação de leitura dos indexadores, passaram por algumas alterações como demonstra o (ANEXO D).

Avaliação de Linguagens Documentárias

- as Linguagens documentárias dos dois sistemas de informação observados (Arquivo do jornal “O Estado de São Paulo” e Departamento de Documentação da Editora Abril) são uma listagem alfabética de termos e que estão longe de ser um tesouro. Por esta razão, avaliamos as Linguagens a partir de uma análise formal de suas estruturas externas, baseando-nos na metodologia de Corrêa (1998), e também de suas estruturas internas utilizando a metodologia de Ginez de Lara (1993). Não foi possível a aplicação da metodologia de Ginez de Lara (1993) à linguagem do Arquivo, limitando a avaliação desta somente a sua estrutura formal;
- foi constatado que ambas as Linguagens não apresentam uma hierarquia entre os termos, que permita identificar termos genéricos e específicos, o que para fins de consistência na identificação de conceitos é uma barreira;

- evidenciou-se que a Linguagem do DEDOC, no item estrutura dos componentes da linguagem, destaca que a existência do termo identificador é um fator que possibilita diminuir a ambigüidade dos termos sinônimos;
- na linguagem do DEDOC, a estrutura formal apresenta-se dividida em três categorias: Descritor (assuntos), Personalidades (pessoas) e Identificador (empresas, instituições e corporações);
- entretanto, na Linguagem do Arquivo esta divisão não aparece, mas notou-se que a identificação de termos considera a existência destas três categorias;
- a avaliação das Linguagens quanto à análise formal das estruturas, proposta por Corrêa (1998), que foi aplicada a ambas, evidenciou-se com a análise dos resultados que , a linguagem do DEDOC apresenta uma estrutura organizada, enquanto que na Linguagem do arquivo notaram-se falhas;
- os resultados obtidos mostraram indícios de que a Linguagem do DEDOC, ao contrário da do Arquivo, mostra-se, provavelmente, como mais adequada a uma boa indexação.

Quanto aos indexadores

- No Arquivo do jornal “O Estado de São Paulo” os indexadores eram jornalistas e bibliotecários;
- Os indexadores do DEDOC são jornalistas, sociólogos, historiadores, entre outros.

Ressaltamos que, em ambas as instituições, os indexadores, em sua maioria, possuem mais de um curso de graduação. E estão fazendo cursos de pós-graduação. (mestrado)

Estudo de observação da leitura do indexador: comparação dos resultados

Introduziremos aqui os resultados da análise das entrevistas, ou dos protocolos com os indexadores do Arquivo do jornal “O Estado de São Paulo” (caracterizados como sujeitos 1 e 2) e do Departamento de Documentação da Editora Abril (caracterizados como sujeitos 3, 4 e 5).

a) Seqüência de operações (aspectos) da leitura dos sujeitos

- As três operações mais utilizadas pelos indexadores das duas instituições foram a “identificação de aspectos importantes da mensagem” e “alocamento de atenção a áreas importantes” e “seleção dos termos de indexação”;

- Os indexadores do DEDOC, todos explicitaram que os objetivos da leitura eram representar o conteúdo e possibilitar a recuperação deste pelos usuários;
- Enquanto os do Arquivo, somente o sujeito 1, no fim do texto 1, menciona que o objetivo da leitura é indexar o conteúdo da matéria visando a recuperação pelos jornalistas;
- As operações, que dizem respeito à monitoração do comportamento, foram realizadas pelos indexadores do DEDOC onde se destaca o questionamento quanto à compreensão do texto;

b) manter em mente o seu objetivo de representar o texto para futura recuperação

- os sujeitos 1, 3, 4 e 5 explicitaram manter em mente o objetivo da leitura durante a realização da tarefa, enquanto que o sujeito 2 não o fez e nem o mencionou;
- foi observado que todos os indexadores do DEDOC expressaram que objetivo da leitura era indexar visando a recuperação pelos usuários;

a) exploração do conhecimento de estruturas textuais

- os indexadores não explicitaram conhecer a estrutura textual, e somente os sujeitos 1 e 5 mencionaram indícios de conhecer a estrutura do caderno com o qual trabalham;
- observou-se que os sujeitos, na maioria dos textos, identificaram conceito no item “Lead” da estrutura textual;
- notou-se, também, que quando foi preciso estender a leitura para além do “Lead” em busca de conceitos, identificaram-se os conceitos no item “Episódio (s)” .

b) Identificação de aspectos importantes da mensagem para identificação de conceitos

Partes da estrutura textual em que os conceitos foram identificados e selecionados

- destaca-se que os sujeitos 2 e 5 foram os que mais exploraram a estrutura textual, enquanto os outros concentraram a leitura no item “Lead”;
- percebemos que o item “ Lead” é o mais significativos para se identificar o conteúdo do texto porque, na maioria dos textos, a exploração desse item foi suficiente para se identificar os assuntos dos textos;

c) Sistemática de identificação de conceitos: verificou-se que os sujeitos realizam a identificação de conceitos utilizando estratégias metacognitivas (conscientes), apresentando indícios de uma sistemática constante:

- retiram os conceitos dos itens “Lead; Cabeçalho e Episódios”;

- os indexadores do DEDOC selecionam os conceitos se questionando quanto à compreensão do texto;
- todos os sujeitos, após a seleção, representam os termos selecionados seguindo padrões estabelecidos pela Linguagem documentaria dos sistemas. Destaca-se que o indexador, sujeito 1, do Arquivo possui conhecimento especializado na área porque a linguagem não o auxilia muito, porque os termos selecionados ela não os possui e, mesmo assim, indexa sem maiores dificuldades;
- a associação com a Linguagem foi pouco mencionada. Percebeu-se que quando ocorreu foi realizada (sujeito 2 texto 3) (depois pelos sujeitos 3, 4 e 5) através da seleção dos conceitos.

Da análise dos resultados observou-se que os indexadores:

- se questionam quanto à compreensão do texto no momento da identificação e seleção dos conceitos;
- exploram a estrutura textual na maioria dos textos nos mesmos itens, mas não explicitam verbalmente conhecer esta estrutura textual, talvez por ser esta atividade tão repetitiva que se tornou um processo automático;
- apresentaram indícios de uma sistemática de identificação de conceitos constantes.

Compatibilidade dos termos selecionados com a linguagem do sistema

- A linguagem do Arquivo não possui a maioria dos termos selecionados. Entretanto, a do DEDOC possui todos os termos selecionados pelos indexadores.

Em síntese, destacamos as considerações obtidas com a análise da parte conceitual e diagnóstica.

Sendo assim, evidenciamos, então, com a análise das entrevistas da leitura dos indexadores do arquivo do jornal “O Estado de São Paulo” e do Departamento de Documentação da Editora Abril que:

- os sujeitos utilizam estratégias metacognitivas como (monitoração do comportamento para ver se está ocorrendo compreensão; explicitação dos objetivos da leitura (ou manutenção dos objetivos na mente, entre outras); realizam a identificação de conceitos utilizando diferentes estratégias; observou-se indícios de que seguem uma sistemática constante;
- e que os indexadores do Arquivo, sujeitos 1 e 2 dominam a Linguagem que apresentou-se estar bem desatualizada e, sendo a eficiência na indexação, mérito dos

indexadores que são especialistas no assunto, à medida que suprem a deficiência da Linguagem;

- em ambas as instituições o conhecimento da estrutura textual não foi expresso verbalmente pelos sujeitos, mas notou-se que exploram determinadas partes do texto para identificar conceitos e, posteriormente, indagam se a matéria está indexada, o que subentende que têm consciência de que a leitura de determinadas partes do texto sejam suficientes;

- entretanto, o sujeito 1, durante a leitura do texto 2, menciona indícios de que a conhece e explora porque afirma “eu não preciso mais ler porque o resto do texto vai tá falando deles”;

Com a análise dos resultados da leitura dos indexadores, evidenciamos que são metacognitivos e estratégicos, e demonstraram indícios de uma metodologia de identificação de conceitos constante, que considera a exploração de determinadas partes da estrutura textual de jornais, de modo a garantir a estratégia de identificação de conceitos conjuntamente com o uso de outras estratégias de leitura como: questionamento quanto a sua compreensão do texto. Sendo assim, será interessante que outros indexadores conheçam como estes sujeitos realizam a indexação. Para isso, estabelecemos, nas páginas seguintes, um modelo provisório de leitura para indexação de jornais que estabelece uma sistemática de identificação de conceitos para a indexação, por meio da exploração da estrutura textual de jornais com o questionamento sobre que conceitos podem identificar nas partes mais significativas do texto. E, nas etapas seguintes desta pesquisa, este modelo provisório de leitura será melhorado, ensinado a outros indexadores e avaliado, visando identificar sua eficácia na leitura de indexadores tidos como não experientes.

Considerando a análise dos resultados levantados na parte conceitual, que indicam que a leitura do indexador visa a compreensão, e que no momento da compreensão o leitor-indexador também utiliza seu conhecimento prévio sobre o assunto do texto, e que isto constitui uma estratégia consciente para se compreender o texto, e que esta pode ser utilizada em maior (top-down) ou menor (bottom-up) grau durante a leitura, entretanto ressaltamos, que sua presença é essencial para que haja compreensão. E que o menor uso de estratégias metacognitivas indica maior legibilidade do texto para o leitor. E como os autores não prevêm o indexador como leitor alvo, este é obrigado, na leitura de um texto cujo assunto não lhe é familiar, realizar uma leitura lenta, detalhada, acionando movimento “bottom-up” para chegar ao movimento “top-down”; entretanto, quando estiver familiarizado com o assunto do texto realizará uma leitura mais fluente realizando

movimento top-down em direção ao conteúdo global para chegar ao movimento “bottom-up”. Mas para agilizar sua compreensão e realizar uma leitura rápida, o indexador deve ser mais estratégico que o leitor normal e isso, seja especialista ou não, é necessário para a identificação de conceitos que é o objetivo da leitura do indexador.

Sendo assim, se faz necessária uma abordagem sistemática de identificação que utilize estratégias conscientes que esperamos obter com a elaboração do modelo de leitura provisório para a indexação de jornais, proposto com base nos resultados da observação da leitura de indexadores experientes do arquivo do jornal “O Estado de São Paulo” e do Departamento de Documentação da Editora Abril e metodologias e recomendações de identificação de conceitos encontrados na revisão de literatura.

6.2 Modelo de Leitura para Artigo de Jornal

Enfatizaremos aqui a elaboração do modelo de estratégia de leitura documentária para indexação de jornais utilizando o domínio da estrutura textual para identificar conceitos evidenciando a análise de conteúdo do texto jornalístico.

Para concretizar esta proposta foi realizado ensaio teórico para definição do modelo baseando-se em revisão de literatura descrita nos três capítulos iniciais (capítulos 1, 2 e 3), nos resultados da avaliação comparada dos dados coletados a partir da observação da leitura dos indexadores do arquivo do jornal “O Estado de São Paulo” e do Centro de Documentação da Editora Abril, metodologia e recomendações de identificação de conceitos encontrados na revisão de literatura considerando as mais relevantes.

E também, no modelo de leitura que Fujita (2000) propôs direcionado ao texto científico elaborado com base no conhecimento das metodologias de análise do documento proposto pela área, e nos modelos de leitura encontrados aqui na revisão de literatura sobre o assunto como os de: Tálamo (1994); Kobashi (1994); Van Dijk & Kintsch (1978) e Van Dijk & Kintsch (1983), e nas suas observações da leitura dos indexadores do Centro de Energia Nuclear e da Sub-rede Nacional de Informações em Ciências da Saúde.

Evidenciamos o modelo de leitura de Fujita (2000) porque nos baseamos nele para propor o modelo provisório para indexação de jornais. Optamos por nos orientar por este modelo, porque, mesmo sendo um modelo experimental, apresenta uma metodologia relevante que evidencia a análise de assunto por meio de questionamento para identificação de conceitos e exploração da estrutura textual do documento observado.

Na elaboração do modelo de leitura para texto científico, Fujita (2000) considerou os resultados obtidos na revisão de literatura sobre leitura e leitura documentária e os resultados das observações da leitura dos indexadores dos centros de estudos de caso citados acima *“considerando que é necessário a elaboração de um modelo de leitura que utilize o conhecimento de estruturas textuais do leitor apoiando em estratégias de abordagem sistemática de identificação de conceitos”*.

Para elaboração do modelo de leitura, Fujita (Op. Cit, 114) abordou:

- concepção orientada para o conteúdo para a análise de assunto (ALBRECHTSEN);
- modelo de leitura para indexação e resumos de Kobashi (1994);
- a identificação do tema do texto pelo objetivo enunciado (TÁLAMO, 1994);
- procedimentos de análise de assunto com a análise conceitual do PRECIS (FUJITA, 1995);
- abordagem sistemática para identificação de conceitos (Norma ISO 5963).

Esse modelo de leitura de Fujita (Op. Cit) consiste em identificar termos por meio das respostas aos questionamentos propostos a determinadas partes da estrutura textual do texto científico. Com a obtenção destas respostas, selecionam-se os conceitos que, posteriormente, são verificadas na Linguagem do sistema, sendo possível ainda a construção de um enunciado de assunto (breve comentário) sobre o texto organizando, de forma lógica, os conceitos identificados.

Aplicação do modelo de leitura: identificação de conceitos por questionamento em partes da estrutura textual (Proposto por FUJITA, 2000, 118)

IDENTIFICAÇÃO DE CONCEITOS (PRECIS) POR QUESTIONAMENTO (NORMA ISO) EM PARTES DA ESTRUTURA TEXTUAL DE ARTIGO CIENTÍFICO

ARTIGO: VIEIRA, L. A. et al. Colonização intestinal de recém-natos por enterobactérias multirresistentes a antimicrobianos em unidade neonatal. *Jornal de Pediatria*, v.75, n.2, p.83-90, mar./abr. 1999.

Quadro 17 : Demonstração da aplicação do modelo de leitura para artigos científicos (FUJITA, 2000)

CONCEITO	QUESTIONAMENTO	PARTES DA ESTRUTURA TEXTUAL	TERMOS IDENTIFICADOS	DeCS
OBJETO	O documento possui em seu contexto um objeto sob efeito de uma atividade?	INTRODUÇÃO O (OBJETIVOS)	Recém-nascido	Recém-nascido
AÇÃO	O assunto contém um conceito ativo (por exemplo, uma ação, uma operação, um processo, etc)?	INTRODUÇÃO O (OBJETIVOS)	Colonização intestinal	Colonização
FUNÇÃO	O objeto é influenciado pela atividade identificada?	INTRODUÇÃO O (OBJETIVOS)	Sim	
AGENTE	O documento possui um agente que praticou esta ação?	INTRODUÇÃO O (OBJETIVOS)	Germes multirresistentes a antimicrobianos	Enterobactérias USE enterobacteriaceae
MÉTODOS DO AGENTE	Este agente refere-se a modos específicos para realizar a ação (por exemplo, instrumentos especiais, técnicas ou métodos)?	METODOLOGIA	• Identificação bacteriana; • Testes de sensibilidade aos antimicrobianos; • Amostras fecais; • Extração de DNA plasmidial bacteriano; • Eletroforese em gel de agarose	• Não tem • Teste de Sensibilidade Bacteriana - USE teste de Sensibilidade Microbiana • Fezes • DNA bacteriano • USE eletroforese em gel agar

LOCAL OU AMBIÊNCIA	Todos estes fatores são considerados no contexto de um lugar específico ou ambiente?	METODOLOGIA •	• Berçários hospitalares; • Hospitais públicos
CAUSA E EFEITO	São identificadas algumas variáveis dependentes ou independentes ?	RESULTADO • S; DISCUSSÃO DE RESULTADO S	• Klebsiella • Pneumoniae; • Baixo peso ao nascer; • Leite de vaca USE substitutos do leite humano; • Alimentação artificial; • Hospitalização
PONTO DE VISTA DO AUTOR; PERSPECTIVA	O assunto foi considerado de um ponto de vista, normalmente não associado com o campo de estudo (por exemplo, um estudo sociológico ou religioso)?	CONCLUSÃO S	

ENUNCIADO DE ASSUNTO: "Colonização intestinal de recém-nascidos por germes multirresistentes a antimicrobianos mediante identificação bacteriana, testes de sensibilidade aos antimicrobianos, amostras fecais, extração de DNA plasmidial bacteriano e eletroforese em gel de agarose em Berçário do HUPE-UERJ cujas causas foram a infecção por *Klebsiella pneumoniae* e baixo peso que resultou na alimentação com leite artificial, hospitalização e uso de antimicrobianos."

Nota: Os dados acima referem-se a continuação do Quadro 17: Demonstração da aplicação do modelo de leitura para artigos científicos (FUJITA, 2000)

O modelo provisório de leitura para artigos de jornais que relatarmos aqui segue alguns dos procedimentos do modelo para texto científico elaborado por FUJITA (2000) citado acima. Devendo destacar que é um modelo teórico e provisório sendo, portanto, um primeiro ensaio em busca de um modelo eficiente para indexadores identificar conceitos em textos jornalísticos.

O modelo teórico de leitura que propomos consiste na identificação de conceitos por meio do questionamento direcionado a determinadas partes da estrutura textual de jornal evidenciando:

- a exploração da estrutura textual do texto jornalístico proposta por Van Dijk (1983) citada aqui no capítulo 2;
- questionamento encontrado no ‘Lead’ da estrutura textual do texto jornalístico (VAN DIJK, 1983 ; SILVA, 2001)
- alguns dos procedimentos de análise de assunto com a análise conceitual do PRECIS (FUJITA, 1989).
- a concepção orientada para o conteúdo para a análise de assunto (ALBRETCHESSEN, 1993);

O modelo para indexação de jornais que propomos optou por considerar a exploração da estrutura textual para texto jornalístico porque o conhecimento da estrutura textual auxilia na identificação de conceitos. E optou-se, ainda, por utilizar ainda o questionamento presente no “Lead” por ser uma metodologia de análise conceitual direcionada ao texto jornalístico, o que vem justificar a nossa opção por descartar de nosso modelo a metodologia de análise conceitual da ISO (5963).

Evidenciamos que utilizamos, nesse primeiro momento, na elaboração do modelo proposto, também alguns dos procedimentos de análise de assunto do sistema Precs de indexação, por ser uma outra maneira de identificar os conceitos explorando a estrutura textual de jornal, por meio de questionamento, possibilitando, assim, verificarmos qual dos questionamentos é mais fácil para os indexadores que testaram o modelo, identificar os conceitos.

O modelo proposto considera a concepção orientada para o conteúdo estabelecida por Albretchtsen (1993), porque direciona-se à identificação do conteúdo do texto jornalístico por meio também da interpretação do texto (conhecimento prévio comparado com a informação do texto).

Assim, o modelo estabelece os seguintes procedimentos para a leitura do texto jornalístico:

- 1- Inicie a leitura do texto buscando prever¹ 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 5 primeiros parágrafos, sem realizar leitura

linear, buscando encontrar respostas aos questionamentos propostos pelo modelo de leitura;

- 3- Identifique as respostas, selecione destas os conceitos e compare com a Linguagem do DEDOC;
- 4- Por fim, elabore um resumo considerando o questionamento do Lead organizando as respostas obtidas em ordem lógica para as questões: O quê?; Quem ?; Quando?; Onde? Como?(ou Para que?); Por quê? (ou Para quê?)

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

Quadro 18: Modelo de leitura para texto jornalístico

CONCEITO (PRECIS)		PARTE DA ESTRUTURA TEXTUAL (VAN DIJK, 1983)	LEAD (itens componentes) VAN DIJK (1983) ; SILVA (2001)
AÇÃO		Sumário/Introdução (Cabeçalho e Lead)	O quê (o que aconteceu, está ou vai acontecer)
AGENTE		Sumário/Introdução (Lead)	Quem (os agentes da Ação)
MÉTODOS DO AGENTE		Sumário/Introdução (Lead) ou Episódio (s)	Como, ou Para quê (as circunstâncias)
LOCAL AMBIÊNCIA	OU	Sumário/Introdução (Cabeçalho e Sumário/Introdução (Cabeçalho e Lead)	Quando (dia da semana e do mês, horas) Onde (o local do acon- Estabelecimento)
CAUSA E EFEITO		Sumário/Introdução (Lead) ou Episódio (s)	Porquê, ou Para quê (os motivo e as razões)

Aplicação do modelo provisório de leitura para artigos de jornais a um texto

Texto utilizado:

MUG, M. *PPB deve concluir hoje alteração em 300 outdoors*. Jornal “O Estado de São Paulo”. Caderno Eleições 2000. 25 de outubro de 2000. P. h3.

Quadro 19: Aplicação do modelo de leitura de texto jornalístico

Conceito (PRECIS)	Parte da estrutura Textual (VAN DIJK)	Conceito (Lead)	Termos Identificados	Linguagem DEDOC
AÇÃO	Lead	O quê (o que aconteceu, está ou vai acontecer)	Crime Eleitoral	Crime Eleitoral
AGENTE	Cabeçalho e Lead	Quem (os agentes da acção)	PPB/Partido	PPB/Partido
MÉTODOS DO AGENTE	Lead	Como (as circunstancias)	Propaganda Eleitoral	Propaganda Eleitoral
LOCAL OU AMBIÊNCIA	Lead	Onde (o local do acontecimento)	São Paulo	
CAUSA E EFEITO	Lead	Porquê (os motivos e as razões)	Assinatura pequena	

Resumo: PPB/PARTIDO é acusado de crime eleitoral em São Paulo devido à propaganda eleitoral em outdoor possuir assinatura pequena.

Em síntese, o modelo provisório de leitura para indexação de jornais, acima, considerou: a revisão de literatura sobre leitura, leitura documentária, identificação de conceitos e estratégias de leitura; os resultados obtidos nas observações de leitura dos indexadores do arquivo do jornal “O Estado de São Paulo” e do Departamento de Documentação da Editora Abril ; concepção orientada para o conteúdo para análise de assunto (ALBRECHETSEN, 1993); alguns dos procedimentos de elaboração do modelo de leitura para texto científico de Fujita (2000); alguns procedimentos de análise de assunto com a análise conceitual do sistema Precís encontrada em (FUJITA, 1989); estrutura textual do texto jornalístico proposta por Van Dijk (1983), e observações de Silva (2001); mesmo assim, destaca-se que necessita ser melhor elaborado seguindo os procedimentos de análise de conteúdo direcionados ao texto jornalístico.

Evidenciamos que com o desenvolvimento dos capítulos 1, 2, 3 e 5 resgatamos informações primordiais que nos possibilitaram nesse capítulo 6 propor um modelo teórico

de leitura para indexação de jornal e alcançar assim o objetivo maior dessa pesquisa que é *fornecer subsídios teórico-metodológicos que contribuam para a elaboração de um modelo de leitura documentária de artigos de jornal.*

O desenvolvimento do próximo tópico desse capítulo 6 se fez necessário porque realizamos neste a avaliação do modelo de leitura proposto utilizando o protocolo verbal com outros dois grupos de indexadores do DEDOC que não participaram da primeira fase de gravação da leitura de indexadores experientes por serem indexadores inexperientes. Ressaltamos que, na avaliação, um indexador caracterizado como sujeito Clara foi orientado, no momento, sobre o funcionamento do modelo proposto e foi solicitado a realizar a leitura de um texto selecionado fazendo uso do modelo enquanto que outros dois indexadores caracterizados como sujeitos Ana e Maria, foram solicitados a realizarem a leitura sem utilizarem o modelo. Destacamos que os dados da observação da leitura dos dois grupos foram gravados em áudio como recomenda a técnica de Protocolo Verbal e, posteriormente, foram transcritos e analisados considerando os mesmos parâmetros de análise levantados na literatura e, que foram utilizados na primeira fase dessa pesquisa na análise da leitura dos indexadores experientes mencionados aqui no capítulo 6; esse procedimento foi estabelecido com o objetivo de comparar o desempenho da leitura na indexação de determinado texto realizada pelos dois grupos em que ocorreu a situação de leitura onde se faz uso do modelo de leitura proposto, e outra na qual não se utilizou o mencionado modelo.

CAPÍTULO 7: AVALIAÇÃO DO MODELO DE LEITURA PARA INDEXAÇÃO DE JORNAIS: metodologia e resultados

Após ter elaborado o modelo, se faz necessário testar a viabilidade de seu funcionamento para fins de indexação no contexto dos centros referenciais que tratam a notícia e que demandam rapidez no tratamento dessa tipologia.

Sendo assim, neste capítulo, evidenciaremos primeiramente a metodologia adotada para se observar a leitura de dois grupos de indexadores inexperientes do DEDOC. E posteriormente, analisamos e discutimos os dados coletados estabelecendo um parâmetro de comparação do desempenho na indexação do texto selecionado entre o indexador que indexou utilizando o modelo proposto e o par de indexadores que não fez uso do modelo.

Portanto a aplicação do Protocolo Verbal para a observação da atividade de leitura dos indexadores inexperientes no contexto do Departamento de Documentação da Editora Abril obedeceu aos seguintes procedimentos:

1- Procedimentos anteriores à aplicação do Protocolo Verbal:

a) Seleção do Texto-Base

Tipo de texto: Para a seleção do texto-base, escolheram-se textos de estruturas textuais semelhantes, inéditos, ou seja, ainda não indexados por nenhum dos indexadores.

Texto A: foi indexado pelos sujeitos caracterizados como Ana e Maria sem fazer uso do modelo de leitura proposto; foi somente pedido para que indexassem como o faziam em seu dia-a-dia, e que pensassem em voz alta.

Texto B: foi indexado pelo sujeito denominado de Clara. Foi pedido para que o mesmo indexasse o texto conforme o modelo de leitura orientava, com a finalidade de saber se o mesmo havia sido compreendido. Nessa fase, houve interação entre o sujeito e a pesquisadora para esclarecer as dúvidas quanto à compreensão de funcionamento do modelo. Entretanto, o texto classificado como texto B, que foi usado no pré-teste com o objetivo de fornecer orientações ao sujeito Clara para uso do modelo de leitura proposto, era um texto cujo conteúdo o indexador disse já conhecer por ter em dias anteriores a aplicação do protocolo, ter auxiliado, ou seja, ajudado um dos indexadores experientes do DEDOC na sua tarefa de indexação, e fez a leitura e indexação do texto selecionado. Evidenciamos aqui que o fato de o indexador Clara ter tido experiência anterior com o conteúdo do texto foi um ponto positivo porque acreditamos que somente veio facilitar a sua compreensão sobre funcionamento do modelo.

Texto A: foi indexado também pelo sujeito Clara fazendo uso do modelo de leitura sem a intervenção da pesquisadora. Esclarecemos que a aplicação de protocolo com o mesmo texto anteriormente indexado pelo par de indexadores (sujeitos Ana e Maria) foi utilizado na indexação do sujeito Clara, com a finalidade de comparar os resultados obtidos na indexação dos indexadores que não usaram o modelo de leitura, com o do indexador que fez uso do modelo.

Temática do texto: os textos selecionados foram escolhidos verificando o que apresentava mais completamente a estrutura textual do texto jornalístico. Quanto à temática, nesse momento, não teve relevância, porque os indexadores que participaram dessa fase de

gravação do protocolo verbal são tidos no DEDOC como inexperientes na atividade de indexação, isto é indexadores que realizam a tarefa de indexar, esporadicamente, com a função de auxiliar os indexadores experientes e permanentes, quando há acúmulo de textos para indexar, independente, das temáticas que se tem para indexar.

Evidenciamos que para a avaliação do modelo de leitura proposto foi realizada a observação da leitura de dois grupos de indexadores do Centro de Documentação, sendo que um grupo de indexador (caracterizados como sujeito Ana e Maria) indexaram somente um texto sem usarem o modelo de leitura e o outro sujeito, caracterizado como Clara, indexou 2 textos, um em que foi orientado a usar o modelo, e outro em que fez uso do modelo baseando-se nas orientações de uso transmitidas pela pesquisadora, portanto, ressaltamos que também indexou o texto A. Assim, nessa fase de avaliação do modelo foi utilizado um total de dois textos (vide ANEXO E – TEXTO A); (vide ANEXO F – TEXTO B) sendo esses:

- Jornal “O Estado de São Paulo” - Caderno Geral-Saúde: “Cientistas desenvolvem moléculas para superar resistência bacteriana” – texto A indexados pelos indexadores caracterizado como Ana, Maria e Clara;
- Jornal Folha de São Paulo - Seção Mundo: “Racismo contamina Justiça, diz anistia” – Texto B que foi indexado pelo sujeito Clara.

b) **Seleção dos Sujeitos:** A seleção dos sujeitos para avaliação do modelo de leitura proposto considerou, como critério, o pouco tempo de permanência no sistema de informação e na atividade de indexação, acrescentando, ainda, “dificuldade em indexação” refletida na situação de que esses indexadores são esporádicos na realização da tarefa porque, segundo informações verbais obtidas pelo responsável do setor, esses não possuem habilidade, experiência na atividade, enfim, encontram dificuldades na realização da tarefa. Dessa forma, foram selecionados, nessa fase da pesquisa, os 3 únicos indexadores inexperientes da instituição que foram classificados como sujeitos Ana, Maria (sem uso do modelo de leitura), sujeito Clara (que recebeu orientações e no segundo momento fez uso do modelo para indexar).

c) **Conversa informal com os sujeitos :** foi realizada uma conversa informal com cada um dos sujeitos em que se mencionaram 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-los

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.

d) Familiarização com a tarefa do “Think Aloud” (“Pensar Alto”) através do texto “Instruções aos sujeitos”. Antes da aplicação do protocolo verbal foi realizada uma familiarização com a tarefa de “Pensar Alto” por meio da leitura do texto que contém “Instruções aos Sujeitos” (ANEXO C) fundamentado em NARDI (1993), tendo o propósito de descontrair e, ao mesmo tempo, apresentar procedimentos que podem auxiliar os sujeitos no desempenho da tarefa.

2- Procedimentos durante a aplicação do protocolo verbal

e) Gravação do “Pensar Alto” durante a leitura do texto – base: Antes de iniciar a gravação foi entregue aos sujeitos o texto-base lembrando que é preciso “Pensar Alto” durante toda a leitura e exteriorizar seus processos mentais, procurando esquecer a presença da pesquisadora que está presente somente para controlar o gravador e lembrar que é preciso “Pensar Alto”. Destacamos que a aplicação do protocolo verbal também nessa fase da pesquisa, foi realizada individualmente com cada um dos indexadores bem como, a conversa informal, familiarização com a realização da tarefa e a leitura do texto “Instruções aos sujeitos”.

f) Depoimento sobre o modelo de leitura

- Foi pedido ao indexador Clara para que apresentasse suas breves considerações sobre o modelo de leitura, a fim de que, se preciso, o mesmo pudesse ser aperfeiçoado.

“ O modelo é bom e me ajudou muito a compreender o texto por meio dessas perguntas, assim fica mais fácil tirar as palavras-chave.”

3- Procedimentos após a aplicação do Protocolo Verbal

g) Transcrições literais das gravações da leitura dos sujeitos (APÊNDICE G): as transcrições foram feitas de maneira a destacar a compreensão dos sujeitos, suas dúvidas, equívocos, identificação e seleção de termos. E para melhor visualização dos processos adotados pelos sujeitos, destacamos cada um deles com um tipo de notação específica como:

[REDACTED]
Itálico: comentários do leitor
[REDACTED]

..PS.. : pausa
[REDACTED]

Como exemplo do uso de algumas notações (grifadas abaixo) e de sua função para uma análise mais facilitada da observação da atividade de leitura, destacamos um trecho da transcrição de um sujeito de um dos estudos de casos:

At a gente eu pelo menos costume ler a matéria toda. [REDACTED]

(PS) [REDACTED]

(PS) [REDACTED]

Remédio, Experiência não tem eu vou procurar mais então eu procuro termos próximos porque ..PS.. [REDACTED].

Evidenciamos agora a análise dos dados coletados com o Protocolo verbal buscando evidenciar na leitura dos indexadores os aspectos da natureza metacognitiva elaborada por Brown (1980) e reformuladas por Kato (1986) anteriormente utilizadas aqui na observação da leitura dos indexadores na primeira fase dessa pesquisa situada na páginas de revisão de literatura.

Primeiramente, foi realizada a análise dos dados coletados com o Protocolo Verbal durante a leitura dos indexadores (sujeitos Ana e Maria sem uso do modelo) e do (sujeito Clara que fez uso do modelo) procurando identificar nas transcrições dos textos selecionados a compreensão dos indexadores comparando o grau de qualidade na indexação do grupo que não usou o modelo e do sujeito que fez uso do modelo de leitura proposto. Durante a leitura, estabelecemos um parâmetro de análise que considerou os seguintes aspectos:

- A evidência de estratégias metacognitivas listadas por Brown (1980) e reformuladas por Kato (1987);
- A natureza normativa de certas estratégias: algumas recomendações da Norma ISO (5963 de 1985) direcionando-as ao texto jornalístico;
- Outras estratégias do comportamento do indexador durante a leitura (identificação e seleção de termos).

Os parâmetros se encontram organizados e visualizados na tabela 2 já mencionada anteriormente no capítulo 6.

Análise dos dados do Sujeito Ana - coletados através da Técnica do Protocolo Verbal, durante a leitura do texto A.

1. Explicitação dos objetivos da leitura

O sujeito durante toda a leitura não mencionou em nenhum momento o objetivo da leitura.

2. Identificação de (e alocamento de atenção a) aspectos importantes da mensagem fazendo uso da estrutura textual.

Logo após identificar um termo e, posteriormente, comprovar que o mesmo não é encontrado na linguagem, o indexador indaga que procurará, ou consultará na Linguagem termos próximos do anteriormente selecionado, demonstrando ter certeza que o assunto é da área identificada restando, somente, ver como a linguagem apresenta a hierarquia e forma de entrada do assunto nessa determinada área.

... Remédio, Experiência não tem, eu vou procurar mais, então eu procuro termos próximos porque ..PS.. . (Nesse momento o sujeito digita novos termos **Experiência Científica**, ...

3. Monitoração do comportamento para ver se está ocorrendo compreensão: Engajamento em revisão e auto-indagação para ver se o objetivo está sendo atingido (a)/ Tomada de ações corretivas quando são detectadas falhas na compreensão (b).

O sujeito não explicitou o objetivo da leitura não tendo, assim, que verificar se o objetivo foi ou não atingido. E também não cometeu falhas a serem corrigidas.

4. Recobrimento de atenção quando a mente se distrai ou faz digressões.

Não evidenciou esse tipo de comportamento porque não encontrou problema de compreensão.

Análise dos dados do Sujeito Maria - coletados através da Técnica do Protocolo Verbal, durante a leitura do texto A.

1. Explicitação dos objetivos da leitura

O sujeito durante toda a leitura não mencionou em nenhum momento o objetivo da leitura.

2. Identificação de (e alocamento de atenção a) aspectos importantes da mensagem fazendo uso da estrutura textual.

O sujeito menciona conhecer a estrutura textual externa do artigo de jornal, quando menciona que, pelo fato de a matéria ter sido publicada em editoria diferente da área à que pertence, pode ser difícil indexar porque como o conteúdo informacional da matéria está resumido, podem faltar informações relevantes, podendo prejudicar a compreensão do texto.

É bom a primeira coisa que eu olho aqui é o chapéu né que é o indica a editoria na verdade aqui ele está fora da editoria dele. É difícil uma boa coisa porque ela está fora do caderno dela da área ...

3. Monitoração do comportamento para ver se está ocorrendo compreensão: Engajamento em revisão e auto-indagação para ver se o objetivo está sendo atingido (a)/ Tomada de ações corretivas quando são detectadas falhas na compreensão (b).

O sujeito não explicitou o objetivo da leitura não tendo, assim, que verificar se o objetivo foi ou não atingido. E também não cometeu falhas a serem corrigidas.

4. Recobrimento de atenção quando a mente se distrai ou faz digressões.

Não evidenciou esse tipo de comportamento porque não encontrou problema de compreensão.

Análise dos dados do Sujeito Clara - coletados através da Técnica do Protocolo Verbal, durante a leitura dos textos A.

1. Explicitação dos objetivos da leitura

O sujeito menciona, no momento de extração dos termos, que o objetivo da leitura é a recuperação pelos usuários, ou seja, atribuir à matéria termos, ou possibilidades de buscas que possibilitem ao usuário recuperar a matéria seja, por termos, ou nome de pasta de assunto, etc.

*Agora eu só tenho que colocar nesse caso da **bactéria** eu nem colocaria termo porque já tá no nome da pasta já vem como termo. Bactéria já funciona como termo e a pesquisa é por palavra livre, se o cara pesquisar o nome do instituto ele vai, por exemplo, no sistema e digita Instituto de pesquisa Scripps, ele recupera a matéria.*

2. Identificação de (e alocamento de atenção a) aspectos importantes da mensagem fazendo uso da estrutura textual.

O sujeito não menciona ter conhecimento da estrutura do texto jornalístico.

3. Monitoração do comportamento para ver se está ocorrendo compreensão: Engajamento em revisão e auto-indagação para ver se o objetivo está sendo atingido (a)/ Tomada de ações corretivas quando são detectadas falhas na compreensão (b).

O sujeito não explicitou o objetivo da leitura não tendo, assim, que verificar se o objetivo foi ou não atingido. E também não cometeu falhas a serem corrigidas.

4. Recobrimento de atenção quando a mente se distrai ou faz digressões.

O sujeito se confunde, um pouco, no momento de encontrar resposta ao questionamento do modelo e corrigiu sua compreensão, imediatamente.

... Local: bactéria eu acho não local no caso seria ah! Tá a cidade La Jolla/Califórnia ...

Discutiremos agora os resultados da observação da leitura do indexador, destacando novamente que, neste estudo utilizou-se a técnica do Protocolo Verbal, e foi possível verificarmos a compreensão dos indexadores do Centro de Documentação da Editora Abril (caracterizados como sujeitos Ana, Maria e Clara), durante a atividade de indexação de jornais porque utilizamos um parâmetro de análise que nos permitiu identificar os aspectos metacognitivos listados por Brown (1980), reformulados por Kato (1986), as recomendações de identificação de conceitos da Norma ISO 5963 adaptadas ao texto jornalístico, e a exploração da estrutura textual de texto jornalístico proposto por Van Dijk (1983), e estabelecendo um parâmetro de comparação entre o indexador que indexou usando o modelo de leitura proposto, e o par de indexadores que não usou o modelo de leitura. Os aspectos que compõem o parâmetro, são os mais representativos para a leitura com fins documentários; foram citados no capítulo 6.

Os quadros apresentados, a seguir, procuram evidenciar os aspectos notados na leitura.

a) Seqüência de operações (aspectos) da leitura dos sujeitos

A verificação de procedimentos da leitura dos indexadores de jornais abordou, primeiramente, a sequência de operações, sendo que essas operações são os aspectos da natureza metacognitiva da leitura, listados por Brown (1980) e reformulados por Kato (1986). Para isso, sintetizamos a sequência de operações obtidas a partir da análise dos dados coletados. O quadro 20, a seguir, demonstra a sequência de operações realizadas pelos 3 sujeitos (indexadores) durante a leitura, criando, também, um parâmetro de comparação entre a sequência da leitura de cada indexador, buscando verificar indícios da presença de um padrão de leitura, e comparando o indexador que usou o modelo de leitura e o par de indexadores que não fez uso do modelo.

Quadro 20: Análise das operações (atividades) desenvolvidas pelos sujeitos Ana, Maria e Clara (indexadores do DEDOC) durante a leitura do texto A:

Sujeito Ana – Texto A (sem uso do modelo)

Identificação de aspectos importantes da mensagem

Seleção de conceitos para indexação

Sujeito Maria – Texto A (Sem uso do modelo)

Identificação de aspectos importantes da mensagem

Seleção de conceitos para a indexação

Sujeito Clara – Texto A (usando o modelo)

Seleção de conceitos para indexação

Explicitação dos objetivos da leitura

Analisando os quadros de sequência das operações notamos que:

- Os sujeitos usam algumas das atividades metacognitivas listadas por Brown (1980);
- destacamos que os sujeitos Ana e Maria não fizeram uso do modelo de leitura proposto;
- sujeito Ana realiza identificação de aspectos importantes da mensagem e seleção de conceitos;
- sujeito Maria também usa as atividades identificação de aspectos importantes da mensagem e seleção de conceitos;
- sujeito Clara, realiza atividades metacognitivas como: explicitação dos objetivos da leitura, seleção de conceitos. Notou-se que os outros indexadores não explicitaram os objetivos da leitura;

- sujeito Maria menciona a necessidade de ler o texto todo o que é comum entre indexadores inexperientes na atividade de indexação;
- o sujeito Ana comenta a importância de conhecer a hierarquia dos conceitos dentro de uma área, e a organização dessa hierarquia dentro da linguagem de indexação; percebemos quando o indexador comenta “deixa eu procurar termos mais próximos...”;
- o sujeito Maria destaca como importante a estrutura textual superficial do jornal que é a organização entre os cadernos que compõem o jornal. Isso foi percebido quando o sujeito mencionou “essa matéria está fora da área dela não teve trazer algo completo” deve ser descrita sucintamente alguma informação relevante pode faltar;
- evidenciamos que a leitura de todos os textos centrou-se, basicamente, no “Lead” da estrutura textual de jornal para a seleção de conceitos pertinentes. Entretanto, os indexadores não mencionaram conhecer a estrutura textual proposta por Van Dijk (1983);
- sujeito Clara, no texto A, fazendo uso do modelo de leitura proposto, mencionou que o objetivo da leitura é indexar considerando o conteúdo do texto e as necessidades informacionais dos usuários; percebeu-se isso porque o indexador procura prever as formas de buscas dos usuários, ou seja, como pode ser recuperado o texto;
- percebeu-se que os três indexadores, ao contrário dos indexadores experientes cujos protocolos verbais foram analisados na primeira fase desta pesquisa, não realizaram explicitamente monitoração do comportamento para verificar se estava ocorrendo a compreensão.

c) Manter em mente o seu objetivo de representar o texto para futura recuperação

- O sujeito Clara menciona o objetivo da leitura que é indexar representando o conteúdo e os interesses dos usuários;
- Enquanto os sujeitos Ana e Maria, deixam transparecer que indexam considerando unicamente o conteúdo do texto. Notamos que essa problemática de identificar, considerando mais o conteúdo que os interesses dos usuários, ocorre porque os indexadores do DEDOC expressam que o sistema de recuperação da instituição possibilita recuperar o documento por vários pontos de acesso como: linguagem livre de qualquer palavra que compõe o texto, nome do jornal podendo ser ou não acompanhado de data, página, título ou qualquer palavra que compõe o título, autor,

nome da pasta de assunto para a qual é enviada a matéria. Essas N possibilidades de garantir a recuperação fazem com que os indexadores não se preocupem em atribuir muitas palavras-chaves que representem o texto, como de vez enquando, deixar de atribuir qualquer palavra-chave, pelo fato de o título já compor essa palavra. A meu ver, são preocupantes a questão do tempo na recuperação e a real certeza da capacidade de memória do usuário em saber, no momento da busca, o título exato, autor, e se isso não restringe a recuperação em termos de quantidade reduzida de matérias sobre o assunto.

- notou-se que somente o sujeito Clara, explicitou o objetivo da leitura, ou seja, que a mesma está direcionada para a representação fidedigna do conteúdo do texto como também para a qualidade na recuperação, considerando as formas de buscas da comunidade usuária. Mas, ressalta-se que, o sujeito Maria comenta que se preocupa em ler o texto todo deixando talvez implícito que indexa considerando o conteúdo do texto.

c) Exploração de seu conhecimento de estrutura textual

O conhecimento e a exploração da estrutura textual pelos indexadores são considerados muito importantes para a leitura documentária, pois proporcionam ao indexador domínio das partes que compõem o texto e do contexto de cada parte, para reconhecimento dos conceitos pertinentes que se pretende representar. A análise dos dados coletados indicou a exploração da estrutura textual seguida pelos 3 sujeitos (indexadores), destacada nos quadros seguintes.

Quadro 21: Seqüência da estrutura textual seguida pelos sujeitos Ana, Maria e Clara:

Sujeito Ana – Texto A
Cabeçalho
Lead
Episódio (atos da fala)
Sujeito Maria – Texto A
Cabeçalho
Lead
Sujeito Clara – Texto A
Cabeçalho
Lead

- Ana iniciou a leitura explorando o elemento Cabeçalho da estrutura textual de jornal onde se localiza (título, subtítulo, nome do jornal, seção, página, data, colaborador (autor da matéria), para preencher a ficha de indexação, e essa é uma atividade desenvolvida por todos os indexadores. Posteriormente, o indexador continua a leitura explorando o Lead da estrutura textual de jornal e concentra a seleção de conceitos, durante a indexação, nessa parte da estrutura textual apesar de ser notado que a leitura do texto por esse indexador se estende até o elemento Episódios (atos da fala) da estrutura textual de jornal;
- Maria inicia leitura também pelo Cabeçalho para preenchimento da ficha de indexação. E, na seqüência, estende a leitura até o Lead e extrai destes os conceitos para a indexação do texto;
- Clara, inicia a leitura pelo Cabeçalho da estrutura textual de jornal para preencher a ficha de indexação. Na continuidade, explora o Lead da estrutura textual de jornal e extrai destes conceitos para representar a matéria que indexa.
- Percebemos que os elementos mais explorados da estrutura textual de jornal foram o *Cabeçalho* e o *Lead*. Nas páginas anteriores, com os resultados das outras observações de leitura dos indexadores experientes, supúnhamos que o modelo de leitura que

propúnhamos deveria considerar que a leitura, para fins de indexação de jornal, deve ser direcionada para exploração até o elemento *Episódio (atos da fala)*, nos casos em que a leitura do *Lead* não fosse suficiente. Diante dessa nova análise dos dados coletados, usando o modelo de leitura proposto, mencionamos que a leitura para fins de indexação de jornal deve considerar a extensão da leitura até o elemento Comentários (expectativas) da estrutura textual de jornal, porque o restante dos itens da estrutura textual de jornal irá tratar de opiniões do jornalista, ou do jornal, não sendo relevantes para fins de indexação. (VAN DIJK, 1983).

d) identificação de aspectos importantes da mensagem para identificação de conceitos

A identificação e seleção de conceitos são, tal como a exploração da estrutura textual, uma importante estratégia de leitura. Em nosso entendimento, é a estratégia vital para atingir o objetivo de representação para recuperação, pois é com esta estratégia que os conceitos são identificados.

Com essa perspectiva em mente, três aspectos da estratégia de identificação serão abordados e verificados na análise dos dados da entrevista:

- d1) em que parte da estrutura textual foram identificados os conceitos;
- d2) qual a sistemática utilizada para a identificação de conceitos e
- d3) identificação e seleção de termos

Para uma melhor visualização dos resultados abordando o procedimento de identificação de conceitos e partes da estrutura textual, foram construídos os quadros apresentados abaixo:

d1) Parte da Estrutura Textual em que foram Identificados os Conceitos

Quadro 22: Identificação de conceitos e Partes do texto - Sujeito Ana

Trechos em que o sujeito identificou os Termos	Parte do Texto
<i>Remédio, Experiência não tem eu vou procurar mais então eu procuro Termos próximos porque..PS..</i>	Lead
(Nesse momento o sujeito digita novos Termos Experiência Científica)	Lead
Instituto de pesquisa Scripps, <i>vou verificar se tem o nome de desse instituto de pesquisa</i> (Nesse momento o sujeito digita o nome do instituto). <i>Ah, não tem então eu vou anotar o nome todo.</i> (Nesse momento digita o instituto como termo) Instituto de Pesquisa Scripps	Lead

Quadro 23: Identificação de conceitos e partes do texto - SUJEITO Maria

Trechos em que o sujeito identificou os Termos	Parte do Texto
<i>Aí então eu colocaria como palavra-chave somente infecção hospitalar</i>	Lead
<i>e também deixa eu ver aqui</i> (Nesse momento o indexador verifica no sistema se existe a palavra-chave medicamento-antibiótico e confirma que existe), então eu coloco medicamento-antibiótico <i>porque o central aqui é o desenvolvimento do antibiótico mesmo, né. Ele é voltado para as bactérias que causam infecção hospitalar eu colocaria isso aí mesmo</i>	Lead

Quadro 24: Identificação de conceitos e partes do texto - SUJEITO Clara

Trechos em que o sujeito identificou os Termos	Parte do Texto
<i>Mas vamos ver com seu roteiro se consigo colocar palavras-chave. Ação: é combater bactéria, Agente: Instituto de pesquisa Scripps, Métodos</i>	Lead
<i>Métodos: <= Criar uma nova classe de moléculas capazes de fazer buracos na parede celular de bactérias,</i>	Lead
<i>Local: bactéria eu acho não local no caso seria ah! Tá a cidade La Jolla/Califórnia</i>	Lead
<i>Causa efeito: é aniquilar bactérias resistentes que causam infecção hospitalar,</i>	Lead

- Notou-se que, na indexação dos três indexadores quase todos os conceitos foram selecionados explorando o Lead.
- Sujeito Ana identificou conceitos no Lead, sem usar o modelo de leitura;
- Sujeito Maria também selecionou os conceitos do Lead, sem usar o modelo de leitura;
- O sujeito Clara, explora unicamente o Lead na indexação do texto e, fazendo uso do modelo de leitura proposto, extrai maior número de conceitos, a nosso ver, mais pertinentes, da representação do conteúdo do texto que os outros dois sujeitos Ana e Maria que não fizeram uso do modelo para indexar o mesmo texto;

d2) Sistemática para identificação de conceitos

A sistemática de identificação foi observada utilizando as transcrições dos textos e a análise dos quadros (18, 19 20, 21 e 22). Observou-se que os sujeitos evidenciaram identificação de conceitos utilizando cinco tipos de comportamentos: destaque de palavras do texto por meio de argolamento, ou digitação; associações com linguagem após reconhecimento de termos; questionamento durante a leitura quanto à escolha dos termos; conhecimento de estrutura textual; e conhecimento prévio, e recorre a memória à longo termo.

- **destaque de palavras do texto** (“argolamento”, ou digitação). Essa atividade não foi expressa verbalmente, a pesquisadora observou e anotou.

Como exemplo temos a seguinte transcrição:

Então vem na tela e costuma digitar o título a gente recorta e cola nesse caso Cientistas desenvolvem moléculas para superar resistência bacteriana. No caso o autor é Herton Escobar a gente marca da mesma maneira. A publicação é o Estado de São Paulo a gente escreve, a data 26, de julho de 2001, aí a pasta ficaria assim Biologia/Bactéria. Então Biologia/Bactéria seria a pasta. Então essa matéria vai para Biologia/Bactéria. Ai seriam as palavras chaves. (Sujeito Ana - Texto A)

- **fazendo associações com Linguagem após reconhecimento de termos:**

Exemplos:

*Sujeito Ana: Remédio, Experiência não tem, eu vou procurar, mais então, eu procuro termos próximos porque ..PS.. . (Nesse momento o sujeito digita novos termos **Experiência Científica***

- **Realiza questionamento durante a Leitura do texto quanto às escolhas dos descritores:**

Exemplos:

Sujeito Ana

*Remédio, Experiência não tem, eu vou procurar, mais então, eu procuro termos próximos porque ..PS.. . (Nesse momento o sujeito digita novos termos **Experiência Científica...***

Sujeito Maria

Os cientistas do Instituto de Pesquisa Scripps, em La Jolla ..PS.. . É, talvez eu coloque o instituto mas não, eu acho que não é importante o instituto mas vou ter que consultar lá (Nesse momento o indexador consulta o sistema para verificar se já existe o nome do instituto como termo controlado de indexação e confirma que não existe o nome do instituto)...

*Aí, então, eu colocaria como palavra-chave somente **infecção hospitalar** e também, deixa eu ver aqui, (Nesse momento o indexador verifica no sistema se existe a palavra-chave medicamento-antibiótico e confirma que existe) então eu coloco **medicamento-antibiótico** porque o central aqui é o desenvolvimento do antibiótico mesmo, né. Ele é voltado para as bactérias que causam infecção hospitalar, eu colocaria isso aí mesmo*

- **Conhecimento de estrutura textual**

Exemplos:

Sujeito Maria

É, bom, a primeira coisa que eu olho aqui é o chapéu né, que é o que indica a editoria, na verdade, aqui, ele está fora da editoria dele. É difícil uma boa coisa porque ela está fora do caderno dela, da área

Os indexadores, no geral, demonstram ficar muito presos à Linguagem, temem criar novos termos se o mesmos não constam da Linguagem do sistema, mesmo tendo consciência que estes representam o conteúdo do texto.

Os indexadores não evidenciam conhecer a estrutura textual proposta por Van Dijk (1983). Notou-se conhecimento da estrutura superficial que compõe os cadernos do jornal, por parte do sujeito Maria no texto A.

d3) identificação e seleção de termos

Apresentamos abaixo quadros referentes aos termos identificados e selecionados esclarecendo que, durante a leitura, o indexador realiza em um único momento, primeiramente, a identificação, e depois, a seleção dos termos, sendo esta última confrontada com a representação do conteúdo do texto e os interesses informacionais dos usuários. Como exemplo, destacamos um trecho de uma transcrição:

Agora eu só tenho que colocar nesse caso da bactéria eu nem colocaria termo porque já tá no nome da pasta já vem como termo. Bactéria já funciona como termo e a

pesquisa é por palavra livre se o cara pesquisar o nome do instituto ele vai por exemplo no sistema e digita Instituto de pesquisa Scripps ele recupera a matéria. Esse aqui é difícil e como se pode recuperar por outras maneiras sem ser somente palavra chave eu salvaria sem termos

Quadro 25: Termos identificados e selecionados - Sujeito Ana - texto A

Sujeito Ana – Texto A (sem uso do modelo)	
Termos Identificados	Termos Selecionados
Remédio	Experiência
Experiência	Científica
Experiência	Instituto de Pesquisa
Científica	Scripps
Instituto de Pesquisa	
Scripps	

Quadro 26: Termos identificados e selecionados - Sujeito Maria - texto A

Sujeito Maria – Texto A (sem uso do modelo)	
Termos Identificados	Termos Selecionados
Medicamento	Infecção hospitalar
Instituto de Pesquisa Scripps	Medicamento- antibiótico
Infecção hospitalar	
Medicamento- antibiótico	

Quadro 27: Termos identificados e selecionados - Sujeito Clara - texto A

Sujeito Clara – Texto A (usando o modelo)	
Termos Identificados	Termos Selecionados
Infecção hospitalar	Infecção hospitalar
Bactéria	Bactéria
Instituto de pesquisa Scripps	Instituto de pesquisa Scripps
Moléculas	Molécula
La Jolla/Califórnia	La Jolla/Califórnia
Infecção hospitalar	Infecção hospitalar

- Sujeito Ana identificou 4 termos e selecionou 2. Deixa transparecer que optou por um termo mais específico que representasse melhor o conteúdo, mas não menciona explicitamente isso, somente deixa claro a questão de ter ou não ter na Linguagem;
- Sujeito Maria identificou 4 termos e selecionou, após consultar a Linguagem, os mais específicos, além de atribuir o nome da pasta como um termo palavra-chave;
- Sujeito Clara, é capaz de identificar e selecionar todos os termos que representa o conteúdo do texto. Impressionante mencionar aqui que o sujeito Clara, durante a indexação, indaga que o texto é difícil, se julga incapaz de extrair conceitos do mesmo, mas com o uso do modelo consegue extrair os termos para o texto.

e) Compatibilidade dos termos selecionados com a Linguagem do Sistema

A análise da compatibilidade dos termos com a linguagem de indexação, considerando os quadros, acima, foi realizada procurando evidenciar se a linguagem possuía os termos selecionados pelos sujeitos durante a leitura.

Para melhor analisar a compatibilidade dos termos com as Linguagens foram elaborados os quadros abaixo:

Quadro 28: Termos identificados e selecionados pelos sujeitos Ana, Maria e Clara no texto A comparados com a linguagem do DEDOC

Sujeito Ana – Texto A (termos identificados)			
Termos Identificados		Manual	do
		DEDOC	
Experiência Científica		Tem	
Instituto de Pesquisa		Não tem	
Scripps			
Sujeito Maria – Texto A (termos identificados)			
Termos Identificados		Manual	do
		DEDOC	
Infecção hospitalar		Tem	
Medicamento- antibiótico		Não Tem	
Sujeito Clara – texto A (termos identificados)			
Termos Identificados		Manual	do
		DEDOC	
Bactéria		Tem	
Instituto de Pesquisa		Não Tem	
Scripps			
Molécula		Tem	
Infecção hospitalar		Tem	
La Jolla/Califórnia		Não Tem	

- Sujeito Ana, na leitura do texto A, identificou e selecionou 2 termos e a Linguagem tinha somente 1 termo;
- Sujeito Maria selecionou 2 termos e a Linguagem tinha somente um dos dois termos;
- Sujeito Clara, usando o modelo de leitura proposto, selecionou 5 termos para representar o texto A, enquanto os outros indexadores, que não usaram o modelo para indexar esse mesmo texto, extraíram, no máximo, 2 termos. E dos 5 termos selecionados, a linguagem possuía 3, deixando de apresentar somente um termo referente ao nome de uma

instituição e o local geográfico do acontecimento e que o indexador incluiu como termo novo na linguagem, com a seleção desses termos durante a indexação do texto.

Percebemos que a linguagem do DEDOC é compatível, apresenta a maioria dos termos selecionados como representantes do conteúdo dos textos indexados.

Os indexadores que não fizeram uso do modelo demonstram dificuldade no momento de explorar o conteúdo do texto com o objetivo de identificar conceitos, e essa é camuflada pelas várias formas de acesso ao texto oferecidas pelo sistema de recuperação do DEDOC. Isso, provavelmente, ocorre porque, segundo os indexadores do DEDOC incluindo aqui as falas descritas nos protocolos, tanto do indexador sujeito Clara que fez uso do modelo como as falas dos trechos dos protocolos de alguns indexadores experientes que participaram da primeira fase dessa pesquisa, que mencionam que as necessidades informacionais dos usuários são satisfeitas por essas várias possibilidades de formas de buscas do texto, ficando a atribuição de termos ao texto, em muitos casos, principalmente naqueles nos quais os indexadores encontram dificuldades em compreender o tema, em segundo plano, porque já está garantida a recuperação do texto por qualquer palavra do título, do texto, autor, nome do jornal, nome da pasta de assunto na qual a matéria é inserida. Portanto, é relevante para o DEDOC verificar o grau de satisfação da sua comunidade usuária com os resultados obtidos, com as buscas realizadas no acervo. Porque, pode-se ter N maneiras de se recuperar um texto, mas tem que se verificar se essa política de indexação é favorável a quem se destina, no caso, aos usuários do DEDOC, e se a recuperação é feita de forma rápida, porque, hoje, o que se almeja é informação em tempo real, ou no mais curto espaço de tempo possível porque *“pensar cansa e três minutos é muito tempo para se lançar a fazer buscas que não são simples tendo apenas uma vaga idéia da estratégia.* (HOUSMAN, A. E. citado por GROGAN, D. 1995).

Evidenciamos que, usando o modelo, Clara indaga que, no dia-a-dia, deixaria de atribuir termos à matéria logo após ter mencionado que se tratava de um texto difícil de se compreender. É relevante mencionar que, fazendo uso do modelo de leitura proposto, Clara consegue identificar conceitos mais pertinentes para representar o conteúdo do texto A, que antes julgava incapaz de compreender, se comparados com os conceitos selecionados pelos sujeitos Ana e Maria que não usaram o modelo de leitura.

Destacando a diferença da compreensão dos sujeitos com relação ao texto A, fazendo uso e não do modelo, exibimos os dados obtidos considerando unicamente os termos selecionados por ambos, nos quadros abaixo:

Quadro 29: 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	Porque (os motivos e as razões)		

Resumo: O Instituto de Pesquisa Scripps realizou uma experiência científica.

Percebemos que o sujeito Ana, com os conceitos que atribuiu ao texto, conseguiria 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 30: 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)		
AGENTE	Cabeçalho e Lead	Quem (os agentes da acção)		
MÉTODOS DO AGENTE	Lead	Como (as circunstâncias)	-Medicamento-antibiótico	
LOCAL OU AMBIÊNCIA	Lead	Onde (o local do acontecimento)		
CAUSA E EFEITO	Lead	Porque (os motivos e as razões)	-Infecção hospitalar	-Infecção hospitalar

Resumo: Vai se desenvolver um medicamento-antibiótico para se combater infecção hospitalar.

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

Quadro 31: 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	Lead	Onde (o local do acontecimento)	La Jolla/California	
CAUSA, E EFEITO	Lead	Porque (os motivos e as razões)	-infecção hospitalar	-Infecção hospitalar

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

Nota-se que, fazendo uso das orientações do modelo de leitura proposto, os conceitos selecionados pelo sujeito Clara, para o texto A, possibilita organizar logicamente uma representação mais fidedigna do tema global comunicado pelo texto do que os conceitos selecionados pelos indexadores que não fizeram uso do modelo de leitura. E ainda, o mais relevante, é que o uso do modelo na indexação, como pode se perceber, comparando os termos selecionados pelos dois grupos de indexadores, nota-se que a indexação do sujeito Clara, para o texto A fazendo uso do modelo, teve, como consequência, uma indexação de boa qualidade porque abrangeu todas as possibilidades de termos para recuperação do texto.

Evidenciamos que a utilização do modelo de leitura proposto para indexação de jornais, obteve êxito em seu objetivo de auxiliar o indexador de jornais ajudando-o a compreender o texto a indexar, por meio do questionamento ao texto, possibilitando-o realizar uma leitura mais rápida porque o questionamento é simples mnemônico, ou seja, de fácil memorização e a exploração dos elementos da estrutura textual de jornal para encontrar as respostas ao questionamento é de fácil identificação.

É relevante mencionar que o modelo de leitura proposto enfatizou a concepção orientada para o conteúdo, e isso é percebido pela análise dos dados do sujeito Clara na

indexação do texto. A o sujeito utilizando o modelo compreendeu o texto e o alto grau de sua compreensão foi expressada nos conceitos pertinentes que selecionou como representantes do conteúdo do texto. Portanto o modelo de leitura para indexação de jornais fornece estratégias adequadas para se realizar uma leitura rápida e compreensiva com fins de indexação, mas, não foi idealizado para acoplar as concepções orientadas para o conteúdo e para a demanda que é primordial para os objetivo da indexação que é fornecer pontos de acesso que possibilite a recuperação precisa de documento para a comunidade usuária a qual o mesmo se destina.

No próximo capítulo, apresentamos as conclusões finais da pesquisa e efetuamos algumas adequações ao modelo e realizamos sugestões ao DEDOC considerando os dados levantados na discussão final dos resultados levantados com a avaliação do modelo de leitura proposto.

CAPÍTULO 8: CONSIDERAÇÕES FINAIS

A atividade de leitura documentária é a fase primordial da análise de conteúdo do documento para fim de indexação, porque a eficácia das outras atividades (elaboração de índices e resumos) depende da qualidade da leitura, ou seja, da compreensão que o indexador tiver da leitura de determinado texto. Essa leitura difere da leitura global porque exige do leitor-indexador uma leitura rápida na qual não se recomenda realizar do começo ao fim do texto um processo linear onde se lê letra por letra, palavra por palavra, tendo em vista que a quantidade de documentos que os sistemas de informações possuem para indexar, é muito grande exigindo, assim, uma leitura estratégica e rápida.

Partindo desse princípio, identificou-se a relevância de estar observando na prática, a leitura dos indexadores de jornais, uma vez que supúnhamos que estes realizavam uma leitura estratégica e rápida devido à grande quantidade de material a indexar e que, para isso, utilizavam estratégias de leitura. Baseando-se nessas suposições, interessou investigar se essas estratégias eram as mesmas estabelecidas pela literatura sobre o assunto. Para verificarmos as suposições levantadas, foi realizado estudo de literatura sobre: leitura e leitura documentária; estratégias de leitura; estrutura textual na compreensão e apreensão do conteúdo do texto; identificação de conceitos na indexação e as metodologias existentes; além de estudo prático de observação de leitura de indexadores de jornais no contexto do Arquivo do jornal “O Estado de São Paulo” e no Departamento de Documentação da Editora Abril, nos quais foi utilizada a técnica do Protocolo Verbal. Portanto, nesta pesquisa de mestrado sobre leitura documentária de artigos de jornais, procurou-se investigar o que a teoria apresentada na revisão de literatura da área recomendava como procedimentos de leitura para análise de conteúdo dos documentos, no caso, artigos de jornais impressos, estabelecendo um parâmetro de comparação com o que realmente acontece na prática da leitura de artigos de jornais do dia-a-dia dos sistemas de informações que tratam a notícia – visando, com essa observação, alcançar três objetivos: *1- proporcionar aos indexadores, subsídios metodológicos para a análise de assunto de textos com estrutura diversificada; 2- identificar estratégias de leitura para artigos de jornal e 3- fornecer subsídios teórico-metodológicos que contribuam para a elaboração de um modelo de leitura documentária de artigos de jornal.*

Os dados levantados na observação prática de leitura foram cruzados com os resultados do estudo de literatura e nos permitiram verificar as suposições levantadas, identificando que os indexadores de jornal são rápidos porque têm conhecimento prévio

sobre o assunto do texto e, também, porque a estrutura textual de jornal facilita a realização de uma leitura rápida, uma vez que apresenta logo no início do texto um resumo do conteúdo que será tratado no texto, denominado “Lead”.

Quanto à outra suposição, percebemos que os indexadores observados utilizam determinadas estratégias de leitura que são também estabelecidas pela revisão de literatura, como exploração da estrutura textual, durante a leitura, e uso do conhecimento prévio sobre o assunto.

Evidenciamos que a revisão de literatura permitiu o resgate de obras relevantes sobre o assunto e, conjuntamente com a observação da prática de leitura dos indexadores experientes do Arquivo do jornal “O Estado de São Paulo” e do Centro de Documentação da Editora Abril Cultural, possibilitou atingir os três objetivos propostos mencionados acima. O objetivo de *identificar estratégias de leitura para artigos de jornal* foi atingido por meio do levantamento de estratégias de leitura na revisão de literatura, como: as atividades metacognitivas estabelecidas por Brown (1980); as estratégias de leitura, ou outras denominações elaboradas por Cintra (1987); Medeiros (1997); Kleiman (1989); Kato (1986); Ahmad (1993); Farrow (1996); as sistemáticas de identificação de conceitos utilizadas pelos indexadores, entre outras estratégias.

O segundo objetivo, *proporcionar aos indexadores, subsídios metodológicos para a análise de assunto de textos com estrutura diversificada*, foi alcançado com a revisão de literatura em que resgatamos metodologias de análise de assunto proposta pelo sistema de indexação PreciS; os Princípios de Indexação do UNISIST (1981); a metodologia de análise do documento da Norma ISO 5963 (1985) - traduzida para a ABNT (12676, 1992); a proposta de análise de assunto de Tálamo (1987) reformuladas por Kobashi (1994) e as dicas de exame do documento propostas por Lancaster (1993). Pela revisão de literatura, observou-se que a estrutura textual é um fator predominante na leitura documentária e que, portanto, o indexador deve conhecer a estrutura textual dos textos com os quais trabalha a fim de realizar uma leitura mais rápida, por meio da exploração de algumas partes que possuem informações relevantes para fins de indexação.

Mencionamos, ainda, as variadas estruturas textuais com as quais o indexador pode se deparar na realização de sua tarefa como: os textos com os quais a Análise Documentária trabalha, citados por Kobashi (1994); a estrutura textual do texto jornalístico, proposta por Van Dijk (1983) e, com algumas modificações quanto aos elementos do LEAD, estabelecida por Silva (2001); outras estruturas textuais mencionadas

por Van Dijk (1992); a estrutura textual do texto científico citado por Galvez & Pinto (s.d.) e por Tálamo (1994).

É relevante destacar que, no capítulo no qual enfocaram-se as estruturas textuais, mencionamos também que o indexador pode fazer uso da Diplomática por meio do conhecimento e utilização do método diplomático elaborado por Guimarães (1998) para identificar as partes mais relevantes da estrutura do documento com que trabalha, independente do suporte e forma de determinado documento.

O terceiro e último objetivo foi de encontrar na revisão de literatura e na observação prática de indexação de jornal, *subsídios teóricos-metodológicos que contribuíssem para a elaboração de um modelo de leitura documentária de artigos de jornal*. Evidenciamos que foi atingido porque os dados coletados na observação da leitura dos indexadores experientes de jornais foram relevantes e significativos e, quando comparados com a literatura sobre o assunto, como: a estrutura textual do texto jornalístico proposta por Van Dijk (1983); com modificações do questionamento do LEAD proposta por Silva (2001); concepção orientada para conteúdo estabelecida por Albrecthsen (1993); metodologia de análise de assunto do sistema de indexação Precís (1989); modelo de leitura para indexação de texto científico elaborado por Fujita (2000); nos possibilitaram elaborar o modelo de leitura para indexação de jornal considerando as estratégias de leitura dos indexadores e as recomendações relevantes da literatura.

O modelo de leitura proposto para indexação de artigos de jornais consiste na identificação de conceitos mediante questionamento, visando localizá-los em determinadas partes da estrutura textual, por meio de interpretação do conteúdo do texto e, posterior verificação de compatibilidade dos conceitos selecionados com a Linguagem do sistema, no caso, DEDOC. Em seguida, a indexação é encerrada com a elaboração de um breve comentário do assunto composto pelos termos selecionados que representam o conteúdo do texto em ordem lógica e permite a compreensão do texto. A elaboração do enunciado visa fornecer uma síntese do conteúdo do texto, que nos textos jornalísticos bem escritos é o próprio Lead, para que o indexador se familiarize com este componente e seus conceitos, e identifique- o mais facilmente durante a leitura.

Após elaborado, o modelo foi testado por um segundo grupo de indexadores do DEDOC que ainda não haviam participado da primeira fase da pesquisa (na qual observamos a leitura dos indexadores experientes visando encontrar procedimentos para elaboração do modelo), por serem indexadores não permanentes na realização da atividade de indexação na instituição, devido ao fato de provavelmente apresentarem dificuldades na

realização da tarefa. Nesse momento, selecionamos três indexadores no sistema, sendo que dois foram solicitados para que indexassem como o fazem no seu dia-a-dia, sem conhecimento e uso do modelo de leitura proposto, e um outro indexador que foi orientado, momentos antes, sobre o funcionamento do modelo de leitura com um outro texto de temática diferente, foi solicitado, agora, a indexar fazendo uso do modelo, utilizando o mesmo texto que os outros dois indexadores trabalharam sem usarem o modelo.

Com as observações do desempenho da leitura de determinado texto, comparando os indexadores que não usaram o modelo com o indexador que fez uso do modelo, ressaltamos que o modelo de leitura é pertinente aos fins de indexação porque proporciona uma leitura mais rápida do texto, por meio da exploração da estrutura textual de jornal que permite uma localização fácil dos elementos que a constitui, por meio de questionamento para identificação de conceitos em determinadas partes dessa estrutura textual de jornal e, esse questionamento, é de fácil memorização. Entretanto, no momento de elaboração do modelo, notamos que o mesmo se concentrou na análise de conteúdo do texto jornalístico porque enfatizava a observação das variáveis leitor/indexador e o texto. E, até então, havia se percebido que a identificação e a seleção de conceitos ocorriam, durante a leitura documentária, a partir da exploração do conteúdo do texto, por isso, a elaboração do modelo ressaltou a concepção orientada para o conteúdo. Na análise dos dados coletados, percebeu-se que a maioria dos indexadores não explicitava o objetivo da leitura que, por sua vez, reflete a política de indexação defendida pelo sistema de informação que, deduzíamos ser: identificar e selecionar conceitos que representem fidedignamente, o conteúdo do texto e que, os mesmos, venham a ser recuperados pelos usuários do sistema como informações relevantes para eles. Mas algumas afirmações dos indexadores gravadas durante a aplicação do Protocolo Verbal, presentes, aqui, em apêndice, evidenciam que identificam e selecionam considerando, muitas vezes, somente o conteúdo porque o sistema de recuperação garante aos usuários outras possibilidade de busca, isto é, de recuperação do texto sem ser somente por palavras-chaves, ou seja, conceitos como: nome da pasta de assunto na qual a matéria é inserida, autor da matéria, título, ou qualquer palavra do título, qualquer palavra do texto e nome do jornal. Essa problemática revelou que devemos propor para o DEDOC que, se preciso, redirecione, ou, simplesmente, avalie a sua política de indexação, por meio de questionário aplicado à sua comunidade usuária, procurando verificar o grau de satisfação dessa comunidade, com o

funcionamento do sistema de busca, incluindo, ainda, no questionário, a alternativa de estratégia de pesquisa mais utilizada, e se o resultado, utilizando-a, satisfaz seus anseios informacionais (modelo questionário proposto, vide seqüência dessa seção).

Portanto, pode-se perceber que o modelo de leitura proposto não se ateve para a combinação entre a representação do conteúdo e necessidade informacional da comunidade usuária. Evidenciamos, também, com a análise dos dados coletados com o protocolo verbal durante a leitura, que os indexadores de jornal identificam conceitos considerando o conteúdo do documento e que uma pequena parcela, principalmente, os indexadores experientes identificam, considerando o conteúdo do texto, e selecionam levando em conta a necessidade ou interesse dos usuários pelo assunto. Assim, percebemos a necessidade de modificar o modelo de leitura na etapa referente à identificação de conceitos para orientar o indexador, para que identifique considerando o conteúdo do texto, e selecione o conceito identificado somente se o mesmo refletir as necessidades informacionais da comunidade usuária, porque o objetivo da indexação é a recuperação do documento pelo usuário a que ele se destina, ou seja, *o documento certo para o usuário certo, no tempo certo*.

Com a proposta de modificação do modelo de leitura por meio da redefinição dos objetivos das etapas que ocorrem durante a leitura que são identificar e selecionar conceitos, o modelo ficou assim:

- 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 5 primeiros parágrafos, sem realizar leitura linear, buscando encontrar respostas aos questionamentos propostos pelo modelo de leitura descritas mais especificamente no quadro abaixo, (alertamos que nem todas as questões estarão presentes num texto)
- 3- Encontre no texto as possíveis respostas ao questionamento: *O quê?; Quem ?; Quando?; Onde? Como?(ou Para que?); Por quê? (ou Para quê?)*
- 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 quê?; Quem ?; Quando?; Onde? Como?(ou Para que?); Por quê? (ou Para quê?)

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

Quadro 32 : Modelo definitivo de leitura para texto jornalístico³

Conceito para análise de assunto do sistema (PRECIS)	Parte da Estrutura Textual (Van Dijk, 1983) onde podemos encontrar o questionamento do LEAD e os elementos do sistema PRECIS	Questionamento ao (LEAD) da estrutura do texto jornalístico segundo VAN DIJK (1983) ; SILVA (2001)	Lembre-se que para atingir o objetivo da leitura documentária o indexador deve:
AÇÃO	Sumário/Introdução (Cabeçalho e Lead)	O quê (o que aconteceu, está ou vai acontecer)	Identificar conceitos considerando o conteúdo do documento.
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 e Sumário/Introdução (Cabeçalho e Lead)	Quando (dia da semana e E do mês, horas) Onde (o local do acontecimento)	
CAUSA E EFEITO	Sumário/Introdução (Lead) ou Episódio (s)	Por quê, ou Para quê (os motivo e as razões)	

³ O leitor para uso do questionamento proposto pelo modelo de leitura pode optar entre os elementos do questionamento para análise de assunto do sistema de indexação Precís, ou o questionamento elaborado por VAN DIJK e SILVA (2001) para o LEAD da estrutura textual do texto de jornal.

Portanto, com as modificações necessárias realizadas, atingimos um modelo eficiente na identificação de conceitos para leitura de artigos de jornais que era o maior objetivo desta pesquisa.

Evidenciamos que formulamos uma proposta à instituição onde ocorreu a observação da leitura dos indexadores nesta pesquisa de mestrado, no caso, Departamento de Documentação da Editora Abril Cultural (DEDOC) solicitando-lhe que verifique sua política de indexação por meio de aplicação de um formulário contendo questões a sua comunidade usuária-alvo, no caso, os jornalistas, como mencionados no capítulo 5. A sugestão baseia-se em algumas idéias de Lancaster (1973 e 1979), citado por Carneiro (1985) sobre avaliação dos sistemas de recuperação da informação. Assim, o modelo do formulário proposto é:

Questionário para verificar a satisfação Da comunidade usuária do DEDOC quanto ao funcionamento do sistema de recuperação

- 1- A estratégia de busca mais utilizada é:** A- pasta de assunto ()
B- Palavras-chave ()
C- -Autor da matéria ()
D- Palavras do Título ou do texto (Termo livre) ()
E- Nome do jornal + termo livre ()
- 2- O resultado obtido lhe satisfaz utilizando essa estratégia?** Sim ()
Não ()
Às vezes ()
- 4- O tempo de resposta do sistema utilizando essa estratégia é tolerável?** Sim ()
Não ()
Às vezes ()
- 5- A forma de saída do dados recuperados apresentada pelo sistema é tolerável considerando a seleção do que realmente interessa?** () Sim
() Não
() Às vezes

Ressaltamos que o questionário proposto para o DEDOC é uma sugestão e se estabeleceram esses parâmetros, considerando como relevante, para a instituição, o conhecimento da estratégia de busca mais utilizada pelos usuários e, se a mesma resulta em informações relevantes para a comunidade usuária, visando que o resultado permitirá ao DEDOC verificar se precisa adequar a política de indexação dos jornais porque,

conforme constatado, os indexadores não se preocupam em atribuir termos de indexação haja visto que confiam que o documento pode e, provavelmente, tem grande possibilidade de ser recuperado por outra estratégia de busca, que afirmam que os usuários conhecem e utilizam. E uma outra problemática levantada no formulário é a questão da satisfação com o tempo de resposta do sistema, com relação a estratégia mais utilizada que, por sua vez, está relacionada com o formato de saída dos dados oferecido pelo sistema de recuperação; essa indagação fornece ao DEDOC a possibilidade de conhecer as exigências de sua comunidade-usuária, porque se o usuário aceita uma demora maior na resposta, demonstra que visa recuperar um maior número de documentos possíveis, e o contrário, se o usuário exige uma resposta mais rápida, é sinal que objetiva documentos específicos, ou seja, seletos, o que realmente lhe interessa no acervo.

Evidenciamos que as pendências quanto à adequação do modelo a considerar no momento de seleção de conceitos, os interesses da comunidade usuária foram resolvidos com as alterações feitas no questionamento para utilização do modelo de leitura proposto, o que resultou num modelo, provavelmente, eficiente para indexação de jornais; e fica aqui a sugestão de avaliação da política de indexação para o DEDOC, porque os dados analisados e discutidos durante toda a pesquisa demonstraram que, talvez, ocorra na instituição uma “camuflada” satisfação dos usuários quanto à política de indexação adotada e, caso seja preciso, uma adequação na política de indexação, é recomendável que o DEDOC documente, por escrito, a nova política de indexação adotada, uma vez que, essa postura *contribuirá para a eficiência do serviço de indexação e servirá de orientação, em caso de mudança na administração do setor.*

Por fim, ressaltamos que a sugestão de avaliação da política de indexação do DEDOC se faz necessária porque, conforme mencionamos no capítulo 3, a política de indexação influencia na realização da leitura do indexador porque estabelece o perfil de usuário a que se destina o serviço de indexação devendo ficar explícito para o leitor-indexador, no momento da seleção dos termos que irão representar o conteúdo de determinado documento, se, deve-se selecionar termos mais genéricos, ou termos específicos, ou ainda, no caso do DEDOC, qual categoria de estratégia de busca mais utilizada pelos usuários, porque a opção de disponibilizar termos genéricos, específicos ou a escolha por disponibilizar essa ou aquela possibilidade de busca, deve refletir os interesses informacionais dos usuários seguindo os padrões de uma linguagem documentária que considerou na sua elaboração, coerentemente, o perfil da comunidade-

usuária que a instituição servirá que foi, por sua vez, anteriormente, determinado na elaboração da sua política de indexação. Sendo assim, a política de indexação estabelece as formas de recuperação do acervo e quando coerentemente definida, isto é, acoplada às concepções de identificação de conceitos direcionadas para o conteúdo do documento e para os interesses dos usuários, direcionará e influenciará positivamente, entre outras variáveis, no resultado da compreensão de leitura do indexador frente ao texto a indexar.

9 REFERÊNCIAS BIBLIOGRÁFICAS

A HISTÓRIA de um jornal. Almanach D" Estado de São Paulo"p.31-47.

ABNT. *NBR 12676: Métodos para análise de documentos – determinação de seus assuntos e seleção de termos de indexação*. Rio de Janeiro. 4 p. 1992.

AHMAD, N. Newspaper indexing: an international overview. *The Indexer*, v.17, n.4, oct., p. 257-266. 1991.

ALBRECHTSEN, H. Subject analysis and indexing: from automated indexing to domain analysis. *The Indexer*, 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*. São Paulo: Universidade de São Paulo. 1991. (Dissertação de Mestrado).

ANDERSON, J. R. A *Spreading activation theory of memory*. Journal of Verbal Learning and Verabl Behavior. p. 261-295. 1983.

ARISTÓTELES. *Arte retórica e arte poética*. Traduzido por: Antonio Pinto de Carvalho . São Paulo: D. E. L. , 1959. 335 p.

BEARE, G. Local newspapaer indexing projects and products. *The Indexer*, v.16, n.4, october. p 227-233. 1989.

BELLOTTTO, H. L. *Arquivos permanentes: tratamento documental*. São Paulo: T. A . Queiroz, 1991. 198 p.

BENITEZ, J. A. *Técnica periodística*. Havana, Editoria Pueblo y Educación, 1980. 160 p.

BERNIER, C. L. *Abstracting*. In: *Encyclopaedia of Library of Information Science*. New York: Marcel Dekker, 1969. p. 18.

BLOOME, D. Reading as a social process. *Advances in Reading/Language Research*, v. 2, 165-95, 1983.

BROWN, A. L. *Metacognitive development and reading*. In: SPIRO, et al (org.) *Theoretical issues in reading comprehension*. Hillsdale: Lawrence Erlbaum, 1980.

BUENO, R. M. *Receitas culinárias como fonte de informação: elementos para seu tratamento temático*. Marília: UNESP. 1998. (Trabalho de Conclusão de Curso).

CALFEE, R. ; CURLEY, R. . *Estructuras discursivas en las diferentes areas del conocimiento*. In: RODRIGUES, E.; LAGER, E. (Comp.) *La lectura*. Santiago de Cali: Universidade del Valle, 1997. p. 53-80.

CARNEIRO, M. V. *Diretrizes para uma política de indexação*. Revista da escola de Biblioteconomia:UFMG, belo Horizonte, v.14, n.2, p. 221-241, set. 1985.

CAVALCANTI, , M. C. I-N-T-E-R-A-Ç-Ã-O leitor – texto: aspectos de interação pragmática. Campinas: editora da UNICAMP. 1989.

CESARINO, M. A. da N.; PINTO, M. C. M. F. Cabeçalho de assunto como linguagem de indexação. *Revista da Escola de Biblioteconomia*. UFMG, Belo horizonte, v.7, n.2, set.p.268-288. 1978.

CESSEL, V. L. *Avaliação de linguagens documentárias para indexação em bases de dados*. Marília: UNESP. 1998. (Trabalho de Conclusão de Curso).

CHARADEAU, P. *Language et discours: éléments de semiolinguistique*. Paris: Hachette Université. 1983.

CINTRA, A. M. M. *Estratégias de leitura em documentação*. In: SMITH, J. W. *Análise documentária: análise da síntese*. Brasília: IBICT, 1987. p. 29-38.

CITELLI, A. *Linguagem e persuasão*. 4 ed. São Paulo: Ática. 1989.

CORREIA, A. O. *A construção de tesouros na perspectiva da metodologia facetada*. Marília: UNESP/FAPESP, 1998. 117 p. (Relatório parcial de pesquisa apresentado Fundação de Amparo à Pesquisa do Estado de São Paulo).

CUNNINGHAM, P. (1987) Are your vocabulary words lunules or lupulins? *Journal of Reading*, v. 30, n. 4, p. 344-350.

ERICSON e SIMON. *Verbal reports on thinking*. 1987. In: FAERCH e KASPER (eds.) *Introspection in second language research*. Clevedon: Multilingual Matters Ltd.

FAERCH, C. ; KASPER, G. (Eds.) *Introspection in second language research*. Clevedon: Multilingual Matters, 1987.

FAGUNDES, S. A. *Leitura em análise documentária de artigos de jornal*. Marília: UNESP. 1997. (Trabalho de Conclusão de Curso).

FARROW, J. F. propositional analysis and macrorules in indexing. *Library Review*, v.45, n.1, p.6-15. 1996.

_____. A cognitive process model of document indexing. *Journal of Documentation*, v. 47, n.2, 149-166. 1991.

_____. All in the mind: concept analysis in indexing. *The Indexer*, v. 19, n. 4, p.245-247. 1995.

FODOR, J. A. *The modularity of mind*. Cambridge, Mass: The MIT Press, 1983.

FUJITA, M. S. L. *Precis na língua portuguesa: teoria e prática de indexação*. Brasília: UNB. 1989. 195 p.

_____. *A leitura em análise documentária*. Marília: UNESP, CNPq. 1998. 184 p. (Relatório de Pesquisa).

_____. *A leitura em análise documentária*. Marília: UNESP, CNPq. 1999. 196 p. (Relatório Parcial de Pesquisa).

_____. *Leitura em Análise Documentária: uma contribuição à formação do indexador*. Marília:UNESP, CNPq. 2000-2002. 185 p. (Relatório Parcial do Projeto Integrado de Pesquisa).

GÁLVEZ, Carmen & PINTO, Maria. (s.d) *Análisis documental de contenido: procesamiento de información*. Madrid: Síntesis, 158 p.

GARZÓN, A. & SEOANE, J. *La memoria desde el procesamiento de información*. 1982. Apud GÁLVEZ, C. & PINTO, M. *Análisis documental de contenido: procesamiento de información*. Madrid: Síntesis, s.d. 158 p.

GIASSON, J. *A compreensão na leitura*. Trad. Maria José frias. Portugal: Edições Asa. 1993. 320 p.

GINEZ DE LARA, M. L. G. de *A representação documentária: em jogo a significação*. São Paulo: Universidade de São Paulo, 1993. Dissertação. (Mestrado em Ciências da Comunicação) - Escola de Comunicações e Artes, Universidade de São Paulo.

GOTOH, T. Cognitive structure in human indexing process. *Library and Information Science*, n. 21, p. 209-226. 1983.

GRACIOSO, L. S. *Avaliação de linguagem documentária em bases de dados*. Marília: UNESP. 1998. (Trabalho de Conclusão de Curso).

GREICE, J. B. *Logic and conversation*. In: COLE, P., MORGAN, P. *Syntax and semantics*. New York: Academic Press. 1975 p. 41-58.

GUIMARÃES, J. A . C. *Análise documentária em jurisprudência: subsídio para uma metodologia de indexação de acórdãos trabalhistas brasileiros*. São Paulo: Universidade de São Paulo. Tese (Doutorado). 1994. 250 p. (Ciência da Comunicação).

_____. O caráter instrumental da diplomática para o tratamento de documentos na área jurídica.. Marília: Universidade Estadual Paulista – Marília publicações: *Cadernos da faculdade de Filosofia e Ciências*, v.7, n; ½, 209 p. 97- 106 p. 1998.

HOSENFELD, C. A preliminary investigation of the reading strategies of successfull and nonsuccessful second language learners. *System*, v. 5, p. 110-123. 1977.

HOUSMAN, A . E. Apud GROGAN, D. *A prática do serviço de referência*. Trad. Antonio Agenor Briquet de Lemos. Brasília: Briquet de Lemos, 1995. 196 p.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION Documentation - methods for examining documents, determining their subjects, and selecting indexing terms. Suíça: ISO. 5 p. (ISO 5963 - 1985)

IRWIN, J. *Teaching reading comprehension processes*. Englewood, New Jersey: Prentice-Hall, 1986.

IWASHITA, K. H. *Estratégias de leitura para elaboração de resumos, com resumidores do Centro de Energia Nuclear, da Comissão nacional de Energia Nuclear*. Marília: UNESP. 1999. (Trabalho de Conclusão de Curso).

KATO, M. *No mundo da escrita: uma perspectiva psicolinguística*. São Paulo: Ática, 1986. 144 p.

KIENTZ, A . *Comunicação de massa: análise de conteúdo*. Rio de Janeiro: Eldorado. 1973.

KINTSH, W. & VAN DIJK T. Toward a model of text comprehension and production. *Psychological Review*, vol. 85. P. 363-394. 1978.

KLEIMAN, A . *Texto e leitor: aspectos cognitivos da leitura*. Campinas: Pontes., 1989. 82 p.

KOBASHI, N. Y. *A elaboração de informação documentária: em busca de uma metodologia*. São Paulo: Universidade de São Paulo, 1994 (Tese Doutorado em Ciências da Comunicação).

KOCHEN, M. & TAGLIACCOZZO, R. A study of cross-referencing. *Journal of Documentation*, v.24, n.3. p. 173-191. 1968

LANCASTER, F. W. & FAYEN, E. G. *Information retrieval online*. Los Angeles, Melville, 1973. 597 p. Apud CARNEIRO, M. V. *Diretrizes para uma política de indexação*. Revista da escola de Biblioteconomia:UFMG, belo Horizonte, v.14, n.2, p. 221-241, set. 1985.

LANCASTER, F. W. *Information retrieval systems: characteristics, testing and evaluation*. 2 ed. New York, J. Willey, 1979. 381. Apud CARNEIRO, M. V. *Diretrizes para uma política de indexação*. Revista da escola de Biblioteconomia:UFMG, belo Horizonte, v.14, n.2, p. 221-241, set. 1985.

_____. *Indexação e resumos: teoria e prática*. Trad. De Antonio Agenor Briquet de Lemos. Brasília: Briquet de lemos/Livros, 1993.

_____. *El control del vocabulario en la recuperación de información*. Trad. Alejandro de la Cueva Martin. Valencia, 1995, p. 17-248. Tradução de: Vocabulary control for information retrieval, led. (Colección Educación Materials, 12).

LEFFA, V. J. *Aspectos da leitura*. Porto Alegre: Sagra, 1996. 104 p..

MARCONDES FILHO, C. (org.) *A linguagem da sedução*. São Paulo: Com-Arte. 1985.

MARSHALL, N. *Discourse analysis as a guide for informal assesment of comprehension*. In: FLOOD (Ed.) *Promoting reading comprehension*. Newark Delaware. International Reading Association. 1984. p. 79-97.

MEDEIROS, J. B. *Redação científica: a prática de fichamentos, resumos, resenhas*. 3 ed. São Paulo: s.n, 1997. 231 p.

MEYER, B. *The structure of prose: effects on learning and memory. Implications for education practice*. In: ANDERSON, R. , RAND, J. y MONTAGUE, W. (Eds.) *Schooling and the acquisition of knowledge*, Hilsdale: Erlbaum. 1977, p. 179-200.

MOLINA, O. *Ler para aprender: desenvolvimento de habilidades de estudo*. São Paulo: EPU. Apud MEDEIROS, J. B. *Redação científica: a prática de fichamentos, resumos, resenhas*. 3 ed. São Paulo: s.n, 1992. 231 p.

MOREIRO GONZÁLEZ, J. A . *Aplicación de las ciencias del texto al resumen documental*. Madrid: Universidad Carlos III de Madrid, Boletín Oficial del Estado. 1993. p. 237.

NARDI, M I. A . *As expressões metafóricas na compreensão de texto escrito em língua estrangeira*. São Paulo: PUC. Dissertação (Mestrado em Linguística Aplicada ao Ensino de Línguas)- Pontifícia Universidade Católica, São Paulo. 1993.

_____. *A metáfora e a prática de leitura como evento social: instrumentos do pensar a Biblioteconomia do futuro*. São Paulo: PUC. Tese (Doutorado Linguística Aplicada e Estudos da Linguagem) – Pontifícia Universidade Católica, São Paulo. 1999. 254 p.

NAVES, M. M. L. *Análise de assunto: concepções*. *Revista de Biblioteconomia de Brasília*, v.20, n.2, jul/dez, 1996. P.215-226.

OKASABE, H. *Argumentação e discurso político*. São Paulo: Kairós. 1979.

OLSHAVSKI, J. E. Reading as problem solving: na investigation of strategies. In: *Reading Research Quarterly*. 12, 1976-1977. P. 654-674. Apud CAVALCANTI, M. C. I-N-T-E-R-A-Ç-Ã-O leitor – texto: aspectos de interação pragmática. Campinas: editora da UNICAMP. 1987.

OXFORD. Looking at language learning strategies. *Language Learning*. New York: Newbury House Publishers, 1989.

PEDROSO, R. N. *Metodologia do ensino de redação jornalística*. Porto Alegre: Revista de Biblioteconomia e Comunicação, v.7, jan/dez. p. 208-218. 1996

PIEDEDE, M. A. R. *Estudo comparativo de algumas linguagens de indexação: eficácia e tempo de pesquisa*. Rio de Janeiro: UFRJ. 1976. (Dissertação de mestrado).

PINTO MOLINA, M. ; GÁLVEZ, C. *Análisis documental: fundamentos y procedimientos*. 2 ed. Ver. Aum. Madrid: Síntesis. 1996. 158 p.

PINTO, M. C. M. F. Análise e representação de assuntos em sistemas de recuperação da informação: linguagens de indexação. *Revista da escola de Biblioteconomia*. UFMG, Belo Horizonte, v.14, n.2, p.169-186, set. 1985.

PIRES, R. A. *Estratégias de leitura para indexação: um estudo de observação dos procedimentos de leitura com indexadores em Agronomia*. Marília: UNESP. 1999. (Trabalho de Conclusão de Curso).

RADFORD, J. & BURTON, A. *Thinkings: its nature and development*. John Wiley & Sons, 1974. Apud CAVALCANTI, M. C. I-N-T-E-R-A-Ç-Ã-O leitor – texto: aspectos de interação pragmática. Campinas: editora da UNICAMP. 1987.

RUBI, M. P. Estratégias de leitura para indexação em energia nuclear. Marília: UNESP, 1999. 86 p. (Relatório Parcial de Pesquisa).

_____. Estratégias de leitura para indexação em energia nuclear. Marília: UNESP, 2000. 167 p. (Relatório Final de Pesquisa).

RUMELHART, D. E. *Hacia una comprensión de la comprensión*. In: RODRIGUEZ, Emma; LAGER, Elisabeth. (Comp.) *La lectura*. Santiago de Cali: Universidad del Valle, 1997. p. 25-51.

RUMELHART, D. E. *Toward an interactive model of reading*. In: DORMICI, S. (org.) *Attention and performance XL*. Lawrence Erlbaum Associates. 1977.

SÁNCHEZ, E.; ORRATIA, J. ; ROSALES, J. Como mejorar la comprensión de textos en el aula. *Comunicación, Lenguaje y education*, n. 14. P.95. 1992.

SILVA, V. J. *Regras de construção*. Disponível em: www.publico.pt/nos/livro_estilo/paginas/1e12_regras.htm, em 27 de março de 2001.

SMITH, F. *Compreendendo a leitura: análise psicolinguística da leitura e do aprender a ler*. Trad. Daise Batista, 3. ed. Porto Alegre: Artes Médicas, 1989.423 p.

SPRADLEY, J. *Participant Observation*. New York: Holt Rinehart & Winston, 1980.

STRAIOTO, A. C. *Análise documentária de embalagens de shampoo: uma alternativa de uso do método diplomático para o profissional da informação*. Marília: UNESP, 1997. (Trabalho de Conclusão de Curso).

TÁLAMO, M. F. G. M. A compreensão literal de textos. São Paulo: Escola de Comunicações e Artes - USP. *Cadernos de análise documentária*, n. 1, maio, 64 p. 1994.

TÁLAMO, M. F. G. M. *Elaboração de resumos*. 1987. Escola de Comunicação e Artes. São Paulo, 14 p. (datilografado)

TESSIER, G. *La diplomatique*. Paris: P. U. F. 1952. Apud GUIMARÃES, J. A. C. O caráter instrumental da diplomática para o tratamento de documentos na área jurídica.. Marília: Universidade Estadual Paulista – Marília Publicações: *Cadernos da faculdade de Filosofia e Ciências*, v.7, n; ½, 1998. 209 p. 97- 106 p.

Um estado de alerta: há 117 anos. Nosso Espaço, n.39, fev. 1992.

UNISIST. Princípios de indexação. *Revista de Biblioteconomia*. UFMG, v.10, n.1. 1981. p.83-94.

VAN DIJK, B. M. Definition of thesauri essential characteristic. 2 vols. Brussels. 1976.

VAN DIJK, T. *La ciencia del texto: un enfoque interdisciplinario*. Barcelona: Paidós. 1992. 303 p..

VAN DIJK, T. A. & KINTSCH, W. (1983) *Strategies of discourse comprehension*. Nueva York: Academic Press. 1983. p.337.

VAN DIJK, T. A. Discourse analysis: its development and application to the structure of news. *Journal of Documentation*, v.33, n.2, p.20-23. 1983.

VEGA, M. de, et al. *Lectura y comprensión: una perspectiva cognitiva*. Madrid: Alianza, 1990, p.17.

APÊNDICE A: GRADE DIPLOMÁTICA DO TEXTO JORNALÍSTICO

GRADE DIPLOMÁTICA DO TEXTO JORNALÍSTICO

Área	Sub área	- Tipologia documentária	Estrutura textual Elementos principais	Elementos temáticos encontrados nas partes principais da estrutura textual ao lado	Elementos descritivos encontrados nas partes principais da estrutura do texto
Mídia	Jornalis mo	Textos jornalísticos impressos	Cabeçalho, Lead, Episódio (s)	Título, Assunto	Coluna de publicação, data, página, colaborador (autor)

**APÊNDICE B: ORGANIZAÇÃO E FUNCIONAMENTO DO ARQUIVO
DO JORNAL “O ESTADO DE SÃO PAULO”**

ORGANIZAÇÃO E FUNCIONAMENTO DO ARQUIVO DO JORNAL “O ESTADO DE SÃO PAULO”

Histórico

Conforme publicação que relata a história do jornal, “O Estado de São Paulo” nasceu no século XIX em 1876 na cidade de São Paulo, quando esta era apenas uma província.

O jornal surgiu com a formação de uma sociedade comanditária formada por um grupo de cavalheiros que estavam insatisfeitos com os jornais da cidade, uma vez que estes publicavam notícias sobre atos oficiais, notícias atrasadas e não possuíam redatores capazes de discutir os problemas nacionais da época.

Essa sociedade era composta pelos cavalheiros: Capitão Bento A. de Almeida Bicudo, Antônio Pompeu de Camargo, Dr. Américo Brasiliense de Almeida Mello, Dr. João Francisco de Paula Souza, João Manuel de Almeida Barbosa, Dr. M. F. de Campos Salles, Dr. Raphael Paes de Barros, Major Diogo de Barros, Dr. João Tobias de Aguiar e Castro, Manuel Elpidio Pereira de Queiroz, João Tibiriçá Piratininga, José de Vasconcellos Almeida Prado, José Pedroso de Moraes Salles, Antônio Carlos de Salles, Francisco de Salles, Dr. Martinho Pedro Júnior, Dr. José Alves de Cerqueira César, Cândido Valle, Francisco Glicério, F. Rangel Pestana e Américo de Campos.

O jornal foi fundado com o nome de “A Província de São Paulo”, mudando para “O Estado de São Paulo” quando a cidade deixou de ser uma província.

O objetivo do jornal era ser uma publicação voltada para a realidade da sociedade da época e isso se constituía um desafio, pois a forma de governo era a Monarquia e a escravidão ainda existia na cidade.

A sociedade comanditária que fundara o jornal foi-se diminuindo diante de cada crise que passara o jornal, terminando essa sociedade em 18 de janeiro de 1882.

A crise de 1885 trouxe para o jornal o advogado Júlio de Mesquita que oferece um desinteressado auxílio ao jornal, conseguindo negociar um acordo para impedir que o jornal caísse nas mãos de qualquer adversário de suas aspirações políticas. (ALMANACH, DO ESTADO DE SÃO PAULO, s.d.). Essa decisão tomada por Júlio de Mesquita o leva a se inserir no jornal e com o passar do tempo passou a Diretor do jornal, o que veio a inovar a posição do jornal em relação ao país.

Em 1915, publicava-se a primeira edição noturna do jornal, chamada de “O Estadinho”, “essa publicação defendia a causa aliada na Primeira Guerra Mundial, procurando sempre conciliar as práticas políticas da época ao seu ideário liberal.”(ALMANACH, DO ESTADO DE S. PAULO)

Após a Revolução de 1924, quando o jornal se preparava para enfrentar mais uma de suas crises, teve sensores do governo em sua redação, começa então a censura.

Nesse período o jornal tinha colaboradores como: Monteiro Lobato, Olavo Bilac e Euclides da Cunha que foi enviado pelo jornal para realizar reportagens sobre a Guerra de Canudos o que fez com que escrevesse o clássico "Os Sertões." (ALMANACH, DO ESTADO DE S. PAULO)

No ano de 1927 Júlio de Mesquita faleceu, assumindo então seus filhos , Júlio de Mesquita Filho e Francisco.

Conforme o ALMANACH DO ESTADO DE S. PAULO (s.d.) o jornal participou e incentivou a Revolução Paulista de 1932. Esse procedimento custou o primeiro exílio imposto aos diretores do jornal.

De acordo com o ALMANACH DO ESTADO DE S. PAULO (Op. Cit) quando Júlio de Mesquita Filho regressou do exílio, transformou em realidade um dos seus sonhos que foi a criação da Universidade de São Paulo, contando com o apoio do jornal " O Estado de São Paulo".

Em 1937, no governo de Getúlio Vargas, instituiu-se "*a supressão de todas as liberdades e a instituição da censura oficial,*" consequentemente, o segundo exílio de Júlio de Mesquita Filho e o confisco do jornal "O Estado de São Paulo", que foi "*transformado, a força, em órgão da Ditadura.*" Essa situação durou até 1945, sendo que esse período de censura de cinco anos não é incluído na sua História oficial do jornal.(ALMANACH DO ESTADO DE S. PAULO)

Com o apoio dado ao Movimento de 1964, o jornal volta a ser censurado, estendendo-se esse período até 1975, sendo proibida a impressão de um total de 2.952 textos que foram substituídos por versos dos livros do escritor Camões.

No ano de 1966 surge o "Jornal da Tarde", criado por Ruy Mesquita, este jornal possuía características de um jornal jovem, cujo objetivo era se tornar a leitura predileta da faixa mais jovem da sociedade.

O jornal da tarde também foi censurado e no lugar das notícias censuradas eram publicadas receitas culinárias.

A censura deixou de existir no jornal em 4 de janeiro de 1975, "*sem explicações , assim como entrou.*"

Após a Ditadura, o jornal recorreu a justiça para receber indenizações pelos danos causados pela censura policial e ganhou a causa.

Com o término da Ditadura o jornal voltou a crescer, com base num projeto de modernização que incluía "modernos recursos, destacar-se entre os grandes órgãos da imprensa mundial e a criação de uma editoria internacional com a criação de uma vasta rede de correspondentes."(NOSSO, ESPAÇO 1992)

Ressaltamos que buscando o desenvolvimento, o que gerou a necessidade de novas instalações, o jornal esteve instalado em vários endereços e hoje encontra-se instalado na Marginal do Tietê e neste local encontram-se as demais empresas do grupo, são impressos o jornal "O Estado de São Paulo", "Jornal da Tarde", publicações de terceiros e as listas telefônicas oficiais de São Paulo, Guarulhos e do interior do estado

do Rio de Janeiro que são distribuídas pela Oesp Gráfica. O jornal encontra-se associado a outras empresas e empregava em 1997 mais de 8.000 pessoas, constituindo um órgão da imprensa e patrimônio nacional.

Descreveremos agora o histórico do arquivo sua infra-estrutura, funcionamento e caracterização dos seus profissionais indexadores e sua política de indexação.

Histórico do Arquivo do Jornal “O Estado de São Paulo”

O arquivo é formado por pastas com recortes de textos, fotografias de pessoas que tenham sido notícias e ainda conta com uma pequena biblioteca onde podemos encontrar livros, enciclopédias, periódicos etc...

Os funcionários recortam nos jornais os textos e fotografias que foram selecionadas pelos indexadores durante a leitura documentária e que comporão as pastas dos assuntos. Consequentemente, é preenchida uma ficha onde se inclui os assuntos extraídos do artigo; um sumário da matéria publicada em 3 linhas, data da publicação; nome da publicação que a estampou; número da página e a coluna em que apareceu, formando assim o índice de assuntos.

O arquivo possuía em 1997 em seu acervo, além das pastas, revistas como: Cruzeiro, Manchete, Life, Time, Squire, Oggi, Illustrated London News, Revista Bancária, Jeune Afrique, Paris Match e livros que constituem uma pequena biblioteca no arquivo.

O arquivo têm como usuários, estudantes a partir do primeiro ano do ginásio, universitários, redatores, reportes, escritores, pesquisadores e curiosos

O objetivo do arquivo é informar aos usuários sobre um determinado assunto de seu interesse, visando assim satisfazer a demanda da informação.

A pesquisa no arquivo antes era aberto ao público, mas em 1997 se encontrava interrompida devido aos problemas conseqüentes da circulação dos usuários pelo acervo como por exemplo furto de páginas de livros; periódicos e das pastas sendo os atendimentos estabelecido via telefone e sendo cobrado pelas pesquisas um custo pela informação fornecida que variava de R\$ 20 a R\$100.

De acordo com informação verbal de funcionária do arquivo, no arquivo assim como, em uma biblioteca o atendente, muitas vezes, tem que inferir o que o usuário quer encontrar, como, por exemplo, o caso de um certo estudante que queria saber se o arquivo de o jornal “O Estado de São Paulo” obtinha alguma notícia sobre “Garrete azul”, quando o que queria era alguma notícia sobre o Presidente Garrastazu Médici.

Área Física do Jornal e do Arquivo

- Jornal: possuía 2 edifícios de 7 andares que constituía a sede e outros edifícios externos;
- Arquivo: possuía uma área destinada a guardar a coleção do jornal desde de 1875, coleção do Jornal da Tarde desde 1966, contava com uma área física de 410 m2.

Relataremos agora a organização da instituição, identificação da clientela, o acervo, serviços executados, serviços prestados, recursos humanos, equipamentos e mobiliários.

Quadro 1: Identificação da Organização

Identificação da Organização
O arquivo está subordinado ao jornal “O Estado de São Paulo”

Ressaltamos aqui que a identificação da instituição é importante para se conhecer os objetivos, no caso o arquivo está subordinado ao jornal e possui a sua própria organização. Quanto aos objetivos do arquivo destacamos que têm por meta atender primeiramente interesses informacionais dos usuários da redação e depois a comunidade externa.

Quadro 2: Identificação da Clientela do arquivo

Usuários Internos	Usuários Externos
Jornalistas da redação	Estudantes
Sucursais do Brasil todo	órgãos públicos
	assinantes do jornal
	público em geral

Destacamos aqui que os usuários internos do arquivo eram no ano de 1997 jornalistas do jornal “O Estado de São Paulo”, “Jornal da Tarde”, da Agência Estado e das Sucursais de todo Brasil.

Os usuários da Redação eram jornalistas do jornal “O Estado de São Paulo”, “Jornal da Tarde”; a Agência Estado e as Sucursais de todo o Brasil que solicitam pesquisas, por telefone, à Redação. Esses usuários procuram a Redação com o objetivo de acrescentar novos dados ao assunto de seu interesse.

Dentre outros públicos do arquivo encontramos os curiosos que consultam o arquivo, como também pessoas que queiram saber informações sobre matérias que foram publicadas pelo jornal. A maior procura de informações na época era sobre tabela do jornal do carro, matéria de saúde, casa de idosos etc...

Ressaltamos que o arquivo atendia em 1997 usuários de todo o Brasil, sendo mais consultado por usuários do Estado de São Paulo e geralmente os usuários do arquivo pesquisavam matérias mais antigas.

Quadro 3: Acervo do Arquivo do jornal “O Estado de São Paulo”

Pastas de Títulos (textos e fotografias)	Biblioteca II: 8.000 volumes em livros
Biblioteca I: 6000 títulos	

Observando a tabela acima percebemos que o acervo englobava em 1997 um total de 54.000 pastas de títulos (textos e fotografias). Esses textos e fotografias eram selecionados por indexadores do jornal sendo estes indexadores jornalistas e uma bibliotecária. Esses indexadores extraíam textos e fotografias do jornal “O Estado de São Paulo”, “Jornal da Tarde”, “Folha de São Paulo”, “Jornal o Globo”, “Jornal do Brasil” e ainda de revistas como: Isto É, Empresa, Time, News Week, Superinteressante entre outras.

O jornal estava sendo microfilmado e sendo previsto microfilmar toda a sua coleção, encontramos então em 1997 uma parte do seu acervo em microfilmes e existia cópias da coleção do jornal que se encontrava na Biblioteca Nacional do Rio de Janeiro e na biblioteca de Chicago nos Estados Unidos.

No acervo encontramos um total de 90% da informação corporativa em outros documentos como: correspondências, normas, documentação, contratos, publicações e outros textos.

O acervo contava com uma biblioteca de 6.000 títulos. Entre esses, encontramos livros de referencia e enciclopédias de assuntos diversos e ainda uma outra biblioteca com 8.000 volumes de livros, instalada no antigo salão nobre, do Sr. Júlio Mesquita e o tratamento desses material estava sendo realizado por pessoal especializado da Universidade de São Paulo (USP).

- Cobertura de Assuntos

Quadro 4: Cobertura de Assuntos do arquivo

Assuntos diversos são indexados no Arquivo
Pessoas que são notícias também são assuntos, ou termos de indexação
Empresas e instituições são indexados como temas

Como se observa acima o arquivo os assuntos são mesmos diversificados até porque nos jornais e nas revistas são publicadas matérias que abordam assuntos, ou temas variados. É importante mencionar que as categorias de assuntos do arquivo vai de encontro as mesmas identificadas na revisão de literatura sobre jornais por AHMAD (1991citado anteriormente).

- Política de seleção e Aquisição dos Documentos-Fonte

Quadro 5: Política de seleção e aquisição no Arquivo

Política de seleção	Política de Aquisição
Seleção consulta catálogo de editoras	Visando os interesses do usuários
Pesquisa na Internet	Opta-se sempre por obras mais gerais

A tabela acima menciona que a seleção de documentos no Arquivo é feita por meio de consulta as obras publicadas via editoras e também pela rede Internet. Quanto a aquisição é realizada baseando-se nos interesses dos usuários e como este possui interesses diversificados sempre procura-se adquirir documentos que abordam assuntos gerais, ou seja, que trate de vários temas.

Quadro 6: Serviços Executados no Arquivo do jornal “O Estado de São Paulo”

Indexação dos jornais
Digitação dos dados selecionados na indexação
Montagem de pastas de assuntos
Descarte de material obsoleto
Microfilmagem dos jornais e das pastas de recortes de artigos de jornais

A tabela acima nos mostra que no ano de 1997 realizava-se no arquivo do jornal os serviços como: indexação de jornais sendo indexados o jornais “ O Estado de São Paulo” e do “Jornal da Tarde”, ou seja, jornais da própria instituição.

Um outro serviço executado era digitação dos dados retirados durante a indexação para fichas que compõem o índice de assuntos procurando preencher os seguintes campos: autor, data, página, coluna de publicação, autor etc....

Evidencia-se ainda os serviços como montagem de pastas de assuntos com base na seleção de assuntos feitas pelos indexadores, descarte de material obsoleto ao final de cada ano. E microfilmagem dos jornais “O Estado de São Paulo” e “Jornal da Tarde”.

Quadro 7: Serviços Prestados no arquivo do jornal “O Estado de São Paulo”

Atendimento de Pesquisa à Redação do jornal
Atendimento de pesquisa a terceiros

No arquivo o serviço prestado é atendimento de pesquisa primeiramente aos funcionários da redação do jornal “O Estado de São Paulo” e a comunidade externa. O atendimento é realizado com horário e data marcada por telefone. Esse atendimento é realizado por qualquer pessoa do arquivo. Inicia-se o atendimento perguntando ao usuário qual o assunto que deseja pesquisar?, à partir de qual período?. Posteriormente à anotação desses dados, o atendente pergunta ao usuário qual a sua necessidade de acesso a esses documentos?. E então são delimitadas as necessidades de cada usuário.

Ressaltamos que no arquivo o atendimento de pesquisa a comunidade externa é cobrado.

E que no ano de 1997 a pesquisa no arquivo era realizada via catálogo de fichas organizados por assunto, autor e título.

Quadro 8: Recursos Humanos do Arquivo

1 gerente geral do arquivo	3 funcionários encarregados da montagem do material nas pastas
3 subgerentes do Banco de Dados	3 indexadores
3 assessores direto da gerência	1 bibliotecária
1 secretária	3 funcionários encarregados da pesquisa
6 digitadores	1 supervisor de fotografia
3 supervisores que coordenam o trabalho	
9 funcionários encarregados da classificação e seleção do material	

O jornal no ano de 1997 possuía 8.000 funcionários. Com relação ao arquivo podemos visualizar os seus funcionários e os serviços que executam na tabela acima. E podemos observar que os funcionários desenvolvem atividades que visam atender aos objetivos do arquivo que é o atendimento aos jornalistas da redação e posteriormente a comunidade externa e se atingirá isso por meio do tratamento e organização dos documentos considerando as necessidades dos usuários.

Quadro 9: Recursos Materiais do arquivo

Mobiliário	Equipamentos
23 mesas de computador	3 impressoras
5 computadores de fábricas diversas	1 scanner
3 mesas para consulta de redatores	1 conjunto de impressora á laser
8 armários de aço	14 computadores
1 sala para reuniões com uma mesa e oito cadeiras	
10 estantes em aço que fazem parte da biblioteca	
28 cadeiras	

Observamos acima como os recursos materiais do arquivo: equipamentos e mobiliários.

APÊNDICE C: CARACTERIZAÇÃO DOS INDEXADORES DO ARQUIVO

Quadro 1: Identificação e Caracterização dos Indexadores do Arquivo

SUJEITO – IC	Indexador Permanente
1- Formação Educacional	1- Formado em Biblioteconomia e Documentação
2- Experiência Profissional anterior no serviço de indexação	2- Não possui experiência anterior
3- Outras atividades que realiza no Arquivo	3- Não realiza outras funções
4- Dificuldades na atividade de indexação	4- Não possui nenhuma dificuldade na indexação
5- Procedimentos de identificação de Conceitos	5- Realiza uma leitura silenciosa do jornal do texto todo, sem ler todos os parágrafos
6- Fontes de experiência de indexação	6- Adquiriu com o serviço de indexação porque trabalha no arquivo há 19 anos
7- Nível de exaustividade e especificidade exigido na indexação	7- Selecionar conceitos que são os mais específicos no texto

Observamos pelos dados da tabela acima que o sujeito IC é bibliotecário. Ressalta que adquiriu experiência com a prática de indexação ao longo dos anos de serviço, além de gostar de ler e acredita que o contato do dia-a-dia com a informação ajuda a formar o conhecimento com os termos da área do jornalismo.

Ressalta que realiza uma leitura silenciosa do texto sem ler tudo, procurando identificar os conceitos mais específicos do texto.

Quadro 2: Identificação e Caracterização dos Indexadores do Arquivo

SUJEITO – IM	Indexador Permanente
1- Formação Educacional	1- Graduado em Jornalismo, História, Pós-graduação em Turismo e meio Ambiente
2- Experiência Profissional anterior no serviço de indexação	2- Não possui experiência anterior
3- Outras atividades que realiza no Arquivo	3- Atendente nos plantões de Domingo e feriado
4- Dificuldades na atividade de indexação	4- Não possui nenhuma dificuldade na indexação
5- Procedimentos de identificação de Conceitos	5- Faz leitura do texto do começo do texto e verifica se o texto continua em outra página
6- Fontes de experiência de indexação	6- Adquiriu com o serviço de indexação porque trabalha no arquivo há 7 anos
7- Nível de exaustividade e especificidade exigido na indexação	7- Busca selecionar os conceitos que serão procurados pelos usuários

Destacamos que IM é jornalista e adquiriu experiência em indexação com a prática diária da atividade pois trabalha no jornal há 7 anos. E que não realizou nenhum curso em indexação.

Além da indexação, IM realiza outras atividades.

Destaca que não tem dificuldades para indexar devido à experiência e gosto pela leitura. entretanto, somente tem dificuldades quando tem barulho e conversas no arquivo.

Quadro 3: Identificação e Caracterização dos Indexadores do Arquivo

SUJEITO – IO	Indexador Permanente
1- Formação Educacional	1-Formado em Jornalismo
2- Experiência Profissional anterior no serviço de indexação	2-Não possui experiência anterior
3- Outras atividades que realiza no Arquivo	3- Não realiza outras atividades
4- Dificuldades na atividade de indexação	4-Não possui nenhuma dificuldade na indexação por trabalhar com isso há 10 anos
5- Procedimentos de identificação de Conceitos	5-Leitura do texto realizando uma leitura rápida
6- Fontes de experiência de indexação	6-Relata que apreendeu com a prática de indexação
7- Nível de exaustividade e especificidade exigido na indexação	7-Retira os assuntos mais importantes nos textos

O indexador IO é formado em Jornalismo, relata que adquiriu experiência em indexação com a prática desses 10 anos de trabalho. Não exerce outras funções no arquivo.

IO menciona não possuir dificuldades em indexação somente agora, porque no começo do serviço as tinha acredita não apresentar dificuldades no momento porque trabalha com o caderno de esportes e gosta, acha fácil. Mas quando encontra alguma dificuldade, discute com os colegas de trabalho para juntos chegarem a uma solução.

Realiza uma leitura rápida dos textos e seleciona os conceitos mais importantes

Quadro 4: Identificação e Caracterização dos Indexadores do Arquivo

1- Formação Educacional	1-Formada em Jornalismo
2- Experiência Profissional anterior no serviço de indexação	2-Não possui experiência anterior
3- Outras atividades que realiza no Arquivo	3-Avaliação de textos que comporão as novas pastas de recortes de jornais, descarte de textos e fotografias, ajuda no atendimento
4- Dificuldades na atividade de indexação	4-Não possui nenhuma dificuldade na indexação
5- Procedimentos de identificação de Conceitos	5-Realiza a leitura do jornal sem ler todo o texto porque não dá tempo
6- Fontes de experiência de indexação	6-Adquiriu com o serviço de indexação
7- Nível de exaustividade e especificidade exigido na indexação	7-Busca selecionar conceitos que serão úteis aos usuários

Observamos na tabela acima que o indexador IS é jornalista e trabalha no arquivo há 7 anos.

Realiza outras atividades além da indexação.

Relata IS não possuir dificuldades porque no curso de graduação em jornalismo apreendeu a distinguir entre uma matéria boa e uma ruim, o que realmente interessa e o que é “release”, ou seja, quando a matéria tem a função de destacar ou, divulgar determinado assunto se propondo, com isso, realizar propaganda de determinada pessoa, por exemplo, e que, com certeza, não será útil para os usuários.

**ANEXO D: ORGANIZAÇÃO E FUNCIONAMENTO DO ARQUIVO DO
DEPARTAMENTO DE DOCUMENTAÇÃO DA EDITORA ABRIL**

ORGANIZAÇÃO E FUNCIONAMENTO DO ARQUIVO DO DEPARTAMENTO DE DOCUMENTAÇÃO DA EDITORA ABRIL

Histórico Breve

A Editora Abril foi fundada em 1950 por Victor Civita.

Naquela época se pensar em fundar uma editora parecia uma loucura porque o contexto demonstrava que tinha tudo para dar errado porque 50 % da população brasileira era de analfabetos. Entretanto Victor Civita traçou um ideal de que a educação era, naquele contexto um compromisso e oportunidade de mudar aquela situação.

Com o passar dos tempos a Editora começa a se desenvolver e logo se destaca no ramo de editoração. E Victor Civita torna-se conhecido por seu caráter empreendedor cujo podemos destacar aqui exemplos de frases que demonstram essas afirmações: *“Sabe quem eu sou? “O senhor é o editor da revistinha que mais vende no Brasil.” “Sabe quantos armazéns frigoríficos há no Brasil? Nenhum. Eu vou fazê-los.”* (DEDOC ELETRÔNICO, 2000)

Hoje a Editora Abril é uma editora moderna que emprega 9.395 funcionários está localizada em 180.840 m² de área.

Relataremos agora o acervo, serviços executados, serviços prestados, a caracterização dos indexadores, equipamentos e mobiliário.

Quadro 1: Identificação da organização e da Clientela

Identificação da Organização	Identificação da Clientela
O DEDOC está vinculado a Editora Abril Cultural como serviço de apoio as necessidades da instituição Abril Editora	Usuários Externos: empresas, políticos, estudantes, crianças, entidades publicas.
	Usuários Internos: funcionários da casa, jornalistas, editores, funcionários da editora Abril Cultural e o presidente da empresa.

Ressaltamos que não foi possível obter o organograma do DEDOC devido ao fato desse possuir cargos e salários incluídos no organograma sendo, por isto, dados confidenciais da instituição. Entretanto evidenciamos que o DEDOC está subordinado a Editora Abril com objetivo de organizar as informações

presentes nos documentos que constituem o acervo documental da Editora como: livros, artigos de jornais, revistas e imagens (fotografias publicadas nos documentos).

Com relação aos usuários do DEDOC mencionamos que o objetivo primeiro é atender as necessidades informacionais dos usuários internos visualizados acima e posteriormente os usuários pertencentes a comunidade externa. E considerando as necessidades informacionais dos usuários internos é que o DEDOC decide sua política de seleção e aquisição de documentos descritos adiante.

Com relação ao acervo destaca-se que está constituído de:

Quadro 2: ACERVO DO DEDOC

Livros e Revistas – 13.000 exemplares	Jornais e Revistas que indexa
Pastas de Recortes – 21.000	O Globo, “O Estado de São Paulo”
40.000 Pastas de Títulos (Texto)	A Folha de São Paulo, Gazeta Mercantil, Jornal da Tarde, JB
150.000 Termos incluídos no DEDOC Eletrônico	Revistas todas as da casa e mais outras

A tabela acima menciona as tipologias documentárias que formam o acervo do DEDOC. Destacamos que o acervo é formado visando os interesses dos usuários internos então, procura-se formar o acervo com documentos cujo, o conteúdo interesse à tais usuários e depois estabelece os interesses dos usuários externos.

- Cobertura de Assuntos

Quadro 3: Cobertura de Assuntos do DEDOC

Assuntos diversos dentro das várias do conhecimento humano
Personalidades: pessoas que viraram notícias
Empresas, instituições

Observando a tabela acima percebemos que no DEDOC a cobertura de assuntos do sistema abrange temas diversos porque no Jornalismo e nas revistas são publicados matérias de temas variados, encontramos

também pessoas que são assuntos porque são sobre elas que determinada matéria relata e temos também as empresas, instituições, corporações que são assuntos porque foram relatadas em alguma matéria publicada.

- Política de Seleção e Aquisição dos Documentos-Fonte

Quadro 4: Política de Seleção e Aquisição do DEDOC

SELEÇÃO DAS OBRAS	AQUISIÇÃO DAS OBRAS
Pesquisa na Rede Internet	Análise em equipe das obras selecionadas
Contato com Livreiros	Interesses dos usuários internos e posteriormente os dos externos
	Obras de referências de conteúdo genérico

No DEDOC para seleção das obras realiza-se pesquisa na Internet e consulta em catálogos adquiridos com livreiros. Posteriormente a consulta, a aquisição das obras é decidida em equipe, visando identificar os interesses dos usuários internos pelas obras a serem compradas e depois os externo. Na aquisição de obras de referência procura-se comprar aquelas que apresentam conteúdo abrangente porque no DEDOC os interesses dos usuários internos é muito diversificado mediante a variedade de assuntos que se publicam nos jornais e revistas.

Quadro 5: Serviços Executados

Indexação de artigos de jornais	Preparação de material DEDOC Eletrônico
Indexação de fotos	Desenvolvimento de sistemas
Classificação de livros e periódicos	

Os serviços executados no DEDOC estão ilustrados acima, e dentre esses destacamos que os relacionados as atividades de Análise Documentária além, da indexação de jornais temos: indexação de fotos, classificação de livros e periódicos.

Quadro 6: Serviços Prestados

Atendimento de pesquisa a todo o grupo Abril
Atendimento de pesquisa à terceiros

O atendimento de pesquisa no DEDOC é agendado pessoalmente tanto para usuários internos como externos. Pode-se pesquisar procurando artigos de jornais, livros, revistas ou fotos no Abril Imagens. Destacamos também que a cada pesquisa é cobrado um valor, sendo que se o usuário optar por realizar a busca é cobrado X e se escolher que a busca seja delegada é cobrado Y. Entretanto mencionamos que dos usuários internos a taxa de cobrança, é direcionada ao departamento, ou editora a que o usuário pertence e não ao jornalista.

Quadro 7: RECURSOS MATERIAIS E EQUIPAMENTOS

EQUIPAMENTOS	MOBILIÁRIOS
Computadores	Cadeiras
Scanners	Mesas
Impressoras	Armários
Máquina Copiadora	

Como se observa acima, o Departamento de Documentação (DEDOC) não possui, uma discriminação dos materiais e equipamentos , mas notamos que estes existem .

Quadro 8: Recursos Humanos do DEDOC

ADMINISTRAÇÃO	ATENDIMENTO DE PESQUISA
1 Gerente	3 pesquisadores
1 Gerente de Pesquisas	1 supervisor
1 Secretária	1 Auxiliar
1 Administrador	

Quadro 9: Recursos Humanos do DEDOC

INDEXAÇÃO DE TEXTOS	BIBLIOTECA DO DEDOC
7 indexadores	1 bibliotecária
1 supervisor	1 auxiliar

Quadro 10: Recursos Humanos do DEDOC

PREPARAÇÃO MATERIAL DEDOC	DESENVOLVIMENTO DE SISTEMAS ELETRÔNICO
1 indexador	2 analistas de sistemas
3 auxiliares	1 supervisor de sistemas

Quadro 11: Recursos Humanos do DEDOC

SETOR DE FOTOS DO DEDOC	ATENDIMENTO DE FOTOS
9 indexadores	6 pesquisadores
1 supervisor	1 supervisor
5 auxiliares	

Observa-se acima o quadro de funcionários do DEDOC, bem como, os setores a que pertencem e podemos perceber que os recursos humanos estão adequados aos objetivos da instituição que é atender as necessidades informacionais dos usuários e isto é atingido na medida em que se realiza tratamento de textos, fotos e material da biblioteca por funcionários capacitados têm consciência dos objetivos institucionais porque realizam as atividades visando tais necessidades dos usuários internos.

APÊNDICE E: CARACTERIZAÇÃO DOS INDEXADORES DO DEDOC

Quadro 1: Identificação e Caracterização dos indexadores do DEDOC

1- Formação Educacional	1-Graduação em Ciências Sociais e Mestrado em Ciências Sociais
2- Experiência Profissional anterior no serviço de indexação	2-Trabalha com atividade de indexação há 10 anos
3- Outras atividades que realiza no DEDOC	3-Atualmente nenhuma outra atividade
4- Dificuldades na atividade de indexação	4-Não possui nenhuma dificuldade para realizar a atividade de indexação
5- Procedimentos de identificação de Conceitos	5-Leitura do texto lendo mais começo do texto
6- Fontes de experiência de indexação	6-Prática da atividade de indexação
7- Nível de exaustividade e especificidade exigido na indexação	7-Não é exigido e procura atribuir termos específicos

O sujeito IM não é formado em Jornalismo. Destacou na entrevista que não realizou nenhum curso de indexação, não possui dificuldades e afirma que obteve experiência com a prática. E o que contribui também para isso foi o conhecimento sobre a estrutura dos conceitos na área de Jornalismo adquirido em serviço anterior, onde realizou um treinamento para conhecer a Linguagem Documentária que a instituição havia elaborado para a indexação de notícias jornalísticas. E que na identificação de conceitos se detém mais no começo do texto. Quanto à atribuição da quantidade, especificidade e generalidade dos termos fica a cargo dos indexadores; busca sempre indexar termos específicos e não se preocupa com quantidade de termos.

Quadro 2: Identificação e Caracterização dos indexadores do DEDOC

1- Formação Educacional	1-Graduação em Comunicação Social.
2- Experiência Profissional anterior no serviço de indexação	2-Trabalha com atividade de indexação há 34 anos
3- Outras atividades que realiza no DEDOC	3-Atendimento à Pesquisa
4- Dificuldades na atividade de indexação	4-Mencionou não possuir nenhuma dificuldade
5- Procedimentos de identificação de Conceitos	5-Realiza uma leitura buscando olhar o texto todo sem ler tudo “não dá tempo”
6- Fontes de experiência de indexação	6-Prática da atividade de indexação
7- Nível de exaustividade e especificidade exigido na indexação	7-Indexa termos bem específicos dependendo do assunto

O sujeito IE, além da indexação, realiza também atendimento à pesquisa de acordo com a sua disponibilidade.

Para IE a fonte de sua experiência de indexação, com certeza, foi a realização diária da atividade de indexação ao longo desses trinta e quatro anos de trabalho.

Quanto à identificação de conceitos, IE menciona que indexa e seleciona os conceitos específicos tratados nas matérias. Entretanto, na indexação da área de esportes sob sua responsabilidade, seleciona conceitos referentes a depoimentos, fatos, recordes, e não se prende à quantidade de conceitos.

Quadro 3: Identificação e Caracterização dos indexadores do DEDOC

1- Formação Educacional	1-Formado em Jornalismo
2- Experiência Profissional anterior no serviço de indexação	2-Trabalha na instituição há 11 anos realizando a indexação
3- Outras atividades que realiza no DEDOC	3-Atendimento à Pesquisa
4- Dificuldades na atividade de indexação	4-Mencionou não possuir nenhuma dificuldade
5- Procedimentos de identificação de Conceitos	5-Realiza uma leitura visando identificar o assunto principal lendo o texto e vendo o que aborda, no seu todo, sem fazer leitura completa
6- Fontes de experiência de indexação	6-Prática da atividade de indexação
7- Nível de exaustividade e especificidade exigido na indexação	7-Não se preocupa com o nível de especificidade e de exaustividade dos conceitos porque o sistema permite cruzar buscas

Como se observa na tabela acima, o indexador IF também além da indexação realiza atendimento à pesquisa.

IF destaca que não possui nenhuma dificuldade na atividade de indexação e o que lhe incomoda, no momento da indexação, é a lentidão da Internet porque como os jornais estão disponíveis no DEDOC Eletrônico, os indexadores têm de acessar o endereço eletrônico para ter acesso aos artigos dos jornais que precisam ser indexados.

Com relação à experiência em indexação adquiriu com a prática da atividade. E no DEDOC indexa as matérias referentes à Arte, Literatura e Política Internacional. Quanto a identificação de conceitos nas matérias referentes a Artes e Literatura sempre procura assuntos que são inéditos e não se preocupa com o nível de especificidade e de exaustividade dos conceitos porque o sistema de busca permite cruzar busca pelo operador AND.

Quadro 4: Identificação e Caracterização dos indexadores do DEDOC

1- Formação Educacional	1-Graduada em Letras e está cursando Letras-Espanhol
2- Experiência Profissional anterior no serviço de indexação	2-Realiza indexação há 2 anos e sete meses
3- Outras atividades que realiza no DEDOC	3-Digitalização de pastas de recortes para o DEDOC Eletrônico
4- Dificuldades na atividade de indexação	4-Material em outro idioma
5- Procedimentos de identificação de Conceitos	5-Realiza uma leitura visando identificar o assunto principal lendo o texto e vendo o que aborda no seu todo sem fazer leitura completa
6- Fontes de experiência de indexação	6-Trabalhou em biblioteca anteriormente
7- Nível de exaustividade e especificidade exigido na indexação	7-Seleciona os assuntos na suas especificidades

Com relação ao indexador IR destacamos que também não é jornalista, novato na atividade de indexação realizando a mesma esporadicamente, pois no momento tem a função de automatizar as pastas de recortes dos jornais, indexando, portanto, somente quando solicitado por alguns dos indexadores permanentes.

Quanto às dificuldades de indexação, destaca material em outro idioma diferente do inglês e o que a irrita é a lentidão do DEDOC eletrônico.

Destaca que tem experiência em indexação porque trabalhou em biblioteca, anteriormente, e adquiriu conhecimento quanto a entradas para nomes pessoais próprios, como também, que na indexação adota-se sempre o assunto no singular, que as palavras são na sua maioria masculinas. Também o vocabulário do DEDOC ajuda muito na seleção dos conceitos de indexação porque, após selecionar o conceito acessa o vocabulário e preenche uma ficha de indexação onde digita no campo palavras-chave o conceito selecionado e, caso ocorra de o conceito atribuído for sinônimo de outro conceito que seja o controlado, o sistema avisa que o termo digitado é um termo proibido e que o termo controlado é tal. Ressalta que seleciona os conceitos específicos abordados no documento.

Quadro 5: Identificação e Caracterização dos indexadores do DEDOC

1- Formação Educacional	1-Graduado em Ciências Sociais
2- Experiência Profissional anterior no serviço de indexação	2-Não possui experiência anterior realiza indexação há 1 ano e 6 meses
3- Outras atividades que realiza no DEDOC	3-Digitalização de pastas de recortes para o DEDOC Eletrônico
4- Dificuldades na atividade de indexação?	4-Às vezes não sabe como representar determinado assunto
5- Procedimentos de identificação de Conceitos	5-Seleciona os conceitos visando a recuperação pelos usuários
6- Fontes de experiência de indexação	6-Está sendo os indexadores mais experientes do DEDOC
7- Nível de exaustividade e especificidade exigido na indexação	7-Seleciona os assuntos tratados na matéria e a quantidade de termos depende do tipo da matéria

Percebemos que IA também não é jornalista e não possui experiência anterior em indexação.

Trabalha no DEDOC há um ano e seis meses, realiza a indexação conforme é solicitada e, atualmente, está digitalizando matérias das pastas de recortes para o DEDOC eletrônico.

Quanto à identificação de conceitos, seleciona os conceitos visando a recuperação pelos usuários. E quando tem alguma dificuldade, ou não sabe como representar determinado conceito, consulta a linguagem objetivando que lhe oriente se o conceito é controlado, proibido ou não existe.

Para IA a sua fonte de experiência em indexação está sendo os indexadores mais experientes do DEDOC que a orienta quando tem dificuldades.

Na identificação de conceitos seleciona assuntos que a matéria aborda e quanto à quantidade a selecionar, depende do tipo de matéria como por exemplo, uma determinada matéria que trata de um político falando de pessoas, seleciona todas as pessoas envolvidas na matéria.

Quadro 6: Identificação e Caracterização dos indexadores do DEDOC

1- Formação Educacional	1-Formado em História e atualmente cursa mestrado em Turismo Cultural.
2- Experiência Profissional anterior no serviço de indexação	2-Não possui experiência anterior e realiza indexação há 6 meses
3- Outras atividades que realiza no DEDOC	3-Digitalização de pastas de recortes para o DEDOC Eletrônico
4- Dificuldades na atividade de indexação	4-Ressalta não ter dificuldades
5- Procedimentos de identificação de Conceitos	5-Seleciona os conceitos geral e específico visando a recuperação pelos usuários.
6- Fontes de experiência de indexação	6-Está sendo os indexadores mais experientes do DEDOC
7- Nível de exaustividade e especificidade exigido na indexação	7-Seleciona os assuntos tratados na matéria e a quantidade de termos depende do tipo da matéria

Notamos que IK relata não possuir experiência anterior no serviço de indexação e realiza a atividade há seis meses.

Atualmente está realizando a digitalização das matérias das pastas de recortes para o DEDOC eletrônico e que realiza a atividade de indexação quando solicitada; ressalta que não possui dificuldades na indexação.

Quanto à identificação de conceitos seleciona os conceitos geral e específico visando a recuperação pelos usuários.

Para IK sua experiência de indexação está sendo construída com a prática de indexação no seu dia-a-dia.

Quadro 7: Identificação e Caracterização dos indexadores do DEDOC

1- Formação Educacional	1-Graduação em Sociologia e Ciência Política
2- Experiência Profissional anterior no serviço de indexação	2-Possui experiência adquirida em outras instituições que trabalhou
3- Outras atividades que realiza no DEDOC	3-Digitalização de pastas de recortes para o DEDOC Eletrônico
4- Dificuldades na atividade de indexação	4-Não possui dificuldades
5- Procedimentos de identificação de Conceitos	5-Leitura do texto mas não todo o texto
6- Fontes de experiência de indexação	6-Prática da indexação
7- Nível de exaustividade e especificidade exigido na indexação	7-Indexa os assuntos importantes da matéria

Percebemos que IC ressalta possuir experiência anterior no serviço de indexação.

Realiza atualmente a digitalização das matérias das pastas de recortes de jornais para o DEDOC eletrônico e que indexa quando solicitado por um dos indexadores permanentes.

Quanto à fonte de experiência em indexação adquiriu com a prática da atividade.

E na identificação de conceitos realiza leitura de alguns partes do texto visando identificar os assuntos relevantes.

Com a caracterização dos indexadores do DEDOC, destacamos que os profissionais caracterizados como IM, IE e IF são indexadores permanentes, enquanto os outros indexam esporadicamente qualquer área de assunto, porque todos têm a função de digitar as pastas de recortes de artigos de jornais para o DEDOC Eletrônico, e também de ajudar qualquer um dos indexadores permanentes que esteja com serviço acumulado.

Após termos caracterizado os indexadores do DEDOC, relataremos a política de indexação da instituição na qual estão inseridos por meio da evidência do processo de indexação e escolha da linguagem.

**APÊNDICE F: TRANSCRIÇÕES DOS PROTOCOLOS VERBAIS DOS
INDEXADORES EXPERIENTES DO ARQUIVO E DO DEDOC**

VERBAIS TRANSCRITOS – INDEXADORES EXPERIENTES

Sujeito 1 – Texto 1: **LOBÃO VOLTA AO ROCK DA MELHOR QUALIDADE**

Notações das Transcrições

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

Sujeito: *então nós vamos iniciar a indexação do Caderno 2, hoje é dia 10 de julho. É o Caderno 2 é o caderno de variedades do Estadão. É o que trata de artes, então ele tem shows, festas, teatro, cinema. É a própria Editoria chama o Caderno de variedades. Então a primeira página do Caderno tá hoje, tá tratando do Lobão é o cantor. Então a primeira coisa que eu faço vou direto no colaborador para não esquecer que existe o colaborador, então eu sempre ficho primeiro o colaborador, sempre vou no fim da matéria par ver si indica a página se continuidade de página porque o Caderno 2 sempre tem essa característica quase sempre vem numa página e continua na outra no caso de hoje continua na página 9. Então eu vou na 9 trago perto da página 1 porque o assunto provavelmente é uma continuidade e é mesmo tá vendo, Lobão, então, uma vez que eu fichei o mesmo colaborador, as vezes sai colaborador diferente na continuidade então por isso que eu vou sempre no mesmo olhar primeiro o colaborador, no caso é o mesmo então eu só argolo tudo e aí eu vou dá uma lida. Então vai trata do, Lobão a gente entra em Lobão entre parêntese cantor **Lobão (Cantor)**. Aí eu vou lê o que fala a matéria. Então tá. Lobão sumiu era o comentário geral no Rio, desde que o cantor e compositor, um dos ícones da chamada Geração 80, gravou um disco de MPB, Nostalgia da Modernidade, criou a maior polêmica e desapareceu. Lobão sumiu de propósito. Foi compor e gravar um novo disco que ele classifica de MPB pesada, mas na verdade é rock da melhor qualidade, da primeira á última faixa. Sujeito: Olha eu não preciso mais ler que eu eu já sei que o resto da matéria tá falando dele, não me interessa mais nada, entendeu eu já paro de ler. É aquilo que a Cidinha te falou que é não dá pra gente lê tudo, se não a gente demora muito, dá uma geral é continuidade. Quando a gente indexa é discos assim música, quando a pessoa que é sem que está sendo indexada não é famosa no caso o Lobão vai se fácil procurar porque o jornalista quando vier recuperar a matéria ele vai procurar direto por Lobão. Se fosse uma pessoa que estiver surgindo no mercado um cantor um cantor novo, aí a gente costuma indexar também no gênero seria rock ou música comentário de disco se o cara é uma pessoa que está surgindo no mercado. Porque na hora de recuperar é mais fácil porque as vezes a pessoa o jornalista lembra do nome da pessoa. Então ele recupera pelo gênero, no caso do Lobão não precisa porque o cara é conhecido então já tô passando adiante eu coloco aqui na primeira página da matéria eu coloca página 1 e 9, porque na hora da digitação da no computador o digitador vai vê que começa na página 1 e que pra chamar também os títulos 9. Certo.*

Sujeito 1 – Texto 2: SEMINÁRIO PROVA QUE ROUPA TAMBÉM É DOCUMENTO

Notações das Transcrições

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

Sujeito: *Então terminei a página 1 passei pra página 2. Seminário prova que roupa também é documento.*
..Sujeito: *Quem escreveu foi a ba Vanessa Baroni colaboradora então eu já vou direto na Vanessa Baroni tem uma outra matéria na mesma página que é dela também então eu já argolo como colaborador o mesmo depois eu vou ver se faz parte da mesma matéria ou se é um assunto diferente* ..PS.. Sujeito: *Então continuando a matéria trata vamos ver*. Longe das passarela iluminadas, a moda é um assunto que desperta o interesse de profissionais de áreas diversas. O que atrai essas pessoas que talvez não dominem as técnicas de modelagem nem sejam hábeis costureiras, mas com certeza podem explicar por que a silhueta feminina muda tanto ao longo dos anos - é o lado documental da moda. Sujeito: *Bom, eu percebo desse primeiro parágrafo que o assunto abordado é moda então eu já vou indexar em moda se for um seminário, vamos ver quem organizou a gente costuma entrar também na pessoa que organiza os eventos quando se trata de moda.* Sob essa nova perspectiva, o psiquiatra, psicanalista e doutor em filosofia Joel Birman e a jornalista e consultora de moda Cristina Franco promovem a Conferência Visão da Figura Feminina, que vai percorrer várias capitais brasileiras e vem a São Paulo no segundo semestre. No Rio, a conferência será no dia 5 de agosto. O projeto é simples, mas rende horas de conversa: cabe a Cristina falar do lado histórico e comportamental da moda, enquanto Birman trata do filosófico e psicanalítico. Sujeito: *Tá é um projeto de moda eu vou entrar no Joel Birman e na Cristina Franco que são os dois que tão promovendo essa conferência, daí par frente vai voltar a falar deles a primeira apresentação pública começo a dar os primeiros frutos a dupla foi convidada tá não tem mais o que fichar ali.*

Sujeito 2 – Texto 3: EXECUÇÃO DE QUADRILHA AFASTA DELEGADO NO MA

Notações das Transcrições

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

Sujeito: *Outra matéria Execução de quadrilha afasta delegado no Maranhão* ..PS.. Sujeito: *Aqui tá falando O secretário de segurança do Maranhão afastou o delegado da regional de Santa Inês. por causa da execução de uma quadrilha de matadores de encomenda e assaltantes de caminhões. O crime ocorreu na quinta-feira. A quadrilha era composta por quatro bandidos. Aqui eu tô lendo mais atentamente porque eu não lembro desse caso não sei se saiu no jornal e também e envolve um problema político foi afastado um delegado então eu preciso saber se vou colocar na Polícia do Maranhão que é um delegado ou só no crime, mas pelo que eu tô entendendo aqui a Polícia que matou essa quadrilha certo não foi briga de quadrilhas então a Polícia aqui no caso é importante tá eu ficho faço dois descritores **Criminalidade Maranhão** e **a Polícia Maranhão**, tá.*

Sujeito 2 – Texto 4: ASSALTANTE DO BB DE CAMPINAS É PRESO

Notações das Transcrições

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

Sujeito: *A outra matéria Assaltante do BB de Campinas é preso. BB é o Banco do Brasil, isso aqui também é um caso que aconteceu faz tempo e a gente tem o título **Banco Assalto** tá, no caso aqui a gente fichou também Banco do Brasil tá porque foi muito importante esse crime tá ele roubou milhões e milhões lá no ..PS.. (apesar do indexador não Ter explicitado a pesquisadora anotou que nesse momento o indexador realiza a leitura do lead e posteriormente comenta:) *Oh! não é o a o assalto que eu estava pensando não foi outro assalto foi no início da semana eu achei que era um outro assalto, tá vendo como que é a gente então a gente sempre tem que fazer a leitura dinâmica sempre tem que fazer viu como como é que é tá, então aqui no caso eu posso tirar o Banco do Brasil que não era um caso que já vinha acontecendo e é sempre assim, se é se é a matéria for enfocar é der continuidade é assaltar o Banco do Brasil eu ficho o Banco do Brasil, não é aqui o caso não é importante Banco do Brasil, Banco do Brasil eu ficharia se alguma resolução do Banco, medidas econômicas, entendeu.**

Sujeito 3 – Texto 5: CONFLITO ENTRE PMS DEIXA CINCO FERIDOS

Notações das Transcrições

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

Pernambuco grevistas trocam tiros com policiais que não aderiram ao movimento em frente à sede do governo; 24 são presos. Conflito entre PMs deixa cinco feridos. Fabio Guibu (A pesquisadora anotou que após ler e identificar o nome do autor da matéria o indexador preenche o campo autor na ficha de indexação) Policiais militares de Pernambuco, em greve há sete dias, trocaram tiros ontem por duas vezes com integrantes da corporação que não aderiram ao movimento. Cinco pessoas ficaram feridas. *A greve dos policiais de fato teve um conflito é mais o conflito é dentro da greve então eu vou colocar greve policia vou colocar como termo **policia militar**.* Três oficiais, um soldado e um negociante de ferro-velho foram baleados ..PS.. Os tiroteios ocorreram em uma das avenidas mais movimentadas de Recife, a Agamenon Magalhães, e em frente ao Palácio do Campo das Princesas, desde do governo do Estado. Os policiais feridos passam bem ..PS.. Os conflitos ocorreram de manhã, horas após o governo divulgar, por meio do “Diário Oficial” do estado, a demissão de 243 grevistas e a abertura de processo para punição de outros 75. A greve é ilegal pela Justiça, por isso, o governo pode demitir ..PS.. *É protesto porque eles falam do protesto aqui e o lugar onde está sendo feita a greve que é Recife é só Recife então **Recife-Pernambuco**, já cobri está tá coberta.*

Sujeito 3 – Texto 6: GAROTINHO AJUDA PEDETISTA EM SÃO GONÇALO

Notações das Transcrições

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

Garotinho ajuda pedetista em São Gonçalo. (A pesquisadora observou: Após ler o título e o indicá-lo na ficha de indexação o indexador seleciona também o nome do autor (Felipe Werneck) da matéria e não menciona porque já o fez nas matérias anteriormente indexadas). Rio – Depois da ausência no início do segundo turno, o governador Anthony Garotinho (PDT) reapareceu esta semana na campanha à prefeitura de São Gonçalo, no Grande Rio, que tem o segundo colégio eleitoral do estado, com 540 mil eleitores. Garotinho tenta ajudar a reeleger o pedetista Edson Ezequiel, que está em segundo lugar, com de 40% dos votos, segundo pesquisa do instituto Gerp, que ouviu 500 pessoas entre os dias 21 e 22. O candidato do PMDB, Henrique Charles, lidera a disputa, com .. PS .. *Garotinho foi visitar aqui São Gonçalo para ajudar a reeleger um pedetista então é eleição municipal São Gonçalo é onde? no Rio, então é a eleição municipal Rio de Janeiro então vai pra esta pasta como é o Garotinho uma figura importante eu vou por como termo o nome dele **Garotinho** e vou por nome da cidade que no nome da pasta não vai constar só consta Rio de Janeiro-Eleição Municipal então eu vou por São Gonçalo barra Rio de Janeiro **São Gonçalo/Rio de Janeiro**. To lendo aqui o resto da matéria Garotinho participaria no fim da tarde de inauguração de obras no bairro Jardim Catarina, feitas em parceria coma prefeitura do PDT. Ele vai inaugurar obras talvez seja alguma coisa pro governo do estado. Também está prevista para esta semana uma caminhada do governador com o candidato pelos principais bairros da cidade, segundo assessores de Ezequiel. Mas eu tô vendo que são coisas irrelevantes então eu não vou por nome do Garotinho então vai ficar vou por PDT porque o candidato é do PDT então vou por PDT barra partido **PDT/Partido** pronto tá indexada.*

Sujeito 4 – Texto 7: AGORA É A VEZ DO CRISTO REDENTOR

Notações das Transcrições

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

Agora é a vez do Cristo Redentor. Depois do Pão de açúcar, prefeitura negocia cessão do Corcovado. Depois do Pão de Açúcar, agora é a vez do Corcovado. *Bondinho do Pão de açúcar. Porque eles estão querendo terminar com a gestão da companhia do bondinho do Pão de Açúcar então tentando vender também o Corcovado esta matéria é uma matéria rotineira é não teria grande importância se não fosse um monumento cartão postal do Rio de Janeiro como um monumento é quando se fala em Rio de Janeiro se fala em Pão de Açúcar se fala em Cristo Redentor essas coisas todas.* O acidente com o bondinho, ocorrido no último Sábado, quando um cabo se rompeu, foi a gota d'água para detonar o processo de cessão dos principais pontos turísticos da cidade .. PS .. *Tá vendo como eu falei eles falam dessas coisas* ⇐ acidente com bondinho ocorrido no último Sábado .. PS .. *aconteceu faz pouco tempo aqui eles falam que o bondinho deixou de funcionar faz pouco tempo.* O prefeito Luiz Paulo Conde conversou ontem pelo telefone com o ministro do planejamento, Martus Tavares. Segundo Conde, além de acertar a transferência da gestão do Pão de açúcar para a Prefeitura, o ministro tratou ainda da cessão do complexo do Corcovado para o município .. PS .. O ministério do planejamento confirmou que está em análise a transferência do corcovado para a prefeitura, embora o estudo ainda não esteja concluído. *Então eles estão tentando tirar a concessão do bondinho e passar pra uma outra companhia e assim com o corcovado também o cristo redentor também então aqui eu pego coisas normais eu ponho bonde ele vai como uma referencia* (Nesse momento o indexador digita o termo atribuído e posteriormente comenta:) *se alguém quiser saber alguma coisa sobre o que aconteceu a época do Luis Conde então eu vou colocar e eles recuperam a matéria a troca da gestão do Pão de Açúcar fala também eles querendo trocar as antigas concessões.* ..PS.. Procon poderá multar empresa de bondinho. O Procon-Rio, órgão vinculado à secretaria estadual de justiça, poderá multar a caminho Aéreo em até R\$ 3 milhões, pelo acidente ocorrido com o bondinho do Pão de Açúcar. O coordenador-geral do órgão, Átila Nunes neto, instaurou ontem processo administrativo contra a empresa, por Ter contrariado o Código de Defesa do Consumidor. Átila explicou que tantos os 70 passageiros como as pessoas que sofreram prejuízos, ao terem seus bens atingidos pelo cabo que se rompeu estão protegidos pelo código. Elas poderão ingressar com ações na justiça pedindo indenização. .. PS .. Segundo Conde, o novo edital de concessão do Pão de Açúcar será lançado pela Secretaria de fazenda assim que a União formalizar a concessão do complexo. Em julho do ano passado, a Prefeitura lançou o primeiro edital, que acabou cancelando no mês seguinte, após o Tribunal de Contas do Município determinar que fossem feitas reformulações no texto. A concessão da caminho aéreo, que durou 90 anos, está vencida desde maio de 1999. No novo edital, o prefeito resolveu suprimir a possibilidade de o futuro concessionário implantar uma ligação por bondinho entre o Morro da Urca e o Morro da babilônia (Leme). Conde alegou que a decisão visa a atender pedido feito por ambientalistas. O presidente da Agência Rio, Humberto Mota, explicou que um dos itens que a prefeitura levará em conta para escolher o futuro concessionário será a proposta que o interessado apresentar para revitalizar o complexo do pão de Açúcar. *A concessão do pão de açúcar tinha noventa anos do bondinho então renovaram até o ano que vem estão renovados então estão querendo dar uma nova estrutura a essas coisas que são coisas tradicionais na cidade do Rio de Janeiro é o Pão de Açúcar, o Corcovado o bonde de Santa Tereza coisas relativas a cultura da cidade. Essa aqui é uma matéria bastante simples não tem muito o que falar sobre ela então a pasta que ela vai ficar é Monumento-Brasil como então Cristo Redentor é um monumento e tá no título não há necessidade de colocar porque o sistema de busca realiza busca por qualquer palavra dentro da matéria inclusive do título então eu coloco o termo referente a gestão em que está sendo trocada sendo o termo CONDE, Luis Paulo e Cristo Redentor já está dentro da matéria e Monumento* (Nesse momento o indexador atribui também o termo monumento aos assuntos da matéria e comenta:) *porque ele é um monumento fechou a matéria quanto a indexação.*

Sujeito 4 - Texto 8: BAILARINA ACUSA O MARIDO ATOR DE ESPANCAMENTO E DÁ QUEIXA NA POLÍCIA.¹³

Notações das Transcrições

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

Bailarina acusa o marido ator de espancamento e dá queixa na polícia. *Bailarina acusa o 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 de anos, registrou queixa contra o marido, o ator Ademir Zanyor, de 26, na 32ª DP (Jacarépagua), no Domingo passado. Ela o acusa de tê-la espancado durante a festa de aniversário de 2 anos de Uriel, filho único do casal. *..PS.. E voce 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 nomes do agressor e da vítima então aqui vemos ⇐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 e também porque nosso sistema trabalha por palavra se você colocar O Globo Jacarépagua todas as matérias que trouxeram a palavra Jacarépagua do jornal O Globo ele vai te jogar na tela. Então por exemplo aqui eu poderia colocar agressão mas não vou colocar porque se eu digitar agressão o sistema me informa que existe uma pasta chamada agressão contra mulher* (Nesse momento o indexador digita a palavra agressão no sistema de busca) *aberta a pasta eu incluo os assuntos: agressor e vítima e não coloco agressão como assunto porque o nome da pasta já informa que houve uma agressão contra mulher.*

Sujeito 5 - Texto 9: "EU, TU, ELES" É O CANDIDATO DO BRASIL AO OSCAR.

Notações das Transcrições

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

"Eu, Tu, Eles" é o candidato do Brasil ao Oscar. *Bom essa matéria é do filme Eu, Tu, Eles indicada para o Oscar da categoria de filme estrangeiro conto. Essa matéria está no caderno cotidiano mas retranca para a ilustrada. Então quando eu for indexar a ilustrada essa matéria é a primeira que eu vou pegar.* (Nesse momento o indexador pega o caderno ilustrada para localizar a matéria e comenta:). *É mais a matéria saiu somente no cotidiano mesmo, mas no jornal eletrônico ele está na seção da ilustrada que é área dela tá. Bom primeiro passo tô registrando a URL da matéria (nesse momento o indexador digita o endereço eletrônico na ficha de indexação do sistema), segundo passo eu copio o título da matéria (nesse momento o indexador digita o título na ficha de indexação), terceiro passo é o autor Sandra Sato tá, (nesse momento o indexador digita o nome do autor da matéria) quarto passo é a data 25 de outubro de 2000, (nesse momento o indexador digita a data de publicação da matéria). E aí depois você coloca os termos tá por exemplo os termos do filme, ou peças de teatro vão para o diretor, autor isso é para a pasta autor – diretor tá Brasília – Eu, Tu, Eles, do diretor Andrucha Waddington .. PS .. então no caso é Andrucha, Waddington né. Então aqui eu vou colocar na pasta dele (nesse momento o indexador atribui a matéria a uma pasta de assunto para posteriormente extrair os assuntos) e qual filme que vem para a pasta dele? "Eu, Tu, Eles" (nesse momento o indexador digita o nome do filme na pasta) eu coloco como termo para na hora que você for procurar dentro da matéria procura colocar termos que possam ter procura no caso aqui ele (autor) e possibilidades de busca pelo título. Eu, Tu, Eles conta a história verídica de uma mulher que vive com três maridos no sertão nordestino (...) pode colocar atores que são casos de buscas como a **Regina Casé** como termo porque as vezes as pessoas podem lembrar dela e não lembrar o nome do filme aquele filme que ela fez e tal, (nesse momento o indexador inclui o nome da atriz como assunto do filme) o outro é o **Lima Duarte** (nesse momento o indexador verifica se o nome do ator Lima Duarte se encontra como termo controlado do sistema confirma que sim e o indexa) e tem um outro cara que é legal colocar como ator o terceiro marido mas nós não temos o nome dele aqui na matéria então eu não vou colocar. (Nesse momento o indexador recorre a memória tentando recordar informações que viveu ou por Ter assistido o filme ou de outros textos já indexados anteriormente sobre o filme e que lhe dão a certeza que existe um outro nome importante a indexar). Então coloquei como termo de busca a **Regina Casé**, nome do filme, o **Lima Duarte** (ator). I é só. Olha aqui Regina Casé foi à sede da Conspiração comemorar a indicação. "Esse filme merece representar o Brasil porque não traz só denúncias nem nos mostra cantando e dançando o tempo todo." .. PS .. vou ter que incluir termos A indicação do Oscar há 40 não 12 matérias sobre o filme desde seu lançamento mas o que marca essa matéria é a indicação pro Oscar. Porque poderia ser uma crítica, lançamento mas o que está marcando aqui essa matéria é a indicação então para facilitar a busca eu vou colocar termos agora como **prêmio Oscar** (nesse momento o indexador digita o termo para verificar se o mesmo é termo controlado do vocabulário do sistema e posteriormente comenta:) nosso vocabulário é um campo controlado por uma moça do sistema que ela coloca todos os termos que são controlados tá, então é para não ficar uma salada mista de colocar Oscar alguém coloca Oscar sistema americano e outro coloca, Oscar de Ficção e o outro Oscar Prêmio tá, agora o outro termo que eu vou colocar é a **indicação** porque não é a premiação quando sair matéria da premiação do Oscar da festa de premiação ele vai tá lá tal mais é uma outra matéria é um outro momento. Indicação do **prêmio Oscar** é só. (Você coloca somente indicação não fica muito geral?) Não porque ãa busca você pode cruzar os termos porque se você colocar somente indicação vai recuperar .. indicação até de efeito piolho mas sempre você vai acabar cruzando porque você vai falar há os filmes brasileiros que foram indicados ao Oscar digita indicação Oscar vêm esse daqui e vai vim aqueles anteriores, há mais é aquele do Andrucha daí você coloca Andrucha né aí Andrucha, indicação já mata tá.*

Sujeito 5 - Texto 10: BECK DIZ QUE AGORA VEM E QUER CAETANO

Notações das Transcrições

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

Beck diz que agora vem e quer Caetano. *Essa aqui é matéria grande porque inclusive tem entrevista e tudo a respeito do Beck (cantor). Então o processo procedimento é o mesmo o título Beck diz que agora vem e quer Caetano, autor Lucio Ribeiro, caderno Ilustrada, data 25 de outubro de 2000, jornal Folha .* (nesse momento o indexador preenche a ficha de indexação referentes a esses campos). *Aqui por exemplo nos campos termos é feito o controle de vocabulário dos dois (termos e nomes das pastas) então ao invés de colocar o Beck eu vou para o vocabulário que vem todos os Beck que tem.* (nesse momento o indexador digita a palavra Beck para verificar se a mesma é um termo controlado do vocabulário e qual a sua forma de entrada) *e no caso é o Beck (cantor) matou e vou até escrever isso aqui* (nesse momento o indexador digita Beck- cantor como nome da pasta de assunto do sistema). *Bom peguei a matéria O artista pop mais "furão" de shows no Brasil é atração do Rock in Rio 3, em janeiro; a Folha, conta seu amor por música brasileira e um plano para o festival. Janeiro está longe. Mas a entrevista concedida à Folha na última Segunda pode ser tomada como bom indicio. O de que o astro pop americano Beck vem mesmo tocar no Brasil, afinal. "desta vez eu vou", prometeu o próprio, ao telefone, de sua casa, em Los Angeles. O cantor é uma das atrações mais festejadas do Rock in Rio 3. Seu nome ocupa a lista do palco principal do dia 13 de janeiro. O que chama atenção nessa matéria é que ele vai vir no Rock in Rio 3 é uma entrevista dele é sobre um show então coloco como termos **entrevista e show** é genérico (nesse momento o indexador realiza uma "olhadela" pelo texto procurando encontrar um termo mais específico que represente a matéria) *mas é só esses mesmos. (pesquisadora: e no caso você não entra pelo nome dele como assunto porque pode ser recuperado pelo nome da pasta?). Pode se você der Beck no campo de pesquisa já vem essa matéria e ainda você pode recuperar por pasta, palavra-chave todas as palavras do texto, por publicação (jornal). Tem que pensar como o pesquisador vai procurar aquilo? Como facilitar a procura para o pesquisador porque para você tá com a matéria na mão tá fácil, tá vendo a matéria tal mas você pensa daqui um mês como você procuraria essa matéria? Então tem que se procurar facilitar com a atribuição de termos para o cara para facilitar a recuperação dele. Então o lance é esse. Porque para você tá fácil tá claro mas é aquela coisa do telefone sem fio onde o último lá pode estar pensando outra coisa. Então você tem que imaginar como o pesquisador vai pesquisar determinada matéria e procurar facilitar a busca.**

**APÊNDICE G: TRANSCRIÇÕES DOS PROTOCOLOS VERBAIS DOS
INDEXADORES INEXPERIENTES DO DEDOC**

Transcrições dos Protocolos Verbais – Avaliação do Modelo de Leitura

Sujeito Ana – sem uso do modelo

Título do texto: **Cientistas Desenvolvem Moléculas para Superar Resistência Bacteriana**

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

*Então vem na tela e costuma digitar o título a gente recorta e cola nesse caso Cientistas desenvolvem moléculas para superar resitência bacteriana. No caso o autor é Herton Escobar a gente marca da mesma maneira. A publicação é o Estado de São Paulo a gente escreve, a data 26, de julho de 2001, aí a pasta ficaria assim Biologia/Bactéria. Então Biologia/Bactéria seria a pasta. Então essa matéria vai para Biologia/Bactéria. Aí seriam as palavras chaves. Aí a gente eu pelo menos costumo ler a matéria toda. Bactérias resistentes a antibióticos podem estar com os dias contados. Cientistas da Califórnia sintetizaram uma nova classe de moléculas capazes de fazer buracos na parede celular de bactérias, matando os micróbios sem danificar tecidos saudáveis. Em experimentos com ratos (PS) <= aniquilaram em poucos minutos grupos (PS) (Nesse instante o indexador digita o termo) Remédio, Experiência não tem eu vou procurar mais então eu procuro termos próximos porque ..PS.. . (Nesse momento o sujeito digita novos termos **Experiencia Cientifica**, instituto de pesquisa Scripps, vou verificar se tem o nome de desse instituto de pesquisa (Nesse momento o sujeito digita o nome do instituto). Ah, não tem então eu vou anotar o nome todo. (Nesse momento digita o instituto como termo) **Instituto de Pesquisa Scripps**. “São agentes extremamente rápidos “ (PS) “As células bacterianas foram destruídas em menos de cinco minutos (PS) Centenas de pa rá r a ra ra. Ah eu só colocaria isso.*

Título do texto: **Cientistas Desenvolvem Moléculas para Superar Resistência Bacteriana**

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

*É bom a primeira coisa que eu olho aqui é o chapéu né que é o indica a editoria na verdade aqui ele está fora da editoria dele. É difícil uma boa coisa porque ela está fora do caderno dela da área. Cientistas desenvolvem moléculas para superar resistência bacteriana. É peptídios sintéticos fazem buracos na parede celular dos micróbios. ..PS.. antibióticos podem estar com os dias contados. Desenvolvimento de antibióticos. ..PS.. Cientistas da Califórnia sintetizaram uma nova classe de moléculas capazes de fazer buracos na parede celular de bactérias, matando os micróbios sem danificar tecidos saudáveis. É desenvolvimento de antibióticos provavelmente é alguma coisa relativa a medicamento. Então deixa eu procurar aqui se tem uma pasta Medicamento/Antibiótico (Nesse momento o indexador realiza busca no sistema para verificar se já existe uma pasta com tal título Medicamento/Antibiótico). Existe eu vou mandar para essa pasta. Agora eu vou ver se tem alguma coisa se me interessa para colocar de palavra-chave. Em experimentos com ratos, esses agentes aniquilaram em poucos minutos grupos de cepas resistentes de Staphylococcus aureus, bactéria responsável pela maioria dos casos de infecção hospitalar. Os cientistas do Instituto de Pesquisa Scripps, em La Jolla ..PS.. É talvez eu coloque o instituto mas não eu acho que não é importante o instituto mas vou ter que consultar lá (Nesse momento o indexador consulta o sistema para verificar se já existe o nome do instituto como termo controlado de indexação e confirma que não existe o nome do instituto). Ai então eu colocaria como palavra-chave somente **infecção hospitalar** e também deixa eu ver aqui (Nesse momento o indexador verifica no sistema se existe a palavra-chave medicamento-antibiótico e confirma que existe) então eu coloco **medicamento-antibiótico** porque o central aqui é o desenvolvimento do antibiótico mesmo né. Ele é voltado para as bactérias que causam infecção hospitalar eu colocaria isso aí mesmo*

Sujeito Clara: com uso do modelo texto

Título do texto: Cientistas Desenvolvem Moléculas para Superar Resistência Bacteriana

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

Cientistas desenvolvem moléculas para superar resistência bacteriana. É esse já é difícil. Vamos fazer todo processo de indexação depois nós vamos se preocupar com os termos principais. (Nesse momento o indexador consulta o texto e retira os dados descritivos para preenchimento da ficha de indexação como: título, nome da publicação, data, autor, página). Bactérias resistentes ..PS.. Esse negócio de ciência eu é uma das coisas mais difícil que eu acho porque se tem aquelas pegadinhas você não sabe se é bactéria se é genética. Sintetizaram moléculas capazes de fazer buracos na parede celular de bactérias ..PS.. Agora vou tentar responder ao questionário quem sabe ajuda. Então tá são bactérias. Quem estão estudando as bactérias? São os cientistas, da Califórnia. Tempo deve ser atualmente. O que eles querem é uma nova classe de moléculas para poder combater as bactérias. Por que? pra acabar com as bactérias. Para matar as bactérias os micróbios. Sem danificar os tecidos saudáveis. Não é bactéria é micróbio. Esse aqui é uma tema difícil. Fazendo buraco na parede celular de bactérias, matando os micróbios ..PS.. Então fica uma coisa estranha. Se quer acabar com as bactérias ou com os micróbios? Ah! Então outra coisa infecção hospitalar. Então a pasta é bactéria porque eles querem alguma coisa que acabem com as bactérias. Acho que bactéria não tem esse termo (Nesse momento o indexador consulta o sistema de informação por meio da digitação do termo bactéria para verificar se o mesmo é um termo autorizado). Tem Biologia/Bactéria.. Uma classe de moléculas ..PS.. Não tinha por exemplo onde que era Califórnia mas no Os cientistas do Instituto de pesquisa Scripps, em La Jolla. Bom eu vou como eu já acertei a pasta a minha compreensão do texto do que faz ou do que deixa de fazer. É fica em segundo plano porque eu não sou cientista eu acho tenho quase certeza que eu acertei a pasta. Onde o tema vai para onde o tema tem que ir. Agora eu só tenho que colocar nesse caso da bactéria eu nem colocaria termo porque já tá no nome da pasta já vem como termo. Bactéria já funciona como termo e a pesquisa é por palavra livre se o cara pesquisar o nome do instituto ele vai por exemplo no sistema e digita Instituto de pesquisa Scripps ele recupera a matéria. Esse aqui é difícil e como se pode recuperar por outras maneiras sem ser somente palavra chave eu salvaria sem termos. Mas vamos ver com seu roteiro se consigo colocar palavras-chave. Ação: é combater bactéria, Agente: Instituto de pesquisa Scripps, Métodos: <= Criar uma nova classe de moléculas capazes de fazer buracos na parede celular de bactérias. Local: bactéria eu acho não local no caso seria ah! Tá a cidade La Jolla/Califórnia, Causa efeito: é aniquilar bactérias resistentes que causam infecção hospitalar. Por que: os motivos dessa ação é aniquilar as bactérias mais resistentes porque causam infecção hospitalar. É só isso.

ANEXO A: EXEMPLOS DA LINGUAGEM DO ARQUIVO

SÃO PAULO (cid.)

Genal

- Agitações
- Alameda...
- Anel Rodoviário (v. Anel Rodoviário)
- Aniversário
- Área Comercial (v. Área Comercial)
- Área Industrial (v. Área Industrial)
- Área Residencial (v. Área Residencial)
- Área Verde (v. São Paulo (cid.) - Parque, Jardim e Área Verde)
- Arquitetura
- Avenida
- Avenida ...
- Bairro
- Bairro de ...
- Bairro Jardins
- Bairro Residencial (v. Área Residencial)
- Chácara (v. Chácara)
- Chácara Flora (v. São Paulo (cid.) - Bairro da Chácara Flora)
- Centro
- Córregos (v. Córrego)
- Economia e Finanças
- Elevado Presidente Artur da Costa e Silva (v. São Paulo (cid.) - Elevado Costa e Silva)
- Esgotos (v. Saneamento)
- Estrada ... (de Parelheiros, das Campinas)
- Estrada de Rodagem (v. Estrada da Rodagem - São Paulo (cid.))
- Galeria
- Galeria Protes Mela
- História
- Inundação (v. Chuva e Inundação - São Paulo (cid.))

of. de. v. 58

(11)

SÃO PAULO (est.)

Gerai

- ~~Alta Paulista (v. Alta Paulista)~~
- ~~Alta Sorocabana (v. Alta Sorocabana)~~
- Corredor de Integração Oeste (v. Corredor Oeste (via expressa))
- Economia e Finanças
- Eleição
- História
- Importação de Equipamentos Israelenses
- Interior
- Limites - Outros Estados
- Mudança da Capital
- Município (v. Município - São Paulo (est.))
- Política e Governo
- Região Alta Paulista
- Região Nordeste
- Região Noroeste
- Região Alta Sorocabana
- Região Sudeste
- Região Sul *Sudeste*

v.t.: Baixada Santista

Litoral

Pontal do Paranapanema

(2)

Abastecimento (alimentos)

- Brasil
- São Paulo (cid.)
- São Paulo (est.)

Acidente

- Brasil
- Estados, Territórios e Brasília
- São Paulo (cid.)

Acidente Aéreo

- Brasil

Acidente Doméstico

Acidente Ecológico

Acidente Especial

Acidente Ferroviário

Acidente Fluvial

Acidente Marítimo

Acidente Nuclear

Acidente Rodoviário

- Brasil (para os Estados/ exceto São Paulo) —D
- São Paulo (est.)

Acidente de Trabalho

Acidente de Trânsito

- Brasil
- São Paulo (cid.)

Açúcar

- Brasil

- Jardim da Aclimação (v. Parque da Aclimação)
- Ladeira ...
- Largo ...
- Loteamento (v. Terrano e Loteamento - São Paulo (cid.))
- Monumento (v. Monumento ...)
- Museu (v. Museu ...)
- Parque (v. Parque ...)
- Parque, Jardim e Área Verde
- Pátio do Colégio
- Pavimentação (v. Pavimentação)
- Periferia
- Política e Governo
- Ponte
- Praça
- Praça ...
- Regional de ... (v. Administração Regional de ...)
- Rua
- Rua ...
- Sistema Viário
- Subúrbio (v. São Paulo (cid.) - Periferia)
- Terrano (v. Terreno e Loteamento - São Paulo (cid.))
- Urbanismo (v. Urbanismo - São Paulo (cid.))
- Vale do Anhangabaú
- Via Arterial
- Viaduto (v. Viaduto ...)
- Via Elevada Presidente Artur da Costa e Silva (v. São Paulo (cid.) -
Elevado Costa e Silva)
- Via Expressa
- Via Expressa ...
- Vila
- Vila ... (v. São Paulo (cid.) - Bairro de Vila ...)
- Zona Comercial

(13)

**ANEXO B : CÓPIA DO TEXTO-BASE UTILIZADO PELOS
INDEXADORES EXPERIENTES**

DATA DE PUBLICAÇÃO: 10/07/1997 E 25/10/2000

*** Grifos presentes nos textos são nossos utilizados no momento da transcrição.**

*** Segue texto utilizado na aplicação do modelo experimental de leitura**

ANEXO C: “INSTRUÇÕES AOS SUJEITOS”

TEXTO ADAPTADO DE NARDI (1993)

Instruções aos sujeitos

O que vamos fazer agora é uma atividade de familiarização com os procedimentos de coleta de dados que serão usados em nossa pesquisa.

É muito simples e natural! Tudo o que você tem a fazer é ler o texto que lhe será entregue, da mesma maneira que você costuma ler outro texto para indexá-lo.

Durante a indexação você precisa "pensar alto". Tente imaginar você sozinho num recinto (seu quarto, por exemplo) lendo um texto que você precisa indexar. Em situações como essa, já não lhe ocorreu começar a falar espontaneamente em voz alta, exteriorizando seus raciocínios, seus mecanismos mentais, para conseguir indexar? Nesse processo, o indivíduo "pensa em voz alta", exteriorizando espontânea e quase inconscientemente seus questionamentos, suas "buscas" de soluções para eventuais problemas que ocorrem durante a indexação, sua maneira singular de extrair descritores de um texto.

Um exemplo bastante claro de exteriorização do pensamento durante a realização de uma tarefa (e que ocorre com a maioria das pessoas) é o "pensar alto" espontâneo durante a realização de um problema matemático. Dá para você ter uma idéia de como funciona a técnica do "PENSAR EM VOZ ALTA"? Corresponde à exteriorização de sua fala interna, seu pensamento.

Agora, a tarefa que você vai realizar é a indexação do texto que vamos lhe apresentar...e, por favor, lembre-se que é preciso "pensar alto" durante toda leitura.

O que esperamos que você faça é "pensar alto" durante toda indexação.

Tente esquecer a presença da pesquisadora. Ela estará apenas para lembrar-lhe que é preciso "pensar alto" o tempo todo. Ignore o gravador. Tente agir tão naturalmente quanto possível, como se você estivesse só.

Atente apenas para a indexação do texto.

ANEXO D: EXEMPLOS DA FICHA DE INDEXAÇÃO DO DEDOC

Envie esta notícia por e-mail para assinantes do UOL ou da Folha



[Texto Anterior](#) | [Próximo Texto](#) | [Índice](#)

PERNAMBUCO

Grevistas trocam tiros com policiais que não aderiram ao movimento em frente à sede do governo; 24 são presos
Conflito entre PMs deixa cinco feridos

FÁBIO GUIBU

DA AGÊNCIA FOLHA, EM RECIFE

Policiais militares de Pernambuco, em greve há sete dias,

Id:	145993	MATERIAS VINCULADAS:	(110º de 200)	Indexar
Assunto:	Conflito entre PMs deixa cinco feridos			Incluir
Assunto:	FÁBIO GUIBU			Excluir
Assunto:	Folha de S. Paulo	Data:	25 10 2000	Salvar
Assunto:	Visualizar			Ver
Assunto:	Pesquisar			Base
Assunto:	POLÍCIA MILITAR RECIFE /PE/			
Assunto:	GREVE-BRASIL-POLÍCIA =			121
Assunto:	http://www.uol.com.br/fsp/brasil/fc2510200011.htm			Arg
Assunto:				

Quarta-feira, 25 de outubro de 2000

Garotinho ajuda pedetista em São Gonçalo

Apoio do governador do Rio inibiu participação da legenda na campanha de Ezequiel

FELIPE WERNECK

RIO - Depois da ausência no início do segundo turno, o governador Anthony Garotinho (PDT) reapareceu esta semana na campanha à prefeitura de São Gonçalo, no Grande Rio, que tem o segundo colégio eleitoral do Estado, com 540 mil eleitores. Garotinho tenta ajudar a reeleger o pedetista Edson Ezequiel, que está em segundo lugar, com de 40% dos votos, segundo pesquisa do Instituto Gerp, que ouviu 500 pessoas entre os dias 21 e 22.

O candidato do PMDR Henri Charles lidera a disputa com 44%. O resultado

Page	145843		
Título	Garotinho ajuda pedetista em São Gonçalo		
Autor	FELIPE WERNECK		
Publicar	O Estado de S. Paulo	Data	25 10 2000
Seção	Especial		
Objeto			
Palavras-chave	ELEIÇÃO-PREFEITURA-RIO DE JANEIRO ==		
URL	http://www.estado.com.br/editorias/2000/10/25/pol679.html		
Formato			

(30 de 1

Incluir

Voltar

SAO GONÇ

☒ Volta à
princípio



SABELA

Independent Ad Serving, Tracking & Analysis



Rio

Bailarina acusa marido ator e registra queixa de agressão Ana Carolina Torres

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

Casos-etc
Aqui tem. Clique.

[Rio da Meada](#)
[Guia Rio](#)
[Poder Público](#)
[Bairros](#)
[Qual é a boa](#)
[Praias](#)

Rec: 145837	(30 de 300)	Indexar
Assunto: Bailarina acusa marido ator e registra queixa de agressão	Incluir	Excluir
Assunto:	Salvar	Excluir
Publico: O Globo	Seleção: 25 10 2000	Verificar
Assunto:	Verificar	Excluir
Assunto:	Verificar	Excluir
Assunto:	Verificar	Excluir
Assunto: AGRESSÃO CONTRA A MULHER	Verificar	Excluir
Assunto: http://www.oglobo.com.br/rio/RIO065.htm	Verificar	Excluir
Assunto:	Verificar	Excluir

BESSA, LUCIANA
ZANYOR, ADEMIR

☒ Mais InformaçõesUm ângulo diferente da cidade**Agora é a vez do Cristo Redentor**
Bernardo de la Peña e Selma Schmidt

Depois do Pão de Açúcar, agora é a vez do Corcovado. O acidente com o bondinho, ocorrido no último sábado, quando um cabo se rompeu, foi a gota d'água para detonar o processo de cessão dos principais pontos turísticos da cidade. O prefeito Luiz Paulo Conde conversou ontem pelo telefone com o ministro do Planejamento, Martus Tavares. Segundo Conde, além de acertar a transferência da gestão do Pão de Açúcar para a Prefeitura, o ministro tratou da cessão do complexo do Corcovado para o município.



Fio da
Guia R
Poder
Bairro:
Qual é
Praias
Cokuna
Pesso:
Ricard
Ao pé
Paulo
Obituá
Links i
E-mail

Num:	145835	
Título:	Agora é a vez do Cristo Redentor	
Autor:	Bernardo de la Peña e Selma Schmidt	
Publica:	O Globo	Data: 25/10/2000
Seção:		Expediente:
Class:		
Fonte:	MONUMENTO-BRASIL	[?]
URL:	http://www.oglobo.com.br/rio/RID10.htm	Arg
Tópicos:		

(33º de 200)

Incluir

Alterar

Voltar

Pesquisar

CONDE, Luiz Paulo

Comissão de cineastas anunciou a escolha ontem "Eu Tu Eles" é indicado a concorrer a uma vaga na disputa pelo Oscar

FREI-LANCE PARA A FOLHA
DA SUCURSAL DO RIO

"Eu Tu Eles", de Andrucha Waddington, foi escolhido ontem, por unanimidade, para concorrer a uma vaga na disputa do prêmio de melhor filme estrangeiro no Oscar 2001.

A indicação foi feita por uma comissão formada pelos cineastas Helvécio Ratton, Sandra Werneck, Carlos Diegues, Nelson Pereira dos Santos e Arnaldo Jabor.

Num:	145876	(00º de 200)	Indexar
Título:	"Eu Tu Eles" é indicado a concorrer a uma vaga na disputa pelo Oscar		
Assunto:			
Publicar:	Folha de S.Paulo	Data: 25 10 2000	Visualizar
Exibir:			
Class:			
Perfil:	WADDINGTON, ANDRUCHA ""		
Endereço:	http://www.uol.com.br/fsp/Ilustrad/tq2510200029.htm		
Visual:			

Incluir Excluir Salvar

Voltar Pesquisar Base

CASE, Regina
Eu tu eles /FILME/
INDICAÇÃO
LIMA DUARTE /ATOR/
PRÊMIO OSCAR

[Texto Anterior](#) | [Próximo Texto](#) | [Índice](#)

O artista pop mais "fuzão" do shows no Brasil é uma das atrações do Rock in Rio 3, em janeiro; a Folha, conta seu amor por música brasileira e sobre um plano para o festival

Beck diz que agora vem e quer Caetano

LÚCIO RIBEIRO
DA REPORTAGEM LOCAL

Janeiro está longe. Mas a entrevista concedida à Folha na última

Id:	145884	(02 de 200)	Indexar
Título:	Beck diz que agora vem e quer Caetano	Incluir	Excluir
Assunto:	LÚCIO RIBEIRO	Salvar	
Fonte:	Folha de S.Paulo	25/10/2000	Visualizar
Assunto:			Votar
Assunto:			Resquisar
Assunto:			Base
Assunto:	BECK /CANTOR/ "	124	
Assunto:	http://www.uol.com.br/fsp/illustrad/tq2510200006.htm	Arq	
Assunto:			

ENTREVISTA
SHOW

**ANEXO E : CÓPIA DO TEXTO-BASE UTILIZADO PELOS
INDEXADORES INEXPERIENTES**

DATA DE PUBLICAÇÃO: 26/07/2001

*** Segue texto utilizado no teste para orientação sobre o funcionamento do modelo.**

Cientistas desenvolvem moléculas para superar resistência bacteriana

*Peptídios sintéticos
fazem buracos na
parede celular dos
micróbios*

HERTON ESCOBAR

Bactérias resistentes a antibióticos podem estar com os dias contados. Cientistas da Califórnia sintetizaram uma nova classe de moléculas capazes de fazer buracos na parede celular de bactérias, matando os micróbios sem danificar tecidos saudáveis. Em experimentos com ratos, esses agentes aniquilaram em poucos minutos grupos de cepas resistentes de *Staphylococcus aureus*, bactéria responsável pela maioria dos casos de infecção hospitalar. Os ratos que não receberam o tratamento morreram em menos de 48 horas.

Os cientistas do Instituto de Pesquisa Scripps, em La Jolla, criaram uma série de peptídios – cadeias de aminoácidos que são a unidade básica das proteínas – em forma de anéis cilíndricos. Em contato com a membrana da bactéria, esses peptídios formam nanotubos, criando poros por onde vaza o líquido intracelular. “São agentes extremamente rápidos”, afirma o químico Reza Gadhiri, que coordena a pesquisa, publicada

na revista *Nature*. “As células bacterianas foram destruídas em menos de cinco minutos, com uma única injeção.”

Centenas de peptídios com propriedades antibióticas já foram identificados em plantas e animais, mas as moléculas naturais, muito grandes e pouco maleáveis, são de difícil aplicação em medicamentos.

A vantagem do peptídio sintético é que sua cadeia de aminoácidos pode ser manipulada para assumir diversas combinações, oferecendo aos cientistas um arsenal de incrível versatilidade. Os antibióticos convencionais precisam encaixar-se em um ponto específico da bactéria, que pode mutar sua estrutura como forma de defesa. Já os agentes peptídicos são atraídos pelas moléculas que compõem a própria membrana da bactéria. “Com tempo, poderemos sintetizar peptídios contra uma grande variedade de microrganismos”, afirma Ghadiri. Em experimentos *in vitro*, os cientistas conseguiram também destruir cepas da bactéria *E. coli*.

Não há garantia de que a técnica vai funcionar em seres humanos, mas as expectativas são boas. “Tudo indica que esses serão os medicamentos do futuro”, aposta Gadhiri. “Os resultados até agora foram todos positivos.”

BREVES

BNDES doa R\$ 1,4 mi a redes voluntárias

O Banco Nacional de Desenvolvimento Econômico e Social (BNDES) vai aplicar R\$ 1,4 milhão em apoio à ampliação da rede de voluntariado no Brasil. A proposta é fortalecer as entidades que fazem trabalho voluntário.

Ibama cria Centro Nacional de Orquídeas

BRASÍLIA – O presidente do Ibama, Hamilton Casara, assinou, ontem, portaria criando o Centro Nacional de Orquídeas. Plantas Medicinais, Ornamentais e Aromáticas. O centro contará com biblioteca e salas de estudos.

Racismo contamina justiça, diz Anistia

RODRIGO PENA MAJELLA

DA REDAÇÃO

Em praticamente todo o planeta, embora em graus diversos, decisões dos sistemas de justiça e procedimentos das forças de segurança do Estado têm desrespeitado os direitos humanos em relação de cor, raça, etnia e nacionalidade, afirma um relatório divulgado em Londres, pela

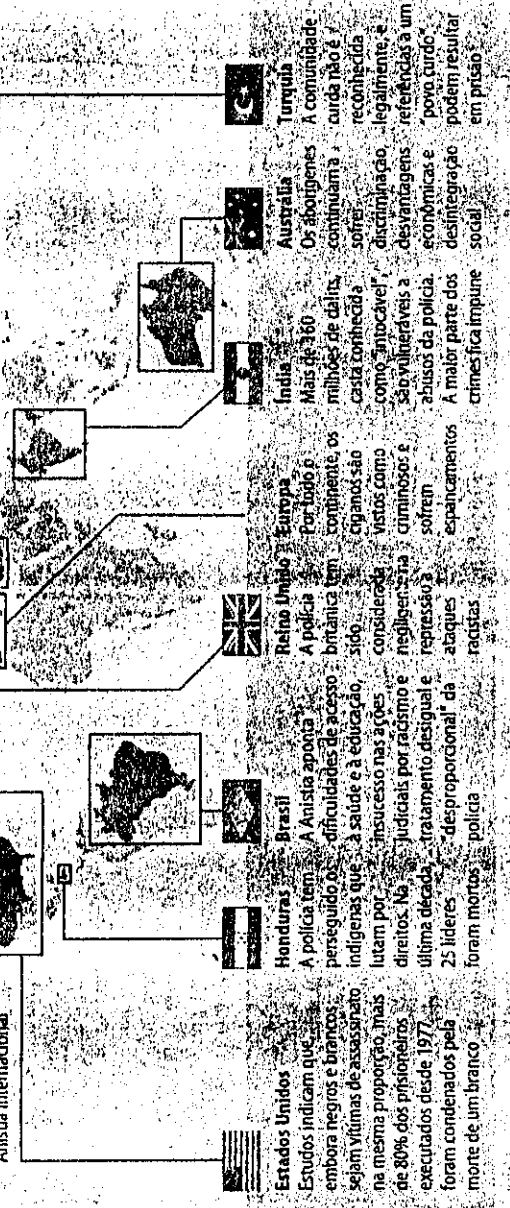
Amnistia Internacional, sobre as violações aos direitos humanos baseadas no racismo, xenofobia e limpeza étnica. O relatório afirma que, embora não receba destaque, são os abusos que ocorrem todos os dias no contexto da administração da justiça parcial ou totalmente devidos ao racismo que a organização de direitos humanos, "A maior dos países, o racismo só pode ser identificado na observação dos padrões de prisão, condenação e sentença em relação à raça do réu da vítima ou dos responsáveis pela justiça".

O documento, que pretende ser uma contribuição da organização à Conferência Mundial Contra o Racismo, Discriminação Racial, Xenofobia e Intolerância Correlata (na África do Sul, de 31 de agosto a 7 de setembro), aponta abusos de países ricos e pobres.

Segundo estudos analisados pela Anistia, nos Estados Unidos a raça é um fator significativo na sentença à pena capital. O número de vítimas de assassinato é praticamente igual entre brancos e negros, contudo mais de 80% dos prisioneiros executados no país

MAPA DO RACISMO

Principais casos apontados pela Anistia Internacional



Editoria de Arte/Folha Imagem

Entidade aponta discriminação e abusos no Brasil

DA REDAÇÃO

O Brasil aparece no terceiro capítulo do relatório da Anistia Internacional, intitulado "Impunidade e Falta de Proteção do Estado".

O documento diz que o Estado muitas vezes faz "vista grossa aos abusos cometidos por seus agentes" e aponta os frequentes ataques às comunidades indígenas, a impunidade, a predominância de negros presos e a grande disparidade social e econômica entre as raças. "Apesar da falta de dados estatísticos, não há dúvidas de que os negros são desproporcionalmente visados pelas forças de segurança e de que, de modo rotineiro, eles não recebem os benefícios concedidos aos suspeitos brancos de classe média", afirma. Diz ainda que as ações movidas com base na lei anti-racismo de 1997 não têm tido muito

longe nos tribunais. Procurado pela Folha, Gilberto Vergne Saboia, responsável pela Secretaria de Estado dos Direitos Humanos (órgão do Ministério da Justiça), disse que não faria comentários antes de ler o relatório. (RPA)

formam outro dos principais grupos sujeitos à discriminação.

Na Austrália, um dos países com os melhores índices de desenvolvimento, os aborígenes continuam a ser marginalizados e vistos por alguns como seres "primitivos". Policiais em Honduras perseguem índios envolvidos em campanhas por direitos — 25 líderes foram mortos em dez anos.

A Anistia pede que todos os governos adotem planos de ação para combater a discriminação e incluam medidas específicas relativas ao sistema de justiça. Sublinha, entretanto, que "a responsabilidade pelo combate ao racismo se estende a todos".

onde jovens brancos e de origem asiática têm se enfrentado em violentos choques, a polícia foi considerada negligente na resposta dada a ataques racistas.

A Turquia, que tenta melhorar sua imagem para entrar na União Europeia, não reconhece legalmente a minoria curda e ameaça com prisão qualquer um que usar a expressão "povo curdo".

Apesar das leis contra a discriminação, membros de castas continuam a ser atacados. Na Índia, abusos contra os dalits (conhecidos também como "intocáveis") raramente são punidos ou sequer investigados.

Os 300 milhões de indígenas

ONU em setembro, os EUA reconheceram que "membros de minorias (especialmente negros e hispânicos) podem ser desproporcionalmente submetidos a tratamento adverso ao longo do processo judicial criminal", e algumas medidas foram tomadas. Os negros continuam, porém, sob maior risco de ir para a prisão ou para o corredor da morte.

Na Europa, ainda é grande o preconceito. Os ciganos são comumente tachados de criminosos e alvo de espantamentos, e jovens de aparência não-europeia são a maioria das vítimas de maus-tratos policiais na França e na Alemanha. No Reino Unido,

desde 1977 foram condenados pela morte de uma pessoa branca.

Negros e outras minorias nos EUA também possuem desproporcionais taxas de "encarceramento". Embora quase 1 em cada 3 americanos pertença a uma minoria, eles são 60% da população de 1,7 milhão de presos. Cerca de sete vez mais negros são detidos do que brancos, e um terço de todos os jovens negros está detido na prisão, em liberdade condicional ou sob investigações. "A Anistia diz ainda que muitos réus que não falam inglês não recebem auxílio de intérpretes, impedindo que tenham um julgamento justo. Num relatório sobre a segurança