

Estudo de fatores interferentes no processo de análise de assunto¹

Madalena Martins Lopes Naves²

No contexto da ciência da informação destaca-se a Análise de assunto como uma das etapas mais importantes da indexação. O responsável pela atividade é o indexador, que dá início ao processo com subjetividade, conhecimento prévio, formação acadêmica e experiência. O objetivo da pesquisa é mostrar que vários fatores interferem em Análise de assunto, tendo sido feito um estudo exploratório com estudo de caso de sete indexadores separados por grupos, por tempo de experiência em indexação. Os resultados mostram que a questão primordial para o desenvolvimento do processo é a compreensão do texto, e que esse processo não deve ser visto isoladamente dentro da ciência da informação, havendo necessidade de se buscar contribuições de outras áreas interdisciplinares.

Palavras-chave: Análise de assunto, Indexação, Indexador

Recebido em 18/04/2001 - Aceito para publicação em: 22/05/2001.

189

Introdução

O aumento da produção científica aliado à crescente interdisciplinaridade entre as áreas do conhecimento que vem ocorrendo nas últimas décadas tem tornado o trabalho de organização e tratamento da informação cada vez mais árduo e complexo.

No campo da ciência da informação, a ação de identificar e descrever o conteúdo de um documento ocorre na *indexação*, na *classificação* ou na *catalogação de assuntos*. A indexação é considerada como processo básico na recuperação da informação e, dentro desse processo, ocorre a compreensão e a interpretação do conteúdo informativo do documento, através de atividade chamada *análise de assunto*.

A complexidade da análise de assunto, sendo esta a primeira etapa da indexação, vem sendo, há muito, apontada e discutida por professores e especialistas da área. Vários são os motivos que levam a esse julgamento em relação a tal disciplina, como a confusão conceitual em que está envolvida, por sofrer interferência de vários fatores, muitas vezes ignorados por quem a executa, e por ser difícil de ser realizada e ensinada. Pode-se, mesmo, verificar que há carência de pesquisas teóricas que se aprofundem sobre o tema, tendo-se dado maior ênfase aos estudos de linguagens de indexação, do que à própria *análise de assunto*.

A definição do tema, objeto desta pesquisa, baseou-se na idéia de que o processo de análise de assunto deve ser independente do vocabulário controlado utilizado no sistema de recuperação da informação e de que a preocupação para com

¹ Artigo extraído da tese de doutorado defendida junto ao Programa de Pós-Graduação em Ciência da Informação da Escola de Ciência da Informação da UFMG.

² Professora Doutora da Escola de Ciência da Informação da UFMG. Email: madalena@eci.ufmg.br

as necessidades do usuário, embora muito importante, só deve ser atendida numa etapa posterior à determinação do assunto do documento. Acredita-se que haja diferença entre os fatores que interferem na análise de assunto de textos de diferentes áreas do conhecimento. Trata-se de constatação surgida gradualmente durante a prática de indexar e ao longo do ensino de indexação, tendo sido referendada pelas leituras feitas no decorrer da revisão de literatura.

Dois são os pressupostos que orientaram a pesquisa:

a) Os fatores interferentes no processo de análise de assunto variam de acordo com o tempo de experiência dos indexadores no trabalho de indexação.

b) A análise de assunto ocorre de forma diferenciada nas diversas áreas do conhecimento, pelas especificidades de cada área, como o tipo de textos produzidos e da terminologia usada pelos especialistas, e a existência, ou não, de uma taxonomia de assuntos, dentre outras.

O objetivo geral proposto para nortear o estudo exploratório da identificação de fatores interferentes no processo de análise de assunto foi o de tentar mostrar que, durante o trabalho de indexadores, diversos fatores interferem no processo de análise de assunto. Como objetivos específicos, pretendeu-se: aprofundar nos estudos teóricos do processo de análise de assunto para tentar delinear melhor certos aspectos que dificultam o entendimento da área; levantar características e opiniões de indexadores sobre análise de assunto; observar o comportamento de indexadores durante o processo em que analisam textos para definição de assuntos; identificar e comparar fatores que interferem no processo de trabalho de indexadores com tempos diversos de experiência em indexação; conhecer a formação acadêmica dos indexadores, e comparar os fatores que interferem nas análises de assunto de textos nas disciplinas sociologia e botânica.

Análise de assunto é um tema que pode ser estudado sob vários aspectos. Neste trabalho, ele é abordado do ponto de vista do indexador, por se acreditar que todo o processo seja iniciado por esse profissional, no momento da leitura do texto e da definição do assunto do documento.

Revisão de literatura

A revisão de literatura feita para esta pesquisa abordou os seguintes temas: a importância do papel do indexador, o processo de análise de assunto como primeira etapa da indexação de assuntos e a interdisciplinaridade em análise de assunto.

O indexador

O indexador é definido como responsável por todo o processo de análise de assunto, tendo a sua figura ocupado um papel de destaque neste trabalho, pois a ele é creditado, em grande parte, o sucesso ou o insucesso de um sistema de recuperação da informação.

Algumas considerações devem ser feitas, inicialmente, sobre o profissional da informação. BARBOSA (1998) acredita que não há uma definição universalmente aceita a respeito de quem se constitui o profissional da informação, podendo este ser um pesquisador, um engenheiro, projetista, desenhista industrial, gerente, contador. Cita STRASSMAN, que adota uma conceituação abrangente e define o profissional da

informação como, simplesmente, aquele que trabalha com informação, em vez de com objetos. O nome *profissional da informação* é cada vez mais aceito na América do Norte e Grã-Bretanha. O autor levanta uma série de questões com relação à formação, à denominação e à imagem do profissional da informação e faz uma crítica à legislação brasileira, que regulamenta o exercício profissional que, segundo ele, não se encontra em sintonia com as novas realidades com as quais vai deparar-se. A grande maioria dos profissionais que exercem atividades relacionadas ao tratamento e à organização da informação são graduados em biblioteconomia, mas o próprio nome do profissional bibliotecário já vem sendo um fator que limita a idéia da vasta abrangência da área de atuação desse profissional.

O profissional da informação tem uma imagem pública pobre e estudos mostram características de introversão, falta de confiança em seu desempenho e de talento na capacidade de comunicar-se. Para tentar melhorar a sua imagem, o profissional da informação deve lutar no sentido de divulgar seu trabalho e a importância do seu papel na eficácia de um sistema de recuperação da informação. Deve investir na educação continuada, procurando estar sempre atualizado, bem como procurar aprofundar-se nas teorias que servem de embasamento às suas atividades, para ter senso crítico e não exercê-las mecanicamente.

São várias as especializações ocupadas pelos profissionais da informação, dependendo da função e da atividade que desenvolvem. Mas, na maioria das vezes, eles são generalistas, desenvolvendo diversos tipos de atividades. No caso do processamento técnico do acervo de documentos, o profissional responsável pela catalogação, classificação e indexação, costuma ter a formação bibliotecária, e receber o nome de *indexador*.

Quando se fala em indexador é preciso estar atendo à precisão conceitual do termo. Nas literaturas inglesa e americana, o termo indexador é aplicado tanto àquela pessoa que elabora índices de textos ou livros quanto àquela que faz a indexação acadêmica. O termo é adotado para referir-se a todos aqueles que fazem o tratamento de assunto, cuja tarefa seria a de analisar o assunto de um item, descrevê-lo em termos próprios e traduzi-los para a linguagem específica do sistema.

Diante das novas tecnologias da informação e da evolução de *softwares* desenvolvidos para tornar a indexação mais eficaz e rápida, certas questões estão sempre presentes nas discussões dos profissionais que lidam com informação: Qual será o futuro do indexador? Diante das bibliotecas virtuais e de textos digitalizados em sofisticados sistemas e bases de dados, haverá lugar para o indexador humano? Se este permanecer, qual será o seu papel?

Observa-se, pelas discussões sobre o tema, que não se conseguiu, pelo menos até hoje, transferir para a máquina o tipo de tarefa em que estão presentes elementos como a abstração, a percepção, a interpretação e outros processos inerentes ao funcionamento da mente humana. Apesar dos inúmeros esforços de equipes interdisciplinares que envolvem analistas de sistemas, profissionais da informação, lingüistas e psicólogos, esses pontos parecem constituir o limite do computador.

Os fatores considerados mais importantes com relação à influência do indexador no processo de análise de assunto são a subjetividade (diferentes indivíduos criam diferentes figuras ou idéias de uma mesma informação externa, por causa de suas inclinações pessoais e afetivas, que certamente interferem no trabalho por eles



desenvolvido), o conhecimento prévio (se refere ao estoque de conhecimento armazenado na memória do indivíduo, assimilado e adquirido em suas vivências) e a sua formação e experiência (um mínimo de conhecimentos da área em que se está indexando é obviamente indispensável).

A atividade desempenhada pelo indexador é a indexação, e o principal processo desenvolvido por ele é a análise de assunto.

O processo de análise de assunto

Para que se possa contextualizar o processo de análise de assunto é preciso que se esclareça como ocorre a indexação, tema que ainda não obteve grandes suportes teóricos.

Compreende duas etapas distintas: a análise de assunto, quando ocorre a extração de conceitos que possam representar o conteúdo de um documento, expressos em linguagem natural, e a tradução desses termos para termos de instrumentos de indexação, que são as chamadas linguagens de indexação. FUGMANN (1985) afirma que a indexação é o processo de: **(a)** discernir a essência de um documento e **(b)** representar essa essência com um grau suficiente de predicabilidade e fidelidade, isto é, num modo de expressão em linguagem de indexação.

A indexação é entendida como o *caráter integrativo* das duas etapas da indexação e acredita que, apesar de distintas, é preciso admitir que se trata de uma única operação realizada em dois momentos. Estudos de indexação se debruçam, principalmente, sobre as questões das linguagens utilizadas no sistema de recuperação. Segundo FROHMANN (1990), muitas pesquisas focalizam apenas aspectos da segunda etapa da indexação, enquanto sobre a primeira etapa, a Análise de assunto, considerada como uma operação intelectual junto à indexação, pouco é feito. Ele se vale de um pensamento de FOSKETT, quando diz que a operação chave da indexação, que é a decisão sobre o que o documento é, ainda é menos discutida e a menos redutível a regras.

Este estudo dá ênfase à primeira etapa da indexação, mostra a divisão do processo em fases e sua interdisciplinaridade com outras áreas do conhecimento.

O processo de ler um documento para extrair conceitos que traduzam a sua essência é conhecido como *análise de assunto*, para alguns, como análise temática, para outros, ou, ainda, como análise documentária, análise conceitual ou, mesmo, análise de conteúdo. Como se pode ver, trata-se também de um processo em que há uma certa confusão conceitual e para o qual aparecem diferentes concepções. NAVES (1996) mostra os diferentes sentidos dados, ao termo, na literatura.

O processo de análise de assunto, do ponto de vista do indexador, é iniciado com a fase de leitura do texto. Para isso, é necessário que se conheçam tipos e estruturas de textos para iniciar-se a sua leitura com fins específicos. Após essa leitura, passa-se à fase da extração de conceitos que possam representar o conteúdo temático do texto, para se chegar ao momento da fase de representação da atinência (*aboutness*), em que são definidos os termos em linguagem natural, denominados por FROHMANN (1990) de *frases de indexação*, que, depois de traduzidos para uma linguagem de indexação, passam a ser chamados de descritores de assunto, cabeçalhos de assunto, palavras-chave, termos de indexação ou enunciados. Todas as

fases do processo sofrem interferência de fatores lingüísticos, cognitivos e lógicos, o que dá ao processo de análise de assunto um caráter interdisciplinar.

Sabe-se que o processo de análise de assunto se realiza a partir de uma figura central: o indexador. Além da política de indexação adotada pelo sistema, tipo de vocabulário utilizado (linguagem natural ou artificial), objetivos da instituição, perfil e necessidades de informação dos usuários, o principal fator que interfere diretamente no processo parece ser, assim, o fator humano. No ato de pensar, quando faz abstrações, interpreta e define o assunto de um documento, o indexador sofre influência de diversos fatores pertencentes a vários campos, principalmente oriundos da lingüística, da ciência cognitiva e da lógica. Duas autoras e especialistas na área de ciência da informação, que tratam em profundidade o tema deste estudo, são Clare Beghtol, pesquisadora canadense, e Maria Pinto Molina, professora espanhola. (BEGHTOL, 1986, PINTO MOLINA, 1994).

Metodologia

Este estudo teve a intenção de analisar o processo de análise de assunto e os fatores interferentes nessa primeira fase da indexação e, para tanto, foi feita a escolha da abordagem pelo ponto de vista do indexador, por ser este considerado aquele que dá início a todo o processo e, por isso, responsável por grande parte de sua eficácia.

Acredita-se que o estudo comparativo entre os casos de indexadores com tempos diferentes de experiência em indexação, bem como a comparação entre o processo de análise de assunto de uma disciplina da área de ciências sociais e humanas e outra da área de ciências biológicas possam comprovar os dois pressupostos levantados inicialmente.

Nesta pesquisa, para fazer um estudo comparativo, são escolhidas as disciplinas sociologia e botânica, por serem áreas com características diferentes, propósitos e linhas de trabalho bem diversas, a segunda possuindo uma taxonomia definida entre suas classes de assuntos, o que já não acontece com a primeira. Consequentemente presume-se que os indexadores procedam de maneiras diferentes, e sofram interferência de fatores diversos durante o processo de análise de assunto dessas duas áreas.

Foi feito um estudo de caso de 6 (seis) indexadores, divididos em três grupos, a saber:

- **GRUPO 1** – 2 (dois) indexadores experientes (com mais de três anos exercendo a atividade de indexação, independente da área de atuação);
- **GRUPO 2** – 2 (dois) indexadores pouco experientes (com experiência de mais de um ano a três anos exercendo a atividade);
- **GRUPO 3** – 2 (dois) indexadores novatos (recém-formados, até um ano de experiência).

Para testar a metodologia a ser utilizada para coleta de dados, foi feito, com um elemento escolhido, um pré-teste, aqui denominado de Estudo Piloto (EP). Devido à riqueza de detalhes com que esse estudo se apresentou, optou-se por inclui-lo nos dados a serem analisados totalizando, assim, 7 (sete) indexadores entrevistados.

Para a coleta de dados foi aplicada a entrevista semi-estruturada, dividida em três partes:



- **1ª Parte** – Dados sobre os indexadores
- **2ª Parte** – Questões sobre análise de assunto
- **3ª Parte** – Análise dos protocolos verbais da análise de assunto de dois textos:

TEXTO 1 – Sociologia e TEXTO 2 – Botânica

A escolha dos textos obedeceu aos seguintes critérios: periódicos nacionais (para evitar a barreira da língua), fascículos ainda não indexados pelas bibliotecas selecionadas, artigos pequenos e com resumos. Os dois textos são os seguintes:

• **TEXTO 1:** REIS FILHO, Daniel Aarão. 1968; o curto ano de todos os desejos. *Tempo Social*; Revista de Sociologia da USP, São Paulo, v. 10, n.2, p.25-35, out. 1998.

• **TEXTO 2:** COELHO, Marcus A . Nadruz, MAYO, Simon J. Cinco espécies novas do gênero *Philodendron Schott (Araceae)* para o Brasil. *Boletim de Botânica da Universidade de São Paulo*, São Paulo, v.17, p.17-60, 1998.

As entrevistas foram feitas no ambiente de trabalho dos indexadores e combinadas com a observação do comportamento dos mesmos durante a análise. Foram gravadas em fitas cassete, com o consentimento dos entrevistados. O roteiro abordou temas como: fases do processo de análise de assunto, requisitos básicos para se fazer análise de assunto, aspectos do conteúdo do texto examinado, considerados importantes na determinação de atenção, fatores interferentes no processo e divergência na análise dos textos de sociologia e de botânica.

Na 3ª Parte da entrevista foi aplicada a Análise de Protocolos, no caso, o Protocolo verbal, com a técnica do *Pensar alto (Think aloud)*. Essa técnica, usada em análise qualitativa, é feita através da aplicação de entrevistas e é discutida por ERICSSON e SIMON (1990), que dão explicações detalhadas da aplicação da metodologia, das dúvidas pertinentes, dos tipos de procedimentos, das formas de comprovação e de detalhes de especificação. O trabalho desses autores analisa ainda a validade dos protocolos verbais como dados. Usam a estrutura teórica do processamento de informação humano para propor um modelo para os processos de verbalização de sujeitos instruídos a pensar alto, para obter registros verbais retrospectivos ou para produzir outros tipos de verbalização em resposta às instruções de experimentadores.

No campo da ciência da informação, a técnica já vem sendo utilizada por alguns pesquisadores, como INGWERSEN (1982), DAVID et al. (1995) e, mais recentemente, na pesquisa de FUJITA (1999).

A transcrição dos protocolos verbais é feita e, para analisar seu conteúdo, é aplicada a técnica de *análise de conteúdo*, um instrumento oferecido pelas ciências humanas, cujo objetivo é promover inferências, tendo ainda uma função heurística, o que enriquece a tentativa exploratória.

Descrição e discussão dos resultados

Conforme a proposta apresentada na Metodologia, a entrevista seguiu um roteiro pré-determinado, tendo sido dividida em três partes; este item descreve e discute os principais resultados obtidos em cada parte da entrevista. Essa separação

foi feita visando-se conhecer, inicialmente, a opinião dos entrevistados sobre análise de assunto (dados obtidos na 2ª parte) para, posteriormente, conhecer-se o que eles pensam durante o processo de análise de assunto (dados obtidos na 3ª parte, durante os protocolos verbais).

Dados sobre os indexadores

Os dados sobre os indexadores têm o intuito de se conhecer o perfil dos mesmos com relação à sua formação acadêmica, sua área profissional de atuação e seu tempo de experiência em Indexação. Os entrevistados foram codificados pela ordem em que as entrevistas foram realizadas, entre novembro de 1999 e janeiro 2000. São eles: EP (indexador experiente), E1 (indexador pouco experiente), E2 (indexador pouco experiente), E3 (indexador novato), E4 (indexador novato), E5 (indexador experiente) e E6 (indexador experiente).

QUADRO 1- Formação acadêmica dos indexadores, por ano de colação de grau

CURSO	EP	E1	E2	E3	E4	E5	E6
BIBLIOTECONOMIA	1980	1978	1980	1998	1998	1984	1976

O QUAD. 1 mostra a época da colação de grau no Curso de Biblioteconomia de cada um dos entrevistados. Defende-se a idéia de que a indexação seja uma atividade que deva ser exercida por profissionais bibliotecários, por ser essencialmente ligada ao saber fazer desse profissional, e ao conteúdo incluído em currículos dessa área.

Baseando-se na consulta feita à Biblioteca Universitária da UFMG e ao Colegiado de Graduação do Curso de Biblioteconomia da UFMG, foram entrevistados indexadores das seguintes áreas: dois da área de saúde, um de arquitetura, um de teologia/filosofia, dois de letras e um de biblioteconomia.

O tempo de experiência em Indexação é um aspecto verificado na pesquisa, a saber: EP – 15 anos, E1 – 3 anos, E2 – 2,5 anos, E3 – 1 ano, E4 – 1 ano, E5 – 14 anos e E6 – 12 anos.

Verifica-se que muitas são as atividades exercidas pelos indexadores, que nem sempre podem dedicar-se em tempo integral à atividade de Indexação. Foram citadas: chefia, processamento técnico, setor de periódicos, comutação bibliográfica, referência, setor de empréstimo, catalogação, classificação, atendimento aos usuários, organização de eventos, projetos, pesquisa, dentre outros.

Essa 1ª Parte da entrevista possibilita uma visão superficial das características de cada indexador entrevistado. A 2ª Parte trata do conhecimento que eles têm sobre análise de assunto, desde a concepção do termo até os fatores que interferem no seu trabalho.

Questões sobre análise de assunto

Como já é mencionado anteriormente, análise de assunto não é um conceito entendido da mesma maneira por todos os que lidam com o tema. Várias concepções são dadas e, para evitar que entre pesquisadora e entrevistados houvesse qualquer tipo de divergência conceitual com relação ao termo, é feita uma primeira questão sobre em que consiste a análise de assunto na opinião de cada um. Observou-se que

a maioria dos entrevistados entendem a concepção dada, neste estudo, ao termo análise de assunto, destacando-se a opinião de E1, quando fala em ...*essência do tema tratado*... Duas opiniões (E3 e E4) dão ênfase ao usuário, aspecto considerado importante mas que, do ponto de vista desta pesquisa, deve ser levado em consideração no momento seguinte ao da análise de assunto, ou seja, na definição de termos, quando vão ser observados os graus de especificidade e de exaustividade. E4 dá a impressão de que está falando de uma análise de assunto visando a seleção de documentos, fase prévia ao tratamento da informação. E3 e E4 são indexadores novatos e, certamente, estão numa fase ainda de insegurança, dando maior importância à satisfação do usuário, mostrando preocupar-se mais com esse, do que com o tema do documento em si.

Uma síntese dos aspectos levantados em relação aos requisitos para que o indexador faça análise de assunto é mostrada no QUAD. 2. Observa-se que apenas um requisito, *conhecer a área de atuação*, é lembrado por cinco indexadores; os demais são apontados isoladamente.

QUADRO 2 - Síntese dos requisitos para o indexador fazer análise de assunto

Requisito	Indexador
Conhecer a área de atuação	EP, E2, E3, E4, E5
Ter maturidade, vivência	E1
Saber um pouco de todas as áreas	E3
Saber o que é análise de assunto	E4
Gostar do que faz	E5
Conhecer as técnicas	E5
Respeitar a opinião do autor	E6
Ter um bom raciocínio	E6

A opinião de INGWERSEN (1982) é que a educação de bibliotecários generalistas, que não têm um *background* em campos de assuntos específicos, é considerada insuficiente, tendo eles um conhecimento de assunto bastante fraco. Afirma que, quanto à prática, somente após longa experiência é que, provavelmente, o bibliotecário desenvolverá métodos de trabalho eficientes.

Outro requisito apontado é a maturidade e vivência do profissional. Sem dúvida, a maturidade traz benefícios em qualquer atividade, formando, juntamente com o conhecimento prévio, a bagagem de conhecimentos adquiridos ao longo da vida do indivíduo. É importante lembrar que a maturidade torna-se benéfica quando aliada ao dinamismo, à vontade de aprender coisas novas, e nunca aliada ao comodismo, ao conformismo, à estagnação ou mesmo ao marasmo.

Os requisitos *saber um pouco de todas as coisas e saber o que é análise de assunto* são citados, respectivamente, por E3 e E4, ambos indexadores novatos. Parece que a idéia de E3 está relacionada a uma cultura geral do indexador já que *saber um pouco de todas as áreas* poderia caracterizar o indexador generalista. Mas E3 acha também importante conhecer a área de atuação do indexador. O outro requisito *saber o que é análise de assunto*, citado por E4, deveria ter sido lembrado por todos os indexadores, pois esse desconhecimento pode comprometer realmente todo o processo, como, por exemplo, em casos em que se analisa o documento, tendo-se

em vista apenas o vocabulário adotado no sistema, o que, como já foi afirmado, consiste numa segunda etapa.

Acredita-se que a reunião de todos os requisitos apontados no QUAD. 2 possa realmente contribuir para o trabalho do indexador.

É solicitado aos entrevistados que apontem os procedimentos necessários para extrair o conteúdo de um documento e o QUAD. 3 mostra uma síntese dos que são citados.

Pode-se observar que os indexadores, ao extrair o conteúdo de um documento, lêem principalmente o título, o resumo e o sumário. Costumam fazer uma leitura dinâmica antes de se ater a tópicos importantes. Essa leitura ajuda a ter uma visão geral do texto, e é abordada na literatura por PINTO MOLINA (1995), que estabelece alguns passos. Segundo a autora, inicialmente, o indexador deve fazer uma leitura rápida para reconhecer as características fundamentais como forma, classe e estrutura da informação.

A leitura do documento na íntegra, citada como procedimento por E3 e E5, torna-se totalmente inviável para a grande maioria dos indexadores. Não há tempo para isso, nem há necessidade de fazê-lo, como lembra FOSKETT (1973). A leitura pode ser inútil, já que se corre o risco de não se entender o conteúdo. Deve-se ter em mente que o que se quer é extrair o assunto do documento, e aí está explicada a necessidade de o indexador ter um raciocínio lógico e rápido, que o auxilie nessa tarefa de síntese.

Todos os itens citados no QUAD. 3 são considerados relevantes na identificação do conteúdo, mas não são lembrados por todos os entrevistados.

QUADRO 3 - Procedimentos a serem seguidos na extração do conteúdo de um documento

PROCEDIMENTOS	INDEXADOR
Leitura do título	EP, E1, E5
Leitura do resumo	EP, E1, E4, E6
Leitura do sumário	EP, E1, E5
Leitura do índice	E5
Exame da bibliografia	EP
Nome do autor e área de atuação	EP
Leitura de capítulos	EP
Orelha do livro	EP
Leitura da introdução	E1, E6
Leitura do documento na íntegra ³	E3, E5
Leitura dinâmica ⁴	E5
Leitura dos inícios dos parágrafos	E4
Leitura do final do texto	E4
Ilustrações	E5
Consulta a documentos do mesmo assunto	E5
Consulta a professores	E5
Observação dos objetivos e métodos do autor	E6
Folheada geral	EP

³ No caso de assuntos novos.

⁴ No caso de assuntos conhecidos.



Uma outra questão, ainda na 2ª parte da entrevista, é pedir-lhes que apontem fatores ou obstáculos que, na sua opinião, interferem no trabalho do indexador. As opiniões estão apresentados no QUAD. 4. Pode-se verificar que o fator mais citado é o da barreira da língua. Sabendo-se que os indexadores lidam com documentos das mais diversas línguas, esse é um aspecto que dificulta seu trabalho. Esse fator é citado tanto por um indexador novato (E3), quanto por um indexador pouco experiente (E1) e por um experiente (E5) e, tendo sido lembrado por indexadores dos três grupos, pode-se afirmar, então, que a barreira da língua independe do tempo de experiência com indexação.

QUADRO 4 - Fatores interferentes no trabalho do indexador

FATOR INTERFERENTE	INDEXADOR
Falta de conhecimento da área	EP, E5
Sistematização da área	EP
Barreira da língua	E1, E3, E5
Assuntos muito técnicos	E1
Contato direto com o usuário	E1
Trabalho no processamento técnico	E1
Familiarização com a terminologia	E1, E3
Falta de especialização	E2
Atualização do vocabulário	E2
Vivência na área de atuação	E3
Tempo para dedicar à Indexação	E4, E5
Execução de outras atividades	E4
Interrupções	E4, E6
Estilo do autor do documento	E6

Dos demais fatores citados, três podem ser considerados interligados: o tempo dedicado à indexação, a execução de outras atividades além da indexação, e interrupções. Esses fatores mostram a ansiedade e frustração (citada por E4) dos indexadores com relação ao tempo para se dedicar à indexação. Por executarem outras inúmeras atividades, não podem se concentrar nesse trabalho, sofrendo interrupções de todo tipo. Em se tratando de uma atividade intelectual, isso certamente consiste em um fator que dificulta o processo, tendo-se constatado que nenhum dos entrevistados se dedica exclusivamente à indexação.

Análise dos protocolos verbais durante a análise de assunto dos textos

Os protocolos verbais são adotados como metodologia de análise qualitativa com o propósito de se tentar conhecer o *processo* da análise de assunto, sem se preocupar com o *produto*, que seriam os termos escolhidos pelos entrevistados para identificar o assunto dos dois textos analisados. Esses são solicitados a identificar termos livres, apenas para a conclusão do processo, mas foram esclarecidos de que não haveria a preocupação com a tradução dos mesmos para um vocabulário controlado.

A transcrição das fitas gravadas do *pensar alto* segue rigorosamente a fala dos entrevistados. Algumas vezes, durante a análise dos textos, os entrevistados fazem curtas ou longas pausas e, tentando-se controlar essas pausas, em alguns casos muito freqüentes, segue-se a estratégia usada por CAVALCANTI (1989), e são feitas

perguntas que forcem a pessoa a dizer, no momento, o que está lendo ou pensando. Essas perguntas variam de entrevista para entrevista, apesar de alguns tópicos se terem repetido em todas elas. Essa variação ocorre porque todos os protocolos verbais registram o processo da análise, ou seja o que as pessoas falam enquanto pensam. Diante disso, algumas perguntas são pertinentes em momentos específicos, não se justificando em outros.

Discussão dos fatores identificados durante os protocolos verbais

Os protocolos verbais registram o pensamento dos entrevistados durante a análise de assunto dos textos. Essa discussão dos fatores traz os temas surgidos durante o processo de análise dos dois textos, de sociologia e de botânica. Os fatores estão separados em três categorias, as saber:

- **CATEGORIA 1** – Fatores relativos ao texto (alíneas **a** a **j**)
- **CATEGORIA 2** - Fatores relativos ao indexador (alíneas **k** a **n**)
- **CATEGORIA 3** – Fatores observados no comportamento do indexador durante os Protocolos verbais (alíneas **o** a **s**).

CATEGORIA 1 – Fatores relativos ao texto

Nesta primeira categoria estão incluídos os fatores identificados como diretamente ligados ao texto e aos procedimentos seguidos pelos indexadores entrevistados durante as análises dos dois textos.

a) Título do artigo

O primeiro fator observado pelos entrevistados durante a análise é o título dos artigos. O título do Texto 1 não pode ser considerado determinante na identificação do assunto e nenhum dos entrevistados extrai o assunto pelo título. Já com relação à análise do Texto 2, pode-se verificar que os comentários feitos mostram uma influência do título na definição do assunto do artigo.

b) Autoria do artigo

No caso do Texto 1, pode-se afirmar que os dados sobre o autor do artigo não ajudam na definição do assunto. Quanto à autoria do Texto 2, apenas um entrevistado faz algum tipo de comentário sobre a autoria do artigo.

c) Resumo do artigo

O resumo é um item presente em todos os artigos de periódicos que seguem normas para publicação. Todos os entrevistados, nas análises dos dois textos, gastam um pouco de seu tempo na leitura dos resumos de ambos.

d) Leitura do texto (estrutura, partes, parágrafos)

As formas de leitura variam de entrevistado para entrevistado, uns fazendo uma leitura geral, outros lendo os inícios dos parágrafos e algumas partes.

e) Referências bibliográficas

As referências bibliográficas são consultadas pelos entrevistados na busca do assunto do documento, pois trazem a relação de publicações consultadas pelo autor para escrever seu texto, refletindo o embasamento teórico utilizado.

f) Terminologia

É considerado um fator que merece atenção, pois entende-se que o domínio dos termos de uma determinada área facilita consideravelmente o trabalho do indexador.

g) Ilustrações

As informações também são citadas como fonte de informação na definição do assunto do documento. No caso em estudo, apenas o Texto 2 – Botânica, traz ilustrações das espécies de plantas mencionadas no texto.

h) Termos definidos como assuntos dos textos

Durante a análise de assunto dos dois textos, é solicitado aos entrevistados que definam alguns termos que representem os assuntos dos mesmos. Como a maior preocupação deste estudo é relativa ao processo de análise de assunto, e não ao produto da análise, ou seja, o termo de indexação, os participantes foram informados de que poderiam definir termos livres sem se aterem a uma determinada linguagem de indexação.

i) Facilidades e dificuldades para definir o assunto

Os entrevistados são solicitados a *pensar alto* sobre as dificuldades enfrentadas para definir o assunto de cada texto. No entanto, alguns apontam fatores não só dificultadores, como também fatores facilitadores.

j) Título do periódico

Um fator muito importante, mas que é esquecido por todos os entrevistados durante suas análises, é o título do periódico onde é publicado o artigo em exame.

CATEGORIA 2 – Fatores relativos ao indexador

Nesta segunda categoria, estão incluídos fatores que também interferem no processo de análise dos textos, mas fazem parte de aspectos pessoais do indexador.

k) Conhecimento prévio

A compreensão de um texto se caracteriza pela utilização do conhecimento prévio do leitor quando ele utiliza, na leitura, o que ele já sabe, o conhecimento adquirido ao longo de sua vida.

l) Formação, experiência e conhecimento na área do texto analisado

Em alguns momentos, durante os protocolos verbais, os entrevistados fazem observações sobre formação profissional e experiência para executar a tarefa de indexar, bem como sobre especialização e prática na área de atuação.

m) Orientação de especialistas na área de atuação

Nos casos em que, sozinho, o indexador não conseguir definir o assunto, o mais comum é a solicitação da orientação de especialistas na sua área de atuação.

n) Interação com o usuário

O trabalho com usuários requer estudos sobre temas específicos, e acredita-se que a atinência de um documento possa ser definida independentemente do tipo de usuário que vai utilizar a informação. Apesar disso, verifica-se que alguns entrevistados citam suas preocupações com seus usuários durante o processo de análise de assunto.

CATEGORIA 3 – Fatores observados no comportamento do indexador durante os protocolos verbais

Como foi previsto no item sobre metodologia, a técnica da entrevista deveria ser combinada com a observação do comportamento do indexador durante a análise.

o) Medo

O fator *medo* é apontado por alguns entrevistados, relacionado ao medo de errar durante a análise de assunto.

p) Ansiedade

Outro fator comportamental observado durante as entrevistas é a *ansiedade*

q) Atenção

O fator *atenção* está presente em todo o processo de Análise de assunto, e aqui é visto como sinônimo de concentração.

r) Auto-crítica

Em determinados momentos do processo de *pensar alto* o que estão fazendo os entrevistados, comentários evidenciam atitude de auto-crítica frente à tarefa que está sendo realizada.

s) Preguiça

Preguiça é um fator aqui considerado, que é mencionado durante a análise por um dos entrevistados e que pode estar relacionado a dificuldades durante a leitura.

Análise comparativa dos resultados

A partir dos dados discutidos no item anterior, traçou-se um perfil de cada indexador, o que possibilitou a identificação e a comparação de fatores que interferem nesse tipo de atividade.

Verifica-se que alguns fatores interferentes em análise de assunto independem do tempo de experiência em indexação, podendo afetar tanto um indexador novato quanto um experiente. Um exemplo disso é o tempo para se dedicar ao trabalho, pois interrupções e falta de concentração podem afetar a execução de uma atividade essencialmente intelectual, que exige reflexões, inferências, deduções, raciocínio, busca de informações na memória e utilização do conhecimento prévio.

Fica também, evidente, a necessidade de se conhecer a área em que se está atuando, e esse é um fator apontado por todos os entrevistados, independentemente do tempo de experiência de cada um em indexação. A maturidade adquirida durante anos de experiência, sem dúvida, auxilia nessa atividade, mas sempre que o profissional se defronta com uma nova área, certamente obstáculos surgem no seu trabalho de indexar.

O domínio da terminologia de uma área é um fator muito importante nesse processo de análise de assunto, e pode ser adquirido com muitas leituras, tempo de experiência e contato com o usuário, que formula suas questões de acordo com a terminologia adotada na sua área. Um fator negativo mencionado é a barreira da língua. A falta do domínio da língua, principalmente do inglês, interfere negativamente no trabalho do indexador. Ele pode deparar-se com documentos em qualquer língua, inclusive em línguas raras, dependendo da área de atuação.

Aos fatores citados, acrescentam-se outros, que são observados no comportamento dos entrevistados durante as entrevistas. São os fatores cognitivos, lingüísticos e lógicos que podem interferir no processo de análise de assunto, e sabe-se que esses são, muitas vezes difíceis de serem identificados, por fazerem parte da personalidade dos indivíduos.

Alguns aspectos foram identificados e comparados entre as análises dos textos de sociologia e de botânica. A estrutura do texto é um desses aspectos de divergência entre as duas áreas. O Texto 1 não tem uma rígida estrutura científica, ou seja, não tem partes e divisões que caracterizam esse tipo de literatura, com introdução,



metodologia, resultados e discussão, o chamado modelo clássico. Já o Texto 2 é considerado, pela maioria dos entrevistados, como mais fácil de analisar, pela estrutura e descrição mais clara, que permite uma melhor visualização do assunto.

Um outro ponto de divergência verificado entre os textos das duas disciplinas é com relação à terminologia usada. Esse é um fator em que as opiniões se dividem entre os entrevistados, e a suposta dificuldade com relação ao Texto 2 surge devido à utilização de nomes científicos em latim, ou de termos em inglês, o que é comum na área e requer familiaridade e vivência com a terminologia para que essa possa ser compreendida e assimilada.

Algumas conclusões da pesquisa

Este estudo pode ser considerado como um primeiro passo na busca de se conhecer em profundidade o processo de análise de assunto mas, como se trata de um estudo exploratório e descritivo, deve ser visto apenas como uma das etapas de uma investigação mais ampla.

Vale lembrar que os resultados obtidos e discutidos podem ser aplicados apenas ao grupo de sete indexadores que foram entrevistados e observados durante a pesquisa pois, em se tratando de um estudo de caso, não é possível generalizar esses resultados a outros indexadores, em situações diversas e examinando outros tipos de documentos para indexação.

Baseando-se nos resultados obtidos na pesquisa, acredita-se que seja possível que se façam algumas reflexões sobre questões básicas levantadas no estudo, tais como:

- A questão primordial para o desenvolvimento do processo de análise de assunto é a compreensão do texto, que faz parte do processamento de informação do ser humano. O estudo do texto é considerado pelos especialistas como o necessário ponto de partida para operações analítico-documentárias.

- Refletindo-se sobre o que foi lido na literatura, bem como aos resultados obtidos na pesquisa, pode-se concluir que o processo de análise de assunto não deva ser visto isoladamente dentro da ciência da informação, havendo necessidade de se buscarem contribuições de outras áreas disciplinares, principalmente da lingüística, da psicologia cognitiva e da lógica.

- Uma questão muito citada por autores da área e que merece ser enfatizada refere-se à deficiência na formação e à falta de especialização do indexador na sua área de atuação. Confessada por alguns entrevistados como origem de suas dificuldades, tornam-se prementes os esforços no sentido de se estudarem formas de superá-la.

- A pesquisa comprova que as dificuldades estão mais relacionadas ao nível de processamento *top-down* do que o *bottom-up*, pois os entrevistados estão sempre ressaltando a importância do conhecimento prévio na área de atuação.

- Após análise comparativa, pode-se afirmar que o tempo maior de experiência do indexador com a técnica de indexação lhe proporciona mais segurança, maior vivência e maior capacidade de decisão ao lidar com situações complexas. No entanto, se lhe for dado para análise o assunto de um texto que nunca tenha sido seu objeto de estudos ou em cuja área não tenha ainda trabalhado, poderão interferir no processo os mesmos fatores que interferem na análise feita por um indexador novato. Isso ocorre porque lhe faltará o conhecimento prévio naquela nova área, o domínio da terminologia

nela adotada e, diante disso, ele poderá enfrentar os mesmos tipos de obstáculos que um novato enfrentaria.

• Durante a pesquisa foi possível verificar que a estrutura de textos (que varia de área para área) e a terminologia usada por seus autores também são características das diversas áreas das ciências, e o conhecimento desses aspectos pode contribuir para o trabalho do indexador, pois pode afetar as fases de leitura, compreensão e extração de conceitos.

Factors that interfere in subject analysis process

In Information Science context, subject analysis is considered one of the most important stages in Indexing. The indexer is the responsible for this activity, and he begins the process with subjectivity, previous knowledge, academic formation and background. The objective of the research is to identify factors that interfere with subject analysis and an exploratory study is done with case study by seven indexers divided into groups according to their different backgrounds. The results show that the primordial question to the development of subject analysis process is text comprehension and that Information Science should try to obtain contributions from interdisciplinary areas.

Key-words: Subject analysis, Indexing, Indexer

Referências bibliográficas

BARBOSA, Ricardo Rodrigues. Perspectivas profissionais e educacionais em biblioteconomia e ciência da informação. *Ciência da Informação*, v. 27, n.1, p. 53-60, jan./abr. 1998.

BEGHTOL, Clare. Bibliographic classification theory and text linguistics: aboutness, intertextuality, and the cognitive act of classifying documents. *Journal of Documentation*, v. 42, n. 2, p.84-113, June 1986.

CAMPOS, Astério Tavares. A indexação. *Revista de Biblioteconomia de Brasília*, v. 15, n. 1, p. 69-72, jan./jun.1987.

CAVALCANTI, Marilda do Couto. *Interação autor-texto: aspectos de interpretação pragmática*. Campinas: Ed. UNICAMP, 1989.

DAVID, Claire et al. Indexing as problem solving: a cognitive approach to consistency. In: ANNUAL MEETING OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE, 58TH., 1995, Washington. *Proceedings...* Washington, DC, 1995. p.49-55.

ERICSSON, K. Anders, SIMON, Herbert. Verbal reports as data. *Psychological Review*, v. 87, n. 3, p. 215-251, May 1980.

FOSKETT, A.C.A *abordagem temática da informação*. São Paulo: Polígono, 1973. 437p.

FROHMANN, B. Rules of indexing: a critique of mentalism in information retrieval theory. *Journal of Documentation*, v. 46, n. 2, p. 81-101, June 1990.

FUGMANN, Robert. The complementarity of natural and indexing languages. In: CHAN, L. M., RICHMOND, P., SVENONIUS, E. *Theory of subject analysis*. Littleton, CO: Libraries Unlmites, 1985. p.390-402.

FUJITA, Mariângela Spotti Lopes. A leitura do indexador: estudos de observação. *Perspectivas em Ciência da Informação*, Belo Horizonte, v. 4, n. 1, p. 101-117, jan./jun. 1999.

INGWERSEN, Peter. Search procedures in the library – analysed from the cognitive point of view. *Journal of Documentation*, v. 38, n. 3, p. 165-191, Sept. 1982.

NAVES, Madalena Martins Lopes. Análise de assunto: concepções. *Revista de Biblioteconomia de Brasília*, Brasília, v. 20, n. 2, p. 215-226, jul./dez. 1996.

PINTO MOLINA, María. Documentary abstracting; toward a methodological model. *Journal of the American Society for Information Science*, v. 46, n.3, p.225-234, 1995.

_____. Interdisciplinary approaches to the concept and practice of Written Documentary Content Analysis (WTDC). *Journal of Documentation*, v. 50, n. 2, p.111-1333, June 1994.

STRASSMANN, P. A. *Information payoff: the transformation of work in the eletronic age*. New York: Free Press, 1985.