

UNIVERSIDADE DO VALE DO RIO DOS SINOS - UNISINOS
ESCOLA DA INDÚSTRIA CRIATIVA: COMUNICAÇÃO, DESIGN E LINGUAGENS
PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA APLICADA
DOUTORADO EM LINGUÍSTICA APLICADA

RITA DO CARMO FERREIRA LAIPELT

METODOLOGIA PARA SELEÇÃO DE TERMOS EQUIVALENTES E
DESCRITORES DE TESAUROS: um estudo no âmbito do Direito do Trabalho e
do Direito Previdenciário

São Leopoldo
2015

Rita do Carmo Ferreira Laipelt

METODOLOGIA PARA SELEÇÃO DE TERMOS EQUIVALENTES E
DESCRITORES DE TESAUROS: um estudo no âmbito do Direito do Trabalho e do
Direito Previdenciário

Tese apresentada como requisito parcial para
obtenção do título de Doutor em Linguística
Aplicada, pelo Programa de Pós-Graduação
em Linguística Aplicada da Universidade do
Vale do Rio dos Sinos – UNISINOS.

Orientadora: Profa. Dra. Maria da Graça
Krieger

São Leopoldo

2015

CIP – Brasil. Dados Internacionais de Catalogação na Publicação

L186m

LAIPELT, Rita do Carmo Ferreira

Metodologia para seleção de termos equivalentes e descritores de tesouros: um estudo no âmbito do Direito do Trabalho e do Direito Previdenciário / Rita do Carmo Ferreira Laipelt. -- 2015.
213 f.

Orientadora: Profa. Dra. Maria da Graça Krieger.

Tese (Doutorado) -- Universidade do Vale do Rio dos Sinos
Programa de Pós-Graduação em Linguística Aplicada, São Leopoldo BR-
RS, 2015.

1. Terminologia. 2. Ciência da Informação. 3. Teoria Comunicativa da Terminologia. 4. Variantes denominativas. 5. Tesouros. I. Krieger, Maria da Graça, orient. II. Título.

CDU 025.43

Elaborada pela autora, CRB-10/1696.

RITA DO CARMO FERREIRA LAIPELT

“METODOLOGIA PARA SELEÇÃO DE TERMOS E DESCRITORES DE
TESAUROS: um estudo no âmbito do Direito do Trabalho e do Direito
Previdenciário”

Tese apresentada como requisito parcial
para obtenção do título de Doutor, pelo
Programa de Pós-Graduação em
Linguística Aplicada da Universidade do
Vale do Rio dos Sinos - Unisinos

Aprovada em 30 de junho de 2015

BANCA EXAMINADORA

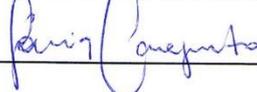
Profa. Dra. Ligia Maria Arruda Café (UFSC)



Profa. Dra. Regina Helena Van Der Laan (UFRGS)



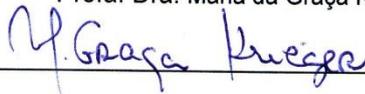
Profa. Dra. Sonia Elisa Caregnato (UFRGS)



Profa. Dra. Rove Luiza de Oliveira Chishman (UNISINOS)



Profa. Dra. Maria da Graça Krieger (UNISINOS)



DEDICATÓRIA

Para os amigos e familiares que estiveram presentes em minha vida nos últimos quatro anos: sem o apoio e carinho de todos vocês essa conquista seria ainda mais difícil. Estejam certos de que minha felicidade só é plena pois posso compartilhá-la com vocês.

AGRADECIMENTOS

Ao Programa de Pós-Graduação em Linguística Aplicada da UNISINOS pela oportunidade de realizar esta pesquisa.

À professora Maria da Graça Krieger por concordar em orientar uma bibliotecária. O discurso interdisciplinar é lindo, mas, vivenciar a interdisciplinaridade na prática, é um desafio tanto para o orientador como para o orientando. Trilhar por esse caminho requer de ambos disponibilidade em dobro para ouvir, explicar e compreender. Muito obrigada professora, por se dispor a me ouvir e pela paciência e generosidade com que compartilhou seus conhecimentos comigo.

À todos os professores, funcionários e colegas do PPGLA com os quais tive a oportunidade de conviver. Especialmente à professora Rove Chishman, pelos conhecimentos partilhados durante o período em que participei do Grupo de Pesquisa Semantec. Aos colegas do PPGLA – Luciana Monteiro, Carla, Eliane, Ana, Deise, Anderson, Carolina, Dênis e Thaís – pela troca de conhecimentos e amizade.

À CAPES pelos vinte e quatro meses em que me concedeu a bolsa Prosup, e à Escola de Desenvolvimento da UFRGS pelos dezoito meses em que me concedeu a bolsa de apoio a qualificação.

Ao coordenador do Projeto LexMI do Senado Federal, senhor João Alberto de Oliveira Lima, pela gentileza e disponibilização dos logs do Portal LexMI para a realização desta pesquisa.

Ao professor Ricardo Camargo da Faculdade de Direito da UFRGS, pela análise de validação de algumas lexias. E à bibliotecária Nariman por ter me colocado em contato com o professor Ricardo.

À professora Regina van der Laan, por ter despertado meu interesse pelo incrível universo de possibilidades de pesquisa da área de organização do conhecimento. Obrigada, também, Regina pelos livros, dicas, ensinamentos e palavras de incentivo.

Às professoras Lígia Café, Regina van der Laan, Rove Chishman e Sonia Caregnato por aceitarem fazer parte de minha banca, mas sobretudo, pelas preciosas contribuições para a realização desta tese.

À todos os colegas do Departamento de Ciências da Informação da Faculdade de Biblioteconomia e Comunicação da UFRGS por terem concordado com a redução de minha carga horária durante um semestre - período fundamental para a realização desta tese. Especialmente à Ana Moura, Sonia Caregnato, Samile Vanz, Rodrigo Caxias, Rita Portela, Jackson Medeiros, Glória Ferreira, Martha Bonotto, Helen Rozados, Bere – pelas diferentes formas de apoio, incentivo e carinho que recebi de todos vocês ao longo do doutorado.

À toda minha família pela torcida e paciência com minha ausência nos últimos anos. Ao Éderson pela tolerância com a minha falta de tempo e finais de semana sempre trabalhando. À minha sogra Luiza, à minhas cunhadas France e Jana, à sobrinha Caroline e meu cunhado Ricardo – por ajudarem a cuidar do Santiago nos finais de semana em que íamos para São Marcos.

À minha mãe, Maria de Lourdes; à minhas irmãs, Angelita e Angélica; ao meu cunhado Volnei; e meus sobrinhos Carolina e Pedro – agradeço pelo incentivo, apoio e compreensão pela minha falta de tempo.

À minha irmã Ângela por estar ao meu lado todos os dias, em todos os momentos, bons e maus, desses quatro anos. Obrigada Ângela, por cuidar do Santiago com tanto carinho e dedicação para que eu pudesse trabalhar e escrever, sem a sua ajuda esta tese não seria possível.

Ao meu filho, Santiago, por me esperar todos os dias com um sorriso largo, acompanhado de um beijinho doce, um abraço apertado e um “adoro você”. Obrigada, San, por recarregar minhas energias com o seu carinho e mostrar que todo meu esforço pelo seu bem-estar vale a pena.

“[...] é o ponto de vista que cria o objeto [...]”
(SAUSSURE, 1999 [1916], p. 15).

RESUMO

A incompatibilidade entre os termos atribuídos por bibliotecários aos documentos, durante o processo de indexação, e a linguagem utilizada pelos usuários, durante o processo de busca de informação, prejudica muitas vezes, a recuperação de documentos pertinentes. Apresenta proposta que visa a inserção do ponto de vista dos usuários na concepção dos tesouros, objetivando minimizar a incompatibilidade entre os descritores e a linguagem dos usuários, Utiliza como objeto de estudo as expressões de busca contidas nos “logs” de pesquisa dos usuários do Portal LexMI do Senado Federal Brasileiro nas áreas do Direito do Trabalho e do Direito Previdenciário. Através desta análise, verifica as expressões (lexias) utilizadas pelos usuários para a recuperação de informação. Cobre um período de quinze dias em 2012 identificando trezentas e oitenta e cinco (385) variantes denominativas verbais, cem (100) lexias alfanuméricas e numéricas e dezenove (19) variantes conceituais empregadas pelos usuários do Portal LexMI. Propõe uma nova perspectiva de elaboração de tesouros considerando aspectos conceituais e linguísticos dos itens lexicais que devem compor a linguagem documentária com base nos resultados do estudo das variantes recorrentes, descritos conforme fundamentos teóricos da Teoria Comunicativa da Terminologia e sistematização de Freixa (2002; 2014). A pesquisa justifica-se, pela abordagem ainda incipiente tanto na literatura como nas normas de elaboração de tesouros, vigentes e antigas, sobre a questão da validação do usuário, as quais destacam a importância dessa validação, mas não mencionam, explicitamente, a metodologia a ser empregada para sua obtenção. O percurso metodológico necessário à implementação da proposta, desenvolveu-se a partir da comparação das informações existentes em um *corpora* formado por: a) leis brasileiras mais relevantes dessas duas áreas do Direito; b) linguagem de indexação utilizada pela Biblioteca do Senado Federal para a descrição do conteúdo dos documentos (Vocabulário Controlado Básico - VCB); c) Base de Dados de Jurisprudência do Portal LexMI; d) Norma ISO 25964-1 (2011) e e) três tesouros da área jurídica. A partir da análise do *corpus* de lexias de buscas dos usuários identificamos três categorias de lexias variantes: a) aspectos linguísticos; b) recursos não verbais e c) aspectos conceituais. Concluí que a tipologia de variantes denominativas mais recorrente pode variar de acordo com a área de conhecimento analisada. No âmbito do Direito do Trabalho e Direito Previdenciário as variantes de maior ocorrência identificadas foram aquelas que compreendem os recursos não verbais, especificamente lexias numéricas e alfanuméricas que ocorrem entre o nome de uma lei e seu respectivo número e data. Sugere a inserção de variantes denominativas como termos equivalentes dos descritores de um tesouro para o aperfeiçoamento desses sistemas de organização do conhecimento tendo em vista o reforço de aspectos linguísticos e comunicativos que elas propiciam.

Palavras-chave: Terminologia. Ciência da Informação. Teoria Comunicativa da Terminologia. Variantes denominativas. Tesouros.

RESUMÉN

La incompatibilidad entre los términos atribuidos por bibliotecarios a los documentos durante el proceso de indización y el lenguaje utilizado por los usuarios durante el proceso de búsqueda de información perjudica, muchas veces, la recuperación de documentos relevantes. Presenta propuesta que pretende insertar el punto de vista de los usuarios en la estructura de los tesauros, con el objetivo de minimizar la incompatibilidad entre los descriptores y el lenguaje de los usuarios. Utiliza como objeto de estudio las expresiones de búsqueda contenidas en los “logs” de investigación de los usuarios del Portal LexMI del Senado Federal Brasileiro en las áreas del Derecho Laboral y Seguridad Social. Mediante este análisis verifica las expresiones (lexías) utilizadas por los usuarios para la recuperación de información. Abarca un período de quince días en 2012, identificando trescientas ochenta y cinco (385) variantes denominativas verbales, cien (100) lexías alfanuméricas y numéricas y diecinueve (19) variantes conceptuales empleadas por los usuarios del Portal LexMI. Propone una nueva perspectiva de elaboración de tesauros que considere aspectos conceptuales y lingüísticos de los ítems lexicales que deben componer el lenguaje documental, tomando como base los resultados del estudio de las variantes recurrentes, descritas conforme fundamentos teóricos de la Teoría Comunicativa de la Terminología y sistematización de Freixa (2002; 2014). La investigación se justifica por el abordaje todavía incipiente tanto en la literatura como en las normas de elaboración de tesauros, vigentes y antiguas, sobre la cuestión de la validación del usuario, las cuales destacan la importancia de esa validación, pero no mencionan, explícitamente, la metodología a ser empleada para su obtención. El recorrido metodológico necesario a la implementación de la propuesta se ha desarrollado a partir de la comparación de las informaciones existentes en un *corpora* formado por: a) leyes brasileras más relevantes de esas dos áreas del Derecho; b) lenguaje de indización utilizado por la Biblioteca del Senado Federal para la descripción del contenido de los documentos (Vocabulario Controlado Básico – VCB); c) Base de Datos de Jurisprudencia del Portal LexMI; d) Norma ISO 25964-1 (2011) y e) tres tesauros del área jurídico. A partir del análisis del *corpus* de lexías de búsquedas de los usuarios hemos identificado tres categorías de lexías variantes: a) aspectos lingüísticos; b) recursos no verbales y c) aspectos conceptuales. Concluye que la tipología de variantes denominativas más recurrente puede variar conforme el área de conocimiento analizado. En el ámbito del Derecho Laboral y Seguridad Social las variantes de mayor ocurrencia identificadas fueron aquellas que comprenden los recursos no verbales, específicamente lexías numéricas y alfanuméricas que ocurren entre el nombre de una ley y su respectivo número y fecha. Sugiere la inserción de variantes denominativas como términos equivalentes de los descriptores de un tesauro para el perfeccionamiento de esos sistemas de organización del conocimiento, teniendo en cuenta el refuerzo de aspectos lingüísticos y comunicativos que propician.

Palabras clave: Terminología. Ciencia de la Información. Teoría Comunicativa de la Terminología. Variantes denominativas. Tesauros.

LISTA DE ILUSTRAÇÕES

Figura 1 – Ontologias.....	64
Figura 2 – Hierarchical refinement of the causal relation	74
Quadro 1 – Termo: configurações morfossintáticas prototípicas	39
Quadro 2 – Classificação formal de variantes denominativas	44
Quadro 3 – Alterações gráficas.....	46
Quadro 4 – Alterações morfossintáticas.....	47
Quadro 5 – Reduções de extensão.....	48
Quadro 6 – Reduções de base	48
Quadro 7 – Alterações léxicas de unidades poliléxicas	49
Quadro 8 – Várias alterações complexas.....	50
Quadro 9 – Comparação entre reduções e várias alterações complexas	52
Quadro 10 – Comparação entre alterações morfossintáticas e várias alterações complexas.....	52
Quadro 11 – Variantes de acordo com a socioterminologia	56
Quadro 12 - Relações associativas	72
Quadro 13 - Cânones para o trabalho no plano das idéias	79
Quadro 14 - Princípios gerais de classificação	81
Quadro 15 – Alterações gráficas: resultados	110
Quadro 16 – Alterações morfossintáticas: resultados	111
Quadro 17 – Reduções de extensão: resultados.....	113
Quadro 18 – Reduções de base: resultados	113
Quadro 19 – Alterações léxicas de unidades poliléxicas: resultados	116
Quadro 20 – Várias alterações complexas: resultados	117
Quadro 21 – Busca por campo semântico	118

Quadro 22 – Alterações gráficas: primeiro corpus de lexias verbais.....	177
Quadro 23 – Alterações morfossintáticas: primeiro corpus de lexias verbais	178
Quadro 24 – Reduções de extensão: primeiro corpus de lexias verbais	179
Quadro 25 – Reduções de base: primeiro corpus de lexias verbais	180
Quadro 26 – Alterações léxicas de unidades monoléticas: primeiro corpus de lexias verbais.....	180
Quadro 27 – Alterações léxicas de unidades poliléticas: primeiro corpus de lexias verbais.....	181
Quadro 28 – Várias alterações complexas: primeiro corpus de lexias verbais	182
Quadro 29 – Alterações gráficas: segundo corpus de lexias verbais	183
Quadro 30 – Alterações morfossintáticas: segundo corpus de lexias verbais	185
Quadro 31 – Reduções de extensão: segundo corpus de lexias verbais	187
Quadro 32 – Reduções de base: segundo corpus de lexias verbais.....	190
Quadro 33 – Alterações léxicas de unidades monoléticas: segundo corpus de lexias verbais.....	192
Quadro 34 – Alterações léxicas de unidades poliléticas: segundo corpus de lexias verbais.....	192
Quadro 35 – Várias alterações complexas: segundo corpus de lexias verbais	193

SUMÁRIO

1 INTRODUÇÃO	14
2 CIÊNCIA DA INFORMAÇÃO: aspectos históricos e aplicados.....	22
2.1 Biblioteconomia, Documentação e Ciência da Informação: o que fomos e o que somos.	22
2.2 A Organização e Representação da Informação e do Conhecimento e os Sistemas de Recuperação da Informação	26
3 FUNDAMENTOS TEÓRICOS DA TERMINOLOGIA PARA A ELABORAÇÃO DE TESAUROS	33
3.1 Teoria Geral da Terminologia e Teoria Comunicativa da Terminologia	34
3.2 Variação	43
3.3 As relações entre termo e descritor	57
4 TESAUROS: o que são e para que servem?	61
4.1 Características técnicas dos tesauros	70
4.2 Fundamentos de elaboração de tesauros.....	75
4.2.1 Teoria Clássica de Aristóteles	76
4.2.2 Teoria da Classificação Facetada	78
4.2.3 Teoria do Conceito	84
4.3 Normas de elaboração de tesauros	88
5 ESTUDO DE CASO NO AMBITO DO DIREITO DO TRABALHO E DO DIREITO PREVIDENCIÁRIO	93
5.1 Portal LexMI	95
5.2 <i>Corpus</i> da pesquisa	95
5.3 Coleta de dados.....	97
5.4 Etapas metodológicas da pesquisa.....	98

5.5 Organização dos dados	100
5.6 Análise dos dados.....	103
6 RESULTADOS.....	104
6.1 Aspectos verbais.....	105
6.1.1 Validação das lexias de busca do primeiro corpus de lexias.....	105
6.1.2 Descrição das tipologias de variantes denominativas do primeiro e do segundo corpus de lexias	108
6.2 Recursos não verbais	118
6.3 Aspectos conceituais.....	120
7 PRINCÍPIOS DE ELABORAÇÃO DE TESAUROS: Metodologia para a Validação da Perspectiva do Usuário para a Seleção de Termos e Descritores de Tesauros	125
8 CONCLUSÃO	131
REFERENCIAS.....	134
APENDICE A – Primeiro <i>corpus</i> de lexias verbais	145
APENDICE B – Lexias alfanuméricas e numéricas	148
APENDICE C - Lexias conceituais	161
APENDICE D - Ocorrência das lexias de buscas nos <i>corpora</i> da pesquisa	162
APENDICE E - Lexias variantes classificadas de acordo com a Classificação Decimal de Direito	169
APENDICE F – Classificação do primeiro <i>corpus</i> de lexias verbais.....	176
APENDICE G - Classificação do segundo <i>corpus</i> de lexias verbais	182
APENDICE H – variantes denominativas validadas em três tesauros.....	194
APENDICE I – Validação do especialista.....	197
APENDICE J – Lexias validadas no primeiro <i>corpus</i> textual especializado	198
APENDICE K - Lexias validadas no VCB.....	200
APENDICE L - Lexias validadas na fonte especializada (Jurisprudência)	202

APENDICE M - Categoria alterações gráficas por supressão de preposição ..	206
APENDICE N - Categoria busca por campo semântico	207
ANEXO A - Classificação Decimal de Direito: Direito do Trabalho.....	208
ANEXO B - Classificação Decimal de Direito: Direito Previdenciário.....	212

1 INTRODUÇÃO

Com os avanços tecnológicos decorrentes da Segunda Guerra Mundial, o tratamento da informação passou a ser fundamental para o controle da literatura e acompanhamento de pesquisas em desenvolvimento. Em função dos problemas gerados pela explosão da informação, a ciência avançou rapidamente e o valor da informação passou a ser percebido em função de interesses políticos, científicos e tecnológicos que se tornaram evidentes. Em 1945, Vannevar Bush, preocupado com a questão da recuperação da informação, falava sobre a associação de conceitos ou palavras para a organização da informação, argumentando que esse é o padrão utilizado pelo cérebro humano para transformar informação em conhecimento. Desde então, a Ciência da Informação tem dedicado maior atenção a esse tema e, com isso, tem desenvolvido pesquisas visando o aprimoramento de técnicas e instrumentos que auxiliem na realização da atividade de organização da informação, tais como as linguagens documentárias.

No contexto atual, com a utilização dos catálogos *on-line*, das bibliotecas e dos repositórios digitais, que permitem aos usuários fazerem suas pesquisas remotamente, os problemas relacionados à recuperação da informação se tornaram ainda mais preocupantes. Hoje, o usuário não precisa ir até a biblioteca para fazer suas pesquisas e, conseqüentemente, não interage ou interage muito pouco com os bibliotecários. Nesse cenário, a interface de pesquisa do Sistema de Recuperação da Informação (SRI), com suas diferentes possibilidades de busca, exerce um papel intermediário entre o usuário e o acervo documental de uma instituição. Desse modo, o papel tradicional de mediador da informação, fortemente exercido pelos bibliotecários quando os catálogos eram manuais, está desaparecendo em função da utilização e disponibilização dos catálogos eletrônicos, que podem ser acessados a distância através da internet. Isso não quer dizer que os bibliotecários vão deixar de exercer seu papel de mediador. Porém, certamente terão de exercer essa função de outra maneira; terão de encontrar uma forma alternativa para auxiliar os diferentes usuários dos SRI a realizarem suas pesquisas com autonomia e sucesso. É pensando nessa alternativa que apresentamos nossa proposta. Acreditamos que a melhor maneira de auxiliar os usuários em suas atividades de busca de informação, no

cenário atual, é investir no aperfeiçoamento de metodologias e ferramentas de representação e recuperação da informação.

As linguagens documentárias são sistemas conceituais auxiliares das atividades de representação e recuperação do conhecimento. São instrumentos que possibilitam o controle do vocabulário utilizado para representação e registro dos assuntos dos documentos de um Sistema de Recuperação da Informação (BONOTTO et al., 2010). De acordo com Sales e Café (2008), tais linguagens são artificiais, constituídas por sistemas simbólicos, utilizadas para indexação, armazenamento e recuperação da informação em sistemas informacionais com a finalidade de descrever sinteticamente o conteúdo dos documentos— razão pela qual também são conhecidas como modelos de representação do conhecimento.

Em função dos diferentes tipos de linguagens existentes, é possível caracterizá-las por sua abrangência (em universais ou especializadas) e por sua forma de apresentação (em codificadas ou alfabéticas). As classificações bibliográficas como a Classificação Decimal de Dewey (CDD) e a Classificação Decimal Universal (CDU) são exemplos de linguagens codificadas universais, pois abordam todas as áreas do conhecimento e apresentam um código numérico (notação) para representar cada assunto. Tesouros e Listas de Cabeçalhos de Assunto são exemplos de linguagens documentárias alfabéticas, pois são organizados por ordem alfabética, podendo ou não apresentar uma estrutura sistemática de organização por classes e subclasses. Entre as linguagens documentárias alfabéticas de indexação existentes, os tesouros são considerados os sistemas mais complexos; são, geralmente, específicos de uma única área de conhecimento e apresentam as relações semânticas (hierárquicas, associativas e de equivalência) entre os termos de um domínio de forma alfabética e sistemática. Os cabeçalhos de assunto, por sua vez, são exemplos de linguagens documentárias alfabéticas universais, pois normalmente referem-se a várias áreas de conhecimento; quanto às relações semânticas, apresentam uma estrutura mais flexível, porém, pouco explícita. Por suas características, existem poucos publicados, sendo o Library of Congress Subject Headings (LCSH) da Library of Congress o cabeçalho de assunto em uso mais importante da atualidade.

As linguagens documentárias alfabéticas surgiram com o objetivo de padronizar a linguagem utilizada na descrição do conteúdo dos documentos e, com isso, garantir a precisão na recuperação da informação. São instrumentos que visam a precisão

terminológica no âmbito da representação do conhecimento e da recuperação da informação (SALES; CAFÉ, 2008). Busca-se, por intermédio das linguagens documentárias, evitar que documentos indesejados, que não correspondem à solicitação do usuário, sejam recuperados. Mas busca-se também, por outro lado, maior possibilidade de revocação a partir da utilização de termos equivalentes, para evitar que documentos desejados pelos usuários em suas buscas deixem de ser recuperados devido a utilização de variantes denominativas dos descritores. Contudo, o que se verifica, na prática, é que nem sempre as linguagens alfabéticas de indexação auxiliam os usuários a recuperar a informação que almejam. Muitas vezes, a incompatibilidade entre os termos atribuídos por bibliotecários aos documentos, durante o processo de representação do conteúdo dos mesmos (indexação), e a linguagem utilizada pelos usuários, durante o processo de busca de informação impossibilita a recuperação de documentos pertinentes. É exatamente com a finalidade de minimizar essa incompatibilidade entre os descritores e a linguagem dos usuários que apresentamos nesta tese uma proposta de metodologia para a seleção e organização de descritores e termos equivalentes para tesouros. Entendemos que os tesouros são constituídos predominantemente por termos, e que apenas alguns desses termos possuem o valor de descritor. Deste modo, os termos que não possuem o valor de descritor, denominados na literatura sobre elaboração de tesouros principalmente por "não descritores", "termos não autorizados" ou "termos não preferidos", são tão importantes quanto os descritores. No entanto, essa forma de expressá-los parece diminuir sua importância dentro da estrutura dos tesouros, por isso, optamos por chamá-los de "termos equivalentes".

É válido destacar, ainda, que esse problema de comunicação, entre bibliotecários e usuários, nos levou a escolher como objeto de estudo as lexias de buscas contidas nos "logs" de pesquisa dos usuários do Portal LexMI do Senado Federal Brasileiro, visto que, toda pesquisa realizada no Portal fica registrada em um arquivo log armazenado no servidor Web da instituição e, através de sua análise, é possível verificar a linguagem utilizada pelos usuários para a recuperação de informação.

Embora os tesouros apresentem em sua estrutura a possibilidade de inserção de remissivas para termos equivalentes, essa incompatibilidade entre a linguagem dos usuários e a linguagem de indexação ainda ocorre, principalmente, por serem privilegiadas a univocidade e a monorreferencialidade nas metodologias de

elaboração de tesouros existentes, o que gera como resultado um frágil sistema de remissivas de um modo geral; razão pela qual a função de controle terminológico nos tesouros tradicionais se sobrepõe à realidade discursiva dos usuários dos Sistemas de Recuperação da Informação, uma vez que a escolha dos termos que fazem parte dessas linguagens se dá predominantemente a partir do ponto de vista dos especialistas do domínio para o qual o tesouro é elaborado. Apesar de os especialistas fazerem parte do público que utiliza os sistemas de recuperação da informação, é preciso considerar que existem diferentes níveis ou graus de especialização, e os sistemas devem estar aptos a atender tanto os usuários mais especializados, como os leigos e os semiespecializados. Conseqüentemente, os tesouros devem ser elaborados visando o atendimento das necessidades informacionais de um público diversificado que possui diferentes níveis de conhecimento em relação à área de conhecimento abordada pelo tesouro.

Esse contexto nos leva às seguintes questões de pesquisa:

- a) Como atender às necessidades de busca de informação de usuários em SRI informatizados, uma vez que as pessoas utilizam diferentes recursos linguísticos e comunicativos para expressar o que desejam?
- b) Como tornar uma linguagem artificial e predominantemente conceitual como os tesouros mais próxima da linguagem natural utilizada pelos usuários de um SRI?

Para responder essas questões, desenvolvemos esta pesquisa com objetivo geral de propor uma metodologia para a identificação e seleção de termos equivalentes e descritores de tesouros que efetivamente considere a linguagem utilizada pelos usuários. Visamos, especificamente, identificar as variantes denominativas empregadas pelos usuários das áreas do Direito do Trabalho e do Direito Previdenciário, e, com isso, propor uma nova perspectiva de elaboração de tesouros que reúna aspectos conceituais e linguísticos da linguagem dos usuários para a seleção de termos equivalentes e descritores. Deste modo, esperamos tornar os tesouros, que são sistemas conceituais, em sistemas mais linguísticos e próximos da linguagem dos usuários. Para alcançar esse objetivo, realizamos a identificação e utilização das variantes denominativas existentes nas lexias de buscas de usuários de um SRI. Processo esse exemplificado, nesta tese, com a aplicação das lexias de buscas das áreas do Direito do Trabalho e do Direito Previdenciário dos usuários do Portal LexMI, administrado pelo Senado Federal do Brasil.

Sabemos que muitos fatores podem gerar problemas para a recuperação da informação, no entanto, nesta pesquisa, escolhemos trabalhar especificamente com os problemas decorrentes da variação terminológica. Esta decisão se deu tanto por influência de nossa prática profissional como bibliotecária, como pela análise da literatura e relatos de casos como o de Baker (2010), que informa que a sinonímia afeta setenta por cento das pesquisas realizadas por usuários no motor de buscas Google nos mais de cem idiomas suportados pela ferramenta. De acordo com o autor, a sinonímia torna os resultados das buscas ambíguos. Desse modo, é necessário verificar os demais termos empregados pelo usuário em sua estratégia de busca para contextualizar e compreender o sentido do termo solicitado e, com isso, oferecer uma resposta adequada às suas necessidades de informação.

A motivação para a elaboração de uma metodologia específica para a seleção e organização de termos equivalentes e descritores de tesouros que efetivamente considere a linguagem dos usuários surgiu a partir da identificação desse problema e da ausência de sua abordagem, tanto na literatura da área de organização do conhecimento como nas normas de elaboração de tesouros. As normas vigentes e antigas abordam a questão da validação do usuário de forma muito superficial, destacam a importância dessa validação para a escolha dos termos que farão parte de um tesouro, mas não mencionam, explicitamente, a metodologia a ser empregada para sua obtenção. Acreditamos que a falta de uma descrição metodológica mais aprofundada em relação a essa etapa de elaboração dos tesouros é uma das principais razões de sua deficiência como ferramenta de apoio à representação e à recuperação de informação. Em vista disto, pretendemos contribuir para o aperfeiçoamento dos tesouros, através da proposição de uma metodologia que explore diretamente a linguagem dos usuários para a obtenção de termos equivalentes e descritores, de modo que, no futuro, seja possível elaborar tesouros que privilegiem tanto aspectos conceituais, como aspectos linguísticos da linguagem especializada.

Outra motivação, decorrente da lacuna identificada, é a dificuldade que encontramos para explicar aos alunos do curso de Biblioteconomia da Universidade Federal do Rio Grande do Sul (UFRGS), no qual ministramos a disciplina de Linguagens Documentárias III, como essa etapa da elaboração de tesouros deve ser realizada. Em função das escassas explicações que encontramos no referencial teórico existente sobre esse aspecto, os alunos ficam dependentes da figura do

professor para a compreensão e desempenho dessa etapa da elaboração de tesouros tendo, conseqüentemente, a autonomia sobre o seu aprendizado afetada.

Isso implica na necessidade de revisão da metodologia de elaboração de tesouros existente. Apesar do desenvolvimento e da evolução das teorias que fundamentam a elaboração de tesouros, observamos que a prática metodológica de elaboração dos mesmos sofreu poucas alterações desde a primeira versão da Norma ISO 2788, que fala sobre a elaboração de tesouros monolíngües publicada em 1974 e atualizada em 1986, bem como da Norma ISO 5964, que fala sobre a elaboração de tesouros multilíngües publicada em 1985. Embora recentemente tenha sido publicada uma nova norma sobre elaboração de tesouros, ISO 25964, cuja primeira parte foi publicada em 2011 e a segunda parte publicada no início de 2013, observamos que as modificações apresentadas na norma ocorrem em função da necessidade de adequação dos tesouros ao contexto da Web Semântica. Porém, muitos aspectos teóricos e recomendações que independem das questões relacionadas à interoperabilidade (capacidade de comunicação entre dois ou mais sistemas para que possam trocar e utilizar informação; no caso dos tesouros, está relacionada, também, à abordagem adotada em relação ao domínio especializado a que os tesouros em questão pertencem) permanecem inalterados na nova norma.

A necessidade identificada de pesquisas que discutam a metodologia de elaboração de tesouros foi reforçada no II CONGRESSO BRASILEIRO DE ORGANIZAÇÃO E REPRESENTAÇÃO DO CONHECIMENTO, realizado pelo capítulo brasileiro da International Society for Knowledge Organization (ISKO Brasil) em 2013¹. Sobretudo no que se refere às necessidades de adequação metodológica para trabalhar com os termos específicos de um domínio durante a elaboração desses instrumentos.

A partir do contexto apresentado, destacamos que nossa pesquisa é de caráter teórico, mas visa a uma aplicação concreta de uso. Por isto, escolhemos o campo do Direito do Trabalho e do Direito Previdenciário para a realização da etapa aplicada do trabalho. Essas ramificações do Direito, por tratarem, respectivamente, da regulamentação das relações de trabalho na sociedade e das regras e direitos dos

¹Nesse evento pesquisadores da área de Ciência da Informação que trabalham com sistemas de organização do conhecimento, tais como Maria Luiza de Almeida Campos, Marisa Brascher e MarildaLopes Ginez de Lara, destacaram, em suas falas, durante os debates da seção sobre a Dimensão Epistemológica da Organização do Conhecimento, a necessidade de pesquisas que discutam a metodologia de elaboração de tesouros.

aposentados e pensionistas, estão mais próximas dos interesses de um amplo número de cidadãos. Logo, constituem-se em um campo bastante produtivo para os nossos interesses de pesquisa, em virtude da riqueza dos itens léxicos que apresentam.

A pesquisa que desenvolvemos se encontra descrita e estruturada ao longo de oito capítulos. A partir da discussão inicial realizada no texto introdutório desta tese sobre os problemas, objetivos e justificativas da pesquisa, apresentamos, no segundo capítulo, o percurso histórico da Biblioteconomia, Documentação e Ciência da Informação no mundo e no Brasil e as características da área de Organização e Representação da Informação e do Conhecimento no âmbito da Ciência da Informação.

No terceiro capítulo, discutimos as Teorias da Terminologia que julgamos pertinentes para a elaboração de tesouros. Destacamos que, embora nossa proposta seja baseada e desenvolvida sob o ponto de vista da Teoria Comunicativa da Terminologia (TCT), também apresentamos alguns aspectos da Teoria Geral da Terminologia (TGT) relevantes para o funcionamento adequado dos tesouros; visto que, buscamos o aperfeiçoamento da metodologia de seleção e organização de termos equivalentes e descritores dos tesouros a partir do estudo detalhado de variantes denominativas, disponibilizado pela TCT, mas não negamos a importância do controle terminológico, realizado pelos tesouros, viabilizado com o auxílio da TGT.

No quarto capítulo, descrevemos a relevância dos tesouros no contexto atual, suas características técnicas e as teorias que influenciam a elaboração dessas ferramentas até a atividade de indexação e análise crítica das normas internacionais de elaboração de tesouros vigentes no Brasil e no mundo, com ênfase na norma ISO 25964-1, publicada em 2011.

No quinto capítulo, descrevemos o percurso metodológico necessário à implementação de nossa proposta, ou seja, à aplicação da metodologia de seleção e organização de descritores e termos equivalentes para tesouros. A metodologia se desenvolve a partir da comparação das informações existentes em um *corpora* de características distintas formado, especificamente, por: a) lexias de buscas dos usuários do Portal LexMI, do Senado Federal Brasileiro, nas áreas do Direito do Trabalho e do Direito Previdenciário, obtidas a partir dos logs de pesquisa desses usuários, disponibilizados pela instituição para a realização desta pesquisa; b) leis brasileiras mais relevantes dessas duas áreas do Direito; e c) linguagem de indexação

utilizada pela Biblioteca do Senado Federal para a descrição do conteúdo dos documentos (Vocabulário Básico Controlado). O confronto de informações também é realizado a partir da Base de Dados de Jurisprudência do próprio Portal LexMI e de três tesouros da área jurídica: Thesaurus de Direito do Trabalho do Tribunal Regional do Trabalho da 4ª Região (TRT4), Thesaurus em Direito do Trabalho elaborado por Regina de Marco Ferreira (TEJUT) e o Vocabulário Jurídico (Tesouro) do Superior Tribunal de Justiça.

O sexto capítulo é constituído pelos resultados obtidos a partir da análise do *corpus* de *lexias* de buscas dos usuários do Portal LexMI. Essa análise nos permitiu identificar, entre as *lexias* variantes, três categorias que compreendem a) aspectos verbais; b) recursos não verbais e c) aspectos conceituais. Para cada categoria, apresentamos uma descrição e exemplos de ocorrência.

No sétimo capítulo apresentamos nossa proposta de metodologia para a validação da perspectiva do usuário para a seleção de termos e descritores de tesouros, constituída a partir do confronto entre aspectos teóricos e os resultados da pesquisa.

Concluimos o texto com algumas considerações sobre os resultados da pesquisa e reflexões, tanto acerca da relevância dos logs como objeto de estudo e fonte para a coleta de dados, quanto acerca da análise de variantes denominativas e suas tipologias para inserção de termos equivalentes e descritores em tesouros. Por fim, apresentamos a lista de referências citadas ao longo desta tese seguida de seus apêndices e anexos.

2 CIÊNCIA DA INFORMAÇÃO: aspectos históricos e aplicados

Para maior compreensão do contexto de aplicação e desenvolvimento desta tese, apresentamos, brevemente, nesse capítulo, o percurso histórico da Biblioteconomia, Documentação e Ciência da Informação no mundo e no Brasil. A partir dessa retrospectiva, descrevemos uma das subáreas específicas da Ciência da Informação, na qual se insere o tema desta pesquisa: a Organização e Representação da Informação e do Conhecimento que trata das questões relacionadas tanto com a própria representação e organização da informação e do conhecimento, como da recuperação da informação. Essa contextualização inicial embasa os demais itens desenvolvidos ao longo desse capítulo, que aborda e discute desde a relevância dos tesouros no contexto atual, suas características técnicas e as teorias que influenciam a elaboração dessas ferramentas, até a atividade de indexação e análise crítica das normas internacionais de elaboração de tesouros vigentes no Brasil e no mundo, com ênfase na norma ISO 25964-1, publicada em 2011.

2.1 Biblioteconomia, Documentação e Ciência da Informação: o que fomos e o que somos

Julgamos necessário discorrer, mesmo que sucintamente, sobre a história da Biblioteconomia, Documentação e Ciência da Informação, pois discutimos nesta pesquisa alguns aspectos da metodologia de elaboração de tesouros, ferramenta predominantemente utilizada pelos bibliotecários no exercício de suas atividades profissionais. Por essa razão, é importante esclarecer quem é esse profissional e como se constitui sua área de atuação. Deste modo, começamos esse percurso pelas seguintes perguntas norteadoras: porque ainda somos chamados de bibliotecários se nossa área de atuação, de acordo com o CNPq, é denominada de Ciência da Informação? Como isso é possível, já que as bibliotecas e os bibliotecários existem desde os anos anteriores a Cristo e a Ciência da Informação, como área de conhecimento, surgiu apenas muitos anos depois na segunda metade do século XX? Como poderemos verificar a seguir, alguns fatos históricos esclarecem essas dúvidas.

Registros históricos apontam a existência de bibliotecas desde os anos anteriores a Cristo. A descoberta da Biblioteca de Ebla, na Síria, em 1975, comprova, inclusive, a existência de coleções organizadas e separadas por assunto em estantes já no terceiro milênio a.C.(SAGREDO; NUÑO, 1994). Durante a Idade Antiga e a Idade Média, o acesso ao acervo das bibliotecas era possível apenas para os integrantes da nobreza e da igreja. Nesse período, as bibliotecas tinham, basicamente, a função de armazenar livros e manuscritos, porém, com o surgimento da imprensa e o conseqüente acúmulo de material gráfico, tornou-se essencial a organização e disposição dos mesmos por assunto. Essas características das bibliotecas como espaços que armazenam e organizam livros permaneceram inalteradas até a Idade Moderna. (SHERA; EGAN, 1961).

Os periódicos científicos e o constante crescimento da bibliografia especializada, cada vez mais difícil de analisar e organizar, em função da necessidade latente de técnicas de descrição e classificação diferentes daquelas utilizadas para a organização de livros e manuscritos, foram problemas abandonados pela Biblioteconomia e assumidos pela Documentação. Aliado ao fato de que as bibliotecas não tinham recursos humanos e financeiros para realizar a organização desses documentos, os bibliotecários foram tomados pela fé na educação das massas que se seguiu por influência do Iluminismo. (SHERA; EGAN, 1961). A necessidade de trabalhadores alfabetizados e treinados em atividades manuais específicas, para atender aos requisitos do mercantilismo crescente e da revolução industrial, reforçou a decisão dos bibliotecários de voltarem seus esforços para o benefício das massas e em favor da cultura popular. De modo que, em “ [...] 1850, a biblioteca pública já era considerada principalmente como uma agência educacional, um instrumento para a ampliação do processo educativo formal durante toda a vida adulta.” (SHERA; EGAN, 1961, p. 25).

Deste modo, enquanto a Biblioteconomia e os bibliotecários se dedicavam à democratização da educação, surgia a área da Documentação para dar conta da organização da literatura científica que crescia rapidamente no final do século XIX. Os advogados belgas Paul Otlet e Henry La Fontaine, fundadores dessa nova área de conhecimento, acreditavam que poderiam “[...] levar ao conhecimento de cientistas e interessados toda a literatura científica e todos os produtos do conhecimento gerados no mundo.” (OLIVEIRA, 2005, p. 10). Para isso, conceberam uma bibliografia universal internacional e planejaram a criação de uma biblioteca universal de

referência com o objetivo de divulgar os dados bibliográficos dos documentos indexados. (OLIVEIRA, 2005; SHERA; EGAN, 1961). Como a biblioteca idealizada por Otlet e La Fontaine era de referência, os mesmos começaram a desenvolver ferramentas para registrar de forma sistemática e padronizada as referências dos documentos. Por isso, inicialmente utilizaram a Classificação Decimal de Dewey (CDD), que foi modificada e ampliada de várias formas, para atender às suas necessidades específicas de análise de assunto. Essas modificações originaram a Classificação Decimal Universal (CDU), que possibilita a organização de outros tipos de documentos além dos livros. (OLIVEIRA, 2005; SHERA; EGAN, 1961). Nesse sentido, é importante destacar a elaboração, por Paul Otlet, do conceito de documento que, de acordo com Oliveira (2005, p. 11), “[...] ampliou o campo de atuação dos profissionais da área ao ultrapassar os limites do espaço da biblioteca e agregar novas práticas de organização e novos serviços de documentação.”.

Logo, especificamente em relação à organização de acervos, o que diferencia a Biblioteconomia da Documentação é que, enquanto a Biblioteconomia desenvolveu técnicas para a organização e classificação apenas de livros e manuscritos, a Documentação expandiu e criou novas técnicas para a organização e classificação de qualquer tipo de documento; visto que, com o crescimento acelerado da literatura científica, o livro já não era mais a principal fonte de informação daqueles que se dedicavam à pesquisa pura na indústria, no comércio e no governo. Diante da necessidade de organizar e usar relatórios especiais manuscritos, cópias fotostáticas, microfilmes, microfichas, fotografias, plantas, esquemas e outros suportes de informação, que surgiam em grande número, foi necessário desenvolver um serviço mais adequado de catalogação e preparo de resumos para atender aos objetivos da Documentação (SHERA; EGAN, 1961).

O desenvolvimento de novos serviços também marca a mudança de paradigmas em relação ao conhecimento registrado. Enquanto a Biblioteconomia ficou presa ao livro e à biblioteca como instituição social, a Documentação voltou sua atenção ao conteúdo dos documentos e, conseqüentemente, às necessidades dos usuários que os utilizam. De acordo com Oliveira (2005), a excessiva preocupação das bibliotecas em armazenar e manter acervos, considerando o livro em si mais importante do que seu conteúdo, e a menor preocupação com as necessidades informacionais dos usuários, fragilizaram a Biblioteconomia. A autora esclarece que, apesar das pesquisas existentes sobre os usuários, a metodologia utilizada sempre

esteve voltada para a avaliação dos serviços da biblioteca e não para os problemas desses usuários. A Documentação, por sua vez, ao se concentrar no conteúdo dos documentos, continuou se expandindo e começou “[...] a trabalhar nos novos métodos de ‘publicação’ e a realizar experiências com máquinas para a organização mecânica e eletrônica desse material.” (SHERA; EGAN, 1961, p. 35-36). Nos anos de 1950, essas experiências, somadas ao êxito das ideias da engenharia de comunicação e teorias cibernéticas para a representação das propriedades dos sistemas de transmissão de sinais em termos matemáticos, foram a base das tentativas de caracterização e modelagem do processo de recuperação da informação e/ou documento (OLIVEIRA, 2005).

O desenvolvimento de novos serviços e o impulso tecnológico propiciado pelo surgimento da automação de sistemas provocaram o crescimento da Documentação como área de conhecimento que, a partir de 1962, passou a se chamar Ciência da Informação. Logo, a Ciência da Informação é uma evolução da Documentação, uma área de conhecimento interdisciplinar, cujos paradigmas estão calcados no conteúdo dos documentos e na recuperação da informação para atender às necessidades de seus usuários.

No Brasil, assim como na história, a Biblioteconomia foi a primeira a surgir, a partir da criação do primeiro curso para a formação de profissionais da área pela Biblioteca Nacional do Rio de Janeiro, em 1911. A Ciência da Informação no País foi introduzida muitos anos depois, em 1970, com a implantação do curso de mestrado em Ciência da Informação pelo Instituto Brasileiro de Bibliografia e Documentação (IBBD), que atualmente se denomina Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT). (RUSSO, 2010). Nos dias atuais, ainda temos o predomínio da formação profissional de bibliotecários em cursos de nível superior, na graduação, e a capacitação formal em Ciência da Informação, no âmbito dos Programas de Pós-Graduação. Porém, é importante destacar que os conhecimentos referentes à Ciência da Informação também fazem parte do currículo dos cursos de graduação em Biblioteconomia existentes. A partir de uma rápida análise dos currículos desses cursos, é possível distinguir facilmente os conteúdos referentes à Biblioteconomia e à Ciência da Informação, razão pela qual é possível afirmar que, no Brasil, bibliotecários e cientistas da informação são duas faces do mesmo profissional que continua sendo chamado apenas de bibliotecário, principalmente em função dos direitos adquiridos a

partir da Lein° 4.084, de 30 de junho de 1962, que protege e regulamenta o exercício dessa profissão em nosso País.

2.2 A Organização e Representação da Informação e do Conhecimento e os Sistemas de Recuperação da Informação

Pinho (2009) explica que a área de Organização e Representação do Conhecimento investiga os fundamentos científicos, as habilidades e os instrumentos que auxiliam o profissional nas atividades de extração, descrição, nomeação e rotulagem do conhecimento que será objeto de sistemas de recuperação da informação. Seu principal objetivo, de acordo com Lima e Alvares (2012, p. 39), “[...] é a recuperação de objetos informacionais, que são as informações registradas nos diferentes suportes existentes (texto, imagem, registro sonoro, mapas, páginas da web, etc.)” Por isso, “[...] as atividades de organização e representação do conhecimento são o cerne da atuação do profissional da informação.” (PINHO, 2009, p. 18).

Dahlberg (2006) considera diferentes o conceito e aplicação dos termos, organização do conhecimento e organização da informação. Para a autora, a organização do conhecimento se dá quando elaboramos sistemas conceituais, como os tesauros; já a organização da informação ocorre quando utilizamos esses sistemas conceituais para descrever o conteúdo dos documentos, ou seja, ocorre quando realizamos a indexação das obras de um acervo. Deste modo, a organização da informação está relacionada aos meios de recuperar a própria informação e, portanto, também ao arranjo de acervos, tradicionais ou eletrônicos, através da descrição do assunto dos documentos (LIMA; ALVARES, 2012). Bräscher e Café (2010, p. 92) definem organização da informação como “[...] um processo que envolve a descrição física e de conteúdo dos objetos informacionais. ” Para as autoras, a representação da informação é o produto dessa descrição que deve ser “[...] entendida como um conjunto de elementos descritivos que representam os atributos de um objeto informacional específico. ” (BRÄSCHER ; CAFÉ, 2010, p. 92). Ou seja, a organização da informação refere-se à organização de suportes físicos nos quais estão contidas as informações. A organização do

conhecimento, por outro lado, refere-se aos conceitos contidos nos documentos, “[...]visa à construção de modelos de mundo que se constituem em abstrações da realidade.” (BRÄSCHER ; CAFÉ, 2010, p. 93). As autoras descrevem a organização do conhecimento como

[...] o processo de modelagem do conhecimento que visa a construção de representações do conhecimento. Esse processo tem por base a análise do conceito e de suas características, para o estabelecimento da posição que cada conceito ocupa num determinado domínio, bem como das suas relações com os demais conceitos que compõem esse sistema nocional. (BRÄSCHER ; CAFÉ, 2010, p. 95).

Por isso, para as autoras, a representação do conhecimento é uma estrutura conceitual

[...] feita por meio de diferentes tipos de sistemas de organização do conhecimento (SOC), que são sistemas conceituais que representam determinado domínio por meio da sistematização dos conceitos e das relações semânticas que se estabelecem entre eles. (BRÄSCHER ; CAFÉ, 2010, p. 96).

Logo, entendemos que, no âmbito da Ciência da Informação, a representação, a organização e a recuperação da informação fazem parte de um fluxo contínuo em que cada etapa depende das demais para o seu sucesso. Representamos a informação contida nos documentos com o objetivo de organizá-la e, posteriormente, recuperá-la. Por esse motivo, não se pode pensar nessas três atividades isoladamente, já que são concomitantes, no sentido de que a decisão tomada durante a representação vai determinar a organização e, conseqüentemente, interferir na recuperação da informação. Em um SRI, a qualidade da recuperação da informação depende dos procedimentos utilizados durante o processo de organização da informação. Por isso, os padrões de organização devem ser definidos desde a concepção do sistema (LIMA, ALVARES, 2012).

Conseqüentemente, quando planejamos um sistema ou um instrumento como os tesouros, cuja função é auxiliar na representação da informação (indexação), é preciso ter em mente que as decisões tomadas (como categorias, subcategorias, relações de equivalência e associação) irão, inevitavelmente, afetar a forma de organização da informação, sua estrutura de apresentação, e a maneira de recuperá-la. Ou seja, ao determinar a forma de entrada dos dados em um SRI, também estamos determinando sua organização e a forma de

recuperação. Esse processo depende e é influenciado por diferentes variáveis, tais como: público-alvo do SRI, contexto socioeconômico, contexto cultural, recursos tecnológicos, recursos humanos, físicos e financeiros disponíveis, entre outros.

Para muitos profissionais e pesquisadores do campo da Biblioteconomia e da Ciência da Informação, a indexação, além de ser muito relevante, é uma das principais atividades que exercem. Guedes e Dias (2010, p. 42) definem indexação como “[...] um conjunto de procedimentos com objetivo de expressar/representar o conteúdo temático de documentos através de linguagens de indexação ou documentárias visando à recuperação posterior.” Lancaster (2004) explica que a indexação é realizada principalmente em duas etapas; são elas: a análise conceitual e a tradução. Durante a análise conceitual, ocorre a identificação do assunto do documento. Já na etapa de tradução, o assunto identificado a partir da análise conceitual é transformado em termos de indexação. A escolha dos termos de indexação, por sua vez, pode ocorrer a partir da extração de palavras que ocorrem no documento ou através da atribuição de termos extraídos de outra fonte. Para Lancaster (2004), o mais frequente é que os termos sejam atribuídos de algum vocabulário controlado, mas também pode acontecer de os mesmos serem selecionados arbitrariamente, conforme a percepção do indexador, o que Rowley (2002) denomina linguagem de indexação livre.

Convencionalmente, a indexação pode ser realizada com o suporte de três tipos de linguagens, descritas por Rowley (2002, p. 169) da seguinte forma:

- a) Linguagem de indexação controlada: definida como um conjunto de termos autorizados para uso na indexação do assunto de documentos. É subdividida em dois tipos: as linguagens alfabéticas de indexação, como os tesouros e listas de cabeçalhos de assunto; e os sistemas de classificação, representados por código ou notação.
- b) Linguagem de indexação natural: refere-se a quaisquer expressões que ocorram em alguma parte do documento. Todos os termos no corpo do documento são candidatos a serem termos de indexação.
- c) Linguagem de indexação livre: para esta linguagem não existem limitações quanto aos termos a serem utilizados no processo de indexação. Sua diferença em relação a linguagem natural é que os

termos utilizados para a indexação são selecionados de maneira arbitrária e não ocorrem no texto do documento.

O tipo de linguagem de indexação a ser adotada por uma unidade de informação pode variar de acordo com a área de conhecimento ou o tipo de instituição, especialmente pelo fato de que em algumas áreas do conhecimento, por exemplo, não existem tesouros ou cabeçalhos de assunto publicados. A vantagem de usar linguagens controladas é que elas conferem maior qualidade à indexação e possibilitam a manutenção da consistência da representação da informação. A desvantagem de utilizá-las, no entanto, é que nem sempre, ou nunca, o usuário emprega esses termos de indexação em sua busca e, com isso, caso o SRI não tenha um sistema de remissivas integrado, ou seja, um dispositivo que liga os termos utilizados na indexação com as variantes dos mesmos, o usuário provavelmente ficará sem resposta para a sua busca. Em relação às linguagens naturais, bastante utilizadas em unidades de informação de um modo geral, como os termos são extraídos direto dos textos dos especialistas, há um grande risco de “poluir” o SRI e torná-lo inconsistente, em função dos diferentes tipos de variações existentes. Pela mesma razão, as linguagens de indexação livres também apresentam um grande risco à consistência do SRI, uma vez que sequer a linguagem dos especialistas da área de conhecimento do documento a ser indexado é consultada.

O processo de indexação envolve tomadas de decisões em diferentes níveis, tais como:

- a) nível de exaustividade – quantidade de termos utilizados para a representação dos assuntos de um documento - pode ser determinada conforme o tipo de documento;
- b) nível de especificidade – refere-se a intensão das características de um termo, de modo que, sua especificidade possibilita, ao mesmo tempo, maior precisão e menor revocação à recuperação da informação;
- c) tipo de linguagem a ser adotada - ou seja, se livre, natural ou controlada (lembrando que o uso de uma linguagem livre, sem padronização, exige mais tempo e trabalho para a busca e recuperação da informação pelo usuário);
- d) recursos disponíveis para a gestão do acervo - capacidade de revocação e precisão, características igualmente importantes, pois expressam a

capacidade de filtragem do sistema em deixar passar o que é solicitado e impedir o que não é solicitado. (CARNEIRO, 1985).

No mesmo sentido, Chaumier (1988) afirma que a indexação é essencial para que os documentos possam ser recuperados de um acervo e, ao mesmo tempo, possam oferecer uma resposta adequada, a toda solicitação dos usuários, sem que haja 'ruídos' (ou seja, sem oferecer como resposta documentos que não correspondem à questão de pesquisa do usuário), nem 'silêncios' (quando o sistema não recupera um documento que existe no acervo). A ausência de precisão, conforme Moreiro González (2004), tem sido comum em qualquer sistema de recuperação baseado em palavras, justamente em função da ambiguidade presente nelas. Para o autor, o problema é que a grande maioria dos sistemas realiza uma busca léxica, pela simples ocorrência das palavras no registro bibliográfico, e não semântica, ou seja, através do conceito representado pelo termo de um domínio, fator que torna ainda mais complexa a indexação e a recuperação da informação.

Para Guedes e Dias (2010), a indexação também pode ser analisada a partir da perspectiva do agente executor do processo. Rafferty e Hilderley² (*apud* GUEDES; DIAS, 2010, p. 45) distinguem três grupos de candidatos a atores no processo de indexação, são eles:

- a) Indexação orientada por especialistas: baseia-se no tratamento da informação através da intervenção de intermediários (bibliotecários, indexadores, editores voluntários), é a indexação feita por especialistas, sendo mais dispendiosa e cara.
- b) Indexação orientada pelo autor: esta abordagem pressupõe que o autor irá utilizar termos que são comumente compreendidos e geralmente aceitos. Um problema que essa abordagem enfrenta é o fato de o autor não ser necessariamente um gestor de informação, com os conhecimentos profissionais de um especialista.
- c) Indexação orientada pelo usuário: esse tipo de indexação possibilita um elevado nível de interação com a comunidade que, provavelmente, não

² RAFFERTY, Pauline; HIDERLEY, Rob. Flickr and democratic indexing: dialogic approaches to indexing. **Aslib Proceedings**, v. 59, Issue 4/5, 2007. p. 397-410. Disponível em: <<http://www.emeraldinsight.com/Insight/viewPDF.jsp?Filename=html/Output/Published/EmeraldFullTextArticle/Pdf/2760590407.pdf>>. Acesso em: 1º maio 2008.

seria possível se tivesse decisões a serem tomadas sobre códigos, convenções e regras que regem qualquer taxonomia controlada.

A indexação orientada por especialistas é utilizada principalmente em SRI de instituições. Contudo, destacamos que a indexação orientada pelo autor e pelos usuários são métodos bastante utilizados na internet, por exemplo, em repositórios digitais e em ambientes caracterizados como pertencentes a *Web 2.0*, embora esse tipo de indexação seja problemático, do ponto de vista da consistência e qualidade da representação das informações contidas nos documentos. Brown³ *et al.* (1996) *apud* Lancaster (2004), chamam a atenção para a necessidade de um tratamento democrático da indexação, em que os usuários acrescentariam aos registros termos de sua própria escolha, quando isso fosse necessário e apropriado. Acreditamos que, diante da riqueza vocabular existente no contexto discursivo de autores e usuários de sistemas de informação, seria uma grande evolução se os SRIs pudessem realizar uma indexação colaborativa entre bibliotecários, autores e usuários. Visto que, termos e conceitos de uma área de conhecimento constituem-se como tais a partir da interação e negociação de sentidos entre os sujeitos. Nesse sentido Gaudin (2005, p. 86, tradução nossa) explica que “[...] não há a palavra certa em si. Há somente palavras apropriadas a interações definidas.”.

A partir dessa breve apresentação, constatamos que, para atender às necessidades de busca de informação dos usuários – considerando os diferentes tipos de variações existentes, em consonância com padrões de controle terminológico, a fim de manter sua consistência e eficácia– é necessário que os SRIs sejam aperfeiçoados. É importante salientar que, em termos de desenvolvimento tecnológico, já existem *softwares*, como o ALEPH, por exemplo, com ferramentas capazes de melhorar consideravelmente a recuperação da informação nos SRIs, tais como as redes de remissivas que ligam os descritores aos termos equivalentes (variantes denominativas) facilitando a recuperação da informação pelo usuário. Contudo, infelizmente, do ponto de vista prático, o uso dessas ferramentas ainda é limitado e muitas instituições, ainda que disponham deste recurso, não o utilizam. A subutilização de *softwares* que permitem a elaboração de redes de remissivas reforça

³ BROWN, Pauline *et al.* The democratic indexing of images. **New Review of Hypermedia and Multimedia**, v. 2, n. 1, p. 107-120, 1996.

a necessidade de maior domínio e compreensão de um conjunto de princípios teóricos que podem contribuir para o aperfeiçoamento dos sistemas de recuperação da informação, tais como as técnicas de elaboração de tesouros e o estudo do fenômeno da variação no âmbito da Terminologia, a título de exemplificação.

3 FUNDAMENTOS TEÓRICOS DA TERMINOLOGIA PARA A ELABORAÇÃO DE TESAUROS

A Terminologia e a Ciência da Informação nasceram e se desenvolveram vinculadas à resolução de problemas práticos da comunicação: no caso da Terminologia, especificamente, à solução de problemas linguísticos relacionados ao emprego de termos técnico-científicos (CABRÉ, 2001a); no caso da Ciência da Informação, à solução de problemas relacionados à recuperação da informação, no âmbito da organização da informação e do conhecimento.

Como resultado de sua evolução teórica, os pesquisadores do campo da Terminologia passaram a se preocupar mais com os aspectos linguísticos dos termos e de seu uso real em oposição à padronização de termos fixados por um suposto consenso de especialistas em âmbito universal. O campo da Ciência da Informação, por sua vez, também avançou, a ponto de, recentemente, passar a valorizar e estimular mais as relações associativas entre os termos dos tesauros, através do incentivo à explicitação das relações semânticas entre os termos; de modo que, nos futuros tesauros, seja possível identificar claramente o tipo de relação associativa existente entre os termos. Em consonância a essa realidade, a norma ISO 25964-1(2011) sugere que essas relações passem a ser explicitadas de forma mais detalhada, possibilitando ao usuário identificar se dois termos são associados em função de uma relação de oposição, de causa e efeito, de produto e produtor, etc. A norma ISO 25964-1(2011) também sugere maior explicitação das relações de equivalência a partir da indicação do tipo de equivalência existente entre os termos tais como: sigla e termo completo, termo corrente e termo obsoleto, entre outros. Os tesauros atuais permitem identificar, claramente, apenas as relações hierárquicas.

Para que a Ciência da Informação possa continuar avançando e aprimorando suas bases teóricas para a elaboração de tesauros, é necessário que ela também busque maior inserção de aspectos linguísticos na estrutura dessas linguagens documentárias e passe a considerar o uso real dos termos pelos usuários dos sistemas de recuperação da informação. Embora a literatura da área mencione a importância da garantia do usuário, também chamada de validação do usuário, conforme identificamos nos textos de Barité (2009); Oliveira e Araújo (2012); Svenius (2000), e nas normas ISO 25964-1 (2011) e ANSI/NISO Z39.19 (2005), entre outros;

não encontramos uma descrição detalhada em relação à metodologia necessária para a efetiva realização dessa atividade. E, ainda que os aspectos linguísticos possam ser contemplados através das relações de equivalência já existentes nos tesouros, o que se observa, na prática, é que não há ênfase sobre a necessidade de explorar a variação dos termos no processo de elaboração dos mesmos.

Como a Terminologia – sendo campo dos estudos linguísticos –, apresenta maior desenvolvimento teórico (comparada à Ciência da Informação) quanto à descrição e ao comportamento dos termos, pode oferecer os subsídios necessários para o desenvolvimento da pesquisa que aqui se delinea. Deste modo, destacamos, a seguir, os aspectos da Teoria Geral da Terminologia, que continuam vigentes para a elaboração de tesouros, e as características da Teoria Comunicativa da Terminologia, que fundamentam nossa proposta de seleção de termos e descritores de tesouros, com o objetivo de ampliar os aspectos linguísticos dessa linguagem específica.

3.1 Teoria Geral da Terminologia e Teoria Comunicativa da Terminologia

A Terminologia, como disciplina científica, que estuda os termos de uma área de especialidade, se delineou apenas em 1931, a partir da tese de doutorado de Eugen Wüster. O autor estabeleceu as bases da chamada Escola (Terminológica) de Viena e, posteriormente, elaborou a Teoria Geral da Terminologia (TGT). (BARROS, 2004).

O propósito de Wüster era normatizar as terminologias, com o objetivo de eliminar as ambiguidades e, desta forma, garantir a perfeita comunicação técnica e científica em âmbito internacional no campo das ciências e das técnicas. Para o autor, tal objetivo poderia ser alcançado através da univocidade obtida pela normatização, ou seja, estabelecendo o uso de um termo e descartando a utilização de outros que referem para o mesmo conceito. Wüster (1998, p. 21) propõe que “Uma unidade terminológica consiste em uma ‘palavra’ à qual se atribui um conceito como seu significado [...]”. Krieger (2001a, p. 120) explica que o autor “[...] compreende o termo como um signo linguístico no sentido saussureano, isto é, uma entidade composta da dupla face de significante / significado”. Por isso, o autor referia-se a termo, compreendendo apenas sua parte significante, designando-o como denominação,

bem como responsabilizando o conceito pela atribuição do estatuto de termo a uma palavra da língua (KRIEGER, 2001a). Essa prevalência do componente conceitual sobre o linguístico explica o direcionamento inicial dos estudos terminológicos, nos quais os conceitos veiculados pelos termos constituem os objetos que interessam às comunicações especializadas, e, conseqüentemente, à teoria da Terminologia (KRIEGER; FINATTO, 2004).

Wüster (1998) preocupou-se em propor princípios de univocidade para o trabalho terminológico com o objetivo de normatizar os termos das áreas técnico-científicas. Para o autor, a padronização, através da escolha consensual de especialistas para fixar um termo a um conceito, evitaria a ambigüidade dos termos e, por sua vez, facilitaria a comunicação em âmbito internacional no domínio das ciências e das técnicas. No entanto, ao considerar que os termos técnico-científicos não são elementos naturais dos sistemas linguísticos, a TGT “[...] anula a dimensão comunicativa das terminologias, a despeito de serem estes elementos das línguas naturais, cujo diferencial é o fato de expressarem conhecimentos especializados no campo das ciências e das técnicas.” (KRIEGER, 2000, p. 215). A negação da naturalidade dos termos aos sistemas linguísticos, de acordo com Krieger (2000), se deve à crença do artificialismo das terminologias marcado pela “[...] forte presença de formantes gregos e latinos nas nomenclaturas das ciências taxionômicas, como é o caso da Botânica, da Zoologia e da Química, entre outras.” (KRIEGER, 2000, p. 215). Ainda de acordo com a autora, a forma particular de designação das ciências fundadoras corrobora para o entendimento equivocado de que as terminologias não pertencem à língua, “[...] muito embora a utilização de nomenclaturas consiste num recurso para provocar efeitos de precisão conceitual e assegurar uma almejada exclusividade designativa nos diferentes campos de conhecimento.” (KRIEGER, 2000, p. 216).

No que se refere à elaboração de tesouros, o caráter normativo dos princípios da TGT permanece válido e ainda é utilizado no âmbito da Ciência da Informação. Na medida em que, os tesouros são ferramentas de controle terminológico e sua metodologia de elaboração recomenda a escolha de apenas um termo para desempenhar a função de descritor quando estamos diante de dois termos variantes entre si, podemos afirmar que suas características classificatórias e normativas ainda são importantes e válidas no contexto atual. Para a organização eficaz de um SRI, a padronização e o controle terminológico são fundamentais para garantir a qualidade

da resposta do sistema durante o processo de recuperação da informação. Porém, essas características precisam ser trabalhadas de diferente forma na metodologia de elaboração de tesouros, para não apagar o potencial comunicativo dessa ferramenta. Dessa forma, percebe-se que os problemas de comunicação não podem ser resolvidos apenas com a padronização. As metodologias tradicionais de elaboração de tesouros priorizam a classificação e a padronização terminológica em detrimento de aspectos linguísticos e comunicativos. Ou seja, a concentração das características classificatórias e normativas dificulta a utilização dos tesouros, afastando-os do seu papel de mediador no processo de comunicação entre usuários e sistemas de organização da informação.

Outro aspecto importante da TGT é a descrição das características dos conceitos para a organização do conhecimento. Wüster (1998, p.39) define conceito como “[...] um elemento do pensamento.” Para o autor, todo o conceito, com exceção dos conceitos de objetos individuais, “[...] corresponde aos elementos comuns que os seres humanos percebem em um grande número de objetos e que utilizam como meio de classificação mental (para entender) e, por conseguinte, também para se comunicar.”

Dahlberg (1992) destaca sua admiração pelo trabalho de Wüster e a importância do estabelecimento de sistemas conceituais terminológicos, como o idealizado pelo autor, já em 1970. Dahlberg (1992) explica que Wüster estudou os sistemas de classificação universal existentes e descobriu que eles estavam repletos de classes que apresentavam uma pré-combinação de conceitos, e não apenas conceitos elementares, mas temas reais, como o que foi encontrado nos títulos de documentos. Wüster (1998) chamou esses sistemas de “Classificações Temáticas” por serem diferentes dos sistemas conceituais. Para o autor, “Um conjunto ordenado de conceitos não constitui um sistema conceitual se, entre os elementos do sistema, não existe uma relação conceitual do tipo lógica ou ontológica, mas somente uma relação temática.” (WÜSTER, 1998, p. 57). Embora os temas também sejam conceitos, para o autor, as relações temáticas são relações de ocorrência (aparição ocasional em um documento), que não afetam a essência dos conceitos de base. Essas “[...] ‘classes’ eram terminologicamente inúteis pois, como eram compostas de várias palavras, sua posição hierárquica não poderia ser definida com clareza. (DAHLBERG, 1992, p.70, tradução nossa).

A descrição de Wüster em relação à natureza dos conceitos é particularmente importante para a elaboração de linguagens documentárias. A noção de intensão e extensão conceitual, e as relações entre os conceitos, caracterizadas por ele como lógicas e ontológicas, são amplamente utilizadas na elaboração de tesouros e estão presentes na Teoria do Conceito de Dahlberg (1978).

Na última década do século XX, a Terminologia iniciou um novo percurso no qual o entendimento sobre a prevalência do componente conceitual sobre o linguístico se modificou. Nesse sentido, Sager (1993, p. 36, tradução nossa) afirma que “A terminologia trata de conceitos e, portanto, de estruturas de conhecimento apenas na medida em que estão representados no léxico da língua.” Logo, a percepção de que os aspectos linguísticos são tão importantes quanto os conceituais, para a Terminologia, fez com que surgissem proposições inovadoras sob o enfoque do funcionamento linguístico dos termos.

De acordo com Cabré (2001b), a terminologia é necessária devido a sua dupla função de representação e transferência do conhecimento especializado. As áreas do conhecimento, ao se constituírem como especializadas, produzem um conjunto de termos que são utilizados com a finalidade representacional e comunicacional de fixação e circulação do conhecimento. Esses termos, encontrados principalmente nos textos especializados e utilizados na comunicação técnico-científica, são considerados pelos especialistas como o reflexo formal da organização conceitual de uma área, bem como um meio de expressão e comunicação profissional (CABRÉ, 1999).

Destaca-se, em oposição à dicotomia entre pensamento e linguagem da TGT, a Teoria Comunicativa da Terminologia (TCT), da qual Maria Teresa Cabré é a principal fonte de referência. Com fundamentos teóricos diferentes, a TCT valoriza os aspectos comunicativos das linguagens de especialidade e compreende as unidades terminológicas como parte da linguagem natural e da gramática das línguas (KRIEGER; FINATTO, 2004).

Para Cabré (1999, p. 119), o que se busca com essa teoria é explicar os termos como unidades singulares, às vezes similares a outras unidades de comunicação, admitindo variação conceitual e denominativa e considerando a dimensão textual e discursiva dos termos. A autora destaca alguns pontos como base para a TCT; são eles:

- a) a concepção de que um termo é poliédrico, ou seja, tem uma dimensão linguística, uma cognitiva e uma social;

- b) o princípio do caráter comunicativo da terminologia, de modo que toda a unidade terminológica apresenta de alguma forma uma finalidade comunicativa;
- c) o reconhecimento da variação, que se baseia no fato de que o exame dos contextos de ocorrência mostra que pode haver mais de uma maneira de denominar um conceito (sinonímia) ou situações de polissemia;
- d) a condição de linguagem natural;
- e) a condição de comunicação especializada, entre outros.

Entre os pontos indicados, destacamos a possibilidade de explorar o princípio da variação, problemática que entendemos como um dos principais causadores de falha na comunicação entre os usuários e os SRIs. Para Faulstich (1998), os termos são signos que encontram sua funcionalidade nas linguagens de especialidade e, com isso, passam por evoluções e variam em diferentes situações comunicativas. Logo, uma abordagem linguística da Terminologia supõe que os termos são estudados em sua dimensão discursiva.

Ao reconhecer o caráter comunicativo do termo, sua poliedricidade e o pertencimento à linguagem natural, é possível ressaltar, através da TCT, a perspectiva comunicacional do termo e, em alguns casos, sua evolução linguística. Um conceito de um domínio especializado mantém relações de diferentes tipos com os demais conceitos do mesmo domínio, constituindo a estrutura conceitual de sua área. Porém, tendo em vista que cada domínio pode ser estruturado de diferentes formas, os conceitos podem participar em mais de uma estrutura com o mesmo ou diferente valor (CABRÉ, 1999). Esse é o motivo pelo qual Cabré (1999) afirma que os termos não pertencem a um domínio, mas eles são usados em uma área com valor singularmente específico.

A Teoria Comunicativa da Terminologia oferece o embasamento teórico necessário para o desenvolvimento de nossa proposta, pois ela explica a linguagem especializada como um conjunto de possibilidades que se concretizam no e conforme o contexto comunicacional. Ao considerar a existência e explicar a funcionalidade da variação terminológica, no âmbito da comunicação especializada, a TCT traz uma nova perspectiva, que pode ser aplicada para a seleção de termos de tesouros; considerando que as variantes são pouco exploradas pela metodologia tradicional de elaboração desses instrumentos, o que, conseqüentemente, gera a subutilização das relações de equivalência dos tesouros, de modo que, todo o potencial de

desenvolvimento e acréscimo de aspectos linguísticos, propiciado pelas relações de equivalência, fica relegado. Especificamente no âmbito dos sistemas de recuperação de informação, é possível perceber rapidamente a importância da variação, tendo em vista que, da mesma maneira que pode ser utilizada como um recurso para facilitar a recuperação de informação, também pode prejudicar e dificultar esse processo para o usuário. Contudo, antes de discutirmos as consequências positivas ou negativas da variação, é necessário entender o que é um termo e como ele se constitui, para que possamos avaliar os impactos da variação para a recuperação da informação.

Os termos possuem duas funções: a de representar o conhecimento e a de possibilitar a sua transferência. De forma resumida, pode-se dizer, então, que um termo representa conceitos de uma área de especialização, transmite conhecimento especializado e possui caráter linguístico em função de sua inserção no discurso. Krieger (2001a) explica que uma unidade lexical pode assumir o valor de termo em função dos fundamentos, princípios e propósitos de uma área, por isso, para uma descrição detalhada das configurações terminológicas, é preciso considerar a diversidade de elementos constitutivos do termo. Além dos aspectos textuais e pragmáticos, a autora destaca a importância de relacionar componentes de natureza morfológica, sintática e semântica, embora a identificação desses, por si só, não seja “[...] suficientemente capaz de oferecer subsídios para dar conta das peculiaridades constitutivas dos termos, dado o papel da textualidade nesse processo.” (KRIEGER, 2001a, p.129). Sob a ótica do sistema e da linguagem, a autora sistematiza as principais características da unidade lexical terminológica sob o prisma de constituintes sistêmicos relativos à morfossintaxe das terminologias, destacando que se trata de configurações teoricamente comuns a qualquer área de conhecimento, conforme se pode observar no quadro a seguir.

Quadro 1 – Termo: configurações morfossintáticas prototípicas

Nomes

Sintagmas terminológicos

Outras configurações sígnicas exemplificadas:

- Siglas (ONU)
- Abreviaturas (set./setembro)
- Fórmulas – H₂O

Fonte: Krieger (2001a, p. 129).

Krieger (2001a) explica que este quadro foi estruturado com base na dimensão linguística dos termos, apresentando dados que correspondem à vertente formal, visto que são aspectos morfológicos e sintáticos que se sistematizam. Quanto às categorias gramaticais, a unidade lexical especializada é principalmente nominal, predominando substantivos com raros adjetivos e verbos, embora estes também possam ser unidades terminológicas (KRIEGER, 2001a).

Outro aspecto importante dos termos é quanto à sua formação, que, tipologicamente, pode ser simples, complexa ou composta. Conforme Barros (2004), os termos simples são aqueles formados por uma única unidade lexical, como, por exemplo, “trabalhador”; os termos complexos são aqueles formados pela união de duas ou mais palavras para representar um único conceito, como, por exemplo, “carteira de trabalho”; e os termos compostos são aqueles formados pela união de duas ou mais palavras através do uso de hífen, como, por exemplo, “mão-de-obra”. Krieger (2001) destaca que as unidades lexicais simples e complexas coexistem nos textos especializados, “Entretanto, há vários estudos que comprovam a prevalência das terminologias instituídas ao modo de sintagmas, num percentual que se situa em torno de 70% das ocorrências terminológicas.” (KRIEGER, 2001a, p. 129).

A formação dos termos é particularmente relevante no âmbito da recuperação da informação, uma vez que, quando temos um sintagma terminológico – definido por Barros (2004, p.102) como “[...] combinatória de elementos que constituem uma unidade linguística maior [...]” –, não podemos separá-lo e/ou fragmentá-lo, sob a pena de perdermos o sentido do mesmo. Para Krieger (2001a) essa constituição é um dos maiores problemas de identificação das unidades terminológicas, visto que se faz necessário “[...] examinar o grau de coesão interna dos componentes que integram a face sintagmática das unidades complexas, bem como seus limites, diferenças e identidades com as fraseologias.” (KRIEGER, 2001a, p. 130).

Do ponto de vista formal, os termos também apresentam outras configurações sógnicas, como: siglas, acrônimos, abreviaturas e fórmulas. Ou seja, “[...] o termo não se configura apenas como unidade verbal, mas como unidade sógnica mais ampla.” (KRIEGER, 2001a, p. 130). Para descrição formal mais detalhada da tipologia terminológica geral, Krieger aponta a necessidade de considerarmos ainda outros elementos, tais como:

- a) os tipos de constituintes, observando-se sua origem (por exemplo, grego, latim);
- b) tipos de processos de formação de termos, examinando-se mecanismos como derivação, justaposição, etc;
- c) natureza dos termos do ponto de vista de sua inserção na língua, como neologismos, empréstimos. (KRIEGER, 2001a, p. 131).

Para a autora, os aspectos morfossintáticos dos termos constituem tipologias prototípicas das terminologias em geral, evidenciando que os termos não apresentam peculiaridades constitutivas opostas às do léxico comum. Portanto, o paradigma terminológico formal não difere daquele que configura o léxico geral da língua, com exceção apenas para o caráter não verbal das fórmulas (KRIEGER, 2001a). Já para as expressões e locuções gerais da língua e os sintagmas terminológicos predominam as seguintes tipologias: N+ADJ (nome mais adjetivo) e N+PREP+N (nome mais preposição mais nome). Essas formações também são próprias das locuções da língua comum, por isso Krieger (2001a, p. 131) destaca que “[...] a limitação analítica de problemas terminológicos a uma perspectiva não é suficiente para dar conta da constituição e comportamento das unidades lexicais terminológicas, sobretudo, as polilexemáticas.” Por isso, a análise do contexto textual e discursivo dos termos é imprescindível para a identificação da constituição e funcionamento dos mesmos.

Como o conteúdo dos tesouros, geralmente, se refere a uma área específica de conhecimento, compreender o funcionamento das linguagens especializadas é fundamental para pensarmos em alternativas de aperfeiçoamento dessas ferramentas. De acordo com Hoffmann (2004), a compreensão do funcionamento das linguagens especializadas só é possível através do estudo do texto especializado, tendo em vista que o mesmo “[...] é entendido como uma totalidade funcional e estrutural.” (HOFFMANN, 2004, p. 79). O autor explica que as linguagens especializadas são sublinguagens que se atualizam nos textos dos âmbitos comunicativos especializados, mas sua especificidade, em relação à linguagem comum e às outras sublinguagens, se expressa pelo léxico. Ou seja, a especificidade das linguagens especializadas se dá através da terminologia, sobretudo, “[...] pela frequência de uso de determinados recursos linguísticos.” (HOFFMANN, 2004, p. 81). Nesse sentido, o estudo do texto especializado não invalida a pesquisa sobre terminologias especializadas, pelo contrário, facilita. Assim, “É no todo do texto que se pode melhor explicar o uso linguístico especializado, a preferência por determinados recursos linguísticos, como também, interpretações funcionais e valores

comunicativos.” (HOFFMANN, 2004, p. 89). Para o autor, ao considerar o texto em sua totalidade, a linguística das linguagens especializadas mudou. “Sua postura agora privilegia linguagens especializadas em funcionamento.” (KALVERKAMPER⁴, 1983 *apud* HOFFMANN, 2004, p. 89). Para a elaboração de tesouros, a análise da linguagem especializada em funcionamento nos textos especializados possibilita, entre outras coisas, identificar as relações semânticas existentes entre os termos.

Ao estudar o funcionamento dos termos em seus reais contextos de ocorrência, a Terminologia contribui para a elaboração dos tesouros, pois além de viabilizar a descrição dos termos, também possibilita a identificação das reais condições de uso dos mesmos pelos especialistas e, conseqüentemente, potenciais usuários dos sistemas de recuperação da informação. De acordo com Krieger (2001b, p. 72), o entendimento de que o termo é componente da linguagem em funcionamento, graças a sua presença em textos e em discursos especializados, significa compreender, também, que os mesmos estão ligados aos elementos “[...] que alicerçam as dimensões semiótica, pragmática e ideológica que presidem os processos comunicacionais.” Ao privilegiarem as dimensões textual e discursiva das linguagens especializadas, as teorias da Terminologia explicam a constituição e funcionamento dos termos, motivo pelo qual, de acordo com Krieger (2001b, p. 76), alguns “[...] problemas de identificação dos termos não podem se restringir ao limite da frase, [...] mas se explicam à luz dos fenômenos da textualidade e da discursividade e, nesse sentido, dentro de um processo pragmático de comunicação.” (KRIEGER, 2001b, p. 76). Por essa razão, a seguir, avançamos no estudo da variação, trazendo, inclusive, o referencial básico no qual baseamos a análise dos dados apresentados nesta tese.

3.2 Variação

Os primeiros passos para o estudo da variação terminológica se deram a partir do surgimento da Socioterminologia, teoria formalizada por Gaudin na década de noventa do século passado, através de sua tese de doutorado “*Pour une*

⁴KALVERKAMPER, Hartwig. Textuelle Fachsprachen-Linguistik als Aufgabe in Fachsprache und Fachliteratur. (La linguistique textuelle des langues spéciales comme domaine d'étude). *LiLi. Zeitschrift für Literaturwissenschaft und Linguistik*, v. 13, n. 51-52, p. 124-166, 1983.

socioterminologie – des problèmes sémantiques aux pratiques institutionnelles”. Para o autor, o objeto de estudo da Socioterminologia é a análise da circulação dos termos em sincronia e diacronia, a qual deve ser realizada considerando a diversidade dos usos sociais dos termos, “[...] que engloba o estudo das condições de circulação e apropriação dos termos considerados como signos linguísticos, e não rótulos de conceito.” (GAUDIN, 2005, p. 81, tradução nossa). Jesus e Barros (2005), explicam que a Socioterminologia, ao fundamentar o estudo dos termos a partir da análise das condições sociais e linguísticas de circulação, tornou necessária a aceitação da variação lexical em contextos especializados.

Em vista disto, os estudos de Terminologia que surgiram a partir da Socioterminologia, além de considerar o fenômeno da variação, também aprofundaram as pesquisas sobre esse aspecto, conforme podemos observar na Teoria Comunicativa da Terminologia. Em trabalhos como o de Ciapuscio (2003), por exemplo, fica em relevo a importância do estudo dos textos especializados para a descrição da variação conceitual. Logo, a partir dos estudos de linguística textual, percebeu-se que a comunicação especializada, em alguns casos, exige uma adequação discursiva e pode apresentar variação em função dos usos e das circunstâncias comunicativas. Deste modo, a variação terminológica faz parte do processo comunicativo, tendo em vista a necessidade de adaptação do discurso dos especialistas aos diferentes níveis de abstração de seus receptores.

A partir da perspectiva poliédrica do termo, Cabré (2008) explica que a denominação de um conceito e sua apropriação pela comunidade científica não inibe o surgimento de diferentes unidades lexicais para representá-lo. De modo que

Os termos são unidades de forma e conteúdo em que o conteúdo é simultâneo à forma. Um conteúdo pode ser expresso com maior ou menor rigor por outras denominações do sistema linguístico, e constituem uma nova unidade linguística de conteúdo especializado relacionado semanticamente com a primeira ou de outros sistemas simbólicos e forma uma unidade linguística de conteúdo especializado. O conteúdo de um termo nunca é absoluto, mas sim relativo segundo cada âmbito e situação de uso (CABRÉ, 1999, p.123).

A variação, portanto, é um fenômeno comunicativo que possibilita a manifestação da diversidade linguística também no contexto especializado. Pode ser usada, inclusive, para evitar a repetição de termos no discurso, tendo em vista que, após a introdução de um conceito, é natural referir-se ao mesmo por meio de expressões alternativas (SUÁREZ; CABRÉ, 2002).

Os textos especializados distinguem-se em relação aos graus de especialização, em função da existência de diferentes níveis sociais e comunicativos. Por esse motivo, utilizam uma terminologia adequada ao contexto a que se destinam, empregando diferentes denominações para tratar um único conceito, o que gera como resultado a variação terminológica (FREIXA, 2002). Considerando o caráter linguístico da variação, Freixa (2002, p. 20, tradução nossa) destaca a importância de sua análise para “[...] compreensão e descrição do funcionamento da linguagem.” As variantes terminológicas podem ser de natureza denominativa ou conceitual.

As variantes conceituais são aquelas em que um mesmo conceito pode ser representado por diferentes denominações. As variantes denominativas, por sua vez, são identificadas através das diferentes formas pelas quais é possível representar um conceito, ou seja, por meio de diferentes denominações para um único conceito (FREIXA, 2002). Esse tipo de variação, conforme Suárez de la Torre (2004), é caracterizado pela presença de formas léxicas diferentes para indicar um mesmo conceito. Para Freixa (2001), variação denominativa é um fenômeno através do qual se denomina, de diferentes formas, a mesma unidade de significação especializada (por exemplo “décimo terceiro salário” e “gratificação natalina”). De modo que, ficam reunidos em uma mesma relação, sinônimos e variantes formais como, por exemplo, variações lexicais, ortográficas, morfossintáticas, etc.

Destacamos que o quadro 2, apresentado a seguir, elaborado por Freixa (2014), sistematiza as possibilidades de variação denominativa, que utilizamos como base analítica para a classificação das variantes identificadas em nosso *corpus* de estudo. Logo após, nos quadros de número 3 até 8, adaptamos o conteúdo original a partir da tradução e inclusão, sempre que possível, de exemplos de variantes das áreas do Direito do Trabalho e Direito Previdenciário.

Quadro 2 – Classificação formal de variantes denominativas

(Continua)

CAMBIOS GRÁFICOS		
1. Término y forma artificial	a) Término y símbolo	<i>cobre / Cu</i>
	b) Término y fórmula química	<i>amoníaco / NH₃</i>
	c) Término y otra forma artificial	<i>ley de residuos industriales de 1983 / ley 6/1983</i>
2. Término y abreviación	d) Término y sigla	<i>clorofluorocarbono / CFC</i>
	e) Término y abreviatura	<i>acero inoxidable / acer inox.</i>
3. Cambio ortográfico		<i>espray / spray</i>

Continuação

CAMBIOS MORFOSINTÁCTICOS		
1. Misma estructura	a) Ausencia / presencia de artículo	<i>gestión de residuos / gestión de los residuos</i>
	b) Cambio de número	<i>contaminación del agua / contaminación de las aguas</i>
	c) Cambio de preposición	<i>condiciones del condensador / condiciones en el condensador</i>
	d) Cambio de género	<i>máxima absoluta / máximo absoluto</i>
2. Diferente estructura	e) [N+A] / [N+SP]	<i>residuos mineros / residuos de la minería</i>
	f) Monoléxico / poliléxico	<i>producto ecológico / ecoproducto</i>
	g) Otros cambios de estructura	<i>materia primera / primera matéria</i>
REDUCCIONES		
1. Reducciones de la extensión	a) Tipo [N+X] = [N] <ul style="list-style-type: none"> ▪ [N+A] = [N] ▪ [N+SP] = [N] ▪ [N₁+N₂] = [N₁] 	<i>aguas residuales / aguas ciclo de vida / ciclo semiconductor-electrolito / semiconductor</i>
	b) Tipo [N+X+Y] = [N+X] <ul style="list-style-type: none"> • [N+SP+A] = [N+SP] • [N+A+SP] = [N+A] • [N+A₁+A₂] = [N+A₁] • [N+SP₁+SP₂] = [N+SP₁] 	<i>ahorro de recursos naturales / ahorro de recursos recogida selectiva de residuos / recogida selectiva metal pesante tóxico / metal pesante tiempo de residencia del gas / tiempo de residencia</i>
	c) Otras reducciones de la extensión	<i>filtro de control de contaminación / filtro</i>
2. Reducciones de la base	d) [N+A] = [A] _N	<i>planta depuradora / depuradora</i>
	e) [N ₁ +N ₂] = [N ₂]	<i>gas CFC / CFC</i>
	f) [N ₁ +SP (de+N ₂)] = [N ₂]	<i>espuma de poliestireno / poliestireno</i>
3. Otras reducciones		<i>emisión de compuestos químicos tóxicos / emisión tóxica</i>
CAMBIOS LÉXICOS		
1. Unidades monoléxicas		<i>contaminación / polución</i>
2. Unidades poliléxicas	a) Cambio de base <ul style="list-style-type: none"> ▪ [N+SP (prep + art + N)] ▪ [N+A] ▪ otros cambios de base 	<i>bien de consumo / producto de consumo abono químico / fertilizante químico destrucción de la capa de ozono / degradación de la capa de ozono</i>
	b) Cambio de extensión <ul style="list-style-type: none"> ▪ [N+SP (pre + art + N)] ▪ [N+A] ▪ otros cambios de extensión 	<i>depósito de asentamiento / depósito de decantación agricultura ecológica / agricultura biológica aguas residuales / aguas de la alcantarilla</i>
VARIOS CAMBIOS COMPLEJOS		
1. Con parentesco formal	a) Monoléxico / poliléxico	<i>alimento / producto alimentario</i>
	b) Poliléxico / poliléxico	<i>lodo digerido / residuo de digestión</i>
2. Sin parentesco formal	c) Monoléxico / poliléxico	<i>vertedero / depósito de residuos</i>
	d) Poliléxico / poliléxico	<i>pesticida de síntesis / plaguicida químico</i>

Fonte: Freixa (2014, p. 308).

(Conclusão)

A proposta de classificação de variantes denominativas elaborada por Freixa (2002; 2014) se divide em cinco blocos conforme o tipo de alteração linguística produzido, como se pode observar no Quadro 3. O primeiro bloco, chamado de alterações gráficas, apresenta três tipos e nove subtipos de variação. O grupo designado por termo e forma artificial é composto por unidades terminológicas provenientes da linguagem natural e outras unidades constituídas, geralmente, por recursos alheios à linguagem natural. Freixa (2002) divide esse tipo de variação em três subtipos, os quais, para melhor compreensão, são apresentados aos pares, ou seja, a unidade terminológica e sua forma artificial correspondente, conforme os exemplos do Quadro 3. O grupo designado por termo e abreviação se subdivide em termo e sigla e termo e abreviação como, também, se pode verificar no Quadro 3.

Quadro 3 – Alterações gráficas

Tipo	Subtipo	Exemplos
Termo e forma artificial	a) Termo e símbolo	Cobre / Cu;
	b) Termo e fórmula química	Amoníaco / NH ₃ ;
	c) Termo e outra forma artificial	Nova lei dos motoristas/ Lei 12.619/2012.
Termo e abreviação	d) Termo e sigla	Carteira de Trabalho e Previdência Social / CTPS;
	e) Termo e abreviação	Horas extras / Hrs extras;
Alteração ortográfica	- presença e ausência de acentuação	Férias / Ferias;
	- presença e ausência de hífen	Auxílio-alimentação / Auxílio alimentação;
	- erros de grafia como o acréscimo ou supressão de letras	Paradigma diferença salarial / Paradiguima diferença salarial;
	- presença ou ausência de espaço entre as palavras	Contracheque / Contra cheque.

Fonte: Adaptado pela autora de Freixa (2002).

Para o item alteração ortográfica, Freixa (2002; 2014) apresenta quatro possibilidades de alterações como presença e ausência de acentuação, presença e ausência de hífen, erros de grafia como o acréscimo ou supressão de letras e presença ou ausência de espaço entre as palavras conforme os exemplos que apresentamos no Quadro 3. De acordo com Freixa (2002), as alterações ortográficas apresentam, principalmente, a alternância entre formas normativas e formas não normativas dos termos, embora também ocorram casos de alternância entre duas formas igualmente normativas, como por exemplo: catorze / quatorze. O segundo bloco, denominado de alterações morfossintáticas, se divide em duas categorias que se desdobram em sete subtipos, conforme se pode observar no Quadro 4.

Quadro 4 – Alterações morfossintáticas

Tipo	Subtipo	Exemplos
Manutenção da estrutura	a) ausência /presença de artigo	Acidente de trabalho / acidente do trabalho); visto que “do” é resultado da contração da preposição “de” + artigo definido “o”;
	b) alteração do substantivo	Execução trabalhista contra herdeiro / Execução trabalhista contra herdeiros;
	c) alteração de preposição	Acidente em Serviço / Acidente de serviço;
	d) alteração de gênero	Empregada doméstica / empregado doméstico.
Alteração da estrutura	e) [Substantivo + Adjetivo] / [Substantivo + Sintagma preposicional]	Vínculo empregatício / vínculo de emprego; nesse caso o adjetivo geralmente é a parte alterada do termo;
	f) Monoléxico / Poliléxico	Produto ecológico / Ecoproduto.
	g) Outras alterações de estrutura.	Primeira matéria / matéria primeira.

Fonte: Adaptado pela autora de Freixa (2002).

É importante esclarecer que na alteração de estrutura entre unidades monoléxicas / poliléxicas, ocorre uma alteração na estrutura [substantivo + adjetivo] na qual o adjetivo da unidade sintagmática é representado por um prefixo na unidade monoléxica resultante, conforme o exemplo apresentado no item f do Quadro 4. (FREIXA, 2002).

O terceiro bloco da classificação de Freixa (2002; 2014), apresenta as variações denominativas decorrentes de reduções. As duas categorias que se desdobram em seis subtipos, foram organizadas pela autora a partir de diferentes critérios: primeiro, conforme a parte da unidade em que ocorre a redução (na extensão ou na base); segundo, de acordo com o número de itens reduzidos; e terceiro, conforme a estrutura do sintagma.

A primeira categoria, constituída por reduções de extensão se divide em três subtipos, são eles:

- a) Tipo [Substantivo + X] = [Substantivo]. Nesse caso, X pode ser um adjetivo, um sintagma preposicional ou outro substantivo. No entanto, o termo resultante da redução será um substantivo, seja qual for a função de X;

- b) Tipo [Substantivo + X + Y] = [Substantivo + X]. Nesse caso, X e Y podem ser adjetivo ou sintagma preposicional, e o termo resultante da redução será formado por um substantivo e um adjetivo ou por um substantivo e um sintagma preposicional;
- c) Outras reduções de extensão. Nesse caso, há uma grande variedade de estruturas, com reduções de mais de um elemento, conforme os exemplos apresentados no Quadro 5.

Quadro 5 – Reduções de extensão

Subtipo	Exemplos
a) Tipo [Substantivo + X] = [Substantivo].	
- [Substantivo + Adjetivo]	Aviso prévio / Aviso;
- [Substantivo + Sintagma preposicional]	Acidente de trabalho / Acidente;
-[Substantivo 1+ Substantivo 2] = Substantivo	Gratificação Assiduidade / Gratificação.
b) [Substantivo + X + Y] = [Substantivo + X].	
- [Substantivo + Sintagma preposicional + Adjetivo] = [Substantivo + Sintagma preposicional]	Equipamento de proteção individual / Equipamento de proteção;
- [Substantivo + Adjetivo + Sintagma preposicional] = [Substantivo + Adjetivo]	Certidão negativa de débitos / Certidão negativa
- [Substantivo + Adjetivo 1 + Adjetivo 2] = [Substantivo + Adjetivo 1]	Certidão negativa trabalhista / Certidão negativa;
- [Substantivo + Sintagma preposicional 1 + Sintagma preposicional 2] = [Substantivo + Sintagma preposicional 1]	Adicional de periculosidade de electricista / Adicional de periculosidade;
c) Outras reduções de extensão	Tabela de lotação de cargos / Tabela.

Fonte: Adaptado pela autora de Freixa (2002).

A segunda categoria, constituída por reduções da base se divide conforme a estrutura morfossintática do sintagma, como nos exemplos apresentados no Quadro 6.

Quadro 6 – Reduções de base

Subtipo	Exemplos
d) [Substantivo + Adjetivo] = [Adjetivo] N	Trabalhador temporário / Temporário
e) [Substantivo 1 + Substantivo 2] = [Substantivo 2]	Adicional insalubridade / Insalubridade
f) [Substantivo 1 + Sintagma preposicional (de + Substantivo 2)] = [Substantivo 2]	Pagamento de diárias / Diárias.
g) Outras reduções	Trabalho doméstico infantil / Trabalho infantil.

Fonte: Adaptado pela autora de Freixa (2002).

Na categoria outras reduções, Freixa (2002) explica que ocorrem diversos tipos de reduções de partes internas do sintagma, especialmente do adjetivo que qualifica a base, conforme se pode verificar no exemplo apresentado no item g do quadro 6.

O quarto bloco da Classificação de Freixa (2002; 2014) é o de alterações léxicas. Para a autora, esse tipo de variação é o que se encontra melhor descrito na literatura, visto que trata dos casos de sinonímia propriamente ditos. Freixa (2002; 2014) apresenta duas categorias de alterações léxicas, a primeira chamada de unidades monoléticas, em que classifica os casos que apresentam alteração lexical entre duas formas monoléticas, como por exemplo: *Diarista / Eventual*; a segunda categoria, chamada de unidades poliléticas, classifica os casos que apresentam alteração lexical entre duas formas poliléticas subdividindo-se de acordo com a parte do sintagma em que a mudança ocorre (base ou extensão) e conforme o tipo de estrutura morfossintática do sintagma. Apresentamos no Quadro 7 alguns exemplos de unidades poliléticas que apresentam alteração de base (a) e alteração de extensão (b).

Quadro 7 – Alterações Léxicas de unidades poliléticas

Subtipo	Exemplos
a) Alteração de base	
- [Substantivo + Sintagma Preposicional (prep. + art. + Substantivo)]	Jornada de trabalho / Horário de trabalho;
- [Substantivo + Adjetivo]	Empregado doméstico / Trabalhador doméstico;
- [Outras alterações de base]	Competência da Justiça do Trabalho / Responsabilidade da Justiça do Trabalho
b) Alteração de extensão	
- [Substantivo + Sintagma Preposicional (prep. + art. + Substantivo)]	Acidente de trabalho/ Acidente de serviço;
- [Substantivo + Adjetivo]	Abono salarial/ abono pecuniário;
- [Outras alterações de extensão]	Consolidação das leis do trabalho / Consolidação das leis trabalhistas;

Fonte: Adaptado pela autora de Freixa (2002).

O quinto bloco da Classificação de Freixa (2002; 2014) é constituído por duas categorias que se subdividem em quatro subtipos. Nessa categoria, encontramos termos que divergem entre si por várias alterações complexas, organizadas pela autora a partir da verificação da manutenção ou não do parentesco formal entre as variantes analisadas e a alternância entre uma unidade monolética e uma unidade polilética ou entre duas unidades poliléticas. No Quadro 8, a seguir, apresentamos alguns exemplos dessas variantes.

Quadro 8 – Várias alterações complexas

Tipo	Subtipo	Exemplos
Com parentesco formal	a) Monoléxica/Poliléxica	Agroecologia / Agricultura ecológica;
	b) Poliléxica/Poliléxica	Doenças ocupacionais / Doença profissional;
Sem parentesco formal	c) Monoléxica/Poliléxica	Sobreaviso / Tempo à disposição;
	d) Poliléxica/Poliléxica	Bombeiro civil / Brigadista particular

Fonte: Adaptado pela autora de Freixa (2002).

A classificação de variantes denominativas de Freixa (2002; 2014) considera aspectos muito específicos de variação e esses, em uma leitura/análise pouco atenta, podem nos deixar em dúvida quanto à possibilidade de classificação de algumas variantes em diferentes categorias. Por esse motivo, ao compararmos duas variantes entre si, é preciso observar com atenção as alterações produzidas, visto que a presença ou ausência de um único aspecto é fator determinante para correta classificação. Por exemplo, qual a diferença entre uma alteração de estrutura morfossintática que ocasiona a mudança de uma unidade monoléxica para uma unidade poliléxica, e as reduções de extensão, visto que na categoria reduções também encontramos unidades poliléxicas que, devido à redução, se tornam monoléxicas? A diferença está no resultado da alteração obtida. No caso das alterações morfossintáticas, um dos termos da unidade poliléxica, geralmente o adjetivo, de acordo com Freixa (2002), se transforma em prefixo da forma monoléxica, sem que o termo perca sua especificidade, como se pode verificar no exemplo a seguir: *produto ecológico / ecoproduto*. No caso das várias alterações complexas, é o radical do substantivo que se torna prefixo da unidade monoléxica resultante, como se pode verificar nesse exemplo: **Agricultura ecológica / Agroecologia**. Quanto às reduções, na maioria dos casos, observamos que ocorre a perda de sentido do termo poliléxico no termo monoléxico resultante, sobretudo nos casos de redução da extensão, que se torna mais genérico, como se pode observar no exemplo a seguir: *Acidente de trabalho / Acidente*.

É importante destacar que, embora os usuários utilizem variantes reduzidas em suas buscas, a inserção das mesmas como termos equivalentes em um tesauro deve

ser analisada com cuidado, visto que para fins de recuperação da informação é imprescindível que as variantes utilizadas como equivalentes na rede de remissivas preservem o sentido do termo original. No âmbito da representação e recuperação da informação, quando um termo perde sua especificidade em função de uma redução e o termo resultante apresenta uma mudança de sentido, passamos a ter dois termos diferentes e não dois equivalentes. É preciso considerar que, em uma base de dados, ao realizarmos uma busca, não temos o contexto discursivo oferecido por um texto completo que possibilita ao leitor perceber que a variante reduzida utilizada pelo autor está sendo empregada com o mesmo sentido do termo original completo. O termo *Acidente* é muito genérico, pois no âmbito do Direito do trabalho existem diferentes tipos de acidente, tais como “acidente de percurso” e “acidente de trabalho”.

Dúvidas, quanto à possibilidade de classificação de algumas variantes em diferentes categorias, também ocorrem em relação à categoria várias alterações complexas e a categoria reduções, quando analisamos a transformação tanto de uma unidade poliléxica em monoléxica como a alteração resultante entre duas unidades poliléxicas. E, da mesma forma que o exemplo anterior, o resultado da alteração indicará a categoria da variante. Nesse caso, observamos que a unidade monoléxica ou poliléxica resultante de uma redução mantém inalterados um dos termos da unidade poliléxica original. Já no caso da categoria de várias alterações complexas, percebemos que tanto a unidade monoléxica como a unidade poliléxica resultante apresentarão alguma modificação na forma do termo resultante embora a unidade poliléxica resultante, conforme Freixa (2002), também possa preservar um, dois ou mais radicais da unidade poliléxica original, como se pode observar nos exemplos em destaque no Quadro 9.

Quadro 9 – Comparação entre reduções e várias alterações complexas

Reduções	Várias alterações complexas
a) Redução de extensão (Polilexa / Monoléxa): Adicional Insalubridade / Adicional	a) Poliléxa / Monoléxica com parentesco formal: Agricultura ecológica / Agroecologia
b) Redução de base (Polilexa / Monoléxa): Adicional insalubridade/ Insalubridade	b) Poliléxica / Poliléxica com parentesco formal Doenças ocupacionais / Doença profissional; Portador de deficiência / Deficiente físico; Repouso compensado semanal / Repouso semanal remunerado
c) Redução de extensão (Polilexa / Polilexa): Equipamento de proteção individual / Equipamento de proteção;	
d) Outras reduções (Polilexa / Poliléxa): Trabalho doméstico infantil / Trabalho infantil.	

Fonte: elaborado pela autora.

Cabe destacar ainda, a diferença entre as unidades monoléxicas resultantes de alterações morfossintáticas e as unidades monoléxicas resultantes de várias alterações complexas. Nesse caso, nas duas categorias, teremos a transformação de uma unidade poliléxica em monoléxica. A diferença é que o termo monoléxico resultante da categoria de alterações morfossintáticas, conforme mencionamos anteriormente, apresentará um prefixo. Já o termo monoléxico resultante de várias alterações complexas pode ou não apresentar um prefixo; e nos casos em que a forma monoléxica apresenta um prefixo, esse não será necessariamente um adjetivo do termo que já estava presente na forma original poliléxica, como ocorre nas alterações morfossintáticas. Observe os exemplos do Quadro 10.

Quadro 10 – Comparação entre alterações morfossintáticas e várias alterações complexas

Alterações de estrutura morfossintática	Várias alterações complexas
Produto ecológico / Eco produto.	Agricultura ecológica / Agroecologia

Fonte: Elaborado pela autora a partir dos exemplos de Freixa (2002).

Com o objetivo de explicar como um conceito pode ser representado no plano linguístico por mais de um termo, Cabré (2011) propõem um estudo para analisar as relações que podem ser estabelecidas entre unidades de categoria conceitual e unidades terminológicas. Nesse trabalho, a autora destaca a existência de variantes denominativas com consequências cognitivas, as quais “[...] podem ser expressas através da forma denominativa em si ou por pistas discursivas que atuam como elementos que evidenciam uma variação conceitual. Estas evidências podem ser de vários tipos e podem ser expressas por diferentes formas gramaticais.” (CABRÉ, 2011, p. 14).

Em vista disto, Cabré (2011) destaca dois casos em que um objeto pode ser representado verbalmente por termos distintos. No primeiro caso, a autora apresenta exemplos que não apresentam consequências cognitivas, ou seja, aqueles em que o objeto corresponde a somente um conceito, expresso por diferentes signos semanticamente coincidentes, ainda que formalmente diferentes. Incluem-se, nesse caso, as variantes gráficas, morfológicas e léxicas, uma vez que o uso de uma ou outra variante não altera a maneira como o conceito se projeta na denominação (CABRÉ, 2011), por exemplo: Exoneração / Rescisão.

Já no segundo caso, são apresentados exemplos que indicam consequências cognitivas, ou seja, o objeto em questão corresponde a somente um conceito; porém esse é expresso por signos semântica e formalmente diferentes. Deste modo, “[...] o uso de uma variante ou outra para designar um mesmo objeto pode estar condicionado por uma intenção cognitiva da parte do locutor e, evidentemente, ter consequências cognitivas no receptor.” (CABRÉ, 2011, p. 13). Um exemplo desse caso se configura em: resíduos biosanitários e resíduos hospitalares. Cabré (2011) explica que o objeto designado por esses dois termos citados como exemplo em relação de variação é o mesmo. No entanto, cada uma das variantes remete a um sentido diferente: no primeiro caso o termo destaca as características da atividade que gera esse tipo de resíduo (atividades bio-sanitárias); já no segundo caso o termo destaca o lugar em que são produzidos os resíduos (hospitais). O que se observa, de acordo com Cabré (2011, p. 13), é que ambos os termos são “[...] expressões sintagmáticas formadas a partir de um núcleo comum (resíduos) cujo significado coincide parcialmente.” Para a autora, esses termos são sinônimos, mas não em sua totalidade, visto que, através da denominação de cada um, é possível perceber uma parte, ou uma faceta diferente do mesmo conceito que representam. Em outras

palavras, os dois termos designam o mesmo objeto, porém as diferentes formas de denominação colocam em evidência uma parte do conceito, de modo que a presença de uma forma ou de outra, em um texto, por exemplo, pode indicar o ponto de vista adotado pelo autor, além de pontos de vista que destaquem aspectos sociais, culturais, econômicos, entre outros.

Cabré (2011, p. 9) destaca que

[...] a mesma unidade de conhecimento pode aparecer nos textos de especialistas não só denominada de forma diferente (sinonímia) mas também explicada ou exposta discursivamente de maneira diferente, em forma de bloco, ou introduzindo progressivamente aspectos diferentes da unidade conceitual.

A autora explica que advérbios ou locuções adverbiais podem ser utilizados para explicar as diferentes dimensões do mesmo conceito, como, por exemplo: “A doença X, a partir do ponto de vista da patologia, está descrita como...; do ponto de vista sintomático, é descrita como...” (CABRÉ, 2011, p. 14). Ou seja, em um texto, é possível identificar traços semânticos que caracterizam um termo.

No contexto de aplicação de nossa pesquisa, a compreensão das características da variação denominativa é fundamental, uma vez que a grafia das unidades lexicais pode interferir no processo de recuperação da informação, “[...] já que os processos de normalização não atuam sobre a pronúncia dos termos, mas precisamente sobre sua forma escrita.” (CABRÉ, 1993, p. 172, tradução nossa). Logo, optamos por utilizar a Classificação de Variantes Denominativas de Freixa (2002; 2014) por dois motivos: primeiro, pelo fato de considerarmos a classificação da autora mais completa do que a classificação de Faulstich (1998); segundo, por julgarmos a classificação de Freixa (2002; 2014) mais adequada à nossa proposta, pois, assim como a autora, desenvolvemos nossa pesquisa a partir dos pressupostos teóricos da TCT.

Freixa (2014) explica que a análise da variação coloca em evidência os dois lados do signo linguístico, ou seja, a forma e o conteúdo. Para a autora, geralmente as classificações de sinônimos e variantes voltam-se para um desses dois sentidos. Quando se concentra, na análise do conteúdo, o significado, é possível falar em graus de sinonímia ou de equivalência; assim, geralmente, a classificação de sinônimos e variantes pode ser interpretada a partir de sua forma. Freixa (2014) menciona que as variações lexicais são exemplos de sinônimos e as variações ortográficas são exemplos de variantes. Porém, argumenta que as variações morfológicas e sintáticas

ou aquelas que combinam várias mudanças não se encaixam nessa dicotomia e, do ponto de vista teórico variacionista, são todos exemplos de variação. Uma mudança na denominação, então, tanto no nível léxico como no nível ortográfico, pode gerar uma nova denominação para o mesmo conceito (Freixa, 2014). A distinção entre sinônimos e variantes é muito sutil e intuitiva, já que os sinônimos são percebidos de maneira direta a partir do uso de diferentes termos e as variantes apresentam, apenas, formas distintas de representar um mesmo termo e, por isso, estão relacionadas com o grau de fixação terminológica (Freixa, 2014). Por entender que a fronteira entre variantes e sinônimos não é nítida, a autora defende a utilização da forma geral variação denominativa, incluindo variantes e sinônimos. Por considerar que tanto as variantes como os sinônimos interferem no processo de recuperação da informação, nesta pesquisa, assim como Freixa (2014), utilizaremos a forma geral variação denominativa, incluindo variantes e sinônimos, sem distinção.

Em sua pesquisa sobre variação, Faulstich (1998, p.2) afirma que “A polifuncionalidade da unidade lexical, no discurso científico, no discurso técnico ou no discurso de vulgarização científica pode produzir mais de um registro ou mais de um conceito para o mesmo termo.” Em decorrência disto, a autora destaca as variantes mais frequentes, em *corpora* de línguas de especialidade, conforme pode-se observar no Quadro 11 a seguir.

Quadro 11 – Variantes de acordo com a Socioterminologia

	Tipos	Descrição	Exemplo
Variantes Linguísticas	Fonológica	O registro pode surgir de formas decalcadas da fala.	<i>Agoa de frol, em relação à água de flor, e queygo em relação a queijo.</i>
	Morfológica	Apresenta alternância de estrutura de ordem morfológica na constituição do termo, sem que o conceito se altere.	<i>Porrada e porreta. No caso nos sufixos -ada e -eta, ambos indicando comida em que entram alhos porros.</i>
	Sintática	Apresenta alternância entre duas construções sintagmáticas que funcionam como predicação de uma UTC. Neste caso, a variação se processa na substituição de uma parte do item lexical por outro com estrutura semelhante, formando uma mesma unidade terminológica.	<i>Agua de frol cõ almísquer e agua de cheiro almisclada, na linguagem da culinária arcaica.</i>
	Lexical	Algum item da estrutura lexical da UTC sofre apagamento ou movimento de posição, mas o conceito do termo não se altera.	<i>Ovos com crara e gemas batidas, ovos batidos cõ crara e jema, gemas dos ovos batidos com a crara, ovos gemas e craras tudo batido.</i>
	Gráfica	Apresenta-se sob forma gráfica diversificada de acordo com as convenções da língua. <i>Este tipo de variação decorre da forma escrita do termo.</i>	<i>Manteiga, mãtejua e mamteyguaxviii e receita, receyta, rrecejta.</i>
Variantes de Registro	Geográfica	Ocorre no plano horizontal de diferentes regiões em que se fala a mesma língua.	<i>Sertãxix, da área de utensílios de culinária. O termo sertã é praticamente ignorado no Brasil, que utiliza frigideira, em lugar daquele.</i>
	Discurso	Decorre da sintonia comunicativa que se estabelece entre elaborador e usuários de textos científicos e técnicos.	<i>Pouquechinho, termo do discurso de vulgarização científica, usado, no séc. XV, nas receitas de cozinha para indicar quantidade. Hoje usamos pitada [de sal, de açúcar].</i>
	Temporal	Aquela que se configura como preferida no processo de variação e mudança, em que duas formas (X e Y) concorrem durante um tempo, até que uma forma se fixe como a preferida.	<i>Adubado (X) substituído por temperado (Y), assim como pão coito (X) que cedeu lugar para biscoito (Y).</i>

Fonte: Adaptado pela autora de Faulstich (1998).

A comparação entre os estudos de Freixa (2002; 2014) e o de Faulstich (1998), nos permite afirmar que ambas apresentam uma classificação das tipologias de variantes. Porém, os pressupostos teóricos utilizados pelas autoras, para o embasamento de suas classificações, são diferentes. Freixa (2002; 2014) realiza suas pesquisas no âmbito da TCT, enquanto Faulstich (1998) utiliza o aporte da Socioterminologia.

3.3 As relações entre termo e descritor

Levando em conta que o foco de nossa proposta está voltado para a seleção e a identificação de termos equivalentes para tesouros, destacamos aqui algumas características de termo e de descritor. Visamos, com isso, tornar mais claro o nosso entendimento sobre a relação entre termo e descritor que, juntos, constituem-se como elementos fundamentais para a viabilidade e sucesso da metodologia que estamos propondo. Lembramos que nosso objetivo é tornar os tesouros, que são sistemas conceituais, em sistemas mais linguísticos e próximos da linguagem dos usuários e que, para alcançar esse objetivo, realizamos a identificação e utilização das variantes terminológicas existentes nas lexias de buscas de usuários de um SRI; processo esse exemplificado, nesta tese, com a aplicação das lexias de buscas das áreas do Direito do Trabalho e do Direito Previdenciários dos usuários do Portal LexMI, administrado pelo Senado Federal do Brasil.

A partir da perspectiva teórica advinda dos estudos de Cabré (1999), de Cabré e Codina (2001) e de Van der Laan (2002), entendemos que tanto o termo como o descritor são, ao mesmo tempo, termo e/ou descritor, propriamente ditos. Para a TCT, *a priori* não existem termos nem palavras, somente unidades lexicais, já que as mesmas adquirem estatuto terminológico no âmbito das comunicações especializadas (KRIEGER; FINATTO, 2004). Kostina (2000, p. 20, tradução nossa) afirma que “[...] os termos ou unidades terminológicas em primeiro lugar são unidades linguísticas que possuem todas as particularidades da linguagem natural, inclusive participando da polissemia, da sinonímia e da homonímia [...]”. Engel e Biscarra Neto (1992) acrescentam que o surgimento de novos conhecimentos implica novas denominações, uma vez que cada novo conceito precisa ser designado por algum tipo

de código. Assim, os termos possuem duas funções: a de representar o conhecimento e a de possibilitar a sua transferência. De forma resumida, pode-se dizer, então, que um termo representa conceitos de uma área de especialização, transmite conhecimento especializado e possui caráter linguístico em função de sua inserção no discurso. No âmbito da recuperação da informação, a formação dos termos é particularmente relevante, uma vez que, quando temos um sintagma terminológico – definido por Barros (2004, p.102) como “[...] combinatória de elementos que constituem uma unidade linguística maior [...]” –, não podemos separá-lo e/ou fragmentá-lo, sob a pena de perdermos o sentido do mesmo.

Quanto aos descritores, no âmbito da Ciência da Informação, pode-se dizer que são unidades capazes de identificar, para fins de indexação, determinado conceito ou o conteúdo de um item ou obra. Dodebei (2002, p. 39) define descritores como “[...] palavras que simbolizam os conceitos organizados em um léxico documentário tais como as linguagens documentárias hierárquicas e /ou facetadas.” Para Van Slype (1991), os descritores representam de maneira unívoca conceitos utilizados para indexação ou consulta de documentos, formados, geralmente, por palavras isoladas (principalmente substantivos), termos compostos ou frases. Os não-descritores, por sua vez, são os sinônimos ou quase-sinônimos dos descritores (em relação de equivalência) e não podem ser utilizados na indexação, podem apenas remeter para os descritores.

Van der Laan (2002), a partir de uma extensa revisão bibliográfica sobre o entendimento, emprego e função dos descritores no âmbito da Biblioteconomia e da Ciência da Informação, afirma que o problema de um descritor, como unidade de representação e/ou recuperação da informação, ocorre pelo fato de o mesmo ser tratado como uma unidade lexical pertencente ao léxico comum; porém, passível de restrições normativas e com o objetivo de controlar a linguagem a ser utilizada para a representação e recuperação das informações contidas em um SRI. Em função desse caráter normativo dos descritores para o controle da linguagem, a autora evidencia certa aproximação dos mesmos com a TGT, para a qual o termo restringe-se à parte significante da unidade lexical.

O termo faz parte do léxico dos especialistas; nesse sentido, todas as formas variantes utilizadas para representar um conceito de um domínio são termos. Por exemplo: “CTPS”, “Carteira de Trabalho e Previdência Social” e “Carteira de Trabalho”. Nesse exemplo, temos três formas diferentes de representar o mesmo

conceito, e todos são termos da área do Direito do Trabalho. Mas, nesse caso, em um tesouro, apenas um dos termos terá o valor de descritor. Por esse motivo, o descritor também é chamado de termo preferido no âmbito da organização do conhecimento, uma vez que ele é escolhido entre outras possibilidades de representação. E, deste modo, os demais termos (variantes) são inseridos nos tesouros como não-descritores, os quais não podem ser utilizados para a representação dos assuntos dos documentos nos registros bibliográficos. Embora não possam ser utilizados nos registros bibliográficos, os não-descritores desempenham o importante papel de indicar ao usuário que o conceito que ele está buscando se encontra representado por outro termo, o descritor. Isso ocorre, pois os tesouros são metalinguagens cuja função é controlar os termos que serão utilizados para a representação do assunto de documentos.

Para Cabré (2001a, p. 12, tradução nossa), os descritores não são termos, pois, mesmo que possam coincidir, “[...] não são unidades da língua, são apenas etiquetas funcionais que servem para descrever a informação [...]”. A autora defende que as similaridades de termos e descritores são apenas aparentes, pois terminólogos e documentalistas possuem uma percepção diferente da terminologia.

Para os terminólogos, os termos são unidades específicas usadas por especialistas em um assunto, ao descrever esse campo. Para o documentalista, os termos são unidades que potencialmente podem se tornar descritores do conteúdo de um documento especializado e elemento de controle em uma classificação. (CABRÉ, 2001a, p. 12, tradução nossa).

Van der Laan (2002) propõe uma aproximação entre termo e descritor ao considerar os últimos como signos representativos de nódulos conceituais. Para a autora, é preciso lançar

[...] um novo olhar para o descritor ampliando sua concepção, não o tratando apenas como uma expressão de representação/recuperação de informações, mas em uma dimensão maior, pois, da mesma forma que uma unidade lexical terminológica, ele é constituído por três dimensões: a cognitiva; a lingüística e a comunicativa. (VAN DER LAAN, 2002, p. 69)

Com essa aproximação entre termo e descritor, Van der Laan (2002) explica que, ao compreendermos que os descritores apresentam uma dimensão lingüística, destacamos, também, a importância dos não-descritores, que passam a ter função em um SRI, visto que eles também possuem uma função representativa para os diferentes

usuários, através de remissivas. Desta forma, as unidades de indexação, utilizadas pelos bibliotecários, poderiam adquirir maior representatividade e, com isso, as informações poderiam ser representadas/recuperadas de forma mais próxima ao vocabulário utilizado pelos usuários.

Concordamos com a proposta de aproximação entre termos e descritores, de Van der Laan (2002), que defende que o ideal é que os descritores sejam termos. Nesse sentido, ao invés de os descritores apenas potencialmente coincidirem com os termos, como destaca Cabré (2001a), os descritores podem passar a serem termos também; assim, essa noção modifica nossa percepção em relação aos descritores. No entanto, para que essa proposta se efetive, é preciso inserir na metodologia de elaboração de tesauros essa perspectiva e a obrigatoriedade de escolhermos termos para serem descritores.

4 TESAUROS: o que são e para que servem?

Currás (2010) menciona que algumas pessoas acreditam que a época do esplendor dos tesouros está chegando ao fim devido aos avanços tecnológicos dos sistemas informatizados. Para a autora, todos os avanços de fato devem vir a melhorar e facilitar cada vez mais a recuperação da informação pelos usuários, no entanto, ainda assim, os profissionais que desenvolvem e planejam sistemas de recuperação da informação informatizados precisam saber muito bem como estabelecer as relações entre os termos e como coordenar conceitos. Afinal, “Não são as máquinas que fazem o esforço intelectual. São os seres humanos que o fazem.” (CURRÁS, 2010, p. 90).

No contexto em que nos encontramos, com diversas pesquisas sobre web semântica e ontologias, é compreensível que surjam questionamentos em relação à relevância ou não de continuarmos investindo esforços para a elaboração de tesouros. Esses questionamentos podem ocorrer, principalmente, pelo fato de existirem inúmeras semelhanças conceituais e metodológicas entre tesouros e ontologias. Porém, é necessário ressaltar que as diferenças, entre eles, também existem e são marcantes. Acreditamos que a comparação entre ambos possibilita uma visualização bastante didática de seus pontos de convergência e divergência, e é por essa razão que apresentamos a seguir os contrastes que envolvem os tesouros e as ontologias. Esperamos que, ao explicitar as semelhanças e diferenças dessas ferramentas, fique, também, em evidência a relevância das duas.

De acordo com Pickler (2007) a Ciência da Informação se propõe a organizar o conhecimento e, para isso, se baseia na classificação de conceitos proposta por Aristóteles. Deste modo, diferentes categorias foram criadas para que todo o tipo de conhecimento existente pudesse ser enquadrado em alguma dessas classes ou categorias. Em seguida, para uma efetiva organização do conhecimento, foi necessário criar “[...] uma linguagem documentária que buscasse uma padronização da linguagem natural eliminando a polissemia característica da mesma.” (PICKLER, 2007, p. 74). Devido à crescente evolução científica e tecnológica e ao conseqüente aumento do volume de documentos gerados, surgiu, no âmbito da Ciência da Informação, a necessidade de controlar a terminologia das áreas especializadas. Deste modo, para facilitar a organização, representação e recuperação de

documentos desenvolveram-se os tesouros, com o objetivo de formalizar e padronizar termos de uma área específica do conhecimento.

A evolução histórica do tesouro de recuperação pode ser traçada a partir de duas vertentes: uma, que toma nitidamente como base o sistema unitermo, idealizado por Mortimer Taube na década de 1950, que defendia a utilização de termos simples (constituídos por uma única palavra), e outra influenciada pela teoria da classificação facetada. De um lado, a influência da América do Norte, caracterizada pela abordagem alfabética, e, de outro, a abordagem sistemática, de influência inglesa, principalmente (CAMPOS; GOMES, 2006).

Durante a década de 60, os tesouros foram sistematicamente aperfeiçoados. De acordo com as diretrizes para a construção de tesouros descritas no padrão norte-americano ANSI/NISO Z39.19:

Um tesouro é um vocabulário controlado organizado em uma ordem preestabelecida e estruturado de modo que os relacionamentos de equivalência, de homografia, de hierarquia, e de associação entre termos sejam indicados claramente e identificados por indicadores de relacionamento padronizados empregados reciprocamente. As finalidades primordiais de um tesouro são (a) facilitar a recuperação dos documentos e (b) alcançar a consistência na indexação dos documentos escritos ou registrados de outra forma e outros tipos, principalmente para sistemas de armazenamento e de recuperação de informação pós-coordenados (SALES; CAFÉ, 2009, p. 3).

Ao observarmos algumas definições de ontologias e tesouros, logo percebemos uma proximidade entre os dois recursos, pois ambos, segundo Sales e Café (2008), são modelos de representação do conhecimento utilizados para representar e recuperar informação através de estruturas conceituais. Ambos têm o objetivo comum de facilitar a representação e recuperação da informação tornando-as mais específicas. A primeira diferença a surgir nessa comparação é que o meio de ação das ontologias é o informático. (SALES; CAFÉ, 2008). Isso quer dizer que as ontologias foram criadas para funcionar especificamente no âmbito digital (internet); foram pensadas e desenvolvidas para serem compreendidas e lidas por máquinas (computadores). Já os tesouros, que também trabalham com o relacionamento existente entre os termos de uma área de conhecimento especializado, não apresentam essa característica, ou seja, não foram idealizados para serem compreendidos/lidos pelos computadores, apenas pelas pessoas.

Gruber (1993, p. 199) define ontologia como “[...] uma especificação formal explícita de uma conceitualização.” Embora essa seja uma das definições mais

clássicas de ontologia, é necessário aprofundá-la um pouco mais para que a mesma possa ser compreendida. Por isso, alguns autores complementam essa definição de Gruber: Prévot, *et al.*(2009) explicam que a conceitualização apontada por Gruber é um conhecimento informal relevante que pode ser extraído e generalizado a partir da experiência, observação ou introspecção e que a especificação é a codificação deste conhecimento em uma representação da linguagem. Sales e Café (2008)⁵ são ainda mais específicos e esclarecem que:

[...]‘*conceitualização*’ se refere a um modelo abstrato dos fenômenos no mundo, identificando os conceitos relevantes daqueles fenômenos. O termo ‘formal’ se refere ao fato de que a ontologia deve ser legível por máquina; ‘explícito’ se justifica porque os tipos de conceitos usados e suas restrições de uso são definidos explicitamente; e ‘compartilhado’ reflete que a ontologia deve capturar o conhecimento consensual aceito pelas comunidades.

Outra definição importante de ontologia é apresentada por Guarino (1998, p. 2, tradução nossa) na qual o autor diz que:

[..] uma ontologia refere-se a um *artefato de engenharia*, constituído por um *vocabulário* específico usado para descrever uma certa realidade, mais um grupo de suposições explícitas considerando o *significado pretendido* das palavras do vocabulário. Esse grupo de suposições tem geralmente a forma de uma teoria lógica de primeira ordem, em que as palavras do vocabulário aparecem como nomes de predicados unários ou binários, respectivamente chamados de conceitos e relações. Nos casos mais simples, uma ontologia descreve a hierarquia de conceitos relacionados por relações de suposição; nos casos mais sofisticados, axiomas convenientes são adicionados a fim de expressar outras relações entre conceitos e para restringir suas interpretações pretendidas.

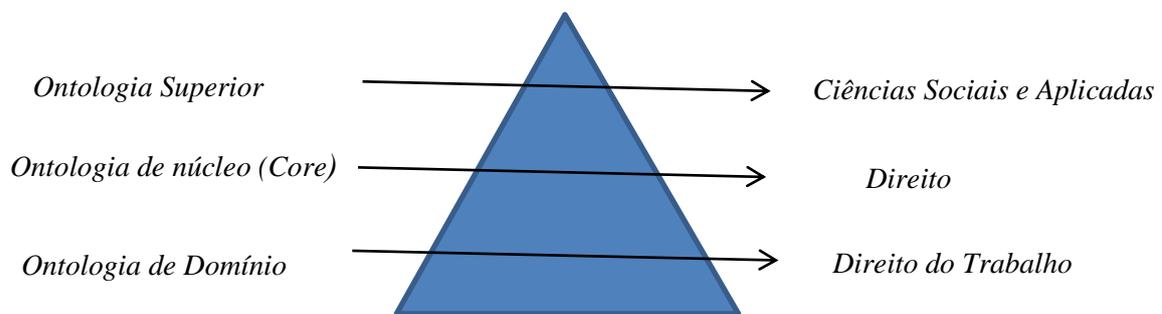
Quanto mais avançam as pesquisas sobre ontologias, mais definições e explicações sobre elas se encontraram na literatura; por isso, selecionamos aquelas que julgamos suficientes para a análise que nos propomos nesse trabalho. A partir delas, é possível afirmar que as inúmeras definições encontradas na literatura sobre ontologias, “[...] convergem em afirmar que a ontologia proporciona um vocabulário

⁵ Documento eletrônico.

formal e comum baseado em uma estrutura de conceitos específicos de um dado domínio.” (SALES; CAFÉ, 2008)⁶.

Prevót et al (2009) apontam a existência de três diferentes níveis de ontologias reconhecidas em pesquisas e práticas sobre o assunto, são elas: de nível superior (*upper*), de núcleo ou referência (*core*), e ontologias de domínio. Para os autores, as ontologias de nível superior dizem respeito a categorias mais gerais que constituem a parte superior da taxonomia do conhecimento, ou seja incluem a instância, a taxonomia simples. A distância existente entre uma ontologia superior e uma ontologia de domínio faz com que exista, necessariamente, uma ontologia intermediária entre elas, que é chamada de ontologia de núcleo (*core*), na qual encontramos conceitos intermediários que fazem a ligação entre a ontologia superior e a de domínio. As ontologias de domínio, por sua vez, são aquelas que abrangem o conhecimento especializado; nelas encontraremos conceitos mais específicos de uma determinada área do conhecimento, de acordo com os objetivos da ontologia elaborada. Didaticamente falando, se escolhêssemos fazer uma ontologia superior, uma *core* e uma de domínio, específicas de uma área do conhecimento, poderíamos ilustrar a relação entre os três tipos de ontologia da seguinte forma:

Figura 1- Ontologias



Fonte: elaborado pela autora.

Em relação aos tesauros, também são inúmeras as definições e explicações encontradas na literatura; por isso, selecionamos aquelas que julgamos mais significativas no âmbito da Ciência da Informação, as quais, analisadas em conjunto, oferecem maior compreensão e clareza sobre os tesauros. Aitchison, Gilchrist e

⁶ Documento eletrônico.

Bawden (2000, p. 1, tradução nossa) definem tesouros como “[...] vocabulário de uma linguagem de indexação controlada, formalmente organizado de modo que, a priori, o relacionamento entre conceitos se torne explícito, para serem utilizados em sistemas de recuperação de informação [...]”. Para Currás (2010, p. 99), trata-se de “[...] uma linguagem especializada, normalizada, pós-coordenada, usada com fins documentários, onde os elementos linguísticos que a compõem – termos simples ou compostos – se relacionam entre si sintática [e] semanticamente.” Já para Sales e Café (2008)⁷, são “[...] vocabulários controlados formados por termos (*descritores*) semanticamente relacionados, que atuam como instrumentos de controle terminológico.” De acordo com a norma ISO 25964-1, tesouro é um

Vocabulário estruturado e controlado em que os conceitos são representados por termos, organizado de modo que as relações entre os conceitos são explícitas, e os termos preferenciais são acompanhados por entradas para sinônimos ou quase-sinônimos. [...] O objetivo de um tesouro é orientar tanto o indexador como o pesquisador para selecionar o mesmo termo preferido ou a combinação dos termos preferidos para representar um determinado assunto. Por essa razão, um tesouro é otimizado para a navegabilidade humana e cobertura terminológica de um domínio. (INTERNATIONAL..., 2011, p. 12, tradução nossa).

A partir das definições de tesouros apresentadas, percebe-se que as principais características dos tesouros são: o controle terminológico; a relação explícita entre conceitos; o fato de serem especializados, estruturados e constituídos por termos; destinados à leitura humana com utilização tanto por indexadores, como por pesquisadores durante o ato de representação de assunto.

Conforme identificamos anteriormente, os tesouros foram idealizados para serem utilizados por bibliotecários no momento da indexação, e também por usuários no momento da busca de informação, embora na realidade, infelizmente, o uso dos tesouros pelos usuários de um SRI raramente se concretize. Logo, pode-se dizer que os tesouros foram idealizados para ser uma ferramenta meio ou intermediária, que atua tanto na entrada como na saída de informações de um sistema, fazendo a mediação entre a linguagem do bibliotecário e a linguagem do usuário. Apóia o bibliotecário durante a representação dos assuntos contidos em um documento que entrará em uma base de dados ou catálogo de biblioteca e auxilia, ou pelo menos

⁷ Documento eletrônico.

deveria auxiliar, os usuários na representação de suas necessidades informacionais durante o processo de recuperação da informação.

As ontologias, por sua vez, são ferramentas fim, elaboradas para fazerem parte de um sistema de informação, ou seja, já estão inseridas no mesmo e atuam como um filtro no momento da busca de informação. Os tesouros até podem ser inseridos no sistema para auxiliar as buscas dos usuários; ainda assim, terão de ser ativados pelo usuário no momento da pesquisa. De forma muito sintética, seria o equivalente a dizer que, tradicionalmente, os tesouros estão fora do sistema, agindo como ferramenta de apoio e as ontologias estão dentro dos sistemas, sendo utilizadas automaticamente no momento da busca, possivelmente sem que o usuário perceba, pois ela faz parte da estrutura existente por trás do sistema, na programação do mesmo.

Sales e Café (2009, p. 14-15) listam as semelhanças entre os tesouros e as ontologias a partir dos termos, dos conceitos e dos objetivos de cada ferramenta, conforme sistematizamos a seguir:

- a) Termos: tesouros e ontologias fazem uso de termos especializados para darem conta de outros conceitos especializados; os termos genéricos e específicos dos tesouros podem ser considerados, nas ontologias, como os termos universais (Tipo) e particulares (Instâncias).
- b) Conceitos: as características do conceito, ou seja, os atributos predicáveis de cada objeto conceitualizado, regem o relacionamento entre os termos; os conceitos são categorizados por semelhança; o relacionamento entre os conceitos, que é regido pelas características que possuem, é hierárquico, podendo se manifestar de forma superordenada (do conceito mais específico para o mais geral) e subordinada (do conceito geral para o específico); a relação do conceito com o termo é que este representa, designa, reflete, expressa, personifica, sintetiza e comunica àquele, ou seja, o conceito é visto por meio do termo.
- c) Objetivos: tanto tesouros, quanto ontologias têm como objetivos controlar terminologias especializadas esclarecendo barreiras linguísticas e concebendo uma estrutura conceitual; e objetivam, também, potencializar a comunicação especializada, gerindo a linguagem específica e concebendo um vocabulário compartilhado. Tesouros e ontologias têm como objetivo

organizar informações especializadas para potencializar a recuperação da informação.

Em relação a esses três itens de semelhanças entre tesouros e ontologias, identificados na literatura a partir de pesquisa bibliográfica realizada por Sales e Café (2009), o único com o qual não concordamos plenamente é o item c, pois as ontologias não têm o objetivo de fazer um controle terminológico dos termos de uma área do conhecimento, como fazem os tesouros. As ontologias apenas realizam uma padronização terminológica dos termos para que a mesma possa funcionar como mecanismo de representação e recuperação da informação. Embora aparentemente controle e padronização sejam coisas semelhantes, elas possuem sentidos diferentes quando estamos falando em tesouros e ontologias. Almeida (2013) explica que uma ontologia pode ser considerada um tipo de vocabulário controlado apenas na medida em que especifica o conjunto de predicados que podem ser utilizados para fazer declarações sobre um recurso.

Em relação às diferenças, entre tesouros e ontologias, Sales e Café (2009) afirmam que as características que os distanciam são maiores que as características que os aproximam. Quanto aos termos, a principal diferença está na função exercida por eles. “Os textos relativos às ontologias focam mais nas funções desempenhadas pelos termos, sem se (pre)ocupar em defini-los, ficando apenas notória a ideia de que um termo é uma etiqueta que se refere a um conceito.” Já a literatura sobre tesouros “[...] atribui aos termos a função de evitar ou diminuir a flexibilidade da linguagem e descrever um conceito de maneira unívoca em um sistema de informação [...]” (SALES; CAFÉ, 2009, p. 16-17).

Os tesouros apresentam três tipos de relações conceituais entre os termos; são elas: hierárquicas (gênero-espécie e todo-parte), associativas (aproximação semântica) e de equivalência (sinonímia). Diferentemente das ontologias, que podem conter essas e outras mais, é uma ferramenta mais flexível do que os tesouros. Outra diferença está na utilização das relações ontológicas (todo-parte) e na utilização das relações de equivalência (sinonímia) adotadas na elaboração de tesouros. A sinonímia é muito relevante para os tesouros, mas não é objeto de interesse em uma ontologia. Ontologias descrevem as entidades reais e as relações entre elas, mas não descrevem seus termos e expressões relacionadas (SCHULZ et al. 2009). Em contrapartida, as diversas possibilidades de estabelecimento de relações parte-todo,

relevantes para as ontologias, não são aplicáveis no contexto de elaboração de tesouros (ALMEIDA; MENDONÇA; AGANETTE, 2013).

Outra diferença marcante entre tesouros e ontologias é o fato de que os tesouros são linguagens documentárias desenvolvidas para serem lidas e compreendidas por pessoas, e as ontologias, em função de seu caráter formal, são desenvolvidas com o objetivo aplicado e específico de serem lidas e compreendidas por computadores.

Em função dos problemas de recuperação da informação identificados no âmbito da internet, os quais as pesquisas relacionadas à web semântica buscam incessantemente resolver, deparamo-nos com algumas questões ambivalentes. Se por um lado, algumas pesquisas apontam que a hierarquização do conhecimento, como é feita nos tesouros, mostra-se ineficaz para atender aos requisitos mais flexíveis da web semântica; por outro lado, há a percepção de que não podemos, no cenário atual, pensar em sistemas de organização do conhecimento universais, tendo em vista as particularidades de cada área de conhecimento, além dos aspectos sociais, culturais, temporais, regionais, etc. (LARA, 2013).

Em artigos recentes, como o de Tredinnick (2013), encontramos a apresentação de propostas que rompem com algumas normas clássicas de organização do conhecimento e buscam na metáfora dos rizomas de Deleuze e Guattari (1987) a resposta para as questões trazidas pela web. Tredinnick (2013, p. 410, tradução nossa) explica que

[...] o rizoma é uma massa de raízes, sem aparente organização ou estrutura, sem centro, nem princípio organizador central, crescendo horizontalmente pela terra disponível. O rizoma é subversivo, colonizador e oportunista, forjando conexões *ad hoc* e estrutura interna.

Deleuze e Guattari (1987) apresentaram a metáfora dos rizomas em contraposição à tradicional árvore do conhecimento. Nesta metáfora, encontramos uma oposição entre a flexibilidade do rizoma e a rígida estrutura hierárquica da árvore do conhecimento. Em relação às ontologias, Tredinnick (2013) explica que o rizoma oferece uma visão equilibrada, que nos permite reconhecer aspectos da web que, de outra forma, seriam negligenciados. No entanto, pelo menos no que se refere à organização do conhecimento, essa ruptura parece contraditória. Considerando que uma das exigências para o funcionamento das ontologias, ou seja, para que as máquinas possam compreender os comandos e, de fato, executar a atividade

esperada, é necessário, justamente, que haja o princípio da obediência e da utilização de uma linguagem formal. Essa linguagem precisa de regras claras e objetivas para funcionar de acordo com o compromisso ontológico estabelecido pela equipe de elaboração da própria ontologia.

Deste modo, é preciso cautela ao interpretar a metáfora dos rizomas, pois sua utilização não precisa implicar necessariamente em um completo abandono da hierarquização. Metaforicamente falando, embora um rizoma apresente uma estrutura mais flexível do que a tradicional árvore do conhecimento, é possível perceber que alguns nódulos dão origem a outros nódulos deixando evidente uma relação hierárquica entre eles. É claro que uma estrutura rizomática é mais fácil de ser atualizada do que uma estrutura arbórea. Porém, é importante lembrar para qual propósito apresentamos o conhecimento de uma área especializada de forma sistemática.

No âmbito da web semântica e da internet, bem como para o funcionamento adequado das ontologias, a estrutura rizomática de fato pode ser mais adequada, já que será lida e interpretada por computadores. No entanto, ao pensarmos especificamente no cotidiano de bibliotecários que buscam um tesouro para a identificação do descritor mais adequado para a representação do assunto de um documento a ser indexado, ou ao pensarmos nos usuários que podem vir a utilizar o tesouro para buscar o termo que melhor representa sua necessidade de informação, perguntamo-nos: qual seria a apresentação gráfica mais clara e compreensível: a flexível e rizomática ou a mais rígida e arbórea? Acreditamos que, para a leitura e compreensão humana, a estrutura hierárquica delimitada de forma clara e objetiva é de mais fácil compreensão, conforme explica Broughton (2006). Talvez não seja necessário utilizar uma estrutura muito rígida ou necessariamente arbórea; porém, é de salutar importância que a natureza das relações entre os termos seja facilmente identificada por olhos humanos.

Não se pode negar completamente a hierarquização, pois sua existência é inerente ao saber. Qualquer área de conhecimento possui termos mais amplos (genéricos) e termos mais específicos. A identificação dessa relação entre os termos de uma área de conhecimento nos ajuda a compreendê-la melhor. Essa é razão pela qual, gostemos ou não, a ciência é dividida em domínios. A consciência em relação aos limites e fronteiras entre os campos é importante para as atividades de ensino e

transferência de informação e conhecimento, por exemplo. E isso não significa que estejamos necessariamente arraigados a uma visão positivista do mundo.

Embora seja evidente a necessidade de rever alguns princípios metodológicos de elaboração de tesouros e a desejável interoperabilidade – harmonização de tesouros diferentes – não podemos perder de vista a função dessas ferramentas. Logo, a adoção de uma estrutura rizomática para ontologias pode ser uma escolha adequada. Porém, para os tesouros nem tanto. A não ser que se desenvolva de modo que as diferentes categorias existentes em uma área do conhecimento possam ser facilmente identificadas, tendo em vista que, por razões pragmáticas do cotidiano profissional, um bibliotecário precisa identificar rapidamente entre o conjunto de termos de um tesouro o descritor que irá utilizar para a representação dos assuntos de um documento. Do mesmo modo, os usuários dos sistemas de informação que venham a utilizá-los, certamente, também desejarão identificar os termos que melhor representam suas necessidades de informação de forma ágil, sem precisar adivinhar o descritor utilizado pelos bibliotecários na representação dos documentos ou refazer suas buscas diversas vezes.

4.1 Características técnicas dos tesouros

Um tesouro apresenta três tipos de relações conceituais entre os termos são elas: hierárquicas (gênero-espécie e todo-parte), associativas (aproximação semântica) e de equivalência (sinonímia). As relações de equivalência são como referências cruzadas, elas remetem de um termo não autorizado no tesouro para outro termo sinônimo ou variante, o descritor e vice versa. Essa relação é apresentada no tesouro através do uso da palavra USE e da sigla UP (usado por). Desta forma, ao consultar um termo não autorizado na lista alfabética do tesouro, o usuário é informado que deve usar outro termo, o descritor. Assim, ao consultar um tesouro da área jurídica, um usuário pode buscar pelo termo “autônomo”; em função de esse termo não ser autorizado, o usuário do tesouro encontrará a seguinte indicação: USE “trabalhador autônomo”. Em contrapartida, junto ao termo “trabalhador autônomo”, também haverá referência ao termo “autônomo”, indicando que o mesmo é considerado um sinônimo do descritor. Para melhor compreensão, coletamos e

adaptamos alguns exemplos do Tesouro Jurídico da Justiça Federal do Brasil que apresenta essas relações de forma bastante clara.

AUTÔNOMO

USE TRABALHADOR AUTÔNOMO

TRABALHADOR AUTÔNOMO

UP AUTÔNOMO

As relações hierárquicas de um tesauro são representadas pelas siglas TG (termo genérico) e TE (termo específico). Ao buscar o termo “trabalhador autônomo” no Tesouro Jurídico da Justiça Federal do Brasil, vamos perceber que esse é um termo específico do termo genérico “trabalhador”. Veja o exemplo:

TRABALHADOR

TE TRABALHADOR AUTÔNOMO

TE TRABALHADOR AVULSO

TE TRABALHADOR NÃO-OPTANTE PELO FGTS

TE TRABALHADOR OPTANTE PELO FGTS

TE TRABALHADOR RURAL

TE TRABALHADOR TEMPORÁRIO

TE TRABALHADOR URBANO

TRABALHADOR AUTÔNOMO

UP AUTÔNOMO

TG TRABALHADOR

TR PRESTAÇÃO DE SERVIÇO

TR PROFISSIONAL LIBERAL

TR TRABALHADOR AVULSO

As relações associativas, por sua vez, são representadas no tesauro pela sigla TR (termo relacionado), que são termos que não apresentam uma hierarquia entre si, são correlatos. No quadro a seguir, podemos visualizar as relações associativas mais comuns de acordo com a International Federation of Library Associations and Institutions (IFLA).

Quadro 12 - Relações associativas

Relações Associativas	Exemplos
Causa / Efeito	Acidentes / Lesões
Processo / Agente	Medição de velocidade / Indicador de velocidade
Ação / Produto da ação	Tecelagem / Tecido
Ação / Paciente ou Objetivo	Ensino / Estudante
Conceito ou Coisa / Propriedades	Liga de aço / Resistência a corrosão
Ação ou Coisa / Contra-Agente	Praga / Praguicida
Coisa / Suas Partes (se não ocorre a relação hierárquica todo-parte)	Carro / Motor
Matéria prima / Produto	Uva / Vinho
Ação/ Propriedade	Comunicação / Habilidades de comunicação
Campo de estudo / Objetos ou fenômenos estudados	Silvicultura / Bosques

Fonte: IFLA (2010, p. 25. tradução nossa).

Se procurarmos pelo termo “industrialário”, por exemplo, no Tesouro Jurídico da Justiça Federal do Brasil, encontraremos a seguinte representação gráfica:

INDUSTRIÁRIO
TR INDÚSTRIA
TR TRABALHADOR

Isso ocorre, pois os termos “indústria” e “trabalhador”, do Tesouro mencionado, são relacionados ao termo “industrialário”.

É importante destacar que os tesouros são elaborados para serem utilizados por bibliotecários no momento da indexação; mas também deveriam ser utilizados por usuários da informação no momento da busca de documentos. Historicamente, esses sistemas conceituais foram idealizados para serem utilizados por bibliotecários e usuários, por isso, pode-se dizer que são uma ferramenta meio, pois ela serve de apoio para a representação dos assuntos contidos em um documento dentro de uma base de dados ou catálogo de biblioteca.

De acordo com Khoo e Jin-Cheon (2006), nos tesouros encontramos, principalmente, cinco relações paradigmáticas (hipônimo/hiperônimo, toponímia, meronímia/holonímia, sinonímia, antonímia), bem como a relação causa-efeito, que é uma importante relação sintagmática para a estrutura do conhecimento humano.

A relação de hiponímia tem sido referida na literatura sob vários nomes, incluindo “é um”, “tipo-de”, taxonômica, superordenada-subordinada, gênero-espécie e classe-subclasse. Hipônimo refere-se ao conceito de um termo mais específico e hiperônimo, ao conceito mais amplo. Em conjunto com a relação parte-todo, é uma relação hierárquica frequentemente encontrada em tesouros, taxonomias e ontologias. (KHOO E JIN-CHEON, 2006).

A meronímia, também conhecida como relação parte-todo, refere-se à relação entre um conceito e suas partes constituintes. Khoo e Jin-Cheon (2006) explicam que a distinção entre meronímia e hiponímia é clara para conceitos concretos, mas confusa para conceitos abstratos. De acordo com os autores, pode-se dizer que os hipônimos estão dentro dos conceitos, mas os merônimos ocorrem entre os conceitos. Outra diferença é que os hipônimos herdam as características dos hiperônimos, mas as partes não herdam a totalidade das características.

Embora seja muito raro encontrarmos sinônimos absolutos, eles também são encontrados nos tesouros. Duas expressões são absolutamente sinônimas se todos os seus significados são idênticos em todos os contextos linguísticos. Os tipos mais comuns de sinônimos encontrados nos tesouros são os quase-sinônimos (que não têm sentidos idênticos, mas apresentam o mesmo significado), e sinônimos parciais (que compartilham alguns sentidos, mas diferem em alguns aspectos, por exemplo, na forma como são utilizados, ou em alguma dimensão de sentido. (KHOO; JIN-CHEON, 2006).

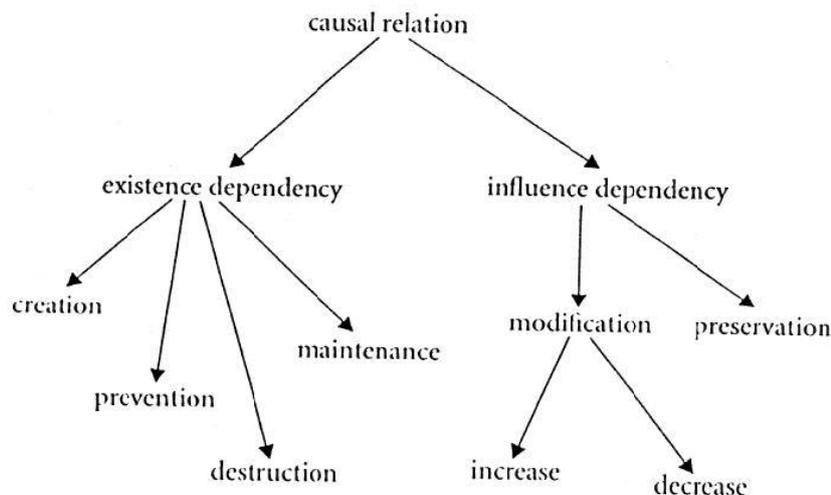
Para Khoo e Jin-Cheon (2006) a antonímia é uma das relações que as pessoas acham de mais fácil compreensão. Ela consiste em analisar um par de conceitos que estabelecem relação de oposição um ao outro, por exemplo: “amor” e “ódio”, “genialidade” e “loucura”.

A relação de causa-efeito, presente nos tesouros, não é tão simples de definir como pode parecer à primeira vista. De acordo com Khoo e Jin-Cheon (2006), o conceito de causalidade é complexo e surpreendentemente difícil de definir. Para os autores, pode-se distinguir entre causas necessárias e suficientes. Os psicólogos

Jaspars, Hewstone e Fincham (1983)⁸ e Jaspars (1983)⁹ (apud KHOO E JIN-CHEON, 2006) encontraram evidências de que tomar uma causa como uma condição necessária e/ou uma condição suficiente é um julgamento que varia conforme o tipo de entidade, sendo considerado de *status* causal. Uma causa pode ser atribuída tanto a uma pessoa, como a uma situação; para os autores, isso ocorre pois ela é atribuída, simultaneamente, como um estímulo e como uma condição necessária e suficiente. Contudo, "[. . .] uma causa pessoal é vista mais como uma condição suficiente, enquanto causas situacionais são concebidas principalmente como condições necessárias" (JASPARS; HEWSTONE; FINCHAM, *apud* KHOO; JIN-CHEON, 2006, p. 179, tradução nossa).

Barriere (2002) divide as relações de causa-efeito em duas categorias: em relação à existência e em relação à influência, como podemos observar na apresentação gráfica elaborada pela autora.

Figura 2 – Hierarchical refinement of the causal relation



Fonte: Barrière (2002, p. 98).

Desta forma, pode-se perceber que a relação causal que depende da existência se subdivide em quatro classes: criação, prevenção, destruição e manutenção. Isso

⁸ JASPARS, J.; HEWSTONE, M.; FINCHAM, F. D. Attribution theory and research: the state of the art. In: JASPARS, J.; FINCHAM, F. D; HEWSTONE, M. (Eds.). **Attribution theory and research: conceptual, developmental and social dimensions**. London: Academic Press, 1983. p. 3-36.

⁹ JASPARS, J. The process of attribution in common-sense. In: HEWSTONE, M. (Ed.). **Attribution theory: social and functional extensions**. Oxford, UK Blackwell, 1983. p. 2844.

significa que uma relação causal pode ocorrer em relação à criação, prevenção, destruição e manutenção de algo. Já a relação causal que depende de influência se subdivide em apenas duas classes: modificação e preservação. Ou seja, uma relação causal poderá ocorrer por influência da modificação e preservação de algo. Além disso, a maior ou menor modificação também são fatores que influenciarão uma relação causal.

A estrutura conceitual dos tesauros é fundamental para o seu desempenho adequado, pois é ela quem vai permitir a visualização das relações semânticas entre os termos e descritores que o constituem. A análise dos conceitos, uns em relação aos outros, possibilita a organização da estrutura conceitual de um determinado domínio. Ao colocá-los dentro de um sistema, através do uso de termos, é possível compreender com mais clareza a estrutura de organização interna de uma área do conhecimento, ou seja, é possível perceber a sua essência e isso vai se refletir, posteriormente, na organização dos acervos documentais sobre esses assuntos. A análise conceitual, portanto, viabiliza a formação de categorias que serão ordenadas dentro de um sistema em classes e subclasses, as quais, no âmbito da Ciência da Informação, são desenvolvidas com o objetivo de organizar documentos. (RANGANATHAN, 1967).

4.2 Fundamentos de Elaboração de Tesauros

Na literatura sobre organização e representação do conhecimento, identificamos algumas teorias que dão suporte ao planejamento e à elaboração de linguagens documentária, tais como: a Teoria Clássica de Aristóteles, a Teoria da Classificação Facetada de Ranganathan, a Teoria do Conceito de Dahlberg, a Teoria Geral da Terminologia de Wüster e a Teoria dos Protótipos de Eleanor Rosch. Por uma questão de organização textual, optamos por discorrer sobre a abordagem de cada teoria em relação aos conceitos e categorias separadamente. Assim, nos itens a seguir apresentamos respectivamente as seguintes teorias: Teoria Clássica de Aristóteles, Teoria da Classificação Facetada, Teoria do Conceito, Teoria dos Protótipos e Semântica de Frames de Fillmore. Destacamos que a Teoria Geral da Terminologia

não é apresentada neste capítulo, pois já foi abordada no capítulo 3 do presente trabalho.

4.2.1 Teoria Clássica de Aristóteles

De acordo com Felber (1984), os primeiros estudos sobre a natureza dos conceitos foram realizados pelas antigas escolas gregas; Platão, com a sua teoria das idéias (Phaidron), deu início a teoria dos conceitos. Aristóteles, com a obra “Organon”, desenvolveu as leis básicas do conceito, características, raciocínio, inferência, definição, etc. Em vista disto, é importante falarmos, mesmo que brevemente, sobre a Teoria Clássica de Aristóteles, pois seu papel histórico figura como ponto de partida dos estudos sobre a organização do conhecimento.

Para Schreiner (1979), os princípios aristotélicos determinaram as diretrizes da classificação do conhecimento por mais de dois mil anos. Por ser baseada na suposição de que o mundo real é estruturado em grupos hierárquicos, que apresentam partes comuns e propriedades inerentes, bem como estabelecem que as categorias apresentam limites sólidos, a Teoria Clássica tem um papel importante nos estudos de classificação. Embora ela apresente inconsistências que, mais tarde, foram discutidas e melhor solucionadas por outras teorias, ela representa o ponto de partida para o aprimoramento da atividade de categorização. Um aspecto importante dos princípios aristotélicos é a divisão por gênero/espécie. Wright (2008) destaca que Aristóteles não criou essa divisão, pois ela já existia, ele simplesmente adaptou e codificou as antigas tradições orais de taxonomias populares. Sua grande contribuição, na verdade, foi tornar as velhas estruturas implícitas em explícitas, ou seja, tornar esse princípio de divisão de classes explícito. Mesmo assim, não podemos ignorar sua importância, até mesmo pelo fato de que ainda hoje utilizamos o princípio de divisão por gênero/espécie para a organização de categorias em tesouros.

Segundo a Teoria Clássica de Aristóteles, “[. . .] as categorias são definidas apenas pelas propriedades comuns a todos os seus membros” (LIMA, 2007, p. 161). Desta forma, um determinado objeto será considerado ou não membro de uma categoria conforme as características necessárias e suficientes que definem a

categoria em questão. Smith e Medin (1981) explicam que o modelo clássico apresenta três pressupostos, são eles:

- a) a representação de conceitos implica descrever sinteticamente uma classe. Para cada conceito, existem algumas características definidoras, que vão determinar se algo pode ou não ser classificado em uma determinada classe. Por exemplo, os gatos são mamíferos? Fazemos uma comparação das características definidoras de um gato com a representação sumária ou com o conjunto de características definidoras dos mamíferos? Observamos que as fêmeas dessa espécie possuem glândulas mamárias e podem alimentar seus filhotes com leite, possuem sangue quente, têm pelos e se reproduzem; se considerarmos essas condições, concluiremos que os gatos possuem tais características; então, sob essa lógica, os gatos são mamíferos.
- b) as características definidoras de uma categoria devem ser tanto individualmente necessárias, quanto conjuntamente suficientes para definir uma categoria. Por exemplo, para uma figura ser categorizada como um retângulo, ela deverá apresentar dois lados com tamanhos e ângulos iguais; se um dos lados for diferente, então a figura não será um retângulo. Porém, outras propriedades do retângulo, como o material utilizado para desenhá-lo ou o local onde o mesmo foi desenhado, não interferem em sua categorização.
- c) as categorias são agrupadas de modo que a categoria subordinada possui todas as características da categoria superordenada. Desta forma, quando julgamos que um objeto apresenta todas as características que o tornam uma "margarida", também sabemos que ele tem todas as características definidoras de "flor", "planta" e "coisa viva".

Outra contribuição importante de Aristóteles são suas dez categorias para a determinação formal do ser (substância, quantidade, qualidade, relação, lugar, tempo, posição, estado, ação e sofrimento da ação). De acordo com Schreiner (1979), essas categorias apresentam um importante caráter classificatório e foram utilizadas em vários sistemas de organização do conhecimento, entre eles a "Colon Classification" (1933), de Ranganathan, e o "Thesaurus of English words and phrases" (1852), de

Peter Mark Roget. Acrescentamos, ainda, que as categorias de Aristóteles também serviram de inspiração para a Teoria do Conceito de Dahlberg (1978 -).

4.2.2 Teoria da Classificação Facetada

A Classificação Facetada estabelece princípios para a organização de conceitos hierarquicamente estruturados e, em função de sua abordagem “universal”, é amplamente citada nos estudos sobre classificação e elaboração de tesouros, desde a publicação da primeira edição do livro “Prolegomena to Library Classification” por Ranganathan, em 1937. Também é bastante significativa sua citação em textos e estudos mais recentes sobre elaboração de ontologias que visam o aperfeiçoamento de sistemas de recuperação da informação no âmbito da web semântica.

Ranganathan (1892-1972), por sua inestimável contribuição na área de classificação de assunto na Biblioteconomia, é considerado o pai da moderna teoria da classificação (DAHLBERG, 1979). Em seu livro “Prolegomena”, “[. . .] Ranganathan formula regras, enuncia cânones e postulados a partir dos quais extrai seus princípios e lança seus chamados “processos” (*devices*) para a formação de representações de conceitos em nível teórico.” (DAHLBERG, 1979, p. 5). A elaboração de uma teoria própria foi motivada por sua insatisfação com os sistemas de classificação até então existentes, especificamente a Classificação Decimal de Dewey (CDD) e a Classificação Decimal Universal (CDU), que o autor julgava falhos em razão da dificuldade desses sistemas em acolher a representação de novos assuntos.

Cabe ressaltar que Ranganathan foi o primeiro autor que se preocupou em formular e descrever detalhadamente as bases teóricas de seu sistema de classificação, uma vez que os sistemas anteriores foram constituídos de forma empírica e arbitrária. Ranganathan descreveu em sua obra “Prolegomena” (1957; 1967) como organizar o conhecimento contido nos documentos. Especificamente, tratava sobre como determinar o assunto de uma obra, estabelecer sua notação (conjunto de símbolos, geralmente constituído por números e letras, que é oferecido a cada documento de acordo com seu assunto; ao ser adicionada na lombada dos livros, a notação permite a organização das obras nas estantes) e, conseqüentemente, reuni-la com outras obras do mesmo assunto. O autor estabeleceu regras básicas para a

divisão e ordenação de classes através de seus cânones, princípios e categorias fundamentais, os quais descrevemos a seguir.

Cânones para o trabalho no plano das idéias

Conjunto de cinco cânones que servem para regulamentar o planejamento de um esquema de classificação. O quadro a seguir apresenta uma visão abrangente do conteúdo de cada cânone que se subdivide com o objetivo de detalhar os diferentes aspectos a serem considerados durante o planejamento de cada classe ou categoria de um sistema de classificação.

Quadro 13 - Cânones para o Trabalho no Plano das Idéias

Característica	Sucessão de características	Renque de classes	Cadeia de classes	Sequência de filiação
Diferenciação	Concomitância	Exaustividade	Extensão decrescente	Classes subordinadas
Relevância	Sucessão relevante	Exclusividade	Modulação	Classes coordenadas
Verificabilidade	Sucessão consistente	Sequência útil		
Permanência		Sequência consistente		

Fonte: Gomes, Motta e Campos (2006).

Cânones para o trabalho no plano verbal

Conjunto de cânones relacionados ao tratamento terminológico que deve ser observado tanto para a elaboração de um esquema de classificação, como para a interpretação dos termos existentes quando vamos utilizá-los no processo de classificação. São eles:

- a) Cânone do Contexto: a denotação de cada termo em um esquema de classificação deverá ser decidida à luz das diferentes classes, ou isolados pela ordem inferior (que pertença à mesma cadeia primária da classe), ou do isolado

denotado pelo termo em questão. [Considerar a rede de conceitos já existentes].

- b) Cãnone da Enumeração: a denotação de cada termo em um esquema de classificação deverá ser decidida à luz das classes enumeradas nas várias cadeias (ligações inferiores) com a classe denotada pelo termo, através do seu elo comum em primeiro lugar;
- c) Cãnone do Uso Corrente: o termo utilizado para designar as classes em um esquema de classificação deve ser o corrente entre os especialistas na área de assunto coberta pelo esquema. Ou seja, deve ser o termo mais atual e aceito pelos especialistas do domínio em questão.
- d) Cãnone da Neutralidade: os termos usados para designar as classes de um esquema de classificação não devem ser “opinativos”, ou seja, não devem deixar transparecer alguma parcialidade do classificador. Exemplo: Utilizar um termo depreciativo ou preconceituoso para designar uma categoria.

Destacamos a semelhança dos cânones para o trabalho no plano verbal de Ranganathan com a Teoria Comunicativa da Terminologia, principalmente em relação ao cânone do uso corrente, no qual o autor sugere que a escolha dos termos de um sistema de classificação deve estar de acordo com os termos utilizados pelos especialistas da área de conhecimento para a qual o sistema é elaborado. Além dos cânones Ranganathan (1967), também estabeleceu-se nove princípios gerais de classificação, os quais sintetizamos no quadro a seguir.

Quadro 14 – Princípios gerais de classificação

Princípio	Descrição
Concretude	Deve ser empregado para a ordenação de classes de modo que aquela que apresentar características mais concretas deve preceder a que apresentar características mais abstratas.
Artificialidade	Deve ser empregado na ordenação de classes de modo que aquela que apresentar características mais próximas da "coisa em si" deve preceder a que apresentar características mais artificiais.
Sequência Útil	Deve ser empregado para determinar a ordem de disposição das classes em um renque.
Contiguidade Espacial	Deve ser empregado quando as classes de um renque são próximas e por isso podem ser dispostas em sequência espacial. Essa sequência pode ser em uma linha direcional, radial ou circular, mas na maioria das vezes ocorrem em pares antitéticos como de cima para baixo, da esquerda para a direita, no sentido horário, da periferia para o centro, etc.
Medida Quantitativa	Deve ser empregado quando o assunto de um renque admite a utilização de medidas quantitativas para sua distinção, nesse caso a ordenação poderá ser em quantidade crescente ou decrescente.
Complexidade Crescente	Deve ser empregado quando as classes de um renque mostram diferentes graus de complexidade. Nesse caso, elas devem ser organizadas na sequência de complexidade crescente, ou seja, do menos complexo para o mais complexo. Como exemplo Ranganathan cita a ordenação que utilizou na Colon Classification para os elementos na área de Linguística: som isolado, sílaba, palavra, frase, oração, período.
Sequência Canônica	Deve ser empregado quando as classes de um domínio são tradicionalmente reconhecidas em uma sequência específica. Nesse caso, é conveniente manter a sequência tradicional.
Garantia Literária	Neste caso, as classes de um renque podem ser dispostas em sequência decrescente da quantidade de documentos publicados sobre eles. Ranganathan adverte que a adoção deste princípio requer cuidado e bom senso para que o ponto de vista do classificador se mantenha neutro, sem favorecer alguns aspectos em detrimento de outros, como por exemplo, o país de origem da classificação.
Sequência Alfabética	Deve ser empregado apenas se nenhuma outra sequência for útil.

Fonte: Adaptado pela autora de Ranganathan (1967).

Ranganathan (1967) destaca que estes princípios podem ser usados conjuntamente sem que haja uma regra definitiva que sirva a todos. Ou seja, devem-se utilizar os princípios que melhor se adaptem ao contexto (domínio) da classificação a ser elaborada. Entre os princípios apresentados pelo autor, discordamos em parte de apenas um aspecto do princípio da garantia literária, especificamente, quanto a possibilidade de organizar as classes de um renque de acordo com o volume de obras publicadas sobre um assunto. Acreditamos que a quantidade de publicações sobre uma temática não é parâmetro suficiente para determinar a relevância e a ordenação dos termos dentro da estrutura de um sistema conceitual. Visto que, alguns assuntos, em função de suas características, são pouco pesquisados ou ainda dizem respeito à uma descoberta recente, razão pela qual apresentam uma quantidade pequena de publicações. Quanto a neutralidade do ponto de vista do classificador, Hjørland (2007) acrescenta que, em função dos diferentes paradigmas existentes, qualquer sistema conceitual pode ser examinado a partir de visões dominantes ou alternativas presentes nas publicações de uma área de conhecimento. Para o autor não existem sistemas neutros,

Qualquer ferramenta semântica pode estar mais ou menos em harmonia, ou em conflito, com os pontos de vista representados na literatura [...] Tudo o que podemos concluir é que uma condição prévia para a concepção de um SOC de qualidade é que o designer [classificador] conheça os diferentes pontos de vista e seja capaz de fornecer uma solução razoavelmente informada e negociada. Além disso, o designer de um determinado SOC deve analisar, a partir de um ponto de vista pragmático, quais os objetivos que o SOC procura cumprir. (HJØRLAND, 2007 b, p. 389, tradução nossa).

Hjørland (2007) destaca, ainda, a importância da garantia literária em comparação à garantia do especialista. Para o autor, “Os cientistas da informação sentem-se mais seguros ao consultar “especialistas” em vez de documentos, mas os documentos relevantes são escritos por especialistas e são igualmente fontes válidas, se não mais.” (HJØRLAND, 2007 b, p. 390, tradução nossa).

Lima (2002, p 190) explica que Ranganathan “[...] sob influência de sua área inicial, a matemática, estruturou o conhecimento de maneira que os assuntos compostos, sinteticamente, surgiam a partir de conceitos elementares”. Dahlberg (1979) complementa que o autor “matematizou” a classificação através de uma abordagem qualitativa similar à do filósofo alemão G. W. Leibniz (1646-1716). Leibniz procurava uma “*characteristica universalis*”, um tipo de linguagem conceitual para a

expressão combinatória de qualquer conceito e assunto existente no mundo; assim, eventualmente influenciou G. Frege em sua lógica predicativa e "Begriffsschrift", bem como a maior parte da lógica moderna. (DAHLBERG, 1979, p. 5).

Em vista disto, especula-se que Ranganathan encontrou a "característica universalis" que Leibniz buscava ao propor a divisão dos assuntos em seus elementos constituintes e depois “[. . .] formular e formalizar, através de suas ‘fórmulas de facetas’, enunciados sobre os assuntos encontrados nos livros e em outros tipos de documentos.” (DAHLBERG, 1979, p. 5). De acordo com Lima (2002, p.190), “[. . .] faceta é a coleção de termos que apresenta um relacionamento hierarquicamente igual com o assunto global, refletindo a aplicação de um princípio básico de divisão.”. Barbosa (1972) explica que, na visão de Ranganathan, a análise de um assunto por suas facetas consiste em verificar as manifestações de certas características que obedecem a postulados pré-determinados. As divisões de uma faceta são chamadas de foco isolado ou apenas isolado.

A Classificação Facetada apresenta cinco categorias fundamentais. São elas: Personalidade (P), Matéria (M), Energia (E), Espaço (S, de *space*) e Tempo (T), também conhecidas, simplesmente, pela sigla PMEST, conforme apresentamos a seguir:

- a) **Tempo** - As Idéias isoladas de Tempo – como milênio, século, década, ano, e assim por diante – são suas manifestações. Idéias de Tempo de outras espécies – como dia e noite, estações como verão e inverno, tempo com qualidade meteorológica como úmido, seco, tormentoso – também são manifestações da Categoria Fundamental "Tempo".
- b) **Espaço** – corresponde ao local de pertencimento de um dado objeto, seja ele indivíduo, coisa, idéia, fenômeno, entre outras entidades. A superfície da terra, o espaço dentro dela e o espaço fora dela são manifestações desta Categoria Fundamental.
- c) **Energia** - sua manifestação é ação de uma espécie ou outra. A ação pode ser entre e por todas as espécies de entidade, inanimada, animada, conceitual, intelectual e intuitiva.
- d) **Matéria** - suas manifestações são de duas espécies – Material e de Propriedade. O material pode ser intrínseco a um objeto, mas não é o próprio objeto, pois o mesmo material pode aparecer em outra entidade. O mesmo ocorre com uma propriedade que pode ser intrínseca a um objeto, mas não ser

o objeto em si. Isto ocorre quando o material é integrante do objeto podendo se constituir como uma de suas características.

- e) **Personalidade** – é uma categoria fundamental de grande dificuldade de identificação. Ranganathan propõe o método residual para identificar sua manifestação: não é "Tempo", não é "Espaço", não é "Energia", ou "Matéria", portanto é considerada uma manifestação da Categoria Fundamental Personalidade.

Ressaltamos que, embora a Teoria da Classificação Facetada se refira à elaboração de sistemas de classificação, que visam, inclusive, à organização física de um acervo, ela é amplamente utilizada para a elaboração de tesouros. Por utilizar categorias universais, pode ser adaptada e aplicada em diferentes domínios de conhecimento. Seu método de categorização trouxe princípios basilares para a organização de conceitos de um domínio. Como consequência, deu origem, no final da década de 70 do século passado, ao Thesaurofacet de Aitchison, que agrega os princípios da Classificação Facetada em sua elaboração. (CAMPOS; GOMES, 2006); o que representou um avanço metodológico devido à possibilidade de organizar os conceitos em categorias, pois, até então, o arranjo dos tesouros era apenas alfabético.

4.2.3 Teoria do Conceito

A Teoria do Conceito de Dahlberg, desenvolvida nos anos 70 do século passado, teve forte influência da Teoria Geral da Terminologia de Wüster, e da Classificação Facetada de Ranganathan. Para a área da Ciência da Informação, as propostas de Dahlberg contribuíram para maior compreensão e determinação do que é considerado conceito, para fins de representação/recuperação da informação “[. . .] apresenta princípios que podem auxiliar na determinação do termo e de suas relações [. . .]” (CAMPOS, 2001, p. 20). Dahlberg, após uma análise mais profunda da definição de conceito como uma unidade do pensamento elaborada por Wüster, define conceito como uma “unidade de conhecimento”. Para a autora, pensamento remete a algo subjetivo, que está na cabeça das pessoas, conhecimento, por outro lado, “[...]”

pressupõe um entendimento mais objetivo de algo observável [...]”. (CAMPOS, 2001, p. 101).

Dahlberg (1979, p. 361) afirma que

[. . .] os conceitos são sínteses rotuladas de enunciados verdadeiros sobre objetos do pensamento: esses enunciados – asserções – levam ao reconhecimento ou à separação das características dos conceitos, que também podem ser consideradas como os elementos dos conceitos.

Motta (1987) explica que o conceito, na teoria de Dahlberg, é formado por um item de referência, predicções verdadeiras sobre esse item de referência e síntese das predicções verdadeiras sobre o item de referência, através de um termo ou nome. Portanto, “[. . .] o conceito não é um todo composto de partes distintas. [. . .] o conceito (ideia) é uno, expresso multiplamente, por um referente, pelos elementos que o definem e pelo que, aqui, chamamos de ‘termo’ que o designa.” (MOTTA, 1987, p. 33).

Dahlberg (1978) divide os conceitos em gerais e individuais, e explica que, com base nesses enunciados, é possível elaborar os conceitos relativos a diversos objetos. Cada enunciado verdadeiro representa um elemento do conceito e a soma de todos os enunciados verdadeiros fornece o conceito de um determinado objeto. Para a autora, o conceito é constituído de elementos (características) que se articulam numa unidade estruturada, de modo que os elementos contidos nos conceitos gerais também estão presentes nos conceitos individuais, o que possibilita o ordenamento dos mesmos. Conceito é definido por Dahlberg (1978, p. 102) como “[. . .] a compilação de enunciados verdadeiros sobre determinado objeto, fixada por um símbolo linguístico.” A autora explica que:

Cada enunciado apresenta (no verdadeiro sentido de predicção) um atributo predicável do objeto que, no nível de conceito, se chama característica. Muitas vezes não se trata de um atributo a que corresponde uma característica, mas de uma hierarquia de características, já que o predicado de um enunciado pode tornar-se sujeito de novo enunciado e assim sucessivamente até atingirmos uma característica tão geral que possa ser considerada uma categoria. (Entende-se aqui por categoria o conceito na sua mais ampla extensão). (DAHLBERG, 1978, p. 102).

Ainda de acordo com a autora, as características podem ser simples (uma única propriedade, exemplo: redondo, colorido) ou complexas (mais de uma característica, exemplo: moldado em metal, pintado com tinta azul, etc). Com base nas categorias aristotélicas, Dahlberg (1978, p. 104) apresenta como exemplo um conjunto de

características que podem ser utilizadas sozinhas ou combinadas com a função classificatória de ordenar conceitos e índices, podem definir conceitos e podem formar o nome dos conceitos; são elas: matéria, qualidade, quantidade (extensão), relação, processo (atividade), modo de ser, posição, localização (lugar) e tempo. Tais características também podem ser divididas em essenciais (necessárias) e acidentais (adicionais ou possíveis).

Em relação às características, a autora esclarece, ainda, que, para os conceitos gerais, as características essenciais têm mais importância do que as acidentais e que a presença de características idênticas em conceitos diferentes indica a existência de relações entre eles. As relações podem ser dos seguintes tipos:

- a) relação hierárquica (gênero e espécie) – ocorre entre dois conceitos diferentes com características idênticas e um deles apresenta uma característica a mais do que o outro, por exemplo: “trabalhador” – “trabalhador doméstico”;
- b) relação partitiva – ocorre entre um todo e suas partes, por exemplo: “jurisprudência” – “acórdãos”;
- c) relação de oposição (negação) – ocorre principalmente entre conceitos que apresentam propriedades e são contrários entre si, tendo, como exemplos: “emprego” – “desemprego”;
- d) relação funcional (intersecção) – ocorre entre conceitos que expressam processos, deve-se considerar as valências semânticas dos verbos e respectivos complementos, por exemplo: “produto” – “produtor” – “comprador”.

A partir da compreensão do que são as características de um conceito, Dahlberg (1978) explica o sentido da intensão e da extensão do conceito. A intensão do conceito é a soma total das suas características. É, também, a soma total dos respectivos conceitos genéricos e das diferenças específicas ou características especificadoras.

[. . .] A extensão do conceito pode ser entendida como a soma total dos conceitos mais específicos que possui. Pode também ser entendida como a soma dos conceitos para os quais a intensão seja verdadeira, ou seja, a classe dos conceitos de tais objetos aos quais pode-se afirmar que possuem aquelas características em comum que se encontram na intensão do mesmo conceito. (DAHLBERG, 1978, p. 105).

Logo, a partir da identificação da intensão e extensão dos conceitos, é possível categorizá-los e combiná-los de diversas formas. Dahlberg (1978) sugere uma categorização semelhante à categorização utilizada nas classificações facetadas, desta forma os conceitos podem ser categorizados por:

- a) Objetivos: por exemplo: plantas, produtos, papel, etc.
 - b) Fenômenos: por exemplo: crescimento, chuva, tráfego, etc.
 - c) Processos: por exemplo: imprimir, sintetizar, etc.
 - d) Propriedades: por exemplo: cego/cegueira, suave/suavidade, etc.
 - e) Relações: por exemplo: causalidade, necessidade.
 - f) Dimensão: por exemplo: espaço, tempo, posição, etc.
- (DAHLBERG, 1978, p. 105).

Por fim, a autora destaca a importância de obtermos a definição correta dos termos. Tendo em vista que as definições fixam o limite de um conceito, elas desempenham papel fundamental na argumentação e nas comunicações verbais, bem como constituem elementos necessários na construção de sistemas científicos. Contudo, a elaboração das mesmas depende do conhecimento que se tem dos respectivos assuntos. (DAHLBERG, 1978).

É importante destacar que Dahlberg reúne em sua teoria aspectos tanto da Teoria Geral da Terminologia, como da Teoria da Classificação Facetada. Suas explicações sobre características essenciais e acidentais, bem como sobre a intensão e a extensão dos conceitos são muito parecidas com as definições de Wüster sobre as características classificatórias dos conceitos, que constam em sua obra póstuma, de 1998. Já a categorização que propõe agrega os princípios de Ranganathan como método de estruturação de um domínio. A partir da influência da Teoria do Conceito, surge o que Campos e Gomes (2006, p. 354) chamam de tesouro conceitual, que “[...] reúne dois princípios fundamentais para sua elaboração: o conceito, como unidade de representação, e o uso de categorias, como base para organização de sistemas de conceitos.” O caráter didático da Teoria do Conceito de Dahlberg justifica sua ampla utilização como fonte de apoio para a elaboração de linguagens documentárias no âmbito da Ciência da Informação.

Diante do exposto, procuramos demonstrar as principais características de cada teoria e suas contribuições para a elaboração de tesouros. Ao compará-las, concluímos que fica evidente o caráter conceitual de cada uma delas, o que é bastante compreensível, já que os tesouros são sistemas conceituais. Nesse sentido, destacamos que reconhecemos a importância da estrutura conceitual dos tesouros, e

não é nosso objetivo contestá-la nesta tese. Porém, para o aperfeiçoamento dessas linguagens documentárias, acreditamos que é necessário contemplar em sua elaboração aspectos linguísticos característicos da língua em uso, os quais poderiam ser aplicados através das relações de equivalência dos tesouros. Essa é a razão pela qual buscamos subsídios na vertente contemporânea da Terminologia, especialmente na Teoria Comunicativa da Terminologia, para a elaboração de nossa proposta.

4.3 Normas de Elaboração de Tesouros

A norma de elaboração de tesouros mais recente e internacionalmente vigente é a ISO 25964, que se divide em duas partes. A primeira parte, publicada em 2011, apresenta recomendações para o desenvolvimento e manutenção de tesouros monolíngues ou multilíngues. Deste modo, substitui as normas anteriores, a ISO 2788, de elaboração de tesouros monolíngues, publicada em 1974 e atualizada em 1986, e a ISO 5964 de elaboração de tesouros multilíngues, publicada em 1985. Ou seja, a primeira parte dessa nova norma, é constituída pela união e reformulação das normas ISO 2788 e ISO 5964 (INTERNATIONAL..., 2013). É importante ressaltar que, embora apresente novo conteúdo para atender as necessidades de recuperação da informação em rede, a primeira parte da norma ISO 25964 também preserva muitas das recomendações contidas nas normas anteriores.

A segunda parte da norma, publicada no início de 2013, trata especificamente da interoperabilidade de tesouros e outros tipos de vocabulários controlados. Seu conteúdo, ao contrário do conteúdo da primeira parte, é completamente novo e foi formulado com o objetivo de atender às demandas de recuperação da informação trazidas pelo amplo desenvolvimento da internet. Em vista disto, a norma ISO 25964-2, além de descrever e comparar as características de diferentes vocabulários, também explica como configurar o mapeamento entre os conceitos desses diferentes vocabulários, para que a interoperabilidade entre eles de fato se realize. (INTERNATIONAL..., 2013).

Ao analisarmos a primeira parte da norma 25964, percebemos que o conceito é definido como uma “unidade de pensamento”. Em nota, a norma apresenta a seguinte explicação:

Os conceitos podem ser expressos de diferentes maneiras. Eles existem na mente como uma entidade abstrata independente dos termos usados para expressá-los. Variam do muito simples, por exemplo, "criança", aos mais complexos, por exemplo, "legislação de proteção da criança". (INTERNATIONAL..., 2013, p. 3, tradução nossa).

Isso significa que a norma ainda emprega a mesma definição de conceito utilizada por Wüster, quando elaborou a Teoria Geral da Terminologia, considerando a existência da primazia dos conceitos sobre os termos e, portanto, o caráter onomasiológico da terminologia. Manter essa definição de conceito parece um retrocesso teórico, tendo em vista que, já em 1978, Dahlberg, ao formular sua Teoria do Conceito, contestou essa definição argumentando que pensamento remete a algo subjetivo e que, por isso, seria melhor definir conceito como uma unidade de conhecimento, tendo em vista que conhecimento “[...] pressupõe um entendimento mais objetivo de algo observável [...]” (CAMPOS, 2001, p. 101). Ao manter essa definição de conceito, a norma ISO 25964 também ignora a norma ISO 704 de 2009 (Terminology work - Principles and methods) que, assim como Dahlberg (1978), define conceito como uma unidade de conhecimento.

Will (2012) explica que a nova norma oferece uma apresentação rigorosa das entidades e relacionamentos que, além de esclarecer e uniformizar as interpretações variadas e conflitantes que existem, também pode ser implementada de forma consistente em sistemas automatizados. A maior parte do modelo apresentado pela norma é compatível com o formato idealizado pelo World Wide Web Consortium (W3C), o SKOS (Simple Sistema de Organização do Conhecimento), para uso na Web Semântica, pois foi projetado em formato adequado para a interpretação de computadores. (WILL, 2012).

Clarke e Zeng (2012) esclarecem que os princípios de elaboração de tesouros não mudaram muito e ressaltam que nas normas anteriores não havia uma diferença clara entre os conceitos de uma busca por informação e os termos que expressam a consulta. As autoras explicam que a confusão causada pela falta de clareza entre termo e conceito já havia sido constatada na norma ISO 2788, tanto que, em sua versão mais atualizada de 1986, o texto chama a atenção dos leitores para o fato, informando que “Para fins práticos, termo e conceito, são por vezes utilizados como sinônimos” (CLARKE; ZENG, 2012, p. 21, tradução nossa).

Essa explicação é bastante relevante, e nos ajuda a compreender as razões pelas quais a norma vigente destaca que, embora as convenções como TG (termo genérico) e TE (termo específico), por exemplo, sejam mantidas, as mesmas devem ser interpretadas como “conceito mais amplo” e “conceito específico” ao invés de “termo genérico” e “termo específico” (WILL, 2012). Isso ocorre, pois, no âmbito da nova norma, “[...] o princípio geral é de que os conceitos não são definidos pelos termos utilizados para classificá-los, mas sim por seu lugar em uma hierarquia e por notas de escopo que especificam critérios para distingui-los entre conceitos semelhantes.” (WILL, 2012, p. 288-289, tradução nossa).

Clarke e Zeng (2012) esclarecem que, em função da Web Semântica, é importante estabelecer a diferença entre termos e conceitos, pois os computadores precisam de um modelo de dados explícito para que também possam fazer essa distinção. Will (2012) acrescenta que o novo modelo é baseado no entendimento de que os tesouros mostram as relações entre os conceitos diferentes dos termos que são utilizados para rotular esses próprios conceitos. Não concordamos com esse aspecto da norma, pois acreditamos no caráter semasiológico dos termos, conforme os pressupostos teóricos da Teoria Comunicativa da Terminologia. Nesse sentido, se os termos são unidades dotadas de forma e conteúdo que, além de denominar, também significam (CABRÉ, 1999; FREIXA, 2002), ao representarem um conceito, não podem ser descolados desses. Embora um único conceito possa ser representado por diferentes termos, esses diferentes termos, ao serem utilizados no discurso dos especialistas, denotam algum significado específico que pode estar relacionado com diversos fatores como: tipo de texto, tipo de público para o qual o texto é dirigido, nível de conhecimento especializado do autor e/ou dos possíveis leitores do texto, etc. Porém, como vimos anteriormente, a norma é elaborada com base nos pressupostos teóricos de Wüster e é por esse motivo que seus elaboradores insistem em fazer essa distinção entre termo e conceito.

Outro aspecto importante da norma, motivado pelo desejo de adequação da mesma ao contexto atual, é seu forte interesse em distinguir os conceitos e suas relações no tesouro da forma mais clara possível. Isso ocorre pois o sucesso da Web Semântica vai depender da atuação de computadores e agentes inteligentes que possam recuperar e manipular informações de diferentes fontes da internet (CLARKE; ZENG, 2012); esse fator também ocasionou outra importante modificação na norma, que é a recomendação de tornar mais explícita a natureza das relações associativas

e de equivalência dos tesouros, a exemplo do que ocorre com as ontologias que especificam a natureza de cada relacionamento em sua estrutura.

A norma sugere, por exemplo, que abreviações e/ou acrônimos podem ser apresentados separadamente dos demais termos equivalentes e, ao invés de utilizarmos a convenção USE e UP (usado por), pode-se utilizar, por exemplo, TC/AB, para termo completo e abreviação, respectivamente. No caso das relações associativas, a norma ressalta que é possível subdividi-la em diferentes tipos. Desta forma, ao invés de simplesmente utilizar a convenção TR (termo relacionado) pode-se utilizar uma convenção que especifique a natureza dessa relação associativa. Por exemplo, no caso de uma relação de causa e efeito, a norma sugere que se utilize como rótulo CAUSA/EFEITO. De acordo com a norma vigente, especificar dessa maneira as relações entre os termos é opcional, tendo em vista o trabalho extra que isso representaria. Porém, se houver o desejo de disponibilizar o tesouro elaborado em rede, visando sua interoperabilidade, a realização dessa atividade extra se torna fundamental.

Acreditamos que, independentemente da interoperabilidade, a especificação da natureza das relações entre os termos equivalentes e associativos deveria ser realizada de forma explícita. O que observamos nos tesouros, de um modo geral, sobretudo nas relações associativas, é que a falta de clareza da relação de fato existente entre os termos associados pode ocasionar a inclusão de termos sem a observação de critérios consistentes para sua escolha.

A norma vigente manteve todos os tipos de relações associativas apresentadas como exemplos representativos de situações típicas que ocorrem na prática. Acrescentou apenas mais um tipo de relação associativa, bem como um exemplo adicional para cada tipo de associação diferente. A relação associativa acrescentada foi a de um objeto e suas partes, caso o mesmo não seja qualificado para pertencer à relação hierárquica do tipo todo-parte. Esse acréscimo não é novidade, considerando que a International Federation of Library Associations and Institutions (IFLA), em 2010, já indicava esta relação associativa. A diferença, no entanto, se deve à explicação apresentada na norma ISO 25964-1 para este fato.

Ao compararmos a norma vigente e o texto prévio de atualização da norma ISO 2788, elaborado por Austin e Dale em 1981(AUSTIN; DALE, 1993) e depois confirmado no texto da norma ISO 2788 de 1986, percebemos que, embora a versão atual continue afirmando que as relações hierárquicas de um tesouro sejam do tipo

gênero e espécie, todo-parte e relação de exemplo, a mesma faz uma importante ressalva quanto à utilização da relação todo-parte. A relação hierárquica do tipo todo-parte abrange um número limitado de situações em que uma parte de uma entidade única ou sistema pertence a um determinado conjunto. É uma relação que continua a ser aplicada a quatro classes principais de termos, são eles: a) sistemas e órgãos do corpo; b) localidades geográficas; c) disciplinas ou áreas do conhecimento e d) estruturas sociais hierarquizadas. A principal diferença entre a norma anterior e a norma vigente, sob esse aspecto, é que a anterior era mais flexível quanto à utilização da relação todo-parte em outras classes de termos, além das quatro principais. A norma vigente sugere o uso apenas das quatro classes principais e desencoraja sua utilização como relação hierárquica em outros casos possíveis do relacionamento todo-parte. De acordo com a ISO 25964-1,

A maioria dos outros casos do relacionamento todo-parte não é elegível para uma ligação hierárquica, porque a parte pode pertencer a mais de um todo. Por exemplo, não deve ser estabelecida uma relação TG / TE entre bicicletas e rodas porque uma roda poderia ser parte de um automóvel, um carrinho de mão ou de muitos outros objetos. (INTERNATIONAL..., p. 60, tradução nossa).

Deste modo, é compreensível a inserção da relação todo-parte como um tipo de relação associativa empregada nos casos em que sua utilização como relação hierárquica não é permitida.

Diante da relevância das relações semânticas entre os termos, parece bastante pertinente a sugestão de maior explicitação das mesmas pela norma ISO 25964. Hjørland (2007) explica que essa demanda por um conjunto mais rico de relações semânticas do que as normalmente utilizadas em tesouros está relacionada com a funcionalidade das mesmas para a recuperação da informação. A maior explicitação das relações semânticas oferece aos bibliotecários mais recursos de representação do conhecimento. Para Hjørland (2007, p. 389), a tarefa mais importante do profissional da informação talvez esteja em “[...] tornar os diferentes interesses e paradigmas visíveis para que o usuário possa fazer uma escolha informada.”

5 ESTUDO DE CASO NO ÂMBITO DO DIREITO DO TRABALHO E DO DIREITO PREVIDENCIÁRIO

Em consonância com nosso objetivo de propor uma metodologia de elaboração de tesouros que considere a linguagem utilizada pelos usuários durante o processo de recuperação da informação, esta pesquisa parte de pressupostos teóricos visando uma aplicação. Sua abordagem reúne aspectos quantitativos, qualitativos e apresenta caráter empírico de análise e observação de dados. Para identificar a terminologia utilizada pelos usuários na recuperação da informação, optamos pela análise das lexias de buscas contidas no arquivo log das pesquisas realizadas pelos usuários do Portal LexMI. Essa escolha se justifica pelo fato de encontrarmos nos logs o registro da interação dos usuários com o sistema de recuperação da informação. Nicholas, Huntington e Watkinson (2005, p. 250) explicam que os logs “[...] representam os usuários, são pegadas de informação digital.” De acordo com os autores, uma das grandes vantagens de trabalharmos com os logs, é que eles fornecem o registro imediato das ações das pessoas, ou seja, são o retrato da realidade dos usuários, pois “Os dados não são filtrados e falam por si [...]” (NICHOLAS; HUNTINGTON; WATKINSON, 2005, p 250). Por essa razão, acreditamos que a análise de logs de pesquisa representa uma nova alternativa de fonte de coleta de termos e descritores para tesouros, pois, desse modo, é possível verificar de forma direta as expressões utilizadas pelos usuários para a recuperação da informação.

No entanto, é importante mencionar que os dados registrados em um arquivo log podem variar de acordo com a configuração utilizada pelo servidor do sistema de recuperação da informação e, também, pela técnica empregada no momento da coleta dos mesmos. No caso do Portal LexML, cada pesquisa realizada pelos usuários gera um registro (log) que informa o número de IP da máquina utilizada para a realização da pesquisa, bem como a data, o horário, o tipo de fonte (Legislação, Jurisprudência, Doutrina, Proposições Legislativas, Publicação Oficial e Outras Manifestações) e as lexias utilizadas pelos usuários em suas buscas.

Diante do exposto, apresentamos, nos tópicos a seguir, os passos seguidos para a seleção e tratamento das lexias de buscas dos usuários. Destacamos que a descrição das etapas metodológicas e da análise dos dados desempenha, neste caso,

uma dupla função: a de descrição da metodologia propriamente dita, e o teste da mesma através de sua aplicação prática.

5.1 Portal LexML

O LexML é um portal especializado em informação jurídica e legislativa, que reúne leis, decretos, acórdãos, súmulas, projetos de leis entre outros documentos das esferas federal, estadual e municipal dos Poderes Executivo, Legislativo e Judiciário de todo o Brasil. De responsabilidade do Senado Federal Brasileiro, seu acervo é o reflexo da cooperação existente entre quatorze bibliotecas jurídicas. O Portal encontra-se disponível para consulta gratuita na internet no seguinte endereço: <www.lexml.gov.br>. O LexML foi escolhido como fonte para a coleta de dados a partir de contato mediado pelo Grupo de Pesquisa Semantec, no âmbito do projeto CNJ Acadêmico da Universidade do Vale do Rio dos Sinos (UNISINOS), coordenado pela Profa. Dra. Rove Chishman. Em videoconferência realizada no final do ano de 2011, os representantes do Portal demonstraram interesse em colaborar com pesquisas acadêmicas e concordaram em disponibilizar os logs de pesquisa dos usuários para análise.

5.2 *Corpus* da Pesquisa

A pesquisa é realizada a partir da análise e comparação de um *corpora* constituído pelos seguintes recursos:

- a) **Corpora de estudo**: constituído por dois conjuntos de lexias de buscas utilizadas pelos usuários do Portal LexML, identificadas nos logs de pesquisa armazenados no servidor do Senado Federal Brasileiro. O primeiro corpus é formado por 217 lexias de busca (98 Lexias verbais, 100 lexias alfanuméricas e 32 lexias conceituais das quais treze também

pertencem a categoria *lexias verbais*) o segundo *corpus* é formado por 385 *lexias* de busca.

b) **Corpus textual especializado e fonte especializada:** o *corpus* textual especializado é constituído por um *corpus* fechado formado pelos textos das seguintes leis: Consolidação das Leis do Trabalho (CLT), instituída pelo Decreto Lei nº 5.452 de 1943; Lei Nº 8.212 de 1991, que dispõe sobre a organização da Seguridade Social; e pela Lei Nº 8.213 de 1991, que dispõe sobre os planos de benefícios da Previdência Social. A fonte especializada, por sua vez, é constituída pelos textos disponíveis na base de dados de jurisprudência do Portal LexML.

c) **Corpus de referência:** constituído pela linguagem documentária utilizada para a indexação de documentos no Portal LexMI. Esse *corpus* é relevante para a pesquisa, pois nos permite verificar em que medida a linguagem empregada pelos usuários para a recuperação de informação corresponde, ou não, à linguagem utilizada para a representação da informação contida nos documentos do Portal. Desse modo, possibilita a identificação pontual dos problemas de comunicação entre usuários e SRI. A linguagem de indexação utilizada em nossa análise foi:

- VCB e AUTR - Vocabulário Controlado Básico e Autorias Padronizadas. Brasília: Senado Federal, Secretaria de Informação e Documentação, [2013]. Disponível em: <http://biblioteca2.senado.gov.br:8991/F/J2P51DT2IKGQT53JCQ9JHPAKYAX76CAC4F5BXB4FIV8Y1P45KS-32194?func=find-b-0&local_base=sen10>. Acesso em: 26 nov. 2013.

Como ferramenta de apoio para o reconhecimento terminológico, foram utilizadas as seguintes obras:

- a) SCHWARZ, R. G. (Org.) **Dicionário de Direito do Trabalho, de Direito Processual do Trabalho e de Direito Previdenciário aplicado ao Direito do Trabalho**. São Paulo: LTR, 2012.
- b) CARVALHO, D. Q. **Classificação Decimal de Direito**. 4.ed. rev. e atual. Brasília: Presidência da República, 2002. Foram utilizadas especificamente

a parte referente ao Direito do Trabalho (ANEXO A) e a parte referente ao Direito Previdenciário (ANEXO B).

Em nossa pesquisa, trabalhamos com um *corpus* textual constituído de acordo com os parâmetros da Linguística de *Corpus*, formado pelas principais Leis do Direito do Trabalho e do Direito Previdenciário, e um conjunto de textos referentes à jurisprudência (fonte especializada) que se encontram disponíveis no Portal LexML. Embora não seja possível chamar esse conjunto de textos de jurisprudência de *corpus*, pois não foram originalmente coletados com o objetivo específico de servir de recurso para a pesquisa linguística, eles atendem às nossas necessidades de pesquisa, visto que, além de disponibilizar o texto original completo desses documentos em formato eletrônico, o Portal também possibilita a verificação da ocorrência ou não das lexias de buscas dos usuários nesses textos, com facilidade e rapidez, informando, inclusive, sua quantidade e o contexto de ocorrência. Por essa razão, julgamos desnecessária a constituição de um *corpus* específico de jurisprudência conforme os requisitos da Linguística de *Corpus*. Tal como entendemos, a base de dados de jurisprudência do Portal LexML mostra-se funcional e suficiente para o estudo pretendido.

5.3 Coleta de Dados

Os Logs foram coletados do servidor do Portal e disponibilizados pelo coordenador do Projeto LexML¹⁰. Ao solicitarmos os dados, fomos informados de que não havia uma política de armazenagem dos mesmos pela instituição. Os logs eram guardados por 30 dias e, depois, apagados do sistema. Em vista disto, a coleta de dados do Portal LexML corresponde ao período de julho a agosto de 2012. A partir da análise do primeiro arquivo de logs, verificamos que 15 dias de coleta, seriam suficientes para atingir nossos objetivos de pesquisa, tendo em vista o volume

¹⁰Senhor João Alberto de Oliveira Lima.

expressivo de dados que recebemos. Especificamente, 276.129 expressões de busca de todas as áreas do Direito, no período de oito de julho até onze de agosto.

É importante mencionar que, embora seja possível identificar o número de endereço IP das máquinas utilizadas para a realização de pesquisas no Portal, a privacidade dos usuários é preservada, pois a coleta realizada não nos permite identificar os dados pessoais dos usuários. Nos logs, visualizamos apenas os dados referentes à interação dos usuários com o sistema de buscas do Portal. Porém, para garantirmos ainda mais a privacidade dos usuários, não divulgaremos, nesta pesquisa, os números corretos de endereço IP das máquinas identificadas nos logs. Optamos, assim, por substituir os primeiros números do endereço IP por letras, conforme os dados do log de pesquisa de um usuário que apresentamos na Figura 3.

Figura 3 – Log de pesquisa

IP	Ano	Mês	Dia	Hora	Minuto	Segundo	País	Lexia de Busca	Tipo de Fonte
xxx.y.241.251	2012	8	11	17	55	4	Brasil	"execução contra herdeiro"	Jurisprudência
xxx.y.241.251	2012	8	11	17	55	15	Brasil	"execução contra herdeiro"	""
xxx.y.241.251	2012	8	11	17	55	19	Brasil	"execução"	""
xxx.y.241.251	2012	8	11	18	25	20	Brasil	"execução trabalhista contra herdeiro"	""
xxx.y.241.251	2012	8	11	18	25	50	Brasil	"execução trabalhista contra executado falecido"	""
xxx.y.241.251	2012	8	11	18	27	55	Brasil	"execução trabalhista de cujus"	""
xxx.y.241.251	2012	8	11	18	29	35	Brasil	"execução trabalhista contra empregador falecido"	""
xxx.y.241.251	2012	8	11	21	18	37	Brasil	"execução trabalhista contra herdeiros"	""
xxx.y.241.251	2012	8	11	21	19	33	Brasil	"execução trabalhista empregador falecido"	""

Fonte: Dados da pesquisa.

5.4 Etapas metodológicas da pesquisa

As etapas metodológicas da pesquisa são as seguintes:

- a) verificar quais são as lexias de buscas utilizadas pelos usuários do Portal LexML nas áreas do Direito do Trabalho e do Direito Previdenciário, para a recuperação de informação (*corpus* de estudo);

- b) identificar variantes denominativas das áreas de conhecimento em análise;
- c) classificar por tipologia as variantes denominativas selecionadas para os dois *corpora* de estudo de lexias;
- d) verificar se as lexias de busca utilizadas pelos usuários do Portal LexML, que constituem o primeiro *corpus* de lexias, ocorrem no *corpus* textual especializado (sincrônico) e na fonte especializada (jurisprudência – diacrônico);
- e) comparar as lexias de buscas utilizadas pelos usuários do Portal LexML, que constituem o primeiro *corpus* de lexias, com os descritores existentes no Vocabulário Controlado Básico (VCB);
- f) comparar os resultados obtidos na classificação por tipologia de variantes denominativas dos dois *corpora* de estudo de lexias;
- g) verificar em três tesouros, dois específicos da área do Direito do Trabalho e um que contempla todas as áreas do Direito, inclusive Direito do Trabalho e Direito Previdenciário, e na norma ISO 25965-1 a ocorrência das tipologias de variantes denominativas identificadas nos dois *corpora* de estudo de lexias;
- h) elaborar proposta metodológica para seleção de termos equivalentes e descritores de tesouros com base nos resultados da pesquisa.

Entendemos que as etapas descritas são necessárias para o alcance dos objetivos propostos, tendo em vista que buscamos uma renovação dos princípios vigentes de seleção e organização de termos e descritores de tesouros. Nesse sentido, destacamos que a etapa **a** é importante para a identificação dos aspectos linguísticos e conceituais da linguagem dos usuários das áreas de estudo que selecionamos para a realização da pesquisa. A etapa **d**, por sua vez, é necessária para o reconhecimento dos contextos de uso dos termos por especialistas. Por essa razão, optamos por trabalhar com um *corpus* e uma fonte especializada, pois o texto das leis, embora vigentes, apresenta uma linguagem estática, semelhante a uma terminologia sincrônica *in vitro*. Os textos da jurisprudência (fonte especializada), por outro lado, apresentam, devido a suas características, uma linguagem em uso e em constante atualização, condizente com uma terminologia diacrônica *in vivo*. A etapa **e**, além de permitir a identificação pontual dos problemas de comunicação entre os usuários e o Portal LexMI, também contribui para comprovação da necessidade de

revisão e proposição de novos princípios de seleção de termos equivalentes e descritores de tesouros (etapa **h**), que visem uma aproximação mais efetiva com o léxico dos usuários dos sistemas de recuperação da informação (etapa **a**) a partir da identificação e classificação de variantes denominativas (etapas **b** e **c** respectivamente). A etapa **f** é importante para o aprofundamento da análise das tipologias de variantes denominativas que ocorrem no âmbito das áreas de conhecimento contempladas nesta pesquisa. A etapa **g**, por sua vez, possibilita identificar quais tipologias de variantes denominativas já são contempladas pelos tesouros da área do Direito do Trabalho e do Direito Previdenciário e pela norma ISO 25965-1 (2011).

No âmbito da elaboração de tesouros, o reconhecimento de variações é particularmente importante para a constituição de um sistema de remissivas que atenda às necessidades dos usuários e facilite a recuperação da informação, e, ao mesmo tempo, possibilite a padronização e o controle terminológico durante o processo de indexação dos documentos (escolha dos descritores mais representativos do conteúdo, do assunto abordado em um documento); visto que, a padronização e o controle terminológico contribuem fortemente para a manutenção da qualidade e consistência de um sistema de recuperação da informação.

5.5 Organização dos dados

Para organização dos dados relacionados à etapa **a**, realizamos os seguintes procedimentos:

- a) Extração dos logs de pesquisa para planilhas do *software* Microsoft Excel, 2013. A extração foi realizada com o auxílio de um extrator de logs elaborado por um profissional da área de informática, especificamente para a realização desta pesquisa. O *software* em questão foi programado para dispor o conjunto de dados relativos à pesquisa de cada usuário do Portal em linhas e colunas. Desta forma, para cada pesquisa é possível identificar o número de IP, a data, o horário, e a lexia de busca utilizada pelo usuário para recuperação da informação.

- b) Após a extração, os logs foram agrupados pelo número de IP, para que pudéssemos fazer uma análise global de todas as lexias de buscas utilizadas pelos usuários no decorrer de sua interação com o Portal.
- c) Foi realizada a listagem das lexias de busca relacionadas às áreas do Direito do Trabalho e do Direito Previdenciário as quais foram identificadas com o auxílio da Classificação Decimal de Direito e da fonte especializada (jurisprudência).
- d) Seleção, entre as lexias identificadas, daquelas que farão parte do primeiro *corpus* de Lexias, de acordo com os seguintes critérios:
- frequência: foram selecionadas as lexias verbais mais frequentes e que apresentaram sinônimos ou algum tipo de variação entre si (APÊNDICE A), tendo em vista que uma das principais razões de fracasso na recuperação da informação se deve aos problemas ocasionados por sinônimos e variantes. As lexias selecionadas ocorreram, pelo menos, quatro vezes na lista de lexias de busca relacionadas às áreas do Direito do Trabalho e do Direito Previdenciário e apresentaram pelo menos mais uma variação.
 - funcionalidade e ocorrência: foram selecionadas lexias que desempenham o papel de termo no âmbito do Direito. É o caso das lexias que convencionamos chamar de lexias alfanuméricas e numéricas, que também podemos chamar de identificadores, visto que designam um conceito único e individual de acordo com Café, Bräscher e Sujii (1990); nesse caso, para a área do Direito, os números de leis, artigos, etc., desempenham o papel de termo. (APÊNDICE B). Tendo em vista que para cada número encontrado no *corpus* foi preciso pesquisar em outra fonte para descobrir se os mesmos pertenciam ou não ao Direito do Trabalho ou Previdenciário, foram selecionadas todas as lexias, dessas duas áreas de interesse, que ocorreram no *corpus* durante o período de coleta de dados. Esses critérios de funcionalidade e ocorrência também foram utilizados para a seleção de lexias com aspectos conceituais ou variantes denominativas com consequências cognitivas, conforme Cabré (2011), (APÊNDICE C). Essas lexias, por suas características,

apresentam indícios de uma busca realizada por um usuário menos especializado que sabe o que quer, mas não domina a terminologia da área. A forma de expressão das lexias conceituais pode ser de dois tipos: a) uma tentativa de definição de um termo; ou b) uma pergunta; nesse caso, destaca-se o fato de serem muito parecidas com as perguntas que os usuários fazem/faziam para o bibliotecário (ver entrevista de referência) quando iam/vão até uma biblioteca para buscar informação.

- e) Classificação de variantes selecionadas para o primeiro *corpus* de lexias verbais de acordo com a Classificação Decimal de Direito (APÊNDICE E).
- f) Constituição de um segundo *corpus* de lexias, formado apenas por lexias de aspectos verbais. Para a seleção dessas lexias, foi considerado somente o critério de ocorrência, a partir da identificação de pares de variantes denominativas, ou seja, a ocorrência de pelo menos duas lexias diferentes para a representação de um mesmo conceito. Esse *corpus* foi constituído especificamente com o objetivo de aprofundar a análise das tipologias de variantes denominativas identificadas a partir da aplicação da Classificação Formal de Variantes Denominativas de Freixa (2002; 2014), motivo pelo qual o mesmo não foi validado no *corpus* textual especializado, na fonte especializada e no Vocabulário Controlado Básico (VCB).

Para a organização dos dados relacionados às etapas **b** e **c**, elaboramos uma planilha para a listagem e anotação da tipologia das variantes denominativas do primeiro *corpus* de lexias verbais (APÊNDICE F). Para o segundo *corpus* de lexias verbais, primeiramente, listamos todas as lexias, depois elaboramos quadros (APÊNDICE G) para a anotação das tipologias de variantes denominativas identificadas; ambos os *corpora* foram classificados de acordo com a Classificação Formal de Variantes Denominativas elaborada por Freixa (2002; 2014). As lexias classificadas de acordo com a proposta de Cabré (2011) foram listadas separadamente e se encontram no Apêndice C.

Para a realização das atividades relacionadas às etapas **d** e **e**, foram necessários os seguintes procedimentos:

- a) elaboração do *corpus* textual especializado com a utilização do *software* de gerenciamento de *corpora*, o Corpógrafo;
- b) elaboração de uma planilha para verificação da ocorrência das lexias de buscas dos usuários no *corpus* textual especializado, na fonte especializada (jurisprudência) e no Vocabulário Controlado Básico (VCB) (APÊNDICE D).

A comparação dos resultados obtidos na classificação por tipologia de variantes denominativas dos *corpora* de estudo de lexias (etapa f) se deu através da análise das planilhas que se encontram nos apêndices F e G. Para a realização da etapa g, também elaboramos uma planilha para anotação das tipologias de variantes denominativas que ocorrem nos tesouros analisados (APÊNDICE H). Os tesouros utilizados para essa análise foram: o Thesaurus de Direito do Trabalho do Tribunal Regional do Trabalho da 4ª Região (TRT4), o Thesaurus em Direito do Trabalho elaborado por Regina de Marco Ferreira (TEJUT) e o Vocabulário Jurídico (Tesouro) do Superior Tribunal de Justiça.

Como a etapa h consiste em desenvolver uma proposta metodológica para seleção de termos equivalentes e descritores de tesouros, a partir da análise dos resultados da pesquisa, não há uma organização de dados específica para sua realização, uma vez que todos os dados utilizados nesta etapa foram organizados nas etapas anteriores.

5.6 Análise dos dados

Para a análise do primeiro *corpus* de lexias no *corpus* textual especializado, optamos por utilizar o recurso de pesquisa KWIC (*Key Word In Context*) do Corpógrafo, a partir do qual é possível verificar a ocorrência ou não das lexias de buscas utilizadas pelos usuários do Portal LexMI nos textos das leis que constituem o *corpus*. Já para a análise das lexias no *corpus* textual de fonte especializada, utilizamos a opção de busca simples do Portal LexMI e filtramos os resultados por jurisprudência. Nos casos em que a pesquisa, devido à natureza da lexia de busca, apresentou como resultado a jurisprudência de vários ramos do Direito, filtramos as

respostas por tipo de autoridade, especificamente pelos Tribunais Regionais de Trabalho e Tribunal Superior do Trabalho.

A validação das *lexias* no VCB foi realizada a partir da verificação da ocorrência ou não das *lexias* de buscas utilizadas pelos usuários do Portal LexMI. O VCB, por ser uma base de dados, está disponível para consulta somente no site da Biblioteca do Senado Federal, razão pela qual é preciso fazer a consulta online de seu conteúdo. Os resultados obtidos foram anotados em uma planilha para posterior comparação e redação da análise dos resultados propriamente ditos.

O Dicionário de Direito do Trabalho, anteriormente mencionado no item 5.2, é utilizado como recurso auxiliar para a identificação de definições de termos e variantes terminológicas identificadas na etapa **b** da pesquisa. Após a obtenção e validação do primeiro *corpus* de *lexias* no *corpus* textual especializado e na fonte especializada (jurisprudência), consultamos o professor Dr. Ricardo Camargo, da Faculdade de Direito da UFRGS, com o objetivo de validar ou não a pertinência das *lexias* de buscas dos usuários que não ocorreram nos *corpora* mencionados. (APÊNDICE I).

A classificação por tipologia das variantes denominativas do primeiro e do segundo *corpus* de *lexias*, nosso *corpora* de estudo, se deu através da análise e comparação de cada *lexia* com as classes elencadas por Freixa (2002; 2014) para a identificação de aspectos formais de variação. Também utilizamos a proposta de variantes denominativas com consequências cognitivas, desenvolvida por Cabré (2011), para a identificação de aspectos conceituais que não são contemplados pela classificação de Freixa (2002; 2014). A partir da identificação das tipologias de variantes denominativas encontradas em nossos *corpora*, verificamos a ocorrência das mesmas em três tesouros da área do Direito.

6 RESULTADOS

Especificamente, em relação à coleta de dados, reiteramos que os logs são uma ótima fonte, tanto para a identificação da linguagem como para a identificação de demandas e necessidades dos usuários de um sistema de informação, visto que, possibilitam a verificação das expressões utilizadas pelos usuários, para a recuperação da informação, de forma direta e sem a interferência do pesquisador. Tendo em vista o grande volume de dados obtidos (276.129 expressões de busca em 32 dias de coleta), optamos por realizar a análise referente a apenas 15 dias de coleta: de 26 de julho a 11 de agosto de 2012. Nesse *corpus*, encontramos informações referentes a 65.536 lexias de buscas de usuários de todas as áreas do Direito. Desse montante, identificamos 2.617 lexias de busca (com ocorrência de lexias repetidas) relacionadas ao Direito do Trabalho e ao Direito Previdenciário dentre as quais selecionamos dois *corpora* de lexias. O primeiro *corpus* é formado por 217 lexias a partir da aplicação dos critérios de frequência de variação, funcionalidade e ocorrência terminológica, conforme explicamos no capítulo anterior (item 5.5 alínea d). O segundo *corpus* é formado por 385 lexias (aspectos verbais) a partir do critério de ocorrência somente, conforme explicamos no capítulo anterior (item 5.5 alínea f).

O exame preliminar do *corpus* de estudo e sua comparação com a classificação de variantes denominativas de Freixa (2002; 2014) nos levou à identificação de três categorias de análise, são elas:

- a) Aspectos verbais;
- b) Recursos não verbais;
- c) Aspectos conceituais.

A categoria aspectos verbais é constituída por todas as tipologias de variantes denominativas elencadas na classificação de Freixa (2002; 2014), com exceção de um subtipo do bloco de alterações gráficas. Optamos por separar em uma nova categoria – que denominamos de recursos não verbais – as variantes relacionadas à alteração gráfica do tipo termo e outra forma artificial, por dois motivos: primeiro, pela quantidade elevada desse tipo de variante em nosso *corpus*; e, segundo, devido à necessidade de um tratamento diferenciado para a identificação de tais variantes, conforme mencionamos no capítulo anterior (tópico 5.5, alínea d). Logo, a categoria recursos não verbais, é formada por lexias de buscas constituídas pela combinação

de números e palavras ou por números puramente, razão pela qual preferimos chamá-las de “lexias alfanuméricas” e “lexias numéricas”, respectivamente. A categoria aspectos conceituais, por outro lado, é constituída por lexias que apresentam partes da definição de um termo ou uma pergunta. Em função das peculiaridades de cada categoria, optamos por apresentá-las separadamente nos tópicos a seguir.

6.1 Aspectos verbais

Conforme mencionamos anteriormente, essa categoria de análise é constituída por variantes denominativas que contemplam aspectos verbais das lexias de buscas empregadas pelos usuários do Portal LexMI. Os resultados obtidos a partir da análise dos dados desta categoria serão apresentados em duas etapas: na primeira etapa, indicamos e comparamos os resultados da validação das lexias de busca do primeiro *corpus* de lexias no *corpus* especializado, na fonte especializada e no VCB; na segunda etapa, descrevemos as tipologias de variantes denominativas identificadas em nosso *corpora* de estudo a partir da aplicação da Classificação Formal de Variantes Denominativas de Freixa (2002; 2014), e verificamos a contemplação ou não dessas tipologias de variantes pela Norma ISO 25964-1 (2011) e por três tesouros jurídicos.

6.1.1 Validação das lexias de busca do primeiro *corpus* de lexias

Observamos que apenas 34 das 98 lexias de busca da categoria aspectos verbais ocorrem no primeiro *corpus* textual especializado (constituído pelas leis) (APÊNDICE J). No Vocabulário Controlado Básico (VCB), identificamos a ocorrência de apenas 28 das 98 lexias de busca (APÊNDICE K). Porém, na fonte especializada (constituída por textos de jurisprudência), identificamos a ocorrência de 76, das 98 lexias de busca (APÊNDICE L).

Acreditamos que a diferença significativa da ocorrência das lexias entre a fonte especializada (jurisprudência), o *corpus* especializado e o VCB, ocorre por diferentes

motivos. No caso do *corpus* especializado, essa diferença se deve ao caráter sincrônico das leis; já em relação ao VCB, deve-se à necessidade de atualização do mesmo, visto que o número de lexias encontradas no VCB foi ainda menor do que o número de lexias identificadas no *corpus* especializado (28 lexias no VCB; 34 lexias no *corpus* especializado). A validação de 76 lexias na fonte especializada, também reforça a necessidade de atualização do VCB em função do uso comprovado e registrado dessas lexias nos documentos de jurisprudência disponíveis no Portal LexMI. Essa validação da ocorrência das lexias de busca do primeiro *corpus* de estudo pode ser verificada no Apêndice D.

Consideramos que vinte e duas (22) das 98 lexias do *corpus* de estudo não foram encontradas/validadas na fonte especializada (jurisprudência), no *corpus* especializado (leis) e no VCB, visto que não encontramos sua ocorrência exatamente da mesma forma que a utilizada pelos usuários do Portal LexMI nos *corpora* da pesquisa. Por esse motivo, as mesmas lexias foram encaminhadas para a análise de um especialista na área do Direito do Trabalho¹¹, para que pudéssemos identificar a razão pela qual tais lexias, apesar de serem utilizadas pelos usuários do Portal LexMI, não ocorrem nas fontes utilizadas para validação dos dados desta pesquisa (APÊNDICE I). O especialista utilizou o corpo dos acórdãos do Tribunal Superior do Trabalho (TST) como fonte de referência para análise dessas lexias. Como resultado, seis lexias foram validadas como termo, visto que ocorrem nos acórdãos do TST, são elas: certidão negativa trabalhista; execução trabalhista de cujus; acidente trabalho; direito trabalho; doenças ocupacionais; hrs extras. Dezesesseis lexias não foram validadas. A análise aprofundada de cada lexia nos leva a crer que os motivos pelos quais elas não podem ser consideradas termo são:

- a) a lexia utilizada pelo usuário representa dois conceitos diferentes – duas lexias (férias dobra; Fundo de garantia e tempo de serviço);
- b) a lexia utilizada pelo usuário apresenta o uso concomitante de duas variantes – uma lexia (gratificação férias abono);
- c) a lexia utilizada pelo usuário apresenta o uso concomitante de sigla e termo na mesma expressão – uma lexia (cnd trabalhista);
- d) a lexia utilizada pelo usuário é uma paráfrase – cinco lexias (execução

¹¹ Professor Ricardo Camargo da Faculdade de Direito da UFRGS.

- trabalhista contra empregador falecido; execução trabalhista contra executado falecido; execução trabalhista contra herdeiro; execução trabalhista contra herdeiros; Dispensa indireta por ato lesivo a honra);
- e) a lexia utilizada pelo usuário representa dois termos e um qualificador (tipo de) – duas lexias (vínculo diarista rural; vínculo eventual rural);
 - f) a lexia utilizada pelo usuário representa dois termos e um especificador de lugar – duas lexias (vínculo diarista fazenda; vínculo eventual fazenda);
 - g) a lexia utilizada pelo usuário apresenta erro de grafia – uma lexia (Trabalador doméstico);
 - h) a lexia utilizada pelo usuário é um neologismo – duas lexias (vínculo empregatista ; carreteiro intermunicipal).

Os resultados obtidos nesta primeira etapa da análise indicam que:

- a) em relação ao *corpus* textual especializado, existe um alto índice de utilização de variantes denominativas pelos usuários dos domínios estudados. A diferença do resultado obtido na análise do *corpus* textual em relação ao resultado obtido na análise da fonte especializada comprova esse fato.
- b) em relação ao VCB, poucos descritores estão de acordo com o léxico dos usuários, o que sugere a necessidade de atualização das remissivas do VCB.
- c) em relação à metodologia utilizada na primeira etapa da análise, os resultados obtidos indicam que essa metodologia possibilita a identificação de termos tanto para a elaboração como para a atualização de linguagens documentárias, como os tesouros. O recorte do *corpus* de lexias de estudo, considerando o critério de frequência, possibilita a identificação dos casos de variação mais frequentes, no entanto, não permite a identificação de todas as tipologias de variação que podem ocorrer em um domínio. Para a identificação de tipologias de variação, que podem ocorrer em uma área de conhecimento, o recorte por frequência não é suficiente. Nesse caso, é preciso ampliar o *corpus* e considerar como critério de seleção a ocorrência de pelo menos duas lexias para a representação de um conceito. A observação desses aspectos nos levou a constituir o segundo *corpus* de

lexias de estudo, cujos resultados são apresentados no tópico a seguir (6.1.2).

6.1.2 Descrição das tipologias de variantes denominativas do primeiro e do segundo *corpus* de lexias

A partir da aplicação da Classificação Formal de Variantes Denominativas de Freixa (2002; 2014) identificamos quais as tipologias de variantes denominativas ocorrem em nossos *corpora* de estudo. Lembramos que os resultados apresentados neste item dizem respeito apenas às lexias da categoria aspectos verbais do primeiro e do segundo *corpus* de estudo de lexias, que correspondem a 98 (APENDICE F) e 385 lexias (APENDICE G), respectivamente. Destacamos, ainda, que o segundo *corpus* de lexias foi constituído com o objetivo específico de aprofundar a análise das tipologias de variantes denominativas identificadas no primeiro *corpus* de estudo de lexias. E, por esse motivo, os critérios de seleção das lexias de busca que compõem o segundo *corpus* de estudo de lexias são diferentes dos critérios utilizados para a formação do primeiro *corpus*, conforme mencionamos anteriormente no item 5.5 (alínea f) desta tese. É importante destacar, ainda, que optamos pela apresentação dos resultados por número de casos. Cada caso se refere a um par de lexias, visto que a análise da variação e sua classificação por tipologia só é possível a partir da comparação de uma lexia em relação a outra lexia.

A classificação de Freixa (2002; 2014) está organizada em cinco categorias, são elas: alterações gráficas, alterações morfossintáticas, reduções, alterações léxicas e várias alterações complexas. Cada categoria dessa classificação se subdivide em pelo menos dois tipos de variantes denominativas, por isso e por uma questão de organização textual, optamos pela apresentação dos resultados de cada categoria separadamente.

A categoria de alterações gráficas apresenta três tipos e nove subtipos de variação. O grupo designado por termo e forma artificial não será apresentado neste item, pois ele diz respeito às lexias da categoria recursos não verbais, que serão discutidas no item 7.2 desta tese. O grupo designado por termo e abreviação se subdivide em termo e sigla, e termo e abreviação; e o grupo chamado de alteração

ortográfica se subdivide em presença e ausência de acentuação, presença e ausência de hífen, erros de grafia (como o acréscimo ou supressão de letras e presença ou ausência de espaço entre as palavras).

Para o primeiro *corpus* de lexias, essa foi a categoria na qual encontramos mais casos de variação – especificamente dezoito casos – em um *corpus* constituído por 98 lexias. Já para o segundo *corpus* de lexias, essa categoria foi a segunda com mais casos, especificamente 41 casos, em um *corpus* com 385 lexias.

Quadro 15 – Alterações gráficas: resultados

Tipo	Subtipo	Primeiro Corpus	Segundo Corpus	Exemplos
Termo e abreviação	d) Termo e sigla	5 casos	12 casos	Consolidação das Leis do Trabalho / CLT
	e) Termo e abreviação	2 casos	2 casos	Hrs. extras
Alteração ortográfica	- presença e ausência de acentuação	2 casos	12 casos	Férias / Ferias
	- presença e ausência de hífen	8 casos	10 caos	Auxílio-alimentação / Auxílio alimentação
	- erros de grafia como o acréscimo ou supressão de letras	1 caso	4 casos	Trabalhador doméstico / Trabalhador doméstico
	- presença ou ausência de espaço entre as palavras	Nenhum caso.	1 caso	Contracheque / Contra Cheque

Fonte: Elaborado pela autora.

Ao realizarmos a verificação das tipologias de variação, que ocorrem nos três tesouros que utilizamos como parâmetro para análise dos dados, observamos a ocorrência de apenas duas dessas tipologias; são elas: termo e sigla e presença e ausência de hífen. O uso de termos e siglas se confirma nos três tesouros e são recorrentes. A aplicação do critério de frequência de uso para a seleção das lexias que constituem o primeiro *corpus* de lexias nos possibilitou identificar que o uso de siglas, pelos usuários do Portal LexMI nas áreas contempladas nesta pesquisa, é mais frequente do que o uso da forma completa do termo. A sigla CLT, por exemplo, ocorre 85 vezes, enquanto suas duas variantes na forma completa (Consolidação das Leis do Trabalho e Consolidação das Leis Trabalhistas) ocorrem apenas cinco vezes cada uma. O mesmo ocorre com as demais siglas identificadas no *corpus*; nesse sentido, observamos que a frequência de uso das siglas sempre é maior do que o uso dos respectivos termos na forma completa.

A apresentação de termos equivalentes, em função da presença e ausência de hífen, ocorre em apenas um dos tesouros observados, o do STJ. As demais tipologias

de variação não ocorrem nos tesouros observados. Considerando a frequência de utilização dessa tipologia de variantes pelos usuários do Portal LexMI, julgamos, de fato, pertinente a inclusão desse tipo de variação em tesouros, visto que a frequência de utilização da forma sem hífen dessas lexias é geralmente maior do que o uso da forma com hífen, embora em nosso *corpus* a forma gramaticalmente correta dessas lexias seja com hífen. A variante “seguro desemprego” ocorre oito vezes em nosso *corpus*, ao passo que a lexia “seguro-desemprego” ocorre apenas uma vez; o mesmo acontece com a lexia “licença maternidade” que ocorre dez vezes em nosso *corpus* enquanto sua variante “licença-maternidade” ocorre apenas uma vez. A norma de elaboração de tesouros ISO 25964-1, prevê a utilização de apenas duas dessas tipologias de variantes mencionadas para a constituição das relações de equivalência de tesouros, que é a relação termo e sigla e uso de hífen.

Na categoria de alterações morfossintáticas, a tipologia de variantes que se destaca, tanto no primeiro como no segundo *corpus* de lexias, é a que prevê alteração do substantivo com manutenção da estrutura, conforme se pode observar no Quadro 16, a seguir.

Quadro 16 – Alterações morfossintáticas: resultados

Tipo	Subtipo	Primeiro Corpus	Segundo Corpus	Exemplos
Manutenção da estrutura	a) ausência /presença de artigo	1 caso	2 casos	Acidente de trabalho / acidente do trabalho;
	b) alteração do substantivo	5 casos	12 casos	Execução trabalhista contra herdeiro / Execução trabalhista contra herdeiros;
	c) alteração de preposição	Nenhum caso	1 caso	Acidente em Serviço / Acidente de serviço;
	d) alteração de gênero	1 caso	2 casos	Empregada doméstica / empregado doméstico.
Alteração da estrutura	e) [Substantivo + Adjetivo] / [Substantivo + Sintagma preposicional]	1 caso	1 caso	Vínculo empregatício / vínculo de emprego;
	f) Monoléxico / Poliléxico	Nenhum caso	Nenhum caso	
	g) Outras alterações de estrutura.	Nenhum caso	3 casos	Trabalho doméstico infantil / Trabalho infantil doméstico.

Fonte: Elaborado pela autora.

Na verificação das tipologias de variação, que ocorrem nos três tesouros que utilizamos como parâmetro para análise dos dados, observamos a ocorrência de apenas duas dessas tipologias em apenas um dos tesouros observados, nesse caso o do STJ. As tipologias encontradas no tesouro do STJ são: ausência/presença de

artigo e alteração de preposição. Ressaltamos que, embora a alteração de preposição não cause alteração de estrutura, a troca de uma preposição por outra pode alterar o sentido do termo. A troca da preposição em pela preposição de na lexia “acidente em serviço”, que é equivalente ao termo “acidente no trabalho”, altera o sentido do termo, visto que a lexia “acidente de serviço” é equivalente ao termo “acidente de trabalho”. Do ponto de vista jurídico o acidente de trabalho não ocorre necessariamente no ambiente de trabalho, ele pode se dar no percurso. O acidente de percurso é um tipo de acidente de trabalho, portanto os termos acidente de trabalho e acidente no trabalho não são equivalentes entre si, são termos relacionados.

Em nossos *corpora* de lexias a ocorrência dessas tipologias de variantes foi bem pequena, como se pode observar no Quadro 16. A norma de elaboração de tesouros ISO 25964-1 (2011), não prevê a utilização de nenhuma dessas tipologias de variantes para a constituição das relações de equivalência de tesouros.

A categoria de variantes denominativas decorrentes de reduções se subdivide em duas categoria; são elas: redução de extensão e redução de base. No primeiro *corpus* de lexias, encontramos poucos casos de variação por redução de extensão. Por outro lado, no segundo *corpus* de lexias, a quantidade de casos de redução por extensão é bastante significativa, concentrando um número elevado de casos. Ao todo, foram identificados 93 casos de variação por redução de extensão no segundo *corpus* de lexias e oito casos (8) no primeiro *corpus*, como se pode observar no Quadro 17, a seguir.

Quadro 17 – Reduções de extensão: resultados

Subtipo	Primeiro Corpus	Segundo Corpus	Exemplos
a) Tipo [Substantivo + X] = [Substantivo].			
- [Substantivo + Adjetivo]	2 casos	24 casos	Aviso prévio / Aviso;
- [Substantivo + Sintagma preposicional]	1 caso	15 casos	Acidente de trabalho / Acidente;
-[Substantivo 1+ Substantivo 2] = Substantivo	Nenhum caso	8 casos	Gratificação Assiduidade / Gratificação.
b) [Substantivo + X + Y] = [Substantivo + X].			
- [Substantivo + Sintagma preposicional + Adjetivo] = [Substantivo + Sintagma preposicional]	Nenhum caso	3 casos	Equipamento de proteção individual / Equipamento de proteção;
- [Substantivo + Adjetivo + Sintagma preposicional] = [Substantivo + Adjetivo]	1 caso	3 casos	Certidão negativa de débitos / Certidão negativa
- [Substantivo + Adjetivo 1 + Adjetivo 2] = [Substantivo + Adjetivo 1]	Nenhum caso	6 casos	Certidão negativa trabalhista / Certidão negativa;
- [Substantivo + Sintagma preposicional 1 + Sintagma preposicional 2] = [Substantivo + Sintagma preposicional 1]	2 casos	4 casos	Adicional de periculosidade de eletricitista / Adicional de periculosidade;
c) Outras reduções de extensão	2 casos	30 casos	Tabela de lotação de cargos / Tabela.

Fonte: Elaborado pela autora.

As variantes denominativas do tipo redução de base ocorrem em pequena quantidade no primeiro *corpus* de lexias – apenas seis casos –, já no segundo *corpus* identificamos 38, conforme se pode observar no Quadro 18, a seguir.

Quadro 18 – Reduções de base: resultados

Subtipo	Primeiro Corpus	Segundo Corpus	Exemplos
d) [Substantivo + Adjetivo] = [Adjetivo] N	1 caso	5 casos	Empregada doméstica / doméstica
e) [Substantivo 1 + Substantivo 2] = [Substantivo 2]	2 casos	2 casos	Adicional insalubridade / Insalubridade
f) [Substantivo 1 + Sintagma preposicional (de + Substantivo 2)] = [Substantivo 2]	3 casos	3 casos	Pagamento de diárias / Diárias.
g) Outras reduções	Nenhum caso	28 casos	Trabalho doméstico infantil / Trabalho infantil.

Fonte: Elaborado pela autora.

Ao analisarmos os três tesouros (STJ, TRT4 e TEJUT), observamos que as reduções de extensão são contempladas nessas linguagens em apenas dois deles: no tesouro do TRT4 e no tesouro do TEJUT. A norma de elaboração de tesouros 25964-1 (2011) fala da possibilidade de se generalizar alguns termos, conforme o nível de especialização dos usuários da instituição para a qual o tesouro será construído. O argumento apresentado na norma é que, se o tesouro a ser elaborado é voltado para um público menos especializado, o uso de termos muito específicos talvez não seja necessário e cabe ao profissional que estiver elaborando essa linguagem decidir se vai ou não usar termos mais específicos ou mais genéricos. Essa flexibilização da norma não nos parece a melhor opção, pois acreditamos que o sentido dos termos sempre deve ser preservado e a troca de um termo específico por um mais genérico implica, necessariamente, em perda de sentido. É relevante mencionar ainda que, mesmo apresentando essa possibilidade de generalização de alguns termos, a norma não coloca o seu uso como uma alternativa nas relações de equivalência.

Destacamos que, ainda que os usuários utilizem variantes reduzidas em suas buscas, a inserção das mesmas como termos equivalentes em um tesouro deve ser analisada com cuidado, visto que, para fins de recuperação da informação, é imprescindível que as variantes utilizadas como equivalentes na rede de remissivas preservem o sentido do termo original. Nesse sentido, quando um termo perde sua especificidade em função de uma redução e o termo resultante apresenta uma mudança de sentido, passamos a ter dois termos diferentes e, não, dois termos equivalentes. O uso de termos complexos, constituídos por duas, três ou até mais palavras, é uma característica do léxico especializado, assim como o uso de reduções desses termos em textos especializados, inclusive por uma questão de economia linguística. Contudo, para fins de recuperação de informação em uma base de dados, devemos priorizar e manter o princípio da especificidade, ou seja, representar o assunto de um documento sempre pelo termo mais específico para evitar ambiguidades e poupar o tempo do usuário em suas pesquisas. É preciso lembrar que, no momento da recuperação da informação, trabalhamos com termos isolados e, por isso, eles devem significar por si só, uma vez que não temos o contexto discursivo de um texto completo para eliminar a ambiguidade causada por termos reduzidos.

Percebemos que há casos em que os termos, mesmo reduzidos, preservam o sentido do termo original, como ocorre com o termo “Carteira de Trabalho”, que apresenta uma redução de extensão do termo “Carteira de Trabalho e Previdência Social” e como ocorre com o termo “Insalubridade” que apresenta uma redução de base do termo “Adicional de Insalubridade”. Por isso, não podemos generalizar dizendo que todos os termos reduzidos sempre apresentam uma perda significativa de sentido em relação ao termo original. Em decorrência disto, recomendamos uma análise cuidadosa das reduções. Consideramos que um termo reduzido pode efetivamente ser inserido como um termo equivalente em um tesouro, desde que o mesmo não apresente uma perda tão significativa de sentido que o leve a se tornar outro termo como ocorre com o termo “Licença”, por exemplo. No âmbito do Direito do Trabalho, existem diferentes tipos de licença, tais como: “licença-interesse” “licença-saúde” “licença-maternidade”. Nesse caso, o termo “Licença” é muito genérico e ambíguo devido à perda da especificidade do termo original, seu uso isolado (reduzido) não pode ser inserido como um termo equivalente em um tesouro. O termo “Licença” isolado (reduzido) só pode ser utilizado em um tesouro como um termo relacionado ou um termo genérico.

A categoria alterações léxicas subdivide-se em duas tipologias, são elas: alterações léxicas de unidades monoléticas e alterações léxicas de unidades poliléticas. Em alterações de unidades monoléticas, são classificados os casos que apresentam alteração lexical entre duas formas monoléticas. A ocorrência desse tipo de variação em nossos *corpora* foi pequena. No primeiro *corpus* de lexias, identificamos apenas um caso; já no segundo *corpus* de lexias foram identificados quatro casos. O par de lexias encontradas no primeiro *corpus* foi Caminhoneiro / Carreteiro; no segundo *corpus*, Caminhoneiro / Carreteiro; Exoneração / Rescisão; Falecimento / Morte; Remuneração / Salário.

Em alterações de unidades poliléticas, são classificados os casos que apresentam alteração lexical entre duas formas poliléticas, as quais se subdividem de acordo com a parte do sintagma em que a mudança ocorre (base ou extensão) e conforme o tipo de estrutura morfossintática do sintagma. Em nossos *corpora*, a quantidade de casos identificados dessa tipologia de variação foi mais significativa: no primeiro *corpus* de lexias identificamos treze casos e, no segundo *corpus* de lexias, 29 casos, conforme se pode observar no Quadro 19 a seguir.

Quadro 19 – Alterações Léxicas de unidades poliléxicas: resultados

Subtipo	Primeiro Corpus	Segundo Corpus	Exemplos
a) Alteração de base			
- [Substantivo + Sintagma Preposicional (prep. + art. + Substantivo)]	Nenhum caso	Nenhum caso	
- [Substantivo + Adjetivo]	3 casos	4 casos	Empregado doméstico / Trabalhador doméstico;
- [Outras alterações de base]	Nenhum caso	1 caso	Competência da Justiça do Trabalho / Responsabilidade da Justiça do Trabalho;
b) Alteração de extensão			
- [Substantivo + Sintagma Preposicional (prep. + art. + Substantivo)]	2 casos	5 casos	Acidente de trabalho/ Acidente de serviço;
- [Substantivo + Adjetivo]	3 casos	7 casos	Abono salarial/ abono pecuniário;
- [Outras alterações de extensão]	5 casos	12 casos	Consolidação das leis do trabalho / Consolidação das leis trabalhistas;

Fonte: Elaborado pela autora.

Observamos a ocorrência de quase todas as tipologias de variação dessa categoria nos três tesouros analisados; a única tipologia que não identificamos foi a alteração de base [substantivo + sintagma preposicional (prep. + art. + substantivo)]. Acreditamos que uma das razões para a ocorrência dessas tipologias de variantes nos três tesouros analisados se deve à previsão das mesmas pelas normas de elaboração de tesouros anteriores e vigente como a norma ISO 25964-1 (2011). De acordo com Freixa (2002; 2014), a categoria de alterações léxicas é a que está mais próxima da discussão sobre sinônimos. A norma ISO 25964-1 (2011) explica quatro situações gerais para o estabelecimento das relações de equivalência, são elas: a) entre termos sinônimos; b) entre termos quase-sinônimos; c) quando um termo é considerado desnecessariamente específico e é representado por uma combinação de dois ou mais termos (conhecido como composto de equivalente); e d) quando um termo é considerado desnecessariamente específico e é representado por outro termo com escopo próximo. As variantes léxicas de unidades monoléxicas e de unidades poliléxicas identificadas em nossos *corpora* e nos tesouros analisados correspondem aos casos de sinônimos e quase-sinônimos previstos pelas normas de elaboração de tesouros.

A categoria que engloba várias alterações complexas se divide em duas

classes: com parentesco formal e sem parentesco formal; em que se dividem, ainda, em monoléxica/poliléxica e poliléxica/poliléxica. Nessa categoria, os termos são classificados a partir da verificação da manutenção ou não do parentesco formal entre as variantes analisadas e a alternância entre uma unidade monoléxica e uma unidade poliléxica ou entre duas unidades poliléxas. No primeiro *corpus* de lexias, identificamos apenas nove casos de alterações complexas; já no segundo *corpus* de lexias, identificamos 33 casos, com destaque para as alterações complexas com parentesco formal entre unidades poliléxicas, conforme se pode observar no Quadro 20, a seguir.

Quadro 20 – Várias alterações complexas: resultados

Tipo	Subtipo	Primeiro Corpus	Segundo Corpus	Exemplos
Com parentesco formal	a) Monoléxica/Poliléxica	1 caso	3 casos	Pessoa com deficiência/ Deficiente.
	b) Poliléxica/Poliléxica	2 casos	19 casos	Doenças ocupacionais / Doença profissional;
Sem parentesco formal	c) Monoléxica/Poliléxica	2 casos	5 casos	Sobreaviso / Tempo à disposição;
	d) Poliléxica/Poliléxica	4 casos	6 casos	Seguro invalidez / Auxílio doença.

Fonte: Elaborado pela autora.

Nos três tesouros analisados, identificamos a ocorrência de todas as tipologias de variação elencadas nessa categoria. Acreditamos que a ocorrência dessas tipologias de variação nos tesouros analisados, se deve à previsão feita pelas normas de elaboração de tesouros – anteriores e vigente – sobre a inserção de sinônimos e quase-sinônimos como termos equivalentes.

É importante mencionar que não foi possível classificar, de acordo com a classificação de Freixa (2002; 2014), alguns pares de lexias (vinte casos). Isso ocorreu, pois a Classificação de Freixa (2002; 2014) não contempla em suas categorias esses casos que identificamos em nossos *corpora*. Em vista disto, criamos duas novas categorias para a organização desses casos. A primeira categoria que criamos foi a de **alterações gráficas por supressão de preposição**. Nesse caso, percebemos que os usuários suprimem o uso das preposições das lexias para a

realização de suas buscas, por exemplo: acidente de trabalho / acidente trabalho. Para essa categoria, identificamos lexias tanto do primeiro como do segundo *corpus* de lexias, totalizando dez casos. Não identificamos a utilização desse tipo de variação nos três tesouros analisados, nem a menção das mesmas pelas normas de elaboração de tesouros antigas e vigentes. As lexias pertencentes a essa categoria se encontram listadas no Apêndice M.

A segunda categoria que criamos foi a de **busca por campo semântico**; nesses casos, há uma proximidade de sentido entre os termos, embora não se possa considerá-los equivalentes (variantes denominativas); são, portanto, termos relacionados, visto que apresentam sentidos diferentes. Essa categoria se subdivide em três tipos: **troca de fenômeno por pessoa**; **troca de fenômeno por objeto e termos antônimos (oposição)**. Identificamos na literatura sobre elaboração de tesouros e na norma ISO 25964-1 (2011) apenas a relação associativa referente a troca de fenômeno por objeto, denominada pela IFLA (2010) como ação / produto da ação e a utilização de termos antônimos, denominada pela ISO 25964-1 (2011) como relação associativa por oposição. Por outro lado, nos três tesouros analisados, identificamos a ocorrência de todas as tipologias de variação elencadas nessa categoria. No Quadro 21, a seguir, apresentamos um exemplo de lexias de cada subtipo, mas a lista completa das lexias classificadas nessa classe pode ser conferida no Apêndice N.

Quadro 21 – Busca por campo semântico

Tipo	Subtipo	Corpora	Exemplos
Busca por campo semântico	a) Troca de fenômeno por pessoa	6 casos	Estabilidade gravidez / Estabilidade gestante
	b) Troca de fenômeno por objeto	2 caso	Pré-contratação / Pré-contrato
	c) Termos Antônimos (oposição)	2 casos	Sentar no trabalho / Trabalhar em pé

Fonte: Elaborado pela autora.

6.2 Recursos não verbais

A categoria recursos não verbais, conforme mencionamos anteriormente, é constituída por número de leis, número de súmulas, número de artigos de leis, número de projetos de lei, etc. Das 2.617 lexias de busca identificadas, cem são alfanuméricas ou numéricas (Apêndice B). Em muitos casos, os usuários utilizam apenas um número isolado como estratégia para recuperação da informação (exemplo: “12401 2011”), fato que torna a análise dos dados mais morosa, tendo em vista que, a cada número identificado no *corpus*, temos que buscar informações em outra fonte (utilizamos o próprio Portal LexMI), para que possamos saber se o mesmo está relacionado com a área do Direito do Trabalho ou Direito Previdenciário. No exemplo apresentado, os números correspondem ao número e ano da Lei que altera a Lei n.º 8.080, de 19 de setembro de 1990, para dispor sobre a assistência terapêutica e a incorporação de tecnologia em saúde no âmbito do Sistema Único de Saúde – SUS.

As lexias de recursos não verbais de nosso *corpus* correspondem ao subitem termo e a outra forma artificial da categoria alterações gráficas, na classificação de variantes denominativas de Freixa (2002; 2014). Não identificamos a ocorrência desse tipo de lexia nos três tesouros que utilizamos como parâmetro para a análise das tipologias de variantes que ocorrem no âmbito do Direito do Trabalho e do Direito Previdenciário. Também não identificamos a previsão de inserção dessa tipologia de variante como termo equivalente em tesouros pelas normas de elaboração de tesouros anteriores e vigente. A norma ISO 25964-1 (2011) indica a pertinência da inserção de termos que representam entidades únicas (identificadores) em tesouros, porém não especifica se os mesmos devem ser inseridos como termos equivalentes ou descritores.

A partir da organização e seleção das lexias de buscas dos usuários do Portal LexMI, percebemos uma significativa utilização de lexias alfanuméricas como estratégia para a recuperação de informação, ou seja, o uso do número de leis, artigos, decretos etc. acompanhados de data, como recurso de representação e expressão de necessidade de informação. No entanto, não identificamos a utilização desses itens como descritores nos registros bibliográficos das obras do Portal, ou seja, como as lexias numéricas e alfanuméricas não são utilizadas como descritores, não constam no campo assunto dos registros. Verificamos que o VCB destaca a possibilidade de

utilização do nome de uma lei como descritor de assunto e, nesses casos o número da lei é considerado um termo equivalente. No entanto, nos registros bibliográficos do portal o uso do nome de leis como descritor de assunto é muito pouco utilizado, buscas específicas pelo nome de leis retornam poucos registros como resultado. Geralmente, o único campo do registro bibliográfico em que é possível encontrar o número de uma lei, por exemplo, é no título – embora o número das leis em discussão numa obra nem sempre faça parte do seu título. No Direito, existem muitas obras explicativas, cujo assunto é, justamente, a discussão de uma determinada lei, decreto, etc. Logo, os usuários buscam por uma informação bastante específica. Querem saber, por exemplo, qual é a interpretação de uma determinada lei, ou artigo de lei, feita por um juiz ou tribunal.

Esse tipo de informação é muito recorrente e existe em grande quantidade na jurisprudência; conseqüentemente, há um volume bastante expressivo de documentos que se encaixam nesse perfil. Porém, como as lexias numéricas e alfanuméricas não são utilizadas como descritores, não é possível filtrar a busca para a obtenção de resultados mais específicos. Esse fato demonstra que é preciso acrescentar esses tipos de lexias nos tesouros, pelo menos como termos equivalentes na rede de remissivas, de modo que o usuário possa realizar uma busca mais específica e obter resultados também específicos. Nossa sugestão, nesse caso, seria a utilização do nome das leis como descritores e o número das mesmas como equivalentes.

Esse entendimento das leis como assunto de uma obra é algo para ser discutido também no âmbito de trabalhos sobre políticas de indexação. A partir de uma observação de catálogos de bibliotecas jurídicas e do próprio LexMI, percebemos que o nome e o número das leis nunca ou raramente são utilizados como descritores. Isso nos leva a crer que entre os bibliotecários da área jurídica não há a cultura de utilização desse item como recurso de representação do conteúdo de documentos. Muitos podem argumentar que não haveria essa necessidade e que, ao representarem esse tipo de obra através dos descritores que representam o assunto das leis, já estariam sendo suficientemente específicos em relação ao conteúdo da obra. Acreditamos, porém, que nesse tipo específico de obra, onde há a discussão de uma determinada lei, é necessário acrescentar o nome da lei como descritor juntamente com os demais descritores tradicionalmente utilizados. Por exemplo, uma obra cujo conteúdo discute a aplicação da lei que regulamenta a profissão de motorista, especificamente em relação à jornada de trabalho e ao período de descanso dessa categoria de

profissionais, não poderia ser representada apenas pelo descritor “jornada de trabalho”, pois, no âmbito do Direito do Trabalho, essa seria uma representação muito genérica. Uma representação mais específica da mesma obra deveria apresentar em seu registro bibliográfico os seguintes descritores: um descritor genérico (também chamado de macro descritor), nesse caso “Direito do Trabalho”; e descritores específicos tais como: “motorista profissional”; “jornada de trabalho”; “Lei 12.619/2012”; “Artigo 67 A. da Lei 12.619/2012”.

Na Biblioteconomia, defende-se o princípio da especificidade para a representação do conteúdo das obras, ou seja, devemos representar os assuntos mais específicos das obras através da utilização de descritores também específicos. No caso do Direito, respeitar e atender o princípio da especificidade implica sobremaneira em passarmos a utilizar as lexias alfanuméricas e numéricas, pelo menos, como termos equivalentes nos tesauros.

O estudo realizado sobre o uso das variantes denominativas mostrou que a classificação de Freixa (2002; 2014) possibilita a identificação e classificação de aspectos formais de variação, mas não contempla aspectos conceituais que foram identificados em algumas lexias que constituem nosso *corpus* de estudo. Por esse motivo, foi necessário criar a categoria aspectos conceituais, na qual encontramos lexias constituídas por partes da definição de um termo ou uma pergunta, conforme apresentaremos no item a seguir.

6.3 Aspectos conceituais

A categoria aspectos conceituais foi criada, por duas razões: a primeira delas surge da identificação, em alguns casos, do conhecimento por parte dos usuários das definições dos conceitos que buscam, em detrimento do termo normalmente utilizado no âmbito do Direito do Trabalho e do Direito Previdenciário para designá-los; e a segunda razão depreende da não identificação de uma categoria em que se pudesse classificar essa tipologia de variantes na Classificação Formal de Variantes Denominativas, de Freixa (2002; 2014). Isso mostra a importância da dimensão

conceitual das pesquisas e a necessidade de também considerarmos esses aspectos no momento da representação dos assuntos identificados em uma obra (indexação).

A quantidade de lexias identificadas nessa categoria não foi muito expressiva no *corpus* analisado – somente 32 (APÊNDICE C) –, no entanto, devido à relevância da dimensão conceitual na recuperação da informação, optamos por apresentá-las em uma categoria própria. A busca por uma parte da definição de um conceito, ou a formulação de uma pergunta como expressão de busca é também um indício sobre o nível de conhecimento do usuário em relação ao assunto pesquisado. Sua utilização nos permite inferir se o usuário que a emprega, como parte da sua estratégia de busca em um sistema de informação, é ou não um especialista das áreas do Direito pesquisadas. Trata-se de uma característica de pesquisa dos usuários, importante de ser destacada, pois é possível identificar um usuário semi-leigo, bem como demonstra a necessidade de pensarmos no tipo de definição a ser elaborada para os termos do glossário, que consideramos como parte fundamental dos tesouros.

O usuário sabe o que quer, mas não sabe expressar de forma objetiva a sua necessidade de informação, em outras palavras, sua forma de expressão ainda não é reconhecida como um descritor pela linguagem documentária utilizada pelo SRI. Um exemplo dessa situação é a lexia “auxílio-alimentação” que, embora seja um tipo de “salário utilidade”, não se encontra como descritor na linguagem documentária utilizada pelo Portal LexMI.

Outro exemplo que se enquadra nessa categoria são as lexias de busca “estabilidade gestante” e “estabilidade gravidez”. Em verificação realizada no VCB e no *corpus* especializado constituído pelas leis do Trabalho e da Previdência, não identificamos a ocorrência dessas lexias. O termo mais próximo que encontramos, nessas duas fontes, foi “estabilidade provisória”. Porém, tanto a lexia “estabilidade gestante”, como “estabilidade gravidez”, ocorrem no *corpus* especializado de jurisprudência. Ao buscarmos a definição dessas lexias no dicionário de Direito, nosso *corpus* de referência, identificamos a ocorrência de variações apenas da lexia “estabilidade gestante” no corpo do texto da definição do termo “estabilidade provisória”. Efetivamente, as lexias utilizadas pelos usuários do Portal LexMI não ocorrem como um termo, com entrada específica no Dicionário de Direito Processual do Trabalho. No entanto, ocorre a utilização de variações da lexia “estabilidade

gestante” no discurso do especialista, conforme podemos observar em alguns trechos da definição do termo “estabilidade provisória”, que transcrevemos a seguir:

Estabilidade Provisória: [...] A estabilidade provisória consiste na restrição transitória e temporária, decorrente de um fato ou evento específico, ao direito *potestativo* do empregador de resilir o contrato de trabalho, do que resulta que o empregado que detém estabilidade provisória não pode ser despedido sem justa causa. [...] A estabilidade provisória encontra-se expressa em lei ou em acordos ou convenções coletivas de trabalho. Modalidades legais de estabilidades temporárias: *Gestante*. Para a empregada, é vedada a dispensa sem justa causa no período gestacional, desde a confirmação da gravidez, estendendo-se a sua estabilidade até cinco meses após o parto. A empregada doméstica faz juz à estabilidade gestacional, nos termos da Lei n. 11.324/2006. [...] Atualmente, encontram-se consolidados no âmbito do Tribunal Superior do Trabalho os seguintes entendimentos sobre a estabilidade da gestante: (SCHWARZ, 2012, p.431-432. Grifo nosso).

Diante dos resultados, fica evidente a relevância da análise das variações dessas duas lexias como termo e/ou descritor de um tesouro, já que constam no discurso do especialista, tendo sido identificadas tanto em textos de jurisprudência (especificamente na Base de Dados de Jurisprudência do Portal LexMI), como no texto de um dicionário especializado nas áreas analisadas. Soma-se a isso o fato de percebermos, através da definição de “estabilidade provisória”, que o termo é genérico e comporta diferentes modalidades de estabilidades temporárias: um claro indício da relação hierárquica gênero/espécie. Identificamos a mesma situação com o termo “auxílio-alimentação”, pois, ao verificarmos o VCB e o *corpus* especializado constituído pelas leis do Trabalho e da Previdência, não identificamos a ocorrência dessa lexia. O termo semanticamente mais próximo que encontramos nessas fontes foi “salário utilidade”. Porém, a lexia “auxílio-alimentação” ocorre tanto no *corpus* especializado de jurisprudência, como no corpo do texto da definição do termo “salário utilidade” do Dicionário de Direito Processual do Trabalho. E, assim como o termo “estabilidade provisória”, o termo “salário utilidade” também é genérico e abrange outros tipos de utilidade, tais como: auxílio transporte, auxílio moradia, etc.

A identificação desses casos colocou em relevo a importância da presença de um glossário com as definições dos termos do tesouro disponíveis para a consulta dos usuários. Acreditamos que, para atender aos objetivos de elaboração e uso desses instrumentos, a definição disponibilizada deveria contemplar tanto os aspectos classificatórios dos termos (gênero/espécie), como linguísticos. Deveria ser, também,

uma definição de caráter enciclopédico, pois esse tipo de definição teria, possivelmente, mais condições de atender satisfatoriamente os diferentes tipos de usuários do sistema – desde o leigo, até o especialista. Entendemos que as lexias identificadas nessa categoria reforçam nosso argumento de que os sistemas de busca precisam estar o mais próximo possível da linguagem dos usuários, visto que a pergunta que antes era feita aos bibliotecários, agora é realizada por meio dos sistemas de busca.

Em síntese, podemos afirmar que das 217 lexias de nosso primeiro *corpus* de estudo, 100 são lexias numéricas ou alfanuméricas pertencentes à categoria recursos não verbais. Dessas, apenas 21 ocorrem no *corpus* textual especializado, nenhuma ocorre no VCB, mas todas ocorrem na fonte especializada (jurisprudência). Isso demonstra que, para os usuários da área jurídica, lexias numéricas e alfanuméricas devem ser consideradas como pontos de acesso à informação e, portanto, a inclusão de tais lexias é relevante como descritores ou termos equivalentes em tesouros. Na categoria aspectos conceituais, foram identificadas 32 lexias, treze delas são também contempladas pela categoria aspectos verbais e apenas duas dessas lexias ocorrem no *corpus* textual especializado. No entanto, assim como na categoria recursos não verbais, nenhuma ocorre no VCB, mas todas podem ser identificadas na fonte especializada.

Não identificamos a ocorrência de variantes das categorias recursos não verbais e aspectos conceituais nos três tesouros utilizados como parâmetro para a análise das variantes denominativas que já são contempladas por tesouros. As normas de elaboração de tesouros anteriores e vigente, como a ISO 25964-1 (2011), não mencionam a possibilidade de inserção dessas tipologias de variantes como termos equivalentes em tesouros. Entendemos que não são todas as tipologias de variantes denominativas que podem ser inseridas nos tesouros; apenas as variantes que preservam o sentido original do termo entre si é que são válidas para inserção nos tesouros, lembrando que uma delas será sempre o descritor e a outra – ou outras – será um termo equivalente.

Em nossos *corpora*, apenas 4 tipologias da classificação de Freixa (2002; 2014) não foram encontradas, são elas: termo e símbolo e termo e fórmula química da categoria alterações gráficas; alteração monoléxica/poliléxica da categoria de alterações morfossintáticas; e a alteração de base tipo [substantivo + sintagma preposicional (prep. + art. + substantivo)] da categoria alterações léxicas de unidades

poliléxicas. Acreditamos que a não ocorrência dessas tipologias de variantes se deve às características do domínio analisado: Direito do Trabalho e Direito Previdenciário.

Tal como podemos constatar, as tipologias de variantes denominativas que ocorrem em uma área de conhecimento variam de acordo com o domínio analisado, por isso nem todas as categorias criadas por Freixa (2002; 2014) em sua classificação vão necessariamente ocorrer em todos os domínios. Nesse sentido, é possível afirmar que, embora seja oriunda de um estudo importante e exaustivo, a classificação de Freixa, não contemplou todas as especificidades da área jurídica; razão pela qual, para classificação dos casos identificados em nossos *corpora*, foi necessário criar duas novas categorias: uma para variantes denominativas (alterações gráficas por supressão de preposição) e uma para termos relacionados (busca por campo semântico), ambas identificadas a partir da análise de pares de lexias.

7 PRINCÍPIOS DE ELABORAÇÃO DE TESAUROS: Metodologia para a Validação da Perspectiva do Usuário para a Seleção de Termos e Descritores de Tesauros

Como alternativa para o aperfeiçoamento desses sistemas conceituais que são os tesauros, buscamos uma metodologia que incorpore mais aspectos linguísticos em sua estrutura, com o objetivo de tornar a linguagem utilizada para a representação dos assuntos dos documentos mais próxima à linguagem dos usuários. Acreditamos que, para a efetivação desta proposição, é essencial acrescentar a perspectiva do usuário da área de conhecimento em questão através da análise e seleção das lexias de buscas utilizadas por esses usuários para a recuperação de informação em um SRI, de modo que a seleção e a validação de termos de um tesouro passem efetivamente a ser realizadas de acordo com as três validações previstas na literatura sobre elaboração de tesauros, isto é, através da garantia literária (formação de um *corpus* com textos especializados), garantia/validação do especialista, e garantia/validação dos usuários, que em nosso contexto de pesquisa se refere, na verdade, à garantia de uso pelo usuário.

Logo, o que propomos, é uma metodologia que efetivamente conte com a garantia/validação do usuário, que preferimos denominar de **metodologia para a validação da perspectiva do usuário para a seleção de termos e descritores de tesauros**. Buscamos, com esta tese, demonstrar como realizar essa validação com a utilização de lexias de buscas empregadas pelos usuários do Portal LexMI para a recuperação de informação das áreas do Direito do Trabalho e Direito Previdenciário.

Entendemos que o tesouro ideal deve preservar suas características de controle terminológico, mas, ao mesmo tempo, acrescentar a perspectiva da variação terminológica, visto que, na prática, para o bom desempenho desses instrumentos, precisamos manter a escolha prescritiva de um termo preferido, o descritor, e ampliar a rede de remissivas, ou seja, as relações de equivalência dos tesauros. A partir dessa escolha é que, de fato, se realiza o controle terminológico dos termos, pois, uma vez escolhido o descritor (termo preferido em relação às demais variantes), apenas esse

poderá ser utilizado para representar o conteúdo das obras nos registros bibliográficos em um sistema de recuperação de informação.

Embora as relações de equivalência já existam nos tesouros há bastante tempo, sempre foram subutilizadas. Por esse motivo, nossa proposta de metodologia também visa à **identificação de variantes denominativas** entre as lexias de buscas empregadas pelos usuários. Acreditamos que, ao inserir na metodologia de elaboração de tesouros a identificação de variantes como uma etapa obrigatória, também estaremos aumentando o potencial linguístico dessas ferramentas. Como resultado, teremos um melhor aproveitamento das relações de equivalência e um tesouro que, além de contemplar os aspectos conceituais, passará a contemplar os aspectos linguísticos da linguagem dos usuários.

Em nossa concepção, a estrutura dos tesouros é constituída, predominantemente, por termos. E, entre esses, apenas alguns terão o valor de descritor. Isso se justifica pelo fato de concordarmos com Van der Laan (2002) que defende que os descritores deveriam ser termos. Desse modo, podemos afirmar que encontramos na estrutura dos tesouros termos e descritores e ambos são igualmente importantes para o funcionamento adequado e desejado dos tesouros, tendo em vista que o descritor, sozinho, não pode contemplar toda a diversidade linguística dos usuários de um domínio. É somente com o auxílio da rede de remissivas constituída pelos demais termos (os não-descritores) que os tesouros podem alcançar seu objetivo maior de auxiliar na representação e recuperação da informação. É por esse motivo, também, que concordamos com Cabré e Codina (2001), quando afirmam que o trabalho documental utiliza as duas concepções da terminologia: a) a diversificada e variada, que reflete as diferenças linguísticas, conceituais e ideológicas dos autores dos documentos; e b) a terminologia normativa, necessária para controlar a variação, reduzir a dispersão formal e facilitar o controle da recuperação da informação.

O atual estágio de desenvolvimento tecnológico possibilita a utilização remota de diversos recursos, tais como bases de dados, catálogos *on-line*, repositórios digitais, etc. Consequentemente, o usuário de uma biblioteca, ou qualquer outro tipo de instituição que trabalhe com sistemas de recuperação da informação, não precisa se deslocar até a instituição para fazer suas pesquisas. Nesse contexto, embora sejam inquestionáveis os avanços trazidos pelo uso das tecnologias, no sentido de facilitar o acesso a informação, também não se pode negar a existência de um afastamento da presença física dos usuários nesses espaços. A consequência negativa desse

afastamento físico dos usuários é a redução da tradicional interação entre os usuários de um SRI e os bibliotecários que realizam a gestão da informação desses sistemas. Com esse distanciamento, os bibliotecários perdem o *feedback* imediato e deixam de acompanhar a linguagem utilizada pelos usuários para a realização de suas pesquisas. Em outras palavras, pode-se dizer que a validação ou garantia do usuário, conforme encontramos na literatura, torna-se seriamente comprometida.

Como alternativa e, principalmente, como estratégia mais viável, no contexto atual, sugerimos a utilização dos *logs* de buscas dos usuários de sistemas de informação para a realização da validação e verificação da perspectiva do usuário. Consideramos que essa é uma etapa fundamental do processo de elaboração de tesouros que, tradicionalmente, é pouco discutida e quase nunca apresentada na prática. Conforme demonstramos nesta tese, os *logs* possibilitam a identificação de termos e variantes denominativas utilizadas pelos usuários para a busca de informação. Através deles, é possível verificar os recursos linguísticos empregados por um usuário e suas estratégias ao reformular uma questão de pesquisa para a obtenção de melhores resultados, com destaque, aqui, para o uso de diferentes variantes denominativas para um mesmo conceito. No entanto, para que essa análise seja possível, é imprescindível ter acesso ao *log* completo de cada busca, ou seja, além dos termos em si (lexias de busca), o número de IP, o dia, o mês, o ano e o horário em que a pesquisa foi realizada também são dados necessários. Sem esses dados, não é possível identificar onde começa e termina a pesquisa de um usuário e, com isso, informações importantes, como as reformulações da busca realizada pelo usuário, são perdidas.

A partir da constituição de um *corpus* de lexias de buscas de usuários, é possível realizar tanto a validação dos usuários, ou seja, a verificação do uso concreto de termos, como a identificação de variantes denominativas e a verificação da forma de representação de um termo mais utilizada pelos usuários de um domínio. A verificação da variante denominativa mais utilizada pelos usuários é um dado importante para a escolha dos termos que serão escolhidos para desempenhar o papel de descritor em um tesouro. Por exemplo, verificamos, nesta pesquisa, que os usuários da área do Direito do Trabalho e Direito Previdenciário preferem utilizar as siglas ao invés do termo completo. Isso explica a razão pela qual os três tesouros da

área do Direito consultados nesta pesquisa consideram siglas como descritores e colocam a forma completa dessas siglas como termos equivalentes.

Outro exemplo significativo identificado em nossa pesquisa, que ilustra como a partir da análise de logs é possível verificar o uso concreto de termos e, ao mesmo tempo, identificar características específicas da linguagem dos usuários de um domínio, foi a constatação do uso de lexias numéricas e alfanuméricas pelos usuários da área do Direito para a realização de suas buscas. A frequência de utilização desse tipo de lexia no âmbito do Direito, de uma forma geral, é muito grande, mas não identificamos nos três tesouros que consultamos, nesta pesquisa, a ocorrência de tais tipos de lexias – nem como descritor, nem como termo equivalente. No entanto, o uso significativo desses recursos não verbais sugere fortemente a necessidade de incorporá-los aos tesouros, se não como descritores, pelo menos como termos equivalentes, até mesmo pelo fato de que números de leis, de artigos de súmulas, etc., podem ser o assunto de uma obra jurídica. Logo, os tesouros desse domínio precisam contemplar alguma forma de representação mais específica desses assuntos, pois uma obra sobre o exercício profissional de motoristas é diferente de uma obra que discute a aplicação da Lei 12.619 de 2012, que regulamenta a jornada de trabalho desses profissionais. Da mesma maneira que identificamos essa característica específica dos usuários da área do Direito, também é possível identificar características específicas de usuários de outras áreas de conhecimento, a partir da aplicação da metodologia que propomos.

Para a identificação de variantes denominativas, adotamos a classificação de Freixa (2002; 2014) por considerar que se trata de um instrumento que possibilita uma análise detalhada das variantes e, conseqüentemente, a escolha daquelas mais adequadas para inserção nos tesouros como descritores ou termos equivalentes. A análise dos pares de variantes possibilita a identificação das alterações de um termo em relação ao outro, o que nos permite identificar com clareza se o sentido do termo permanece ou se altera. Observamos que a norma de elaboração de tesouros ISO 25964-1 (2011) contempla, pelo menos, um caso de quatro das cinco categorias de Freixa (2002; 2014).

As categorias alterações léxicas e várias alterações complexas são contempladas pela norma como sinônimos ou quase-sinônimos, a categoria alterações gráficas é contemplada pela norma com a possibilidade de inserção de siglas como termos equivalentes. Consideramos que a categoria reduções é

contemplada pela norma, tendo em vista que a mesma sugere que, em alguns casos, o elaborador do tesouro deve verificar se realmente é necessário incluir termos muito específicos – conforme os propósitos do próprio tesouro –, indicando a utilização de um termo próximo mais genérico no lugar de um termo específico. Não concordamos com essa flexibilização da norma, visto que a troca de um termo específico por um termo genérico pode trazer prejuízos para o usuário, em função da elevação da quantidade de documentos desnecessários oferecidos como resposta para sua busca. Esse tipo de resposta, conseqüentemente, dificulta a recuperação da informação desejada pelo usuário, pois demanda que se filtre um volume maior de resultados. Na categoria alteração morfosintática, identificamos alterações como o uso de termos no plural e no singular, por exemplo; essas variações não são recomendadas pela norma que sugere a preferência pelo uso de descritores no singular, sem mencionar como possibilidade a inserção das variantes na forma plural como termos equivalentes.

Concluimos que não são todas as tipologias de variantes denominativas da classificação de Freixa (2002; 2014) que podem ser inseridas como termos equivalentes em tesouros. Algumas categorias precisam de uma análise mais atenta, visando à preservação do sentido do termo original, tendo em vista que, quando identificamos uma alteração significativa no sentido do termo, estamos diante de um novo termo e, não, de um termo equivalente. Essa questão da preservação do sentido dos termos coloca em evidência a relevância e a necessidade de contarmos, também, com as definições dos termos na própria estrutura dos tesouros, na parte alfabética, por exemplo, ou com a inserção de um glossário como apêndice.

Embora a norma ISO 25964-1 (2011) não mencione, especificamente, a elaboração de um glossário, nem a literatura da área seja muito enfática em relação a essa etapa, o fato é que, sem as definições precisas dos termos, não há como fazer um tesouro. Há tesouros que apresentam definições, apenas, de alguns termos, há outros – na verdade, muitos –, que inclusive confundem Nota de Escopo com definição. Isso é um problema, visto que a Nota de Escopo é um recurso relacionado com a política de indexação e serve para dar orientações sobre a utilização de um determinado termo para o bibliotecário no momento da escolha dos descritores que vão representar uma obra, o que é muito diferente de uma definição.

Na descrição da metodologia de elaboração de tesouros recentemente publicados, como o Tesouro de Ciência da Informação elaborado por Pinheiro e Ferrez (2014), publicado pelo IBICT, não verificamos a realização da garantia do

usuário. Em publicação anterior, as autoras descrevem os princípios utilizados para a elaboração do tesauro publicado em 2014, que contemplou apenas a garantia literária e a consulta a especialistas. Para garantia literária, Pinheiro e Ferrez (2010), utilizaram como método técnicas bibliométricas, as quais, resumidamente falando, são realizadas a partir da verificação da frequência de ocorrência dos termos em um *corpus* textual cuidadosamente construído. A garantia do especialista, por sua vez, foi realizada a partir do contato direto com pesquisadores da área. É compreensível essa opção apenas pela garantia literária e pela garantia do especialista para a elaboração de um tesauro da área de Ciência da Informação, devido à dupla função desses especialistas que, ao mesmo tempo, são produtores de conteúdo através de suas publicações e usuários do conteúdo produzido por seus pares. Contudo, ainda assim, os especialistas não representam toda a diversidade de usuários dessa área de conhecimento.

Acreditamos que a comparação dos termos obtidos a partir da garantia literária com os termos obtidos a partir da perspectiva dos usuários possibilita uma escolha de descritores mais adequada à realidade dos recursos de buscas dos usuários da informação sobre um domínio. Afinal, é para eles, os usuários, que os acervos são organizados e, portanto, em última análise, é para eles também que os tesauros devem ser elaborados.

8 CONCLUSÃO

No contexto atual, é preciso contar com diferentes estratégias para acompanhar as necessidades informacionais e as características específicas dos usuários de informação. Entre as estratégias possíveis, destacamos o estudo de logs, em função das diferentes possibilidades de pesquisa que eles viabilizam. No âmbito da CI, há um grande potencial para a utilização dos logs como objeto de estudo e fonte de coleta de dados, visto que ainda há poucos estudos que os utilizem.

Em nosso levantamento, encontramos apenas um trabalho no Brasil que menciona a análise de logs. Pontes (2006), em sua monografia de especialização em Gestão de Unidades de Informação pela Universidade Federal da Paraíba, fala sobre a contribuição do estudo de logs em pesquisas sobre gestão da informação, destacando a importância desses dados como ferramentas de apoio para a tomada de decisões e aperfeiçoamento dos Catálogos Online de Acesso Público (OPAC). Isto é resultado, também, da baixa frequência com que as unidades de informação coletam os logs referentes às pesquisas dos usuários. Geralmente, os *softwares* de gestão da informação das instituições estão configurados para apagar esses dados automaticamente dos sistemas, para que não ocupem espaço em seus servidores. Entendemos que, em função do espaço ocupado nos servidores, não seria necessário guardar todos os logs de acesso dos usuários ao sistema. No entanto, é possível configurar os *softwares* para coletarem e guardarem os logs apenas por alguns meses, de modo que o espaço ocupado nos servidores seja mínimo. Para isso, basta determinar uma política para a coleta e armazenagem desses dados e configurar os *softwares* de gestão de acervos, realizando, então, essa tarefa automaticamente.

Os resultados obtidos nesta tese confirmam, portanto, a relevância dos logs de pesquisa como fonte para a coleta de dados referentes ao léxico empregado pelos usuários para a recuperação da informação disponibilizada em sistemas de recuperação da informação, tais como: catálogos de bibliotecas, repositórios, bases de dados, etc. A observação da interação dos usuários com o SRI, durante o processo de recuperação da informação, sobretudo do léxico empregado, possibilita o reconhecimento de características importantes desses usuários; tanto para a tomada de decisão em relação à determinação de descritores, como para a constituição e

ampliação da rede de remissivas, que será formada pelas variantes denominativas identificadas.

A análise dos *corpora* de lexias de buscas possibilitou a identificação de três categorias que compreendem: a) aspectos linguísticos; b) recursos não verbais e c) aspectos conceituais. Essas categorias foram analisadas e classificadas de acordo com a Classificação de Freixa (2002; 2014), com exceção da categoria aspectos conceituais que, apesar de ser o resultado de uma observação sobre aspectos linguísticos, mostrou-se indissociável de aspectos predominantemente conceituais.

Entendemos que, no âmbito dos estudos de representação e recuperação da informação, o reconhecimento de variantes denominativas é fundamental para orientar a elaboração das relações de equivalência dos tesouros. Essas relações constituem um sistema de remissivas que deve ser condizente tanto com o léxico dos usuários, como com a linguagem utilizada para a representação do conteúdo dos documentos em sistemas de gerenciamento da informação, tais como catálogos de bibliotecas, repositórios digitais, entre outros.

Em vista disto, especificamente em relação à Classificação Formal de Variantes Denominativas de Freixa (2002; 2014) destacamos que a análise de variantes (com base nessa classificação), sem dúvida, pode auxiliar na identificação de variantes denominativas para inserção como termos equivalentes em tesouros. Contudo, é importante ressaltar que nem todas as variantes identificadas e classificadas de acordo com a classificação de Freixa (2002; 2014) podem ser inseridas como termos equivalentes em tesouros. Algumas categorias necessitam de uma avaliação mais cuidadosa, como a categoria referente às reduções de extensão e de base, e a categoria de alteração morfossintática referente à troca de preposição. Nesses casos, é preciso verificar se a alteração produzida preserva o sentido do termo original, visto que, se o sentido do termo se modifica, então, teremos um novo termo e não um termo equivalente.

Constatamos que as tipologias de variantes denominativas que ocorrem em uma área de conhecimento variam de uma área para outra, por isso nem todas as categorias criadas por Freixa (2002; 2014) em sua classificação vão ocorrer em outros domínios, tampouco poderão contemplar as especificidades de todas as áreas de conhecimento. Em nossos *corpora*, por exemplo, embora a classificação utilizada tenha se mostrado importante e exaustiva, não abarca todas as especificidades da área estudada. E, por essa razão, para a classificação dos casos não contemplados

pela Classificação de Freixa (2002; 2014), em nossos *corpora*, criamos duas categorias: alterações gráficas por supressão de preposição e busca por campo semântico.

Dessa forma, além de considerar aspectos estritamente linguísticos, observamos as variações na relação com o domínio de conhecimento estudado. Entendemos, assim, que a consideração pela linguagem de busca do usuário, na perspectiva que adotamos nesta tese, soma-se e justifica-se no âmbito de uma proposta de elaboração de tesouros, sem, contudo, esquecer dos propósitos de controle terminológico e de padronização, característicos da área de organização do conhecimento no âmbito da Biblioteconomia e Ciência da Informação. Diante do volume de informação produzida e disseminada atualmente, se não tivermos critérios de organização claros e padronizados, dificilmente teremos uma recuperação da informação de qualidade e, ainda, corremos o risco de perder essas informações nas *nuvens*. Logo, não podemos abreviar os caminhos da organização da informação, pois todo o tempo despendido nessa etapa é recompensado através de uma rápida e qualificada recuperação da informação.

REFERÊNCIAS

AITCHISON, J.; GILCHRIST, A.; BAWDEN, D. **Thesaurus construction and use: a practical manual**. 4th. ed. London : Aslib IMI, 2000. 201 p.

ALMEIDA, M. B. Revisiting ontologies: a necessary clarification. **Journal of the American Society for Information Science and Technology**, v. 64, n. 8, p. 1682-1693, 2013.

ALMEIDA, M. B.; MENDONÇA, F. M.; AGANETTE, E. C. Interfaces entre ontologias e conceitos seminais da ciência da informação: em busca de avanços na organização do conhecimento. In.: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 14., 2013. Florianópolis. **Anais**. Florianópolis: ANCIB/UFSC, 2013. Disponível em: <http://mba.eci.ufmg.br/downloads/OntologiasAltoMedioNivel_Enacib2013_camera_ready.pdf>. Acesso em: 30 abr. 2014.

ANSI/NISO Z39.19. **Guidelines for the construction, format, and management of monolingual controlled vocabularies**. Bethesda, Maryland: NISO Press, 2005. Disponível em: <http://www.niso.org/apps/group_public/download.php/6487/>. Acesso em: 10 de nov. 2013.

BAKER, S. Helping computers understand language. **Google Oficial Blog**. [S.l], Google, 2010. Disponível em: <<http://googleblog.blogspot.com.br/2010/01/helping-computers-understand-language.html>>. Acesso em: 15 nov. 2013.

BARITÉ, M. Garantía literaria y normas para construcción de vocabularios controlados: aspectos epistemológicos y metodológicos. **Scire: Representación y Organización del Conocimiento**, Zaragoza, v. 15, n. 2, p. 13-24, 2009. Disponível em: <<http://web.a.ebscohost.com/ehost/pdfviewer/pdfviewer?sid=6e64a3d3-6689-401d-96bf-4cd8dedffd94%40sessionmgr4005&crlhashurl=login.aspx%253fdirect%253dtrue%2526profile%253dehost%2526scope%253dsite%2526authtype%253dcrawler%2526jrnl%253d11353716%2526AN%253d47781138&hid=4114&vid=0>>. Acesso em: 15 nov. 2014.

BARRIERE, C. (2002). Hierarchical refinement and representation of the causal relation. **Terminology**, v.8, n.1. 2002. p. 91-111.

BARROS, L. A. **Curso Básico de Terminologia**. São Paulo: EDUSC, 2004. 285 p.

BONOTTO, M. *et al.* **Glossário da Área de Organização e Tratamento da Informação**. [Porto Alegre: UFRGS/DCI, 2013?]. Disponível em: <
<http://www.ufrgs.br/fabico/documentos-comgrads/glossario-de-ciencia-da-informacao>>. Acesso em: 30 abr. 2014.

BRASCHER, M. ; CAFÉ, L. Organização da Informação ou Organização do Conhecimento?. In: Marilda Lopes Ginez de Lara; Johanna Wilhelmina Smit. (Org.). **Temas de pesquisa em Ciência da Informação no Brasil**. São Paulo: Escola de Comunicações e Artes/USP, 2010, p. 85-102.

BROUGHTON, V. **Essential thesaurus construction**. Facet Publishing. 2013. 296p.

CABRÉ, M. T. El principio de poliedricidad: La articulación de ló discurso, ló cognitivo y ló lingüístico em terminología. **Ibérica**, v.1, n.16. Castelló: AELFE, 2008.

_____. Introducció. In: JORNADA DE TERMINOLOGIA I DOCUMENTACIÓ, 1., 2000, Barcelona. **Terminologia i Documentació**. Barcelona: Universitat Pompeu Fabra, Institut Universitari de Lingüística Aplicada, 2001a. p. 11 – 12.

_____.La Terminología: representación y comunicación. **Elementos para una teoría de base comunicativa y otros artículos**. Barcelona: IULA; Universitat Pompeu Fabra, 1999.

_____. **La Terminologia**: teoria, metodologia, aplicaciones. Barcelona: Antártica/Empúries,1993.

_____. Sumario de principios que configuran la nueva propuesta teórica y consecuencias metodológicas. In: CABRÉ, M. T.; FELIU, J. (edt.). **La terminología científico-técnica**: reconocimiento, análisis y extracción de información formal y semántica. Barcelona: Universitat Pompeu Fabra, Institut Universitari de Lingüística Aplicada, 2001b. p. 17-25.

CABRÉ, M. T. ; CODINA, L. Terminologia i documentació: necessitats recíproques i camps d'aplicació. In: JORNADA DE TERMINOLOGIA I DOCUMENTACIÓ, 1., 2000, Barcelona. **Terminologia i Documentació**. Barcelona: Universitat Pompeu Fabra, Institut Universitari de Lingüística Aplicada, 2001. p. 13 – 29.

CAFÉ, L. ; BRASCHER, M.; SUJII, M. K.. Elaboração de tesouros utilizando-se o programa de elaboração de tesouros em microcomputador (Tecer). **Revista de Biblioteconomia de Brasília**, Brasília, v. 18, n. 2, 1990. Disponível em:

<<http://repositorio.ibict.br/bitstream/123456789/339/1/CafeRBB1990.pdf>>. Acesso em: 25 Nov. 2014.

CAMPOS, M. L. A. **Linguagem documentária: teorias que fundamentam sua elaboração**. Niterói: EDUFF, 2001. 133p.

CROFT, W.; CRUSE, D. A. **Cognitive linguistics**. New York: Cambridge University Press, 2004.

CAMPOS, M. L. A.; GOMES, H. E. Metodologia de elaboração de tesouro conceitual: a categorização como princípio norteador. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 11, n. 3. 2006. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-99362006000300005&lng=en&nrm=iso>. Acesso em: 26 Set. 2014.

CARNEIRO, M. V. Diretrizes para uma política de indexação. **Revista da Escola de Biblioteconomia da UFMG**, Belo Horizonte, v.14. n.2, p. 221- 241, set. 1985.

CHAUMIER, J. Indexação: conceito, etapas e instrumentos. **Revista Brasileira de Biblioteconomia e Documentação**, São Paulo, v.21, n1/2, p.63-79, jan./jun. 1988.

CIAPUSCIO, G. E. **Textos especializados y terminología**. Universitat Pompeu Fabra, 2003. 149p.

CLARKE, S. G. D.; ZENG, M. L. From ISO 2788 to ISO 25964: The evolution of thesaurus standards towards interoperability and data modeling. **Information Standards Quarterly**, n. 24, vol.1. 2012. p. 20-26. Disponível em: <<http://www.niso.org/publications/isq/2012/v24no1/clarke/>>. Acesso em: 22 jul. de 2013.

CURRÁS, E. Ontologias, taxonomia e tesouros em teoria de sistemas e sistemática. Brasília: Thesaurus, 2010. 182p.

DAHLBERG, I. Teoria da Classificação, Ontem e Hoje. In.: Conferência Brasileira de Classificação Bibliográfica, Rio de Janeiro, 12-17 set., 1976. **Anais**. Brasília, IBICT/ABDF, 1979. v. 1, p. 352-370.

_____. Teoria do conceito. **Ciência da Informação**, Brasília, v. 7, n. 2, p. 101-07, 1978.

_____. Knowledge organization: a new science? **Knowledge Organization**, Würzburg: Ergon-Verlag, v. 33, n.1, p. 11-19, 2006.

_____. Knowledge organization and Terminology: philosophical and linguistic bases. **International Classification**, Frankfurt, v. 19, n. 2, p. 65-71, 1992.

DELEUZE, G.; GUATTARI, F. **A thousand plateaus**. Minneapolis, MN: University of Minnesota Press. 1987.

DODEBEI, V. L. D. **Tesouro**: linguagem de representação da memória documentária. Niterói: Intertexto; Rio de Janeiro: Interciência, 2002.

DUQUE, P. H. Teoria dos Protótipos, Categoria e Sentido Lexical: primeira parte. **Revista Philologus**, Rio de Janeiro, ano 7, n. 21. [2001?]. Disponível em: <[http://www.filologia.org.br/revista/artigo/7\(21\)13.htm](http://www.filologia.org.br/revista/artigo/7(21)13.htm)> Acesso em: 02 ago. de 2011.

ENGEL, G. I.; BISCARRA NETO, B. Terminologia e Transferência de Tecnologia. In: **Simpósio Latino-americano de Terminologia**, 2; Encontro Brasileiro de Terminologia Técnico-científica, 10-14 set. 1990, Brasília. Anais. Brasília : IBICT,1992. p.22-38.

FAULSTICH, E. Termo e Variação: Tendências no Português do Brasil. In: **Socioterminologia**. Brasília: UnB, [1998]. (Excerto, parte II).

FELBER, H. **Terminology Manual**. Paris, UNESCO, 1984. 426 p.

FILLMORE, C. J. Frames and the semantics of understanding. **Quaderni di Semantica**, [S.l.], v.5, n.2, p.222-54, Dec. 1985.

FINATTO, M. J. B. Terminologia e Ciência Cognitiva. In.: KRIEGER, M. G. ; MACIEL, A. M. B. (orgs). **Temas de Terminologia**. Porto Alegre : Editora da Universidade/UFRGS, 2001. p. 141 – 149.

FREIXA, J. Reconocimiento de unidades denominativas: incidência de La variación em El reconecimiento de las unidades terminológicas. In.: CABRÉ, M.T.; FELIU, J. (Ed.) **La terminología científico-técnica:reconocimiento, análisis y extracción de información formal y semántica**. Barcelona: IULA, UPF: 2001.

FREIXA, J. **La variació terminològica**: anàlisi de la variació denominativa en textos de diferent grau d'especialització de l'àrea de medi ambient. Tese (Doutorado) - Universitat Pompeu Fabra. Barcelona, 2002.

FREIXA, J. La variación denominativa en terminología: tipos y causas. In: ISQUERDO, Aparecida Negri e Dal Corno; MANTOVANI, G. O. **As ciências do léxico**: lexicologia, lexicografia e terminologia. Campo Grande: UFMS, 2014. Vol. VI.

GARFIELD, E. A tribute to S.R. Ranganathan, the father of Indian library science. Part 1. Life and works. **Current Contents** (6):3-10.1984.

GAUDIN, F. La socioterminologie. **Langages**, v. 39, n. 157, p. 81-93, 2005. Disponível em: <http://www.persee.fr/web/revues/home/prescript/article/lgge_0458-726x_2005_num_39_157_976>. Acesso em: 01 Out. 2014.

GUEDES, R. M. ; DIAS, E. J. W. Indexação social: abordagem conceitual. **Revista ACB**: Biblioteconomia em Santa Catarina, Florianópolis, v. 15, n. 1, p. 39-53, 2010. Disponível em : <http://www.brapci.inf.br/_repositorio/2010/06/pdf_fcb17df2cd_0010808.pdf >. Acesso em : 12 nov. 2012.

GOMES, H. E.; MOTTA, D. F.; CAMPOS, M. L. A. **Revisitando Ranganathan: a classificação na rede**. Rio de Janeiro: Conexario, 2006. Disponível em: <<http://www.conexario.com/bit/revisitando/revisitando.htm>>. Acesso em: 05 Dez. 2011.

GUARINO, N. Formal ontology and information systems. In: GUARINO, N. (Ed.). **Formal Ontology in Information Systems**, 1., 1998. Trento, Italy. **Proceedings of FOIS'98**. Amsterdam: IOS Press, 1998, p.3-15. Disponível em:<<http://uosis.mif.vu.lt/~donatas/Vadovavimas/Temos/OntologiskaiTeisingasKoncepcinisModeliavimas/papildoma/Guarino98-Formal%20Ontology%20and%20Information%20Systems.pdf>>. Acesso em: 22 nov. 2011.

GRUBER, T. R. A translation approach to portable ontology specifications. **Knowledge Acquisition**, v. 5, n. 2, p.199-220, 1993a. Disponível em: <<http://tomgruber.org/writing/ontolingua-kaj-1993.pdf>>. Acessado em: 05 jun. 2011.

HJØRLAND, B. Semantic and Knowledge organization. **ARIST**, v.41, p. 367-405, 2007.

HOFFMANN, L. Conceitos básicos da lingüística das linguagens especializadas. **Cadernos de Tradução**, Porto Alegre, n. 17, p. 79-90, 2004.

IFLA. **Requisitos Funcionales para Datos de Autoridad de Materia (FRSAD): un modelo conceptual**. Netherlands: IFLA, 2012. 66 p. Disponível em: <<http://www.ifla.org/news/now-available-frsad-in-spanish>>. Acesso em: 07 Ago. 2012.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 25964-1: Thesauri and interoperability with other vocabularies**. Part 1: Thesauri for information retrieval. Geneve: International Standard Organization, 2011.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 25964-2: Thesauri and interoperability with other vocabularies**. Part 2: Interoperability with other vocabularies. Geneve: International Standard Organization, 2013.

IYER, H. **Classificatory structures: Concepts, relations and representations**. Germany: Index Verlag, 1995.

IYER, H.; D'AMBROSIO, A. Archetypes, Idealized Cognitive Models and Prototype Effect: a discussion of images and cognition in categorization. **Advances in Knowledge Organization**, Würzburg, v. 13. 2012. p. 236-241.

JESUS, A. M. R. de; BARROS, L. A. Variação terminológica no domínio da Dermatologia: os termos que fazem referência a nomes de animais. **Estudos Lingüísticos**, v. 34, p. 1384-1389, 2005. Disponível em: <<http://www.gel.org.br/estudoslinguisticos/edicoesanteriores/4publica-estudos-2005/4publica-estudos-2005-pdfs/variacao-terminologica-na-area-697.pdf?SQMSESSID=a38ffc79c82bcbe561e1c641326fd16c>>. Acesso em: 10 jul. 2010.

KOSTINA, I. Y. **Dinamicidad de lós conceptos especializados em lós textos de diferente nível de especialización: trabajo de investigación**. 2000. Tese (Doctorado em Teoría de La Traducción) – Programa Del Doctorado em Teoría de La Traducción, Universidad Autónoma de Barcelona, Bella Terra, 2000.

KREMER, J. M. (org.) **Fontes de informação para pesquisadores e profissionais**. Belo Horizonte: UFMG, 2000. p. 246-262.

KRIEGER, M. G. O termo: questionamentos e configurações. **Tradterm**, São Paulo, v. 7, p. 111-140, 2001a. Disponível em:

<<http://www.revistas.usp.br/tradterm/article/download/49146/53228> >. Acesso em: 20 nov. 2014.

KRIEGER, M. G. O termo: questionamentos e configurações. In.: KRIEGER, M. G., MACIEL, A. M. B. *Temas de Terminologia*. Porto Alegre: Editora da Universidade, 2001b. 454p.

KRIEGER, M. G. Terminologia revisitada. **DELTA**, São Paulo , v. 16, n. 2, p. 209-228, 2000 . Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-44502000000200001&lng=en&nrm=iso>. Acesso em: 25 nov. 2014.

KRIEGER, M. G.; FINATTO, M. J. B. **Introdução à Terminologia: teoria & prática**. São Paulo: Contexto, 2004. 223 p.

KHOO, C. S. G.; JIN-CHEON, Na. Semantic relations in information science. **Annual Review of Information Science and Technology**, v. 40. 2006. p. 157–228.

LANCASTER, F.W. **Indexação e Resumos: teoria e prática**. Brasília: Briquet de Lemos Livros, 2004.

LARA, M. L. G. Problemas da organização do conhecimento na contemporaneidade. In: **Complexidade e organização do conhecimento**: desafios de nosso século. Rio de Janeiro : ISKO-Brasil ; Marília : FUNDEPE, 2013. p. 238-245.

LIMA, G. A. B. Categorização como um processo cognitivo. **Ciências e Cognição**, Rio de Janeiro, v. 11, 2007. p. 156-167. Disponível em: <<http://cienciasecognicao.tempsite.ws/revista/index.php/cec/article/view/663/444>>. Acesso em: 10 dez. 2011.

LIMA, J. O.; ALVARES, L. Organização e representação da informação e do conhecimento. In: ALVARES, L. (Org.) **Organização da Informação e do Conhecimento**: conceitos, subsídios interdisciplinares e aplicações. São Paulo: B4 Editores, 2012. 248p.

MCILWAINE, I. C. Trends in Knowledge Organization Research. **Knowledge Organization**, Würzburg, v. 30, n. 2. 2003. p. 75-86.

MOREIRA, A.; SALOMAO, M. M. Análise ontológica aplicada ao desenvolvimento de Frames. **Alfa**: Revista de Linguística (São José do Rio Preto), São Paulo , v. 56,

n. 2, Dec. 2012 . Disponível em:

<[http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1981-](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1981-57942012000200007&lng=en&nrm=iso)

57942012000200007&lng=en&nrm=iso>. Acesso em: access on 02 Oct. 2013.

MOREIRO GONZÁLEZ, J. A. **El contenido de los documentos textuales: su análisis y representación mediante el lenguaje natural**. Gijón: Trea, 2004.

MOTTA, D. F. **Método Relacional como Nova Abordagem para a Construção de Tesouros**. Rio de Janeiro, 1987. 77 p. Dissertação (Mestrado em Ciência da Informação) – Instituto Brasileiro de Informação em Ciência e Tecnologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1987.

MÜLLER, C.; CHISHMAN, R. Ontologia dos Eventos Jurídicos: contribuições da semântica verbal. In.: ONTOBRAS-MOST, Recife, Brazil, 19-21 set. 2012. **Proceedings**. Recife, CEUR Workshop, 2012. v. 938, p. 248-253.

NICHOLAS, D.; HUNTINGTON, P.; WATKINSON; A. Scholarly journal usage: the results of deep log analysis. **Journal of Documentation**.v. 61, n. 2, p.248-280, 2005.

OLIVEIRA, D. A.; ARAÚJO, R. F. Construção de Linguagens Documentárias em Sistemas de Recuperação da Informação: a importância da garantia do usuário. **Encontros Bibli**: revista eletrônica de biblioteconomia e ciência da informação, Florianópolis, v. 17, n.34, p. 17-30, maio/ago., 2012. Disponível em: <<https://periodicos.ufsc.br/index.php/eb/article/view/23994>>. Acesso em: 12 jan. 2014.

OLIVEIRA, M. Origens e Evolução da Ciência da Informação. In:_____. **Ciência da Informação e Biblioteconomia**: novos conteúdos e espaços de atuação. Belo Horizonte, Editora UFMG, 2005. 139p.

PETRUCK, M. R. L. Frame semantics. **Handbook of pragmatics**,1996. p. 1-13.

PICKLER, M. E. V. Web Semântica: ontologias como ferramentas de representação do conhecimento. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 12, n. 1, 2007. p. 65-83. Disponível em:

<<http://www.scielo.br/pdf/%0D/pci/v12n1/05.pdf>>. Acesso em: 03 dez. 2011.

PICKLER, M. E. V. Web Semântica: ontologias como ferramentas de representação do conhecimento. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 12, n. 1, 2007. p. 65-83. Disponível em:

<<http://www.scielo.br/pdf/%0D/pci/v12n1/05.pdf>>. Acesso em: 03 dez. 2011.

PINHEIRO, L. V. R.; FERREZ, H. D. Rede metodológica integrando epistemologia, organização do conhecimento, bibliometria e tesouros: concepção e construção do Tesouro brasileiro de ciência da informação. In: ROBREDO, J.; BRÄSCHER, M. (Orgs.). **Passeios pelo bosque da informação: estudos sobre a representação e organização da informação e do conhecimento**. Brasília, DF: Ibict, 2010. p. 259-276. (Edição comemorativa dos 10 anos do Grupo de Pesquisa EROIC). Disponível em: <<http://www.ibict.br/publicacoes/eroic.pdf>>.

PINHO, F. A. Fundamentos da Organização e Representação da Informação e do Conhecimento. Recife : UFPE, 2009. 156p.

PRÉVOT, L. et al. Ontology and lexicon: a multidisciplinary perspective. In.: HUANG, Chu-Ren et al. (edited). **Ontology and the Lexicon: a natural language processing perspective**. Cambridge: Cambridge University Press, 2009. p. 3- 24.

RANGANATHAN, S. R. **Prolegomena to Library Classification**. 3d. ed. Bombay: Asia Publ. House, 1967. 640 p.

RIEMER, N. **Introducing Semantics**. Cambridge: Cambridge University Press, 2010. 478p.

ROWLEY, Jennifer. *A biblioteca eletrônica*. Brasília: Briquet de Lemos, 2002. 399 p.

RUSSO, M. **Fundamentos de Biblioteconomia e Ciência da Informação**. Rio de Janeiro: E-papers Serviços Editoriais, 2010. 178p.

SAGER, J. C. **Curso práctico sobre el procesamiento de la terminología**. Madrid: Pirámide, 1993. 442p.

SAGREDO, F.; NUÑO, M. V. En los orígenes de la Biblioteconomía y Documentación: Ebla. **Documentación de las Ciencias de la Información**, Madrid, n. 17. 1994. Disponível em: <<http://revistas.ucm.es/index.php/DCIN/article/viewFile/DCIN9494110123A/20046>>. Acesso em: 09 ago. 2014.

SALES, R; CAFÉ, L. Semelhanças e Diferenças entre Tesouros e Ontologias. **DataGramaZero - Revista de Ciência da Informação**, Rio de Janeiro, v.9 n.4. 2008. Disponível em: <http://www.dgz.org.br/ago08/Art_02.htm>. Acesso em: 20 nov. 2011.

SALES, R.; CAFÉ, L. Tesouros e ontologias sob olhar da Teoria Comunicativa da Terminologia. In: NUEVAS PERSPECTIVAS PARA LA DIFUSIÓN Y ORGANIZACIÓN DEL CONOCIMIENTO, 9., 2009, Valência. **Actas del Congreso**.

Valencia: ISKO-ESPNA, 2009. p. 1-23. Disponível em:
<<http://dialnet.unirioja.es/descarga/articulo/2923286.pdf>>. Acesso em: 20 nov. 2011.

SAUSSURE, F. **Curso de Linguística Geral**. São Paulo: Cultrix, 1999. [Versão de: Cours de Linguistique Générale. Paris, Payot, 1916].

SCHREINER, H. B. Considerações Históricas Acerca do Valor das Classificações Bibliográficas. In.: Conferência Brasileira de Classificação Bibliográfica, Rio de Janeiro, 12-17 set. 1976. **Anais**. Rio de Janeiro, IBICT/ABDF, 1979. v. 1, p. 190-207.

SHERA, J. H.; EGAN, M. E. Exame do estado atual da biblioteconomia e documentação. In: BRADFORD, S. C. **Documentação**. Rio de Janeiro: Fundo de Cultura, 1961. p. 15-64.

SCHULZ, S. et al. Strengths and limitations of formal ontologies in the biomedical domain. **Reciis-Electronic Journal of Communication, Information and Innovation in Health**. v. 3, n. 1, p. 31-45, 2009.

SMITH, E. E.; MEDIN, D. L. **Categories and Concepts**. Cambridge, Harvard University Press, 1981.

SOUZA, R. R.; ALVARENGA, L. A Web Semântica e suas contribuições para a ciência da informação. **Ciência da Informação**, Brasília, v. 33, n. 1, p. 132-141, jan./abr., 2006.

SUÁREZ DE LA TORRE, M. M. La Variación Denominativa Explícita: propuesta de tipología de casos. **Organon**, Porto Alegre, v. 18, n. 37, p. 187-211, 2004. Disponível em: <<http://seer.ufrgs.br/organon/article/view/31188>>. Acesso em: 31 out. 2011.

SVENONIUS, E. **The intellectual foundation of information organization**. Cambridge: The MIT Press, 2000. 255p.

TAYLOR, J. R. Prototypes in Cognitive Linguistics. In.: ROBINSON, P.; ELLIS, N. C. (orgs). **Handbook of Cognitive Linguistics and Second Language Acquisition**. New York : Routledge, 2008. p. 39 – 65.

TREDINNICK, L. Each One of us was Several: Networks, Rhizomes and Web Organisms. **Knowledge Organization**, v. 40, n. 6, 2013.

VAN DER LAAN, R. **Tesouro e terminologia: uma relação lógica**. 173 p. Tese (Doutorado em Lingüística) – Instituto de Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2002.

VAN SLYPE, G. **Los Lenguajes de indexación: concepción, construcción y utilización en los sistemas documentales**. Madrid: Fundación Germán Sánchez Ruipérez, 1991.

VICKERY, B.C. **Classificação e indexação nas ciências**. Rio de Janeiro: BNG/Brasilart, 1980.

WILL, L. The ISO 25964 data model for the structure of an information retrieval thesaurus. In: **Categories, contexts and relations in knowledge organization**. Proceedings of the Twelfth International ISKO Conference, 6-9 August 2012, Mysore, India. Würzburg : Ergon, 2012, p.284-290. Disponível em: <http://www.willpowerinfo.co.uk/Will_ISKO2012_paper.doc>. Acesso em 24 jul. 2013.

WRIGHT, A. **Glut : mastering information through the ages**. Washington: Joseph Henry, 2008. 266 p.

WÜSTER, E. **Introducción a la teoría general de la terminología y a la lexicografía terminológica**. Barcelona : Institut Universitari de Lingüística Aplicada/Universitat Pompeu Fabra, 1998.

APÊNDICE A – Primeiro *corpus* de lexias verbais

1	Doméstica
2	Domésticas
3	Empregada doméstica
4	Empregado doméstico
5	Trabalador doméstico
6	Diarista
7	Carteira de Trabalho
8	CTPS
9	Certidão negativa de débitos trabalhistas
10	Certidão negativa trabalhista
11	certidão negativa de débitos trabalhista
12	cnd trabalhista
13	Cndt
14	CND
15	Licença Maternidade
16	Licença gestante
17	licença-maternidade
18	Décimo terceiro salário / 13º salário
19	Décimo terceiro
20	décimo-terceiro salário / 13º salário
21	"gratificação natalina"
22	abono de férias
23	férias
24	ferias
25	gratificação férias abono
26	acidente de serviço
27	Acidente de trabalho
28	Acidente do trabalho
29	acidente trabalho
30	salário-maternidade
31	Salário maternidade
32	Auxílio Doença
33	Auxílio Doenças
34	Auxílio-doença
35	auxílio-acidente
36	auxílio acidente
37	auxílio-alimentação
38	"auxílio alimentação
39	adicional de insalubridade
40	adicional insalubridade
41	insalubridade
42	adicional de periculosidade
43	adicional periculosidade
44	periculosidade

APÊNDICE A – Primeiro corpus de lexias verbais

45	Aviso prévio
46	aviso previo
47	consolidação das leis do trabalho
48	consolidação das leis trabalhistas
49	CLT
50	Direito trabalhista
51	direito trabalho
52	direito do trabalho
53	direito processual do trabalho
54	escala de plantão
55	escala de serviço
56	estabilidade gestante
57	estabilidade gravidez
58	execução trabalhista contra empregador falecido
59	execução trabalhista contra executado falecido
60	execução trabalhista contra herdeiro
61	execução trabalhista contra herdeiros
62	execução trabalhista de cujus
63	férias dobra
64	férias em dobro
65	FGTS
66	Fundo de garantia e tempo de serviço
67	fundo de garantia
68	horas extras
69	horas-extras
70	hrs extras
71	INSS
72	instituto nacional do seguro social
73	Doença ocupacional
74	DOENÇAS OCUPACIONAIS
75	seguro desemprego
76	seguro-desemprego
77	sobreaviso
78	tempo a disposicao
79	vinculo eventual fazenda
80	vinculo eventual rural
81	vinculo diarista fazenda
82	vinculo diarista rural
83	vínculo empregatício
84	vínculo empregatista
85	Vínculo de emprego
86	motorista de caminhão
87	motorista intermunicipal
88	motorista

APÊNDICE A – Primeiro *corpus* de lexias verbais

89	Motoristas
90	caminhoneiro
91	carreteiro intermunicipal
92	carreteiro
93	bombeiro civil
94	bombeiros
95	brigadista particular
96	rescisão indireta
97	Dispensa indireta por ato lesivo a honra

APÊNDICE B- Lexias alfanuméricas e numéricas

Lexias	Contexto
"11788"	Lei do Estágio - Lei 11788/08
"12275"	Altera a redação do inciso I do § 5º do art. 897 e acresce § 7º ao art. 899, ambos da Consolidação das Leis do Trabalho
"12401 2011"	inclusão de medicamentos e a incorporação de novos produtos e tecnologias no Sistema Único de Saúde (SUS)
"2530/2011"	Projeto de Lei. Acrescenta o inciso XII ao art. 10 da Lei nº 7.783, de 28 de junho de 1989, para considerar os serviços prestados pelos bancários como essenciais para os idosos.
"258724"	Acórdão nº 258724 do Processo nº20050110737745apc. CARREIRA DE MAGISTÉRIO DO DISTRITO FEDERAL. REESTRUTURAÇÃO
"347 clt engenheiro ambiental"	Art. 347 - Aqueles que exercerem a profissão de químico sem ter preenchido as condições do art. 325 e suas alíneas, nem promovido o seu registro, nos termos do art. 326, incorrerão na multa de 200 cruzeiros a 5.000 cruzeiros, que será elevada ao dobro, no caso de reincidência.
"396 CLT estar amamentando"	Art. 396 - Para amamentar o próprio filho, até que este complete 6 (seis) meses de idade, a mulher terá direito, durante a jornada de trabalho, a 2 (dois) descansos especiais, de meia hora cada um. Parágrafo único - Quando o exigir a saúde do filho, o período de 6 (seis) meses poderá ser dilatado, a critério da autoridade competente.
"445"	CLT -Art. 445 - O contrato de trabalho por prazo determinado não poderá ser estipulado por mais de 2 (dois) anos, observada a regra do art. 451. (Redação dada pelo Decreto-lei nº 229, de 28.2.1967)
"451 CLT"	Art. 451 - O contrato de trabalho por prazo determinado que, tácita ou expressamente, for prorrogado mais de uma vez passará a vigorar sem determinação de prazo. (Vide Lei nº 9.601, de 1998)
"aplicação 745-a cpc processo trabalho"	Como a <u>CLT</u> tem dispositivos específicos para tratar de liquidação e execução de sentença (artigos 876 a 892), queda inaplicável o artigo <u>475-J</u> do <u>CPC</u> , sob pena de se afrontar o comando do artigo celetista.

APÊNDICE B- Lexias alfanuméricas e numéricas

Lexias	Contexto
"477 reconhecimento de vínculo juízo"	CLT- Art. 477 - É assegurado a todo empregado, não existindo prazo estipulado para a terminação do respectivo contrato, e quando não haja êle dado motivo para cessação das relações de trabalho, o direito de haver do empregador uma indenização, paga na base da maior remuneração que tenha percebido na mesma empresa. (Redação dada pela Lei nº 5.584, de 26.6.1970)
"482 j clt"	Art. 482 - Constituem justa causa para rescisão do contrato de trabalho pelo empregador:
"607173"	REsp 607173 / RJ RECURSO ESPECIAL 2003/0184465-6. Fgts, Correção Monetária, Sucumbência Recíproca.
"727"	CLT- Art. 727 - Os vogais das Juntas de Conciliação e Julgamento, ou dos Tribunais Regionais, que faltarem a 3 (três) reuniões ou sessões consecutivas, sem motivo justificado, perderão o cargo, além de incorrerem nas penas do artigo anterior.
"8213"	Lei de Benefícios da Previdência Social - Lei 8213/91
"8541"	Lei nº 8.541 de 23 de Dezembro de 1992. Altera a legislação do Imposto de Renda e dá outras providências.
"8542"	Dispõe sobre a política nacional de salários.
"8742 de 1993"	Dispõe Sobre a Organização da Assistência Social e dá outras providências.
"9715"	Lei nº 9.715, de 25 de Novembro de 1998. Dispõe sobre as contribuições para os Programas de Integração Social e de Formação do Patrimônio do Servidor Público - PIS/PASEP, e dá outras providências.
"art 21, III, d 8.213/91"	Previdência. Art. 21. Equiparam-se também ao acidente do trabalho, para efeitos desta Lei:
"art 892 CLT"	CLT - Art. 892 - Tratando-se de prestações sucessivas por tempo indeterminado, a execução compreenderá inicialmente as prestações devidas até a data do ingresso na execução.
"art. 137 clt"	Art. 137 - Sempre que as férias forem concedidas após o prazo de que trata o art. 134, o empregador pagará em dobro a respectiva remuneração. (Redação dada pelo Decreto-lei nº 1.535, de 13.4.1977)

APÊNDICE B- Lexias alfanuméricas e numéricas

Lexias	Contexto
"art. 195, I salário-maternidade"	O art. 195 da Instrução Normativa INSS/PRES nº 45/10 estabelece que a renda mensal do salário-maternidade para a segurada empregada, consiste numa renda mensal igual à sua remuneração no mês do seu afastamento, ou se for o caso, de salário total ou parcialmente variável, na igualdade da média aritmética simples dos seus seis últimos salários, apurada de acordo com a lei salarial ou o dissídio coletivo da categoria, excetuando-se o 13º salário, adiantamento de férias e as rubricas constantes do § 9º do art. 214 do RPS.
LEI N. 11.960/2009	obre parcelamento de débitos de responsabilidade dos Municípios, decorrentes de contribuições sociais de que tratam as alíneas a e c do parágrafo único do art. 11 da Lei nº 8.212, de 24 de julho de 1991
"art. 467 clt"	Art. 467. Em caso de rescisão de contrato de trabalho, havendo controvérsia sobre o montante das verbas rescisórias, o empregador é obrigado a pagar ao trabalhador, à data do comparecimento à Justiça do Trabalho, a parte incontroversa dessas verbas, sob pena de pagá-las acrescidas de cinquenta por cento". (Redação dada pela Lei nº 10.272, de 5.9.2001)
"art. 482, j, clt assédio moral"	Art. 482 - Constituem justa causa para rescisão do contrato de trabalho pelo empregador:
"art.652, "d" da CLT"	Art. 652 - Compete às Juntas de Conciliação e Julgamento: (Vide Constituição Federal de 1988)
"artigo 1º F Lei 9494 princípio simetria"	Disciplina a aplicação da tutela antecipada contra a Fazenda Pública, altera a Lei nº 7.347, de 24 de julho. Art. 1º-A. Estão dispensadas de depósito prévio, para interposição de recurso, as pessoas jurídicas de direito público federais, estaduais, distritais e municipais. (Incluído pela Medida provisória nº 2.180-35, de 2001) de 1985, e dá outras providências.
"artigo 764 da CLT"	Art. 764 - Os dissídios individuais ou coletivos submetidos à apreciação da Justiça do Trabalho serão sempre sujeitos à conciliação.
"ato 317/2012"	Altera a Resolução Administrativa 1.470/2011, que regulamenta o Banco Nacional de Devedores Trabalhistas.
"ato segjud/gp 491"	

APÊNDICE B- Lexias alfanuméricas e numéricas

Lexias	Contexto
"aviso prévio com redução de sete dias 477 CL"	Art. 477 - É assegurado a todo empregado, não existindo prazo estipulado para a terminação do respectivo contrato, e quando não haja êle dado motivo para cessação das relações de trabalho, o direito de haver do empregador uma indenização, paga na base da maior remuneração que tenha percebido na mesma empresa. (Redação dada pela Lei nº 5.584, de 26.6.1970)
"aviso prévio lei 12506"	nova lei do Aviso Prévio
"Decreto 3.597"	Promulga Convenção 182 e a Recomendação 190 da Organização Internacional do Trabalho (OIT) sobre a Proibição das Piores Formas de Trabalho Infantil e a Ação Imediata para sua Eliminação, concluídas em Genebra, em 17 de junho de 1999.
"decreto 73626/74"	É aprovado o anexo Regulamento, assinado pelo Ministro do Trabalho e Previdência Social, disciplinando a aplicação das normas concernentes às relações individuais e coletivas de trabalho rural, estatuídas pela Lei nº 5.889, de 8 de junho de 1973.
"dedutibilidade e remuneração e 357"	<p>implicações quanto à legislação do imposto de renda e à legislação previdenciária. Art. 357. Serão dedutíveis na determinação do lucro real as remunerações dos sócios, diretores ou administradores, titular de empresa individual e conselheiros fiscais e consultivos (Lei nº 4.506, de 1964, art. 47).</p> <p>Parágrafo único. Não serão dedutíveis na determinação do lucro real (Decreto-Lei nº 5.844, de 1943, art. 43, § 1º, alíneas b e d):</p> <p>I - as retiradas não debitadas em custos ou despesas operacionais, ou contas subsidiárias, e as que, mesmo escrituradas nessas contas, não correspondam à remuneração mensal fixa por prestação de serviços (Decreto-Lei nº 5.844, de 1943, art. 43, § 1º, alíneas b e d);</p>
"dispensa arbitrária lei 9029"	LEI Nº 9.029, DE 13 DE ABRIL DE 1995. Proíbe a exigência de atestados de gravidez e esterilização, e outras práticas discriminatórias, para efeitos admissionais ou de permanência da relação jurídica de trabalho, e dá outras providências.
"Dissídio Coletivo 0155"	Não foi possível identificar o significado do número, existem muitas possibilidades.

APÊNDICE B- Lexias alfanuméricas e numéricas

Lexias	Contexto
"empregado domestico multa 477 clt"	CLT- Art. 477 - É assegurado a todo empregado, não existindo prazo estipulado para a terminação do respectivo contrato, e quando não haja êle dado motivo para cessação das relações de trabalho, o direito de haver do empregador uma indenização, paga na base da maior remuneração que tenha percebido na mesma empresa. (Redação dada pela Lei nº 5.584, de 26.6.1970)
"engenheiro piso tst lei 4.950-a"	PISO SALARIAL DE ENGENHEIRO. LEI Nº 4.950-A/66.
"imunidade e artigo 14 e CTN e contribuições p	da imunidade tributária das entidades beneficentes de assistência social
"IN 16 do TST"	IN 16/TST. AGRAVO DE INSTRUMENTO NÃO CONHECIDO.
"IN 16"	
"Instrução Normativa 16 TST"	
"instrução normativa 03/93"	Ainda, conforme a Instrução Normativa do TST n. 3/93, os depósitos de que trata o art. 40, e seus parágrafos, da Lei n. 8.177/91, com a redação dada pelo art. 8º da Lei nº 8.542/92, não tem natureza jurídica de taxa de recurso, mas de garantia do juízo recursal , que pressupõe decisão condenatória ou executória de obrigação de pagamento em pecúnia, com valor líquido ou arbitrado. O depósito recursal está disciplinado no artigo 899 da Consolidação das Leis do Trabalho – CLT.
"instrução normativa 27/05"	Instrução Normativa n. 27, que dispõe sobre normas procedimentais aplicáveis ao processo do trabalho em decorrência da ampliação da competência da Justiça do Trabalho pela Emenda Constitucional n. 45, de 8 de dezembro de 2004.
"instrução normativa 30/2007"	A parte que protocoliza seu recurso por intermédio do sistema e-DOC está dispensada de apresentar as vias originais dos comprovantes de recolhimento do depósito recursal e do pagamento de custas processuais, nos termos do art. 7º da Instrução Normativa n. 30 /2007 do TST.
"intervalo art. 66, da CLT"	O desrespeito ao intervalo mínimo de onze horas entre as jornadas de trabalho enseja a recomposição do prejuízo causado ao empregado, remunerando-o com horas extraordinárias quando não observado o intervalo interjornada estabelecido no artigo 66 da CLT

APÊNDICE B- Lexias alfanuméricas e numéricas

Lexias	Contexto
"justiça do trabalho 268 cpc"	A preempção prevista no CPC não se aplica ao processo do trabalho, pois a CLT contém penalidade específica no art. 732 para o reclamante que der causa ao arquivamento de reclamação ajuizada por duas vezes seguidas. Destarte, não se pode falar em aplicação do artigo 268 do CPC ao processo do trabalho, pois o direito processual comum somente é aplicável subsidiariamente quando a CLT for omissa, nos termos do art. 769 da CLT
"laudo de insalubridade art 195 da clt"	Para que seja reconhecido o labor em ambiente insalubre, a realização de perícia técnica é obrigatória, nos termos do art. 195 da CLT, isso porque o laudo pericial deve ser elaborado de acordo com os critérios técnicos definidos nos anexos da NR-15 do MTE, nos quais se estabelecem os agentes insalubres, os limites de tolerância e os critérios técnicos e legais para avaliar e caracterizar as atividades e operações insalubres, bem como o adicional devido para cada caso. Dessa forma, apenas após realizada a perícia, com as especificidades técnicas elencadas na NR-15, poderá ser caracterizado o exercício em condições insalubres, ficando assegurado ao empregado a percepção do referido adicional.
"LEI 4.348 1964"	Estabelece normas processuais relativas a mandado de segurança.
"lei 10406 2012"	ITUI A SEMANA DE APOIO AO EMPREGO E À GERAÇÃO DE TRABALHO E RENDA. O Povo do Município de Belo Horizonte, por seus representantes, decreta e eu sanciono a seguinte Lei:
"Lei 12.619/2012"	Lei 12.619/2012 regulamenta o exercício da profissão dos motoristas profissionais que laboram mediante vínculo empregatício. Na prática, as regras proíbem os profissionais de dirigir por um período superior a quatro horas sem descanso mínimo de 30 minutos.
"lei 12551"	Altera o art. 6º da Consolidação das Leis do Trabalho (CLT), aprovada pelo Decreto-Lei 5452, de 1º de maio de 1943, para equiparar os efeitos jurídicos da subordinação exercida por meios telemáticos e informatizados à exercida por meios pessoais e diretos. Não se distingue entre o trabalho realizado no estabelecimento do empregador, o executado no domicílio do empregado e o realizado a distância, desde que estejam caracterizados os pressupostos da relação de emprego

APÊNDICE B- Lexias alfanuméricas e numéricas

Lexias	Contexto
"lei 7394/85"	Regula o Exercício da Profissão de Técnico em Radiologia, e dá outras providências.
"lei 7701/81"	Encontramos apenas a lei 7701 de 1988. Dispõe sobre a especialização de Turmas dos Tribunais do Trabalho em processos coletivos e dá outras providências
"lei 8.460 de 17 de setembro de 1992"	Concede antecipação de reajuste de vencimentos e de soldos dos servidores civis e militares do Poder Executivo e dá outras providências.
"Lei 8742 de 07 de dezembro de 1993"	Dispõe sobre a organização da Assistência Social e dá outras providências.
"lei 9.796 de 05 de maio de 1999"	LEI HAULY - ALTERADA. Dispõe sobre a compensação financeira entre o Regime Geral de Previdência Social e os regimes de previdência dos servidores da União, dos Estados, do Distrito Federal e dos Municípios, nos casos de contagem recíproca de tempo de contribuição para efeito de aposentadoria, e dá outras providências.
"lei 9.796 de 5 maio de 1999"	
"lei federal 6.194"	DPVAT foi instituído pela Lei Federal 6.194, de 19 de dezembro de 1974, que estabeleceu diferentes coberturas para as pessoas vitimadas por acidentes, quais sejam: indenização por morte, indenização por invalidez permanente e reembolso das despesas de assistência médica e suplementares.
"Lei n 7.701, de 21 de dezembro de 1988"	Dispõe sobre a especialização de Turmas dos Tribunais do Trabalho em processos coletivos, e dá outras providências.
"lei nº 1.771 de 28 de outubro de 1952"	Dispõe sobre o Estatuto dos Funcionários Públicos Civis da União. Revogada pela Lei nº 8.112, de 1990
"lei no 9.717 de 27 de novembro de 1998"	Dispõe sobre regras gerais para a organização e o funcionamento dos regimes próprios de previdência social dos servidores públicos da União, dos Estados, do Distrito Federal e dos Municípios, dos militares dos Estados e do Distrito Federal e dá outras providências.
"MEDIDA PROVISÓRIA 563"	

APÊNDICE B- Lexias alfanuméricas e numéricas

Lexias	Contexto
"MP 563/2012"	Altera a alíquota das contribuições previdenciárias sobre a folha de salários devidas pelas empresas que especifica, institui o Programa de Incentivo à Inovação Tecnológica e Adensamento da Cadeia Produtiva de Veículos Automotores, o Regime Especial de Tributação do Programa Nacional de Banda Larga para Implantação de Redes de Telecomunicações, o Regime Especial de Incentivo a Computadores para Uso Educacional, o Programa Nacional de Apoio à Atenção Oncológica, o Programa Nacional de Apoio à Atenção da Saúde da Pessoa com Deficiência, restabelece o Programa Um Computador por Aluno, altera o Programa de Apoio ao Desenvolvimento Tecnológico da Indústria de Semicondutores, instituído pela Lei nº 11.484, de 31 de maio de 2007, e dá outras providências.
"multa art 477 § 8"	É assegurado a todo empregado, não existindo prazo estipulado para a terminação do respectivo contrato, e quando não haja êle dado motivo para cessação das relações de trabalho, o direito de haver do empregador uma indenização, paga na base da maior remuneração que tenha percebido na mesma empresa. § 8º - A inobservância do disposto no § 6º deste artigo sujeitará o infrator à multa de 160 BTN, por trabalhador, bem assim ao pagamento da multa a favor do empregado, em valor equivalente ao seu salário, devidamente corrigido pelo índice de variação do BTN, salvo quando, comprovadamente, o trabalhador der causa à mora.
"multa do 475j justiça do trabalho"	uma multa de 10% sobre o valor total da condenação para os casos em que há o descumprimento da sentença judicial. Essa multa, prevista no artigo 475-J do Código de Processo Civil
"nao violacao dos artigos 114 e 202 da cf comp"	Art. 114. Compete à Justiça do Trabalho processar e julgar: (Redação dada pela Emenda Constitucional nº 45, de 2004). Art. 202. É assegurada aposentadoria, nos termos da lei, calculando-se o benefício sobre a média dos trinta e seis últimos salários de contribuição, corrigidos monetariamente mês a mês, e comprovada a regularidade dos reajustes dos salários de contribuição de modo a preservar seus valores reais e obedecidas as seguintes condições

APÊNDICE B- Lexias alfanuméricas e numéricas

Lexias	Contexto
"nr 17"	NORMA REGULAMENTADORA 17. Esta Norma Regulamentadora visa a estabelecer parâmetros que permitam a adaptação das condições de trabalho às características psicofisiológicas dos trabalhadores, de modo a proporcionar um máximo de conforto, segurança e desempenho eficiente.
"oj 142"	Orientação Jurisprudencial nº 142 da SBDI-1 do TST, é passível de nulidade a decisão que acolhe embargos declaratórios com efeito modificativo sem conferir oportunidade para a parte contrária manifestar-se, ante a vulneração dos princípios constitucionais do devido processo legal e da ampla defesa.
"oj 398 acordo"	CONTRIBUIÇÃO PREVIDENCIÁRIA. ACORDO HOMOLOGADO EM JUÍZO SEM RECONHECIMENTO DE VÍNCULO DE EMPREGO. CONTRIBUINTE INDIVIDUAL. RECOLHIMENTO DA ALÍQUOTA DE 20% A CARGO DO TOMADOR E 11% A CARGO DO PRESTADOR DE SERVIÇOS. (DEJT divulgado em 02, 03 e 04.08.2010)
"OJ 70"	Nos acordos homologados em juízo em que não haja o reconhecimento de vínculo empregatício, é devido o recolhimento da contribuição previdenciária, mediante a alíquota de 20% a cargo do tomador de serviços e de 11% por parte do prestador de serviços, na qualidade de contribuinte individual, sobre o valor total do acordo, respeitado o teto de contribuição. Inteligência do § 4º do art. 30 e do inciso III do art. 22, todos da Lei n.º 8.212, de 24.07.1991.
"Opção sexual e ato lesivo a honra e boa fama Opção sexual no trabalho art.483 "e""	Art. 483 - O empregado poderá considerar rescindido o contrato e pleitear a devida indenização quando: e) praticar o empregador ou seus prepostos, contra ele ou pessoas de sua família, ato lesivo da honra e boa fama;
"pec 89 aposentadoria"	
"pec 89"	Proposta de Emenda Constitucional 89/03, de autoria da senadora Ideli Salvatti (PT-SC), que tem como objetivo autorizar a perda de cargo do juiz ou membro do Ministério Público por decisão de dois terços dos membros do tribunal ou conselho ao qual estiver vinculado.
"pis 9.715 vacancia"	LEI Nº 9.715, DE 25 DE NOVEMBRO DE 1998. Dispõe sobre as contribuições para os Programas de Integração Social e de Formação do Patrimônio do Servidor Público - PIS/PASEP, e dá outras providências.

APÊNDICE B- Lexias alfanuméricas e numéricas

Lexias	Contexto
"provimento 01/2012 cgjt"	Dispõe sobre os procedimentos a serem adotados pelos MM. Juízos do Trabalho relativamente a credores trabalhistas de Empresa Falida ou em Recuperação Judicial e dá outras providências
"provimento 01/2012 TST"	
"regulamenta a profissão de médico decreto "	LEI Nº 12.842, DE 10 DE JULHO DE 2013. Dispõe sobre o exercício da Medicina. Ato médico.
"rescisão injusta art. 474"	CLT -Art. 474 - A suspensão do empregado por mais de 30 (trinta) dias consecutivos importa na rescisão injusta do contrato de trabalho.
"rescisão e lei 9.656/98"	Dispõe sobre os planos e seguros privados de assistência à saúde
"subsidiaria artigo 467 clt"	CLT - Art. 467. Em caso de rescisão de contrato de trabalho, havendo controvérsia sobre o montante das verbas rescisórias, o empregador é obrigado a pagar ao trabalhador, à data do comparecimento à Justiça do Trabalho, a parte incontroversa dessas verbas, sob pena de pagá-las acrescidas de cinquenta por cento". (Redação dada pela Lei nº 10.272, de 5.9.2001)
"sumula 364"	Súmula nº 364 - TST
"sumula 113"	Adicional de Periculosidade - Exposição Eventual, Permanente e Intermitente
"sumula 16"	A conversão em tempo de serviço comum, do período trabalhado em condições especiais, somente é possível relativamente à atividade exercida até 28 de maio de 1998 (art. 28 da Lei nº 9.711/98).
"SUMULA 228"	ADICIONAL DE INSALUBRIDADE. BASE DE CALCULO. A partir de 9 de maio de 2008, data da publicação da Súmula Vinculante nº 4 do Supremo Tribunal Federal, o adicional de insalubridade será calculado sobre o salário básico, salvo critério mais vantajoso fixado em instrumento coletivo.
"sumula 236"	Súmula nº 236 do TST HONORÁRIOS PERICIAIS. RESPONSABILIDADE (cancelada) - Res. 121/2003, DJ 19, 20 e 21.11.2003 A responsabilidade pelo pagamento dos honorários periciais é da parte sucumbente na pretensão relativa ao objeto da perícia.

APÊNDICE B- Lexias alfanuméricas e numéricas

Lexias	Contexto
"Súmula 278, item II do TST "	<p>Súmula nº 278 do TST EMBARGOS DE DECLARAÇÃO. OMISSÃO NO JULGADO (mantida) - Res. 121/2003, DJ 19, 20 e 21.11.2003 A natureza da omissão suprida pelo julgamento de embargos declaratórios pode ocasionar efeito modificativo no julgado. Súmula nº 278 do TST</p>
"súmula 331"	<p>Súmula nº 331 do TST CONTRATO DE PRESTAÇÃO DE SERVIÇOS. LEGALIDADE (nova redação do item IV e inseridos os itens V e VI à redação) - Res. 174/2011, DEJT divulgado em 27, 30 e 31.05.2011 I - A contratação de trabalhadores por empresa interposta é ilegal, formando-se o vínculo diretamente com o tomador dos serviços, salvo no caso de trabalho temporário (Lei nº 6.019, de 03.01.1974).</p>
"súmula 331, IV, do TST"	<p>IV - O inadimplemento das obrigações trabalhistas, por parte do empregador, implica a responsabilidade subsidiária do tomador dos serviços quanto àquelas obrigações, desde que haja participado da relação processual e conste também do título executivo judicial.</p>
"SÚMULA 338"	<p>Súmula nº 338 do TST JORNADA DE TRABALHO. REGISTRO. ÔNUS DA PROVA (incorporadas as Orientações Jurisprudenciais nºs 234 e 306 da SBDI-1) - Res. 129/2005, DJ 20, 22 e 25.04.2005 I - É ônus do empregador que conta com mais de 10 (dez) empregados o registro da jornada de trabalho na forma do art. 74, § 2º, da CLT. A não-apresentação injustificada dos controles de frequência gera presunção relativa de veracidade da jornada de trabalho, a qual pode ser elidida por prova em contrário. (ex-Súmula nº 338 – alterada pela Res. 121/2003, DJ 21.11.2003)</p>

APÊNDICE B- Lexias alfanuméricas e numéricas

Lexias	Contexto
"sumula 353 tst"	Súmula nº 353 do TST EMBARGOS. AGRAVO. CABIMENTO (nova redação da letra "f" em decorrência do julgamento do processo TST-IUJ-28000-95.2007.5.02.0062) – Res. 189/2013, DEJT divulgado em 13, 14 e 15.03.2013 Não cabem embargos para a Seção de Dissídios Individuais de decisão de Turma proferida em agravo, salvo:
"súmula 60 do tst jornada mista a partir de 0 ho	Súmula nº 60 do TST ADICIONAL NOTURNO. INTEGRAÇÃO NO SALÁRIO E PRORROGAÇÃO EM HORÁRIO DIURNO (incorporada a Orientação Jurisprudencial nº 6 da SBDI-1) - Res. 129/2005, DJ 20, 22 e 25.04.2005 I - O adicional noturno, pago com habitualidade, integra o salário do empregado para todos os efeitos. (ex-Súmula nº 60 - RA 105/1974, DJ 24.10.1974) II - Cumprida integralmente a jornada no período noturno e prorrogada esta, devido é também o adicional quanto às horas prorrogadas. Exegese do art. 73, § 5º, da CLT. (ex-OJ nº 6 da SBDI-1 - inserida em 25.11.1996)
"SUMULA SDI 341"	Súmula 341/TST SDI I É de responsabilidade do empregador o pagamento da diferença da multa de 40% sobre os depósitos do FGTS, decorrente da atualização monetária em face dos expurgos inflacionários.»
"suspensão 482 j clt "	Art. 482 - Constituem justa causa para rescisão do contrato de trabalho pelo empregador:

APÊNDICE B- Lexias alfanuméricas e numéricas

Lexias	Contexto
"XII da instrução n. 01/82 do TST"	<p>XII - após calculada a recomposição salarial, serão compensados os aumentos salariais, espontâneos ou compulsórios, concedidos durante o prazo da vigência do acordo, da convenção ou da sentença anterior, exceto os provenientes de:</p> <p>a) término de aprendizagem (Decreto n.º 31.456, de 6 de outubro de 1953);</p> <p>b) implemento de idade;</p> <p>c) promoção por antigüidade ou merecimento;</p> <p>d) transferência de cargo, função, estabelecimento ou de localidade;</p> <p>e) equiparação salarial determinada por sentença transitada em julgado.</p>
"475-J processo do trabalho"	<p>EMENTA: MULTA DO ARTIGO 475-J DO CPC - APLICABILIDADE NO PROCESSO DO TRABALHO. A multa do artigo 475-J, do CPC, é aplicável subsidiariamente no Processo do Trabalho, com fulcro no artigo 769, da CLT. A aplicabilidade da multa fica condicionada à prévia intimação da executada para cumprimento da obrigação no prazo assinado. Todavia, em se tratando de execução provisória, não haverá incidência da multa, visto que ainda não consolidado o débito.</p>
13º	<p>Ocorre 13º em acordãos de diversos TRT, mas como a busca é muito poluída não dá para ver o número de ocorrências. Para 13º salário há 5319 ocorrências.</p>
13º salário	

APENDICE C – Lexias aspectos conceituais

1.	Falta de depósito de fundo de garantia
2.	Rscisão indireta;
3.	Licença maternidade estabilidade
4.	Estabilidade provisória;
5.	Legislação motoboy
6.	Nova lei para motoboys;
7.	Lei do estágio
8.	Lei do idoso
9.	Lei do trabalho
10.	Qual a forma de cálculo dos depósitos em atraso do FGTS
11.	Qual a forma de cálculo do FGTS;
12.	Lei que exeta imposto de pessoas portadoras de doenças grave
13.	Projeto de lei que quer aumentar jornada dos aeronautas
14.	Projeto de lei que quer aumentar jornada dos pilotos
15.	Pec da obrigatoriedade do curso de jornalista
16.	Pec dos jornalistas;
17.	Pec da exigencia de curso de jornalista;
18.	regulamentação taxistas
19.	regulamenta a profissão de físico
20.	lei que fala sobre aposentadoria por invalides
21.	lei sobre doenças ocupacionais
22.	lei sobre medidas preventivas acidentes de trabalho
23.	lei trabalhista rural
24.	LEI DAS DOMESTICAS
25.	Lei Esteticista
26.	Lei Barbeiro, Cabelereiro, Esteticista, Manicuri
27.	Lei Massagem
28.	Lei Massoterapia
29.	auxílio-alimentação
30.	estabilidade gestante
31.	estabilidade gravidez
32.	Falta de depósito de fundo de garantia

Fonte: Elaborado pela autora.

APÊNDICE D - Ocorrência das lexias de buscas nos corpora da pesquisa

Lexias de busca	Corpus Textual Especializado	VCB	LexMI-Jurisprudência
Doméstica	Não ocorre	Não ocorre	7422 ocorrências
Domésticas	Não ocorre	Não ocorre	7422 ocorrências
Empregada doméstica	Ocorre	Não ocorre	1056 ocorrências
Empregado doméstico	Ocorre	ocorre como autorizado	1056 ocorrências
Trabalador doméstico	Não ocorre	Não ocorre	Não ocorre Trabalhador domestico tem 166 ocorrências
Diarista	Não, mas ocorre Empregado diarista	Não. Apenas empregado doméstico diarista como autorizado.	439 ocorrências
Carteira de Trabalho	Ocorre	ocorre como não	1159 ocorrências.
CTPS	Ocorre	Não ocorre	8186 ocorrências
Certidão negativa de débitos trabalhista	Ocorre	Não ocorre	4 ocorrências.
Certidao negativa trabalhista	Não ocorre	Não ocorre	Não ocorre
certidao negativa de debitos trabalhista	Não ocorre	Não ocorre	4 ocorrências.
cnd trabalhista	Não ocorre	Não ocorre	Não ocorre
Cndt	Ocorre	Não ocorre	2 ocorrências
CND	ocorre	Não ocorre	8 ocorrencias relacionadas aos tribunais de Trabalho
Licença Maternidade	Não ocorre	ocorre como não	773 ocorrências
Licença gestante	Não ocorre	ocorre como autorizado Licença à gestante	129 ocorrências
licença-maternidade	Ocorre	Não ocorre	773 ocorrências
Décimo terceiro salário / 13º salário	Não ocorre	Ocorre como autorizado	3399 ocorrências

APÊNDICE D - Ocorrência das lexias de buscas nos corpora da pesquisa

Lexias de busca	Corpus Textual Especializado	VCB	LexMI-Jurisprudência
Décimo terceiro	Ocorre e décimo-terceiro sal	Não ocorre	3709 ocorrências e décimo-terceiro salário
décimo-terceiro salário / 13º salário	Ocorre	Não ocorre	3399 ocorrências
"gratificação natalina"	ocorre	Não ocorre	8940 ocorrencias
abono de férias	Ocorre	Abono de férias é o autorizado	563 ocorrencias
férias	Ocorre	ocorre como autorizado	29557 ocorrencias
ferias	Não ocorre	Não ocorre	29557 ocorrencias
gratificação férias abono	Não ocorre	não ocorre	não ocorre
acidente de serviço	Não ocorre	Não ocorre	ocorre, porém em apenas uma decisão do TST
Acidente de trabalho	Não ocorre	Não ocorre	18779 ocorrências.
Acidente do trabalho	Ocorre	Ocorre como autorizado	14308 ocorrências.
acidente trabalho	Não ocorre	Não ocorre	Não ocorre
salário-maternidade	Ocorre	ocorre como autorizado	353 ocorrências
Salário maternidade	Não ocorre	Não ocorre	353 ocorrências
Auxilio Doença	Não ocorre	Não ocorre	6871 ocorrências.
Auxilio Doenças	Não ocorre	Não ocorre	6871 ocorrências.
Auxilio-doença	Ocorre	Ocorre	6871 ocorrências.
auxílio-acidente	Ocorre	Não ocorre	3690 ocorrências
auxílio acidente	Não ocorre	Não ocorre	3690 ocorrências
auxílio-alimentação	Não, ocorre apenas salário-utilidade	Não ocorre	13494 ocorrências.
"auxílio alimentação	Não ocorre	Não ocorre	13494 ocorrências.
adicional de insalubridade	ocorre	ocorre	48623 ocorrencias
adicional insalubridade	não ocorre	não ocorre	149 ocorrencias

APÊNDICE D - Ocorrência das lexias de buscas nos corpora da pesquisa

Lexias de busca	Corpus Textual Especializado	VCB	LexMI-Jurisprudência
insalubridade	ocorre	ocorre	51195 ocorrencias
adicional de periculosidade	fica subentendido adicinoal de insalubridade ou de preiculosidade	ocorre	47132 ocorrencias
adicional periculosidade	não ocorre	Não ocorre	178 ocorrencias
periculosidade	ocorre	ocorre	64503 ocorrencias
Aviso prévio	ocorre	ocorre	26846 ocorrencias
aviso previo	não ocorre	Não ocorre	22426 ocorrencias
consolidação das leis do trabalho	Ocorre	Ocorre como autorizado.	44613 ocorrências.
consolidação das leis trabalhistas	Não ocorre	Não ocorre	560 ocorrências
CLT	Ocorre	ocorre como não autorizado para a forma	654365 ocorrencias.
Direito trabalhista	Não ocorre	ocorre como não autorizado em direito	ocorre 8027 ocorrências.
direito trabalho	Não ocorre	Não ocorre	Não ocorre
direito do trabalho	Ocorre	Ocorre	Ocorre 12081 ocorrências.
direito processual do trabalho	Ocorre	Não ocorre	1298 ocorrências.
escala de plantão	Não, mas ocorre escala de prontidão	Não ocorre	90 ocorrências
escala de serviço	Ocorre	Não ocorre	91 ocorrências.
estabilidade gestante	Não, ocorre apenas estabilidade provisória	Não ocorre	831 ocorrencias
estabilidade gravidez	Não ocorre	Não ocorre	9 ocorrencias
execução trabalhista contra empregador falecido	Não, mas ocorre apenas morte do empregador	Ocorre morte do	ocorre empregador falecido

APÊNDICE D - Ocorrência das lexias de buscas nos corpora da pesquisa

Lexias de busca	Corpus Textual Especializado	VCB	LexMI-Jurisprudência
execução trabalhista contra executado	Não ocorre	Não ocorre	Não ocorre
execução trabalhista contra herdeiro	Não ocorre	Não ocorre	Não ocorre
execução trabalhista contra herdeiros	Não ocorre	Não ocorre	Não ocorre
execução trabalhista de cujus	Não ocorre	Não ocorre	2243 ocorrências para De cujus. Tribunal Regional do Trabalho (193) Tribunal Superior do Trabalho (218)
férias dobra	Não ocorre	Não ocorre	203 itens como termos separados férias. Dobra . Dobra de férias 344 ocorrencias
férias em dobro	Não ocorre	Não ocorre	926 ocorrências
FGTS	Ocorre	ocorre como termo não autorizado para a forma completa do	125140 ocorrências.
Fundo de garantia e tempo de serviço	Ocorre Fundo de Garantia do tempo de Serviço	Ocorre Fundo de Garantia do tempo de	1 ocorrência fundo de garantia do tempo de serviço
fundo de garantia	não ocorre	Não ocorre	5263 ocorrências
horas extras	Ocorre	hora extra	Ocorre mesmo resultado para o termo com hífen e também no singular 167996 ocorrências
horas-extras	Não ocorre	Não ocorre	ocorre
hrs extras	Não ocorre	Não ocorre	Não ocorre

APÊNDICE D - Ocorrência das lexias de buscas nos corpora da pesquisa

Lexias de busca	Corpus Textual Especializado	VCB	LexMI-Jurisprudência
INSS	Ocorre	ocorre como não autorizado para a forma	Ocorre. 25755 ocorrências.
instituto nacional do seguro social	Ocorre	Ocorre	Ocorre. 1857 ocorrências.
Doença ocupacional	não ocorre	ocorre não autorizado	6133 ocorrencias
DOENÇAS OCUPACIONAIS	Não, mas ocorre doença profissional e doença do trabalho	Não ocorre	Doença profissional 7063 ocorrências. Doença do trabalho 886 ocorrências.
seguro desemprego	Não ocorre	Não ocorre	5731 ocorrências
seguro-desemprego	Ocorre	ocorre	5731 ocorrências
sobreaviso	Ocorre	ocorre	4704 ocorrências.
tempo a disposicao	Não ocorre	Não ocorre	ocorre. 5099 ocorrências.
vinculo eventual fazenda	Não ocorre	Não ocorre	Não ocorre
vinculo eventual rural	Não ocorre	Não ocorre	Não ocorre
vinculo diarista fazenda	Não ocorre	Não ocorre	Não ocorre
vinculo diarista rural	Não ocorre	Não ocorre	diarista rural. 1 ocorrência em Acórdão do TRT 12°.
vínculo empregatício	Ocorre	Ocorre	ocorre. 44640 ocorrências
vínculo empregatista	Não ocorre	Não ocorre	Não ocorre
Vínculo de emprego	Ocorre	Não ocorre	Ocorre. 42846 ocorrências.
motorista de caminhão	Não, mas ocorre motorista de transporte rodoviário de cargas e motorista profissional	Ocorre como autorizado	1008 ocorrencias
motorista intermunicipal	Não ocorre	Não ocorre	13 ocorrencias

APÊNDICE D - Ocorrência das lexias de buscas nos corpora da pesquisa

Lexias de busca	Corpus Textual Especializado	VCB	LexMI-Jurisprudência
motorista	Não ocorre apenas motorista profissional	ocorre	11729 ocorrencias
Motoristas	não ocorre	Não ocorre	15129 ocorrencias
caminhoneiro	não ocorre	ocorre	98 ocorrencias
carreteiro intermunicipal	Não ocorre	Não ocorre	não ocorre
carreteiro	Não, mas ocorre motorista de transporte rodoviário de cargas e motorista profissional	Não ocorre	247 ocorrencias
bombeiro civil	Não ocorre	Não ocorre	70 ocorrencias
bombeiros	Não ocorre	ocorre Bombeiro	3778 ocorrencias
brigadista particular	Não ocorre	Não ocorre	1 ocorrência. Projeto de Lei do Senado nº 5358, de 2009. Altera o nome do cargo da profissão de "bombeiro civil" para "brigadista particular"
rescisão indireta	Não ocorre	Não ocorre	Ocorre. 2799 ocorrências.
Dispensa indireta por ato lesivo a honra	Não, mas ocorre despedida indireta	Não ocorre	35 ocorrências de dispensa indireta. 271 ocorrências para despedida indireta.

APÊNDICE D - Ocorrência das lexias de buscas nos corpora da pesquisa

Lexias de busca	Corpus Textual Especializado	VCB	LexMI-Jurisprudência
	De 98, Apenas 34 ocorrem exatamente como utilizado pelo usuário. Para 11 Lexias foi possível identificar variantes. E em um caso encontramos parte da lexia utilizada pelo usuário.	De 98, Apenas 28 lexias ocorrem exatamente da mesma forma empregada pelo usuário no VCB. Em 6 casos foram identificadas variantes	77 lexias ocorreram exatamente como utilizado pelo usuário. Apenas 7 não ocorrem exatamente igual a lexia do usuário. Logo em apenas 14 lexias não foram obtidos resultados para a busca.

APÊNDICE E – Lexias variantes classificadas de acordo com a Classificação Decimal de Direito

Lexias de busca	342.6 DIREITO DO TRABALHO	341.6 DIREITO PREVIDENCIÁRIO. SEGURIDADE SOCIAL
Doméstica	342.65183 Trabalho doméstico	341.612152 Empregado Doméstico
Domésticas	342.65183 Trabalho doméstico	341.612152 Empregado Doméstico
empregada doméstica	342.65183 Trabalho doméstico	341.612152 Empregado Doméstico
trabalhador doméstico	342.65183 Trabalho doméstico	341.612152 Empregado Doméstico
trabalhador doméstico	342.65183 Trabalho doméstico	341.612152 Empregado Doméstico
empregado domestico.	342.65183 Trabalho doméstico	341.612152 Empregado Doméstico
Diarista	342.65422 Diaristas	341.612153 Trabalhador avulso
Carteira de Trabalho	342.6521 Admissão. Carteira de Trabalho. Anotações	
CTPS	342.6521 Admissão. Carteira de Trabalho. Anotações	-
" CERTIDAO NEGATIVA TRABALHISTA"	342.6814 Tribunal Superior do Trabalho (TST)	-
" certidao negativa de debitos trabalhista"	342.6814 Tribunal Superior do Trabalho (TST)	-
" certidao negativa de debitos trabalhistas"	342.6814 Tribunal Superior do Trabalho (TST)	-
" cnd trabalhista"	342.6814 Tribunal Superior do Trabalho (TST)	-

APÊNDICE E – Lexias variantes classificadas de acordo com a Classificação Decimal de Direito

Lexias de busca	342.6 DIREITO DO TRABALHO	341.6 DIREITO PREVIDENCIÁRIO. SEGURIDADE SOCIAL
Cndt	342.6814 Tribunal Superior do Trabalho (TST)	-
""licença gestante"	342.655 Trabalho da mulher. Proteção à maternidade	
""LICENÇA MATERNIDADE""	342.655 Trabalho da mulher. Proteção à maternidade	
"licença-maternidade"	342.655 Trabalho da mulher. Proteção à maternidade	
"13° salario"	342.65438 13° Salário	
"decimo terceiro"	342.65438 13° Salário	
"abono pecuniario "	342.65432 Abonos	
"abono salarial"	342.65432 Abonos	
"acidente de serviço"	342.61 Acidentes do Trabalho	
"Acidente de trabalho"	342.61 Acidentes do Trabalho	
"Acidente percurso"	342.61 Acidentes do Trabalho	
"acidente trabalho"	342.61 Acidentes do Trabalho	
"salário-maternidade"	-	341.623533 Auxílio Maternidade
"Salário maternidade"	-	341.623533 Auxílio Maternidade
"Auxilio Doença"	-	341.623532 Auxílio Doença
"auxilio doenças"	-	341.623532 Auxílio Doença
"AUXÍLIO-DOENÇA"	-	341.623532 Auxílio Doença

APÊNDICE E – Lexias variantes classificadas de acordo com a Classificação Decimal de Direito

Lexias de busca	342.6 DIREITO DO TRABALHO	341.6 DIREITO PREVIDENCIÁRIO. SEGURIDADE SOCIAL
"auxílio-acidente"	-	341.623531 Auxílio Acidente. Auxílio Acidente de trabalho
"auxilio acidente"	-	341.623531 Auxílio Acidente. Auxílio Acidente de trabalho
"auxílio-alimentação"	-	342.65439 Salário-Educação. Auxílio -- alimentação
"auxílio alimentação"	-	342.65439 Salário-Educação. Auxílio - alimentação
"CLT"	342.69 Legislação do Trabalho	-
"consolidação das leis do trabalho"	342.69 Legislação do Trabalho	-
"consolidação das leis trabalhistas"	342.69 Legislação do Trabalho	-
"controle de horario"	342.622 Jornada de trabalho. Duração da jornada de trabalho. Interrupção obrigatória	-
"controle de ponto"	342.622 Jornada de trabalho. Duração da jornada de trabalho. Interrupção obrigatória	-
"Direito trabalhista"	342.6 DIREITO DO TRABALHO	-
"direito trabalho"	342.6 DIREITO DO TRABALHO	-
"direito do trabalho"	342.6 DIREITO DO TRABALHO	-
"direito processual do trabalho"	342.6 DIREITO DO TRABALHO	-
"epi"	342.617 Medidas preventivas dos acidentes de trabalho. Higiene e segurança do trabalho	-
"equipamento proteção"	342.617 Medidas preventivas dos acidentes de trabalho. Higiene e segurança do trabalho	-
"escala de plantão"	342.626	-

APÊNDICE E – Lexias variantes classificadas de acordo com a Classificação Decimal de Direito

Lexias de busca	342.6 DIREITO DO TRABALHO	341.6 DIREITO PREVIDENCIÁRIO. SEGURIDADE SOCIAL
	Repouso semanal remunerado. Duração da semana de trabalho	
"escala de serviço"	342.626 Repouso semanal remunerado. Duração da semana de trabalho	-
"escalas de revezamento"	342.626 Repouso semanal remunerado. Duração da semana de trabalho	-
"estabilidade gestante"	342.655 Trabalho da mulher. Proteção à maternidade	-
"estabilidade gravidez"	342.655 Trabalho da mulher. Proteção à maternidade	-
"execução trabalhista contra empregador falecido"	342.654 Obrigações do empregador. Cumprimento das condições do contrato. Salário. Remuneração do trabalho	-
"execução trabalhista contra executado falecido"	342.654 Obrigações do empregador. Cumprimento das condições do contrato. Salário. Remuneração do trabalho	-
"execução trabalhista contra herdeiro"	342.654 Obrigações do empregador. Cumprimento das condições do contrato. Salário. Remuneração do trabalho	-
"execução trabalhista contra herdeiros"	342.654 Obrigações do empregador. Cumprimento das condições do contrato. Salário. Remuneração do trabalho	-
"execução trabalhista de cujus"	342.654 Obrigações do empregador. Cumprimento das condições do contrato. Salário. Remuneração do trabalho	-
"férias dobra"	342.6253 Indenização de férias não gozadas. Pagamento em dobro	-
"férias em dobro"	342.6253 Indenização de férias não gozadas. Pagamento em dobro	-
"fgts"	342.6528 Fundo de Garantia por Tempo de Serviço (FGTS)	-

APÊNDICE E – Lexias variantes classificadas de acordo com a Classificação Decimal de Direito

Lexias de busca	342.6 DIREITO DO TRABALHO	341.6 DIREITO PREVIDENCIÁRIO. SEGURIDADE SOCIAL
"horas extras"	342.623 Trabalho extraordinário. Hora extra	-
"horas-extras"	342.623 Trabalho extraordinário. Hora extra	-
"INSS"	-	341.6242 Instituto Nacional do Seguro Social (INSS)
"instituto nacional do seguro social"	-	341.6242 Instituto Nacional do Seguro Social (INSS)
"pdv"	342.6526 Dissolução do contrato de trabalho. Proteção contra a despedida	-
"prêmio assiduidade"	342.65432 Abonos	-
"gratificação assiduidade"	342.65432 Abonos	-
"abono assiduidade"	342.65432 Abonos	-
"prescrição lesão de ofício"	342.617 Medidas preventivas dos acidentes de trabalho. Higiene e segurança do trabalho	-
"DOENÇAS OCUPACIONAIS"	342.617 Medidas preventivas dos acidentes de trabalho. Higiene e segurança do trabalho	-
"previdencia social"	-	341.624 Instituições de Previdência 341.6241 Instituto Nacional de Previdência Social (INPS)
"previdencia"	-	341.624 Instituições de Previdência 341.6241 Instituto Nacional de Previdência Social (INPS)
"seguro invalidez"	-	341.623526 Aposentadoria por Invalidez
"seguro desemprego"	-	341.6214 Benefício assistencial de um salário mínimo. Renda mínima

APÊNDICE E – Lexias variantes classificadas de acordo com a Classificação Decimal de Direito

Lexias de busca	342.6 DIREITO DO TRABALHO	341.6 DIREITO PREVIDENCIÁRIO. SEGURIDADE SOCIAL
"seguro-desemprego"	-	341.6214 Benefício assistencial de um salário mínimo. Renda mínima
"sobreaviso"	342.622 Jornada de trabalho. Duração da jornada de trabalho. Interrupção obrigatória	-
"tempo a disposicao"	342.622 Jornada de trabalho. Duração da jornada de trabalho. Interrupção obrigatória	-
"vinculo eventual fazenda"	342.65182 Trabalho rural 342.65187 Trabalho eventual. Trabalho avulso	-
"vinculo eventual rural"	342.65182 Trabalho rural 342.65187 Trabalho eventual. Trabalho avulso	-
"vinculo diarista fazenda"	342.65182 Trabalho rural 342.65187 Trabalho eventual. Trabalho avulso	-
"vinculo diarista rural"	342.65182 Trabalho rural 342.65187 Trabalho eventual. Trabalho avulso	-
"vínculo empregatício"	342.65 Contrato individual de Trabalho	-
"vínculo empregatista"	342.65 Contrato individual de Trabalho	-
"horas in itinere"	342.622 Jornada de trabalho. Duração da jornada de trabalho. Interrupção obrigatória	-
"horas itinere"	342.622 Jornada de trabalho. Duração da jornada de trabalho. Interrupção obrigatória	-
"motorista de caminhão"	Não há notação específica	Não há notação específica
carreteiro	Não há notação específica	Não há notação específica

APÊNDICE E – Lexias variantes classificadas de acordo com a Classificação Decimal de Direito

Lexias de busca	342.6 DIREITO DO TRABALHO	341.6 DIREITO PREVIDENCIÁRIO. SEGURIDADE SOCIAL
Motorista de ônibus	Não há notação específica	Não há notação específica
"bombeiro civil"	Não há notação específica	Não há notação específica
"bombeiros"	Não há notação específica	Não há notação específica
"brigadista particular"	Não há notação específica	Não há notação específica
professor de academia.	Não há notação específica	Não há notação específica
professor de ginastica.	Não há notação específica	Não há notação específica
OJ	Não há notação específica	Não há notação específica
Orientação Jurisprudencial	Não há notação específica	Não há notação específica
IN	Não há notação específica	Não há notação específica
Instrução Normativa	Não há notação específica	Não há notação específica

APÊNDICE F – Classificação do primeiro corpus de lexias com aspectos verbais

Quadro 22 – Alterações gráficas: primeiro corpus de lexias verbais

Tipo	Subtipo	Exemplos
	c) Termo e outra forma artificial	Nova lei dos motoristas/ Lei 12.619/2012.
Termo e abreviação	d) Termo e sigla	Carteira de Trabalho e Previdência Social / CTPS; Certidão negativa de débitos trabalhistas / CNDT; Instituto Nacional do Seguro Social / INSS; Consolidação das Leis do Trabalho / CLT; Fundo de Garantia do Tempo de Serviço / FGTS
	e) Termo e abreviação	Horas extras / Hrs extras; Certidão negativa de débitos trabalhistas / CND trabalhista;
Alteração ortográfica	- presença e ausência de acentuação	Férias / Ferias; Aviso prévio / aviso previo;
	- presença e ausência de hífen	Auxílio-alimentação / Auxílio alimentação; Licença-maternidade/ licença maternidade; Auxílio-doença / Auxílio doença; Auxílio-acidente / Auxílio acidente/ Salário-maternidade / Salário maternidade; Horas-extras / Horas extras; Seguro-desemprego / Seguro desemprego; Décimo-terceiro salário / Décimo terceiro salário;
	- erros de grafia como o acréscimo ou supressão de letras	Trabalhador doméstico / Trabalhador doméstico;
	- presença ou ausência de espaço entre as palavras	

Fonte: Elaborado pela autora.

APÊNDICE F – Classificação do primeiro corpus de lexias com aspectos verbais

Quadro 23 – Alterações morfosintáticas: primeiro corpus de lexias verbais

Tipo	Subtipo	Exemplos
Manutenção estrutura	a) ausência /presença de artigo	Acidente de trabalho / acidente do trabalho); visto que “do” é resultado da contração da preposição “de” + artigo definido “o”;
	b) alteração do substantivo	Execução trabalhista contra herdeiro / Execução trabalhista contra herdeiros; Doméstica/Domésticas; Auxílio doença / Auxílio doenças; Motorista / Motoristas; Doença ocupacional / Doenças ocupacionais
	c) alteração de preposição	Troca de uma preposição por outra preposição.
	d) alteração de gênero	Empregada doméstica / empregado doméstico.
Alteração estrutura	e) [Substantivo + Adjetivo] / [Substantivo + Sintagma preposicional]	Vínculo empregatício / vínculo de emprego; nesse caso o adjetivo geralmente é a parte alterada do termo;
	f) Monoléxico / Poliléxico	
	g) Outras alterações de estrutura.	

Fonte: Elaborado pela autora.

APÊNDICE F – Classificação do primeiro corpus de lexias com aspectos verbais

Quadro 24 – Reduções de extensão: primeiro corpus de lexias verbais

Subtipo	Exemplos
a) Tipo [Substantivo + X] = [Substantivo].	
- [Substantivo + Adjetivo]	Motorista intermunicipal / Motorista; Carreteiro intermunicipal / Carreteiro;
- [Substantivo + Sintagma preposicional]	Motorista de caminhão / Motorista;
-[Substantivo 1+ Substantivo 2] = Substantivo	
b) [Substantivo + X + Y] = [Substantivo + X].	
- [Substantivo + Sintagma preposicional + Adjetivo] = [Substantivo + Sintagma preposicional]	
- [Substantivo + Adjetivo + Sintagma preposicional] = [Substantivo + Adjetivo]	Certidão negativa de débitos/ Certidão negativa
- [Substantivo + Adjetivo 1 + Adjetivo 2] = [Substantivo + Adjetivo 1]	
- [Substantivo + Sintagma preposicional 1 + Sintagma preposicional 2] = [Substantivo + Sintagma preposicional 1]	Fundo de Garantia do Tempo de Serviço / Fundo de Garantia Certidão negativa trabalhista / Certidão negativa
c) Outras reduções de extensão	Carteira de Trabalho e Previdência Social / Carteira de Trabalho. Décimo terceiro salário / Décimo terceiro;

Fonte: Elaborado pela autora.

APÊNDICE F – Classificação do primeiro corpus de lexias com aspectos verbais

Quadro 25 – Reduções de base: primeiro corpus de lexias verbais

Subtipo	Exemplos
d) [Substantivo + Adjetivo] = [Adjetivo] N	
e) [Substantivo 1 + Substantivo 2] = [Substantivo 2]	Adicional insalubridade / Insalubridade; Adicional periculosidade / Periculosidade;
f) [Substantivo 1 + Sintagma preposicional (de + Substantivo 2)] = [Substantivo 2]	Adicional de periculosidade/ periculosidade; Adicional de insalubridade / insalubridade; Abono de férias / férias;
g) Outras reduções	

Fonte: Elaborado pela autora

Quadro 26 – Alterações Léxicas de unidades monoléticas: primeiro corpus de lexias verbais

Subtipo	Exemplos
Unidades Monoléticas	Caminhoneiro/Carreteiro

Fonte: Elaborado pela autora.

APÊNDICE F – Classificação do primeiro corpus de lexias com aspectos verbais

Quadro 27 – Alterações Léxicas de unidades poliléxicas: primeiro corpus de lexias verbais

Subtipo	Exemplos
a) Alteração de base	
- [Substantivo + Sintagma Preposicional (prep. + art. + Substantivo)]	
- [Substantivo + Adjetivo]	Empregado doméstico / Trabalhador doméstico; Rescisão indireta / Dispensa indireta; Motorista intermunicipal / carreteiro intermunicipal;
- [Outras alterações de base]	Competência da Justiça do Trabalho / Responsabilidade da Justiça do Trabalho
b) Alteração de extensão	
- [Substantivo + Sintagma Preposicional (prep. + art. + Substantivo)]	Acidente de trabalho/ Acidente de serviço; Escala de plantão / escala de serviço;
- [Substantivo + Adjetivo]	Licença maternidade / licença gestante; Estabilidade gestante / estabilidade gravidez; Vínculo empregatício / vínculo empregatista;
- [Outras alterações de extensão]	Consolidação das leis do trabalho / Consolidação das leis trabalhistas; Vínculo diarista fazenda / vínculo diarista rural; Vínculo eventual fazenda/ vínculo eventual rural; Execução trabalhista contra empregador falecido / Execução trabalhista contra executado falecido; Acidente de trabalho / Acidente de serviço;

Fonte: Elaborado pela autora.

APÊNDICE F – Classificação do primeiro corpus de lexias com aspectos verbais

Quadro 28 – Várias alterações complexas: primeiro corpus de lexias verbais

Tipo	Subtipo	Exemplos
Com parentesco formal	a) Monoléxica/Poliléxica	Bombeiro civil / Bombeiros;
	b) Poliléxica/Poliléxica	Gratificação férias abono / Abono de férias. Consolidação das leis do trabalho / Consolidação das leis trabalhistas;
Sem parentesco formal	c) Monoléxica/Poliléxica	Sobreaviso / Tempo à disposição; Caminhoneiro / Motorista de caminhão
	d) Poliléxica/Poliléxica	Execução trabalhista contra empregador falecido / execução trabalhista de cujos; Bombeiro civil / Brigadista particular. Gratificação natalina /Décimo-terceiro salário; Férias dobra / Férias em dobro;

Fonte: Elaborado pela autora.

APÊNDICE G – Classificação do segundo corpus de lexias verbais

Quadro 29 – Alterações gráficas: segundo corpus de lexias verbais

(continua)

Tipo	Subtipo	Exemplos
Termo e abreviação	d) Termo e sigla	Carteira de Trabalho e Previdência Social / CTPS; Certidão negativa de débitos trabalhistas / CNDT; Instituto Nacional do Seguro Social / INSS; Consolidação das Leis do Trabalho / CLT; Fundo de Garantia do Tempo de Serviço / FGTS Orientação Jurisprudencial / OJ; 12c Tribunal Regional do Trabalho / TRT; Tribunal Superior do Trabalho / TST; Certidão negativa de débitos trabalhistas / CND trabalhista; Devolução da Carteira de Trabalho / Devolução da CTPS; Mínimo regional Rio Grande do Sul / Mínimo regional RS; Presidente sociedade anônima empregado / Presidente S/A empregado;
	e) Termo e abreviação	Horas extras / Hrs extras; Piso salarial dos agentes comunitários de saúde / Piso salarial agente com. de saúde;
Alteração ortográfica	- presença e ausência de acentuação	Férias / Ferias; Aviso prévio / aviso previo; Certidão negativa de débitos trabalhistas / Certidao negativa de debitos trabalhistas; Assédio / Assedio; Aviso prévio / Aviso previo; Carga horária / Carga horaria; 12c Contribuição previdenciária / Contribuição previdenciaria; Contribuição assistencial / Contribuição assistencial; Desvio de função / Desvio de funcao; Diárias / Diarias; Rescisão indireta / Rescisao indireta; Rescisão / Rescisao;
	- presença e ausência de hífen	Auxílio-alimentação / Auxílio alimentação;

APÊNDICE G – Classificação do segundo corpus de lexias verbais

Tipo	Subtipo	Exemplos
Alteração ortográfica	-presença e ausência de hífen	Licença-maternidade/ licença maternidade; Auxílio-doença / Auxílio doença; Auxílio-acidente / Auxílio acidente/ Salário-maternidade / Salário maternidade; Horas-extras / Horas extras; Seguro-desemprego / Seguro desemprego; Décimo-terceiro salário / Décimo terceiro salário; Contra cheque / Contra-cheque ; Trabalho infanto-juvenil / Trabalho infantojuvenil; 10c
	- erros de grafia como o acréscimo ou supressão de letras	Trabalhador doméstico / Trabalador doméstico; Enquadramento sindical / Encuadramento sindical; Licença / Licenca; 4c Lei que exeta imposto de pessoas portadoras de doenças grave / Lei que esta imposto de pessoas portadoras de doenças grave;
	- presença ou ausência de espaço entre as palavras	Contracheque / Contra cheque;

Fonte: Elaborado pela autora.

(Fim)

APÊNDICE G – Classificação do segundo corpus de lexias verbais

Quadro 30 – Alterações morfossintáticas: segundo corpus de lexias verbais

(continua)

Tipo	Subtipo	Exemplos
Manutenção estrutura	a) ausência /presença de artigo	Acidente de trabalho / acidente do trabalho); visto que “do” é resultado da contração da preposição “de” + artigo definido “o”; Fundo de Garantia do Tempo de Serviço / Fundo de Garantia e Tempo de Serviço; 2c
	b) alteração do substantivo	Execução trabalhista contra herdeiro / Execução trabalhista contra herdeiros; Doméstica/Domésticas; Auxílio doença / Auxílio doenças; Motorista / Motoristas; Doença ocupacional / Doenças ocupacionais Certidão negativa de débitos trabalhista / Certidão negativa de débitos trabalhistas; Acidente de trabalho / Acidentes de trabalho; Acumulo de função / Acumulo de funções; Lei do trabalho / Leis trabalhistas; Plano de saúde / Planos de saúde; Reembolso de despesa / Reembolso de despesas; Salário funcionário / Salário funcionários; 12c
	c) alteração de preposição	Troca de uma preposição por outra preposição. Segurança do trabalho / Segurança no trabalho; 1c
	d) alteração de gênero	Empregada doméstica / empregado doméstico. Ajuda de custo / Ajuda de custa; 2c

APÊNDICE G – Classificação do segundo corpus de lexias verbais

Alteração estrutura	da	e) [Substantivo + Adjetivo] / [Substantivo + Sintagma preposicional]	Vínculo empregatício / vínculo de emprego; nesse caso o adjetivo geralmente é a parte alterada do termo; 1c
		f) Monoléxico / Poliléxico	
		g) Outras alterações de estrutura.	Jornada de trabalho habitual / habitualidade da jornada de trabalho; Trabalho doméstico infantil / trabalho infantil doméstico; TST vale transporte dinheiro salarial / TST vale transporte salarial dinheiro; 3c

Fonte: Elaborado pela autora.

(Fim)

APÊNDICE G – Classificação do segundo corpus de lexias verbais

Quadro 31 – Reduções de extensão: segundo corpus de lexias verbais

(Continua)

Subtipo	Exemplos
a)Tipo [Substantivo + X] = [Substantivo].	
- [Substantivo + Adjetivo]	Motorista intermunicipal / Motorista; Carreteiro intermunicipal / Carreteiro; Gratificação assiduidade / Gratificação; Assédio moral / Assédio; Assédio sexual / Assédio; Aviso prévio / Aviso; Bombeiro civil / Bombeiro; Deficiente físico / Deficiente; Dissídio coletivo / Dissídio; Empregada doméstica / Empregada; Empregado rural / Empregado; Equiparação salarial / Equiparação; Estabilidade provisória / Estabilidade; Férias atrasadas / Férias; Intervalo intrajornada / Intervalo; Justa causa verbal / Justa causa; Multa rescisória / Multa; Reajuste salarial / Reajuste; Rescisão indireta / Rescisão; Segurança privada / Segurança; Trabalhador rural / Trabalhador; Trabalho doméstico / trabalho; Trabalho infantil / trabalho; Trabalho infanto-juvenil / Trabalho;
- [Substantivo + Sintagma preposicional]	Motorista de caminhão / Motorista; Acidente de trabalho / Acidente;

24c

APÊNDICE G – Classificação do segundo corpus de lexias verbais

	<p>Acidente de percurso / Acidente; Adicional por tempo de serviço / Adicional; Aposentadoria por tempo de serviço / Aposentadoria; Assistente de juiz / Assistente; Carteira de trabalho / Carteira; Consignação em pagamento / Consignação; Contrato de trabalho / Contrato; Desvio de função / Desvio; Jogador de futebol / Jogador; Jornada de Trabalho / Jornada; Lei do trabalho / Lei; Reembolso de despesas / Reembolso; Vigilância em saúde / Vigilância;</p>	15c
-[Substantivo 1+ Substantivo 2] = Substantivo	<p>Aprendiz estabelecimento / Aprendiz; Auxílio-alimentação / auxílio; Auxílio-acidente / auxílio; Auxílio-doença / auxílio; Idoso Legislação / Idoso; Seguro-desemprego / Seguro; Vale transporte dinheiro / Vale transporte; Vale transporte / Vale;</p>	8c
b) [Substantivo + X + Y] = [Substantivo + X].		
- [Substantivo + Sintagma preposicional + Adjetivo] = [Substantivo + Sintagma preposicional]	<p>Adicional de periculosidade motoboy / adicional de periculosidade Férias em dobro doméstico / Férias em dobro; Jornada de trabalho habitual / Jornada de trabalho;</p>	3c
- [Substantivo + Adjetivo + Sintagma preposicional] = [Substantivo + Adjetivo]	<p>Certidão negativa de débitos/ Certidão negativa; Dano moral por calúnia / Dano moral; Rescisão indireta contrato de trabalho / Rescisão indireta;</p>	3c

APÊNDICE G – Classificação do segundo corpus de lexias verbais

<p>- [Substantivo + Adjetivo 1 + Adjetivo 2] = [Substantivo + Adjetivo 1]</p>	<p>Estabilidade provisória gestante / Estabilidade provisória; Estabilidade provisória gravidez / Estabilidade provisória; FGTS prescrição trintenar / FGTS prescrição; Recurso ordinário trabalhista / Recurso ordinário; Trabalho doméstico infantil / trabalho doméstico; Trabalho infantil doméstico / Trabalho infantil;</p> <p style="text-align: right;">6c</p>
<p>- [Substantivo + Sintagma preposicional 1 + Sintagma preposicional 2] = [Substantivo + Sintagma preposicional 1]</p>	<p>Fundo de Garantia do Tempo de Serviço / Fundo de Garantia Certidão negativa trabalhista / Certidão negativa; Aposentadoria por invalidez contrato de trabalho / Aposentadoria por invalidez; Plano de saúde para aposentados / Plano de saúde;</p> <p>4c</p>
<p>c) Outras reduções de extensão</p>	<p>Carteira de Trabalho e Previdência Social / Carteira de Trabalho. Décimo terceiro salário / Décimo terceiro; Certidão negativa de débitos trabalhistas / Certidão Ações conexas trabalhistas / Ações conexas; Aposentadoria especial deficiente físico / Aposentadoria especial; Aposentadoria segurado especial rural / Aposentadoria segurado especial; Adicional insalubridade vigilante / adicional insalubridade; Assédio moral no trabalho / Assédio moral; Assédio sexual no trabalho / Assédio sexual; Assédio sexual responsabilidade da empresa / Assédio sexual no trabalho; Benefício assistência social / Benefício; Cálculo em prática trabalhista / Cálculo trabalhista; Consolidação das leis do trabalho / Consolidação das leis; Dano moral cárcere privado em local de trabalho / Dano moral cárcere privado; Dano moral por calúnia quantum / Dano moral por calúnia;</p>

APÊNDICE G – Classificação do segundo corpus de lexias verbais

	Despedida discriminatória restritiva / Despedida discriminatória; Fraude na marcação de ponto falta grave / Fraude na marcação de ponto; Fundo de Garantia do Tempo de Serviço / Fundo Trabalho; Fundo de Garantia do Tempo de Serviço / Fundo; Jornada de trabalho inferior CTPS / Jornada de trabalho inferior; Justa causa verbal reversão / Justa causa verbal Licença não remunerada / Licença; Ônus da prova trabalho / Ônus da prova; Paradigma diferença salarial / Paradigma; Piso salarial dos agentes comunitários de saúde / Piso; Portador de HIV deficiente físico / Portador de HIV deficiente; Reajuste benefício Regime próprio / Reajuste benefício; Rescisão indireta por não depositar FGTS / Rescisão indireta; Vínculo empregatista de pastor / Vínculo empregatista; Licença maternidade natimorto / licença maternidade 30c
--	--

Fonte: Elaborado pela autora.

(Fim)

Quadro 32 – Reduções de base: segundo corpus de lexias verbais

(continua)

Subtipo	Exemplos	
d) [Substantivo + Adjetivo] = [Adjetivo] N	Trabalhador temporário / Temporário; Empregada doméstica / Doméstica; Intervalo intrajornada / Intrajornada ; Jornada de motoristas / Motoristas; Médico veterinário / Veterinário	5c
e) [Substantivo 1 + Substantivo 2] = [Substantivo 2]	Adicional insalubridade / Insalubridade; Adicional periculosidade / Periculosidade;	2c
f) [Substantivo 1 + Sintagma preposicional (de + Substantivo 2)] = [Substantivo 2]	Adicional de periculosidade/ periculosidade; Adicional de insalubridade / insalubridade; Abono de férias / férias;	3c
g) Outras reduções	Certidão negativa de débitos trabalhista / Certidão negativa trabalhista;	

APÊNDICE G – Classificação do segundo corpus de lexias verbais

	<p>Abandono de emprego empregado domestico / Abandono domestico; Abandono de posto vigilante justa causa / Abandono de posto justa causa; Aposentadoria especial portadores de deficiência física / Aposentadoria especial deficiente físico; Assistência médico hospitalar reembolso / Assistência hospitalar reembolso; Dano moral cárcere privado em local de trabalho / Cárcere privado no trabalho; Estabilidade provisória gestante / Estabilidade gestante; Estabilidade provisória gravidez / Estabilidade gravidez; Falta injustificada e desconto dsr / Falta injustificada dsr; Férias em dobro pagamento a menor dobra / Férias pagamento a menor dobra; Jornada de trabalho doméstico / Jornada doméstico; Natureza jurídica indenizatória do intervalo intrajornada / Natureza jurídica do intervalo intrajornada; Nulidade de pedido de demissão coação / Nulidade pedido de demissão; Penhora em conta salario / Penhora salario; Rescisão indireta – não fornecimento de vale transporte / Rescisão indireta – vale transporte; Rescisão indireta improcedência pedido de demissão / Rescisão indireta pedido de demissão; Rescisão indireta contrato de trabalho / Rescisão indireta; Resgates do seguro aposentado por invalidez / Resgates do seguro invalidez; Súmula vinculante, empregada doméstica / Súmula vinculante, doméstica; Trabalho doméstico infantil / Trabalho infantil; TST vale transporte dinheiro salarial / TST vale transporte salarial; TST vale transporte dinheiro salarial / Vale transporte dinheiro; Adicional periculosidade eventual / periculosidade eventual; Adicional de periculosidade motoboy / Periculosidade motoboy Fornecimento de EPI / EPI; Qual a forma de cálculo dos depósitos em atraso do FGTS / Qual a forma de cálculo do FGTS; Licença maternidade natimorto / Licença maternidade; 28c Qual a forma de cálculo dos depósitos em atraso do FGTS / Qual a forma de cálculo do FGTS;</p>
--	--

Fonte: Elaborado pela autora.

APÊNDICE G – Classificação do segundo corpus de lexias verbais

Quadro 33 - Alterações léxicas de unidades monoléxicas: segundo corpus de lexias verbais

Subtipo	Exemplos
Unidades Monoléxicas	Caminhoneiro/Carreteiro Exoneração / Rescisão; Falecimento/ Morte; Remuneração / Salário

Fonte: Elaborado pela autora.

Quadro 34 – Alterações léxicas de unidades poliléxicas: segundo corpus de lexias verbais

Subtipo	Exemplos
a) Alteração de base	
- [Substantivo + Sintagma Preposicional (prep. + art. + Substantivo)]	
- [Substantivo + Adjetivo]	Empregado doméstico / Trabalhador doméstico; Rescisão indireta / Dispensa indireta; Motorista intermunicipal / carreteiro intermunicipal; Abono assiduidade / Prêmio assiduidade; 4c
- [Outras alterações de base]	Competência da Justiça do Trabalho / Responsabilidade da Justiça do Trabalho 1c
b) Alteração de extensão	
- [Substantivo + Sintagma Preposicional (prep. + art. + Substantivo)]	Acidente de trabalho/ Acidente de serviço; Escala de plantão / escala de serviço; Controle de horário / Controle de ponto; Professor de ginástica / Professor de academia; Uso de sanitários / Uso do banheiro; 5c
- [Substantivo + Adjetivo]	Licença maternidade / licença gestante; Vínculo empregatício / vínculo empregatista; Abono pecuniário / Abono salarial Gratificação Assiduidade / Prêmio assiduidade; INSS safra / INSS safreiros; INSS safra / INSS safristas; Previdência Social / Previdência; 7c

APÊNDICE G – Classificação do segundo corpus de lexias verbais

- [Outras alterações de extensão]	Consolidação das leis do trabalho / Consolidação das leis trabalhistas; Vínculo diarista fazenda / vínculo diarista rural; Vínculo eventual fazenda/ vínculo eventual rural; Execução trabalhista contra empregador falecido / Execução trabalhista contra executado falecido; Escala de plantão / Escala de serviço; Acidente de trabalho / Acidente de serviço; Adicional de periculosidade de eletricitista / Adicional de periculosidade e eletricitista; Discriminação racial no emprego / Discriminação racial no trabalho; Exercia duas funções / Exercia dupla função; Licença maternidade morto / Licença maternidade natimorto; Presunção de insalubridade dentista / Presunção de insalubridade odontólogo; Projeto de lei que quer aumentar jornada dos aeronautas / Projeto de lei que quer aumentar jornada dos pilotos;	12c
-----------------------------------	---	-----

Fonte: Elaborado pela autora.

Quadro 35 – Várias alterações complexas: segundo corpus de lexias verbais

(continua)

Tipo	Subtipo	Exemplos
Com parentesco formal	a) Monoléxica/Poliléxica	Mão de obra temporária / temporário; Pessoa com deficiência / deficiente; Bombeiro Civil / Bombeiros;
	b) Poliléxica/Poliléxica	Gratificação férias abono / Abono de férias. Férias dobra / Férias em dobro; Pessoa com deficiência / Deficiente físico; Portador de deficiência / Deficiente físico; Consolidação das leis do trabalho / Consolidação das leis trabalhistas; Direito do Trabalho / Direito trabalhista Assistência médico hospitalar reembolso / Assistência médica reembolso;

APÊNDICE G – Classificação do segundo corpus de lexias verbais

		<p>Discriminação do negro no trabalho / Discriminação racial no emprego; Discriminação do negro no trabalho / Discriminação racial no trabalho; Licença sem remuneração interesse / Licença não remunerada; Pec da obrigatoriedade do curso de jornalista / Pec dos jornalistas; Pec da obrigatoriedade do curso de jornalista / Pec da exigencia de curso de jornalista; Pessoa com deficiência / Portador de deficiência; Repouso compensado semanal / Repouso semanal remunerado Periculosidade vigilante / Risco de vida vigilantes. 19c Legislação motoboy / Nova lei para motoboys; Dano moral por cárcere privado / Dano moral por ficar preso; Licença sem remuneração interesse / Licença não remunerada Projeto de lei que quer aumentar jornada dos aeronautas / Projeto de lei que quer aumentar jornada dos pilotos;</p>
Sem parentesco formal	c) Monoléxica/Poliléxica	<p>Sobreaviso / Tempo à disposição; Caminhoneiro / Motorista de caminhão Caminhoneiro / Carreteiro intermunicipal Carreteiro / Motorista de caminhão Falecimento / Licença nojo; 5c</p>
	d) Poiléxica/Poliléxica	<p>Execução trabalhista contra empregador falecido / execução trabalhista de cujos; Oficial de justiça horário de trabalho / oficial de justiça jornada; Bombeiro civil / Brigadista particular. Gratificação natalina /Décimo-terceiro salário; Gratificação natalina / Décimo terceiro; Aumento da jornada de trabalho habitual / Aumento carga horária; 6c</p>

Fonte: Elaborado pela autora.

APÊNDICE H – Variantes denominativas validadas em três tesouros

(Continua)

ALTERAÇÕES GRÁFICAS		Tesouro STJ	TEJUT 2012	TRT 4
1. Termo e forma artificial	a) Termo e símbolo			
	b) Termo e fórmula química			
	c) Termo e outra forma artificial	<i>Décimo terceiro salário/ 13º salário</i>	<i>Decimo terceiro salario/13 salario</i>	<i>Decimo terceiro salario/13 salario</i>
2. Termo e abreviação	d) Termo e Sigla	<i>CTPS/ Carteira de Trabalho e Previdência Social</i>	<i>CTPS/ Carteira de Trabalho e Previdencia Social</i>	<i>CTPS/ Carteira de Trabalho e Previdencia Social</i>
	e) Termo e abreviatura			
3. Alteração ortográfica		auxílio-alimentação/auxílio alimentação;		
ALTERAÇÕES MORFOSINTÁTICAS				
1. Mesma estrutura	a) Ausência / presença de artigo	<i>Acidente de trabalho/ acidente do trabalho.</i>		
	b) Alteração do substantivo			
	c) Alteração de preposição	<i>Acidente em Serviço/Acidente de serviço;</i>		
	d) Alteração de gênero			
2. Alteração de estrutura	e) [N+A] / [N+SP]			
	f) Monoléxico / poliléxico			
	g) Outras alterações de estrutura			

APÊNDICE H – Variantes denominativas validadas em três tesouros

		Tesouro STJ	TEJUT 2012	TRT 4
REDUÇÕES				
1. Reduções da extensão	a) Tipo $[N+X] = [N]$ <ul style="list-style-type: none"> ▪ $[N+A] = [N]$ ▪ $[N+SP] = [N]$ $[N_1+N_2] = [N_1]$		Carteira do Trabalho	Carteira do Trabalho
	b) Tipo $[N+X+Y] = [N+X]$ <ul style="list-style-type: none"> • $[N+SP+A] = [N+SP]$ • $[N+A+SP] = [N+A]$ • $[N+A_1+A_2] = [N+A_1]$ $[N+SP_1+SP_2] = [N+SP_1]$			
	c) Outras reduções da extensão			
2. Reduções da base	d) $[N+A] = [A]_N$			
	e) $[N_1+N_2] = [N_2]$			
	f) $[N_1+SP (de+N_2)] = [N_2]$			
3. Outras reduções				
ALTERAÇÕES LÉXICAS				
1. Unidades monoléticas		<i>Salário/ordenado</i>	<i>Salario/ordenado</i>	<i>Salario/ordenado</i>
2. Unidades poliléticas	a) Alteração na base <ul style="list-style-type: none"> ▪ $[N+SP (prep + art + N)]$ ▪ $[N+A]$ outras alterações na base	<i>Relação de emprego/vinculo de emprego</i>	<i>Relação de emprego/vinculo de emprego</i>	<i>Relação de emprego/vinculo de emprego</i>
		<i>Empregado doméstico /trabalhador doméstico</i>	<i>Empregado domestico /trabalhador domestico</i>	<i>Empregado domestico /trabalhador domestico</i>

APÊNDICE H – Variantes denominativas validadas em três tesouros

		Tesouro STJ	TEJUT 2012	TRT 4
	b) Alteração na extensão <ul style="list-style-type: none"> ▪ [N+SP (pre + art + N)] ▪ [N+A] ▪ Outras alterações na extensão 		<i>Licença gestante/ licença maternidade</i>	<i>Licença gestante/ licença maternidade</i>
VARIAS ALTERAÇÕES COMPLEXAS				
1. Com parentesco formal	a) Monoléxico / poliléxico			
	b) Poliléxico / poliléxico	<i>Licença à gestante/ licença-maternidade; Licença à gestante/ licença-gestante</i>	<i>Doença profissional/ doença ocupacional</i>	<i>Doença profissional/ doença ocupacional</i>
2. Sem parentesco formal	c) Monoléxico / poliléxico		<i>Sobreaviso/horário a disposição Trabalhador eventual/ biscateiro</i>	<i>Sobreaviso/horário a disposição; Trabalhador eventual/ biscateiro</i>
	d) Poliléxico / poliléxico	<i>Décimo terceiro salário/Gratificação de Natal; Décimo terceiro salário/Gratificação Natalina; Relação de emprego/vinculo empregatício</i>	<i>Relação de emprego/vinculo empregatício;</i>	<i>Horário a disposição/regime de sobreaviso; Relação de emprego/vinculo empregatício;</i>

(Fim)

APÊNDICE I - Validação do especialista

	Lexias de busca
1	Certidão negativa trabalhista
2	cnd trabalhista
3	gratificação férias abono
4	acidente trabalho
5	direito trabalho
6	execução trabalhista contra empregador falecido
7	execução trabalhista contra executado falecido
8	execução trabalhista contra herdeiro
9	execução trabalhista contra herdeiros
10	execução trabalhista de cujus
11	férias dobra
12	Fundo de garantia e tempo de serviço
13	hrs extras
14	DOENÇAS OCUPACIONAIS
15	vinculo eventual fazenda
16	vinculo eventual rural
17	vinculo diarista fazenda
18	vinculo diarista rural
19	vínculo empregatista
20	carreteiro intermunicipal
21	Dispensa indireta por ato lesivo a honra

APÊNDICE J - Lexias validadas no primeiro corpus textual especializado

	Lexias de busca	Corpus Textual Especializado
1	Empregada doméstica	Ocorre
2	Empregado doméstico	Ocorre
3	Carteira de Trabalho	Ocorre
4	CTPS	Ocorre
5	Certidão negativa de débitos trabalhistas	Ocorre
6	Cndt	Ocorre
7	CND	ocorre
8	licença-maternidade	Ocorre
9	décimo-terceiro salário / 13º salário	Ocorre
10	gratificação natalina	Ocorre
11	abono de férias	Ocorre
12	férias	Ocorre
13	Acidente do trabalho	Ocorre
14	salário-maternidade	Ocorre
15	Auxílio-doença	Ocorre
16	auxílio-acidente	Ocorre
17	auxílio-alimentação	Não, ocorre apenas salário-utilidade
18	adicional de insalubridade	ocorre
19	insalubridade	ocorre
20	Adicional de periculosidade	Não ocorre, mas fica subentendido adicional de insalubridade ou de periculosidade
21	periculosidade	ocorre
22	Aviso prévio	ocorre
23	consolidação das leis do trabalho	Ocorre
24	CLT	Ocorre
25	direito do trabalho	Ocorre
26	direito processual do trabalho	Ocorre
27	escala de plantão	Não, mas ocorre escala de prontidão
28	escala de serviço	Ocorre

APÊNDICE J - Lexias validadas no primeiro corpus textual especializado

	Lexias de busca	Corpus Textual Especializado
29	estabilidade gestante	Não, ocorre apenas estabilidade provisória
30	execução trabalhista contra empregador falecido	Não, mas ocorre apenas morte do empregador
31	FGTS	Ocorre
32	Fundo de garantia e tempo de serviço	Ocorre Fundo de Garantia do tempo de Serviço
33	horas extras	Ocorre
34	INSS	Ocorre
35	instituto nacional do seguro social	Ocorre
36	DOENÇAS OCUPACIONAIS	Não, mas ocorre doença profissional e doença do trabalho
37	seguro-desemprego	Ocorre
38	sobreaviso	Ocorre
39	vínculo empregatício	Ocorre
40	Vínculo de emprego	Ocorre
41	motorista de caminhão	Não, mas ocorre motorista de transporte rodoviário de cargas e motorista profissional
42	motorista	Não ocorre apenas motorista profissional
43	carreteiro	Não, mas ocorre motorista de transporte rodoviário de cargas e motorista profissional
44	Orientação Jurisprudencial	Ocorre
45	Dispensa indireta por ato lesivo a honra	Não, mas ocorre despedida indireta
		De 98, Apenas 34 ocorrem exatamente como utilizado pelo usuário. Para 11 Lexias foi possível identificar variantes.

APÊNDICE K - Lexias validadas no VCB

	Lexias de busca	VCB
1	Empregado doméstico	ocorre como autorizado
2	Diarista	autorizado.
3	Carteira de Trabalho	ocorre como não autorizado
4	Licença Maternidade	ocorre como não autorizado
5	Licença gestante	ocorre como autorizado Licença à gestante
6	Décimo terceiro salário / 13º salário	Ocorre como autorizado
7	abono de férias	Ocorre é o autorizado
8	férias	ocorre como autorizado
9	Acidente do trabalho	Ocorre como autorizado
10	salário-maternidade	ocorre como autorizado
11	Auxílio-doença	Ocorre
12	adicional de insalubridade	ocorre
13	insalubridade	ocorre
14	adicional de periculosidade	ocorre
15	periculosidade	ocorre
16	Aviso prévio	ocorre
17	consolidação das leis do trabalho	Ocorre como autorizado.
18	CLT	termo.
19	Direito trabalhista	ocorre como não autorizado em direito do trabalho
20	direito do trabalho	Ocorre
21	execução trabalhista contra empregador falecido	Não, Ocorre morte do empregador
22	FGTS	ocorre como termo não autorizado para a forma completa do termo.
23	Fundo de garantia e tempo de serviço	Não Ocorre Fundo de Garantia do tempo de Serviço
24	horas extras	Não, mas ocorre no singular hora extra
25	INSS	termo.
26	instituto nacional do seguro social	Ocorre

APÊNDICE K - Lexias validadas no VCB

	Lexias de busca	VCB
27	Doença ocupacional	ocorre não autorizado
28	seguro-desemprego	ocorre
29	sobreaviso	ocorre
30	vínculo empregatício	Ocorre
31	motorista de caminhão	Ocorre como autorizado
32	motorista	ocorre
33	caminhoneiro	ocorre
34	bombeiros	Não mas ocorre no singular
35		De 98, Apenas 28 lexias ocorrem exatamente da mesma forma empregada pelo usuário no VCB. Em 6 casos foram identificadas variantes dos termos utilizados pelos usuários

APÊNDICE L - Lexias validadas na fonte especializada (Jurisprudência)

1	Lexias de busca	LexMI-Jurisprudência
2	Doméstica	7422 ocorrências
3	Domésticas	7422 ocorrências
4	Empregada doméstica	1056 ocorrências
5	Empregado doméstico	1056 ocorrências
6	Diarista	439 ocorrências
7	Carteira de Trabalho	1159 ocorrências.
8	CTPS	8186 ocorrências
9	Certidão negativa de débitos trabalhistas	4 ocorrências.
10	certidao negativa de debitos trabalhista	4 ocorrências.
11	Cndt	2 ocorrências
12	CND	8 ocorrencias relacionadas aos tribunais de Trabalho
13	Licença Maternidade	773 ocorrências
14	Licença gestante	129 ocorrências
15	licença-maternidade	773 ocorrências
16	Décimo terceiro salário / 13º salário	3399 ocorrências
17	Décimo terceiro	3709 ocorrências e décimo- terceiro salário
18	décimo-terceiro salário / 13º salário	3399 ocorrências
19	abono de férias	563 ocorrencias
20	férias	29557 ocorrencias
21	ferias	29557 ocorrencias
22	acidente de serviço	ocorre, porém em apenas uma decisão do TST
23	Acidente de trabalho	18779 ocorrências.
24	Acidente do trabalho	14308 ocorrências.
25	salário-maternidade	353 ocorrências
26	Salário maternidade	353 ocorrências
27	Auxilio Doença	6871 ocorrências.

APÊNDICE L - Lexias validadas na fonte especializada (Jurisprudência)

	Lexias de busca	LexMI-Jurisprudência
28	Auxilio Doenças	6871 ocorrências.
29	Auxilio-doença	6871 ocorrências.
30	auxílio-acidente	3690 ocorrências
31	auxílio acidente	3690 ocorrências
32	auxílio-alimentação	13494 ocorrências.
33	"auxílio alimentação	13494 ocorrências.
34	adicional de insalubridade	48623 ocorrencias
35	adicional insalubridade	149 ocorrencias
36	insalubridade	51195 ocorrencias
37	adicional de periculosidade	47132 ocorrencias
38	adicional periculosidade	178 ocorrencias
39	periculosidade	64503 ocorrencias
40	Aviso prévio	26846 ocorrencias
41	aviso previo	22426 ocorrencias
42	consolidação das leis do trabalho	44613 ocorrências.
43	consolidação das leis trabalhistas	560 ocorrências
44	CLT	654365 ocorrencias.
45	Direito trabalhista	ocorre 8027 ocorrências.
46	direito do trabalho	Ocorre 12081 ocorrências.
47	direito processual do trabalho	1298 ocorrências.
48	escala de plantão	90 ocorrências
49	escala de serviço	91 ocorrências.
50	estabilidade gestante	831 ocorrencias
51	estabilidade gravidez	9 ocorrencias
52	execução trabalhista contra empregador falecido	ocorre empregador falecido

APÊNDICE L - Lexias validadas na fonte especializada (Jurisprudência)

	Lexias de busca	LexMI-Jurisprudência
53	execução trabalhista de cujus	2243 ocorrências para De cujos. Tribunal Regional do Trabalho (193) Tribunal Superior do Trabalho (218)
54	férias dobra	203 itens como termos separados férias. Dobra . Dobra de férias 344 ocorrências
55	férias em dobro	926 ocorrências
56	FGTS	125140 ocorrências.
57	Fundo de garantia e tempo de serviço	1 ocorrência fundo de garantia do tempo de serviço
58	fundo de garantia	5263 ocorrências
59	horas extras	Ocorre mesmo resultado para o termo com hifem e também no singular 167996 ocorrências
60	horas-extras	ocorre
61	INSS	Ocorre. 25755 ocorrências.
62	instituto nacional do seguro social	Ocorre. 1857 ocorrências.
63	Doença ocupacional	6133 ocorrências
64	DOENÇAS OCUPACIONAIS	Doença profissional 7063 ocorrências. Doença do trabalho. 886 ocorrências.
65	seguro desemprego	5731 ocorrências
66	seguro-desemprego	5731 ocorrências
67	sobreaviso	4704 ocorrências.
68	tempo a disposicao	ocorre. 5099 ocorrências.
69	vinculo diarista rural	diarista rural. 1 ocorrência em Acórdão do TRT 12°.
70	vínculo empregatício	ocorre. 44640 ocorrências
71	Vínculo de emprego	Ocorre. 42846 ocorrências.
72	motorista de caminhão	1008 ocorrências
73	motorista intermunicipal	13 ocorrências

APÊNDICE L - Lexias validadas na fonte especializada (Jurisprudência)

	Lexias de busca	LexMI-Jurisprudência
		11729 ocorrencias
74	motorista	
75	Motoristas	15129 ocorrencias
76	caminhoneiro	98 ocorrencias
77	carreteiro	247 ocorrencias
78	bombeiro civil	70 ocorrencias
79	bombeiros	3778 ocorrencias
80	brigadista particular	1 ocorrência. Projeto de Lei do Senado nº 5358, de 2009. Altera o nome do cargo da profissão de "bombeiro civil" para "brigadista particular"
81	rescisão indireta	Ocorre. 2799 ocorrências.
82	Dispensa indireta por ato lesivo a honra	35 ocorrências de dispensa indireta. 271 ocorrências para despedida indireta.
		De 98, 77 lexias ocorreram exatamente como utilizado pelo usuário. Para 5 lexias encontramos parte da lexia utilizada pelo usuário. E para 2 lexias encontramos variantes da lexia do usuário. Logo para penas 14 lexias não foram obtidos resultados para a busca

APÊNDICE M – Supressão de preposição

Quadro 36 – Supressão de preposição

Tipo	Exemplos
Supressão de preposição	Acidente de trabalho / Acidente trabalho; Adicional de insalubridade / Adicional insalubridade; Adicional de periculosidade / adicional periculosidade; Direito do Trabalho / Direito Trabalho; Assédio moral no trabalho / Assédio moral trabalho; Dano moral por cárcere privado / Dano moral cárcere privado; Horas in itinere / Horas itinere; Multa do FGTS / Multa FGTS; Rescisão indireta pelo não fornecimento de vale transporte / Rescisão indireta – não fornecimento de vale transporte; Súmulas do TST / Súmulas TST

Fonte: Elaborado pela autora.

APÊNDICE N - Busca por campo semântico

Quadro 37 – Busca por campo semântico

Tipo	Subtipo	Exemplos
Busca por campo semântico	a) Troca de fenômeno por pessoa	Gravidez / Gestante; Estabilidade gravidez / Estabilidade gestante ; Renda aposentadoria /Renda aposentado; Deficiência física / Deficiente físico; Licença maternidade morte / Licença maternidade morto; Estabilidade provisória gestante / Estabilidade provisória gravidez;
	b) Troca de fenômeno por objeto	Pré-contratação / Pré-contrato Rasurar a CTPS / Rasura na CTPS
	c) Termos Antônimos (oposição)	Sentar no trabalho / Trabalhar em pé Jornada de trabalho habitual / mudança de horário no trabalho;

Fonte: Elaborado pela autora.

OBS.: Busca por campo semântico (nesses casos há uma proximidade de sentido entre os termos mas não podemos considerá-los como termos equivalentes (variantes denominativas) do ponto de vista terminológico, visto que apresentam definições diferentes.

ANEXO A – Classificação Decimal de Direito: Direito do Trabalho

342.6	DIREITO DO TRABALHO
342.601	Conceito. Teoria
342.6011	Fontes
342.6013	Interpretação
342.609	(Divisão geográfica, como 930-999)
342.61	Acidentes do Trabalho
342.611	Conceito de acidente de trabalho. Elementos para sua caracterização
342.612	Responsabilidade
342.613	Conseqüências do acidente de trabalho
342.6131	Morte
342.6132	Incapacidade total e permanente
342.6133	Incapacidade parcial e permanente
342.6134	Incapacidade temporária
342.614	Beneficiários
342.615	Indenização. Garantia da indenização
342.616	Seguro contra acidentes de trabalho
342.617	Medidas preventivas dos acidentes de trabalho. Higiene e segurança do trabalho
342.62	Duração do Trabalho
342.621	Regulamentação da duração do trabalho. Justificativas
342.622	Jornada de trabalho. Duração da jornada de trabalho. Interrupção obrigatória
342.623	Trabalho extraordinário. Hora extra
342.624	Trabalho noturno
342.625	Férias. Conceito
342.6251	Periodicidade. Anualidade
342.6252	Condições para aquisição do direito a férias. Período aquisitivo
342.6253	Indenização de férias não gozadas. Pagamento em dobro
342.6254	Férias coletivas
342.626	Repouso semanal remunerado. Duração da semana de trabalho
342.6261	Feriados e dias festivos
342.627	Lazer. Organização do lazer
342.6271	Colônias de férias
342.6272	Centros de recreação
342.6273	Centros culturais para utilização do lazer. Teatros. Bibliotecas
342.63	Acordo e Convenção coletiva de trabalho
342.6301	Conceito. Natureza jurídica. Teoria
342.631	Elementos da convenção coletiva. Eficácia, formação e conteúdo
342.632	Forma da convenção coletiva. Celebração e partes
342.633	Publicidade da convenção
342.634	Duração e vigência
342.635	Extensão da convenção coletiva
342.636	Comissões de conciliação prévia

ANEXO A – Classificação Decimal de Direito: Direito do Trabalho

342.64	Associações profissionais (Sindicatos, uniões, federações, confederações, coligações patronais e operários)
342.6401	Conceito. Natureza jurídica do sindicato. (Subdividir por países 930-999) Exemplo: 342.640981 Sindicalismo no Brasil
342.641	Liberdade sindical. Unidade e pluralidade sindical
342.6411	Sindicato livre
342.6412	Sindicato obrigatório
342.6413	Sindicato por profissões
342.6414	Sindicato por indústrias
342.6415	Sindicatos mistos
342.6416	Organização sindical
342.64161	Estatutos. Poderes do sindicato
342.64162	Direito de filiação
342.64163	Eleições sindicais
342.64164	Contribuições sindicais. Mensalidades
342.642	Associações profissionais. Órgão de fiscalização profissional
342.65	Contrato individual de Trabalho
342.6501	Teoria. Conceito. Natureza jurídica de relação de emprego
342.651	Modalidades de contrato de trabalho
342.6511	Contrato por tempo determinado
342.6512	Contrato por tempo indeterminado
342.6513	Contrato de trabalho manual
342.6514	Contrato de trabalho intelectual
342.6515	Contrato individual
342.6516	Contrato coletivo ou de equipe
342.6518	Contratos especiais de trabalho
342.65181	Aprendizagem
342.65182	Trabalho rural
342.65183	Trabalho doméstico
342.65184	Contrato de trabalho com entidades públicas
342.65185	Trabalho marítimo
342.65186	Trabalho aeronáutico
342.65187	Trabalho eventual. Trabalho avulso
342.65188	Trabalho autônomo
342.652	Início e fim do contrato de trabalho
342.6521	Admissão. Carteira de Trabalho. Anotações
342.6522	Alteração do contrato de trabalho
342.6523	Interrupção do contrato de trabalho
342.6524	Suspensão do contrato de trabalho
342.6525	Tempo de serviço
342.6526	Dissolução do contrato de trabalho. Proteção contra a despedida
342.6527	Estabilidade do empregado. Indenização
342.6528	Fundo de Garantia por Tempo de Serviço (FGTS)
342.6529	Aviso prévio
342.653	Obrigações do empregado
342.6531	Obrigação da prestação do trabalho

ANEXO A – Classificação Decimal de Direito: Direito do Trabalho

342.6532	Diligência, fidelidade, obediência
342.654	Obrigações do empregador. Cumprimento das condições do contrato. Salário. Remuneração do trabalho
342.6541	Conceito de salário. Composição. Caracteres
342.65411	Princípios do salário igual. Isonomia salarial
342.6542	Salário-base. Formas de pagamento de salário
342.65421	Horistas
342.65422	Diaristas
342.65423	Mensalistas
342.65424	Tarefeiros
342.65425	Empreiteiros
342.6543	Elementos suplementares do salário
342.65431	Indenizações por insalubridade ou periculosidade
342.65432	Abonos
342.65433	Participação nos lucros
342.65434	Gratificações. Gorjetas
342.65435	Comissões
342.65436	Prêmios de produtividade
342.65437	Salário-família
342.65438	13º Salário
342.65439	Salário-Educação. Auxílio -alimentação
342.6544	Salário-mínimo. Processo de determinação
342.65441	Salário-mínimo profissional
342.65442	Piso salarial
342.6545	Programa de Integração Social (PIS)
342.6546	Proteção geral do salário
342.65461	Irredutibilidade
342.65462	Proibição de cessão
342.65463	Impenhorabilidade
342.65464	Salário como crédito preferencial
342.655	Trabalho da mulher. Proteção à maternidade
342.656	Trabalho do menor
342.66	Conflitos coletivos do Trabalho
342.661	Greve. Direito de greve
342.6611	Greve legal
342.6612	Greve ilegal
342.662	Lock-out
342.663	Outras formas de conflitos
342.6631	Sabotagem
342.6632	Boicotagem
342.6633	Piquetagem
342.664	Solução dos conflitos coletivos
342.6641	Conciliação
342.6642	Arbitragem
342.6643	Mediação
342.6644	Processo coletivo de trabalho
342.6645	Sentença coletiva
342.68	Direito Processual do Trabalho
342.681	Organização Judiciária do Trabalho

ANEXO A – Classificação Decimal de Direito: Direito do Trabalho

342.6811	Juntas de conciliação e julgamento. Varas do trabalho
342.6812	Tribunais regionais
342.6813	Tribunais superiores
342.6814	Tribunal Superior do Trabalho (TST)
342.6815	Juizes do Trabalho
342.6816	Jurisdição e competência
342.6817	Penalidades
342.682	Ministério Público do Trabalho
342.6821	Procuradoria-Geral do Trabalho
342.6822	Procuradorias Regionais do Trabalho
342.69	Legislação do Trabalho (Dividir por países, como 930-999)

ANEXO B – Classificação Decimal de Direito – Direito Previdenciário

341.6	DIREITO PREVIDENCIÁRIO. SEGURIDADE SOCIAL
341.601	Conceito. Teoria
341.6011	Fontes
341.6012	Interpretação
341.609	(Divisão geográfica, como 930-999)
341.61	Seguridade Social
341.611	Organização da Seguridade Social
341.6111	Sistema Nacional de Seguridade Social
341.6112	Órgãos Colegiados
341.61121	Conselho Nacional de Previdência Social
341.61122	Conselho de Recursos da Previdência Social
341.611221	Composição
341.611222	Recursos
341.611223	Divulgação dos Atos e Decisões da Previdência Social
341.611224	Convênios. Contratos. Acordos
341.61123	Juntas de Recurso da Previdência Social
341.612	Custeio da Seguridade Social
341.6121	Financiamento da Seguridade Social
341.61211	Contribuição Facultativa
341.61212	Contribuição do Produtor Rural
341.61213	Contribuição da União
341.61214	Contribuição do Segurado Especial
341.61215	Contribuição do Segurado
341.612151	Empregado
341.612152	Empregado Doméstico
341.612153	Trabalhador avulso
341.612154	Contribuição Individual
341.612155	Contribuição sobre a Receita de Concursos, Prognósticos e Sweepstake
341.612156	Salário de Contribuição
341.612157	Contribuição do Empregador
341.612158	Contribuição do Empregado Doméstico
341.612159	Isenção de Contribuição
341.62	Previdência Social
341.621	Beneficiários
341.6214	Benefício assistencial de um salário mínimo. Renda mínima
341.622	Segurados
341.6221	Dependentes
341.6222	Manutenção e Perda da Qualidade de Segurado
341.623	Regime Geral da Previdência
341.6231	Plano de Benefícios da Previdência Social
341.6232	Prestações em Espécie
341.6233	Prestações em Geral
341.6234	Inscrições do segurado e dependente
341.6235	Benefícios
341.62351	Carência

ANEXO B – Classificação Decimal de Direito – Direito Previdenciário

341.62352	Aposentadorias e Pensões
341.623521	Aposentadoria Especial
341.623522	Aposentadoria por Idade
341.623523	Aposentadoria por velhice
341.623524	Aposentadoria por tempo de serviço
341.623525	Aposentadoria por tempo de Contribuição
341.623526	Aposentadoria por Invalidez
341.623527	Pensão por morte
341.62353	Auxílios
341.623531	Auxílio Acidente. Auxílio Acidente de trabalho
341.623532	Auxílio Doença
341.623533	Auxílio Maternidade
341.623534	Auxílio Reclusão
341.623535	Auxílio funeral
341.623536	Salário-família
341.623537	Abono Anual - equivalente ao 13º salário
341.623538	Habilitação e Reabilitação Profissional
341.623539	Pecúlio
341.6236	Reconhecimento da Filiação
341.62361	Reconhecimento do Tempo de Filiação
341.62362	Indenização
341.62363	Retroação da data do início da contribuição
341.62364	Contagem recíproca de tempo de contribuição
341.62365	Justificação administrativa
341.624	Instituições de Previdência
341.6241	Instituto Nacional de Previdência Social (INPS)
341.6242	Instituto Nacional do Seguro Social (INSS)
341.6243	Outras instituições de previdência social
341.625	Previdência Privada. Previdência Fechada
341.6251	Fundos de pensão
341.6252	Previdência Aberta. Seguros e mercados de capitais. Superintendência de Seguros Privados (SUSEP)
341.63	Assistência Social
341.631	Formas de Assistência
341.6311	Assistência Alimentar
341.6312	Assistência Financeira. Empréstimos
341.6313	Assistência Habitacional
341.63131	Banco Nacional da Habitação (BNH)
341.64	Saúde
341.641	Assistência Médica
341.6411	Sistema Único de Saúde (SUS)
341.6412	Fornecimento de medicamentos