



Programa de
Pós-Graduação em
Linguística

**ASPECTOS SEMÂNTICOS DA DEFINIÇÃO
TERMINOLÓGICA (DT): DESCRIÇÃO LINGÜÍSTICA E PROPOSTA DE
SISTEMATIZAÇÃO**

Dayse Simon Landim Kamikawachi

SÃO CARLOS

2009



Universidade Federal de São Carlos

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE EDUCAÇÃO E CIÊNCIAS HUMANAS
PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA

**ASPECTOS SEMÂNTICOS DA DEFINIÇÃO
TERMINOLÓGICA (DT): DESCRIÇÃO LINGUÍSTICA E PROPOSTA DE
SISTEMATIZAÇÃO**

Dayse Simon Landim Kamikawachi
Bolsista FAPESP

Dissertação apresentada ao Programa de Pós-Graduação em Linguística da Universidade Federal de São Carlos, como parte dos requisitos para a obtenção do Título de Mestre em Linguística.

Orientadora: Profa. Dra. Gladis Maria de Barcellos Almeida

São Carlos, São Paulo, Brasil
Abril – 2009

**Ficha catalográfica elaborada pelo DePT da
Biblioteca Comunitária da UFSCar**

K15as

Kamikawachi, Dayse Simon Landim.

Aspectos semânticos da definição terminológica (DT) :
descrição linguística e proposta de sistematização / Dayse
Simon Landim Kamikawachi. -- São Carlos : UFSCar, 2010.
161 f.

Dissertação (Mestrado) -- Universidade Federal de São
Carlos, 2009.

1. Linguística - processamento de dados. 2. Terminologia.
3. Linguística de *corpus*. 4. Linguagem humana e tecnologia.
5. Análise semântica. I. Título.

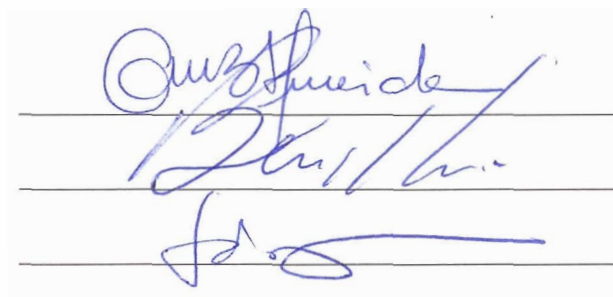
CDD: 410.285 (20ª)

BANCA EXAMINADORA

Profa. Dra. Gladis Maria de Barcellos Almeida

Prof. Dr. Bento Carlos Dias da Silva

Profa. Dra. Sandra Maria Aluisio



The image shows three handwritten signatures in blue ink on a white background with horizontal lines. The first signature is 'Gladis Maria de Barcellos Almeida', the second is 'Bento Carlos Dias da Silva', and the third is 'Sandra Maria Aluisio'.

À minha querida Anna Clara.

Agradecimentos

À Prof.^a Gladis Maria de Barcellos Almeida, por me apresentar ao fascinante mundo da Terminologia. Agradeço-lhe pela confiança depositada, pela orientação e principalmente pela amizade e apoio em vários momentos da minha trajetória.

À Prof.^a Sandra Maria Aluísio e ao Prof. Bento Carlos Dias da Silva, pela leitura atenta do trabalho e pelas contribuições muito produtivas dadas no exame de qualificação e na defesa.

Aos professores do PPGL: Prof. Oto Vale, pela indicação de autores e por assistir a exposição deste trabalho de mestrado por muitas vezes, sem perder, ainda assim, o olhar crítico, sempre disposto a contribuir com a pesquisa e com o meu crescimento acadêmico; e à Prof.^a Lucia Rino, pelo olhar de informata quanto à língua e pela exímia e desafiante aula.

Ao Marcelo Adriano Amancio, pesquisador do NILC, pela elaboração de dois programas computacionais específicos para esta pesquisa.

Aos meus colegas do GETerm - Jacqueline Souza, Daniela Mattos, Fernando Alves, Solange Bazzon, Leandro Oliveira - pelo diálogo e aprendizado constantes. Agradeço especialmente ao Douglas Pino, pela “velha” amizade, pelas discussões acerca da Linguagem e pela revisão deste trabalho de mestrado; e ao paciente e didático Joel Coleti, pela presteza na assessoria computacional fornecida especialmente quanto à metodologia empregada na pesquisa.

À secretária do PPGL, Nani, pela disposição e engajamento de levar o programa adiante e pela seriedade dedicada ao trabalho diário realizado, tornando a nossa passagem pelo Mestrado mais suave e consciente das nossas responsabilidades como aluno e pesquisador.

Aos meus amigos, pelo apoio e pela compreensão quanto à minha ausência nas reuniões, viagens, defesas, por diversas vezes. Em especial, às minhas amigas de disciplinas, risos e riscos: Maria Cristina Andrade, Cláudia Dias de Barros e Lívia Cucatto.

À FAPESP, pelo auxílio financeiro que possibilitou a realização deste trabalho.

Aos meus pais, Cleide e Edilson, e aos meus irmãos David e Debora, pelo amor dedicado e pela torcida em mais uma etapa da minha vida. Sou eternamente grata por tudo que fizeram e que ainda fazem por mim.

À minha sogra Aparecida por me propiciar condições de estudo e pela afeição dedicada ao longo dos anos.

Ao meu querido Everton, pela paciência, carinho, cumplicidade e dedicação durante todo esse tempo. Obrigada pelo seu sorriso, por me ensinar e por me amar.

À minha sempre Clarinha, a quem devo mais que um agradecimento. Dedico este trabalho a ela, que, embora seja uma pequenina, acompanhou muito de perto o percurso para a realização deste Mestrado.

Àquele que mesmo antes de nascer, me fez companhia.

Finalmente, agradeço a Deus, por cuidar de mim e por ter me dado forças para a conclusão deste trabalho.

A constituição de uma terminologia própria marca, em toda ciência, o advento ou o desenvolvimento de uma conceitualização nova, assinalando, assim, um momento decisivo de sua história. Poder-se-ia mesmo dizer que a história particular de uma ciência se resume na de seus termos específicos. Uma ciência só começa a existir ou consegue se impor na medida em que faz existir e em que impõe seus conceitos, através de sua denominação. Ela não tem outro meio de estabelecer sua legitimidade senão por especificar seu objeto denominando-o, podendo este constituir uma ordem de fenômenos, um domínio novo ou um modo novo de relação entre certos dados.

Benveniste, 1989.

'When I use a word,' Humpty Dumpty said, in rather a scornful tone, 'it means just what I choose it to mean -- neither more nor less.'
'The question is,' said Alice, 'whether you can make words mean so many different things.'
'The question is,' said Humpty Dumpty, 'which is to be master -- that's all.'

Lewis Carroll (*Alice's adventures in Wonderland*).

Resumo

A elaboração da definição é a etapa mais complexa e custosa numa pesquisa terminológica que objetiva a elaboração de dicionários especializados, já que um bom dicionário se avalia, principalmente, pela qualidade das suas definições. Assim como há diferença entre termo e palavra, também há diferenças entre os modos de definir termos e palavras. Enquanto a definição lexicográfica descreve e delimita os sentidos das palavras enquanto unidades do sistema, oferecendo todas as acepções possíveis, a definição terminológica (DT) oferece a identificação do termo somente com referência ao sistema conceitual do qual faz parte. É esse segundo tipo de definição, portanto, que é objeto desta pesquisa de mestrado. Na experiência adquirida durante nossa pesquisa terminológica realizada durante a Iniciação Científica, junto ao Grupo de Estudos e Pesquisas em Terminologia (GETerm), deparamo-nos com diversas dificuldades para a elaboração da DT. O método de trabalho foi por diversas vezes alterado, adaptado, ampliado, revisto, de forma a dar conta dos desafios que iam surgindo. A partir dessa experiência, chegamos a um conjunto de procedimentos metodológicos úteis para a redação da definição, em qualquer domínio de conhecimento. Pretendeu-se, pois, nesta pesquisa de mestrado, sistematizar esses procedimentos, de modo a proporcionar conhecimento linguístico que será implementado num ambiente computacional baseado na *Web*, de maneira a auxiliar demais grupos que desenvolvem pesquisas terminológicas em língua portuguesa a elaborar a DT de forma mais sistemática.

Palavras-chave: definição terminológica, *corpus*, relações semânticas

Abstract

The making of the definitions is the most complex step in terminological research which aims to compile specialized dictionaries, for the good ones are evaluated mainly as to the quality of their definitions. As there are differences between terms and words, there are also differences between defining them. While the lexicographic definition specifies and describes the meanings of words as language units, giving all their possible senses, the terminological definition (TD) identifies and describes terms only in reference to the specific conceptual system to which they belong. Therefore, the TD is the object of this Master's study. Our undergraduate experience in terminological research in the Group of Studies and Research in Terminology (GETerm), revealed several difficulties associated with the TD process. The defining method was changed, adapted, extended, and revised to handle the challenges of the research. From that experience, we came to a set of methodological procedures to aid the TD routines in any domain knowledge. Accordingly, this Master's dissertation research aims to systematize those procedures to be implemented into a Web-based, computational environment (E-Termos) to aid research groups in their terminological endeavor.

Keywords: terminological definition, *corpus*, semantic relations

Lista de Figuras

Figura 1 – Exemplo de DT no modelo GPDE.....	23
Figura 2 - Tipologia das relações conceituais possíveis (extraída de Seppälä, 2004).....	33
Figura 3 - Tela da Ontologia do DiRC	43
Figura 4 - Tela da Base Definicional do DiRC	45
Figura 5 - Tela do Scipo (Resumo – Seleção da Estrutura)	55
Figura 6 - Tela do e-Termos (Principal).....	58
Figura 7 - Ficha Terminológica do e-Termos.....	59
Figura 8 - Editor de Definição do e-Termos	60
Figura 9 - Exemplo de armazenamento do <i>corpus</i>	64
Figura 10 – Padrão de nomeação de arquivos	65
Figura 11 - Tela inicial do Editor de Cabeçalho.....	67
Figura 12 - Cabeçalho com etiquetas “xml” gerado pelo Editor de Cabeçalho	67
Figura 13 - Anotação linguística em “xml” gerado pelo Editor de Cabeçalho	69
Figura 14 - Telas “Concord” e “WordList” do programa <i>WordSmith Tools</i>	74
Figura 15 - Telas sobre “tags” do programa <i>WordSmith Tools</i>	75
Figura 16 - Extrator semiautomático de “xml”	76
Figura 17 – Arquivos referentes ao Anotador semiautomático de “xml”	77
Figura 18 - Sistematização proposta para o Editor de Definição do e-Termos.....	104
Figura 19 - Como ler os dados?.....	120

Lista de Quadros

Quadro 1 - Tipologia das relações propostas por Sager (1993)	30
Quadro 2 - Tipologia das relações conceituais utilizadas em Seppälä (2004)	34
Quadro 3 - Relações conceituais propostas por Feliu (2004).....	35
Quadro 4 - Protocolo da relação “Causalidade”	36
Quadro 5 - Papéis <i>Qualia</i> de Pustejovsky (1995).....	37
Quadro 6 – Tipologia das relações propostas por Jordan (1992)	40
Quadro 7- Adequabilidade do tipo de ferramenta, quanto ao conhecimento do usuário	54
Quadro 8 - Tipologia proposta (1).....	61
Quadro 9 - Tipologia proposta (2).....	61
Quadro 10 - Composição do <i>corpus</i>	63
Quadro 11 - Etiquetas utilizadas na anotação linguística	68
Quadro 12 - Expressões linguísticas: AGENTIVO	90
Quadro 13 - Expressões linguísticas: ASSOCIAÇÃO	91
Quadro 14 - Expressões linguísticas: CIRCUNSTÂNCIA	91
Quadro 15 - Expressões linguísticas: CONSEQUÊNCIA	92
Quadro 16 - Expressões linguísticas: CONSTITUTIVO	92
Quadro 17 - Expressões linguísticas: CONSTITUTIVO (continuação)	93
Quadro 18 - Expressões linguísticas: EXEMPLO.....	94
Quadro 19 - Expressões linguísticas: MEIO	94
Quadro 20 - Expressões linguísticas: SEMELHANÇA	95
Quadro 21 - Expressões linguísticas: TÉLICO	95
Quadro 22 - Expressões linguísticas: TÉLICO (continuação)	96
Quadro 23 – Termos com sufixo -OSE	98
Quadro 24 - Termos com sufixo -ITE	98
Quadro 25 - Termos com sufixo -METRO	99
Quadro 26 - Termos com sufixo -GRAFIA.....	99
Quadro 27 - Termos com sufixo -AGEM.....	100
Quadro 28 - Termos com sufixo -MENTO	100
Quadro 29 - Termos com sufixo -ÇÃO	100
Quadro 30 - Termos com sufixo -IDADE	101

Lista de Tabelas

Tabela 1 - Quantidade de DE _n s analisadas.....	79
Tabela 2 - Quantidade de DE _n s por DT.....	80
Tabela 3 - Relações semânticas no <i>corpus</i> de acordo com a frequência.....	80
Tabela 4 - Sequências produtivas das relações semânticas.....	81
Tabela 5 - Relações semânticas na posição DE1.....	82
Tabela 6 - Relações semânticas na posição DE2.....	82
Tabela 7 - Dados sobre as relações ASSOCIAÇÃO e SEMELHANÇA.....	83
Tabela 8 - Relações semânticas na posição DE3.....	83
Tabela 9 - Relações semânticas na posição DE4.....	84
Tabela 10 - Relação x + SEMELHANÇA.....	84
Tabela 11 - Relação x + ASSOCIAÇÃO.....	84
Tabela 12 - Relações semânticas na posição DE5.....	85
Tabela 13 - Relações semânticas do campo DEFEITO.....	85
Tabela 14 - Relações semânticas do campo DISFUNÇÃO.....	86
Tabela 15 - Relações semânticas do campo ENSAIO.....	86
Tabela 16 - Relações semânticas do campo TESTE E MEDIDA.....	86
Tabela 17 - Relações semânticas do campo EXAME COMPLEMENTAR.....	87
Tabela 18 - Relações semânticas do campo INSTRUMENTO OU EQUIPAMENTO.....	87
Tabela 19 - Relações semânticas do campo INSTRUMENTO DE MEDIDA.....	87
Tabela 20 - Dados quanto à quantidade de itens léxicos por relação semântica.....	88
Tabela 21 - Dados quanto à quantidade de itens léxicos por DE _n	89

Sumário

1	INTRODUÇÃO	13
2	FUNDAMENTOS TEÓRICOS.....	16
2.1	ORIGENS E DESENVOLVIMENTO DA TERMINOLOGIA	16
2.2	A DEFINIÇÃO.....	19
2.3	A DEFINIÇÃO TERMINOLÓGICA (DT)	21
2.4	RELAÇÕES SEMÂNTICAS	28
2.5	ALGUNS MODELOS E APLICAÇÕES SEMÂNTICAS.....	31
2.5.1	<i>Selja Seppälä (2004)</i>	32
2.5.2	<i>Judit Felíu (2004)</i>	34
2.5.3	<i>James Pustejovsky (1995)</i>	36
2.5.4	<i>Michael P. Jordan (1992)</i>	38
3	FUNDAMENTOS PRÁTICOS	42
3.1	MÉTODO DE ELABORAÇÃO DA DT NO ÂMBITO DO GETERM.....	42
3.1.1	<i>Ponto de partida: ontologia</i>	42
3.1.2	<i>Base definicional</i>	44
3.1.3	<i>Observação e seleção dos traços semânticos</i>	45
3.1.4	<i>Redação da DT</i>	47
3.1.5	<i>Validação da DT</i>	48
3.2	INFORMÁTICA E ESCRITA	49
3.2.1	<i>AMADEUS (Amiable Article Development for User Support)</i>	51
3.2.2	<i>SciPo (Scientific Portuguese)</i>	54
3.3	E-TERMOS (TERMOS ELETRÔNICOS).....	57
4	METODOLOGIA.....	61
4.1	PROPOSTA DE TIPOLOGIA DAS RELAÇÕES SEMÂNTICAS	61
4.2	COMPOSIÇÃO DO CORPUS	62
4.3	ARMAZENAMENTO DO CORPUS	64
4.4	NOMEAÇÃO DOS ARQUIVOS	65
4.5	MANIPULAÇÃO DO CORPUS (EDITOR DE CABEÇALHO)	65
4.6	PROTOCOLO DE ANOTAÇÃO DO CORPUS.....	69

4.6.1	<i>Algumas regras quanto à segmentação das DE_ns</i>	71
4.7	PROGRAMAS UTILIZADOS NA ANÁLISE DO CORPUS	73
4.7.1	<i>Wordsmith Tools</i>	73
4.7.2	<i>Extrator semiautomático de etiquetas “xml”</i>	75
4.7.3	<i>Anotador semiautomático de etiqueta “xml”</i>	76
5	ANÁLISE DO CORPUS	79
6	SISTEMATIZAÇÃO DO CONHECIMENTO OBTIDO NA PESQUISA	102
7	CONSIDERAÇÕES FINAIS	106
	REFERÊNCIAS	109
	ANEXO I	115
	ANEXO II	116
	ANEXO III	117
	APÊNDICE A – RELAÇÕES SEMÂNTICAS I E II	118
	APÊNDICE B – ANÁLISE DAS DTS POR CAMPO SEMÂNTICO	120

1 Introdução

A Terminologia tem como foco privilegiado de investigação o termo, porém, o desenvolvimento teórico e aplicado levou os estudos terminológicos a tratarem também, por exemplo, das unidades fraseológicas, textos de especialidades e da própria definição, até mesmo porque cada um desses objetos, distintamente, projeta os fundamentos do conhecimento especializado.

Nos últimos 40 anos, vários estudiosos, linguistas e não linguistas, vêm se dedicando a esses objetos inerentes às línguas de especialidade. Consequentemente, com o avanço das áreas de PLN (Processamento Automático de Língua Natural) e da Linguística de *Corpus*, observações de um grande volume de dados estão sendo realizadas, o que torna as análises mais ricas e descritivas.

Ao mesmo tempo, o uso de terminologias sistematizadas contribui para tornar a transferência de conhecimento e a comunicação entre seus usuários mais eficazes. Com a associação entre Terminologia e a Informática, é possível gerar com mais rapidez produtos terminológicos mais fiáveis, como vocabulários, dicionários, etc. No Brasil, a aproximação entre essas duas áreas ainda é incipiente, e os poucos terminólogos que desenvolvem esses produtos utilizam ferramentas computacionais não elaboradas especificamente para este fim.

Nesse cenário, caminha para a fase de finalização, um ambiente computacional colaborativo baseado na *web*, denominado e-Termos, que prevê seis etapas (desde compilação do *corpus* até edição do verbete), cujo propósito é semiautomatizar as tarefas de criação e gerenciamento do trabalho terminológico. Uma das etapas que será acoplada ao *e-Termos* constitui um “Editor de Definição”, que tem como objetivo auxiliar o terminólogo a redigir a definição terminológica (DT) de forma sistemática.

A partir da hipótese de que as DTs seguem determinados padrões semânticos, nossa proposta é observar e descrever esses padrões e, conseqüentemente, sistematizá-los a fim de servirem como conhecimento lingüístico a ser implementado no “Editor de Definição” do e-Termos.

A fim de que a etapa dedicada à DT, de fato, facilite a tarefa da redação da definição de modo eficiente, é necessário o cumprimento de alguns objetivos específicos, os quais são propostos para essa pesquisa:

- I. identificação das relações semânticas constitutivas das DTs do tipo gênero próximo e diferença específica (GPDE), cujos termos estão organizados em campos semânticos pertencentes a áreas de conhecimento distintas: Revestimento Cerâmico (domínio das Engenharias) e Fisioterapia (domínio das Ciências da Saúde)¹;
- II. observação e seleção das expressões linguísticas empregadas para expressar as relações semânticas identificadas;
- III. observação e seleção de determinados morfemas que constituem os termos-entradas para a identificação e recuperação do seu campo semântico e, por conseguinte, do seu hiperônimo;
- IV. sistematização dos aspectos linguísticos envolvidos na elaboração da DT.

Para a realização desse estudo descritivo das DTs, foi necessário o cumprimento de algumas etapas metodológicas, quais sejam: i) composição do *corpus*, o qual foi formado por 500 DTs e seus respectivos termos, seguindo determinados critérios de escolha, referentes a sete campos semânticos da área de Revestimento Cerâmico (tais como “defeito”, “propriedade”, “equipamento”, etc.), e a oito campos da área da Fisioterapia (“articulação”, “disfunção”, “teste e medida”, etc.); ii) tratamento do *corpus* para que este pudesse ser manipulado computacionalmente. Esta etapa compreendeu o armazenamento e a limpeza do *corpus*, a nomeação padronizada de todos os arquivos que constituem o *corpus* e a anotação estrutural e linguística do *corpus*; iii) observação e descrição dos aspectos semânticos das DTs, segundo uma tipologia de relações semânticas elaborada nesta pesquisa; iv) descrição de determinados aspectos morfológicos pertinentes para a identificação a qual campo semântico o termo pertence. Como resultado, foi realizada uma proposta geral de sistematização quanto à redação da DT de forma a integrar o “Editor de Definição”.

O trabalho está organizado na forma descrita a seguir.

No Capítulo 2, são apresentadas a área dos estudos terminológicos e a teoria a qual nos filiamos – a Teoria Comunicativa da Terminologia (TCT). Os tópicos seguintes são sobre a definição e a definição terminológica. Em seguida, são apresentadas as relações semânticas presentes na Terminologia e, ao final do capítulo, as aplicações semânticas na Terminologia:

¹ O *corpus* da pesquisa será detalhado no item 4.1.

Feliu (2004) e Seppälä (2004); as relações semânticas no nível do léxico: Pustejovsky (2005), e no nível do texto: Jordan (1992) são abordadas.

No Capítulo 3, são expostos os aspectos práticos que integram essa pesquisa. A saber: o método de trabalho na redação da DT no âmbito do GETerm; a relação informática e escrita, por meio de dois ambientes computacionais de auxílio à escrita acadêmica (AMADEUS e SciPo) e o e-Termos.

O Capítulo 4 faz referência à metodologia empregada na pesquisa que teve por fim a observação e a descrição das DTs: proposta de tipologia das relações semânticas empregada na anotação do *corpus*, composição e armazenamento do *corpus*, nomeação dos arquivos, manipulação do *corpus*, protocolo de anotação do *corpus*, além da descrição dos programas computacionais utilizados: Editor de Cabeçalho, *Wordsmith Tools* e programas criados especialmente para essa pesquisa (Extrator semiautomático de etiquetas “xml” e Anotador semiautomático de etiquetas “xml”).

O Capítulo 5 destina-se à análise do *corpus*. No Capítulo 6, é exposta a sistematização do conhecimento obtido na pesquisa. No Capítulo 7, são apresentadas as considerações finais.

O Apêndice A apresenta as relações semânticas utilizadas como etiqueta no *corpus* e por fim, o Apêndice B apresenta informações das DTs de acordo com seu campo semântico.

2 Fundamentos teóricos

2.1 Origens e desenvolvimento da Terminologia

A Terminologia, como prática, data do século XVIII, com os trabalhos de Lavoisier e Berthold no domínio da Química e Linné nos domínios da Botânica e Zoologia. O objetivo desses especialistas concentrava-se, sobretudo, na fixação das denominações para os conceitos científicos. O interesse dos especialistas pelos conceitos e denominações continuou no século seguinte. Devido ao desenvolvimento progressivo das ciências, os próprios especialistas manifestaram particular interesse em sistematizar regras de formação de termos para cada domínio de especialidade (CABRÉ, 1993, p. 21).

No século XX, período em que o progresso das ciências e das técnicas se intensificou, assim como o desenvolvimento crescente da tecnologia, os cientistas passam a se preocupar com questões inerentes às linguagens de especialidade: denominação e harmonização de conceitos novos (CABRÉ, 1993, p. 22).

Nesse cenário, surgem, na Europa, os primeiros trabalhos que se propuseram a estabelecer a Terminologia como uma disciplina: D.S. Lotte (1889-1950), pai da Escola Soviética de Terminologia, e o engenheiro austríaco E. Wüster (1898-1977), fundador da Escola de Viena. Esses autores constituem os precursores da disciplina (CABRÉ, 1993, p.22). Embora seja discutido qual dos dois inaugurou os estudos terminológicos, de fato, a tese de doutoramento de Wüster foi a primeira exposição sistematizada de uma teoria da Terminologia, intitulada de *A normalização internacional da terminologia*, em 1931.

Essa tese é a primeira obra de Wüster que apresenta sinais da elaboração da chamada Teoria Geral de Terminologia (TGT). Mais tarde, na década de 70, quando ele lecionou na Universidade de Viena, no curso de “Introdução à Teoria Geral de Terminologia e à Lexicografia Terminológica”, Wüster apresentou os pressupostos da TGT, os quais foram publicados, após seu falecimento, no livro *Introdução à Teoria Geral de Terminologia e à Lexicografia Terminológica*.

A TGT tem como primazia o **conceito** e apresenta como proposta a compilação de conceitos e termos para a normalização, com o objetivo de assegurar a univocidade da comunicação profissional. Para garantir esse método prescritivo de tratar as terminologias, questões de sinonímia, homonímia e polissemia não são consideradas. Isso se relaciona com

o fato de que Wüster postula que as linguagens de especialidade são dissociadas das línguas naturais. Desta forma, conforme Cabré (in Prefácio da obra de Wüster, 1998, p. 11), “a TGT é internamente coerente, pois se baseia no logicismo, na busca da língua universal e na uniformidade da comunicação” (tradução nossa)². Contudo, a TGT é válida para resolver somente um tipo de comunicação: a **estandardizada**, ou seja, a comunicação padronizada, a qual é apenas uma das possibilidades de comunicação real (CABRÉ, 1999, p.129).

As últimas décadas foram determinantes para que ocorresse uma reflexão acerca das insuficiências da TGT, já que houve um aumento significativo nas descrições das terminologias, elaboração de obras terminográficas, desenvolvimento de tecnologia adequada ao tratamento computacional de terminologias e ao planejamento linguístico.

A partir dos anos 80, junto às críticas à TGT, começam a surgir perspectivas teóricas novas, as quais apresentam como intersecção o mesmo ponto de origem – os estudos da linguagem, tais como:

- I. **Socioterminologia** (GAMBIER, 1987; BOULANGER, 1991, 1995; GAUDIN 1993, 2003): tem como foco a descrição dos termos em seus mais diferentes contextos de usos. A variação terminológica é estudada em movimentos sincrônico e diacrônico;
- II. **Terminologia de Base Textual** (HOFFMANN, 1998; CIAPUSCIO, 2003): tem como objetivo a compreensão minuciosa dos textos especializados, tanto em nível macro como microestrutural;
- III. **Teoria Comunicativa da Terminologia (TCT)** (CABRÉ, 1999, 2003): tem como princípio a visão linguística sobre a linguagem especializada. Nessa abordagem, o termo integra um determinado âmbito específico, sem perder as características próprias de qualquer unidade pertencente ao sistema linguístico das línguas naturais. Desta forma, um termo é considerado um signo linguístico em funcionamento numa situação de comunicação especializada. A seguir, a TCT será mais detalhada, já que é a perspectiva adotada nesta investigação;

² *la TGT es internamente coherente, porque se basa en el logicismo, en la búsqueda de la lengua universal y en la uniformidad de la comunicación.*

IV. **Teoria Sociocognitiva da Terminologia** (TEMMERMANN, 2000): apresenta como destaque o papel dos modelos cognitivos, procurando mostrar as relações entre os processos de categorização e a linguagem. A abordagem é fundamentada na semântica cognitiva e questiona a centralidade da standardização em detrimento de uma descrição autêntica dos significados dos termos tal como aparecem nos textos.

As quatro perspectivas citadas se destacam por conceber as terminologias³ como campos de investigação pertinentes à linguagem, contudo, cada qual possui um foco específico, se servindo de um aparato (não somente linguístico) que forneça contribuições apuradas de acordo com o que se está pesquisando; o mesmo ocorre com várias correntes linguísticas (Análise do Discurso, Sociolinguística, Semântica da Enunciação) e com os distintos níveis de análise (morfológico, sintático, semântico lexical, textual).

Como mencionado, dentre esses paradigmas, optamos pela Teoria Comunicativa da Terminologia (TCT) apresentada por Cabré (1999, 2003) e sua equipe do Instituto Universitário de Linguística Aplicada (IULA) da Universidade Pompeu Fabra (UPF) em Barcelona/Espanha. Ao contrário da TGT, a TCT articula-se baseada na valorização dos aspectos comunicativos das linguagens especializadas em detrimento dos propósitos normalizadores.

De acordo com Almeida (2006, p. 86 e 87), um projeto terminológico vinculado à TCT deve refletir na sua prática os seguintes pressupostos gerais:

- I. o objeto central da Terminologia são as unidades terminológicas e não os conceitos. Eleger as unidades como objeto central significa reforçar uma perspectiva linguística e uma abordagem semiológica;
- II. não há uma diferença *a priori* entre termo e palavra, o que há são signos linguísticos que podem realizar-se no discurso como termo ou palavra dependendo da situação comunicativa;
- III. os níveis lexical, morfológico, sintático e textual podem veicular conhecimento especializado;
- IV. os termos devem ser observados no seu ambiente natural de ocorrência, ou seja, nos discursos especializados;
- V. a variação conceitual e denominativa deve ser considerada;
- VI. do ponto de vista cognitivo, as unidades terminológicas: i) estão subordinadas a um contexto temático; ii) ocupam um lugar preciso num mapa conceitual, iii) o

³ É consenso na área grafar *terminologia* com *t* minúsculo no sentido de “conjunto de termos”, e com *T* maiúsculo quando se refere ao “campo de estudos ou disciplina”.

seu significado específico é determinado pelo lugar que ocupam nesse mapa (CABRÉ, 2003, apud ALMEIDA, 2006, p. 86 e 87).

Desta forma, na realização desta pesquisa, fundamentamo-nos nos pressupostos teóricos e práticos da TCT, pois julgamos que essa é a abordagem mais apropriada na observação dos termos, suas definições e todas as suas implicações (variação denominativa, ontologia, percurso semasiológico ao onomasiológico, significado, etc.).

É importante assinalar que a Terminologia, como uma área da Linguística Aplicada, pode apoiar-se em teorias linguísticas diversas, desde que essas propiciem subsídios para a construção da sua própria teoria. Ainda sim, os estudos terminológicos alimentam-se de outras grandes áreas como Ciência da Informação, Computação, Tradução, Cognitivismo e, sobretudo, da própria Linguística.

2.2 A Definição

A definição está presente em diversos níveis do nosso cotidiano: em um nível menos formal, ao definirmos um dado objeto, sentimento ou ação para um simples interlocutor; como também é comum em materiais didáticos, manuais, textos de divulgação científica, etc.; até chegar ao seu alto nível de formalização, quando inserida em verbetes de dicionários, glossários, etc.

Uma definição “natural” e uma definição “dicionarística” diferenciam-se muito mais pela forma de apresentação do que pelo seu conteúdo. Enquanto na primeira é permitido o uso de uma metalinguagem como “é definido como”, “poder ser considerado”, “é designado como”, “significa”, ou seja, expressões que evidenciam que se trata de um excerto definitório; a definição “dicionarística” se caracteriza por determinadas convenções que regem sua estrutura.

No âmbito da Terminologia e da Terminologia Computacional (ou Terminótica), alguns pesquisadores analisaram de que forma a definição apresenta-se nos textos que constituem o *corpus* de pesquisa em um trabalho terminográfico. No geral, essas pesquisas têm como foco a extração semiautomática de contextos definitórios, os quais podem corroborar com as etapas da elaboração do mapa conceitual (ontologia) e da própria redação do texto definitório⁴.

⁴ São exemplos desse tipo de pesquisa, os trabalhos de Auger (1998) e Sierra & Alarcón (2003).

Quanto a definições “dicionarizadas”, essas podem ser classificadas em **definição lexicográfica, definição enciclopédica e definição terminológica**. Larivière (1996, p. 409) discrimina os três tipos:

- I. **definição lexicográfica (DL):** utilizada nos dicionários de língua e nos dicionários enciclopédicos, propõe-se explicitar os significados distinguindo os sentidos e os empregos dos signos (ou palavras) de uma língua;
- II. **definição enciclopédica (DE):** utilizada nas enciclopédias e nos dicionários enciclopédicos, propõe-se fornecer um conjunto de conhecimento sobre uma coisa;
- III. **definição terminológica (DT):** utilizada nos vocabulários especializados, propõe-se caracterizar (delimitar e distinguir de outras noções) as noções denominadas por um termo e que representam uma coisa no interior de um sistema organizado (tradução nossa)⁵.

Portanto, é de se esperar que essas diferenças influenciem na constituição do texto definitório. Pois, como afirma Barros (2004, p.159): “não existe uma definição válida para dois dicionários, uma vez que a cada tipo de obra correspondem algumas características específicas que determinam o conteúdo e a organização do enunciado definicional”.

Com essa afirmação, reitera-se a noção de que, assim como ocorre nos diversos gêneros textuais citados (materiais didáticos, manuais, textos de divulgação científica, etc.), a etapa que corresponde à definição exige um certo planejamento anterior à sua redação, principalmente quanto ao público-alvo e à finalidade da obra em questão. Ou seja, o autor de um manual ou livro didático tem como objetivo transmitir um dado assunto ou fazer com que o usuário execute uma dada tarefa, e para tal, o autor, necessariamente, apresenta uma linguagem que seja adequada a esse usuário, em termos estruturais e de conteúdo. O mesmo acontece na redação da DT, em que terminólogos preveem as necessidades dos consulentes em potencial e, a partir disso, adequam o texto definitório.

⁵ *la définition lexicographique (DL): utilisée dans les dictionnaires de langue et les dictionnaires encyclopédiques, qui se propose d'expliciter des signifiés en distinguant les sens et les emplois des signes (ou mots) d'une langue; la définition encyclopédique (DE): utilisée dans les encyclopédies et les dictionnaires encyclopédiques, qui se propose de fournir un ensemble de connaissances sur une chose; la définition terminologique (DT): utilisée dans les vocabulaires spécialisés, qui se propose de caractériser (i.e. de délimiter et de distinguer des autres notions) des notions dénomimées par un terme et représentant une chose à l'intérieur d'un système organisé.*

2.3 A Definição Terminológica (DT)

Mas afinal, como a DT pode ser definida? Alguns pontos de vista sobre a DT, os quais serão descritos a seguir, devem ser considerados como produtos da movimentação da Terminologia de se firmar cada vez mais como disciplina linguística.

Entre as visões menos linguísticas sobre a DT, podemos destacar Wüster, fundador dos estudos terminológicos, que afirma: “em seu sentido amplo, uma definição é a descrição de um conceito mediante conceitos conhecidos, expressos geralmente por meio de palavras.”⁶ (WÜSTER⁷, 1998, p. 65, tradução nossa). A TGT **tem como foco principal o componente conceitual, em detrimento do significado**. Se os conceitos, nesta perspectiva, são estáveis, paradigmáticos e universais, como idealiza o lógico-positivismo, a DT não é realizada como um texto, e sim como um espelho do *definiendum*.

As insuficiências e inviabilidades da TGT em situações reais de comunicação especializada despertaram o interesse de grupos de pesquisadores, sobretudo do Canadá, da Espanha e da França. Robert Dubuc é um dos representantes mais renomados da famosa Escola Canadense de Terminologia e foi professor de Terminologia na Universidade de Montreal durante quase 30 anos. O seu livro intitulado *Manuel pratique de terminologie* (1992), que em 1999 foi traduzido para a língua espanhola, apresenta, em 16 capítulos, temas relevantes como: “O que é terminologia? Termos e noções”, “A terminologia bilíngue e plurilíngue”, “A ficha terminológica”, “Definição”, entre outros.

No capítulo referente à Definição, o autor afirma que o objetivo da definição é “dar uma imagem mental exata de um conceito (...) que permita sua identificação, baseando-se em seus traços essenciais” (DUBUC, 1999, p. 119, tradução nossa)⁸. Com relação às características da DT, o autor afirma que um texto definatório satisfatório deve ter clareza, adequação e concisão. Dubuc (1999) também aponta o que deve ser evitado numa DT: o texto não deve ser tautológico, nem circular, e preferencialmente não deve ser escrito na forma negativa. Mais adiante, Dubuc (1999) dedica um subitem a *fórmulas de início* que servem para começar a definição de determinadas categorias morfosintáticas. Um exemplo

⁶ *en su sentido amplio, una definición es la descripción de un concepto mediante conceptos conocidos, expresados generalmente por medio de palabras.*

⁷ A primeira edição dessa obra foi escrita em alemão e publicada em 1979.

⁸ *dar una imagen mental exacta de un concepto (...) que permita su identificación, basándose en sus rasgos esenciales.*

que consta do capítulo (DUBUC, 1999, p.125) refere-se à definição do termo “destilação”:
“Ação e efeito de separar por meio do calor (...)” (tradução nossa)⁹.

Outro autor que merece destaque é Juan Carlos Sager, estudioso argentino radicado na Inglaterra. Vale a pena mencionar que inicialmente ele se esforçou no ideário da Escola de Viena para só depois chegar a contestá-la. Em seu livro *Curso práctico sobre el procesamiento de la terminología* (1993), tradução espanhola do inglês, obra originalmente publicada em 1990, no capítulo “La Dimensión Cognoscitiva”, o autor aborda sequencialmente a “teoria da referência”, a “teoria dos conceitos” e, em seguida, “definições e alternativas” e “definições terminológicas e relações”.

Na obra, o autor afirma que a DT, enquanto produto, “é uma descrição linguística de um conceito, constituída pela enumeração de um conjunto de características que dão conta do significado do conceito” (Sager, 1993, p. 68) (tradução nossa)¹⁰. O autor faz referência ao tipo clássico de definição composto pelo gênero próximo e diferença específica (GPDE), que se constituiu como referência para a elaboração de definição desde os estudos filosóficos, passando pela Terminologia Clássica, até alcançar lugar de destaque nos estudos terminológicos nos dias de hoje.

A definição que segue esse modelo GPDE, apresenta-se da seguinte forma: a DT é encabeçada pelo gênero próximo (GP), ou seja, um hiperônimo (ou termo-pai), o qual geralmente é recuperado observando-se o termo superordenado na ontologia. Imediatamente, segue a diferença específica (DE), que estabelece a diferença entre os termos-entradas que possuem o mesmo GP. Segue, na Figura 1, exemplo contendo 2 DTs que apresentam o mesmo GP¹¹.

⁹ *Acción y efecto de separar por medio del calor (...)*

¹⁰ *es una descripción lingüística de un concepto. Basada en el listado de un número de características que transmiten el significado del concepto.*

¹¹ DTs extraídas do DiRC (Dicionário de Revestimento Cerâmico), em fase de finalização no âmbito do GETerm.

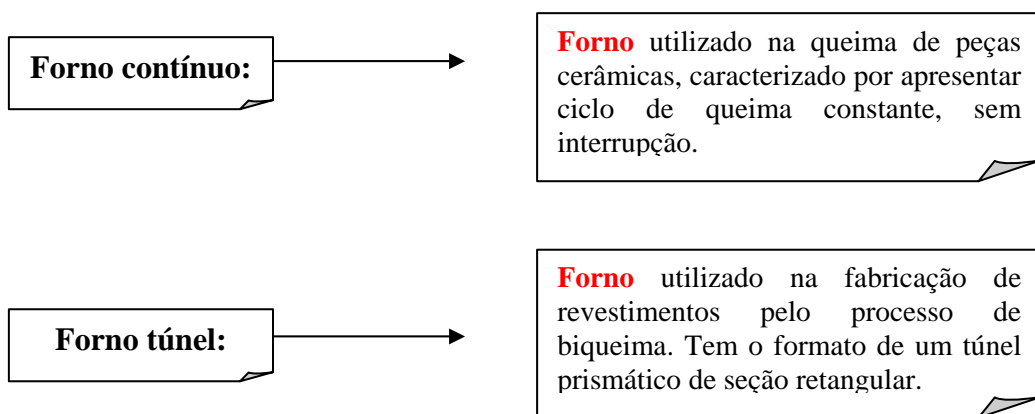


Figura 1 – Exemplo de DT no modelo GPDE

Porém, como Sager acredita que esse modelo não é adequado para definir todos os tipos de termos, propõe a seguinte tipologia¹²:

- I. **definição mediante análise.** Exemplo: estomatite = inflamação da boca;
- II. **definição mediante sinônimos.** Exemplo: margarida do campo = margarida-anual (*Novo Dicionário Eletrônico Aurélio versão 5.0*);
- III. **definição mediante paráfrase.** Exemplo: brancura = estado de branco;
- IV. **definição mediante síntese** (mediante relação identificativa, mediante descrição). Exemplo: margarida₂ = em algumas máquinas de escrever e impressoras, disco com palhetas dispostas radialmente, na extremidade das quais há tipos em relevo para a impressão de caracteres (*Novo Dicionário Eletrônico Aurélio versão 5.0*);
- V. **definição mediante implicação** (mediante o uso da palavra em um contexto explicativo). Exemplo: diagnóstico = se faz um diagnóstico quando se identificam certos sintomas como característicos de uma condição especial;

¹² Alguns exemplos foram modificados, outros foram retirados do dicionário eletrônico *Novo Dicionário Eletrônico Aurélio versão 5.0* (2004).

- VI. **definição mediante denotação** (mediante a enumeração de exemplos, mediante extensão). Exemplo: oceano = os oceanos são o Atlântico, o Pacífico e o Índico;
- VII. **definição mediante demonstração** (definição ostensiva). Exemplo: desenhos, fotografias, referência situacional.

Algumas definições podem ser mistas, como, por exemplo: mediante **análise e descrição**; mediante **sinônimo e descrição**; mediante **sinônimo e análise**.

Para Sager (1993), as definições terminológicas cumprem três funções:

- I. situam corretamente o termo em seu lugar dentro da estrutura de conhecimento. Isso pressupõe certo conhecimento da área, que se obteve a partir da leitura das definições já existentes, dos contextos, das consultas a especialistas;
- II. fixam o significado especializado do termo. Deve ser flexível e menos rígida em certas áreas do conhecimento como as recentes tecnologias; em oposição às taxonomias da Biologia, Química, Medicina;
- III. oferecem aos usuários não-especializados a compreensão do termo.

Segundo Cabré (1993), as definições devem ser elaboradas de acordo com uma série de princípios e convenções, que são divididos em três grupos: I. Adequações gerais; II. Adequações específicas e III. Expressão. Observem-se a seguir:

I. Adequações gerais

As definições devem:

- a) ser verdadeiras;
- b) permitir a distinção dos outros conceitos;
- c) reconhecer as dimensões pertinentes de cada campo de especialidade;
- d) situar-se na perspectiva do campo nocional a que pertence o conceito;
- e) adequar-se a finalidades do trabalho de que faz parte.

II. Adequações específicas

As definições devem:

- a) acoplar-se ao sistema de definições de um campo concreto, partindo da estrutura prévia deste campo;
- b) explicitar todas as características (traços) essenciais de cada conceito, de acordo com a estrutura nocional estabelecida;
- c) refletir as relações sistemáticas que cada conceito mantém com os demais conceitos do mesmo campo;
- d) incluir todas as características que, ainda que não sejam essenciais, são importantes para uma descrição completa do conceito;
- e) apresentar os traços dos conceitos de forma sistemática, especialmente quando se trata de conceitos que pertencem a um mesmo campo do mapa conceitual.

III. Expressão

As definições devem:

- a) ser expressas corretamente;
- b) ser formalmente adequadas, de acordo com as normas de construção de definições;
- c) utilizar a expressão adequada aos destinatários do trabalho;
- d) constar de uma só oração, evitando pontos internos¹³;
- e) respeitar os princípios da lexicografia no que diz respeito a sua apresentação formal, que se resumem aos seguintes pontos:
 - o descritor inicial deve ser da mesma categoria gramatical do termo descrito;
 - devem-se utilizar palavras conhecidas pelos usuários em geral e, ao utilizar vocábulos mais específicos, eles também têm de ser entradas;
 - a definição não deve ser circular;
 - evitar a negação, a utilização de paráfrases desnecessárias e de fórmulas metalinguísticas.

¹³ Embora nem sempre seja possível dar a exata “fotografia” do conceito com uma única oração.

A tese de doutoramento de Maria José Bocorny Finatto (Universidade Federal do Rio Grande do Sul - UFRGS), denominada *Definição terminológica: fundamentos teórico-metodológicos para sua descrição e explicação* (2001), se propõe observar as condições de subjetividade reveladas do texto definitório, utilizando, para isso, os conceitos basilares da semântica enunciativa cunhados por Émile Benveniste. A pesquisa demonstra “a viabilidade e produtividade do reconhecimento de relações mantidas quer entre o enunciado-texto da definição e seu ambiente de significação, quer entre o enunciado definitório e seus sujeitos enunciativos” (FINATTO, 2001, p. 14).

Conforme Almeida et al. (2007) “Ao trazer à tona reflexões acerca da Terminologia e principalmente da DT, Finatto (2001) oferece um renovado paradigma teórico, abrindo novos olhares sobre a DT como um objeto linguístico multifacetado”:

é preciso voltar a atenção para o texto definitório em si mesmo, sem tomar um modelo prévio de observação, de modo a inscrevê-lo relacionamente em algo mais amplo, o que poderia ser algo como um *frame de significação*, simultaneamente construído pelo indivíduo-autor e pela coletividade que ele representa, recuperando-se, enfim, sua autoria, ambiência e, fundamentalmente, sua natureza linguística (FINATTO, 2001, p.139).

Para além das características desejáveis que deve ter a DT, as quais já foram apontadas neste capítulo, como afirma Almeida et al. (2007) “parece-nos que cada vez mais é necessário admitir a presença de um sujeito enunciativo no texto definitório”. Embora o redator da definição tenha uma certa liberdade quanto à redação da DT (o que faz, aliás, com que dois redatores redijam a mesma DT de formas distintas), ele está circunscrito num dado contexto de trabalho (público-alvo, objetivo da obra terminográfica e etc.), e isso faz com que a tarefa da redação da DT possa ser otimizada.

Como o objetivo final dessa pesquisa é fornecer conhecimento linguístico acerca da DT para que se componha o Editor de Definição do ambiente e-Termos, os principais princípios e convenções da redação da definição terminológica serão estruturados a fim de servirem como guia aos usuários do Editor. A seguir, serão apresentadas algumas noções que estão implicadas na DT e que são imprescindíveis para a sua compreensão.

Os conceitos/termos são descritos e se distinguem por meio da presença ou ausência de traços ou características¹⁴. Ordená-los em campos (como numa ontologia) é um início para

¹⁴ Traço e característica, neste trabalho, são utilizados como termos sinônimos. Conforme assinala Barros (2004, p. 107): “a identificação, a distinção e a descrição dos diferentes conceitos são feitos por meio de traços de sentido, chamados em Terminologia de *característica*...”.

que se compreenda o objeto, porém, não é suficiente para que um termo seja discriminado dos demais.

Em Cabré (1993, p. 198), as **características**, segundo a relevância que manifestam na conformação de um conceito, podem ser:

- I. **essenciais:** em que é descrita a essência, a natureza do termo;
- II. **não-essenciais ou complementares:** são as que acrescentam elementos não relevantes à descrição. Porém, a autora lembra que dependendo do campo de especialidade e dos objetivos do trabalho terminológico, uma característica não-essencial pode converter-se em essencial.

Cabré (1993, p. 199) ainda assinala que as **características** podem ser intrínsecas e extrínsecas¹⁵, segundo a relação que mantém com o objeto que representam:

- I. **intrínsecas:** são inerentes ao objeto descrito. Exemplos: forma, cor, medida, material, etc.;
- II. **extrínsecas:** são externas ao objeto descrito. Exemplos: função, origem, destino, localização, inventor, etc.

Em muitos casos (especialmente quando se trata de substantivos concretos), podemos afirmar que uma característica é extrínseca ou intrínseca. Por exemplo, é uma característica intrínseca de um CD – *compact disc* – ser uma “circunferência com um furo no meio e metalizado” (forma e material), ou seja, características constituintes da própria natureza do termo. Ao mesmo tempo, possui característica extrínseca de “tocar música e ser produzido na China” (função e origem).

Contudo, muitas vezes as características intrínsecas não são suficientes para a descrição e distinção de um termo, por isso faz-se necessário que também se agreguem à definição as características extrínsecas. Logo, se é essencial ou não-essencial, vai depender se o termo exige (essencial) ou não (não-essencial) as características para ser definido.

¹⁵ Em Terminologia, a noção de características intrínseca e extrínseca é iniciada por Wüster (1998, p. 46), e que está presente na ISO 740/1987.

Assim, no âmbito do GETerm, no que concerne à prática terminográfica, consideramos que:

(...) não seja possível, nem produtivo, insistir na cisão entre traços [características] intrínsecos e traços extrínsecos, propondo que os primeiros figurem na definição e os últimos na informação enciclopédica (IE¹⁶). Ora, nem sempre se pode classificar com segurança o que são traços distintivos intrínsecos e extrínsecos do conceito cujo termo está sendo definido. Além do mais, é possível que aquilo que julgamos ser um traço extrínseco é, muitas vezes, a informação que faltava para o consulente compreender o termo-entrada. (Almeida et al., 2007)

Além de não ser produtivo considerar essa cisão entre características intrínsecas e extrínsecas, é relevante, num projeto terminográfico, que a equipe estabeleça um modelo de definição do termo, seguindo determinados critérios, e que esse modelo seja eficiente aos propósitos da pesquisa. Quanto a essa questão, Barros (2004, p. 109) destaca que “como alternativa, a TCT defende a livre escolha do modelo linguístico para a análise dos termos em todos os níveis, inclusive o semântico”.

2.4 Relações semânticas

Conforme Bodson (2004), os termos “relação semântica” e “relação conceitual” são duplamente empregados nos estudos da área. A escolha de um termo em detrimento do outro reflete, entre outras coisas, o ponto de vista adotado. O autor afirma que “na verdade, a relação semântica explicita uma ligação entre dois significados, enquanto que a relação conceitual põe em relevo a estrutura de conhecimento [ontologia do domínio]”¹⁷ (Bodson, 2004, p. 36, tradução nossa).

Nesse sentido, assumimos, juntamente com Jackendoff (1983, p. 95), que, apesar de haver os dois níveis de representação mental (conceitual e semântico), essa distinção não é tão produtiva, pois acaba denotando o mesmo fenômeno. Em razão disso, o autor vai preferir o termo “estrutura semântica”. Adotaremos, então, em nosso trabalho, o termo “semântico” para compor as lexias “traço semântico”, “relação semântica” e “campo semântico”.

¹⁶ A informação enciclopédica é um campo da ficha terminológica, utilizada pelos terminólogos para inserir informações que auxiliam na compreensão do termo, mas que não são consideradas imprescindíveis.

¹⁷ *En effet, la relation sémantique exprime un lien entre deux sens, alors que la relation conceptuelle met plutôt l'accent sur la structure de la connaissance.*

As relações semânticas estão presentes nos estudos terminológicos desde a tipologia de relações hierárquicas inaugurada por Wüster (1998), a qual até hoje é considerada como referência na área. Essas relações, sintetizadas por Sager (1993, p. 56 e 59), podem ser de dois tipos:

I. **relação genérica (espécie-gênero):** que pode ser descrita como “tipo de”, mediante fórmulas:

X é um tipo de *A*

X, Y, Z são tipos de *A*

A contém *X, Y, Z*

A contém o subtipo *X*

II. **relação partitiva (parte-todo):** que indica ligação entre os conceitos, os quais consistem em mais de uma parte e suas partes constitutivas. Ocorre mediante fórmulas:

X é um componente de *A*

X, Y, Z são componentes de *A*

A consiste em *X*

A consiste em *X, Y, Z*

Essas relações, embora sejam relevantes para a Terminologia, ainda assim são insuficientes, fato esse que levou autores como Arnzt e Pitch (1989), e posteriormente Sager (1993), a considerarem também as relações denominadas não-hierárquicas ou complexas, pois frequentemente os conceitos se inter-relacionam por suas proximidades situacionais, o que exige uma estruturação diferente das relações genéricas e partitivas. As relações possíveis propostas por Sager (1993, p. 62 e 63) podem ser:

Relações		
1	Causa	Efeito
2	Material	Produto
3	Material	Propriedade
4	Material	Estado
5	Processo	Produto
6	Processo	Instrumento
7	Processo	Método
8	Processo	Paciente
9	Fenômeno	Medida
10	Objeto	Contra-agente
11	Objeto	Recipiente
12	Objeto	Material
13	Objeto	Qualidade
14	Objeto	Operação
15	Objeto	Característica
16	Objeto	Forma
17	Atividade	Lugar

Quadro 1 - Tipologia das relações propostas por Sager (1993)

As relações complexas entre os conceitos não se limitam a essa lista. Outras criações são possíveis e vão depender da área que está sendo descrita. Os tipos de relações podem ser evidenciados nos textos definitórios, por meio de expressões linguísticas que propiciem o estabelecimento das relações propostas acima. Exemplo: *causada por* (1), *propriedade de* (3), *instrumento para* (6), *lugar para* (17), etc.

No caso da Terminótica, pesquisas têm sido realizadas a partir da eleição de determinadas relações semânticas na extração de informação semiautomatizada em *corpus*, tanto para a geração de ontologia como para a extração de enunciados definitórios. Essas investigações revelam que existem quatro categorias de relações muito produtivas, a saber: **relação hiperonímica, relação meronímica, relação causal e relação de finalidade** (Davidson et al. (1998), Marshman (2002), Marshman et al. (2002) e Condamines (2003)).

Como tais relações já foram objetos de estudo de várias pesquisas da área e, portanto, foram amplamente definidas, pretendemos não nos aprofundarmos nesse tópico. Contudo, alguns esclarecimentos acerca de cada relação são importantes para que, posteriormente, seja possível eleger nossa tipologia de relações a ser aplicada no *corpus*.

- I. **relação hiperonímica/hiponímica**¹⁸: relação que vincula um item específico ao seu correspondente genérico ou um genérico ao seu específico;
- II. **relação meronímica**¹⁹: relação que expressa vínculo entre um todo e suas partes;
- III. **relação causal**: relação que vincula uma causa a seu efeito. As seguintes fórmulas (genéricas) podem ser empregadas para validar essa relação: *X causa Y*, em que X é o agente causal e Y é o efeito; ou *X foi causado por Y*, em que X é o efeito e Y é o agente causal;
- IV. **relação de finalidade**: relação que expressa a utilidade de uma entidade. A fórmula (genérica) *X serve para Y*, em que X é a entidade e Y é a função.²⁰

O estudo acerca das relações semânticas revela-se útil na descrição, geração e extração de informação, sendo que para cada propósito de pesquisa relacionada ao tema é empregado um modelo ou um rol de relações semânticas que permita atingir determinados objetivos. Desse modo, no subcapítulo 2.5, serão apresentados e discutidos sobre alguns modelos e aplicações semânticas no âmbito da Linguística e da própria Terminologia que contribuíram para esta pesquisa.

2.5 Alguns modelos e aplicações semânticas

Ao longo do primeiro ano de pesquisa, em apresentações realizadas do projeto de mestrado, alguns questionamentos foram feitos sobre a validade das relações semânticas elencadas na nossa descrição dos termos referentes às duas áreas do conhecimento utilizadas no estudo, quais sejam: Fisioterapia e Revestimento Cerâmico. Dito de outro modo: as relações elencadas estão de fato presentes em textos definitórios de outros domínios do conhecimento, de forma que os resultados obtidos nesta pesquisa possam ser replicados em outros domínios?

¹⁸ O mesmo que a relação genérica proposta por Wüster.

¹⁹ O mesmo que a relação partitiva proposta por Wüster.

²⁰ Ressalte-se que cada uma dessas relações, especialmente as três últimas, podem ainda apresentar alguns desdobramentos.

Percebemos que relações como: **causa, finalidade, consequência** são mais usuais²¹, enquanto que traços como **recomendação, diagnóstico, indicação** (DicFisio²²) são mais específicos para determinados campos semânticos de algumas áreas de especialidade. Ao mesmo tempo, graças às disciplinas cursadas e às leituras realizadas, pudemos entrar em contato com alguns autores que tratam de relações semânticas, ou ainda, que aplicaram tais relações em suas pesquisas, sob diferentes perspectivas.

Deste modo, este capítulo destina-se a discutir as relações semânticas e/ou conceituais²³ presentes em obras de determinados autores da Terminologia e da Linguística Geral, e de que forma essas pesquisas podem contribuir com o estudo acerca da descrição da DT.

2.5.1 Selja Seppälä (2004)

Entre os autores pesquisados, Seppälä (2004), da *Université de Genève*, foi a única que tratou das relações conceituais no âmbito da definição terminológica. Em seu trabalho *Composition et formalisation conceptuelles de la définition terminographique*, a autora propõe, a partir da estruturação das relações conceituais presente em Simmen (2002)²⁴, uma tipologia de relações conceituais utilizada para assinalar as diferenças específicas (DE) que compõem as DTs do seu *corpus*, a fim de descrever padrões recorrentes no texto definitório, tais como: tipos de relações, quantidades de DEs por DT e ordem da informação. Ela parte das seguintes relações, explicitadas na Figura 2 (tradução nossa)²⁵:

²¹ Dado esse já demonstrado pelos estudos terminológicos (conforme capítulo 2.1 desse trabalho).

²² Dicionário de Fisioterapia produzido no âmbito do GETerm.

²³ Ainda que já tenhamos postulado nossa posição quanto a essa cisão, neste capítulo será respeitada a terminologia empregada por cada autor, até mesmo como modo de recuperar o contexto de produção de cada um.

²⁴ Simmen (2002), por sua vez, resgata relações conceituais propostas por Arntz et Picht (1995), Dahlberg (1983), Gaudin (2003), Gouadec (1990), ISO (1987), Otman, (1989) e Sager (1993).

²⁵ O original consta do Anexo I.

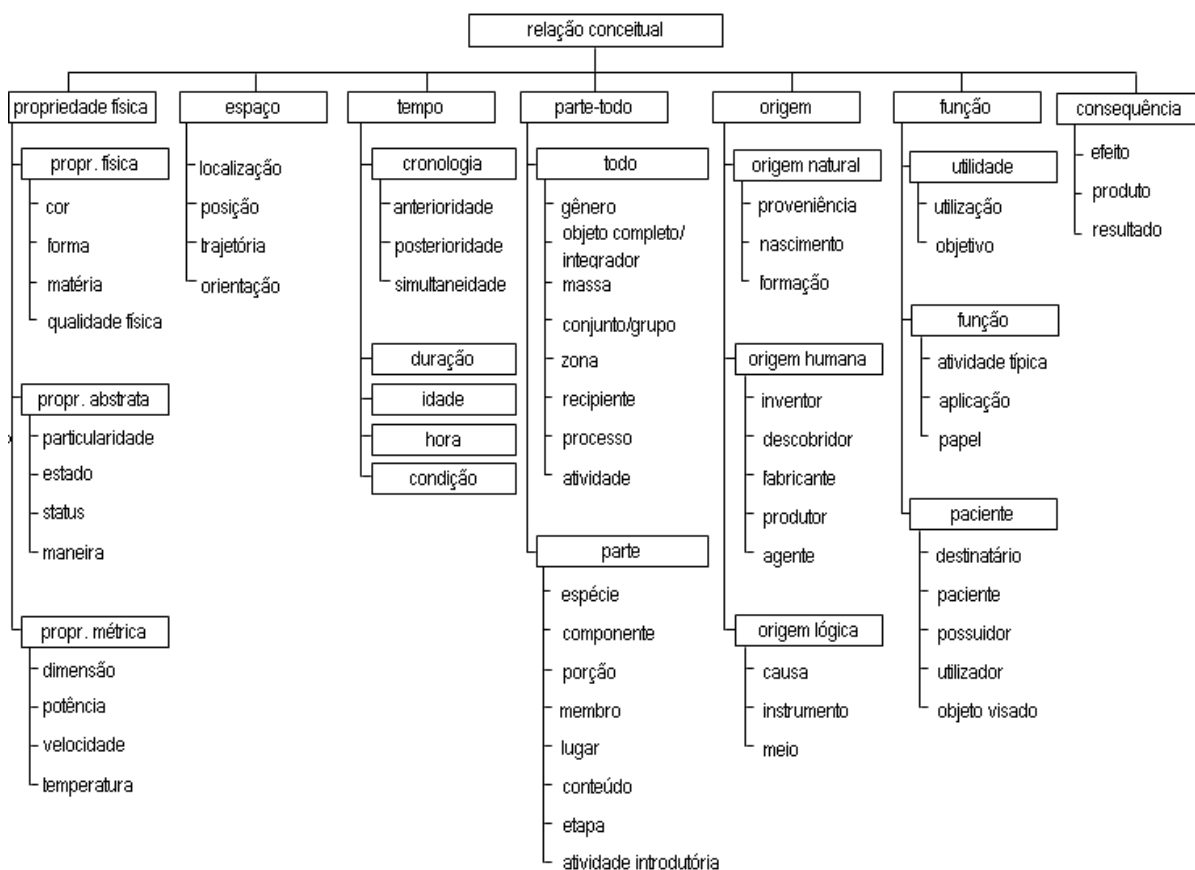


Figura 2 - Tipologia das relações conceituais possíveis (extraída de Seppälä, 2004)

A partir dessas relações e da observação do *corpus*, a pesquisadora propôs uma tipologia de relações conceituais utilizadas na anotação das diferenças específicas (DEs). Ela ainda afirma que essa lista foi modificada e completada à medida que a análise do *corpus* avançava.

O elenco de relações conceituais proposto foi:

Relações conceituais
Agente
Beneficiário
Causa
Condição
Consequência
Conteúdo
Destinatário
Domínio
Extensão
Função
Gênero
Instrumento
Objeto visado
Parte
Paciente
Propriedade abstrata
Propriedade métrica
Propriedade física
Espacial
Temporal
Todo
Utilidade

Quadro 2 - Tipologia das relações conceituais utilizadas em Seppälä (2004)

A pesquisa de Seppälä (2004) mostra-se extremamente relevante para este trabalho de mestrado, pois assim como se constitui a nossa proposta, ela partiu de um *corpus* constituído por textos definitórios, analisou os traços e sistematizou-os com a finalidade de serem formalizados computacionalmente. Embora a nossa pesquisa tenha especificidades próprias, o método de investigação dos textos definitórios se inspira no percurso de trabalho pelo qual Seppälä (2004) passou (tais como: estudo acerca das relações semânticas, elaboração e manipulação do *corpus*, listagem das relações utilizadas, análise dos dados).

2.5.2 Judit Feliu (2004)

Em sua tese de doutoramento, intitulada *Relacions conceptuals i terminologia: anàlisi i proposta de detecció semiautomàtica*, da *Universitat Pompeu Fabra* (UPF) em

Barcelona/Espanha, Feliu (2004) pesquisou as relações semânticas com o propósito de observar em que medida essas relações estão configuradas no *corpus*, a fim de elaborar um protótipo de sistema de detecção semiautomática das relações conceituais, que poderá alimentar banco de dados e ontologias. Assim como Seppälä (2004), o estudo baseou-se em relações semânticas de obras terminológicas até chegar a uma tipologia de relações definitivas. A pesquisadora enfatizou o uso das expressões linguísticas pertinentes a cada relação e criou também um protocolo para cada uma.

As relações contempladas foram:

Relações conceituais
Semelhança (positiva ou negativa)
Inclusão (de classe ou hiponímica)
Sequencialidade (espacial ou temporal)
Causalidade (causal)
Instrumental
Meronímia (parte-todo)
Associação (geral ou especializada)

Quadro 3 - Relações conceituais propostas por Feliu (2004)

Como a nossa análise das DTs implicará, posteriormente, apresentar essas mesmas relações aos usuários do Editor de Definição, o rol de Feliu (2004), embora seja aparentemente restrito (com 7 relações principais), apresenta em cada relação uma noção mais abrangente e, conseqüentemente, uma variedade de expressões linguísticas arroladas em cada item, o que pode demandar maior formalização computacional do Editor de Definição. Como exemplo, segue, no Quadro 4, o protocolo da relação “Causalidade” (tradução nossa)²⁶.

²⁶ O original consta do Anexo II.

SÍMBOLO: CAUS

DESCRIÇÃO: Relação de causalidade: relação que se estabelece entre uma causa e o seu efeito.

PROPRIEDADES: Não simétrica e não necessariamente transitiva

OBSERVAÇÕES: Apresenta subtipos: causal e procedimental

ARGUMENTOS: 2 o > 2 segundo os subtipos

RECURSIVIDADE: sim

CARACTERÍSTICAS a: entidade. atividade. propriedade

CARACTERÍSTICAS b: entidade. atividade. propriedade

OBSERVAÇÕES (a/b): não necessariamente idênticas

EXPRESSÕES DE RELAÇÃO: depender de; produzir; ser a causa (principal) de; deve-se a; implicar; aparecer; contribuir para; ocultar, dar lugar a; <cond.> encontrar:: reforçar; provocar; aumentar; produzir; transformar-se em; aumentar; obter-se (+ger); dar lugar a; permitir (+inf)

Quadro 4 - Protocolo da relação “Causalidade”

Por outro lado, um aspecto muito útil do trabalho de Feliu (2004), do qual nos valeremos na formalização das relações, diz respeito à organização de um protocolo (semelhante ao Quadro 4), onde deverá conter a definição, propriedades da relação e a descrição das expressões linguísticas.

2.5.3 James Pustejovsky (1995)

Uma perspectiva que vem sendo aproveitada pelos estudos da linguagem em geral, em especial pela Terminologia e pela Lexicologia²⁷, é a Estrutura *Qualia* proposta por Pustejovsky (1995), em sua obra intitulada *The Generative Lexicon*. Nesse livro, o autor postula que o significado de uma unidade léxica seja estruturado, na sua gramática, em quatro níveis interligados. E um desses níveis é a Estrutura *Qualia*, na qual estão contidas informações semânticas básicas de uma palavra. Conforme Pustejovsky (1995), a Estrutura *Qualia* define atributos essenciais de objetos, eventos e relações, em função dos *quale*:

²⁷ Como em trabalhos de Höfling (2000), Zavaglia (2002) e Seppälä (2004).

Formal: que distingue um objeto dentro de um conjunto;

Agentivo: que apresenta fatores envolvidos na origem de um objeto;

Télico: que apresenta finalidade e função de um objeto;

Constitutivo: que apresenta relação entre um objeto e suas partes constituintes.

Nesta pesquisa, esse construto da Estrutura *Qualia* parece apropriado na observação preliminar das DTs, já que esses papéis constituem-se questionamentos básicos que se faz, geralmente, acerca de unidades lexicais, e que portanto tornam-se pertinentes constarem das DTs.

Papéis <i>Qualia</i>
Formal: o que é <i>x</i> ?
Constitutivo: <i>x</i> é feito de quê?
Télico: qual a função de <i>x</i> ?
Agentivo: qual foi a causa de <i>x</i> ?

Quadro 5 - Papéis *Qualia* de Pustejovsky (1995)

Segue um exemplo da definição de “neuropraxia” e da mesma DT anotada de acordo com os Papéis *Qualia*:

Neuropraxia: disfunção neuromuscular caracterizada pela diminuição de alguns tipos de sensação, pela preservação distal do nervo e pela ausência de potenciais de fibrilação nos músculos, causada por bloqueio, interrupção ou retardamento de condução nervosa.

<QUALIA FORMAL >

disfunção neuromuscular

<QUALIA CONSTITUTIVO>

caracterizada pela diminuição de alguns tipos de sensação, pela preservação distal do nervo e pela ausência de potenciais de fibrilação nos músculos,

<QUALIA AGENTIVO>

causada por bloqueio, interrupção ou retardamento de condução nervosa.

Se pensarmos nas duas etapas que andam juntas na redação da definição – a ontologia e a base definicional – conseguimos as respostas para as perguntas acima. De fato, são os questionamentos primários que são feitos acerca de um termo, contudo, como afirma

Pustejovsky (1995): i) toda categoria expressa uma Estrutura *Qualia* e; ii) nem todos os itens lexicais carregam consigo valores para cada papel.

Sobre esse dois pontos, Zavaglia (2002) declara que “o primeiro item é importante para se entender como um léxico gerativo sustenta uma representação semântica uniforme composicionalmente de todos os elementos de uma frase. Já o segundo é aplicável e específico para classes semânticas particulares” (Zavaglia (2002, p. 114).

Porém, a DT não se esgota aqui. Ela é, acima de tudo, um texto em que estão implicadas outras relações, no qual, geralmente, extrapola-se o conjunto categorial proposto pelo pesquisador. Compartilhamos da mesma opinião de Höfling (2000, p. 128), estudiosa que aplicou a Estrutura *Qualia* na análise de definições lexicográficas de nomes concretos:

Como vemos, a descrição da estrutura *qualia* pode ser aproveitada na definição lexicográfica. Porém, não podemos esquecer que, ao contrário de Pustejovsky – que procura chegar a um núcleo significativo geral que permita a inserção do item em diferentes contextos – o lexicógrafo quer chegar a uma descrição dos diferentes usos e acepções. Por isso, o verbete do dicionário não pode limitar-se à apresentação da estrutura *qualia*, que deve ser complementada por especificações adicionais que permitam ao consultante o emprego do item nos diferentes contextos de uso.

Desta forma, julgamos que esse modelo de Pustejovsky (1995) possa ser viável na descrição inicial das DTs, entretanto, não se considera que somente os papéis *qualia* sejam suficientes na descrição da maioria das DTs.

2.5.4 Michael P. Jordan (1992)

Para Jordan²⁸ (1992), uma relação semântica constitui uma “noção semântica textual de conexão binária entre quaisquer duas partes de um texto” (*apud* Pardo, 2005, p. 18) e “estas relações capturam a forma como os conhecimentos contidos em um texto se relacionam, sendo completamente desvinculadas das intenções do escritor” (Pardo, 2005, p. 19).

Em seu artigo *An integrated three-pronged analysis of a fund-raising letter*, publicado em 1992, no livro *Discourse Description – Diverse linguistic analyses of a fund-raising text*, Jordan (1992, p. 174) afirma que “as orações e sentenças são ligadas entre si por

²⁸ Pesquisador do qual tomamos conhecimento na disciplina do mestrado sobre RST (*Rhetorical Structure Theory*) ministrada pela Profa. Dr. Lucia Helena Machado Rino. Jordan é linguista da *Queen's University at Kingston*. Suas pesquisas mais conhecidas são nas áreas de Teorias Discursivas (da linha americana) e de PLN. Seu trabalho poderá ser bastante aproveitado nesta pesquisa na parte de descrição das DTs.

significados específicos”. O autor ainda oferece um exemplo, a fim de demonstrar como suas relações se configuram num texto:

1. *(Christ) has now reconciled you (to God)*
2. *by means of dying physically* **MEANS** de 1
3. *in order that you will be holy* **PURPOSE** de 1

As relações propostas por Jordan (1992) podem ser observadas no Quadro 6 (tradução nossa)²⁹.

²⁹ O original consta do Anexo III.

Tipo de Relação	Relações	
Detalhe	identificação	
	Classificação	
	Especificação	
	Aparência	
	Características	
	Função	
	Material	
	Partes	
	Geral	Ativo
	Passivo	
Lógica	Agente	
	Fonte	
	Avaliação	
	Base	
	Causa	
	Efeito	
	Efeito emocional	
	Finalidade	
	Meios	
	Problema	
	Solução	
	Modal	Possibilidade
	Capacidade	
	Precisão	
	Exatidão	
	Necessidade	
	Falta	
	Conclusão	
	Realização	
	Futuro	
	Intenção	
	Ordem	
	Autoridade	

Tipo de Relação	Relações	
Tempo	Anterior	
	Posterior	
	Simultâneo	
	Tempo invertido	
	Tempo	
	Manipulação textual	Elaboração
		Resumo
		Repetição
		Paráfrase
Especial	Previ	
	Transição	
	Inversão colateral	
	Concessão	
	Compatibilidade	
	Contraste	
	Comparação	
	Condicionais	
	Estrutura documental	
	Real-hipotético	
Outras	Transição de dísticos	
	Acompanhamento	
	Circunstância	
	Circunstância invertida	
	Conexão	
	Habilidade	
	Exemplo	
	Extensão	
Localização		
	Localização invertida	
	Maneira	
	Verdade	

Quadro 6 – Tipologia das relações propostas por Jordan (1992)

Enquanto Seppälä (2004) e Feliu (2004) apresentam uma representação mais conceitual, e Pustejovsky (1995) aproxima-se mais do paradigma categorial, a perspectiva de Jordan diferencia-se das demais por apresentar um caráter mais textual, ou seja, o que está em jogo para o pesquisador são as relações implicadas entre as partes do texto. Como se observa no Quadro 6, as relações podem ser diversas, sendo que algumas delas têm

correspondência com os outros autores explorados anteriormente (como “causa”, “efeito” ou “consequência”, “características”, “função”, e etc.).

Como já foi mencionado, a DT constitui-se um texto e, além das relações semânticas e categoriais, ela apresenta características inerentes ao próprio texto. Desta forma, serão consideradas nessa pesquisa também as relações de Jordan (1992) que se mostrarem profícuas.

A proposta de tipologia das relações utilizadas na descrição das DTs será apresentada no Capítulo 4 da dissertação, o qual é destinado à metodologia aplicada na pesquisa.

3 Fundamentos práticos

Esse capítulo será dedicado à explicitação da prática de elaboração da DT no âmbito do GETerm, à apresentação dos ambientes *web* de auxílio à escrita (AMADEUS e SciPo) e, finalmente, à exposição do e-termos, ambiente colaborativo no qual serão implementados os resultados desta pesquisa. Esses itens explorados são de suma importância para o desenvolvimento desse trabalho, pois a partir da perspectiva: de quais são os passos para a redação da DT (cap. 3.1), de que forma pode ser dado um suporte linguístico a um ambiente de auxílio à escrita (cap. 3.2) e, de como o e-Termos se realiza (cap. 3.3), é que direcionaremos a metodologia quanto às análises das DTs.

3.1 Método de elaboração da DT no âmbito do GETerm

O GETerm vem desenvolvendo projetos terminográficos desde 1999, em conformidade com a TCT. A DT sempre foi objeto de estudo do grupo, já que de todas as etapas que compõem a elaboração de uma obra terminográfica, como dicionários, glossário, etc., a definição se destaca por ser a mais custosa e complexa. Como para o GETerm a DT constitui-se um tema de pesquisa e uma prática constante, o método da sua redação foi por diversas vezes testado e alterado (quando preciso), com o propósito de tornar a prática mais sistemática, menos morosa e, sobretudo, com o menor índice de erros.

Neste capítulo, serão apresentados os passos metodológicos quanto à elaboração da DT, os quais são seguidos pelo GETerm em seus projetos terminográficos. São eles: i) Ponto de partida: ontologia; ii) Base definicional; iii) Observação e seleção dos traços semânticos; iv) Redação da DT e v) Validação da DT

3.1.1 Ponto de partida: ontologia

Inicialmente se parte da ontologia, a qual permite a visualização das relações de hiperonímia, hiponímia e co-hiponímia³⁰ entre os termos que compõem o domínio que está

³⁰ Relação entre dois termos que estão no mesmo nível da ontologia, tendo, portanto, o mesmo hiperônimo. Exemplo: dentro do campo *fruta* (hiperônimo), encontramos as unidades *maçã*, *laranja*, *banana*, etc.,

sendo trabalhado. É interessante selecionar o campo semântico com seus respectivos termos, que será definido. Esse procedimento torna a atividade terminográfica mais ágil, pois proporciona: i) recuperação imediata do hiperônimo (ou gênero próximo); ii) buscas por contextos explicativos e definitórios³¹ sobre o termo de forma mais otimizada, devido à aproximação semântica entre os termos do mesmo campo; iii) sistematização dos traços semânticos que compõem a DT devido a uma compreensão mais minuciosa do campo. Observe-se, na Figura 3, um trecho da ontologia do Dicionário de Revestimento Cerâmico (DiRC) armazenada no formato *Folio Views*³².

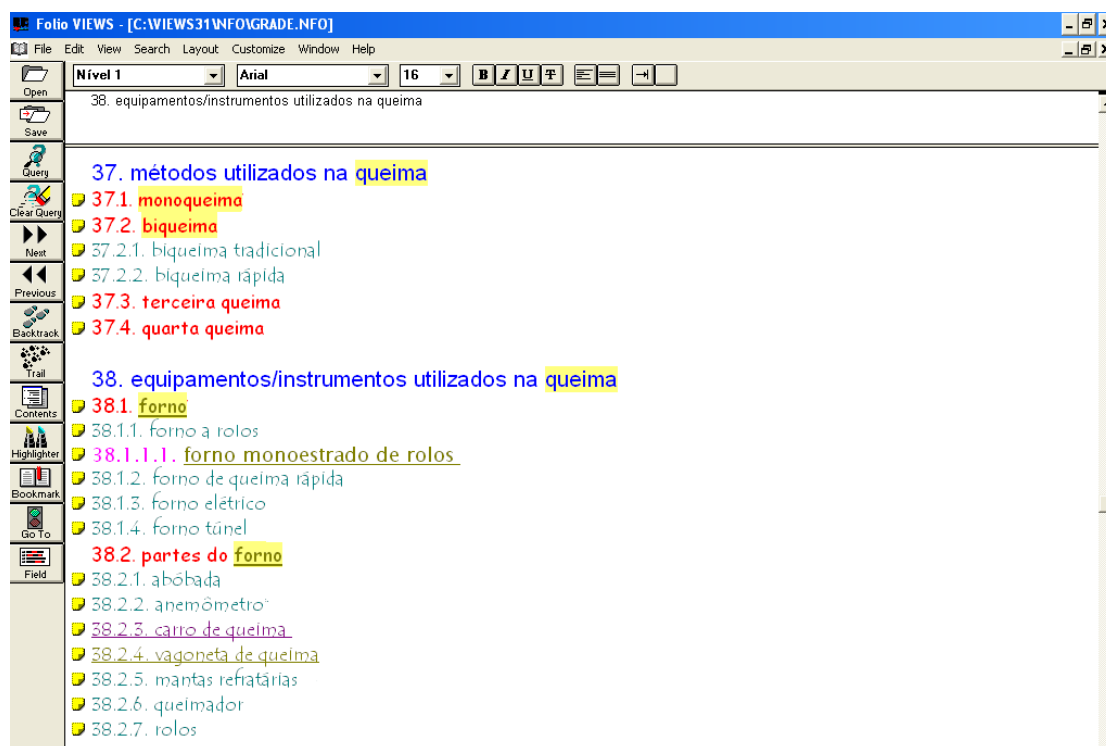


Figura 3 - Tela da Ontologia do DiRC

Na ontologia acima, poder-se-ia, por exemplo, a partir do campo semântico **Método utilizado na queima**, selecionar seus termos hipônimos **monoqueima**, **biqueima** (com os

constituindo todos hipônimos de *fruta*. Assim, a relação entre *maçã*, *laranja* e *banana* é uma relação de co-hiponímia.

³¹ Segundo Aubert (1996), os excertos extraídos de textos especializados podem ser classificados das seguintes formas: **contexto associativo**: apresenta o termo, mas não indica os traços semânticos dele; **contexto explicativo**: apresenta alguns traços semânticos como material, finalidade, funcionamento, similares; **contexto definitório**: conjunto complexo dos traços semânticos distintos do termo, com certa abstração.

³² Trata-se de uma infobase na qual as informações são organizadas em registros, níveis, campos e grupos. O *Folio Views* permite, desta forma, a estruturação hierárquica dos termos.

subitens **biqueima tradicional, biqueima rápida**), **terceira queima e quarta queima**, para que sejam definidos.

3.1.2 Base definicional

O passo seguinte diz respeito à base definicional. Tal denominação foi criada no próprio GETerm, a fim de nomear um repositório de contextos definitórios e/ou explicativos referentes ao termo, compilados de diversas e variadas fontes, tais como: o próprio *corpus* da pesquisa, como também outras fontes não incluídas no *corpus*, mas que podem revelar-se fontes úteis para a obtenção de contextos, a saber: manuais, revistas científicas, dicionários de áreas conexas, dicionários de língua geral, sítios da Internet e demais fontes que se mostrarem úteis para a obtenção de textos pertinentes sobre os termos que serão definidos.

A base definicional é composta por uma tabela com 4 colunas, podendo ser executada no *Microsoft Word* ou no *Microsoft Excel*, por exemplo.³³ As 4 colunas contêm as seguintes informações:

- I. **1ª coluna:** o termo;
- II. **2ª coluna:** os contextos explicativos ou definitórios extraídos do repertório citado anteriormente;
- III. **3ª coluna:** as fontes a partir das quais os contextos foram obtidos;
- IV. **4ª coluna:** a data da inserção do contexto na base.

A base definicional deve ser compreendida como uma etapa importante do trabalho, já que é dela que são extraídas as informações para elaborar as definições. Observe-se, na Figura 4, um trecho retirado da base definicional do DiRC.

³³ No ambiente e-termos, há uma etapa prevista para o preenchimento da base definicional, o que torna o uso dos programas citados (*Word* e *Excel*) desnecessário.

	A	B	C	D
532	Carbonato de Cálcio /Calcium carbonate/ Carbonato cálcico	1) Precipitated calcium carbonate, in low micron sizes, is used as an inorganic filler in "basing cements". These cements consist of a two-stage phenol-formaldehyde resin, calcium carbonate filler and enough hexamethylenetetramine to catalyze the reaction of the resin with heat. Sometimes various organic dyes are added. Material can also be used for insulating coatings for ceramic capacitors and printed circuits. (See Lime and Whiting)	MTC	20/5/2004
533		2) Sal cálcica del ácido carbónico, de fórmula CaCO3. Polvo blanco, insoluble en agua. Por efecto del calor (825 graus Celsius) desprende CO2, transformándose en óxido cálcico (cal viva). Se encuentra en la naturaleza como aragonito, calcita, piedra caliza, mármol, etc. Se emplea en construcción, fabricación de cementos y cal viva, y en la industria química.	VCT	20/5/2004
534		3) (CaCO3): es la metéria prima más usada para introducir calcio em pastas y esmaltes . Se expende em dos formas: natural y artificial. La primera proviene del mineral calcita o rocas calizas molidas a malla No 200, y es la preferida para la fabricación de pastas (suele llamarse carbonato de calcio "pesado"). Existe también una forma más pura, obtenida químicamente por precipitación (es el carbonato de calcio "liviano")(...) [p. 113]	DC_J	21/5/2004
535		4) Carbonato de Cálcio. Óxido de Cálcio. Fundente, insolúvel, refratário e branqueador, torna o esmalte mais duro e resistente além de baixo coeficiente de expansão. Usado na composição da maioria dos esmaltes. Fusão: 2095 a 2485° C.	GCNR	21/5/2004
536		5) CARBONATO DE CÁLCIO : um sólido branco, de fórmula CaCO3, que é pouco solúvel na água. O carbonato de cálcio decompõe-se por aquecimento formando-se óxido de cálcio (cal viva) e dióxido de carbono. Ocorre na natureza como os minerais calcita e aragonita. As rochas contendo carbonato de cálcio dissolvem-se lentamente sob a ação de chuvas ácidas (contendo CO2 dissolvido) provocando dureza temporária. No laboratório, o carbonato de cálcio é precipitado borbulhando dióxido de carbono na solução aquosa de cal viva. O carbonato de cálcio é usado na produção de cal (óxido de cálcio) por aquecimento (decomposição térmica), segundo a reação: CaCO3 CaO + CO2	RQ	21/5/2004
537		6) (...) un incremento en el porcentaje de carbonato cálcico de la pasta trae consigo un aumento en la absorción de agua y un retraso en el inicio de la sinterización por formación de fase líquida, aunque, cuando ésta tiene lugar lo hace de forma mucho más rápida. Este hecho es debido a que, a temperaturas del orden de los 800-900°C, se produce la descomposición de la calcita en óxido cálcico y dióxido de carbono, reaccionando posteriormente el óxido cálcico con la sílice y alumina, procedentes de la deshidroxilación de los minerales arcillosos, y con el cuarzo y feldespato de pequeño tamaño de partícula para formar fases cristalinas cálcicas (silicatos y silicoaluminatos de calcio). El elevado porcentaje de fases cristalinas (cuarzo residual y fases cálcicas) formadas durante la cocción y la escasez de fases vítreas, hasta temperaturas en muchos casos del orden de 1100°C, confieren a las piezas obtenidas a partir de pastas clacáreas una porosidad y coeficiente de dilatación elevados, y un intervalo de cocción	MPF	21/5/2004
538	Carbono/carbon ou matéria	1) símbolo C. Um elemento não metálico do grupo 14 da tabela periódica. O carbono puro apresenta-se na formas principais: diamante e grafite, sendo o diamante muito duro e o grafite um bom condutor de eletricidade e calor. Usado também como lubrificante. Sua	RQ	4/3/2005

Figura 4 - Tela da Base Definicional do DiRC

Almeida et al. (2007) destacam algumas diferenças entre a base definicional e o *corpus*, comprovando que ambos não se confundem:

A base pode receber excertos em distintas línguas, porque o que interessa é a informação sobre o termo e não tanto o contexto de uso. Além disso, não se exige que a base seja representativa, ou seja, que contenha os distintos usos do termo, nem balanceada, isto é, que contemple distintos gêneros textuais (ALMEIDA et al., 2007).

Portanto, a base definicional é uma etapa adicional à composição do *corpus*, na qual são inseridos contextos provenientes do *corpus* e de outras fontes, como já mencionado.

3.1.3 Observação e seleção dos traços semânticos

Na base definicional, a partir da leitura desses contextos referentes a vários termos de mesmo campo semântico, são observados quais são os traços recorrentes na explicação

desses termos. Em seguida, é realizada a sistematização desses traços, a fim de estabelecer quais são importantes para compor a DT e qual a melhor ordem. Observe-se, a seguir, um conjunto de contextos do termo *carbonato de cálcio*. Em cada um dos excertos, foram observados os traços semânticos recorrentes, assinalados nos textos entre colchetes vermelhos³⁴:

[Precipitated calcium carbonate, in low micron sizes]2, [is used as an inorganic filler in "basing cements"]5.

[These cements consist of a two-stage phenol-formaldehyde resin, calcium carbonate filler and enough hexamethylenetetramine to catalyze the reaction of the resin with heat]2. Sometimes various organic dyes are added. [Material can also be used for insulating coatings for ceramic capacitors and printed circuits]5.

[Sal cálcica del ácido carbónico, de fórmula CaCO₃]2. [Polvo blanco, insoluble en agua. Por efecto del calor (825 graus Celsius) desprende CO₂, transformándose en óxido cálcico (cal viva).]3 [Se encuentra en la naturaleza como aragonito, calcita, piedra caliza, mármol, etc.]4 [Se emplea en construcción, fabricación de cementos y cal viva, y en la industria química.]5

(CaCO₃): [es la materia prima más usada para introducir calcio en pastas y esmaltes]5 . [Se expende em dos formas: natural y artificial.]3 [La primera proviene del mineral calcita o rocas calizas molidas a malla No 200]2, [y es la preferida para la fabricación de pastas]5 (suele llamarse carbonato de calcio "pesado").

[Existe también una forma más pura, obtenida químicamente por precipitación (es el carbonato de calcio "liviano").]2(...)

Óxido de Cálcio. [Fundente, insolúvel, refratário e branqueador, torna o esmalte mais duro e resistente além de baixo coeficiente de expansão.]3 [Usado na composição da maioria dos esmaltes.]5 [Fusão: 2095 a 2485 graus Celsius.]3

³⁴ Exemplo extraído de Almeida et al. (2007).

[um sólido branco, de fórmula CaCO₃, que é pouco solúvel na água.]2 [O carbonato de cálcio decompõe-se por aquecimento formando-se óxido de cálcio (cal viva) e dióxido de carbono.]3 [Ocorre na natureza como os minerais calcita e aragonita. As rochas contendo carbonato de cálcio dissolvem-se lentamente sob a ação de chuvas ácidas (contendo CO₂ dissolvido) provocando dureza temporária.]4 No laboratório, o carbonato de cálcio é precipitado borbulhando dióxido de carbono na solução aquosa de cal viva. [O carbonato de cálcio é usado na produção de cal (óxido de cálcio) por aquecimento]5 (decomposição térmica), segundo a reação: CaCO₃ CaO + CO₂

Os traços recorrentes permitiram-nos agrupar um conjunto de quatro traços: constituição, propriedades, origem (como ou onde se obtém, ou como se origina a matéria-prima) e emprego. Observem-se:

[1] o que é (ou termo superordenado/hiperônimo)

[2] constituição

[3] propriedades

[4] origem

[5] emprego

A partir do conjunto composto por esses traços, estabelece-se uma ordem de relevância e de sequência dos traços mais ou menos fixa. É elaborada uma espécie de modelo ou gabarito que deverá orientar a redação da DT de todos ou da maioria dos termos do campo semântico escolhido.

Tal modelo se mostrou tão útil para a definição de substantivos, que ele é aplicado em todos os projetos desenvolvidos no GETerm.

3.1.4 Redação da DT

Em seguida, a DT é redigida, de acordo com os seguintes parâmetros:

- I. a adequação do texto definitório conforme o público-alvo e a finalidade da obra terminográfica em questão;
- II. os traços semânticos estabelecidos;

- III. os contextos inseridos na base definicional;
- IV. as convenções quanto à estrutura textual³⁵.

Quanto à complexidade da DT, Almeida et al. (2007) afirma:

(...) a tarefa da redação da DT é das mais complexas numa pesquisa terminográfica, já que o terminólogo precisa dominar uma multiplicidade de conhecimentos e habilidades. Primeiramente, é fundamental que o pesquisador conheça a área-objeto para a qual elabora o dicionário ou glossário. É necessário, também, que domine aspectos teóricos e metodológicos da Terminologia enquanto disciplina. Além disso, espera-se que o terminólogo tenha noções de Linguística, posto que são acionados conhecimentos de Linguística Textual, Análise do Discurso e demais subáreas que têm o texto como objeto de estudo, afinal, a DT é, antes de tudo, um texto. Por fim, o pesquisador deve conhecer a língua em que desenvolve o produto terminológico, já que a tarefa de elaboração da DT é um verdadeiro exercício de redação (Almeida et al., 2007)

Observe-se, a seguir, a definição de *carbonato de cálcio* elaborada no GETerm para o DiRC, de acordo com os procedimentos descritos acima:

carbonato de cálcio: [carbonato]1 [sólido branco, insolúvel em água, decompõe-se por aquecimento, formando-se óxido de cálcio (cal viva) e dióxido de carbono]2. [Branqueador, torna o esmalte mais duro e resistente, além de propiciar baixo coeficiente de expansão térmica.]3 [Ocorre na natureza como os minerais calcita e aragonita.]4 [É a matéria-prima mais utilizada para introduzir cálcio em massas e esmaltes. Empregado na composição da maioria dos esmaltes (fusão: 2095 a 2485°C).]5

3.1.5 Validação da DT

Depois de redigida, a DT é analisada e revista pelo coordenador do projeto em pauta. Ressalte-se a importância da ontologia e da base definicional na revisão das DTs, já que o coordenador, muitas vezes, vale-se desses dois instrumentos para se certificar se os possíveis problemas no texto são de ordem conceitual ou redacional, pois é possível recorrer

³⁵ Destacamos que as seguintes convenções não foram utilizadas: a DT deve ser composta por uma única oração; deve-se evitar a utilização de paráfrase ou glosa, como também as definições negativas. Anteriormente à etapa da redação da DT, é importante que a equipe envolvida tome decisões acerca de quais convenções serão consideradas, levando em conta o público-alvo e a finalidade da obra terminográfica.

facilmente às informações originais. Depois dessa revisão, as DTs são submetidas aos especialistas de domínio. Somente após a sua aprovação é que ela é inserida na ficha terminológica.

Essa sistematização do trabalho na elaboração de DT revela-se vantajosa, pois o terminólogo não irá definir termos de modo aleatório (como em ordem alfabética, por exemplo), mas sim seguindo uma linha coerente que favorece a compreensão do termo e, conseqüentemente, a qualidade da DT.

Como se observa pela descrição acima, os recursos informáticos utilizados são adaptados a cada objetivo, de acordo com a etapa a qual se destina:

- I. ontologia: aplicativo *Folio Views*;
- II. base definicional: *Word* ou *Excel*;
- III. observação e seleção dos traços e redação da DT: editor de texto (***Word, Bloco de notas, ou qualquer outro editor de texto***);
- IV. validação da DT: editor de texto que tenha recursos semelhantes a “inserir comentários”, por exemplo, como é o caso do *Word*.

Contudo, o método de trabalho deve ser constantemente aperfeiçoado. Como será exposto mais adiante, o e-Termos, está sendo desenvolvido justamente para auxiliar a equipe de projetos terminográficos em língua portuguesa a desenvolver todas essas etapas com recursos específicos e de modo integrado e colaborativo, facilitando o intercâmbio entre as informações.

Como será observado no capítulo 3.2, a área de PLN tem contribuído com o desenvolvimento de recursos, ferramentas e ambientes que auxiliem os usuários na elaboração, correção e aprimoramento de seus textos.

3.2 Informática e escrita

Atualmente, existem ferramentas que auxiliam usuários no processo da escrita a melhorarem a qualidade do texto produzido. De acordo com Aluísio (1995, p. 8), elas podem ser classificadas de acordo com diferentes critérios:

- I. com relação ao **grupo de usuários** a que são destinadas (escritores acadêmicos, editores, secretárias, estudantes);
- II. com relação **ao uso** (para melhorar vários aspectos de um texto, para escrever um texto que segue uma especificação, para aprender como se escreve um tipo de texto);
- III. com relação **à tarefa** a que são destinadas (auxílio à escrita ou ensino da escrita);
- IV. com relação **ao tipo de texto** a que servem (comerciais, acadêmicos, instrucionais);
- V. com relação a qual **estágio do processo da escrita** são adequadas.

A pesquisadora ainda sublinha que, sobre o estágio do processo de escrita (item “V”), três classes de ferramentas podem ser discriminadas (Aluísio, 1995, p.8). São elas:

- I. **ferramentas de pré-processamento do texto:** auxiliam na fase de geração de ideias e compreendem os editores gráficos de redes de notas, os editores gráficos de estrutura em forma de árvore e os *outliners* textuais que permitem expandir ou esconder os diferentes níveis de organização de um texto;
- II. **ambientes para a escrita:** compreendem um conjunto de ferramentas destinadas a dar suporte a uma grande parte do processo de escrita (geralmente, da fase de geração e agrupamento de ideias até a composição de um texto contínuo), ou a todo o processo. Combinam, geralmente, um editor gráfico para as notas/ideias, um editor de estruturas e um editor de textos em um único ambiente; e
- III. **ferramentas para pós-processamento do texto:** são destinadas a melhorar a qualidade do texto e compreendem ferramentas para checagem de estilo, corretores ortográficos e gramaticais, ferramentas estatísticas e corretores de erros mecânicos.

Interessa-nos, nesta investigação, conhecer esses recursos, assim como compreender o percurso (sobretudo o de aspecto linguístico) realizado pelos envolvidos nos projetos. Ressalte-se que as ferramentas que serão apresentadas nos itens 3.2.1 e 3.2.2 nos serviram como motivação para a proposição do futuro Editor de Definição. Aqui, destacaremos dois

ambientes, os quais foram desenvolvidos no âmbito do Núcleo Interinstitucional de Linguística Computacional (NILC)³⁶.

3.2.1 AMADEUS (*Amiable Article Development for User Support*)³⁷

Projeto coordenado pelos pesquisadores Sandra Maria Aluísio e Osvaldo Novais de Oliveira Jr. Trata-se de um ambiente computacional de auxílio ao ensino e produção da escrita acadêmica em língua inglesa para não-nativos. O AMADEUS foi o desdobramento de um estudo feito com alunos brasileiros no exterior, realizado por Fontana et al. (1993), cujo resultado demonstrou que alguns problemas na escrita dos alunos estavam relacionados ao uso inadequado ou ausência de expressões específicas do gênero acadêmico em língua inglesa. Uma saída para esse problema, e que foi implementado no ambiente, é a disponibilização de expressões-padrão utilizadas em textos reconhecidos como bons para que, no caso de dúvida na escrita, o escritor tivesse uma base de bons exemplos da qual ele pudesse se valer. Para a geração dessa estratégia, foi necessária a compilação de expressões e orações extraídas de artigos e livros científicos.

O ambiente, em seu módulo linguístico, conta com quatro ferramentas inter-relacionadas (ferramenta de referência, ferramenta de suporte, ferramenta de crítica e ferramenta tutorial), as quais são ativadas pelos usuários de acordo com seu nível de dificuldade em relação à língua inglesa e à escrita acadêmica. A seguir, cada uma das ferramentas será detalhada:

I. **Ferramenta de Referência**

Essa ferramenta tem como finalidade servir de referência para o escritor, no caso, um não-nativo, provendo uma base de expressões-padrão que podem ser utilizadas na escrita científica, especificamente em artigos acadêmicos. Na compilação dessas expressões, foi utilizado como *corpus* artigos do domínio da Física, a fim de verificar as expressões

³⁶ O Núcleo Interinstitucional de Linguística Computacional (NILC) está sediado no Instituto de Ciências Matemáticas e de Computação (ICMC) da Universidade de São Paulo (USP), campus São Carlos. O grupo é formado por pesquisadores – cientistas da computação e linguistas - da USP/São Carlos, da Universidade Federal de São Carlos (UFSCar) e da Universidade Estadual Paulista (UNESP), campus Araraquara. Desde 1993 vem desenvolvendo projetos nas áreas de Linguística Computacional e Processamento de Linguagem Natural, tais como sumarização e tradução automática, ferramentas de suporte à escrita e desenvolvimento de léxicos e *corpora*. Pode se acessado em: <http://www.nilc.icmc.usp.br/nilc/index.html>

³⁷ <http://www.nilc.icmc.usp.br/nilc/projects/amadeus.htm>

linguísticas pertinentes aos componentes e subcomponentes da estrutura esquemática de um artigo. De acordo com Feltrim (2004, p. 45), a ferramenta utiliza quatro recursos/bases:

- ◆ a base de expressões-padrão, chamada de Base de Padrões;
- ◆ uma lista de componentes e subcomponentes da estrutura esquemática para artigos proposta por Deyes (1982), com ponteiros para a Base de Padrões;
- ◆ uma lista de funções retóricas (p.e. definições, exemplificações, etc.) que também apontam para as entradas correspondentes na Base de Padrões;
- ◆ uma lista de palavras-chave características do gênero (p.e. *report*, *paper*, *model*, *technique*, etc.) que permite acesso rápido às entradas da Base de Padrões, por meio de uma busca simples que retornará todas as expressões contendo a palavra-chave pesquisada.

II. Ferramenta de Suporte

Como afirma Aluísio (1995), embora a ferramenta de Referência seja útil para quebrar o bloqueio inicial com relação à escrita, ainda assim se faz necessária a orientação quanto aos problemas textuais de escolha e ordenação dos componentes do artigo e do melhor padrão de oração para o texto que está sendo redigido. Dessa forma, a ferramenta de Suporte cumpre esse papel de auxiliar o escritor a escrever com certa fluência introduções de artigos do domínio da Física.

O material linguístico assim como a estrutura esquemática que compõe esse estágio, devido ao grau de detalhamento, foi restrito à seção “Introdução” de artigos científicos do domínio da Física, os quais compõem o *corpus*, sendo que a granularidade adotada para a análise foi a oracional. Para tanto, segundo Feltrim (2004), Aluísio (1995) utiliza dois mecanismos principais: i) instanciação de um esquema detalhado de introdução com material linguístico de textos autênticos; ii) aplicação de adaptações interativas que flexibilizam as opções do esquema instanciado, sendo que o primeiro mecanismo fornece a estrutura retórica do texto real para que o escritor veja como os mecanismos de coesão são expressos na língua. O segundo adapta essa estrutura às necessidades atuais do novo escritor.

Contudo, como afirma Aluísio (1995), a ferramenta só deve fornecer suporte à escrita de um texto inicial que, muito provavelmente, deve ser novamente revisado. A decisão final sobre a qualidade do texto está nas mãos dos usuários.

III. Ferramenta de Crítica

Segundo Aluísio (1995, p. 5), pode-se dizer que **críticas** são “sistemas que trabalham em cooperação com humanos para resolver algum problema, sendo que o papel desses sistemas é apresentar uma opinião/crítica, não necessariamente negativa, sobre o produto ou ação gerada pelo humano” (Aluísio, 1995, p. 5).

Assim, o usuário inicia a interação gerando um texto que será criticado pelo sistema; em seguida, com as sugestões fornecidas, faz as alterações que ele considera adequadas, ou ainda, pode procurar obter explicações adicionais.

Assim como as outras ferramentas, a crítica também se baseou na análise da seção introdutória de artigos do mesmo *corpus* e suas estruturas retóricas e mais outras informações retóricas particulares da área.

IV. Ferramenta Tutorial

Esse módulo foi idealizado para o ambiente AMADEUS, porém, ainda se encontra em fase de estudo. Ele difere das outras ferramentas do ambiente por ser adequada quando o usuário é completamente novato ou quando existe algum tópico que ele desconheça inteiramente. Nesse caso, é desejável uma interação tutor/aprendiz, pois o sistema tem mais conhecimento que o usuário (Feltrim, 2004).

No Quadro 7, é exibida a adequabilidade do tipo de ferramenta, de acordo com o conhecimento do usuário (Barros, 2000, apud Feltrim, 2004).

Inglês	Escrita Técnica	Boa Experiência	Alguma Experiência	Nenhuma Experiência
Boa recepção		<i>Ferramenta de Referência</i>		
Problemas de coesão em nível de parágrafo			<i>Ferramenta de Suporte</i>	
Problemas na escrita para um propósito e audiência específicos			<i>Ferramenta de Crítica</i>	
Problemas de coesão em vários níveis				<i>Ferramenta Tutorial</i>

Quadro 7- Adequabilidade do tipo de ferramenta, quanto ao conhecimento do usuário

3.2.2 SciPo (*Scientific Portuguese*)³⁸

O SciPo constituiu-se pesquisa de doutorado de Valéria Delisandra Feltrim, sob coordenação de Maria das Graças Volpe Nunes e Sandra Maria Aluísio, sendo que as três são pesquisadoras do NILC-ICMC/USP - São Carlos. Trata-se de um ambiente *web* de auxílio à escrita acadêmica, baseado em *corpus* e em sistemas de críticas que visa auxiliar, sobretudo escritores iniciantes, na tarefa de estruturação e redação de duas seções de textos acadêmicos – Resumo e Introdução.

O uso da abordagem baseada em *corpus* foi herdada do AMADEUS e parte do princípio de que os textos acadêmicos são passíveis de reuso do ponto de vista de sua estrutura retórica ou de seu material linguístico (expressões padronizadas). Como *corpus* foram utilizados resumos e introduções de teses e dissertações da Ciência da Computação nas seguintes subáreas: Banco de Dados, Inteligência Computacional, Engenharia de Software, Hipermídia, Sistemas Digitais, Sistemas Distribuídos e Programação Concorrente, Computação Gráfica e Processamento de Imagens.

O SciPo apresenta uma lista ordenada de componentes de estrutura esquemática (como Contexto, Lacuna, Propósito, etc.) e suas respectivas estratégias retóricas³⁹ para a elaboração textual de um resumo ou de uma introdução, os quais serão criticados pelo sistema. Observe-se, na Figura 5, uma tela do ambiente:

³⁸ O SciPo pode ser acessado no endereço: <http://www.nilc.icmc.usp.br/nilc/projects/scipo.htm>

³⁹ Por exemplo, na seção “Contexto”, como estratégias retóricas apresentam-se os itens: “Declarar proeminência do tópico”, “Familiarizar termos e conceitos” e “Familiarizar a pesquisa a partir da grande área”.

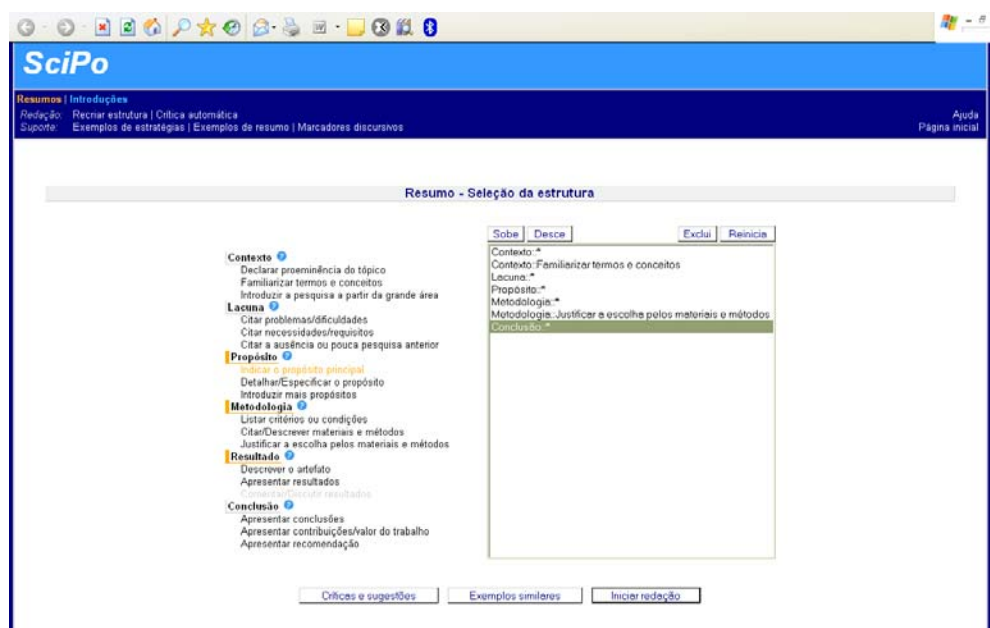


Figura 5 - Tela do Scipo (Resumo – Seleção da Estrutura)

Inicialmente, quanto à Redação, há duas formas distintas de interação, as quais alcançam a mesma saída – o texto pronto. A partir dos *links* **Recrutar estrutura** e **Crítica automática**, o usuário é direcionado para a produção e a avaliação do seu texto:

- I. **recrutar estrutura**: local onde o usuário iniciará seu texto a partir da escolha dos componentes e estratégias que irão compor a estrutura textual (Figura 5). Essa orientação – do planejamento da estrutura para a redação do texto – é denominada *top-down*. Após a seleção dos componentes e estratégias, ao clicar em “Iniciar redação”, o usuário é direcionado à tela “Criação de texto”, onde ele inicia a redação do texto, em formulários, cada qual para um componente, de acordo com as escolhas realizadas. Caso o usuário deixe de selecionar algum componente indispensável, ou ainda, que altere a ordem dos componentes (Exemplo: metodologia – conclusão – propósito), ao clicar em “Iniciar redação”, abrir-se-á uma janela com “Críticas” e “Sugestões”. Entende-se por “Críticas” as alterações necessárias que o usuário necessita fazer em sua estrutura, de acordo com a estrutura mínima aceita pelo sistema. Já “Sugestões”, como o próprio nome diz, são dicas que visam ao enriquecimento da estrutura esquemática, as quais o usuário pode aceitar ou não;

- II. **crítica automática:** local onde o usuário é encaminhado para uma tela denominada “Classificação e Crítica automática”. Aqui, o texto já pronto (ou rascunho) é submetido a um campo a fim de ser classificado pelo sistema (nível sentencial) quanto aos componentes da estrutura, sendo possível também a edição do texto. Essa orientação – partindo do texto pronto – é denominada *bottom-up*. Ressalte-se que o classificador automático apresenta uma precisão média de 70% e que, dessa forma, é possível que haja uma classificação equivocada de algumas sentenças, contudo, o sistema permite que o usuário corrija eventuais equívocos. Em seguida, o texto é submetido a “Críticas”.

O ambiente apresenta ainda outros recursos que podem auxiliar o usuário durante o processo da escrita acadêmica. São eles:

- I. **base de exemplos:** exemplos autênticos (resumos ou introduções) de teses e dissertações em Ciências da Computação (anotados e comentados), que podem ser acessados em qualquer momento da interação com o ambiente. O usuário ainda tem a opção de reutilizar alguma estrutura apresentada pela base em seu próprio texto;
- II. **definição de componentes e exemplos de estratégias:** cada componente e suas respectivas estratégias estão dispostas na ordem em que devem configurar o texto acadêmico. Ao clicar no ponto de interrogação (?), uma janela com a explicação da Estratégia abrir-se-á; clicando nas estratégias, o usuário é direcionado a uma base com exemplos da estratégia solicitada, com expressões-padrão destacadas;
- III. **marcadores discursivos:** expressões organizadas pela função que desempenham no texto (adição, contraste, restrição, etc.);
- IV. **revisor gramatical:** pode ser acionado pelo *link* “Revisar Texto” durante a fase de edição do texto. O usuário pode ainda configurar o revisor de acordo suas preferências. O sistema, ao detectar o “erro”, pode fornecer sugestões que o usuário pode aceitar ou ignorar.

Ao final da redação e revisão do texto, o sistema ainda oferece três modos de armazenamento do arquivo:

- I. com o texto e a estrutura, no formato ASCII;
- II. com o histórico de críticas e sugestões, no formato HTML;
- III. somente com o texto, no formato RTF.

Semelhante aos dois ambientes de auxílio à escrita apresentados neste trabalho, é que prevemos o Editor de Definição. Ainda assim, é imprescindível que sejam expostas também as principais características do e-termos (Capítulo 3.3), especialmente quanto às funcionalidades que se relacionam diretamente com a redação da DT.

3.3 e-Termos (Termos Eletrônicos)

Amparado pelos pressupostos teóricos da TCT, o e-Termos⁴⁰ é um ambiente *web* colaborativo que contempla as atividades de desenvolvimento de terminologias. Como uma aplicação *Computer Supported Collaborative Work* (CSCW), o ambiente prevê seis etapas (módulos) de trabalhos independentes, mas inter-relacionadas, cujo propósito é automatizar ou semiautomatizar as tarefas de criação e gerenciamento do trabalho terminológico.

Sobre o e-Termos, Oliveira (2006) afirma:

O *e-Termos* como um Ambiente Colaborativo Web é, sob uma perspectiva genérica, um sistema Web cuja entrada principal é um *cópus* de especialidade de um determinado domínio do conhecimento, e a saída um produto terminológico (glossário, dicionário, lista de termos, ontologia, etc.) do domínio em questão. Todas as funcionalidades que o sistema fornecerá têm a função primária de servir como suporte a todas as tarefas envolvidas na criação destes produtos. No entanto, tais funcionalidades serão abrigadas em uma única ferramenta computacional (Ambiente) sendo transparente aos usuários. Os usuários, por sua vez, têm a participação colaborativa, interagindo com o sistema em diferentes fases do processo de elaboração do produto, a partir das interfaces Web que o Ambiente fornecerá, tendo acesso compartilhado aos subprodutos gerados pelos módulos, se assim for conveniente ou necessário (Oliveira, 2006, p. 94 e 95).

O e-Termos contempla todo o processo de geração de um produto terminológico: desde a criação automática de *corpora* especializados (etapa 1) até a distribuição e intercâmbio do conjunto de verbetes (etapa 6).

⁴⁰ Disponível em: <http://www.etermos.ufscar.br>

e-Termos Ambiente Colaborativo Web para Criação de Produtos Terminológicos.

O e-Termos

- Principal
- O Projeto e-Termos
- Objetivos
- Etapas de Trabalho
- Ferramentas
- Usuários
- Acesso aos Produtos
- Equipe
- Fale Conosco
- Cadastre-se!

Principal

O **e-Termos**, acrônimo de Termos Eletrônicos, é um ambiente computacional colaborativo web dedicado exclusivamente à pesquisa e à prática terminológica. Seu principal objetivo é viabilizar a criação de produtos terminológicos, sejam eles para os fins de pesquisa acadêmica ou de divulgação, por meio da (semi-)automatização das etapas do trabalho terminológico.

Apoiado nos pressupostos teóricos de uma teoria descritiva de base lingüística, o **e-Termos** implementa 6 etapas de trabalho que representam as fases de criação dos produtos terminológicos. Cada etapa de trabalho abriga tarefas específicas e inerentes ao processo de criação desses produtos, sendo atrelados a eles diferentes ferramentas de análise lingüística, que terão a função de dar suporte às tarefas de Processamento de Língua Natural (PLN) envolvidas neste processo.

O **e-Termos** será desenvolvido para atender as necessidades dos diferentes perfis de usuários constantes neste processo, de maneira que o fluxo de dados entre as etapas de trabalho seja garantido e aconteça de forma transparente.

Para começar a utilizar o **e-Termos**, efetue seu cadastro no link Cadastre-se.

Acesso

e-mail:

Senha:

[Esqueci minha senha!](#)

Links Relacionados

- ▶ AIT
- ▶ Bureau De La Traduction
- ▶ CITRAT
- ▶ Corpógrafo
- ▶ EAFI
- ▶ GT - Lexicologia, Lexicografia e Terminologia
- ▶ ILTEC
- ▶ IULA
- ▶ LEXTERM
- ▶ Office québécois de la langue française
- ▶ Realiter
- ▶ RITERM
- ▶ TERMCAT
- ▶ Terminology Forum
- ▶ Terminometro
- ▶ TERMIP
- ▶ TERMISUL
- ▶ União Latina

NILC - Núcleo Interinstitucional de Lingüística Computacional e GETerm - Grupo de Estudos e Pesquisas em Terminologia

Figura 6 - Tela do e-Termos (Principal)

O grande diferencial do e-Termos é que ele

permite ainda que os diferentes integrantes de uma mesma equipe de pesquisa possam acessar, editar, atualizar, inserir e retirar informações de todos os módulos (*corpus*, ontologia, fichas terminológicas, base definicional, redação de definições, edição de verbetes), bastando conectar-se à internet, buscar a *URL* e utilizar uma senha de acesso. O mesmo procedimento pode ser utilizado para a interação com os especialistas do domínio, ou seja, especialistas previamente selecionados podem opinar, criticar, sugerir alterações, ratificar os dados (lista de termos, definições, etc.) por meio de acesso à internet (Almeida et al., 2006, p.44).

De modo sintetizado, a cada etapa do e-Termos são atribuídas as seguintes atividades:

Etapa 1: compilação automática de *corpus*

Etapa 2: compilação e suporte para análise de *corpus*

Etapa 3: extração automática de candidatos a termos

Etapa 4: edição da ontologia e inserção de termos

Etapa 5: criação e gerenciamento da base de dados terminológica

Etapa 6: edição dos verbetes e intercâmbio dos produtos terminológicos

Cada uma das etapas citadas acima prevê uma série de ferramentas de PLN que irão auxiliar em cada atividade do trabalho terminológico. Destaca-se que a Etapa 5 é a responsável pela **criação e preenchimento da ficha terminológica, da elaboração da base**

definicional e da redação da DT. E é exatamente com essa última tarefa (redação da DT) que este trabalho de mestrado visa colaborar.

Nesta etapa 5, após o preenchimento da Base Definicional, o linguista ou terminólogo segue para a tarefa da Criação e Preenchimento dos campos da Ficha Terminológica (Figura 7). No campo “Definição”, tem-se a opção de inserir logo o texto definitório ou é possível ser direcionado para o “Editor de Definição” (Figura 8).

Principal | Etapa 1 | Etapa 2 | Etapa 3 | Etapa 4 | Etapa 5 | Etapa 6

Projeto **Glossário de Cerâmica Vermelha** **Quinta Etapa**
Perfil Gerente de Projeto

Projeto Recado Mail Base Definicional **Ficha Terminológica** Fórum Ajuda Sair

Preencha os campos da Ficha Terminológica e clique em **Salvar**. Se desejar excluir clique em **Excluir Dados**

Dados do Termo

Termo: **Cerâmica Branca**
[Ver Genealogia](#) - [Ver Relações](#)

Código Termo: 233

Data:

Definição:
Editor de Definição **Link para o Editor de Definição**

Categorial Gramatical: Adjetivo - [adj.]

sinônimo:

Campos em vermelho são obrigatórios.

Salvar Excluir Dados

Chat do e-Termos

Membros da Equipe

- Dayse - você
- Leandro Henrique

NILC - Núcleo Interinstitucional de Linguística Computacional e GETerm - Grupo de Estudos e Pesquisas em Terminologia

Figura 7 - Ficha Terminológica do e-Termos

O Editor de Definição atualmente conta somente com o campo “Definição” ao lado esquerdo da tela, e os contextos definitórios acerca do termo a ser definido (cerâmica branca, como no exemplo da Figura 8), os quais foram coletados e armazenados na Base Definicional do ambiente.

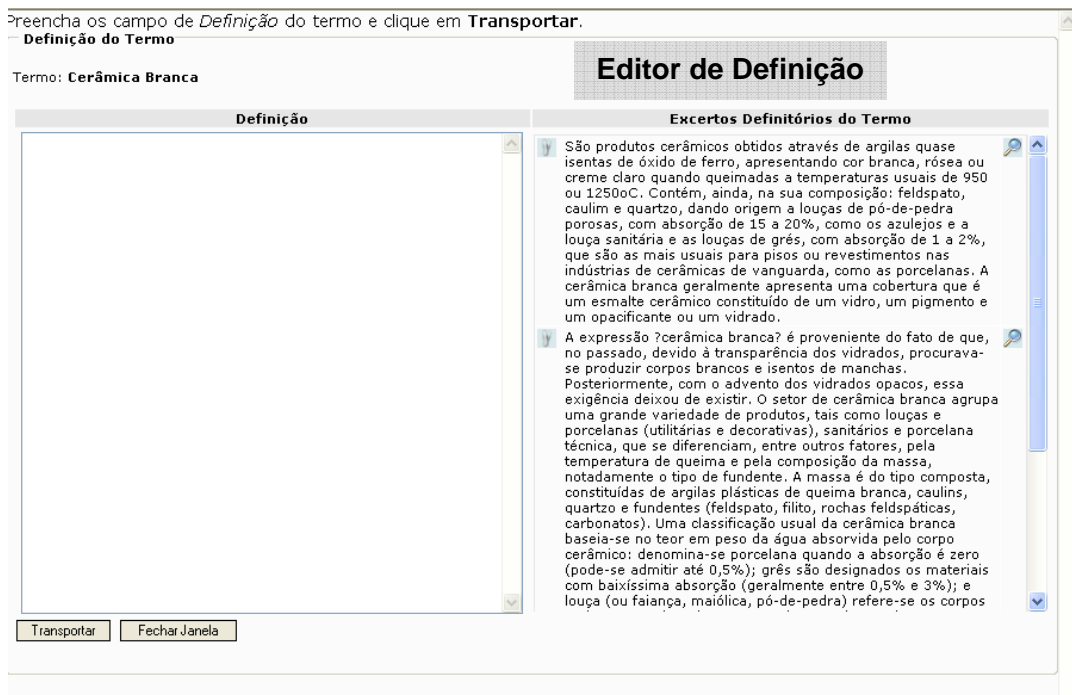


Figura 8 - Editor de Definição do e-Termos

Assim, o desdobramento desta pesquisa de Mestrado será a sistematização dos procedimentos referentes à redação da DT e sua posterior implementação⁴¹ no Editor de Definição do e-Termos.

O próximo capítulo será dedicado à explicitação de toda a metodologia empregada nesta pesquisa, incluindo a apresentação do *corpus*.

⁴¹ Lembramos que a implementação será efetuada pelo informata Leandro Henrique Mendonça de Oliveira, idealizador do e-termos.

4 Metodologia

4.1 Proposta de tipologia das relações semânticas

Após a pesquisa realizada com os quatro autores apresentados no subcapítulo 2.5 e com a análise preliminar do *corpus*, chegamos à seguinte tipologia de relações semânticas utilizadas na descrição das DTs. Primeiramente, são considerados os papéis *qualia* de Pustejovsky e, em seguida, outras relações semânticas que se apresentam no texto definitório.

Como se observa no Quadro 8, em todos os papéis *qualia*, há exemplos de relações semânticas que os constituem, provenientes de Seppälä (2004).

Papéis	Formal	Constitutivo	Télico	Agentivo
Exemplos de relações semânticas	Hiperonímia	Propriedades	Função	Origem
	Sinonímia	Partes	Objetivo	Causa
		Material	Utilização	
		Aparência		

Quadro 8 - Tipologia proposta (1)

Em seguida, são consideradas outras relações que também se mostram produtivas na DT. Foram extraídas de Feliu (2004) e Jordan (1992). Ressalte-se que, em cada uma, podem ainda ser incluídas determinadas relações, observem-se no Quadro 9.

Relações semânticas ⁴²						
Consequência	Meio	Circunstância	Semelhança	Exemplo	Associação	Elaboração
Efeito	Instrumento	Tempo	Co-hiponímia	Hiponímia		Descrição
Produto	Método	Espaço				Indicação
Resultado						Recomendação
						Deteccção

Quadro 9 - Tipologia proposta (2)

⁴² As relações CONSEQUÊNCIA, MEIO, CIRCUNSTÂNCIA, SEMELHANÇA, EXEMPLO, ASSOCIAÇÃO, ELABORAÇÃO, CONSTITUTIVO, AGENTIVO, TÉLICO, com exceção da FORMAL, serão efetivamente anotadas na análise das DTs.

As relações semânticas utilizadas na descrição do *corpus*, assim como a definição e exemplo de cada uma consta do Apêndice A desta dissertação.

Para a realização da análise das DTs, foi preciso, inicialmente, constituir um *corpus* do nosso objeto de estudo. A seguir, são apresentadas as etapas metodológicas envolvidas: **composição do *corpus*, armazenamento do *corpus*, nomeação dos arquivos e manipulação do *corpus* (incluindo a anotação)**. Tais etapas foram cumpridas respeitando uma sequência, contudo, por falhas em geral, muitas vezes, foi necessário retornar a uma etapa anterior da metodologia, a fim de que o erro fosse sanado, impedindo, assim, equívocos na análise dos dados.

4.2 Composição do *corpus*

De acordo com Berber Sardinha (2004, p. 17), “*corpus* é uma coletânea de porções de linguagem que são selecionadas e organizadas de acordo com critérios linguísticos explícitos, a fim de serem usadas como uma amostra da linguagem”. Embora seja formado por porções de linguagem autênticas, o *corpus* é, em si, artificial, criado para cumprir fins específicos.

O *corpus* utilizado na pesquisa é formado por 500 DTs e seus respectivos termos ou entradas, referentes a sete campos semânticos do Dicionário de Revestimento Cerâmico e a oito campos do Dicionário de Fisioterapia. O *corpus* possui no total 28.117 itens léxicos (entre termos e DTs), sendo que, desse número: 26.911 são relativos às DTs constituídas por gênero próximo e diferença específica⁴³ (GPDE) e 25.592 relativos apenas à diferença específica (DE). Ressalte-se que as DTs utilizadas como *corpus* são textos que já passaram pela revisão linguística/terminológica e pela apreciação minuciosa dos especialistas de ambas as áreas.

A eleição desses dois domínios (Revestimento Cerâmico e Fisioterapia) se deu por três motivos: o primeiro pela facilidade ao acesso às DTs, já que foram obtidas no âmbito do GETerm, não havendo, portanto, dificuldades quanto à obtenção das DTs por parte dos autores-pesquisadores. Isso permitiu que não se demandasse tempo em pedido de autorização a pesquisadores-autores de DTs externos. A outra razão, mais terminológica, reside no fato de que as áreas escolhidas – Revestimento Cerâmico e Fisioterapia – são completamente distintas. Enquanto a primeira se insere no domínio maior das Engenharias de Materiais, a

⁴³ Conceito já explicitado no item 2.3 deste trabalho.

segunda se insere no da Saúde. E, certamente, esse dado provoca uma riqueza na análise do *corpus*, já que a proposta final desta pesquisa é a elaboração de um Editor de Definição que auxilie os terminólogos redigirem DTs **em geral**. O último motivo e, sobretudo, o mais importante, é que o *corpus* é constituído de DTs que foram redigidas de acordo com os procedimentos descritos anteriormente. Tais procedimentos envolvem etapas do processo terminográfico que envolvem a redação da DT, conforme já apresentados no item 3.1. Esses dados serão retomados na análise do *corpus* e subsidiarão a construção do Editor.

As DTs estão distribuídas conforme apresentado no Quadro 10.

Domínio	DTs	Campo semântico	Sigla
Revestimento Cerâmico	21	Defeito	DE
	36	Ensaio	EN
	35	Etapa ou Método de Fabricação	EP
	51	Instrumento ou Equipamento	IN
	60	Matéria-Prima	MP
	13	Produto Acabado	PA
	34	Propriedade	PR
Fisioterapia	4	Ação Muscular	AM
	8	Articulação	AR
	91	Disfunção	DI
	9	Exame Complementar	EC
	8	Instrumento de Medida	IM
	13	Movimento Articular	MA
	43	Teste e Medida	TM
	74	Técnica e Recurso Terapêutico	TR
Total	500		

Quadro 10 - Composição do *corpus*

Como se pode observar, superficialmente o *corpus* não está balanceado, já que há campos com 4, 8, ou mesmo 13 DTs; e outros com 60, 74, 91 DTs. Isso ocorre devido ao fato de que o *corpus* reflete duas situações acerca das duas áreas: uma é que os campos semânticos possuem termos em quantidades diferentes; e a outra é quanto à quantidade de DTs já elaboradas pelos terminólogos e revisadas pelos especialistas das áreas.

As DTs foram escolhidas de acordo com os seguintes critérios:

- I. DTs revisadas quanto aos aspectos **linguístico e semântico**;
- II. DTs concernentes a termos da classe morfológica **substantivo**;
- III. DTs elaboradas que seguem o padrão **gênero próximo e diferença específica**;
- IV. DTs elaboradas de acordo com os **traços semânticos previstos**.⁴⁴

Ressalte-se que a análise foi realizada por grupos de DTs de mesmo campo semântico, e que a geração de conhecimento realizada nesta pesquisa se aplicará a definições relativas a termos da classe gramatical **SUBSTANTIVO**, as quais seguem o padrão **GPDE**, pois este se revelou como o mais produtivo na elaboração das DTs dos domínios eleitos.

4.3 Armazenamento do *corpus*

Após a seleção das DTs, a etapa seguinte foi a de armazenamento do *corpus*. Cada termo e sua respectiva DT foi armazenado individualmente em um arquivo “txt” (*notepad* ou bloco de notas), pois como o *corpus* seria tratado posteriormente por programas computacionais, esse é o formato de arquivo ideal para a manipulação do *corpus*. Segue um exemplo na Figura 9, a título de ilustração.

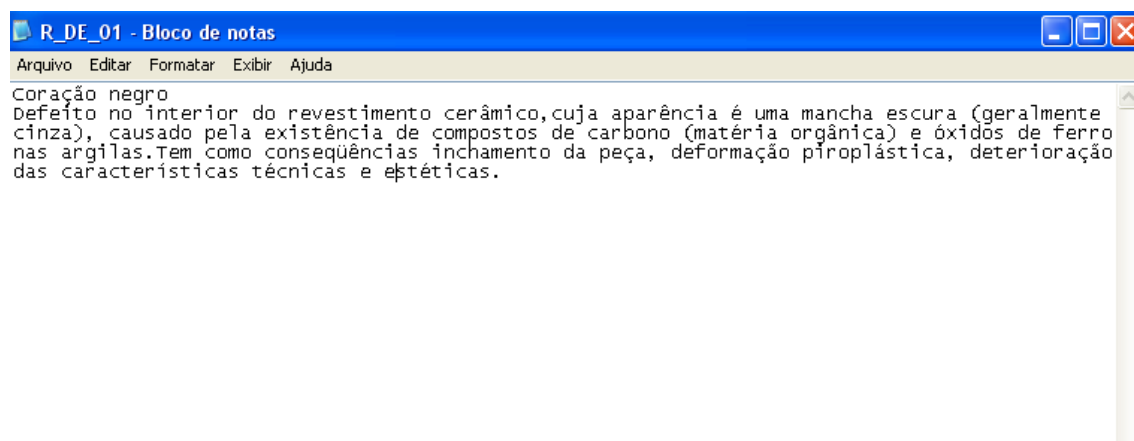


Figura 9 - Exemplo de armazenamento do *corpus*

⁴⁴ Esse critério se refere à eleição de DTs que foram elaboradas segundo o **modelo proposto para cada campo semântico** (como o exemplo do campo MP do domínio de Revestimento Cerâmico apresentado no subcapítulo 3.1.3). Embora tais modelos direcionem a redação das DTs, nem sempre todos os traços que os constituem se configuram no texto definitório. Ainda assim, esse critério é relevante, pois se determinados traços não são contemplados na DT, pode ser devido a fatores como: ausência de contextos definitórios ou explicativos melhores, revisão do especialista, traço não considerado como consenso pelas linhas de pesquisa da área, etc.

Dessa forma, ao final do armazenamento, o *corpus* foi distribuído em 500 arquivos de extensão “txt”, sendo que, nesse momento, o formato eleito foi o *ANSI* (*American National Standards Institute*), pois é uma exigência do Editor de Cabeçalho e do *Wordsmith Tools*, programas computacionais utilizados na pesquisa.

4.4 Nomeação dos arquivos

Em seguida, os arquivos passaram pelo processo de nomeação. Seguiu-se o padrão abaixo:

DOMÍNIO⁴⁵_CAMPO SEMÂNTICO⁴⁶_NÚMERO DO ARQUIVO

sendo que, para os dois primeiros itens acima, foram criadas siglas, a fim de que o nome do arquivo não ficasse muito extenso. Exemplo:

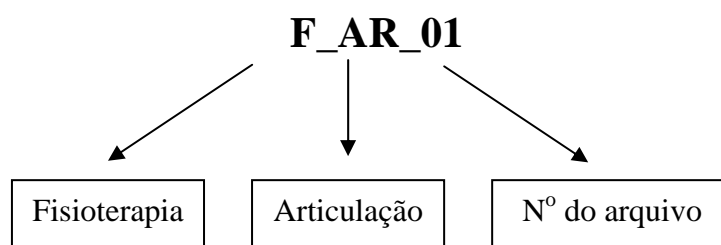


Figura 10 – Padrão de nomeação de arquivos

Esse padrão de nomeação facilitou sobremaneira a recuperação da informação.

4.5 Manipulação do *corpus* (Editor de Cabeçalho)

Após o armazenamento do *corpus* e a nomeação dos arquivos, o passo seguinte foi a realização da anotação do *corpus*, de forma a otimizar o armazenamento e a recuperação das informações contidas em cada arquivo quando fossem realizadas as análises. Nessa etapa, a

⁴⁵ Fisioterapia ou Revestimento Cerâmico.

⁴⁶ Defeito, Disfunção, Técnica e Recurso Terapêutico, Propriedade, etc.

anotação estrutural (nome do arquivo, domínio, subdomínio, campo semântico, etc.) e anotação linguística do *corpus* (termo, definição terminológica, gênero próximo, diferença específica, papéis *qualia*, relações semânticas, etc.) foram efetuadas.

O tipo de anotação utilizada foi o padrão “xml”, acrônimo de *eXtensible Markup Language*, que é formado por uma linguagem que permite fazer dos documentos uma estrutura hierárquica, além de permitir a criação de qualquer tipo de etiqueta de anotação.

Inicialmente, foi gerado um cabeçalho para cada arquivo “txt”. Isso foi feito por meio do Editor de Cabeçalho⁴⁷, um programa computacional que “auxilia o linguista a especificar diversas informações sobre os textos” (Aluísio & Almeida, 2006, p. 174). O programa sofreu algumas pequenas adaptações e simplificações da versão utilizada em outro projeto do GETerm⁴⁸.

Segue, na Figura 11, uma tela do Editor. Note-se, acima, o menu principal, por onde é possível abrir o arquivo desejado. Nesta tela também se encontra um espaço (à esquerda) contendo a relação de arquivos que poderão ser abertos, uma área (à direita) de visualização e edição do texto e o acesso ao editor de cabeçalho, que permitirá o seu preenchimento.

⁴⁷ O Editor de Cabeçalho foi originariamente elaborado para o Projeto Lácio-Web (LW), coordenado pela Profa. Sandra Maria Aluísio, é um projeto organizado pelo NILC, em parceria com o Instituto de Matemática e Estatística (IME) e com a Faculdade de Filosofia, Letras e Ciências Humanas (FFLCH), todos pertencentes à USP. O LW disponibiliza livremente na Web: a) vários *corpora* do português brasileiro escrito contemporâneo, representando bancos de textos adequadamente compilados, catalogados e codificados em um padrão que possibilite fácil intercâmbio, navegação e análise; e b) ferramentas linguístico-computacionais, tais como contadores de frequência, concordanciadores e etiquetadores morfossintáticos. (Aluísio & Almeida, 2006, p. 162). Disponível no site: <http://www.nilc.icmc.usp.br/lacioweb/>

⁴⁸ O projeto em questão é o NanoTerm (Terminologia em Língua Portuguesa da Nanociência e Nanotecnologia: Sistematização do Repertório Vocabular e Elaboração de Dicionário-Piloto), financiado pelo CNPq e iniciado em 2006. O projeto é uma parceria entre a UFSCar, a USP/São Carlos e a UNESP/São José do Rio Preto, e tem como objetivos: 1) a constituição de um *corpus* em língua portuguesa da Nanociência e Nanotecnologia (N&N); 2) uma ontologia em língua portuguesa da área de N&N; 3) a elaboração do primeiro dicionário-piloto de N&N em língua portuguesa (Almeida e Aluísio, 2006, p.174).

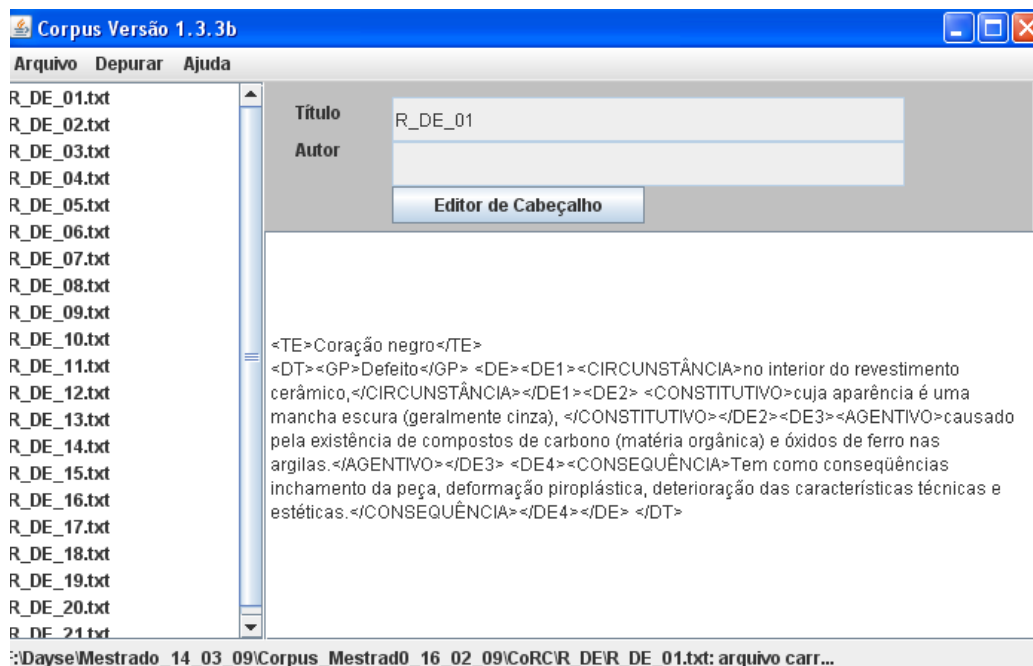


Figura 11 - Tela inicial do Editor de Cabeçalho

O cabeçalho (*header*) utilizado no Editor é formado pelas etiquetas que se encontram em destaque na Figura 12.

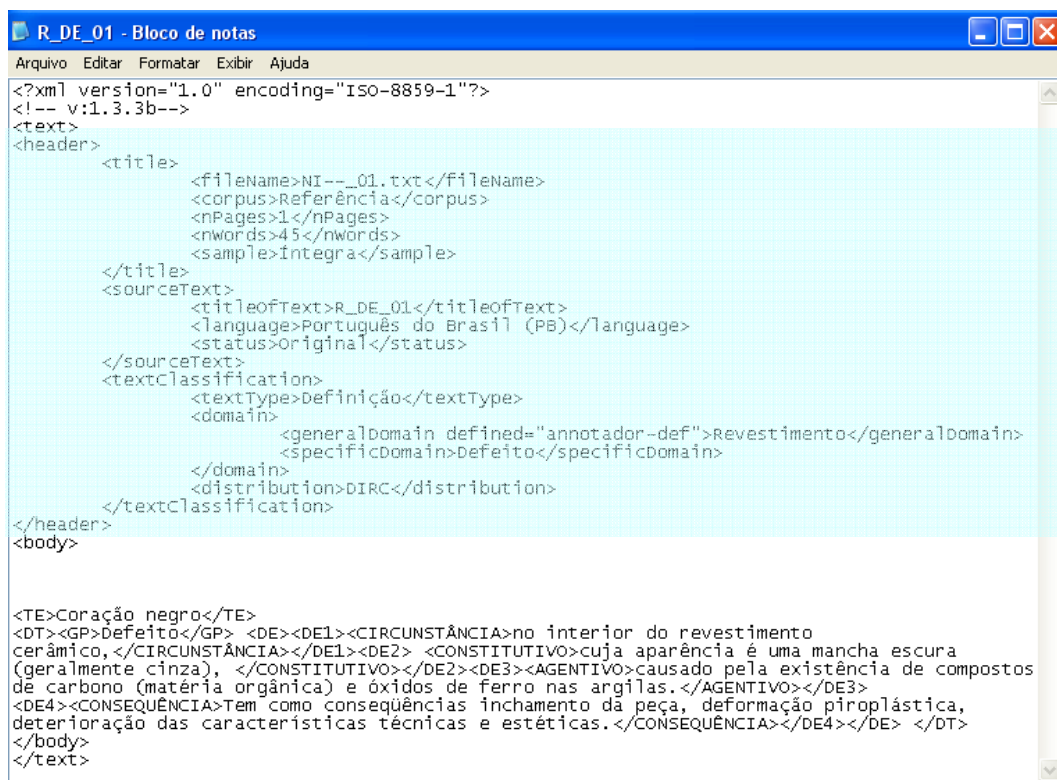


Figura 12 - Cabeçalho com etiquetas “xml” gerado pelo Editor de Cabeçalho

Em relação à anotação linguística, foram substituídas as etiquetas do projeto NanoTerm por etiquetas que se adequassem aos propósitos deste projeto, conforme se pode observar no Quadro 11.⁴⁹

Etiquetas	Etiquetas
<TE>	<AGENTIVO>
<DT>	<CONSTITUTIVO>
<GP>	<TÉLICO>
<DE>	<CIRCUNSTÂNCIA>
<DE1>	<CONSEQUÊNCIA>
<DE2>	<ELABORAÇÃO>
<DE3>	<EXEMPLO>
<DE4>	<LOCALIZAÇÃO>
<DE5>	<MEIO>
<DE6>	<SEMELHANÇA>

Quadro 11 - Etiquetas utilizadas na anotação linguística

Segue, em destaque na Figura 13, a anotação linguística no corpo do texto (*body*).

⁴⁹ Essas etiquetas serão explicadas mais à frente.

```

<?xml version="1.0" encoding="ISO-8859-1"?>
<!-- v:1.3.3b-->
<text>
<header>
  <title>
    <fileName>NI--_01.txt</fileName>
    <corpus>Referência</corpus>
    <nPages>1</nPages>
    <nWords>45</nWords>
    <sample>Integra</sample>
  </title>
  <sourceText>
    <titleOfText>R_DE_01</titleOfText>
    <language>Português do Brasil (PB)</language>
    <status>Original</status>
  </sourceText>
  <textClassification>
    <textType>Definição</textType>
    <domain>
      <generalDomain defined="annotador-def">Revestimento</generalDomain>
      <specificDomain>Defeito</specificDomain>
    </domain>
    <distribution>DIRC</distribution>
  </textClassification>
</header>
<body>

<TE>Coração negro</TE>
<DT><GP>defeito</GP> <DE><DEL><CIRCUNSTÂNCIA>no interior do revestimento
cerâmico,</CIRCUNSTÂNCIA></DEL><DE2> <CONSTITUTIVO>cuja aparência é uma mancha escura
(geralmente cinza), </CONSTITUTIVO></DE2><DE3><AGENTIVO>causado pela existência de compostos
de carbono (matéria orgânica) e óxidos de ferro nas argilas.</AGENTIVO></DE3>
<DE4><CONSEQUÊNCIA>Tem como consequências inchamento da peça, deformação pirolástica,
deterioração das características técnicas e estéticas.</CONSEQUÊNCIA></DE4></DE> </DT>
</body>
</text>

```

Figura 13 - Anotação linguística em “xml” gerado pelo Editor de Cabeçalho

4.6 Protocolo de anotação do corpus

Para a anotação do *corpus*, foi criado um protocolo a ser seguido quanto a cada item. Ressalte-se que esse protocolo foi por diversas vezes alterado, conforme explorávamos o *corpus*. Ao final dessa etapa, ele apresentou a seguinte configuração:

- I. **anotação do CABEÇALHO em geral:** nome do arquivo, domínio geral, campo semântico, entre outros. Nessa etapa, seria necessária a ontologia de cada domínio (Fisioterapia e Revestimento), contudo, a própria nomeação do arquivo já indicava a qual domínio (área) e campo semântico o termo e sua respectiva DT pertenciam;
- II. **anotação do TERMO (<TE>):** foi anotado com essa etiqueta o termo (ou os termos) que encabeça a DT;

Exemplo: <TE>Contração isométrica</TE>

- III. **anotação da DEFINIÇÃO (<DT>)**

Exemplo: <DT><GP>contração</GP> <DE>muscular estática, onde há um aumento de tensão sem mudança no comprimento muscular. É realizado contra um peso que não pode ser removido como, por exemplo, uma parede. Sua finalidade é de manutenção da postura e estabilização das articulações.</DE></DT>

- IV. **anotação do GÊNERO PRÓXIMO (<GP>):** foi considerado como GP a expressão, formada por uma ou mais palavras, que inicia a DT. Com o gênero próximo recupera-se, geralmente, o termo superordenado (hiperônimo) alocado na ontologia, ou ainda um termo sinônimo. Foi considerado como GP também adjetivo(s) ou locução adjetiva que acompanham a primeira unidade lexical da DT, caso do exemplo abaixo;

Exemplo: <DT><GP>articulação cartilaginosa</GP><DE>formada por hialina (cartilagem translúcida, cristalina, branco-azulada) que antes da idade adulta transforma-se em osso. O exemplo mais típico é a sincondrose eseno-occipital que pode ser visualizada na base do crânio.</DE></DT>

- V. **anotação da DIFERENÇA ESPECÍFICA (<DE>):** foi considerada como DE toda a DT com exceção do GP, haja vista que o restante do texto é, na verdade, a ordenação das relações semânticas;

Exemplo: <DT><GP>articulação cartilaginosa</GP><DE>formada por hialina (cartilagem translúcida, cristalina, branco-azulada) que antes da idade adulta transforma-se em osso. O exemplo mais típico é a sincondrose eseno-occipital que pode ser visualizada na base do crânio.</DE></DT>

- VI. **anotação das DIFERENÇAS ESPECÍFICAS (<DE₁>, <DE₂>, <DE₃>, <DE_n>):** cada uma das relações constitui uma diferença específica, como dito acima. Assim, foi considerada como DE_n fragmentos identificados como predicação do GP. Podem variar em extensão e serem constituídos por uma ou mais orações, e até mesmo sentenças;

Exemplo: <DT><GP>disfunção músculo-esquelética</GP> <DE><DE₁> caracterizada por uma infecção inflamatória crônica das articulações e ligamentos da coluna vertebral; </DE₁> <DE₂> tem início nas articulações sacro-ilíacas</DE₂><DE₃>causando alterações vertebrais lentas no sentido de baixo para cima da coluna vertebral. </DE₃></DE></DT>

VII. **anotação das RELAÇÕES SEMÂNTICAS I (papéis *qualia*):** dentro das DE_ns, anotou-se qual papel *qualia* cada uma desempenha. Ressalte-se que nem todos os papéis *qualia* irão se manifestar numa mesma DT. Os papéis podem ser:

- ◆ Formal (o que é?): representado por <FORMAL>⁵⁰
- ◆ Téliico (qual é a finalidade?): representado por <TÉLICO>
- ◆ Constitutivo (do que é feito?): representado por <CONSTITUTIVO>
- ◆ Agentivo (qual é a origem?): representado por <AGENTIVO>

Exemplo: <DT> <GP> disfunção neuromuscular </GP> <DE> <DE1> <CONSTITUTIVO>caracterizada pela incapacidade do indivíduo em realizar ações normais na execução da marcha</CONSTITUTIVO></DE1> <DE2> <AGENTIVO> provocada por uma lesão frontal bilateral.<AGENTIVO> </DE2> </DE> </DT>

VIII. **anotação das RELAÇÕES SEMÂNTICAS II (todas as demais relações).**

Exemplo: <DT> <GP> exame </GP> <DE><DE1> <CIRCUNSTÂNCIA> realizado quando há suspeita de fratura </CIRCUNSTÂNCIA></DE1><DE2> <TÉLICO>e que se constitui em registrar em filme especialmente sensibilizado as estruturas internas do corpo,<TÉLICO> </DE2><DE3><MEIO>através da passagem de raios X ou gama sobre a área examinada.</MEIO> </DE3></DE></DT>

4.6.1 Algumas regras quanto à segmentação das DE_ns

Durante a anotação linguística do *corpus*, surgiram algumas dúvidas a respeito de como proceder quanto à segmentação das DEs e das suas relações semânticas, haja vista que tratamos de linguagem natural e que ela é, por natureza, passível de vários arranjos⁵¹. Desta forma, este capítulo se destina à descrição de algumas regras elaboradas para essa etapa da pesquisa.

⁵⁰ Devido ao fato de o Papel "Formal" ser idêntico ao GP, decidimos não anotá-lo.

⁵¹ Aqui podemos referir a hipótese do inatismo de Chomsky que, entre outras ideias fundadoras, admite a capacidade de o falante produzir infinitos enunciados com finitas estruturas (conferir a obra inaugural do autor intitulada *Syntactic Structures*, publicada em 1957).

- I. cada DE_n é considerada **INTEGRALMENTE** como uma relação. Ou seja, cada DE_n corresponde a uma única relação semântica, não havendo portanto uma DE_n com duas ou mais relações;
- II. devem constituir as DE_ns fragmentos que funcionam como orações subordinadas, paráfrases, parênteses, **JUNTO** com a oração a qual se refere, pois esses fragmentos dizem respeito à oração principal e não ao GP em si. A função desses trechos implicaria em outro estudo, que pode ser realizado mais à frente. Como é o caso de:
- ◆ </DE2> </MEIO> **É realizado por meio da avaliação dimensional, na qual se mede o material antes e após a secagem.** </MEIO> </DE2>
 - ◆ <DE1><CONSTITUTIVO>**constituída essencialmente por silicatos hidratados de alumínio, que podem vir acompanhados de outros minerais (feldspato, quartzo, carbonatos e metais pesados) (...)** </CONSTITUTIVO></DE1>
 - ◆ <DE1><TÉLICO>**que avalia a capacidade que a superfície da placa cerâmica possui de não alterar sua aparência quando em contato com determinados produtos químicos.** </TÉLICO></DE1>
 - ◆ <DE2><CONSTITUTIVO>**Apresenta ciclo de secagem de intervalo de tempo reduzido, ou seja, mais rápido que outros secadores industriais.**</CONSTITUTIVO> </DE2>
 - ◆ <DE1><TÉLICO> (...) **Tem como finalidade transferir tinta ou esmalte pelos orifícios da tela (que são coincidentes com o formato do desenho) até a superfície cerâmica, formando uma estampa.**</TÉLICO></DE1>
- III. DE_ns seguidas, que apresentam a mesma relação semântica, são marcadas cada uma **INDIVIDUALMENTE**, como em:
- ◆ <DE2><AGENTIVO>provocada pela destruição progressiva dos tecidos que compõem as articulações.</AGENTIVO></DE2>
<DE3><AGENTIVO>Ocorre devido ao uso ou movimento excessivo (...)
</AGENTIVO></DE3>

IV. há casos que apresentam sequência de orações (coordenadas) dentro de uma mesma relação. Nessas situações, optamos por marcar todo o conjunto com uma mesma etiqueta. Segue um exemplo:

- ◆ <DE1><TÉLICO>realizado **para identificar** as fases que se encontram sólidos na massa cerâmica, e **evitar** que apareçam fendas na peça queimada.</TÉLICO></DE1>

Observamos ainda que a anotação do *corpus* foi realizada de acordo com os critérios listados acima, subsidiados e revisados por ferramentas computacionais e trabalho humano. Apesar disso, considera-se que a tarefa de atribuição dessas etiquetas, sobretudo a descrição do que considerar como DE_n e a qual relação ela se vincula, é muitas vezes subjetiva, o que significa dizer que um outro anotador pode ter uma outra interpretação sobre a mesma relação.

4.7 Programas utilizados na análise do *corpus*

Após as etapas de anotação e revisão manual do *corpus*, os arquivos foram submetidos a programas computacionais que pudessem dar suporte às nossas análises. Nos próximos subcapítulos, os 3 programas utilizados serão resumidamente descritos.

4.7.1 *Wordsmith Tools*

Na análise do *corpus*, foi utilizado o programa computacional *WordSmith Tools*, versão 3, da autoria de Mike Scott, da Universidade de Liverpool. O programa é composto basicamente de ferramentas (*Wordlist*, *Keywords* e *Concord*), e em cada uma dessas ferramentas, há uma gama de outros recursos que auxiliam na descrição linguística. A escolha por esse programa se deu pelo fato de que ele apresenta:

- I. **bom desempenho estatístico:** por meio da ferramenta *Wordlist* foi possível recuperar informações acerca da quantidade de palavras, sentenças, índice de riqueza vocabular do *corpus* de modo integral, por campo semântico ou ainda individualmente (cada DT);
- II. **manipulação de vários arquivos ao mesmo tempo:** o programa permite manipular vários arquivos simultaneamente, fornecendo assim informações (estatísticas e linguísticas) sobre cada arquivo em específico;
- III. **leitura de etiquetas “xml” (tags):** recurso que permite a leitura de etiquetas, e por conseguinte, a criação de buscas específicas por meio delas.

Foram utilizadas, especificamente, as ferramentas *Wordlist* (sobretudo a janela dos dados estatísticos referentes aos arquivos selecionados) e *Concord* (para realização de buscas específicas das relações). Observem-se, na Figura 14, duas telas do programa, uma da ferramenta *Wordlist* e outra do *Concord*, a título de ilustração.

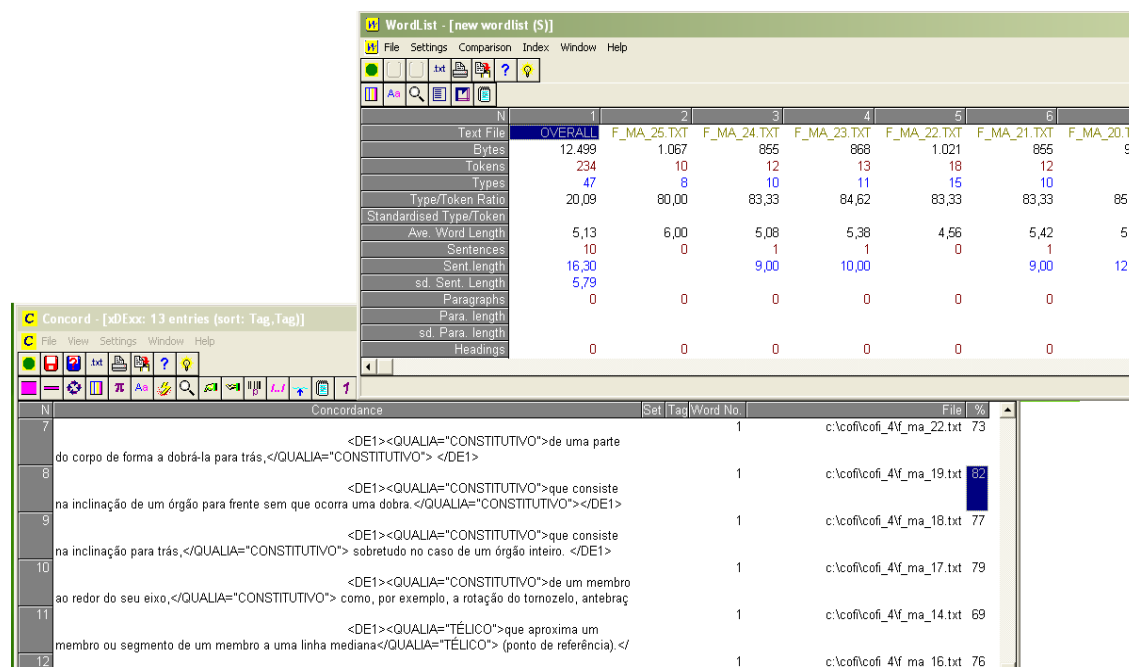


Figura 14 - Telas “Concord” e “WordList” do programa *WordSmith Tools*

Além disso, todas as buscas tiveram como filtro as etiquetas “xml”. Para isso, foi muito produtivo o recurso *Tags* (inserido em *WordSmith Tools Settings*), que permite

fundamentalmente incluir ou não as etiquetas na visualização das ocorrências e selecionar partes do texto a serem analisadas. Observem-se, na Figura 15, telas ilustrativas desse recurso.

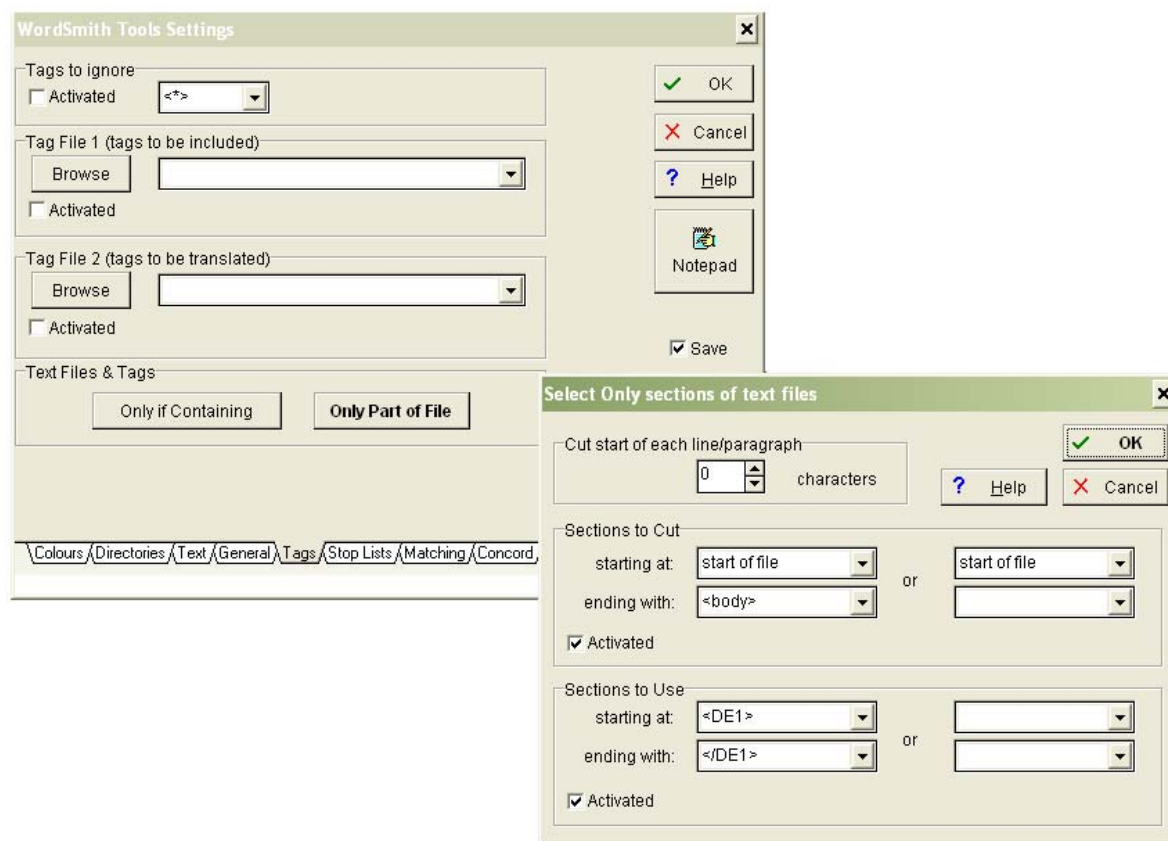


Figura 15 - Telas sobre “tags” do programa *WordSmith Tools*

De modo resumido, o *WordSmith Tools* foi útil quanto à recuperação de dados estatísticos das DTs e quanto à descrição das relações presentes nas DTs. Essas informações foram organizadas e serão demonstradas no Capítulo 5 do trabalho.

4.7.2 Extrator semiautomático de etiquetas “xml”

O programa escrito em *Perl*⁵² idealizado, especialmente para esta pesquisa, pelo pesquisador de Mestrado Marcelo Adriano Amancio do âmbito do NILC (USP/São Carlos), tem a finalidade de extrair de modo semiautomático fragmentos anotados com as etiquetas em “xml” para um arquivo que agrupe os fragmentos pertinentes a uma mesma etiqueta.

⁵² Acrônimo de *Practical Extraction And Report Language*. Para saber mais, acesse: <http://www.perl.com>

A partir de: i) um arquivo que contém todas as 500 DTs⁵³ anotadas (chamado de **Corpus.txt**) e de ii) um arquivo que contém o nome das dez etiquetas que nos interessam (chamado de **Etiquetas.txt**), são gerados os fragmentos agrupados em seus respectivos arquivos, segundo a sua etiqueta.

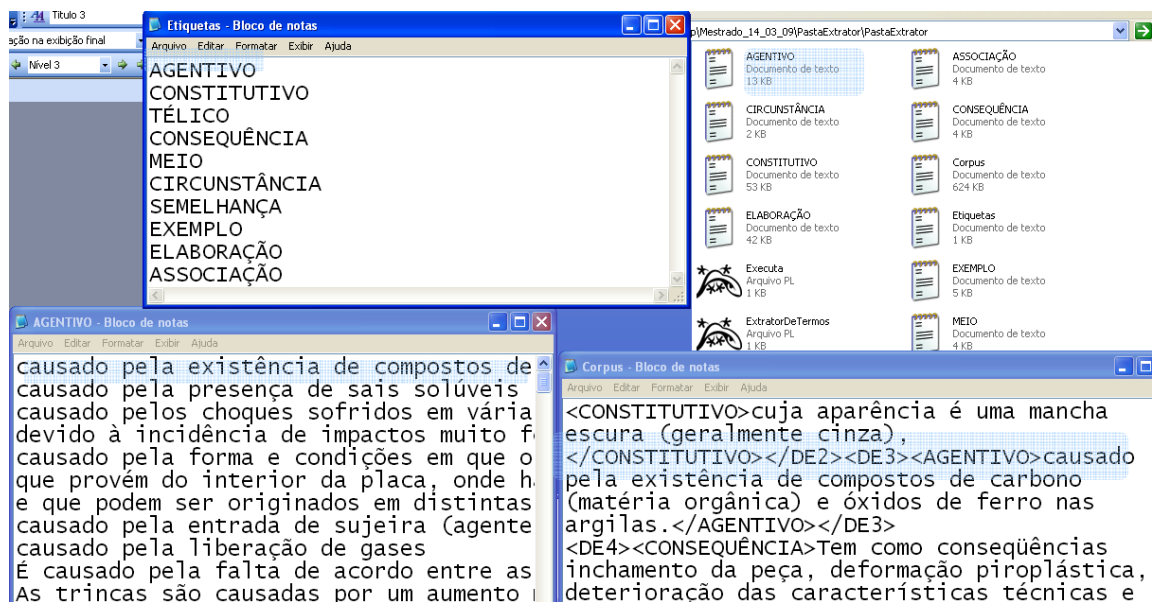


Figura 16 - Extrator semiautomático de “xml”

Esse programa foi utilizado porque, além de possibilitar outra disposição dos dados (por agrupamentos de fragmentos que apresentam a mesma etiqueta semântica), foi um modo produtivo encontrado para que eles fossem visualizados e para que, a partir dos arquivos gerados, as expressões linguísticas relativas a cada etiqueta de relação semântica fossem anotados em “xml” também (com o auxílio do programa que será demonstrado em 4.6.3).

4.7.3 Anotador semiautomático de etiqueta “xml”

O anotador foi outro programa idealizado para essa pesquisa pelo mesmo pesquisador do NILC, Marcelo Adriano Amancio, a fim de realizar anotação semiautomática em *xml* das expressões linguísticas selecionadas a partir dos arquivos gerados pelo programa citado anteriormente.

⁵³ A fim de juntar os 500 arquivos em “txt” num único arquivo, foi utilizado um recurso em **DOS** (sigla para **Disk Operating System**) que realiza a geração automática de um único arquivo a partir de vários outros arquivos.

Trata-se de um anotador na linguagem de programação *python*⁵⁴ que trabalha do seguinte modo: a partir do: i) arquivo “txt” nomeado com a etiqueta “xml” que deverá ser usada na anotação (Exemplo: Marcador_agentivo, Marcador_exemplo, Marcador_télico, etc.), contendo no corpo do texto todas as expressões que deverão ser localizadas e anotadas posteriormente e de ii) um ou mais arquivos que constam do *corpus* que deverão ser anotados com as etiquetas, é gerado arquivo anotado com as etiquetas equivalentes às expressões encontradas.

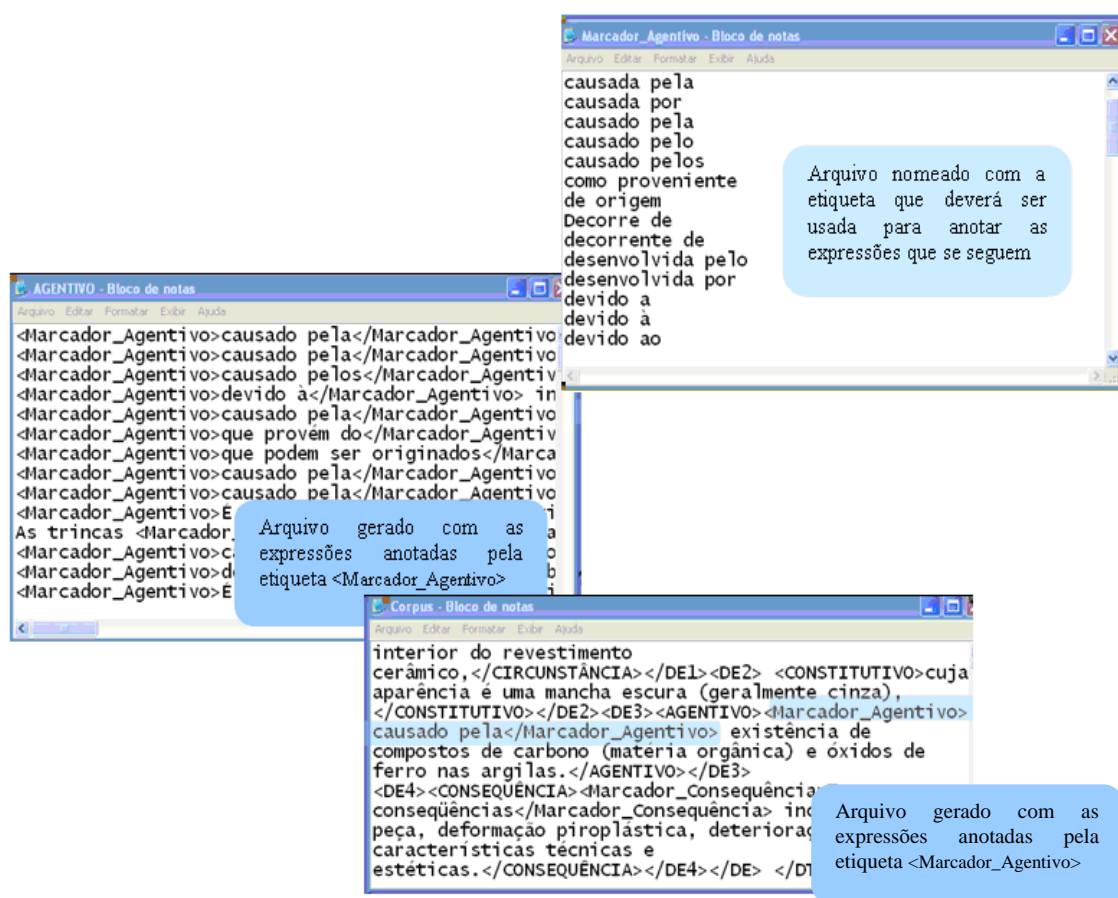


Figura 17 – Arquivos referentes ao Anotador semiautomático de “xml”

Nessa etapa de anotação das expressões linguísticas, optamos por gerar dez arquivos referentes às relações semânticas utilizadas na pesquisa (AGENTIVO, CONSTITUTIVO, TÉLICO, ELABORAÇÃO, MEIO e etc.), conforme a segunda tela exibida na Figura 17; além de realizar essa anotação das expressões linguísticas no próprio *corpus* integral com as 500 DTs, conforme à terceira tela da Figura 17.

⁵⁴ <http://www.python.org>

Em suma, os dois programas elaborados para essa pesquisa nos permitiram realizar anotação das expressões linguísticas de um modo otimizado e, além disso, possibilitaram a geração de arquivos (ou *subcorpora*) que poderão ser aproveitados tanto nesta pesquisa como em outras.

5 Análise do *Corpus*

Após a anotação do *corpus*, e por meio dos programas apresentados nos subcapítulos 4.6.1; 4.6.2 e 4.6.3, foi iniciada a análise das DTs. Conforme já mencionado, as análises foram realizadas por grupos de DTs de cada campo semântico, e a seleção do que era relevante constar dessa análise foi realizada a partir das perguntas pertinentes à pesquisa em si, assim como outros dados que podem ainda serem aproveitados em trabalhos futuros⁵⁵.

A análise do *corpus* foi efetuada em diferentes níveis. Inicialmente partimos de algumas **informações gerais** (item I); em seguida abordamos aspectos **semânticos** da DT (item II), como as combinações de DE_ns mais produtivas, posição de determinadas relações e vínculo entre a escolha de certas DE_ns nas DTs e o campo semântico no qual o termo a ser definido se insere. No item III, expomos a quantidade de **itens léxicos por DE** e **por relação semântica**; no item IV, são apresentadas as **expressões linguísticas** empregadas pertinentes em cada relação; e no item V, quanto à produtividade de alguns **morfemas** e seus respectivos **campos semânticos**.

I. Informações gerais:

Considerando os dois domínios analisados, foram analisadas 500 DTs agrupadas em 15 campos semânticos diferentes (média de 33 DTs por campo semântico). Quanto à quantidade de DE_ns analisadas, tem-se a seguinte configuração:

QTDE. DE DE _n s ANALISADAS	FREQUÊNCIA ABSOLUTA ⁵⁶	FREQUÊNCIA RELATIVA ⁵⁷
DE1	500	40,3%
DE2	455	36,7%
DE3	221	17,8%
DE4	56	4,5%
DE5	9	0,7%
TOTAL	1241	100%

Tabela 1 - Quantidade de DE_ns analisadas

⁵⁵ Devido à quantidade de resultados obtidos, optamos por inserir neste capítulo as informações de um modo geral, ou seja, englobando as 500 DTs. No Apêndice B, são apresentados os dados referentes às DTs por campo semântico.

⁵⁶ Frequência absoluta (freq.abs.) corresponde à quantidade de vezes que um valor aparece como dado da mostra.

⁵⁷ Frequência relativa (freq. rel.) corresponde ao quociente entre a frequência absoluta do valor da variável e o número total de observações.

Todas as DTs analisadas do *corpus* apresentam no mínimo 1 DE e no máximo 5 DEs (totalizando a frequência de 1241 DE_ns), o que comprova a afirmação de Bessé (1996, apud Seppälä, 2004, p. 149) quanto ao aspecto estrutural de que “uma definição terminográfica não comporta mais de 5 diferenças específicas” (tradução nossa).

Quanto à quantidade de DE_ns que compõe as DTs, os números mais expressivos correspondem a DTs formadas por 2 DEs e 3 DEs (46,8% e 33,0% das DTs, respectivamente).

QTDE. DE _n s POR DT	FREQ. ABS.	FREQ. REL.
DT=GP+DE(DE1)	45	9,0%
DT=GP+DE(DE1+DE2)	234	46,8%
DT=GP+DE(DE1+DE2+DE3)	165	33,0%
DT=GP+DE(DE1+DE2+DE3+DE4)	47	9,4%
DT=GP+DE(DE1+DE2+DE3+DE4+DE5)	9	1,8%
TOTAL	500	100%

Tabela 2 - Quantidade de DE_ns por DT

Do total de 1241 DE_ns analisadas, as 10 relações semânticas utilizadas para anotar as DE_ns apresentam a seguinte frequência:

RELAÇÃO	FREQ. ABS.	FREQ. REL.
CONSTITUTIVO	388	31,3%
TÉLICO	288	23,2%
ELABORAÇÃO	248	20,0%
AGENTIVO	121	9,8%
EXEMPLO	47	3,8%
MEIO	46	3,7%
CIRCUNSTÂNCIA	41	3,3%
CONSEQUENCIA	37	3,0%
ASSOCIAÇÃO	20	1,6%
SEMELHANÇA	5	0,4%
TOTAL	1241	100,0%

Tabela 3 - Relações semânticas no *corpus* de acordo com a frequência

As relações semânticas mais frequentes no *corpus* são **CONSTITUTIVO** (31,3%), **TÉLICO** (23,2%), **ELABORAÇÃO** (20,0%) e **AGENTIVO** (9,8%). Ressalte-se que a relação **ELABORAÇÃO** não integra os papéis *qualia* propostos por Pustejovsky (1995). Ainda assim, a alta ocorrência das outras três relações confirma uma das hipóteses de pesquisa deste trabalho, ou seja, os papéis *qualia*, que aqui são considerados como relações semânticas, são compreendidos como “informações semânticas básicas de uma palavra”. De fato, esses papéis se mostraram

recorrentes no corpus, totalizando 64,3% das relações utilizadas, confirmando que as relações que integram os papéis *qualia* de Pustejovsky (1995) coincidem com aquelas que os terminólogos apontam como as mais frequentes.

No que se refere à frequência da relação **ELABORAÇÃO** (20,0%), entendemos que se deve ao fato de que essa relação abarca outras relações que decidimos não anotar individualmente, como **DESCRIÇÃO**, **INDICAÇÃO**, **RECOMENDAÇÃO**, **DETECÇÃO**, pois constituem relações muito específicas de determinados domínios. Ressaltamos aqui que a nossa proposta é a sistematização de relações mais gerais que possam ser empregadas em distintos campos do conhecimento.

II. Informações semânticas:

Quanto à ordem na qual as relações se apresentam nas DTs, observamos que, considerando o critério de 2 DE_{ns} em sequência (desprezando a posição se a relação integrava a DE1, DE2, DE_n), temos as seguintes sequências produtivas:

SEQUÊNCIA DE RELAÇÕES	FREQ. ABS.	FREQ. REL.
TÉLICO + ELABORAÇÃO	110	18,0%
CONSTITUTIVO + TÉLICO	98	16,1%
CONSTITUTIVO + AGENTIVO	88	14,4%
ELABORAÇÃO + ELABORAÇÃO	54	8,9%
CONSTITUTIVO + ELABORAÇÃO	49	8,0%
TÉLICO + CONSTITUTIVO	44	7,2%
CONSTITUTIVO + EXEMPLO	30	4,9%
CIRCUNSTÂNCIA + CONSTITUTIVO	25	4,1%
CONSTITUTIVO + CONSEQUÊNCIA	22	3,6%
CONSTITUTIVO + MEIO	21	3,4%
CONSTITUTIVO + ASSOCIAÇÃO	17	2,8%
TÉLICO + MEIO	15	2,5%
AGENTIVO + CONSEQUÊNCIA	14	2,3%
ELABORAÇÃO + TÉLICO	13	2,1%
AGENTIVO + CONSTITUTIVO	10	1,6%
TOTAL	610	100,0%

Tabela 4 - Sequências produtivas das relações semânticas

As três sequências mais recorrentes no *corpus* são de **TÉLICO + ELABORAÇÃO** (18,0%), **CONSTITUTIVO + TÉLICO** (16,1%) e **CONSTITUTIVO + AGENTIVO** (14,4%), ou seja, como já esperado, referem-se às relações mais produtivas do *corpus*. Essas combinações podem ser úteis como sugestões para o usuário do Editor de Definição. Por exemplo, na redação da DT

de um substantivo x , a sequência **CONSTITUTIVO + TÉLICO** sugere ao redator que descreva “do que é constituído” e “qual a função” do termo.

Também foram analisadas as relações segundo o lugar (DE_n) que as relações semânticas ocupam no texto da DT. Chegamos aos seguintes resultados:

DE1	FREQ. ABS.	FREQ. REL.
CONSTITUTIVO	292	58%
TÉLICO	151	30%
CIRCUNSTÂNCIA	35	7%
AGENTIVO	12	2%
ELABORAÇÃO	5	1%
MEIO	4	1%
EXEMPLO	1	0%
TOTAL	500	100%

Tabela 5 - Relações semânticas na posição DE1

As relações semânticas que mais ocorrem como DE1, ou seja, como a primeira diferença específica a aparecer na DT, são **CONSTITUTIVO** e **TÉLICO** (58% e 30% respectivamente), sendo que essas duas relações somam 443 (88%) das 500 DE1s analisadas. Em paralelo, relações como **AGENTIVO**, **ELABORAÇÃO**, **MEIO** e **EXEMPLO** ocorrem nesta posição poucas vezes, com 22 ocorrências (4%). Relações como **CONSEQUÊNCIA**, **ASSOCIAÇÃO** e **SEMELHANÇA** não ocorrem nenhuma vez como DE1.

Já na posição DE2, todas as 10 relações apareceram pelo menos duas vezes (Tabela 6). Ao contrário do que ocorre com a DE1, a DE2 apresenta uma distribuição mais equilibrada quanto às relações que ocorrem nessa posição (de 23% a 0,2%):

DE2	FREQ. ABS.	FREQ. REL.
ELABORAÇÃO	106	23%
TÉLICO	89	20%
AGENTIVO	79	17%
CONSTITUTIVO	78	17%
EXEMPLO	32	7%
MEIO	30	7%
CONSEQUÊNCIA	20	4%
ASSOCIAÇÃO	16	4%
SEMELHANÇA	3	1%
CIRCUNSTÂNCIA	2	0%
TOTAL	455	100%

Tabela 6 - Relações semânticas na posição DE2

As relações **ASSOCIAÇÃO** e **SEMELHANÇA** são as que menos aparecem no *corpus* (com 20 e 5 ocorrências, conforme apresentado na Tabela 3); e quando elas ocorrem nas DTs é, na maioria das vezes, na posição DE2 (com 16 e 3 ocorrências, respectivamente, o que equivale a 80% e 60% das vezes que ocorrem no total no *corpus*). Esses dados são apontados na Tabela 7.

RELAÇÃO	TOTAL (FREQ. ABS.)	DE2 (FREQ. ABS.)	FREQ. REL.
ASSOCIAÇÃO	20	16	80%
SEMELHANÇA	5	3	60%

Tabela 7 - Dados sobre as relações ASSOCIAÇÃO e SEMELHANÇA

Todas as 10 relações aparecem também no mínimo duas vezes na DE3, e a ordem do mais frequente para o menos frequente praticamente é a mesma apresentada na Tabela da DE2. A relação **ELABORAÇÃO** se apresenta com 47% das DE3 do *corpus*.

DE3	FREQ. ABS.	FREQ. REL.
ELABORAÇÃO	104	47%
TÉLICO	39	18%
AGENTIVO	25	11%
CONSTITUTIVO	13	6%
EXEMPLO	12	5%
CONSEQUÊNCIA	11	5%
MEIO	8	4%
ASSOCIAÇÃO	4	2%
CIRCUNSTÂNCIA	3	1%
SEMELHANÇA	2	1%
TOTAL	221	100%

Tabela 8 - Relações semânticas na posição DE3

No grupo de relações referentes à DE4 (Tabela 9), a **ELABORAÇÃO** aparece em 50% das 56 DE4 do *corpus*, e a relação **CONSTITUTIVO** deixa de figurar entre as mais frequentes, aparecendo com 3 ocorrências apenas. Acreditamos que isso se deva ao fato de que a caracterização ou constituição de um termo seja mais frequente no início do texto definitório (conferir Tabela 5).

DE4	FREQ. ABS.	FREQ. REL.
ELABORAÇÃO	28	50%
TÉLICO	8	14%
CONSEQUÊNCIA	6	11%
AGENTIVO	4	7%
MEIO	4	7%
CONSTITUTIVO	3	5%
EXEMPLO	2	4%
CIRCUNSTÂNCIA	1	2%
TOTAL	56	100%

Tabela 9 - Relações semânticas na posição DE4

SEMELHANÇA e **ASSOCIAÇÃO** não apresentaram nenhuma ocorrência na posição DE4. Provavelmente, essas duas relações sejam mais pertinentes quando se inserem poucas relações semânticas na DT, pois ao redigir “semelhante à x”, ou “está associado à x”, geralmente não é no final de uma longa descrição, mas sim no começo e, na maioria das vezes, numa DT que contemple menos DE_ns (ver Tabela 7).

Seguem as Tabelas 10 e 11, contendo as sequências produtivas que contêm as relações **SEMELHANÇA** e **ASSOCIAÇÃO**.

SEQUÊNCIA (SEMELHANÇA)	FREQ. ABS.	FREQ. REL.
CONSTITUTIVO+SEMELHANÇA	2	40%
TÉLICO+SEMELHANÇA	2	40%
EXEMPLO+SEMELHANÇA	1	20%
TOTAL	5	100%

Tabela 10 - Relação x + SEMELHANÇA

SEQUÊNCIA (ASSOCIAÇÃO)	FREQ. ABS.	FREQ. REL.
CONSTITUTIVO+ASSOCIAÇÃO	17	85%
CONSEQUÊNCIA+ASSOCIAÇÃO	2	10%
ELABORAÇÃO+ASSOCIAÇÃO	1	5%
TOTAL	20	100%

Tabela 11 - Relação x + ASSOCIAÇÃO

E por fim apresentamos os dados referentes à DE5:

DE5	FREQ. ABS.	FREQ. REL.
ELABORAÇÃO	5	56%
CONSTITUTIVO	2	22%
AGENTIVO	1	11%
TÉLICO	1	11%
TOTAL	9	100%

Tabela 12 - Relações semânticas na posição DE5

Fazem parte da DE em posição 5 os papéis *qualia* de Pustejovsky (1995), entretanto, a relação **ELABORAÇÃO** novamente se mostra como a mais produtiva.

No que se refere ao campo semântico no qual o termo a ser definido se insere, agrupamos os campos dos dois domínios que apresentam características semelhantes, a fim de demonstrar: a) em que medida o campo semântico interfere na escolha das relações semânticas das DTs; b) o quanto as relações são produtivas, mesmo em se tratando de domínios de conhecimento distintos.

Defeito (DE) e Disfunção (DI)

REVESTIMENTO CERÂMICO	RELAÇÃO	FREQ. ABS.	FREQ. REL.
DE	CONSTITUTIVO	19	29%
	CIRCUNSTÂNCIA	18	28%
	AGENTIVO	16	25%
	CONSEQUÊNCIA	7	11%
	ELABORAÇÃO	4	6%
	EXEMPLO	1	2%
TOTAL	65	100%	

Tabela 13 - Relações semânticas do campo DEFEITO

FISIOTERAPIA	RELAÇÃO	FREQ. ABS.	FREQ. REL.
DI	CONSTITUTIVO	91	42%
	AGENTIVO	85	40%
	CONSEQUÊNCIA	26	12%
	ELABORAÇÃO	8	4%
	CIRCUNSTÂNCIA	3	1%
	ASSOCIAÇÃO	1	0%
	EXEMPLO	1	0%
	TOTAL	215	100%

Tabela 14 - Relações semânticas do campo DISFUNÇÃO

Esses dois campos, um do domínio da Fisioterapia e outro do Revestimento Cerâmico, se caracterizam pela presença das relações semânticas comuns: **CONSTITUTIVO**, **AGENTIVO** e **CONSEQUÊNCIA**, pois são dois campos semânticos semelhantes e que exigem praticamente os mesmos predicados.

Ensaio (EN), Teste e Medida (TM) e Exame Complementar (EC)

REVESTIMENTO CERÂMICO	RELAÇÃO	FREQ. ABS.	FREQ. REL.
EN	TÉLICO	38	40%
	ELABORAÇÃO	35	37%
	MEIO	19	20%
	CONSTITUTIVO	2	2%
	TOTAL	94	100%

Tabela 15 - Relações semânticas do campo ENSAIO

FISIOTERAPIA	RELAÇÃO	FREQ. ABS.	FREQ. REL.
TM	ELABORAÇÃO	81	64%
	TÉLICO	44	35%
	MEIO	1	1%
	TOTAL	126	100%

Tabela 16 - Relações semânticas do campo TESTE E MEDIDA

FISIOTERAPIA	RELAÇÃO	FREQ. ABS.	FREQ. REL.
EC	ELABORAÇÃO	10	42%
	TÉLICO	9	38%
	CONSTITUTIVO	4	17%
	CIRCUNSTÂNCIA	1	4%
	TOTAL	24	100%

Tabela 17 - Relações semânticas do campo EXAME COMPLEMENTAR

Nesses três campos analisados, as relações que se mostraram profícuas foram **TÉLICO** e **ELABORAÇÃO**, pois se trata de termos que expressam técnicas ou ensaios, cuja finalidade (**TÉLICO**) precisa ser explicitada, bem como a explicação (**ELABORAÇÃO**) de como são realizados.

Instrumento ou Equipamento (IN) e Instrumento de medida (IM)

REVESTIMENTO CERÂMICO	RELAÇÃO	FREQ. ABS.	FREQ. REL.
IN	TÉLICO	58	43%
	CONSTITUTIVO	34	25%
	ELABORAÇÃO	34	25%
	EXEMPLO	7	5%
	CIRCUNSTÂNCIA	1	1%
	TOTAL	134	100%

Tabela 18 - Relações semânticas do campo INSTRUMENTO OU EQUIPAMENTO

FISIOTERAPIA	RELAÇÃO	FREQ. ABS.	FREQ. REL.
IM	TÉLICO	11	55%
	CONSTITUTIVO	6	30%
	ELABORAÇÃO	2	10%
	SEMELHANÇA	1	5%
	TOTAL	20	100%

Tabela 19 - Relações semânticas do campo INSTRUMENTO DE MEDIDA

Os campos semânticos apresentados nas Tabelas 18 e 19 também trazem como mais frequentes as relações **TÉLICO**, **CONSTITUTIVO** e **ELABORAÇÃO**.

Essa descrição demonstra que apesar de serem áreas do conhecimento distintas, a estrutura de organização da DT apresenta certa prototipicidade, comprovando que essa orientação pode ser útil aos usuários do Editor de Definição.

III. **Informações quanto à quantidade de itens léxicos por relação semântica e por DE_n:**

RELAÇÃO	ITENS LÉXICOS	QTDE. DE VEZES QUE A RELAÇÃO APARECE	MÉDIA DE ITENS LÉXICOS POR RELAÇÃO
ELABORAÇÃO	7052	248	28
ASSOCIAÇÃO	465	20	23
TÉLICO	6012	288	21
CONSTITUTIVO	8086	388	21
EXEMPLO	739	47	16
AGENTIVO	1861	121	15
CONSEQUENCIA	530	37	14
MEIO	572	46	12
CIRCUNSTÂNCIA	247	41	6
SEMELHANÇA	27	5	5
MÉDIA TOTAL	2559	124	16

Tabela 20 - Dados quanto à quantidade de itens léxicos por relação semântica

O *corpus* revelou que as relações **ELABORAÇÃO**, **ASSOCIAÇÃO**, **TÉLICO** e **CONSTITUTIVO** apresentam uma média maior de itens léxicos do que as demais relações. Em contrapartida, as relações **MEIO**, **CIRCUNSTÂNCIA** e **SEMELHANÇA** apresentaram ocorrência inferior à média total.

Acreditamos que, no caso, a relação **ELABORAÇÃO** se mostrou com uma média superior às demais porque por meio dessa relação são realizadas descrições que, por sua vez, abarcam geralmente mais que duas orações e até mesmo sentenças inteiras.

Segue na Tabela 21, a relação quantidade de itens léxicos por DE_n.

DE _n	ITENS LÉXICOS	QTDE. DE DE _n s	MÉDIA DE ITENS LÉXICOS POR DE _n
DE1	8743	500	17
DE2	10175	455	22
DE3	5187	221	23
DE4	1191	56	21
DE5	279	9	31
MÉDIA TOTAL	5115	248	23

Tabela 21 - Dados quanto à quantidade de itens léxicos por DE_n

Nota-se que a média de itens léxicos por DE_n é de 23. A DE5 se caracteriza por ser mais extensa que as demais (com média de 31 itens léxicos), e a DE1 como a mais curta (média de 17). Enquanto que a DE2, DE3 e DE4 aproximam-se da média.

IV. Informações quanto às expressões linguísticas e suas respectivas relações semânticas⁵⁸:

Apresentamos, em ordem alfabética, os quadros que contêm as relações semânticas e suas respectivas expressões linguísticas. Destacamos que as expressões descritas estão na forma como ocorreram no *corpus*, ou seja, não estão em sua forma lematizada e apresentam diferenças entre início em maiúsculo e minúsculo.

⁵⁸ Ressalte-se que as expressões linguísticas referentes à relação ELABORAÇÃO não foram anotadas, pois conforme mencionado, o Editor de Definição, num primeiro momento, se dedicará a apresentação de relações mais gerais que possam ser empregadas em distintos campos do conhecimento.

Expressões linguística – AGENTIVO
As causas mais freqüentes são
As principais causas ou fatores de risco são
Causada pela
Causada por
Causado pela
Causado pelo
causado pelos
como proveniente
de origem
Decorre de
decorrente de
desenvolvida pelo
Desenvolvida por
Devido a
Devido à
devido ao
É causado pela
É ocasionada por
É originada pela
É proveniente de
É provocada por
É provocada por
É provocado pela
e são geralmente provocadas
Em consequência
em decorrência
Em decorrência de
Obtém-se por
obtida pelo

Expressões linguística – AGENTIVO
Obtida por
Obtido a partir
Ocorre devido ao
Pode ocorrer devido à
pode ocorrer pela
Pode ocorrer por
pode ser causada por
Pode ser causado por
Pode ser de origem
pode ser desencadeado por
Pode ser obtido por
Pode ser provocada por
Pode ser resultado de
Pode ter origem
podem ser provocadas por
provocada geralmente por
provocada pela
Provocada por
que ocorre devido à
que podem ser originados
que provém do
realizada pelo
realizado pelo
resultante de
São causadas por
São fatores que propiciam
Suas causas mais freqüentes são
Suas causas possíveis são
trazida da

Quadro 12 - Expressões linguísticas: AGENTIVO

Expressões linguística – ASSOCIAÇÃO
depende da
É comprometida devido à
Está diretamente ligada à
está relacionada à
está relacionada ao
está relacionada às
está relacionado com
Os fatores que interferem
Pode estar associada à
Pode estar associado à
pode estar relacionada com
relacionadas à
relacionado à
sofre influência principalmente

Quadro 13 - Expressões linguísticas: ASSOCIAÇÃO

Expressões linguística – CIRCUNSTÂNCIA
aplicado antes e depois de
aplicado no
Efetuada em
empregada na
na placa cerâmica
na superfície
no interior
no revestimento cerâmico
ocorre no momento da
Pode ocorrer
Pode se desenvolver
quando
que atua no
que ocorre após
que ocorre durante
que ocorre em
realizada na
realizado quando
Surge
utilizada quando

Quadro 14 - Expressões linguísticas: CIRCUNSTÂNCIA

Expressões linguística – CONSEQUÊNCIA
afeta significativamente
causando
Como consequência
compromete
comprometendo
leva à
levando a
levando à
levando ao
o que leva à
pode causar
Pode levar à
podendo comprometer
podendo levar à
podendo ocasionar
provoca
Provoca
provocando
que provoca
reduzindo
Tem como consequências
tendo como consequência

Quadro 15 - Expressões linguísticas: CONSEQUÊNCIA

Expressões linguística – CONSTITUTIVO
Apresenta
Apresenta propriedades como
Apresenta-se como
Apresenta-se na
baseada na
Baseia-se fundamentalmente
caracterizada pela
caracterizada pelo
Caracterizada por
caracterizado pela
caracterizado pelo
Caracterizado por
composto de
composto essencialmente por
Consiste em
Consiste na
Consiste num
consiste numa
Constituída basicamente de
Constituída especialmente de

Expressões linguística – CONSTITUTIVO
Constituída essencialmente de
constituída essencialmente pelo
Constituída essencialmente por
Constituídas por
Constituído de
Constituído por
cujas aparências são
cujos constituintes principais são
É baseada na
é composta por
é composto, basicamente
é constituída de
é constituído por
É formado por
em que há
em que ocorre
em que são efetuados
em que se
em que se realiza
Engloba

Quadro 16 - Expressões linguísticas: CONSTITUTIVO

Expressões linguística – CONSTITUTIVO
essencialmente constituído por
feito de
Formada por
Formado por
na qual
na qual as
na qual um
no qual
no qual a
no qual o
no qual ocorre o
no qual os
por meio da qual se observa
Possui forma
que apresenta
Que apresentam características
que consiste
que consiste em
que consiste na
que consiste num
que consiste numa
que descreve
que envolve
que inclui
que indica
que possui
que se apresenta como
que se caracteriza pela
que se caracteriza pelo
que se caracteriza por
que se constitui em
relacionada à
relacionada ao
se baseia
Seu material pode ser
tem formato
Tem o formato de

Quadro 17 - Expressões linguísticas: CONSTITUTIVO (continuação)

Expressões linguística – EXEMPLO
Alguns exemplos são
As formas de aplicação mais conhecidas são
As formas mais conhecidas de aplicação são
como por exemplo
como, por exemplo
Há dois tipos:
mais usados são
O exemplo mais típico é
Os exemplos mais comuns são:
Os principais tipos
Os principais tipos são:
pode ser classificada em
pode ser:
podem ser classificadas em:
podem ser classificados em
podem-se dividir em
podem-se dividir nos seguintes tipos:
são classificadas em
São exemplos
Um caso comum é
Um exemplo comum é
Um exemplo desse tipo
Um exemplo seria

Quadro 18 - Expressões linguísticas: EXEMPLO

Expressões linguística – MEIO
é analisada por meio do
é realizada tanto pelo método
É realizado através do
É realizado em
É realizado por meio da
É realizado visualmente,
na qual é empregada
O equipamento utilizado nesta técnica é
Os aparelhos utilizados
Pode ser realizado pelo
Pode ser realizado por
podendo ser medido
podendo ser realizado por dois métodos principais
por meio da
por meio de
por meio do
realizada por meio de
tendo como parâmetro
Utiliza-se o método

Quadro 19 - Expressões linguísticas: MEIO

Expressões linguística – SEMELHANÇA
Conhecida como
Contrário a
Contrário à
Espécie de
Opõe-se ao

Quadro 20 - Expressões linguísticas: SEMELHANÇA

Expressões linguística – TÉLICO
a fim de
a fim de que
a qual possibilita dar
Além da utilização como
Atua como
atuam como
com a finalidade de
com o fim de
com o objetivo de
com o propósito de
com objetivo de
conferindo à
cuja função é
cuja principal característica é propiciar
Cumprir função
Dentre os objetivos, encontram-se
destinado ao
É empregada
é empregada com o objetivo de
É empregada como
É empregado em
É empregado na
É geralmente empregada como
possibilita
É preferencialmente usado
É responsável por
é utilizada nas
é utilizada no
É utilizada para
é utilizada para
É utilizado em
É utilizado para
Empregado em
empregada na
Empregado na

Expressões linguística – TÉLICO
Empregado nas
Empregado nos
geralmente utilizada para
muito utilizada como
Muito utilizado como
Muito utilizado na
no qual se controla
Normalmente se utiliza
O objetivo é
Os objetivos são
Os principais efeitos
para evitar
Para fins
para obter
pode ser empregado como
pode ser utilizada na
pode ser utilizada para
Pode ser utilizado
possibilitando a
promove a
Propicia
propiciando
próprio para
provoca a
que avalia
que avalia a
que controla
que determina
que fornece
que identifica
que permite
que produz
Que quantifica
que registra
que se dedica a

Quadro 21 - Expressões linguísticas: TÉLICO

Expressões linguística – TÉLICO
que serve para
que tem como finalidade
que tem como funcionalidade
que tem por finalidade
que têm por objetivo
que tem por objetivo
que tem por objetivo atingir
Que tem por objetivo aumentar
Que tem por objetivo avaliar
Que tem por objetivo avaliar
que tem por objetivo detectar
que tem por objetivo determinar
Que tem por objetivo identificar
Que tem por objetivo mensurar
que tem por objetivo mostrar
que tem por objetivo realizar
que tem por objetivo verificar
que visa
realizado para
Responsável por
São utilizadas
serve como
Sua aplicação favorece
Sua finalidade é
Tal recurso objetiva
Tem como finalidade
Tem como objetivo
Têm diversas aplicações
tem por finalidade
Tem por objetivo
Utilizada em
utilizada na
utilizada no
utilizado com a finalidade de
Utilizado em
utilizado junto à
Utilizado na
utilizado na
utilizado no
utilizado para
Utilizado para
Visa

Quadro 22 - Expressões linguísticas: TÉLICO (continuação)

V. Informações quanto aos aspectos morfológicos:

Há projetos terminográficos que não fazem uso de ontologia do domínio que está sendo descrito. O uso desse recurso é essencial na compreensão do campo semântico no qual o termo a ser definido aloca-se, assim como a ontologia auxilia o terminólogo na recuperação do seu gênero próximo ou hiperônimo de modo a encabeçar o texto definitório. Uma alternativa da qual o redator da definição, que não utiliza ontologia, pode valer-se é observar a estrutura morfológica dos termos do domínio em questão, pois as áreas do conhecimento se caracterizam pela presença de determinados morfemas, tanto radicais quanto afixos que podem ser “pistas” a fim de iniciar a DT.

Nos Quadros 23 a 29, são apresentados: i) na primeira coluna, alguns termos com seus respectivos sufixos em destaque (exclusivamente esse tipo de morfema foi analisado) que exemplificam a produtividade de certas estruturas no *corpus* de DT constituído para a pesquisa⁵⁹; ii) na segunda coluna, o nome do campo semântico que abrange o termo e; iii) na terceira, o gênero próximo utilizado no início do texto definitório referente ao respectivo termo⁶⁰.

Primeiramente são exibidos alguns sufixos produtivos no domínio da Fisioterapia.

- OSE

Esse prefixo é empregado pela medicina para formar os substantivos de diversos **processos patológicos e doenças**. Devido ao fato de que a Fisioterapia integra a área de saúde, são bastante recorrentes as formações eruditas (radicais e afixos de origem grega ou latina) em sua terminologia.

⁵⁹ Lembramos que devido à anotação do *corpus* em partes que integram a DT (termo-entrada, gênero próximo, diferença específica, etc.), a detecção dos padrões analisados foram realizados por meio de buscas específicas no concordanciador do *WordSmith Tools*.

⁶⁰ Foram consultados os dicionários Houaiss eletrônico e o Aulete Digital na definição dos morfemas que se seguem.

TERMO	CAMPO SEMÂNTICO	GÊNERO PRÓXIMO
OSTEOPOROSE	DISFUNÇÃO	DISFUNÇÃO
ATEROSCLEROSE		DISFUNÇÃO
NECROSE		DISFUNÇÃO
ESQUIMOSE		DISFUNÇÃO
HERMATROSE		DISFUNÇÃO
ARTROSE		DISFUNÇÃO
ANQUILOSE		DISFUNÇÃO
PSEUDOARTROSE		DISFUNÇÃO

Quadro 23 – Termos com sufixo -OSE

- ITE

No contexto da área médica, o sufixo – **ITE** significa **inflamação**.

TERMO	CAMPO SEMÂNTICO	GÊNERO PRÓXIMO
PERICARDITE	DISFUNÇÃO	DISFUNÇÃO
TROMBOFLEBITE		DISFUNÇÃO
ARTRITE		DISFUNÇÃO
VASCULITE		DISFUNÇÃO
ESPONDILITE		DISFUNÇÃO
BRONQUITE		DISFUNÇÃO
FLEBITE		DISFUNÇÃO
DERMATITE		DISFUNÇÃO
ENCEFALITE		DISFUNÇÃO
ESPONDILITE		DISFUNÇÃO
OSTEOARTRITE		DISFUNÇÃO
ARTRITE		DISFUNÇÃO
TENOSSINOVITE		DISFUNÇÃO
MIOSITE		DISFUNÇÃO
BURSITE		DISFUNÇÃO
TENDINITE		DISFUNÇÃO
SINOVITE		DISFUNÇÃO

Quadro 24 - Termos com sufixo -ITE

- METRO

Elemento de composição, de origem grega, que apresenta o sentido de **medida**, **que mede**, **medidor**. É amplamente empregado nas terminologias de caráter científico.

TERMO	CAMPO SEMÂNTICO	GÊNERO PRÓXIMO
ERGÔMETRO	INSTRUMENTO DE MEDIDA	INSTRUMENTO DE MEDIDA
DINAMÔMETRO		INSTRUMENTO DE MEDIDA
ESTESIÔMETRO		INSTRUMENTO DE MEDIDA
ESFIGNOMANÔMETRO		INSTRUMENTO DE MEDIDA
GONIÔMETRO		INSTRUMENTO DE MEDIDA
OXÍMETRO		INSTRUMENTO DE MEDIDA

Quadro 25 - Termos com sufixo -METRO

- GRAFIA

Pospositivo⁶¹, de origem grega, que significa **escrita, descrição, registro** geralmente em substantivos relativos a ciências e artes e técnicas antigas ou modernas

TERMO	CAMPO SEMÂNTICO	GÊNERO PRÓXIMO
RADIOGRAFIA	EXAME COMPLEMENTAR	EXAME
ULTRASSONOGRAFIA		EXAME
ANGIOGRAFIA		EXAME
MIELOGRAFIA		EXAME
ELETRONEUROMIOGRAFIA		EXAME

Quadro 26 - Termos com sufixo -GRAFIA

Nas tabelas 27, 28 e 29, são apresentados alguns sufixos produtivos no domínio do Revestimento Cerâmico.

- AGEM

Sufixo formador de substantivos a partir de base nominal ou verbal, com as noções de **ação ou resultado da ação**.

⁶¹ Elemento que é colocado em posição posterior. [Cf. sufixo.] (Aulete Digital).

TERMO	CAMPO SEMÂNTICO	GÊNERO PRÓXIMO
GRETAGEM	ETAPA DO PROCESSO	PROCESSO
BRITAGEM		PROCESSO
COLAGEM		PROCESSO
PESAGEM		PROCESSO
MOAGEM		PROCESSO

Quadro 27 - Termos com sufixo -AGEM

- MENTO

Sufixo que caracteriza **ação** ou **resultado de ação**; processo.

TERMO	CAMPO SEMÂNTICO	GÊNERO PRÓXIMO
ACABAMENTO	ETAPA DO PROCESSO	PROCESSO
ENVELHECIMENTO		PROCESSO
POLIMENTO		PROCESSO
PENEIRAMENTO		PROCESSO
ESQUADRAMENTO		PROCESSO
ARMAZENAMENTO		PROCESSO
RESFRIAMENTO		PROCESSO

Quadro 28 - Termos com sufixo -MENTO

- ÇÃO

Sufixo que indica **ato**, **ação**, **resultado** ou **efeito** de x.

TERMO	CAMPO SEMÂNTICO	GÊNERO PRÓXIMO
PULVERIZAÇÃO	ETAPA DO PROCESSO	PROCESSO
ATOMIZAÇÃO		PROCESSO
RETIFICAÇÃO		PROCESSO
DEFERRIZAÇÃO		PROCESSO
COMINUIÇÃO		PROCESSO
COMPACTAÇÃO		PROCESSO
CONFORMAÇÃO		PROCESSO
SINTERIZAÇÃO		PROCESSO
GRANULAÇÃO		PROCESSO

Quadro 29 - Termos com sufixo -ÇÃO

- IDADE

Sufixo que integra especialmente substantivos abstratos. Significa **qualidade, caráter** ou **característica**.

TERMO	CAMPO SEMÂNTICO	GÊNERO PRÓXIMO
TONALIDADE	PROPRIEDADE DA PLACA CERÂMICA	CARACTERÍSTICA
COMPACIDADE		PROPRIEDADE
LIMPABILIDADE		PROPRIEDADE
EXPANSÃO POR UMIDADE		PROPRIEDADE
PSEUDOPLASTICIDADE		PROPRIEDADE
ORTOGONALIDADE		PROPRIEDADE
PLASTICIDADE		PROPRIEDADE
PLANARIDADE		PROPRIEDADE
TRABALHABILIDADE		PROPRIEDADE

Quadro 30 - Termos com sufixo -IDADE

Estamos de acordo com Correia & Almeida (2008, p. 32) quanto à seguinte afirmação: “O conhecimento morfológico não serve apenas como um conhecimento mais ou menos abstrato de características da terminologia em análise, mas tem também aplicações práticas sobre várias etapas do trabalho terminológico” (tradução nossa)⁶².

Desta forma, destacamos que essa pequena descrição se propôs a demonstrar que os morfemas podem contribuir com a escolha do gênero próximo (ou hiperônimo) da DT e que esse conhecimento linguístico pode ser aproveitado como mais um recurso a ser implementado no Editor de Definição.

No capítulo 6, será apresentada a sistematização proposta acerca do conhecimento gerado pela pesquisa a fim de ser utilizada no e-termos.

⁶² *Le savoir morphologique ne sert donc pas seulement à une connaissance plus ou moins abstraite des caractéristiques de la terminologie sous analyse, mais il a aussi d'applications pratiques sur plusieurs tâches du travail terminologique.* (CORREIA & ALMEIDA, 2008, p. 32)

6 Sistematização do conhecimento obtido na pesquisa

Inicialmente, a sistematização do conhecimento adquirido na pesquisa irá contribuir para que o Editor de Definição seja utilizado como um ambiente de auxílio no processo da redação da DT. Além das considerações acerca da DT apresentadas nesse trabalho, faz-se importante manifestar alguns critérios que devem guiar o projeto da implementação dos recursos no Editor. São eles:

- I. **perfil de usuário** a quem se destina o Editor de Definição: o sistema e-terms⁶³ foi criado, considerando 3 perfis de usuários, a saber:
 - i. **gerente de projeto**: criador do projeto e é o principal responsável pelo controle das etapas e pela equipe de trabalho. Executa qualquer tarefa/etapa do e-terms (menos o especialista).
 - ii. **terminólogo/linguista**: responsável pelas atividades terminológicas e linguísticas do e-terms. É o principal perfil em potencial de usuário do Editor de Definição.
 - iii. **especialista do domínio**: responsável pelas tarefas de validação das etapas do ambiente (candidatos a termos, ontologia e a própria DT).

Nesse sentido, considera-se que esses três perfis terão acesso ao Editor de Definição e que eles podem, ainda, quanto **ao estágio da escrita**, ter um conhecimento nulo, básico, intermediário ou avançado da redação da DT.

- II. **uso e tarefa** que será atribuído ao Editor de Definição: o editor se destinará a auxiliar os usuários na redação de um **tipo de texto** específico, a DT, que segue padrões quanto à sua estrutura. O editor poderá favorecer o ensino da redação da DT àqueles que são iniciantes na tarefa.

⁶³ O **sistema e-terms** se diferencia do **portal e-terms**, o qual tem acesso público e irrestrito, destina-se ao público geral e onde se alocam os produtos terminológicos gerados a partir do sistema e-terms. Enquanto que o sistema é caracterizado pelo acesso autorizado, usuários cadastrados, criação dos produtos e local das ferramentas e etapas.

- III. A orientação fornecida pelo Editor diz respeito ao planejamento da estrutura para a redação da DT (*top down*).
- IV. a utilização do Editor de Definição independe da classe morfológica do termo que será definido, contudo, nosso trabalho se restringiu a termos classificados como **substantivo**, justamente por ser a categoria mais recorrente em grande parte das terminologias, e que fosse definido de acordo com o padrão de texto definitório **GPDE**, reconhecido como o modelo ideal de definição. Dessa forma, o desenvolvimento das análises assentou-se nesses dois princípios: **Substantivos** definidos a partir do **padrão GPDE**;
- V. a partir da DE, pudemos fragmentá-la em várias DE_n s de modo a observar qual **número mínimo e máximo de DE_n s** que normalmente constituem a DT, assim como examinar quais **relações semânticas estão implicadas em cada uma**, as **expressões linguísticas** que as explicitam e quanto à produtividade da **ordem em que as relações aparecem** no texto definitório;
- VI. e por fim, quanto à ocorrência de certos morfemas que sinalizam qual(is) gênero(s) próximo(s) podem ser empregados na DT.

A sistematização compreende “sugestões” que serão agregadas ao Editor e compete ao usuário a decisão de usá-las.

EDITOR DE DEFINIÇÃO

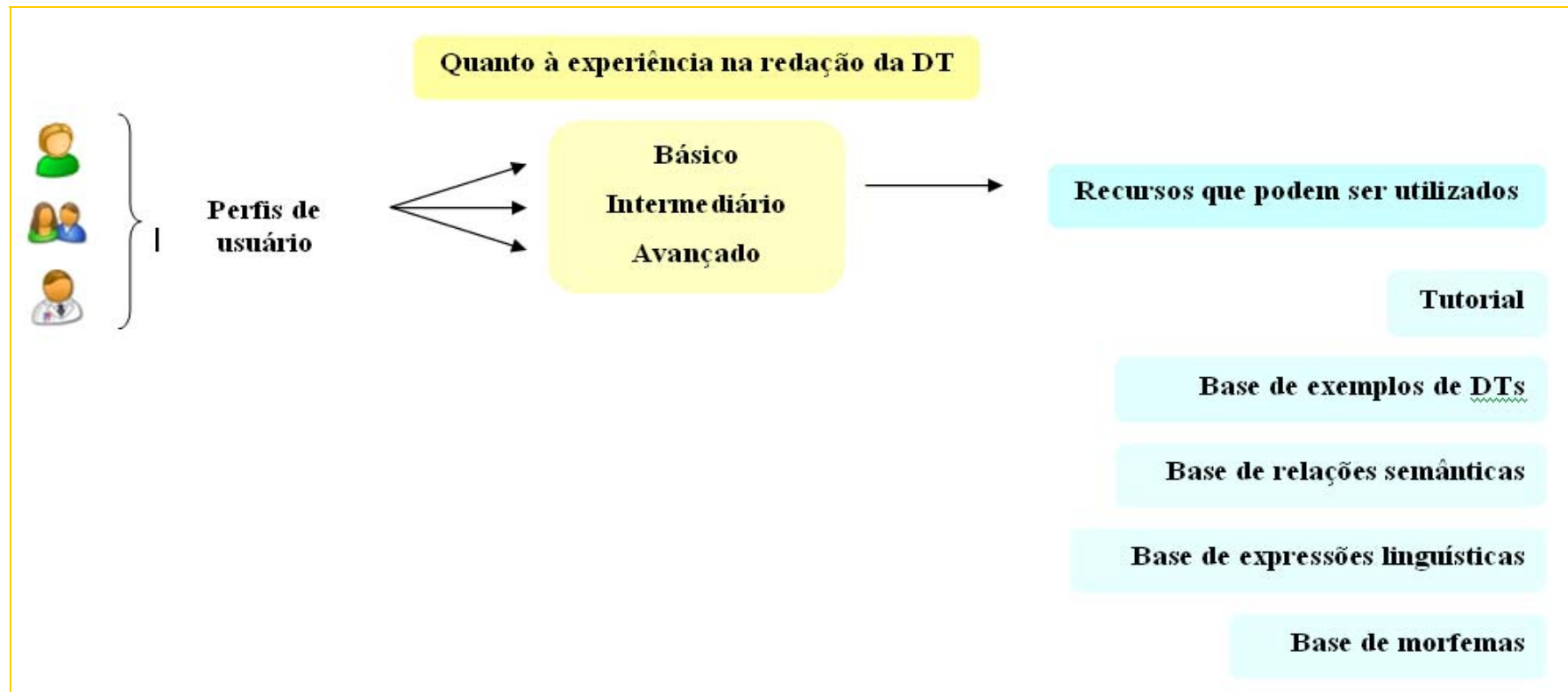


Figura 18 - Sistematização proposta para o Editor de Definição do e-Termos

Tutorial: manual sobre princípios e convenções da DT elaborado especialmente para os que estão iniciando na redação do texto definitório. A partir desse tutorial, cabe ao gerente do projeto eleger quais desses princípios e convenções serão adotados pela equipe, levando-se em conta o público-alvo e a finalidade da obra terminográfica.

Base de exemplos de DTs: base que contemplará DTs de substantivos (anotadas conforme TE, GP, DEs, DE_ns), que servirão de exemplos para o usuário compor seu próprio texto definitório. Lembramos aqui que as DTs que compõem essa base foram redigidas segundo a tipologia GPDE.

Base de relações semânticas e suas expressões linguísticas: base que reunirá as relações semânticas propostas nesse trabalho, definição ou explicação e suas expressões linguísticas correspondentes. Serão ainda demonstradas quais as combinações e a ordem de relações mais produtivas na DT. Destaca-se que os redatores da DT podem inserir nova relação e expressão linguística, caso sintam necessidade, a fim de que os membros da equipe possam se beneficiar.

Base de morfemas: base na qual serão apresentados os morfemas analisados (no item 5 da dissertação) e seus significados. Nessa base, os redatores da DT poderão editá-la de modo a inserir informações sobre os morfemas empregados na terminologia do projeto em pauta.

No capítulo 7 serão apresentadas as considerações finais da pesquisa.

7 Considerações finais

A trajetória percorrida para a realização deste trabalho apresentou como eixo central a Definição Terminológica. Para tecer o fio condutor do nosso objeto de estudo, foram apresentadas: i) a fundamentação teórica acerca do percurso da Terminologia até os dias de hoje; ii) a TCT, perspectiva na qual compreendemos as linguagens de especialidades; iii) a DT e os elementos que a constituem (características e relações semânticas). No decorrer da pesquisa, observamos que era necessário recorrer a outros autores, de linhas teóricas distintas, que tivessem tratado de alguma forma de relações semânticas, pois esse era o olhar que deveria ser dado ao nosso *corpus* de pesquisa. Algumas considerações sobre as relações semânticas foram feitas nessa etapa da investigação:

- I. a princípio, as relações semânticas deveriam ser compreendidas por humanos; logo, elas devem ser claras e objetivas;
- II. o conjunto de relações não necessitava ser muito extenso, a ponto de haver grandes divergências acerca de qual relação deve ser contemplada na anotação. Por outro lado, um rol que abrangesse as relações semânticas mais tipificadas em qualquer domínio do conhecimento já era o suficiente para o nosso propósito de pesquisa.

Ao final do levantamento das relações utilizadas por cada pesquisador, pudemos propor o nosso rol de relações semânticas de modo a empregá-las nas análises das DTs. Em seguida, foi imprescindível considerar uma metodologia de pesquisa para a manipulação do *corpus*. Embora o objetivo central da pesquisa fosse sistematizar o conhecimento linguístico, sobretudo quanto à porção semântica do texto definitório, houve uma certa dedicação, que demandou mais tempo que o previsto, quanto a essa etapa da pesquisa, de modo a cumprir dois papéis importantes:

- I. empregar uma metodologia que pudesse “cercar” a DT, de modo que fosse possível extrair várias informações dela de modo rápido e fiável;
- II. asseverar que os passos metodológicos seguidos pudessem ser replicados em outras análises de cunho linguístico-textual.

Chegamos, então, a uma sequência de etapas, as quais compreenderam: composição e armazenamento do *corpus*; nomeação dos arquivos e anotação do *corpus* (Editor de Cabeçalho) e utilização de programas computacionais nas análises (*WordSmith Tools*, Anotador e extrator semiautomático de etiqueta “xml”). Ressalta-se que todas as decisões tomadas em relação à preparação do *corpus* tiveram que ser conforme com:

- I. os programas que seriam utilizados posteriormente na análise;
- II. os dados que considerássemos relevantes constar da pesquisa.

Quanto à etapa de análise, foi possível recuperar as informações linguísticas e estatísticas de um modo otimizado, porém, quais e quantos desses dados deveriam, de fato, ser apresentados no trabalho de dissertação? Percebemos, então, que a discussão acerca das informações linguísticas é tão ou mais importante que a descrição do *corpus*. Os recursos informáticos estão à disposição dos linguistas, de modo que um número muito grande de dados possa ser manipulado simultaneamente. Contudo, a capacidade de transformar aquelas tantas informações depende muito mais do analista do que qualquer outro fator. Sem dúvida, essa é uma das lições aprendidas neste trabalho de mestrado.

A sistematização das informações adquiridas nesta pesquisa, que se tornarão posteriormente recursos disponibilizados aos usuários do Editor de Definição do ambiente computacional e-terms, está estritamente relacionada com os dados que se mostraram como mais produtivos nas análises (quanto à quantidade de DE_s por DT; relações semânticas mais recorrentes e suas expressões linguísticas, a ordem de apresentação das DE_s e a observação do aspecto morfológico da terminologia em pauta).

Como projetos futuros, além de auxiliarmos o informata na implementação do Editor de Definição, pretendemos:

- I. ampliar a descrição morfológica, de maneira a enriquecer a base de morfemas e melhor atender ao usuário;
- II. estudar outros tipos definitórios sugeridos por Sager (1993), já citados no item 2.3, tais como: definição mediante análise, definição mediante sinônimos, definição mediante paráfrase, definição mediante síntese;
- III. observar quais aspectos linguísticos estão implicados quando se vai definir as seguintes classes de palavras também existentes em terminologias: verbo e adjetivo.

No que diz respeito aos aspectos terminológicos, acreditamos ter oferecido uma contribuição, na medida em que sistematizamos o tipo definitório mais produtivo, tornando a tarefa do terminólogo menos morosa e mais controlada.

Em relação aos aspectos semânticos e linguísticos, pudemos apresentar as relações semânticas mais frequentes nas terminologias e as expressões linguísticas que as explicitam no português do Brasil.

E finalmente, acreditamos que nossa pesquisa poderá incrementar o e-Termos, tornando o ambiente ainda mais completo para a tarefa de desenvolvimento e gestão de terminologias.

A propósito deste último item, cumpre ressaltar aqui a necessária e enriquecedora interdisciplinaridade entre PLN e Terminologia, uma vez que o conhecimento linguístico é imprescindível para melhorar o desempenho de ferramentas e programas computacionais, mas, em contrapartida, todo esse aparato computacional funciona como lentes para melhor enxergar e descrever a língua. Nesse sentido, a Terminologia não só se beneficia da Computação, mas dela necessita.

Referências

ALMEIDA, G.M.B. A Teoria Comunicativa da Terminologia e a sua prática. **Alfa**, Araraquara, v. 50, p. 81-97, 2006.

Disponível em: <<http://www.alfa.ibilce.unesp.br/download/v50-2/06-Almeida.pdf>>. Acesso em: fev/2008.

ALMEIDA, G.M.B.; OLIVEIRA, L.H.M.; ALUÍSIO, S.M. A Terminologia na era da Informática. **Ciência e Cultura**, Campinas, v. 58, n. 2, p. 42-45, 2006.

Disponível em: <http://cienciaecultura.bvs.br/scielo.php?script=sci_arttext&pid=S0009-67252006000200016&lng=pt&nrm=iso>. Acesso em: fev/2008.

ALMEIDA, G.M.B.; SOUZA, D.S.L.; PINO, D.H.P. A definição nos dicionários especializados: proposta metodológica. **Debate Terminológico**, v. 3, p. 1-20, 2007.

Disponível em: <http://www.riterm.net/revista/n_3/index.htm>. Acesso em: fev/2008.

ALPÍZAR CASTILLO, R. **Cómo hacer un diccionario científico técnico?** Buenos Aires: Editorial Memphis, 1997. 184p.

ALUÍSIO, S.M. **Ferramentas para auxiliar a escrita de artigos científicos em inglês como língua estrangeira.** 1995. 216 f. Tese (Doutorado em Física) – Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 1995.

ALUÍSIO, S.M.; ALMEIDA, G.M.B. O que é e como se constrói um *corpus*? Lições aprendidas na compilação de vários corpora para pesquisa linguística. **Calidoscópico**, Porto Alegre, v. 4, n. 3, p. 155-177, 2006.

Disponível em:

<http://www.unisinos.br/publicacoes_cientificas/images/stories/pdfs_calidoscopio/vol4n3/art_04_Aluísio.pdf>. Acesso em: fev/2008.

ARNZT, R.; PICHT, H. **Introducción a la Terminología.** Madrid: Fundación Germán Sánchez Ruipérez, 1995. 382p.

AUBERT, F. H. Introdução à Metodologia da Pesquisa Terminológica Bilíngüe - **Cadernos de Terminologia**, 2ª.ed. São Paulo: Humanitas Publicações - FFLCH/ USP, 2001. 104p.

AUGER, A. **Repérage des énonces d'intérêt définitoire dans les bases de données textuelles.** 1997. 224f. Tese (Doutorado em Linguística) - Faculté des Lettres, Université de Neuchâtel, Genebra, 1997.

Disponível em: <http://doc.rero.ch/record/473/files/these_AugerA.pdf> Acesso em: fev/2009

BARROS, L. A. **Curso básico de Terminologia**. São Paulo: EdUSP, 2004. 285p.

BIDERMAN, M. T. C. **Teoria Lingüística**. Teoria Lexical e Lingüística Computacional. 2^a ed. São Paulo: Martins Fontes, 2001. 356p.

BODSON, C. **Termes et relations sémantiques en corpus spécialisés : rapport entre patrons de relations sémantiques (PRS) et types sémantiques (TS)**. 2005. 298f. Tese (Doutorado em Linguística) - Faculté des études supérieures, Université de Montréal, Montréal, 2005.

Disponível em: <<http://olst.ling.umontreal.ca/pdf/bodson2005.pdf>>. Acesso em: fev/2009.

BOULANGER, J.C. Une lecture socio-culturelle de la terminologie. **Cahiers de linguistique sociale**. Rouen, p. 13-30, 1991.

BOULANGER, J.C. Présentation: images et parcours de la socioterminologie. **Meta**, Montréal, v. 40, n. 2, p. 194-205, 1995.

Disponível em: <<http://www.erudit.org/revue/meta/1995/v40/n2/002117ar.pdf>>. Acesso em fev/2009.

BOUTIN-QUESNEL et al. Vocabulaire systématique de la terminologie. **Cahiers de l'Office de la Langue Française**. Québec: Publications du Québec, 1985.

CABRÉ, M. T. **La terminología: teoría, metodología, aplicaciones**. Tradução de Carles Tebé. Barcelona: Editorial Antártida/Empúries, 1993. 529p.

CABRÉ, M. T. **La terminología: representación y comunicación. elementos para una teoría de base comunicativa y otros artículos**. Barcelona: IULA/Universitat Pompeu Fabra, 1999. 369p.

CABRÉ, M. T. Theories of terminology: their description, prescription and explanation. **Terminology**, v. 9, n. 2, p. 163-200, 2003.

CIAPUSCIO, G. E. **Textos especializados y Terminología**. Barcelona: IULA/ Universitat Pompeu Fabra, 2003. 115p. (Sèrie Monografies, 6).

CONDAMINES, A. Sémantique et corpus spécialisés: constitution de bases de connaissances terminologiques, **Carnets de grammaire**, Relatório interno da ERSS (Équipe de Recherche en Syntaxe et Sémantique), n. 13. Toulouse: CNRS et Université de Toulouse-le Mirail, 2003.

CORREIA, M.; ALMEIDA, G. M. B. Le rôle du savoir morphologique dans l'organisation de l'arbre conceptuel d'un domaine spécialité. In: **Terminologie: discours, technologie et acteurs sociaux**, 2008, Brest - Bretagne - France. Actes de GLAT-BREST 2008. Brest - França: Groupe de Linguistique Appliquée des Télécommunications - GLAT, 2008. v. 1. p. 23-34.

DAHLBERG, I. Terminological definitions: characteristics and demands. In: Association internationale de terminologie *et al.*, in: **Problemes de la definition et de la synonymie en terminologie: actes du Colloque international de terminologie**, Université Laval, Quebec, 23-27 mai 1982, Girsterm, Quebec, 1983, p. 15-34.

DAVIDSON, L. et al. Semi-automatic extraction of knowledge-rich contexts from corpora. In: D. Bourigault, C. Jacquemin et M. C. L'Homme (eds.), in: **Proceedings of Computerm'98: First Workshop on Computational Terminology**. Montréal, Université de Montréal, 1998.

DEYES, T. Discourse, Science and Scientific Discourse. **The Raw Material of Comprehension in ESP**. Working Paper 6, Brazilian ESP Project, PUC/SP, São Paulo, 1982

DROUIN, P. **Acquisition automatique des termes: l'utilisation des pivots lexicaux spécialisés**. 2002. 307f. Tese (Doutorado em Linguística) - Faculté des études supérieures, Université de Montréal, Montréal, 2002.

DUBUC, R. **Manual práctico de terminología**. Tradução de Ileana Cabrera. Chile: RiL Editores, 1999. 236p.

FELIU, J. **Relacions conceptuais i variació funcional: elements per a un sistema de detecció automàtica**. (Trabalho de pesquisa) - UPF/IULA, Barcelona, 2000.

FELTRIM, V. D. **Uma abordagem baseada em córpus e em sistemas de crítica para a construção de ambientes Web de auxílio à escrita acadêmica em português**. 2004. 169f. Tese (Doutorado em Computação) – Instituto de Ciência Matemática e Computação, Universidade de São Paulo, São Carlos, 2004.

FERREIRA, A. B. H. **Novo Dicionário Eletrônico Aurélio**. 3ª edição. Curitiba: Editora Positivo, 2004.

FINATTO, M. J. B. **Definição terminológica: fundamentos teórico-metodológicos para sua descrição e explicação.** 2001. 395f. Tese (Doutorado em Letras) - Instituto de Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2001.

FONTANA, N. et al. Computer Assisted Writing - Applications to English as a Foreign Language. **CALL**, v. 6, n. 2, p. 145-161, 1993.

GAMBIER, Y. Problèmes terminologiques des pluies acides: pour une socio-terminologie. **Meta**, Montréal, v.32, n. 3, p. 314-320, 1987.

Disponível em : < <http://www.erudit.org/revue/meta/1987/v32/n3/002791ar.pdf> >

Acesso em: fev/2009

GAUDIN, F. **Socioterminologie: des problèmes semantiques aux pratiques institutionnelles.** Rouen: Université de Rouen, 1993. 255p. (Publications de l'Université de Rouen, 182).

GAUDIN, F. **Socioterminologie: une approche sociolinguistique de la terminologie.** Bruxelles: Duculot De Boeck, 2003. 286p.

GOUADEC, D. **Terminologie: constitution des données.** Paris: AFNOR, 1990. 219p.

HOFFMAN, L. **Els llenguatges d'especialitat: selecció de textos.** Barcelona: IULA, 1998. 284p. (Sèrie Monografies)

HÖFLING, C. **Da análise crítica de definições de nomes concretos em dicionários para uma proposta de definição padrão.** 2000. 376f. Dissertação (Mestrado em Linguística) - Faculdade de Ciências e Letras de Araraquara, UNESP, Araraquara, 2000.

ISO, Principes et methodes de la terminologie: **ISO 704**, Organisation Internationale de Normalisation, Genebra, 1987.

JACKENDOFF, R. **Semantics and Cognition.** Cambridge: MIT Press, 1983. 283p.

JORDAN, M.P. An Integrated Three-Pronged Analysis of a Fund-Raising Letter. In: W.C. Mann and S.A. Thompson (eds.), **Discourse Description: Diverse Linguistic Analyses of a Fund-Raising Text**, 1992. p. 171-226.

LARIVIÈRE, L. Comment formuler une définition terminologique. **Meta**, Montréal, v. 41, n. 3, 405-418, 1996.

Disponível em: < <http://www.erudit.org/revue/meta/1996/v41/n3/003401ar.pdf>> Acesso em: fev/2009.

MARSHMAN, E. **The Cause Relation in Biopharmaceutical Corpora: English and French Patterns for Knowledge Extraction**. 2002. 166f. Dissertação (Mestrado em Linguística) - University of Ottawa, Ottawa, 2002.

Disponível em: <<http://www.ruor.uottawa.ca/fr/handle/10393/6385>>. Acesso em: fev/2009.

MARSHMAN, E.; MORGAN, T.; MEYER, I. French patterns for expressing concept relations. **Terminology**, 2002, v. 8, n. 1, p. 1-29. 2002.

OLF. Le grand dictionnaire terminologique. Disponível em: <http://www.granddictionnaire.com/btml/fra/r_motclef/index800_1.asp>. Acesso em: março/2009.

OLIVEIRA, L. H. M. **e-Termos: um ambiente colaborativo web para a criação de produtos terminológicos**. 2006. Monografia de qualificação (Doutorado em Computação) – Instituto de Ciência Matemática e Computação. Universidade de São Paulo, São Carlos, 2006.

OTMAN, G. Terminologie et intelligence artificielle, in **La Banque des mots**, n° special, 1989, p. 63-95.

PUSTEJOVSKY, J. **The generative lexicon**. Londres: Cambridge/MIT Press, 1998. 312p.

REY, A. **La terminologie: noms et notions**. Paris: Presses universitaires de France, 1979. 127p.

REY-DEBOVE, J. **Étude linguistique et sémiotique des dictionnaires français contemporains**. The Hague/Paris: Mouton, 1971. 329p.

SAGER, J. C. **Curso práctico sobre el procesamiento de la terminología**. Tradução de Laura C. Moya. Madrid: Fundación Germán Sánchez Ruipérez/Pirámide, 1993. 442p.

SEPPÄLÄ, S. **Composition et formalisation conceptuelles de la définition terminographique**. 2004. 200f. Tese (Doutorado em Tratamento Informático Multilíngue), École de traduction et d'interprétation, Université de Genève, Genebra, 2004.

SIERRA MARTÍNEZ, G.; ALARCÓN, R. El rol de las predicaciones verbales en la extracción automática de conceptos. **Estudios de Lingüística Aplicada**, n. 38, 129-144. Distrito Federal, México, 2003.

SIMMEN, F. **La représentation des systemes conceptuels en terminologie**, Université de Genève, École de traduction et d'interpretation, Genève, 2002.

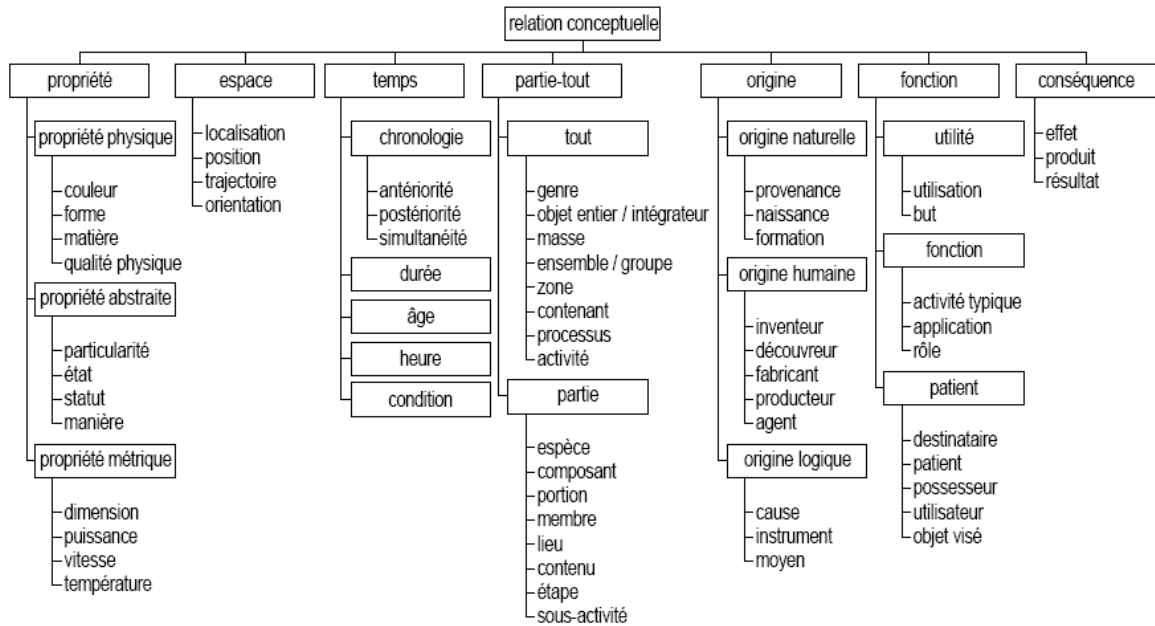
TEMMERMANN, R. **Towards new ways of terminology description. The sociocognitive approach**. Philadelphia: John Benjamins, 2000. 258p.

WÜSTER, E. **Introducción a la teoría general de la terminología y a la lexicografía terminológica**. Tradução de Anne-Cécile Nokerman. Barcelona: IULA/Universitat Pompeu Fabra, 1998. 227p.

ZAVAGLIA, C. **Análise da homonímia no português: tratamento semântico com vistas a procedimentos computacionais**. 2002. 360f., v.I; 199f.; v.II Tese (Doutorado em Linguística) - Faculdade de Ciências e Letras de Araraquara, UNESP, Araraquara, 2002.

Anexo I

Tipologia das relações conceituais possíveis (extraída de Seppälä, 2004)



Anexo II

Protocolo da relação de “causalidade” proposta por Feliu (2004)

SÍMBOL: **CAUS.**

DESCRIPCIÓ: /Relació de causalitat: relació que s'estableix entre una causa i el seu efecte/

PROPIETATS: *No simètrica i no necessàriament transitiva*

OBSERVACIONS: /Presenta subtipus: causal i procedural/

ARGUMENTS: 2 o > 2 segons els subtipus

RECURSIVITAT: sí

CARACTERÍSTIQUES a: *entitat, activitat, propietat*

CARACTERÍSTIQUES b: *entitat, activitat, propietat*

/Observacions (a/b): *no necessàriament idèntiques*

Expressió de la relació: *dependre de; fer que; ser la causa (principal) de; deure's a; implicar; aparèixer; contribuir a; dependre de; emmascarar; donar lloc a; <cond.> trobar.; reforçar; provocar; augmentar; produir; transformar-se en; obtenir-se (+ ger.); donar lloc a; permetre (+inf.).*

Anexo III

Tipologia das relações semânticas propostas por Jordan (1992)

Tipo de Relação	Relações
Detail	Identification
	Classification
	Specification
	Appearance
	Characteristics
	Function
	Material
	Parts
General	Active
	Passive
	Agent
	Source
Logical	Assessment
	Basis
	Cause
	Effect
	Emotive Effect
	Purpose
	Means
	Problem
	Solution
Modal	Possibility
	Capability
	Correctness
	Propriety
	Necessity
	Need
	Completion
	Achievement
	Future
	Intention
	Mandate
	Authority

Tipo de Relação	Relações
Time	Before
	After
	Simultaneous
	Inverted time
	Time
Text manipulation	Elaboration
	Summary
	Repetition
	Paraphrase
	Forecast
	Transition
Special	Collateral inversion
	Concession
	Compatibility
	Contrast
	Comparison
	Conditionals
	Document structures
	Hypothetical-Real
	Transition couplets
Other	Accompaniment
	Circumstance
	Inverted circumstance
	Connection
	Enablement
	Example
	Extent
	Location
	Inverted Location
	Manner
True	

Apêndice A – Relações Semânticas I e II

A seguir, são apresentadas as relações semânticas utilizadas na anotação do *corpus* e suas principais características.

RELAÇÃO	ETIQUETA	EXEMPLOS	DEFINIÇÃO	EXEMPLO	EXPRESSÕES LINGUÍSTICAS
SEMANTICA I	Constitutivo	Propriedades	Relação semântica básica do termo-entrada que apresenta do que ele é feito ou constituído.	caracterizado por fissuras (...) que se estende da superfície externa até a interface do esmalte com a placa cerâmica.	possui forma; formado por; que se caracteriza por...
		Partes			
		Material			
		Aparência			
	Agentivo	Causa	Relação semântica básica do termo-entrada que apresenta sua causa ou origem.	É causado pela falta de acordo entre as expansões térmicas do vidrado e suporte (...)	ocorre devido ao; pode ter origem; causado por...
		Origem/Agente			
	Télico	Utilização	Relação semântica básica do termo-entrada que apresenta sua função ou objetivo.	que determina a resistência ao desgaste das placas cerâmicas não-esmaltadas,	empregado como; a fim de que; cuja função é...
		Função			
		Objetivo			

RELAÇÃO	ETIQUETA	EXEMPLOS	DEFINIÇÃO	EXEMPLO	EXPRESSÕES LINGUÍSTICAS
SEMÂNTICA II	Consequência	Efeito	Relação semântica do termo-entrada que apresenta sua consequência.	tendo como consequência a perda da impermeabilidade e (...)	provoca; o que leva à; comprometendo...
		Produto			
		Resultado			
	Meio	Instrumento	Relação semântica do termo-entrada que apresenta o meio ou instrumento com que ele é realizado.	É realizado por meio da avaliação dimensional (...)	por meio de; utiliza-se o método...
		Método			
	Circunstância	Espaço	Relação semântica do termo-entrada que apresenta o espaço ou tempo no qual ele se situa.	que ocorre após o assentamento.	quando; que ocorre durante...
		Tempo			
	Semelhança	Co-hiponímia	Relação semântica do termo-entrada que apresenta associação positiva ou negativa dele com outro termo da ontologia de igual campo semântico, ou com qualquer outro termo/palavra devido à semelhança.	Opõe-se ao estado sólido ou cristalino.	opõe-se ao; contrário ao...
	Exemplo	Hiponímia	Relação semântica do termo-entrada que apresenta seu termo hipônimo, classificação ou exemplo.	Segundo (...), podem ser classificadas em: (...)	um exemplo comum é; os principais tipos são...
		Classificação			
Elaboração	Descrição	Relação semântica do termo-entrada que apresenta elaboração dele, ou seja, dados ou detalhes sobre ele que não se encaixam nas outras relações, os quais servem para deixar a DT mais rica, quase enciclopédica.	Encontra-se também na maior parte das variedades de argila, em forma de quartzo ou outra forma cristalina ou amorfa de sílica.		
	Indicação				
	Recomendação				
	Deteção				
Associação		Relação semântica do termo-entrada que apresenta relação de dependência com algum termo ou conceito de dentro ou fora da ontologia.	A pseudoplasticidade está relacionada às características físicas das partículas que compõem o fluido (...)	está diretamente ligado à; está relacionado com...	

Apêndice B – Análise das DTs por Campo Semântico

O presente Apêndice é constituído das análises realizadas com grupos de DTs de mesmo campo semântico. A partir dessa descrição, foi possível observar algumas características pertinentes aos distintos grupos.

Como ler os dados?

Seguem, na Figura 19, explicações acerca de como devem ser lidos os dados, sendo que, a título de ilustração, usamos o campo semântico Ação Muscular (AM) do domínio da Fisioterapia, devido à sua extensão. Salientamos ainda que esses dados foram distribuídos em gráficos e tabelas distintas a fim de que pudéssemos salientar as informações de modo mais eficaz, assim como torná-las mais claras aos leitores, contudo esses dados estão relacionados, sendo que é possível compará-los de modo que a leitura se torne mais produtiva.

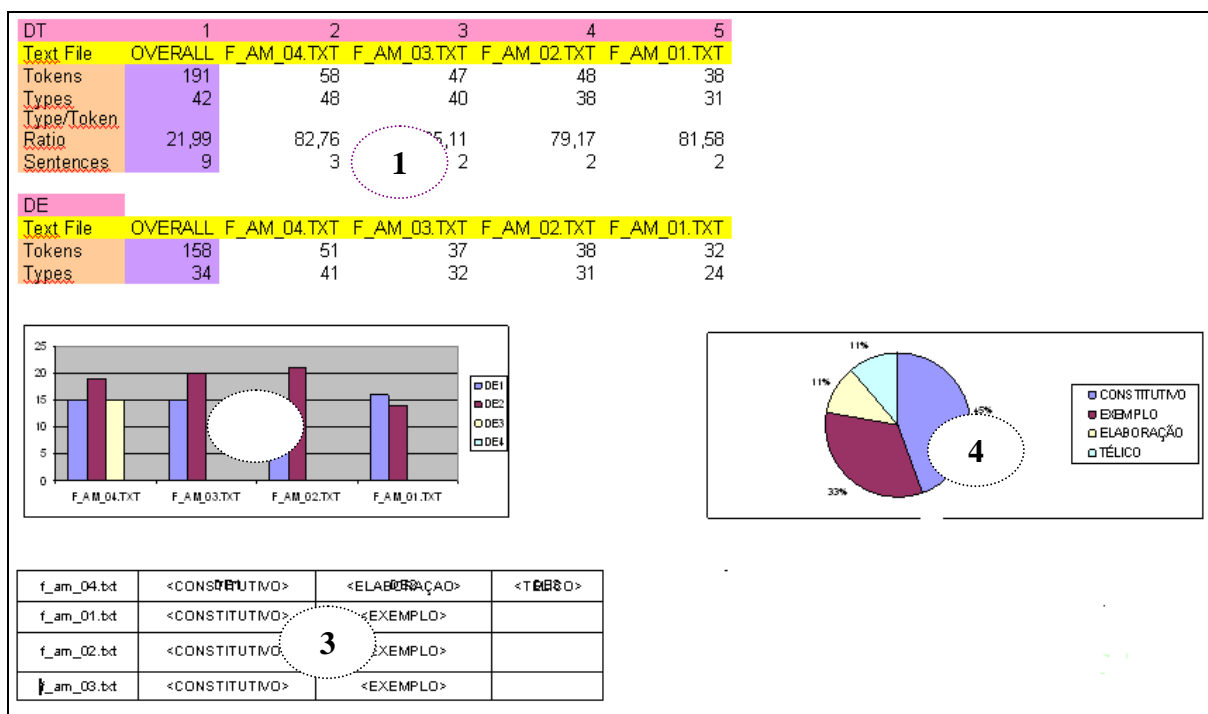


Figura 19 - Como ler os dados?

Os dados da Figura 19 representam:

Dados 1: na primeira tabela, estão elencados alguns dados estatísticos sobre a DT (GP+DE) e a DE referente a todas as DTs que compreendem o campo semântico selecionado, além de serem exibidos os dados no total, de todos os arquivos pertencentes ao campo (em *overall*). Os dados são:

- a) *Text file*: nome do arquivo;
- b) *Tokens*: quantidade de unidades lexicais presentes no arquivo;
- c) *Types*: quantidade de unidades lexicais diferentes presentes no arquivo;
- d) *Type/Token Ratio*: divisão de *types* por *tokens*⁶⁴;
- e) *Sentences*: quantidade de sentenças.

Ressaltamos que esses dados estatísticos puderam nos revelar, se existe relação entre quantidade de *tokens* e de relações presentes na DT, tamanho médio de DTs e de DEs.

Dados 2: no primeiro gráfico, estão elencados a quantidade de DEs (DE1, DE2, DE_n) por cada DT do campo semântico selecionado e a quantidade de *tokens* relativa a cada DE.

Dados 3: na segunda tabela, está elencada a ordem na qual as relações semânticas aparecem na DT. Na primeira coluna consta o nome do arquivo, e da segunda coluna em diante, é exibida, abaixo de cada DE correspondente, a relação representada.

Dados 4: no segundo gráfico, são exibidas as relações e suas porcentagens, de modo que possamos visualizar as relações mais significativas do campo semântico selecionado.

Esses dados puderam ser recuperados devido à anotação do *corpus* e aos recursos estatísticos e de concordância do *WordSmith Tools*.

Primeiramente, serão apresentados os dados referentes aos campos semânticos do domínio da Fisioterapia; e em seguida, dos campos do domínio do Revestimento Cerâmico.

⁶⁴ Conforme Biderman (2001): “Na língua inglesa os estatísticos do léxico costumam opor o *token* (ocorrência no texto) ao *type* (lexema referido pela ocorrência formal).”

campo AÇÃO MUSCULAR (AM)

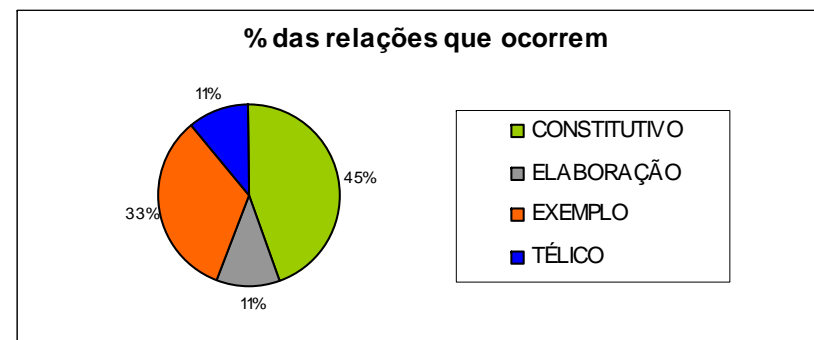
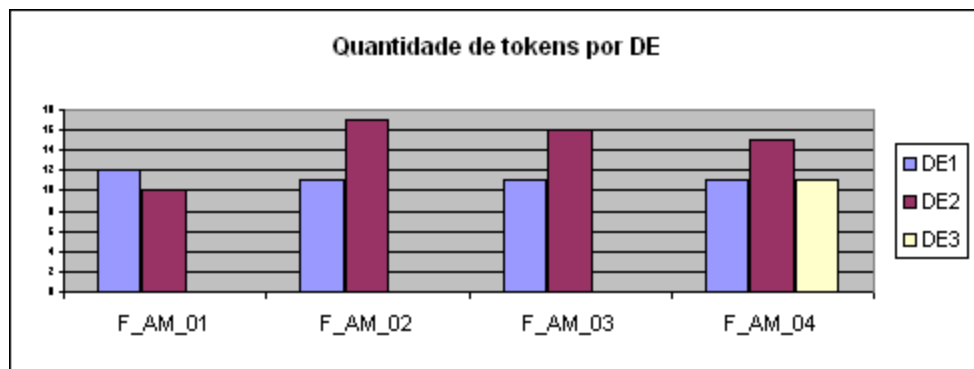
DT

Text File	OVERALL	F_AM_01	F_AM_02	F_AM_03	F_AM_04
Tokens	131	24	34	33	40
Types	32	21	29	30	36
Type/Token Ratio	24,43	87,5	85,29	90,91	90
Sentences	9	2	2	2	3

DE

Text File	OVERALL	F_AM_01	F_AM_02	F_AM_03	F_AM_04
Tokens	114	22	28	27	37
Types	29	19	26	26	34

	DE1	DE2	DE3
F_AM_01	<constitutivo>	<exemplo>	
F_AM_02	<constitutivo>	<exemplo>	
F_AM_03	<constitutivo>	<exemplo>	
F_AM_04	<constitutivo>	<elaboração>	<télico>



campo ARTICULAÇÃO (AR)

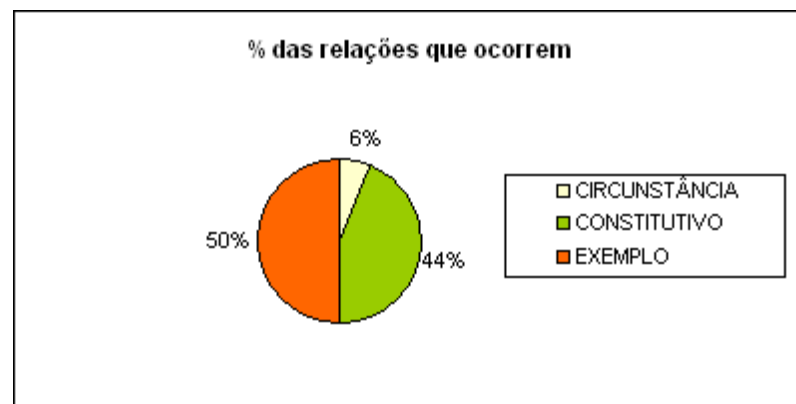
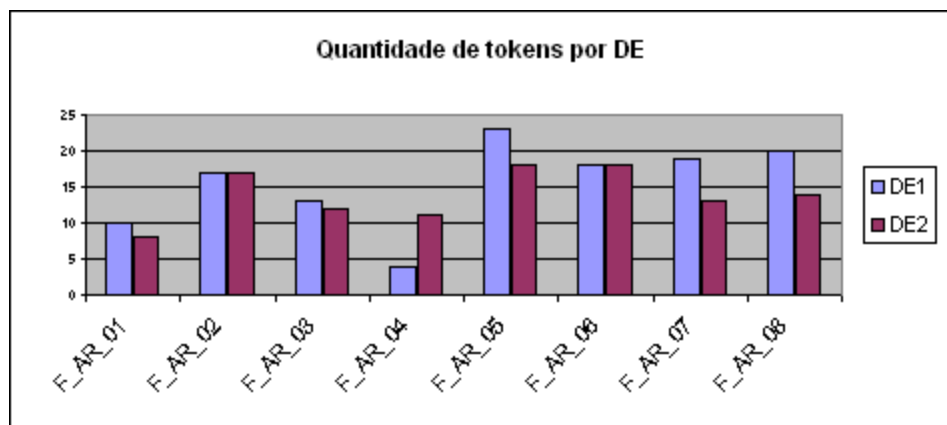
DT

Text File	OVERALL	F_AR_01	F_AR_02	F_AR_03	F_AR_04	F_AR_05	F_AR_06	F_AR_07	F_AR_08
Tokens	255	24	36	26	18	43	38	34	36
Types	42	22	35	25	17	39	31	29	31
Type/Token Ratio	16,47	91,67	97,22	96,15	94,44	90,7	81,58	85,29	86,11
Sentences	16	2	2	2	2	2	2	2	2

DE

Text File	OVERALL	F_AR_01	F_AR_02	F_AR_03	F_AR_04	F_AR_05	F_AR_06	F_AR_07	F_AR_08
Tokens	235	18	34	25	15	41	36	32	34
Types	39	17	33	24	14	38	30	28	30

	DE1	DE2
F_AR_01	<constitutivo>	<exemplo>
F_AR_02	<constitutivo>	<exemplo>
F_AR_03	<constitutivo>	<exemplo>
F_AR_04	<circunstância>	<exemplo>
F_AR_05	<constitutivo>	<exemplo>
F_AR_06	<constitutivo>	<exemplo>
F_AR_07	<constitutivo>	<exemplo>
F_AR_08	<constitutivo>	<exemplo>



campo DISFUNÇÃO (DI)

DT

Text File	OVERALL	F_DI_01	F_DI_02	F_DI_03	F_DI_04	F_DI_05	F_DI_06	F_DI_07	F_DI_08	F_DI_09	F_DI_10	F_DI_11	F_DI_12	F_DI_13	F_DI_14	F_DI_15	F_DI_16	F_DI_17	F_DI_18
Tokens	3.442	19	60	36	67	35	39	37	52	46	39	41	48	33	52	22	30	26	34
Types	380	16	51	27	48	32	34	30	45	39	37	40	41	32	47	20	25	23	31
Type/Token Ratio	11,04	84,21	85	75	71,64	91,43	87,18	81,08	86,54	84,78	94,87	97,56	85,42	96,97	90,38	90,91	83,33	88,46	91,18
Sentences	126	1	2	1	2	1	1	2	2	2	1	2	2	2	2	1	1	1	2

DE

Text File	OVERALL	F_DI_01	F_DI_02	F_DI_03	F_DI_04	F_DI_05	F_DI_06	F_DI_07	F_DI_08	F_DI_09	F_DI_10	F_DI_11	F_DI_12	F_DI_13	F_DI_14	F_DI_15	F_DI_16	F_DI_17	F_DI_18
Tokens	3.223	17	58	34	65	33	37	35	50	44	37	39	45	31	50	20	28	23	32
Types	374	14	49	25	46	30	32	28	43	37	35	38	38	30	45	18	23	20	29

DT

Text File	F_DI_19	F_DI_20	F_DI_21	F_DI_22	F_DI_23	F_DI_24	F_DI_25	F_DI_26	F_DI_27	F_DI_28	F_DI_29	F_DI_30	F_DI_31	F_DI_32	F_DI_33	F_DI_34	F_DI_35	F_DI_36	F_DI_37
Tokens	37	19	64	54	64	51	30	35	47	20	63	31	26	33	53	39	20	35	13
Types	32	18	55	44	55	43	26	29	41	18	52	24	25	30	44	34	20	31	12
Type/Token Ratio	86,49	94,74	85,94	81,48	85,94	84,31	86,67	82,86	87,23	90	82,54	77,42	96,15	90,91	83,02	87,18	100	88,57	92,31
Sentences	2	1	2	2	2	2	1	1	2	0	2	1	2	2	2	1	1	0	1

DE

Text File	F_DI_19	F_DI_20	F_DI_21	F_DI_22	F_DI_23	F_DI_24	F_DI_25	F_DI_26	F_DI_27	F_DI_28	F_DI_29	F_DI_30	F_DI_31	F_DI_32	F_DI_33	F_DI_34	F_DI_35	F_DI_36	F_DI_37
Tokens	35	17	62	52	62	49	28	33	44	18	61	29	24	31	51	37	18	33	11
Types	30	16	53	42	54	41	24	27	38	16	50	22	23	28	42	32	18	29	10

DT

Text File	F_DI_38	F_DI_39	F_DI_40	F_DI_41	F_DI_42	F_DI_43	F_DI_44	F_DI_45	F_DI_46	F_DI_47	F_DI_48	F_DI_49	F_DI_50	F_DI_51	F_DI_52	F_DI_53	F_DI_54	F_DI_55	F_DI_56
Tokens	34	32	34	33	49	43	30	21	18	31	32	50	29	45	29	45	23	21	14
Types	32	29	30	27	40	40	28	21	17	30	29	40	25	42	26	38	20	21	13
Type/Token Ratio	94,12	90,63	88,24	81,82	81,63	93,02	93,33	100	94,44	96,77	90,63	80	86,21	93,33	89,66	84,44	86,96	100	92,86
Sentences	2	1	2	1	1	1	1	1	1	2	2	2	1	1	1	1	1	1	1

DE

Text File	F_DI_38	F_DI_39	F_DI_40	F_DI_41	F_DI_42	F_DI_43	F_DI_44	F_DI_45	F_DI_46	F_DI_47	F_DI_48	F_DI_49	F_DI_50	F_DI_51	F_DI_52	F_DI_53	F_DI_54	F_DI_55	F_DI_56
Tokens	32	30	32	31	47	41	28	19	16	29	31	48	27	43	27	43	20	18	12
Types	30	27	28	25	38	38	26	19	15	28	28	38	23	40	24	36	17	18	11

DT

Text File	F_DI_57	F_DI_58	F_DI_59	F_DI_60	F_DI_61	F_DI_62	F_DI_63	F_DI_64	F_DI_65	F_DI_66	F_DI_67	F_DI_68	F_DI_69	F_DI_70	F_DI_71	F_DI_72	F_DI_73	F_DI_74	F_DI_75
Tokens	21	24	24	49	48	23	23	20	26	30	31	44	36	44	34	50	36	44	42
Types	21	22	20	41	41	21	21	19	23	30	26	39	28	36	30	42	31	38	38
Type/Token Ratio	100	91,67	83,33	83,67	85,42	91,3	91,3	95	88,46	100	83,87	88,64	77,78	81,82	88,24	84	86,11	86,36	90,48
Sentences	1	1	1	1	1	1	1	1	1	2	1	1	2	2	1	2	1	1	1

DE

Text File	F_DI_57	F_DI_58	F_DI_59	F_DI_60	F_DI_61	F_DI_62	F_DI_63	F_DI_64	F_DI_65	F_DI_66	F_DI_67	F_DI_68	F_DI_69	F_DI_70	F_DI_71	F_DI_72	F_DI_73	F_DI_74	F_DI_75
Tokens	18	21	21	46	45	20	20	17	23	27	28	41	33	41	31	47	33	41	39
Types	18	19	17	38	38	18	18	16	20	27	23	36	25	33	27	40	28	35	35

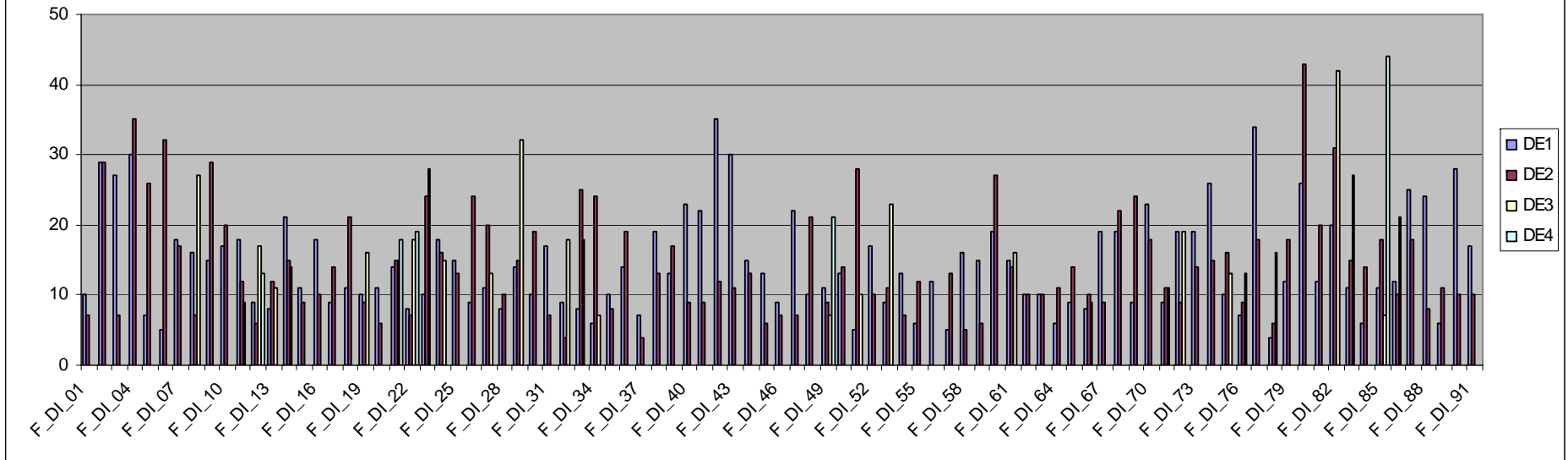
DT

Text File	F_DI_76	F_DI_77	F_DI_78	F_DI_79	F_DI_80	F_DI_81	F_DI_82	F_DI_83	F_DI_84	F_DI_85	F_DI_86	F_DI_87	F_DI_88	F_DI_89	F_DI_90	F_DI_91
Tokens	32	55	29	33	72	35	96	56	23	83	44	47	35	20	41	29
Types	27	46	27	30	58	27	73	42	18	66	38	40	28	19	36	26
Type/Token Ratio	84,38	83,64	93,1	90,91	80,56	77,14	76,04	75	78,26	79,52	86,36	85,11	80	95	87,8	89,66
Sentences	1	2	1	2	2	1	2	1	1	2	2	2	2	1	1	1

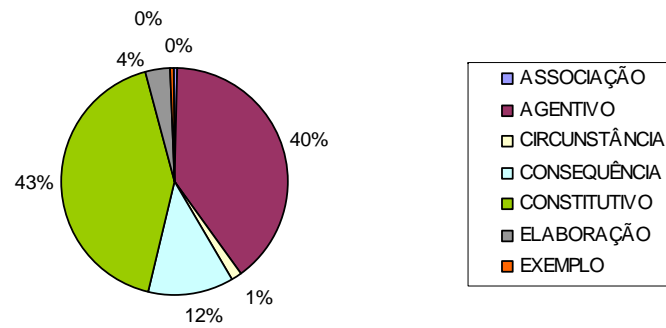
DE

Text File	F_DI_76	F_DI_77	F_DI_78	F_DI_79	F_DI_80	F_DI_81	F_DI_82	F_DI_83	F_DI_84	F_DI_85	F_DI_86	F_DI_87	F_DI_88	F_DI_89	F_DI_90	F_DI_91
Tokens	29	52	26	30	69	32	93	53	20	80	43	43	32	17	38	27
Types	24	43	24	27	55	24	71	40	15	63	37	38	26	16	33	24

Quantidade de tokens por DE



% das relações que ocorrem



	DE1	DE2	DE3	DE4
F_DI_01	<constitutivo>	<agentivo>		
F_DI_02	<constitutivo>	<agentivo>		
F_DI_03	<constitutivo>	<agentivo>		
F_DI_04	<constitutivo>	<agentivo>		
F_DI_05	<constitutivo>	<agentivo>		
F_DI_06	<constitutivo>	<agentivo>		
F_DI_07	<constitutivo>	<agentivo>		
F_DI_08	<constitutivo>	<agentivo>	<agentivo>	
F_DI_09	<constitutivo>	<agentivo>		
F_DI_10	<constitutivo>	<agentivo>		
F_DI_11	<constitutivo>	<elaboração>	<circunstância>	
F_DI_12	<constitutivo>	<agentivo>	<elaboração>	<consequência>
F_DI_13	<constitutivo>	<exemplo>	<elaboração>	
F_DI_14	<constitutivo>	<circunstância>	<elaboração>	
F_DI_15	<constitutivo>	<agentivo>		
F_DI_16	<constitutivo>	<consequência>		
F_DI_17	<constitutivo>	<consequência>		
F_DI_18	<constitutivo>	<consequência>		
F_DI_19	<constitutivo>	<agentivo>	<circunstância>	
F_DI_20	<constitutivo>	<agentivo>		
F_DI_21	<constitutivo>	<agentivo>	<agentivo>	<elaboração>
F_DI_22	<constitutivo>	<agentivo>	<consequência>	<agentivo>
F_DI_23	<constitutivo>	<agentivo>	<agentivo>	
F_DI_24	<constitutivo>	<consequência>	<agentivo>	
F_DI_25	<constitutivo>	<agentivo>		
F_DI_26	<constitutivo>	<agentivo>		

	DE1	DE2	DE3	DE4
F_DI_27	<constitutivo>	<consequência>	<agentivo>	
F_DI_28	<constitutivo>	<agentivo>		
F_DI_29	<constitutivo>	<agentivo>	<consequência>	
F_DI_30	<constitutivo>	<agentivo>		
F_DI_31	<constitutivo>	<agentivo>		
F_DI_32	<constitutivo>	<consequência>	<agentivo>	
F_DI_33	<constitutivo>	<agentivo>	<elaboração>	
F_DI_34	<constitutivo>	<agentivo>	<consequência>	
F_DI_35	<constitutivo>	<agentivo>		
F_DI_36	<constitutivo>	<agentivo>		
F_DI_37	<constitutivo>	<agentivo>		
F_DI_38	<constitutivo>	<agentivo>		
F_DI_39	<constitutivo>	<agentivo>		
F_DI_40	<constitutivo>	<agentivo>		
F_DI_41	<constitutivo>	<agentivo>		
F_DI_42	<constitutivo>	<agentivo>		
F_DI_43	<constitutivo>	<agentivo>		
F_DI_44	<constitutivo>	<agentivo>		
F_DI_45	<constitutivo>	<agentivo>		
F_DI_46	<constitutivo>	<agentivo>		
F_DI_47	<constitutivo>	<associação>		
F_DI_48	<constitutivo>	<agentivo>		
F_DI_49	<constitutivo>	<agentivo>	<consequência>	<agentivo>
F_DI_50	<constitutivo>	<consequência>		
F_DI_51	<constitutivo>	<agentivo>	<consequência>	
F_DI_52	<constitutivo>	<agentivo>		

	DE1	DE2	DE3	DE4
F_DI_53	<constitutivo>	<agentivo>	<elaboração>	
F_DI_54	<constitutivo>	<agentivo>		
F_DI_55	<constitutivo>	<agentivo>		
F_DI_56	<constitutivo>			
F_DI_57	<constitutivo>	<agentivo>		
F_DI_58	<constitutivo>	<agentivo>		
F_DI_59	<constitutivo>	<consequência>		
F_DI_60	<constitutivo>	<consequência>		
F_DI_61	<constitutivo>	<agentivo>	<consequência>	
F_DI_62	<constitutivo>	<agentivo>		
F_DI_63	<constitutivo>	<agentivo>		
F_DI_64	<constitutivo>	<agentivo>		
F_DI_65	<constitutivo>	<consequência>		
F_DI_66	<constitutivo>	<agentivo>	<agentivo>	
F_DI_67	<constitutivo>	<agentivo>		
F_DI_68	<constitutivo>	<agentivo>		
F_DI_69	<constitutivo>	<agentivo>		
F_DI_70	<constitutivo>	<agentivo>		
F_DI_71	<constitutivo>	<agentivo>	<agentivo>	
F_DI_72	<constitutivo>	<consequência>	<agentivo>	
F_DI_73	<constitutivo>	<consequência>		

	DE1	DE2	DE3	DE4
F_DI_74	<constitutivo>	<agentivo>		
F_DI_75	<constitutivo>	<agentivo>	<consequência>	
F_DI_76	<constitutivo>	<agentivo>	<consequência>	
F_DI_77	<constitutivo>	<agentivo>		
F_DI_78	<constitutivo>	<consequência>	<agentivo>	
F_DI_79	<constitutivo>	<agentivo>		
F_DI_80	<constitutivo>	<agentivo>		
F_DI_81	<constitutivo>	<agentivo>		
F_DI_82	<constitutivo>	<agentivo>	<agentivo>	
F_DI_83	<constitutivo>	<consequência>	<agentivo>	
F_DI_84	<constitutivo>	<agentivo>		
F_DI_85	<constitutivo>	<agentivo>	<consequência>	<elaboração>
F_DI_86	<constitutivo>	<consequência>	<agentivo>	
F_DI_87	<constitutivo>	<agentivo>		
F_DI_88	<constitutivo>	<agentivo>		
F_DI_89	<constitutivo>	<agentivo>		
F_DI_90	<constitutivo>	<consequência>		
F_DI_91	<constitutivo>	<agentivo>		

campo EXAME COMPLEMENTAR (EC)

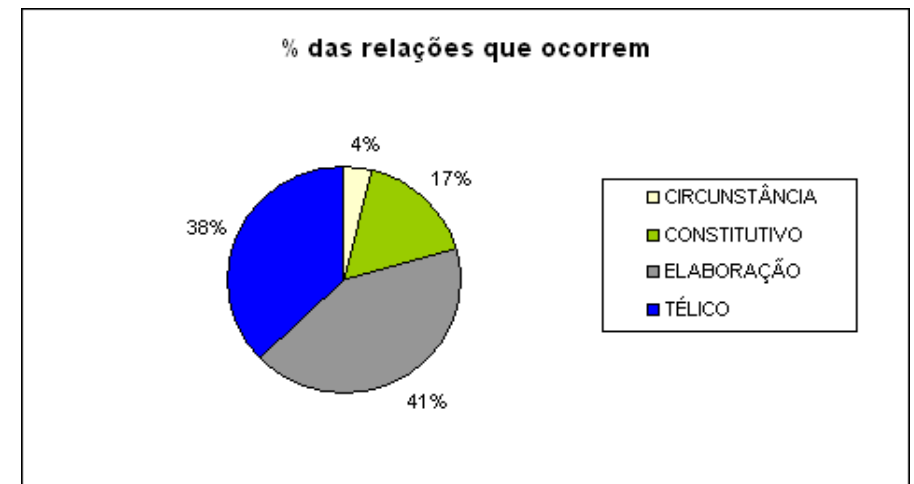
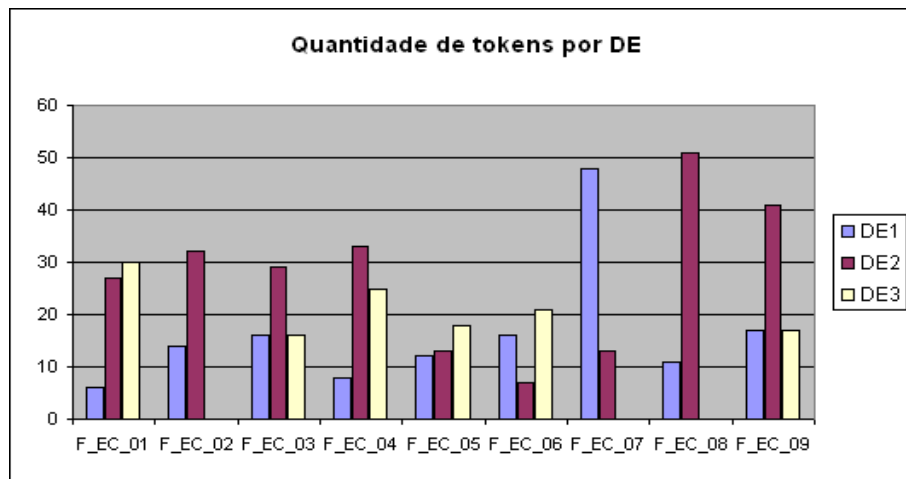
DT

Text File	OVERALL	F_EC_01	F_EC_02	F_EC_03	F_EC_04	F_EC_05	F_EC_06	F_EC_07	F_EC_08	F_EC_09
Tokens	563	64	53	65	70	47	53	66	66	79
Types	74	52	42	49	56	38	44	55	53	59
Type/Token Ratio	13,14	81,25	79,25	75,38	80	80,85	83,02	83,33	80,3	74,68
Sentences	26	2	1	2	4	3	3	5	3	3

DE

Text File	OVERALL	F_EC_01	F_EC_02	F_EC_03	F_EC_04	F_EC_05	F_EC_06	F_EC_07	F_EC_08	F_EC_09
Tokens	521	63	46	61	66	43	44	61	62	75
Types	69	52	38	46	54	34	40	51	50	56

	DE1	DE2	DE3
F_EC_01	<circunstância>	<constitutivo>	<elaboração>
F_EC_02	<constitutivo>	<télico>	
F_EC_03	<constitutivo>	<télico>	<elaboração>
F_EC_04	<télico>	<elaboração>	<télico>
F_EC_05	<télico>	<constitutivo>	<elaboração>
F_EC_06	<elaboração>	<télico>	<elaboração>
F_EC_07	<elaboração>	<télico>	
F_EC_08	<télico>	<elaboração>	
F_EC_09	<télico>	<elaboração>	<elaboração>



campo INSTRUMENTOS E MEDIDAS (IM)

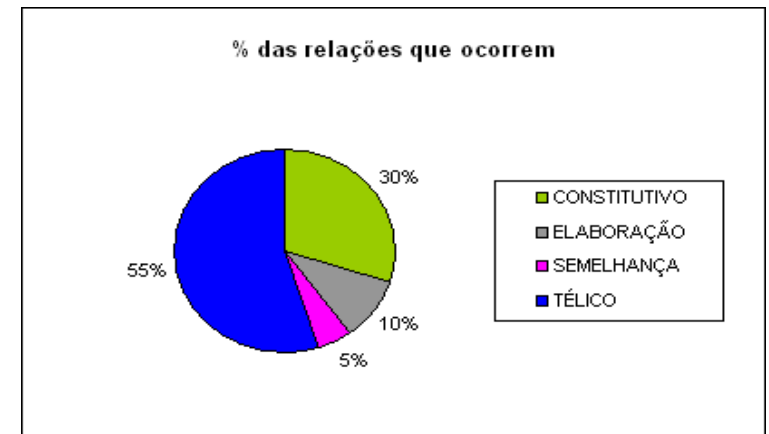
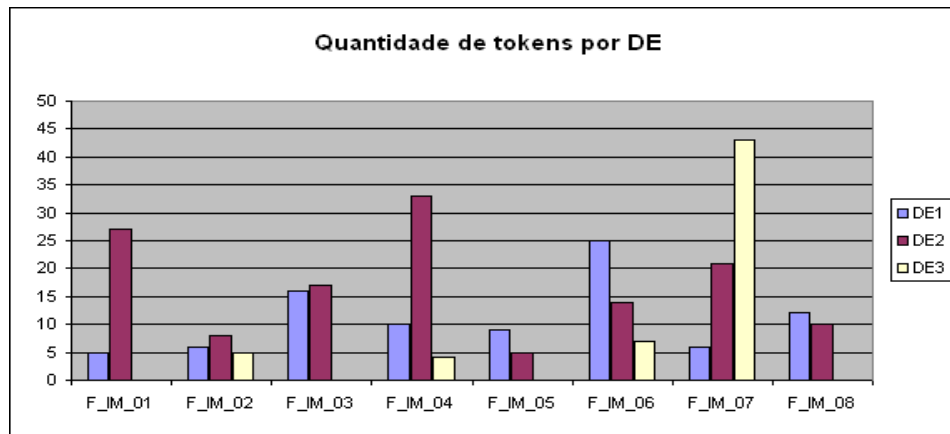
DT

Text File	OVERALL	F_IM_01	F_IM_02	F_IM_03	F_IM_04	F_IM_05	F_IM_06	F_IM_07	F_IM_08
Tokens	307	35	22	36	50	17	49	73	25
Types	42	29	20	31	38	17	41	50	21
Type/Token Ratio	13,68	82,86	90,91	86,11	76	100	83,67	68,49	84
Sentences	22	3	3	2	3	2	3	4	2

DE

Text File	OVERALL	F_IM_01	F_IM_02	F_IM_03	F_IM_04	F_IM_05	F_IM_06	F_IM_07	F_IM_08
Tokens	283	32	19	33	47	14	46	70	22
Types	40	27	18	29	36	14	39	48	19

	DE1	DE2	DE3
F_IM_01	<télico>	<constitutivo>	
F_IM_02	<télico>	<semelhança>	<télico>
F_IM_03	<constitutivo>	<elaboração>	
F_IM_04	<télico>	<constitutivo>	<télico>
F_IM_05	<télico>	<télico>	
F_IM_06	<télico>	<constitutivo>	<télico>
F_IM_07	<télico>	<constitutivo>	<elaboração>
F_IM_08	<constitutivo>	<télico>	



campo MOVIMENTO ARTICULAR (MA)

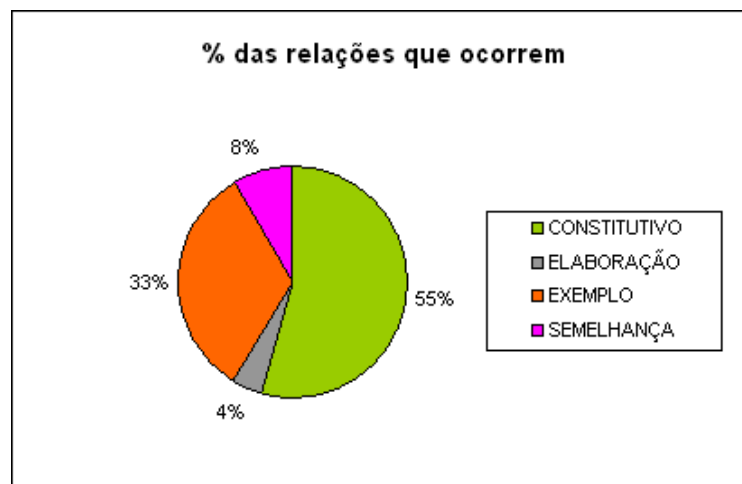
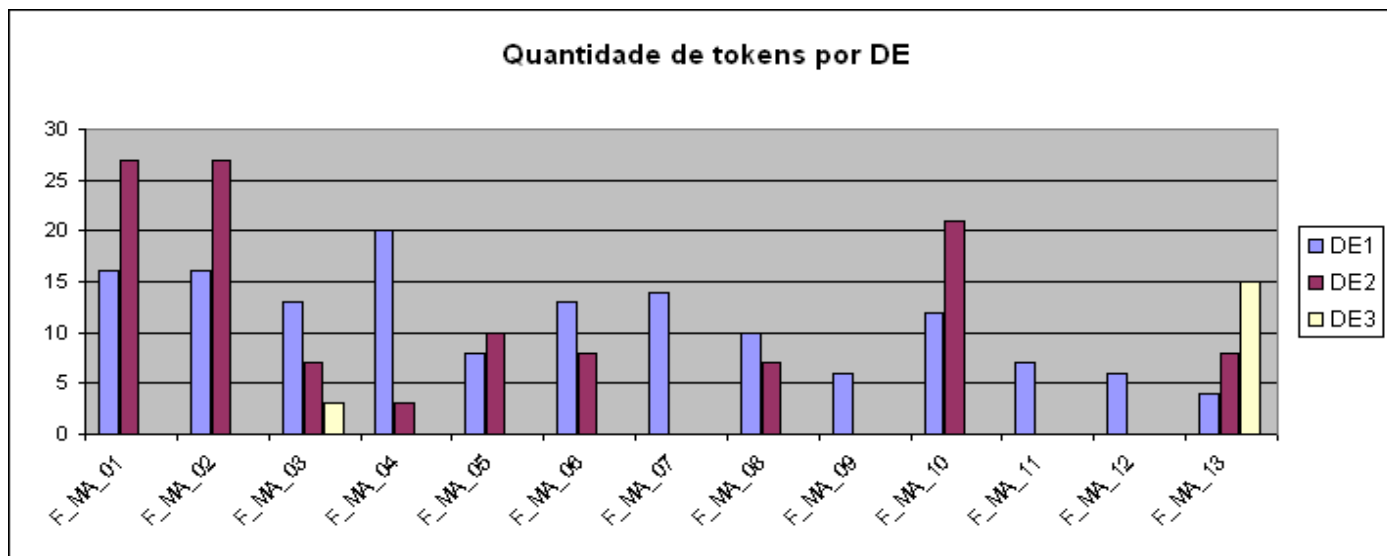
DT

Text File	OVERALL	F_MA_01	F_MA_02	F_MA_03	F_MA_04	F_MA_05	F_MA_06	F_MA_07	F_MA_08	F_MA_09	F_MA_10	F_MA_11	F_MA_12	F_MA_13
Tokens	294	44	44	24	24	19	22	15	18	7	34	8	7	28
Types	63	32	33	19	20	18	20	14	18	7	27	8	7	25
Type/Token Ratio	21,43	72,73	75	79,17	83,33	94,74	90,91	93,33	100	100	79,41	100	100	89,29
Sentences	20	2	2	2	2	1	2	1	2	1	1	1	1	2

DE

Text File	OVERALL	F_MA_01	F_MA_02	F_MA_03	F_MA_04	F_MA_05	F_MA_06	F_MA_07	F_MA_08	F_MA_09	F_MA_10	F_MA_11	F_MA_12	F_MA_13
Tokens	281	43	43	23	23	18	21	14	17	6	33	7	6	27
Types	63	31	32	18	19	17	19	13	17	6	27	7	6	25

	DE1	DE2	DE3
F_MA_01	<constitutivo>	<exemplo>	
F_MA_02	<constitutivo>	<exemplo>	
F_MA_03	<constitutivo>	<exemplo>	<semelhança>
F_MA_04	<constitutivo>	<semelhança>	
F_MA_05	<constitutivo>	<exemplo>	
F_MA_06	<constitutivo>	<exemplo>	
F_MA_07	<constitutivo>		
F_MA_08	<constitutivo>	<exemplo>	
F_MA_09	<constitutivo>		
F_MA_10	<constitutivo>	<exemplo>	
F_MA_11	<constitutivo>		
F_MA_12	<constitutivo>		
F_MA_13	<constitutivo>	<exemplo>	<elaboração>



campo TÉSTE E MEDIDA (TM)

DT

Text File	OVERALL	F_TM_01	F_TM_02	F_TM_03	F_TM_04	F_TM_05	F_TM_06	F_TM_07	F_TM_08	F_TM_09	F_TM_10	F_TM_11	F_TM_12	F_TM_13	F_TM_14	F_TM_15	F_TM_16	F_TM_17	F_TM_18
Tokens	3.193	86	82	93	147	59	120	75	80	81	54	76	71	99	110	101	47	93	42
Types	287	61	66	71	98	50	76	51	57	58	46	58	53	66	72	63	38	59	38
Type/Token Ratio	8,99	70,93	80,49	76,34	66,67	84,75	63,33	68	71,25	71,6	85,19	76,32	74,65	66,67	65,45	62,38	80,85	63,44	90,48
Sentences	161	3	4	4	6	3	7	3	5	4	3	4	4	3	4	5	3	4	3

DE

Text File	OVERALL	F_TM_01	F_TM_02	F_TM_03	F_TM_04	F_TM_05	F_TM_06	F_TM_07	F_TM_08	F_TM_09	F_TM_10	F_TM_11	F_TM_12	F_TM_13	F_TM_14	F_TM_15	F_TM_16	F_TM_17	F_TM_18
Tokens	2.992	81	80	88	142	54	115	70	75	76	49	71	66	94	105	96	42	88	37
Types	282	57	64	67	95	47	72	46	52	54	41	54	50	62	68	60	35	55	34

DT

Text File	F_TM_19	F_TM_20	F_TM_21	F_TM_22	F_TM_23	F_TM_24	F_TM_25	F_TM_26	F_TM_27	F_TM_28	F_TM_29	F_TM_30	F_TM_31	F_TM_32	F_TM_33	F_TM_34	F_TM_35	F_TM_36	F_TM_37
Tokens	38	45	67	72	67	85	53	54	62	62	52	75	63	90	101	96	81	45	53
Types	33	38	49	54	48	63	45	47	46	49	44	59	53	69	73	77	64	31	44
Type/Token Ratio	86,84	84,44	73,13	75	71,64	74,12	84,91	87,04	74,19	79,03	84,62	78,67	84,13	76,67	72,28	80,21	79,01	68,89	83,02
Sentences	3	3	3	3	2	4	4	3	3	4	3	4	3	4	5	5	3	4	3

DE

Text File	F_TM_19	F_TM_20	F_TM_21	F_TM_22	F_TM_23	F_TM_24	F_TM_25	F_TM_26	F_TM_27	F_TM_28	F_TM_29	F_TM_30	F_TM_31	F_TM_32	F_TM_33	F_TM_34	F_TM_35	F_TM_36	F_TM_37
Tokens	33	41	62	67	63	80	49	50	58	58	47	71	58	85	96	91	77	41	48
Types	29	35	44	50	45	60	41	44	43	45	40	56	50	66	70	74	62	27	40

DT

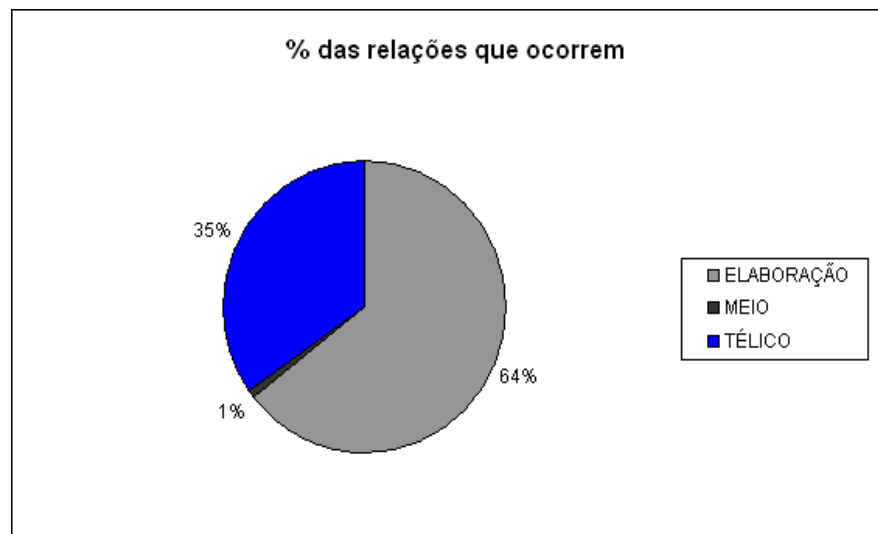
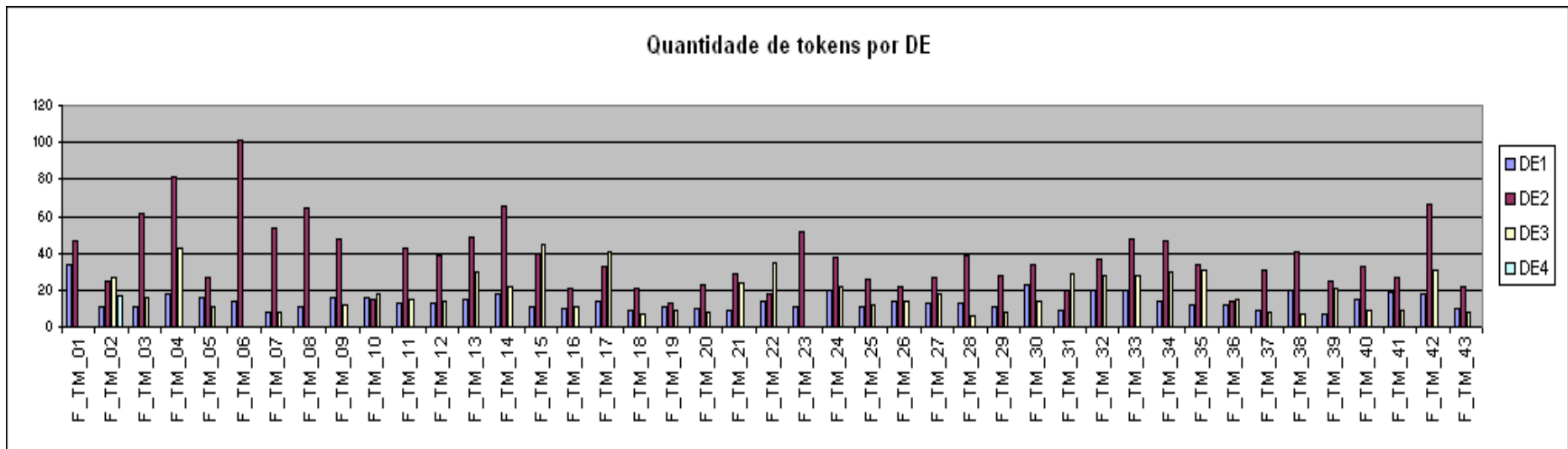
Text File	F_TM_38	F_TM_39	F_TM_40	F_TM_41	F_TM_42	F_TM_43
Tokens	73	58	61	59	120	45
Types	59	49	48	46	74	37
Type/Token Ratio	80,82	84,48	78,69	77,97	61,67	82,22
Sentences	5	3	4	3	5	3

DE

Text File	F_TM_38	F_TM_39	F_TM_40	F_TM_41	F_TM_42	F_TM_43
Tokens	68	53	57	55	115	40
Types	55	45	46	44	70	34

	DE1	DE2	DE3	DE4
F_TM_01	<télico>	<elaboração>		
F_TM_02	<télico>	<meio>	<télico>	<elaboração>
F_TM_03	<télico>	<elaboração>	<elaboração>	
F_TM_04	<télico>	<elaboração>	<elaboração>	
F_TM_05	<télico>	<elaboração>	<elaboração>	
F_TM_06	<télico>	<elaboração>		
F_TM_07	<télico>	<elaboração>	<elaboração>	
F_TM_08	<télico>	<elaboração>		
F_TM_09	<télico>	<elaboração>	<elaboração>	
F_TM_10	<télico>	<elaboração>	<elaboração>	
F_TM_11	<télico>	<elaboração>	<elaboração>	
F_TM_12	<télico>	<elaboração>	<elaboração>	
F_TM_13	<télico>	<elaboração>	<elaboração>	
F_TM_14	<télico>	<elaboração>	<elaboração>	
F_TM_15	<télico>	<elaboração>	<elaboração>	
F_TM_16	<télico>	<elaboração>	<elaboração>	
F_TM_17	<télico>	<elaboração>	<elaboração>	
F_TM_18	<télico>	<elaboração>	<elaboração>	
F_TM_19	<télico>	<elaboração>	<elaboração>	
F_TM_20	<télico>	<elaboração>	<elaboração>	
F_TM_21	<télico>	<elaboração>	<elaboração>	
F_TM_22	<télico>	<elaboração>	<elaboração>	

	DE1	DE2	DE3	DE4
F_TM_23	<télico>	<elaboração>		
F_TM_24	<télico>	<elaboração>	<elaboração>	
F_TM_25	<télico>	<elaboração>	<elaboração>	
F_TM_26	<télico>	<elaboração>	<elaboração>	
F_TM_27	<télico>	<elaboração>	<elaboração>	
F_TM_28	<télico>	<elaboração>	<elaboração>	
F_TM_29	<télico>	<elaboração>	<elaboração>	
F_TM_30	<télico>	<elaboração>	<elaboração>	
F_TM_31	<télico>	<elaboração>	<elaboração>	
F_TM_32	<télico>	<elaboração>	<elaboração>	
F_TM_33	<télico>	<elaboração>	<elaboração>	
F_TM_34	<télico>	<elaboração>	<elaboração>	
F_TM_35	<télico>	<elaboração>	<elaboração>	
F_TM_36	<télico>	<elaboração>	<elaboração>	
F_TM_37	<télico>	<elaboração>	<elaboração>	
F_TM_38	<télico>	<elaboração>	<elaboração>	
F_TM_39	<télico>	<elaboração>	<elaboração>	
F_TM_40	<télico>	<elaboração>	<elaboração>	
F_TM_41	<télico>	<elaboração>	<elaboração>	
F_TM_42	<télico>	<elaboração>	<elaboração>	
F_TM_43	<télico>	<elaboração>	<elaboração>	



campo TÉCNICA E RECURSO TERAPÊUTICO (TR)

DT

Text File	OVERALL	F_TR_01	F_TR_02	F_TR_03	F_TR_04	F_TR_05	F_TR_06	F_TR_07	F_TR_08	F_TR_09	F_TR_10	F_TR_11	F_TR_12	F_TR_13	F_TR_14	F_TR_15	F_TR_16	F_TR_17	F_TR_18
Tokens	3.939	35	60	81	35	29	32	17	45	20	19	30	93	154	130	184	91	34	54
Types	456	30	46	59	32	25	30	17	35	19	17	28	74	110	95	118	70	29	47
Type/Token Ratio	11,58	85,71	76,67	72,84	91,43	86,21	93,75	100	77,78	95	89,47	93,33	79,57	71,43	73,08	64,13	76,92	85,29	87,04
Sentences	173	2	2	2	1	2	2	1	2	1	1	2	4	6	6	7	4	3	3

DE

Text File	OVERALL	F_TR_01	F_TR_02	F_TR_03	F_TR_04	F_TR_05	F_TR_06	F_TR_07	F_TR_08	F_TR_09	F_TR_10	F_TR_11	F_TR_12	F_TR_13	F_TR_14	F_TR_15	F_TR_16	F_TR_17	F_TR_18
Tokens	3.742	33	58	79	33	27	30	15	43	18	17	28	88	149	129	179	89	33	53
Types	444	28	44	57	30	23	28	15	34	17	15	26	69	106	95	117	69	28	47

DT

Text File	F_TR_19	F_TR_20	F_TR_21	F_TR_22	F_TR_23	F_TR_24	F_TR_25	F_TR_26	F_TR_27	F_TR_28	F_TR_29	F_TR_30	F_TR_31	F_TR_32	F_TR_33	F_TR_34	F_TR_35	F_TR_36	F_TR_37
Tokens	40	26	38	49	64	66	55	24	87	33	87	66	75	66	44	72	30	36	83
Types	33	25	33	41	51	57	48	23	73	29	66	54	61	54	38	54	29	33	70
Type/Token Ratio	82,5	96,15	86,84	83,67	79,69	86,36	87,27	95,83	83,91	87,88	75,86	81,82	81,33	81,82	86,36	75	96,67	91,67	84,34
Sentences	1	2	2	3	2	2	3	1	3	1	4	4	4	3	3	3	2	1	4

DE

Text File	F_TR_19	F_TR_20	F_TR_21	F_TR_22	F_TR_23	F_TR_24	F_TR_25	F_TR_26	F_TR_27	F_TR_28	F_TR_29	F_TR_30	F_TR_31	F_TR_32	F_TR_33	F_TR_34	F_TR_35	F_TR_36	F_TR_37
Tokens	38	24	36	46	57	64	53	22	85	29	85	62	71	64	42	70	28	34	81
Types	31	23	31	39	46	55	46	21	71	26	64	51	58	53	36	52	27	31	68

DT

Text File	F_TR_38	F_TR_39	F_TR_40	F_TR_41	F_TR_42	F_TR_43	F_TR_44	F_TR_45	F_TR_46	F_TR_47	F_TR_48	F_TR_49	F_TR_50	F_TR_51	F_TR_52	F_TR_53	F_TR_54	F_TR_55	F_TR_56
Tokens	41	31	66	55	43	55	54	60	63	42	18	91	74	32	43	30	44	66	89
Types	36	25	52	47	37	47	47	51	55	39	16	73	58	31	38	25	38	52	62
Type/Token Ratio	87,8	80,65	78,79	85,45	86,05	85,45	87,04	85	87,3	92,86	88,89	80,22	78,38	96,88	88,37	83,33	86,36	78,79	69,66
Sentences	2	1	3	2	2	2	2	3	3	1	1	4	3	2	2	2	1	3	3

DE

Text File	F_TR_38	F_TR_39	F_TR_40	F_TR_41	F_TR_42	F_TR_43	F_TR_44	F_TR_45	F_TR_46	F_TR_47	F_TR_48	F_TR_49	F_TR_50	F_TR_51	F_TR_52	F_TR_53	F_TR_54	F_TR_55	F_TR_56
Tokens	38	28	65	53	41	53	52	58	61	40	15	87	70	28	39	26	40	62	85
Types	34	23	51	45	35	45	45	49	53	37	14	70	55	28	35	22	35	50	60

DT

Text File	F_TR_57	F_TR_58	F_TR_59	F_TR_60	F_TR_61	F_TR_62	F_TR_63	F_TR_64	F_TR_65	F_TR_66	F_TR_67	F_TR_68	F_TR_69	F_TR_70	F_TR_71	F_TR_72	F_TR_73	F_TR_74
Tokens	46	41	49	29	56	51	28	29	29	22	31	39	64	42	36	76	43	47
Types	39	36	41	27	44	39	24	25	26	21	30	31	53	32	31	58	34	38
Type/Token Ratio	84,78	87,8	83,67	93,1	78,57	76,47	85,71	86,21	89,66	95,45	96,77	79,49	82,81	76,19	86,11	76,32	79,07	80,85
Sentences	2	2	2	1	1	1	2	2	2	2	2	2	2	2	1	4	1	1

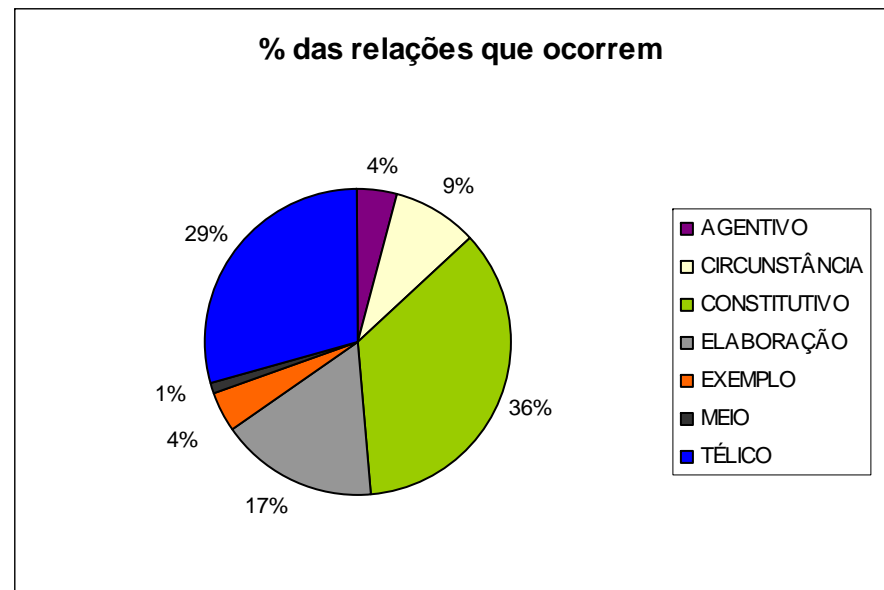
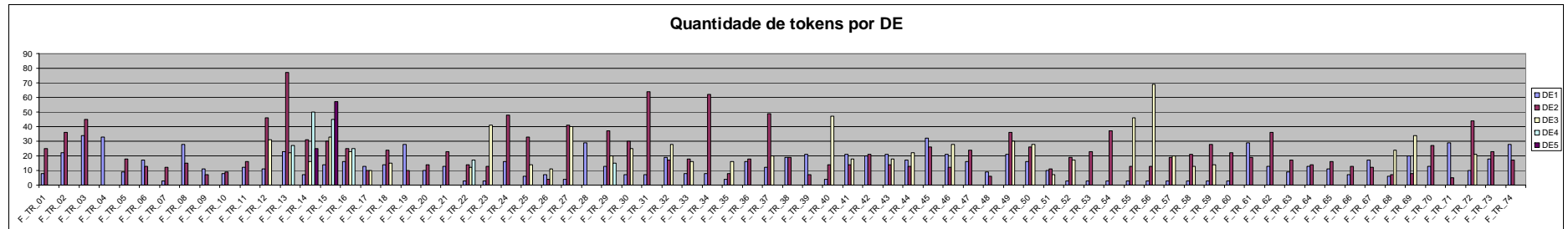
DE

Text File	F_TR_57	F_TR_58	F_TR_59	F_TR_60	F_TR_61	F_TR_62	F_TR_63	F_TR_64	F_TR_65	F_TR_66	F_TR_67	F_TR_68	F_TR_69	F_TR_70	F_TR_71	F_TR_72	F_TR_73	F_TR_74
Tokens	42	37	45	25	48	49	26	27	27	20	29	37	62	40	34	75	41	45
Types	36	32	38	24	39	37	22	23	24	19	28	29	51	30	29	57	32	36

	DE1	DE2	DE3	DE4	DE5
F_TR_01	<circunstância>	<télico>			
F_TR_02	<constitutivo>	<télico>			
F_TR_03	<constitutivo>	<télico>			
F_TR_04	<constitutivo>				
F_TR_05	<constitutivo>	<télico>			
F_TR_06	<constitutivo>	<elaboração>			
F_TR_07	<agentivo>	<télico>			
F_TR_08	<constitutivo>	<télico>			
F_TR_09	<constitutivo>	<agentivo>			
F_TR_10	<constitutivo>	<télico>			
F_TR_11	<constitutivo>	<consequência>			
F_TR_12	<télico>	<constitutivo>	<elaboração>		
F_TR_13	<constitutivo>	<elaboração>	<télico>		<elaboração>
F_TR_14	<agentivo>	<constitutivo>	<elaboração>	<télico>	<elaboração>
F_TR_15	<télico>	<constitutivo>	<consequência>	<exemplo>	<elaboração>
F_TR_16	<constitutivo>	<constitutivo>	<télico>	<elaboração>	
F_TR_17	<constitutivo>	<elaboração>	<exemplo>		
F_TR_18	<constitutivo>	<constitutivo>	<télico>		
F_TR_19	<télico>	<meio>			
F_TR_20	<constitutivo>	<télico>			
F_TR_21	<constitutivo>	<télico>			
F_TR_22	<agentivo>	<constitutivo>	<télico>	<elaboração>	
F_TR_23	<agentivo>	<constitutivo>	<télico>		
F_TR_24	<constitutivo>	<télico>			
F_TR_25	<circunstância>	<constitutivo>	<elaboração>		
F_TR_26	<meio>	<agentivo>	<télico>		
F_TR_27	<agentivo>	<constitutivo>	<elaboração>		

	DE1	DE2	DE3	DE4	DE5
F_TR_28	<télico>				
F_TR_29	<agentivo>	<constitutivo>	<télico>	<elaboração>	
F_TR_30	<circunstância>	<constitutivo>	<télico>		
F_TR_31	<circunstância>	<constitutivo>			
F_TR_32	<constitutivo>	<constitutivo>	<constitutivo>		
F_TR_33	<télico>	<constitutivo>	<télico>		
F_TR_34	<télico>	<elaboração>			
F_TR_35	<circunstância>	<télico>	<constitutivo>		
F_TR_36	<constitutivo>	<télico>			
F_TR_37	<constitutivo>	<elaboração>	<télico>		
F_TR_38	<constitutivo>	<télico>			
F_TR_39	<constitutivo>	<télico>			
F_TR_40	<circunstância>	<télico>	<elaboração>		
F_TR_41	<constitutivo>	<télico>	<elaboração>		
F_TR_42	<constitutivo>	<télico>			
F_TR_43	<constitutivo>	<télico>	<elaboração>		
F_TR_44	<constitutivo>	<télico>	<elaboração>		
F_TR_45	<constitutivo>	<elaboração>			
F_TR_46	<constitutivo>	<télico>	<elaboração>		
F_TR_47	<constitutivo>	<télico>			
F_TR_48	<constitutivo>	<télico>			
F_TR_49	<télico>	<elaboração>	<elaboração>		
F_TR_50	<télico>	<elaboração>	<elaboração>		
F_TR_51	<télico>	<constitutivo>	<télico>		
F_TR_52	<circunstância>	<constitutivo>	<elaboração>		
F_TR_53	<circunstância>	<constitutivo>			
F_TR_54	<circunstância>	<constitutivo>			

	DE1	DE2	DE3	DE4	DE5
F_TR_55	<circunstância>	<télico>	<elaboração>		
F_TR_56	<circunstância>	<télico>	<elaboração>		
F_TR_57	<circunstância>	<constitutivo>	<télico>		
F_TR_58	<circunstância>	<constitutivo>	<télico>		
F_TR_59	<circunstância>	<constitutivo>	<elaboração>		
F_TR_60	<circunstância>	<constitutivo>			
F_TR_61	<constitutivo>	<télico>			
F_TR_62	<constitutivo>	<télico>			
F_TR_63	<constitutivo>	<exemplo>			
F_TR_64	<constitutivo>	<exemplo>			
F_TR_65	<constitutivo>	<exemplo>			
F_TR_66	<constitutivo>	<exemplo>			
F_TR_67	<circunstância>	<exemplo>			
F_TR_68	<constitutivo>	<télico>	<elaboração>		
F_TR_69	<constitutivo>	<télico>	<exemplo>		
F_TR_70	<télico>	<constitutivo>			
F_TR_71	<constitutivo>	<télico>			
F_TR_72	<circunstância>	<constitutivo>	<elaboração>		
F_TR_73	<constitutivo>	<télico>			
F_TR_74	<constitutivo>	<télico>			



Campo DEFEITO (DE)

DT

Text File	OVERALL	R_DE_01	R_DE_02	R_DE_03	R_DE_04	R_DE_05	R_DE_06	R_DE_07	R_DE_08	R_DE_09	R_DE_10	R_DE_11
Tokens	922	43	84	69	38	24	51	63	49	49	25	37
Types	115	40	65	52	34	23	47	51	43	44	25	34
Type/Token Ratio	12,47	93,02	77,38	75,36	89,47	95,83	92,16	80,95	87,76	89,8	100	91,89
Sentences	35	2	3	1	1	1	2	2	2	2	1	1

DE

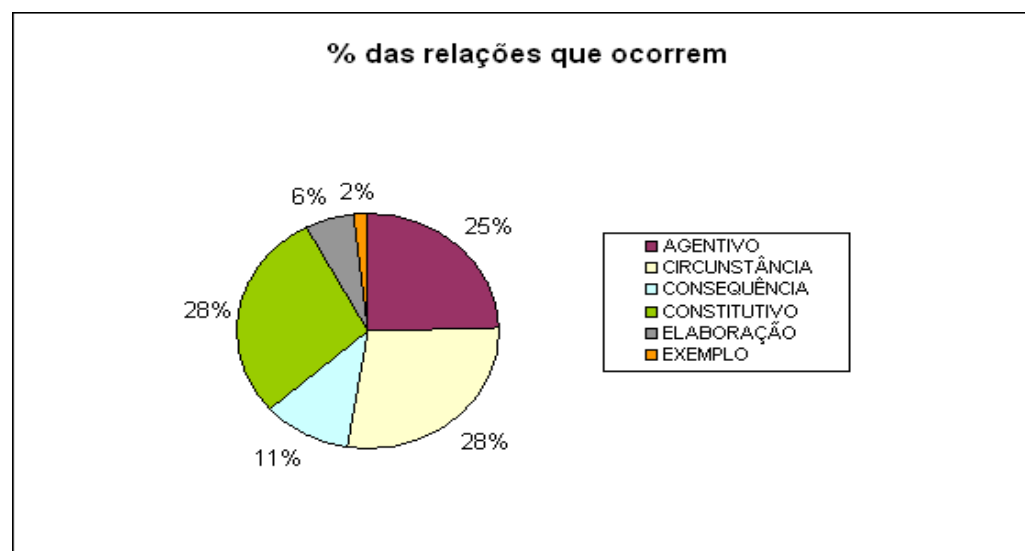
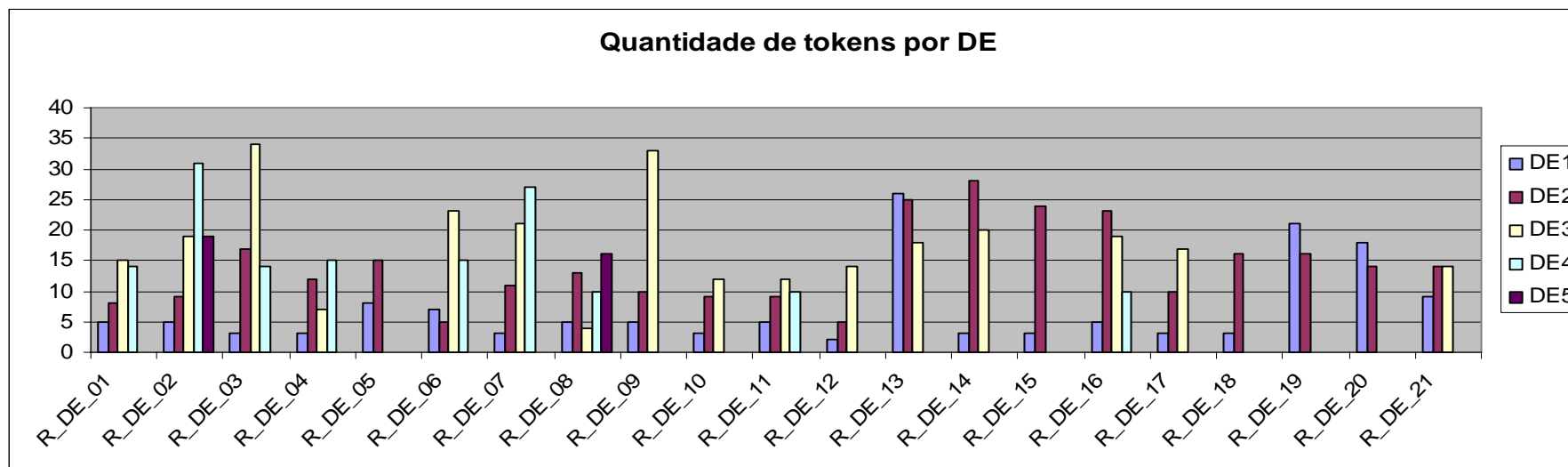
Text rle	OVERALL	R_DE_01	R_DE_02	R_DE_03	R_DE_04	R_DE_05	R_DE_06	R_DE_07	R_DE_08	R_DE_09	R_DE_10	R_DE_11
Tokens	901	37	32	37	19	30	57	27	51	69	21	36
Types	114	30	25	31	19	28	47	25	44	52	20	33

DT

Text File	R_DE_12	R_DE_13	R_DE_14	R_DE_15	R_DE_16	R_DE_17	R_DE_18	R_DE_19	R_DE_20	R_DE_21
Tokens	22	70	52	28	58	31	20	38	33	38
Types	21	53	45	26	48	28	20	32	26	31
Type/Token Ratio	95,45	75,71	86,54	92,86	82,76	90,32	100	84,21	78,79	81,58
Sentences	1	2	2	1	2	2	1	2	2	2

DE

Text rle	R_DE_12	R_DE_13	R_DE_14	R_DE_15	R_DE_16	R_DE_17	R_DE_18	R_DE_19	R_DE_20	R_DE_21
Tokens	24	48	48	62	50	23	37	68	83	42
Types	24	43	42	50	46	22	33	31	64	39



	DE1	DE2	DE3	DE4	DE5
R_DE_01	<circunstância>	<constitutivo>	<agentivo>	<consequência>	
R_DE_02	<circunstância>	<constitutivo>	<agentivo>	<elaboração>	<elaboração>
R_DE_03	<circunstância>	<constitutivo>	<agentivo>	<consequência>	
R_DE_04	<circunstância>	<constitutivo>	<agentivo>	<consequência>	
R_DE_05	<constitutivo>	<agentivo>			
R_DE_06	<constitutivo>	<circunstância>	<agentivo>	<elaboração>	
R_DE_07	<circunstância>	<constitutivo>	<agentivo>	<circunstância>	
R_DE_08	<circunstância>	<constitutivo>	<consequência>	<elaboração>	<agentivo>
R_DE_09	<circunstância>	<constitutivo>	<agentivo>		
R_DE_10	<circunstância>	<constitutivo>	<agentivo>		
R_DE_11	<circunstância>	<constitutivo>	<agentivo>	<consequência>	
R_DE_12	<circunstância>	<agentivo>	<circunstância>		
R_DE_13	<constitutivo>	<agentivo>	<exemplo>		
R_DE_14	<circunstância>	<constitutivo>	<agentivo>		
R_DE_15	<circunstância>	<constitutivo>			
R_DE_16	<circunstância>	<constitutivo>	<agentivo>	<consequência>	
R_DE_17	<circunstância>	<constitutivo>	<associação>		
R_DE_18	<circunstância>	<agentivo>			
R_DE_19	<constitutivo>	<associação>			
R_DE_20	<constitutivo>	<associação>			
R_DE_21	<constitutivo>	<consequência>	<associação>		

campo ENSAIO (EN)

DT

Text File	OVERALL	R_EN_01	R_EN_02	R_EN_03	R_EN_04	R_EN_05	R_EN_06	R_EN_07	R_EN_08	R_EN_09	R_EN_10	R_EN_11	R_EN_12
Tokens	2.191	43	89	86	143	130	45	74	69	78	44	50	48
Types	259	31	65	60	75	70	35	51	49	55	36	40	35
Type/Token Ratio	11,82	72,09	73,03	69,77	52,45	53,85	77,78	68,92	71,01	70,51	81,82	80	72,92
Sentences	104	3	3	4	8	5	2	4	3	3	3	3	3

DE

Text rfile	OVERALL	R_EN_01	R_EN_02	R_EN_03	R_EN_04	R_EN_05	R_EN_06	R_EN_07	R_EN_08	R_EN_09	R_EN_10	R_EN_11	R_EN_12
Tokens	2.130	42	88	85	142	129	44	73	68	77	43	49	47
Types	257	30	64	59	75	69	34	50	48	54	35	39	34

DT

Text File	R_EN_13	R_EN_14	R_EN_15	R_EN_16	R_EN_17	R_EN_18	R_EN_19	R_EN_20	R_EN_21	R_EN_22	R_EN_23	R_EN_24	R_EN_25
Tokens	66	96	78	54	36	41	20	39	57	62	112	38	93
Types	47	67	55	43	28	36	19	34	38	49	76	31	65
Type/Token Ratio	71,21	69,79	70,51	79,63	77,78	87,8	95	87,18	66,67	79,03	67,86	81,58	69,89
Sentences	3	4	3	3	2	2	1	2	3	3	5	2	4

DE

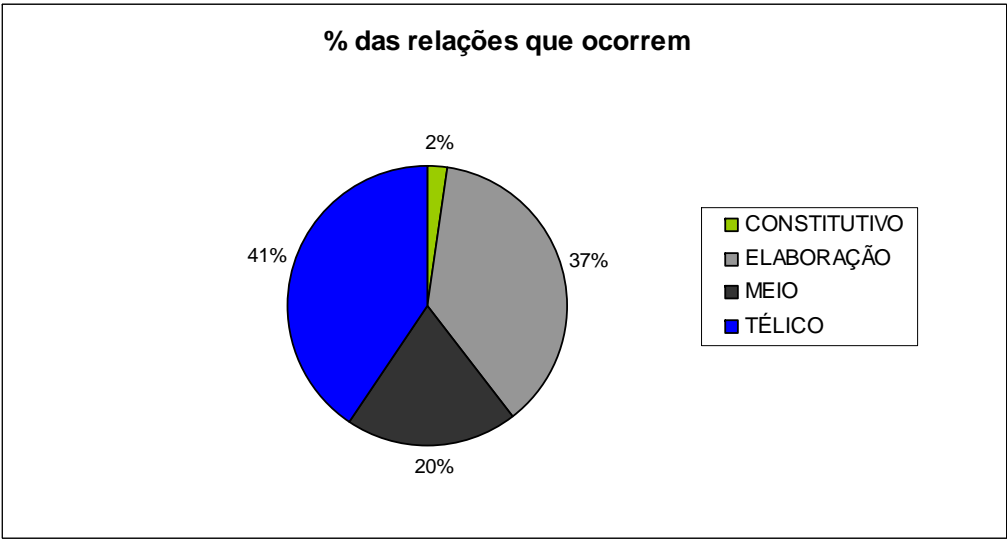
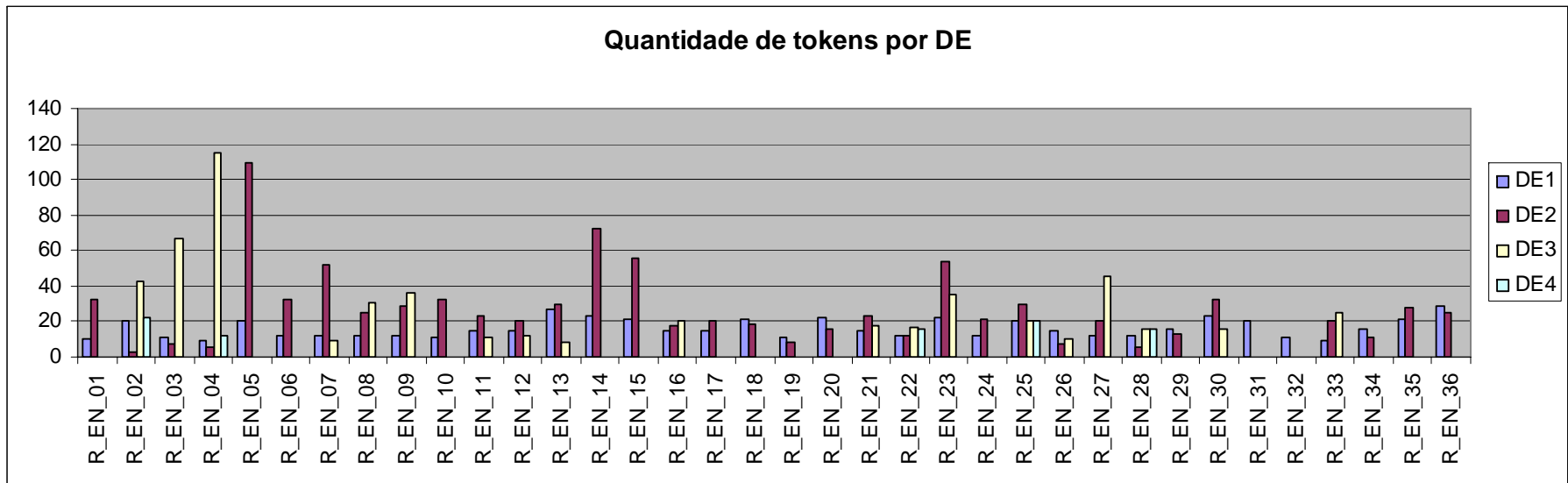
Text rfile	R_EN_13	R_EN_14	R_EN_15	R_EN_16	R_EN_17	R_EN_18	R_EN_19	R_EN_20	R_EN_21	R_EN_22	R_EN_23	R_EN_24	R_EN_25
Tokens	65	95	77	53	35	40	19	38	56	57	111	33	90
Types	46	66	54	42	27	35	18	33	37	45	75	27	63

DT

Text File	R_EN_26	R_EN_27	R_EN_28	R_EN_29	R_EN_30	R_EN_31	R_EN_32	R_EN_33	R_EN_34	R_EN_35	R_EN_36
Tokens	33	78	51	30	72	26	17	55	31	52	55
Types	25	54	40	26	51	24	16	43	27	40	43
Type/Token Ratio	75,76	69,23	78,43	86,67	70,83	92,31	94,12	78,18	87,1	76,92	78,18
Sentences	2	3	3	2	3	1	1	2	2	2	2

DE

Text rfile	R_EN_26	R_EN_27	R_EN_28	R_EN_29	R_EN_30	R_EN_31	R_EN_32	R_EN_33	R_EN_34	R_EN_35	R_EN_36
Tokens	32	77	50	29	71	20	11	54	27	49	54
Types	24	53	39	25	51	18	11	43	24	39	43



	DE1	DE2	DE3	DE4
R_EN_01	<télico>	<elaboração>		
R_EN_02	<télico>	<meio>	<elaboração>	<elaboração>
R_EN_03	<télico>	<meio>	<elaboração>	
R_EN_04	<télico>	<meio>	<elaboração>	<elaboração>
R_EN_05	<télico>	<elaboração>		
R_EN_06	<télico>	<elaboração>		
R_EN_07	<télico>	<elaboração>	<meio>	
R_EN_08	<télico>	<elaboração>	<elaboração>	
R_EN_09	<télico>	<elaboração>	<meio>	
R_EN_10	<télico>	<elaboração>		
R_EN_11	<télico>	<elaboração>	<meio>	
R_EN_12	<télico>	<elaboração>	<meio>	
R_EN_13	<télico>	<elaboração>	<meio>	
R_DE_14	<télico>	<elaboração>		
R_EN_15	<télico>	<elaboração>		
R_EN_16	<télico>	<meio>	<elaboração>	
R_EN_17	<télico>	<elaboração>		
R_EN_18	<télico>	<meio>		
R_EN_19	<télico>	<meio>		
R_EN_20	<télico>	<elaboração>		

R_EN_21	<télico>	<elaboração>	<elaboração>	
R_EN_22	<télico>	<meio>	<meio>	<elaboração>
R_EN_23	<télico>	<elaboração>	<elaboração>	
R_EN_24	<télico>	<elaboração>		
R_EN_25	<télico>	<elaboração>	<constitutivo>	<elaboração>
R_EN_26	<télico>	<meio>	<télico>	
R_EN_27	<télico>	<meio>	<elaboração>	
R_EN_28	<télico>	<meio>	<elaboração>	<elaboração>
R_EN_29	<télico>	<elaboração>		
R_EN_30	<télico>	<elaboração>	<télico>	
R_EN_31	<télico>			
R_EN_32	<télico>			
R_EN_33	<télico>	<constitutivo>	<meio>	
R_EN_34	<télico>	<meio>		
R_EN_35	<télico>	<elaboração>		
R_EN_36	<télico>	<meio>		

campo ETAPA OU MÉTODO DE FABRICAÇÃO (EP)

DT

Text File	OVERALL	R_EP_01	R_EP_02	R_EP_03	R_EP_04	R_EP_05	R_EP_06	R_EP_07	R_EP_08	R_EP_09	R_EP_10	R_EP_11	R_EP_12	R_EP_13	R_EP_14	R_EP_15	R_EP_16	R_EP_17
Tokens	1.369	33	20	20	22	20	27	25	0	50	19	67	53	56	58	41	56	61
Types	172	30	18	19	21	18	24	24	0	42	19	53	41	44	45	31	41	44
Type/Token Ratio	12,56	90,91	90	95	95,45	90	88,89	96	84	100	79,1	77,36	78,57	77,59	75,61	73,21	72,13	78,26
Sentences	56	1	1	1	1	1	1	1	0	2	1	2	2	2	2	1	2	3

DE

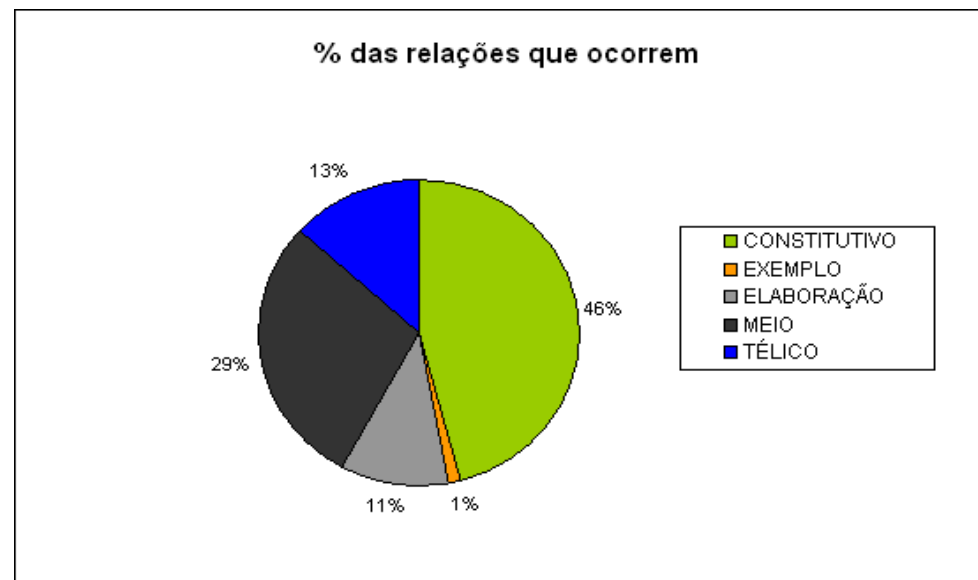
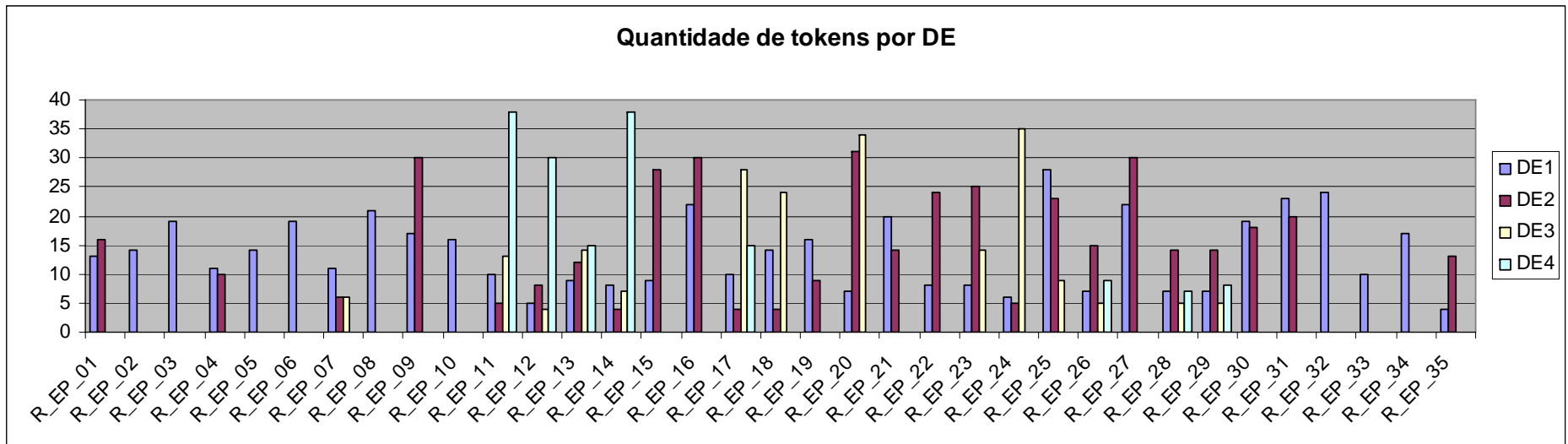
Text rfile	OVERALL	R_EP_01	R_EP_02	R_EP_03	R_EP_04	R_EP_05	R_EP_06	R_EP_07	R_EP_08	R_EP_09	R_EP_10	R_EP_11	R_EP_12	R_EP_13	R_EP_14	R_EP_15	R_EP_16	R_EP_17
Tokens	1.250	29	14	19	21	14	19	23	21	47	16	66	47	50	57	37	52	57
Types	169	26	13	18	20	13	17	23	19	40	16	52	38	40	44	27	37	40

DT

Text File	R_EP_18	R_EP_19	R_EP_20	R_EP_21	R_EP_22	R_EP_23	R_EP_24	R_EP_25	R_EP_26	R_EP_27	R_EP_28	R_EP_29	R_EP_30	R_EP_31	R_EP_32	R_EP_33	R_EP_34	R_EP_35
Tokens	46	32	76	43	39	54	55	64	37	56	34	35	41	47	28	14	20	20
Types	36	26	51	33	33	43	46	50	25	41	24	26	35	42	23	14	17	18
Type/Token Ratio	81,25	67,11	76,74	84,62	79,63	83,64	78,13	67,57	73,21	70,59	74,29	85,37	89,36	82,14	100	85	90	
Sentences	2	2	4	2	2	3	2	3	1	2	1	0	2	2	1	1	1	1

DE

Text rfile	R_EP_18	R_EP_19	R_EP_20	R_EP_21	R_EP_22	R_EP_23	R_EP_24	R_EP_25	R_EP_26	R_EP_27	R_EP_28	R_EP_29	R_EP_30	R_EP_31	R_EP_32	R_EP_33	R_EP_34	R_EP_35
Tokens	42	25	72	34	32	47	46	60	36	52	33	34	37	43	24	10	17	17
Types	32	22	49	27	28	39	40	47	24	37	23	25	32	39	19	10	15	16



	DE1	DE2	DE3	DE4
R_EP_01	<constitutivo>	<télico>		
R_EP_02	<constitutivo>			
R_EP_03	<constitutivo>			
R_EP_04	<constitutivo>	<meio>		
R_EP_05	<constitutivo>			
R_EP_06	<constitutivo>			
R_EP_07	<constitutivo>	<meio>	<télico>	
R_EP_08	<constitutivo>			
R_EP_09	<constitutivo>	<elaboração>		
R_EP_10	<constitutivo>			
R_EP_11	<constitutivo>	<meio>	<télico>	<elaboração>
R_EP_12	<meio>	<constitutivo>	<meio>	<télico>
R_DE_13	<constitutivo>	<meio>	<télico>	<meio>
R_DE_14	<constitutivo>	<meio>	<télico>	<elaboração>
R_EP_15	<constitutivo>	<télico>		
R_EP_16	<constitutivo>	<meio>		
R_EP_17	<constitutivo>	<meio>	<elaboração>	<elaboração>
R_EP_18	<constitutivo>	<meio>	<elaboração>	

R_EP_19	<constitutivo>	<meio>		
R_EP_20	<constitutivo>	<meio>	<elaboração>	
R_EP_21	<constitutivo>	<elaboração>		
R_EP_22	<meio>	<constitutivo>		
R_EP_23	<meio>	<constitutivo>	<elaboração>	
R_EP_24	<constitutivo>	<meio>	<télico>	
R_EP_25	<constitutivo>	<télico>	<exemplo>	
R_EP_26	<constitutivo>	<meio>	<constitutivo>	<meio>
R_EP_27	<constitutivo>	<meio>		
R_EP_28	<constitutivo>	<meio>	<constitutivo>	<meio>
R_EP_29	<constitutivo>	<meio>	<constitutivo>	<meio>
R_EP_30	<constitutivo>	<télico>		
R_EP_31	<constitutivo>	<télico>		
R_EP_32	<constitutivo>			
R_EP_33	<constitutivo>			
R_EP_34	<constitutivo>			
R_EP_35	<constitutivo>	<meio>		

campo INSTRUMENTO OU EQUIPAMENTO (IN)

DT

Text File	OVERALL	R_IN_01	R_IN_02	R_IN_03	R_IN_04	R_IN_05	R_IN_06	R_IN_07	R_IN_08	R_IN_09	R_IN_10	R_IN_11	R_IN_12	R_IN_13	R_IN_14	R_IN_15	R_IN_16	R_IN_17
Tokens	2.866	65	44	142	30	50	25	90	49	41	85	35	42	42	86	36	44	75
Types	318	52	36	83	27	44	23	62	36	37	58	31	35	35	62	28	38	58
Type/Token Ratio	11,1	80	81,82	58,45	90	88	92	68,89	73,47	90,24	68,24	88,57	83,33	83,33	72,09	77,78	86,36	77,33
Sentences	116	3	2	4	1	2	2	3	2	1	3	1	1	2	3	2	1	3

DE

Text File	OVERALL	R_IN_01	R_IN_02	R_IN_03	R_IN_04	R_IN_05	R_IN_06	R_IN_07	R_IN_08	R_IN_09	R_IN_10	R_IN_11	R_IN_12	R_IN_13	R_IN_14	R_IN_15	R_IN_16	R_IN_17
Tokens	2.781	64	41	141	29	49	24	89	48	40	84	34	41	41	85	35	43	74
Types	310	51	35	83	26	43	22	61	35	36	57	30	34	34	62	27	37	58

DT

Text File	R_IN_18	R_IN_19	R_IN_20	R_IN_21	R_IN_22	R_IN_23	R_IN_24	R_IN_25	R_IN_26	R_IN_27	R_IN_28	R_IN_29	R_IN_30	R_IN_31	R_IN_32	R_IN_33	R_IN_34	R_IN_35
Tokens	74	75	75	80	61	57	99	110	52	55	54	66	33	39	44	55	67	85
Types	55	59	56	60	45	43	75	69	41	40	41	47	29	32	35	41	53	62
Type/Token Ratio	74,32	78,67	74,67	75	73,77	75,44	75,76	62,73	78,85	72,73	75,93	71,21	87,88	82,05	79,55	74,55	79,1	72,94
Sentences	3	2	3	3	2	3	6	3	2	2	2	2	2	2	3	2	3	4

DE

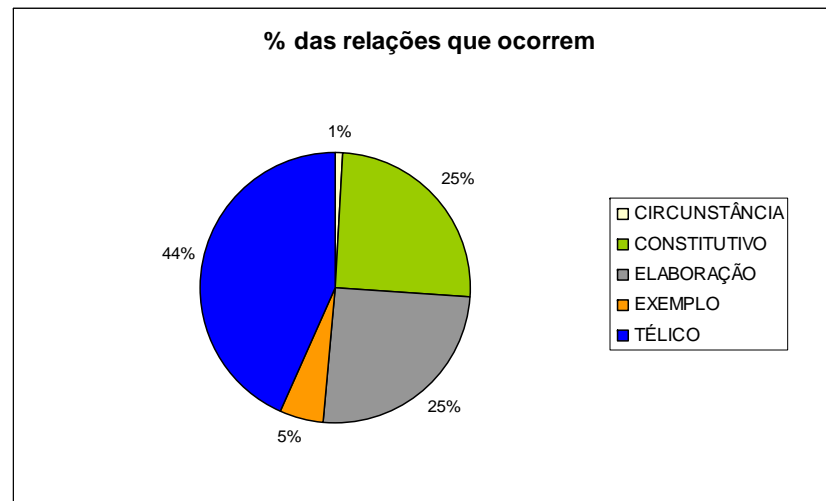
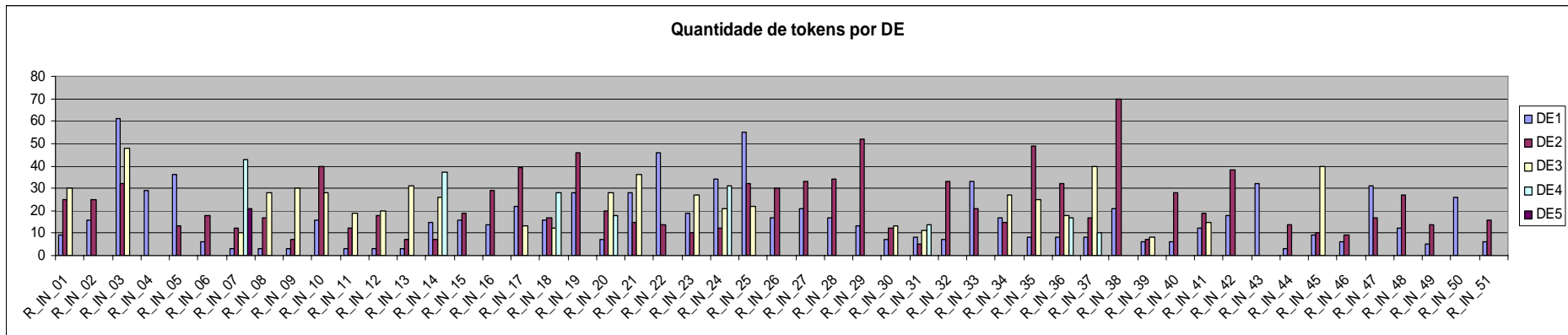
Text File	R_IN_18	R_IN_19	R_IN_20	R_IN_21	R_IN_22	R_IN_23	R_IN_24	R_IN_25	R_IN_26	R_IN_27	R_IN_28	R_IN_29	R_IN_30	R_IN_31	R_IN_32	R_IN_33	R_IN_34	R_IN_35
Tokens	73	74	73	79	60	56	98	109	47	54	51	65	32	38	40	54	59	82
Types	54	58	54	59	45	42	74	68	38	39	38	46	28	31	31	40	48	60

DT

Text File	R_IN_36	R_IN_37	R_IN_38	R_IN_39	R_IN_40	R_IN_41	R_IN_42	R_IN_43	R_IN_44	R_IN_45	R_IN_46	R_IN_47	R_IN_48	R_IN_49	R_IN_50	R_IN_51
Tokens	78	77	98	22	35	47	57	33	18	61	16	49	40	20	27	26
Types	59	61	70	21	31	43	45	26	17	52	14	40	35	18	24	24
Type/Token Ratio	75,64	79,22	71,43	95,45	88,57	91,49	78,95	78,79	94,44	85,25	87,5	81,63	87,5	90	88,89	92,31
Sentences	4	4	3	2	1	2	2	1	1	3	1	2	2	1	1	1

DE

Text File	R_IN_36	R_IN_37	R_IN_38	R_IN_39	R_IN_40	R_IN_41	R_IN_42	R_IN_43	R_IN_44	R_IN_45	R_IN_46	R_IN_47	R_IN_48	R_IN_49	R_IN_50	R_IN_51
Tokens	75	75	91	21	34	46	56	32	17	59	15	48	39	19	26	22
Types	56	60	64	20	30	42	44	25	16	50	13	40	35	17	23	20



	DE1	DE2	DE3	DE4	DE5
R_IN_01	<constitutivo>	<télico>	<elaboração>		
R_IN_02	<constitutivo>	<télico>			
IN_03	<constitutivo>	<télico>	<elaboração>		
R_IN_04	<télico>				
R_IN_05	<constitutivo>	<télico>			
R_IN_06	<constitutivo>	<télico>			
R_IN_07	<télico>	<constitutivo>	<télico>	<elaboração>	<elaboração>
R_IN_08	<télico>	<constitutivo>	<elaboração>		
R_IN_09	<télico>	<constitutivo>	<télico>		
R_IN_10	<télico>	<constitutivo>	<elaboração>		
R_IN_11	<télico>	<constitutivo>	<télico>		
R_IN_12	<télico>	<constitutivo>	<télico>		
R_IN_13	<télico>	<constitutivo>	<elaboração>		
R_IN_14	<télico>	<constitutivo>	<télico>	<elaboração>	
R_IN_15	<télico>	<télico>			
R_IN_16	<télico>	<elaboração>			
R_IN_17	<télico>	<constitutivo>	<elaboração>		
R_IN_18	<télico>	<constitutivo>	<exemplo>	<elaboração>	
R_IN_19	<télico>	<constitutivo>			
R_IN_20	<constitutivo>	<télico>	<elaboração>	<elaboração>	
R_IN_21	<télico>	<elaboração>	<exemplo>		
R_IN_22	<télico>	<elaboração>			
R_IN_23	<télico>	<exemplo>	<elaboração>		
R_IN_24	<télico>	<elaboração>	<constitutivo>	<elaboração>	
R_IN_25	<télico>	<constitutivo>	<elaboração>		
R_IN_26	<télico>	<elaboração>			
R_IN_27	<télico>	<elaboração>			

R_IN_28	<télico>	<elaboração>			
R_IN_29	<télico>	<elaboração>			
R_IN_30	<constitutivo>	<télico>	<télico>		
R_IN_31	<télico>	<exemplo>	<elaboração>	<elaboração>	
R_IN_32	<télico>	<elaboração>			
R_IN_33	<télico>	<elaboração>			
R_IN_34	<télico>	<constitutivo>	<elaboração>		
R_IN_35	<télico>	<constitutivo>	<elaboração>		
R_IN_36	<télico>	<constitutivo>	<elaboração>	<elaboração>	
R_IN_37	<télico>	<constitutivo>	<elaboração>	<exemplo>	
R_IN_38	<télico>	<constitutivo>			
R_IN_39	<télico>	<constitutivo>	<exemplo>		
R_IN_40	<constitutivo>	<télico>			
R_IN_41	<télico>	<constitutivo>	<exemplo>		
R_IN_42	<télico>	<elaboração>			
R_IN_43	<télico>				
R_IN_44	<constitutivo>	<télico>			
R_IN_45	<télico>	<constitutivo>	<elaboração>		
R_IN_46	<télico>	<constitutivo>			
R_IN_47	<télico>	<constitutivo>			
R_IN_48	<télico>	<constitutivo>			
R_IN_49	<circunstância>	<télico>			
R_IN_50	<télico>				
R_IN_51	<constitutivo>	<télico>			

campo MATÉRIA-PRIMA (MP)

DT

Text File	OVERALL	R_MP_01	R_MP_02	R_MP_03	R_MP_04	R_MP_05	R_MP_06	R_MP_07	R_MP_08	R_MP_09	R_MP_10	R_MP_11	R_MP_12	R_MP_13	R_MP_14	R_MP_15	R_MP_16	R_MP_17
Tokens	5.242	15	38	153	79	93	44	44	59	68	61	22	21	80	72	105	22	40
Types	599	15	35	103	64	72	27	39	45	59	51	19	19	68	55	75	22	34
Type/Token Ratio	11,43	100	92,11	67,32	81,01	77,42	61,36	88,64	76,27	86,76	83,61	86,36	90,48	85	76,39	71,43	100	85
Sentences	286	1	2	7	4	6	2	2	3	2	3	2	2	4	4	4	3	3

DE

Text File	OVERALL	R_MP_01	R_MP_02	R_MP_03	R_MP_04	R_MP_05	R_MP_06	R_MP_07	R_MP_08	R_MP_09	R_MP_10	R_MP_11	R_MP_12	R_MP_13	R_MP_14	R_MP_15	R_MP_16	R_MP_17
Tokens	5.075	13	36	150	77	92	41	43	56	65	60	20	20	76	70	101	21	39
Types	589	13	33	102	62	71	25	38	42	57	50	17	18	64	53	72	21	33

DT

Text File	R_MP_18	R_MP_19	R_MP_20	R_MP_21	R_MP_22	R_MP_23	R_MP_24	R_MP_25	R_MP_26	R_MP_27	R_MP_28	R_MP_29	R_MP_30	R_MP_31	R_MP_32	R_MP_33	R_MP_34	R_MP_35
Tokens	65	65	104	64	42	102	75	169	258	23	123	121	123	308	333	103	154	99
Types	57	50	77	50	38	68	61	111	165	21	83	78	89	197	196	78	106	76
Type/Token Ratio	87,69	76,92	74,04	78,13	90,48	66,67	81,33	65,68	63,95	91,3	67,48	64,46	72,36	63,96	58,86	75,73	68,83	76,77
Sentences	3	4	4	4	4	7	6	8	14	2	7	3	5	18	17	7	9	7

DE

Text File	R_MP_18	R_MP_19	R_MP_20	R_MP_21	R_MP_22	R_MP_23	R_MP_24	R_MP_25	R_MP_26	R_MP_27	R_MP_28	R_MP_29	R_MP_30	R_MP_31	R_MP_32	R_MP_33	R_MP_34	R_MP_35
Tokens	64	64	94	62	41	101	71	166	257	20	120	120	120	304	330	98	152	97
Types	56	49	70	48	37	67	58	110	165	19	81	78	86	194	195	74	104	75

DT

Text File	R_MP_36	R_MP_37	R_MP_38	R_MP_39	R_MP_40	R_MP_41	R_MP_42	R_MP_43	R_MP_44	R_MP_45	R_MP_46	R_MP_47	R_MP_48	R_MP_49	R_MP_50	R_MP_51	R_MP_52	R_MP_53
Tokens	50	15	86	97	49	140	313	62	19	109	85	91	75	92	77	39	63	45
Types	41	14	65	68	36	100	180	48	18	79	68	73	60	63	63	33	51	40
Type/Token Ratio	82	93,33	75,58	70,1	73,47	71,43	57,51	77,42	94,74	72,48	80	80,22	80	68,48	81,82	84,62	80,95	88,89
Sentences	4	1	5	3	2	10	13	5	1	3	5	4	5	5	4	2	5	3

DE

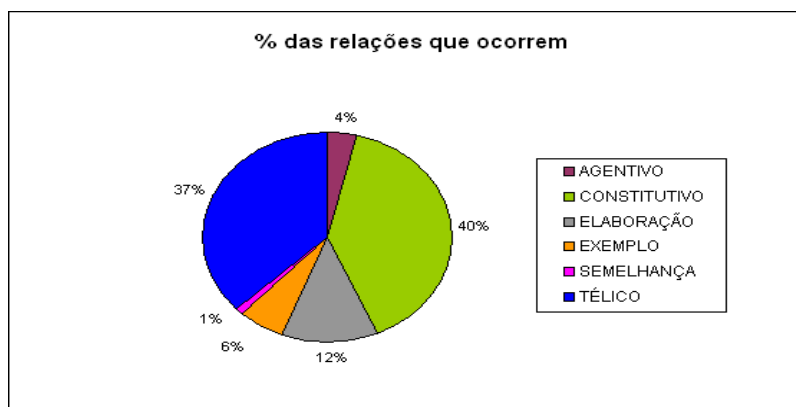
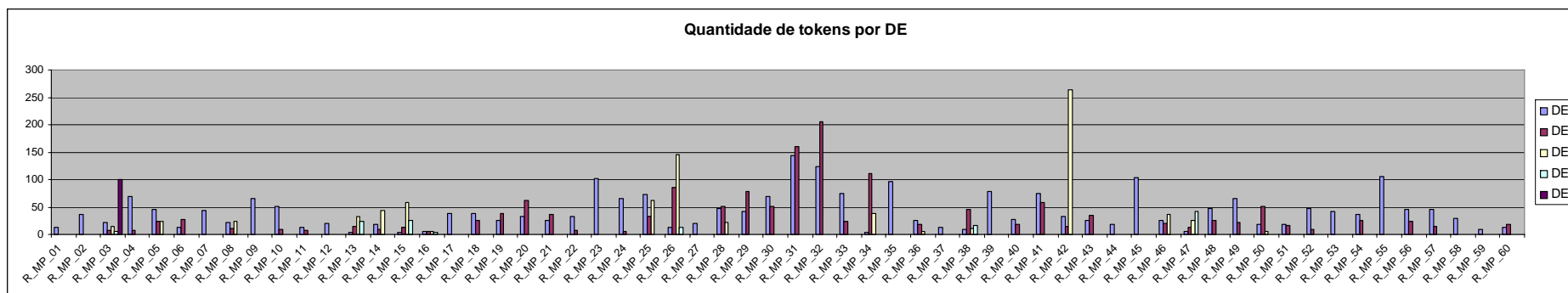
Text File	R_MP_36	R_MP_37	R_MP_38	R_MP_39	R_MP_40	R_MP_41	R_MP_42	R_MP_43	R_MP_44	R_MP_45	R_MP_46	R_MP_47	R_MP_48	R_MP_49	R_MP_50	R_MP_51	R_MP_52	R_MP_53
Tokens	49	13	82	96	45	134	310	59	18	103	83	85	73	88	75	36	57	41
Types	41	12	61	67	34	97	178	46	17	75	66	68	58	62	61	30	45	36

DT

Text File	R_MP_54	R_MP_55	R_MP_56	R_MP_57	R_MP_58	R_MP_59	R_MP_60
Tokens	66	110	70	60	30	18	34
Types	56	80	54	46	29	17	28
Type/Token Ratio	84,85	72,73	77,14	76,67	96,67	94,44	82,35
Sentences	4	9	3	3	1	1	2

DE

Text File	R_MP_54	R_MP_55	R_MP_56	R_MP_57	R_MP_58	R_MP_59	R_MP_60
Tokens	62	106	69	59	29	10	32
Types	54	78	53	45	28	10	26



	DE1	DE2	DE3	DE4	DE5
R_MP_01	<télico>				
R_MP_02	<télico>				
R_MP_03	<constitutivo>	<agentivo>	<exemplo>	<constitutivo>	<télico>
R_MP_04	<constitutivo>	<télico>			
R_MP_05	<constitutivo>	<télico>	<exemplo>		
R_MP_06	<constitutivo>	<exemplo>			
R_MP_07	<constitutivo>				
R_MP_08	<constitutivo>	<télico>	<exemplo>		
R_MP_09	<constitutivo>				
R_MP_10	<constitutivo>	<télico>			
R_MP_11	<constitutivo>	<télico>			
R_MP_12	<constitutivo>				
R_MP_13	<elaboração>	<exemplo>	<elaboração>	<télico>	
R_MP_14	<constitutivo>	<exemplo>	<télico>		
R_MP_15	<constitutivo>	<télico>	<elaboração>	<constitutivo>	
R_MP_16	<elaboração>	<constitutivo>	<télico>	<constitutivo>	
R_MP_17	<constitutivo>				
R_MP_18	<constitutivo>	<télico>			
R_MP_19	<constitutivo>	<télico>			
R_MP_20	<constitutivo>	<télico>			
R_MP_21	<agentivo>	<constitutivo>			
R_MP_22	<constitutivo>	<télico>			
R_MP_23	<constitutivo>				
R_MP_24	<constitutivo>	<télico>			
R_MP_25	<constitutivo>	<elaboração>	<télico>		
R_MP_26	<elaboração>	<constitutivo>	<télico>	<elaboração>	
R_MP_27	<constitutivo>				

R_MP_28	<constitutivo>	<télico>	<elaboração>		
R_MP_29	<constitutivo>	<télico>			
R_MP_30	<constitutivo>	<télico>			
R_MP_31	<constitutivo>	<télico>			
R_MP_32	<constitutivo>	<télico>			
R_MP_33	<constitutivo>	<télico>			
R_MP_34	<agentivo>	<constitutivo>	<télico>		
R_MP_35	<constitutivo>				
R_MP_36	<constitutivo>	<télico>	<semelhança>		
R_MP_37	<télico>				
R_MP_38	<agentivo>	<télico>	<elaboração>	<télico>	
R_MP_39	<constitutivo>	<télico>			
R_MP_40	<constitutivo>	<elaboração>			
R_MP_41	<constitutivo>	<télico>			
R_MP_42	<télico>	<exemplo>	<télico>		
R_MP_43	<constitutivo>	<elaboração>			
R_MP_44	<télico>				
R_MP_45	<télico>				
R_MP_46	<télico>	<constitutivo>	<elaboração>		
R_MP_47	<constitutivo>	<télico>	<constitutivo>	<elaboração>	
R_MP_48	<constitutivo>	<télico>			
R_MP_49	<constitutivo>	<télico>			
R_MP_50	<agentivo>	<elaboração>	<télico>		
R_MP_51	<constitutivo>	<télico>			
R_MP_52	<constitutivo>	<télico>			
R_MP_53	<télico>				
R_MP_54	<constitutivo>	<télico>			
R_MP_55	<constitutivo>				

	DE1	DE2	DE3	DE4	DE5
R_MP_56	<télico>	<elaboração>			
R_MP_57	<télico>	<elaboração>			
R_MP_58	<constitutivo>				

R_MP_59	<TÉLICO>				
R_MP_60	<CONSTITUTIVO>	<EXEMPLO>			

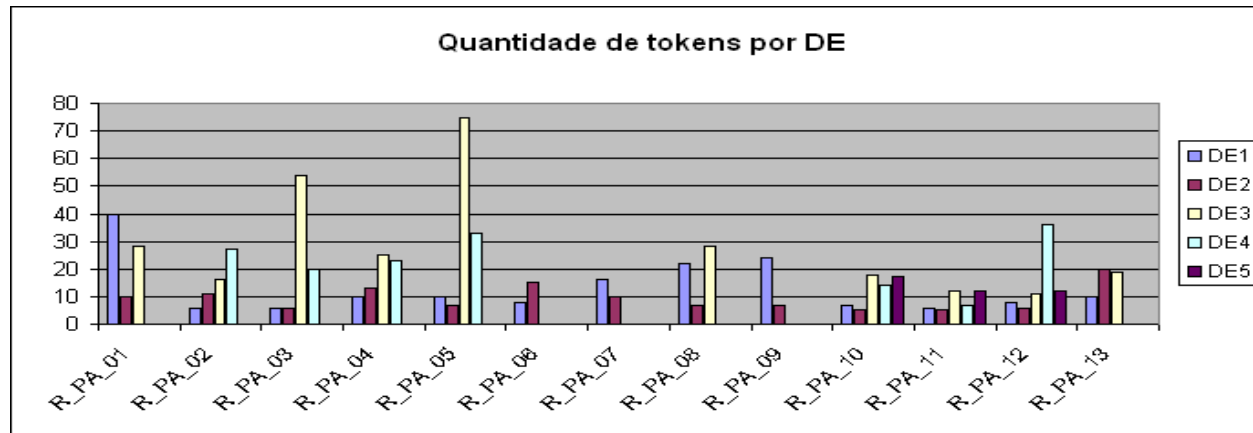
campo PRODUTO ACABADO (PA)

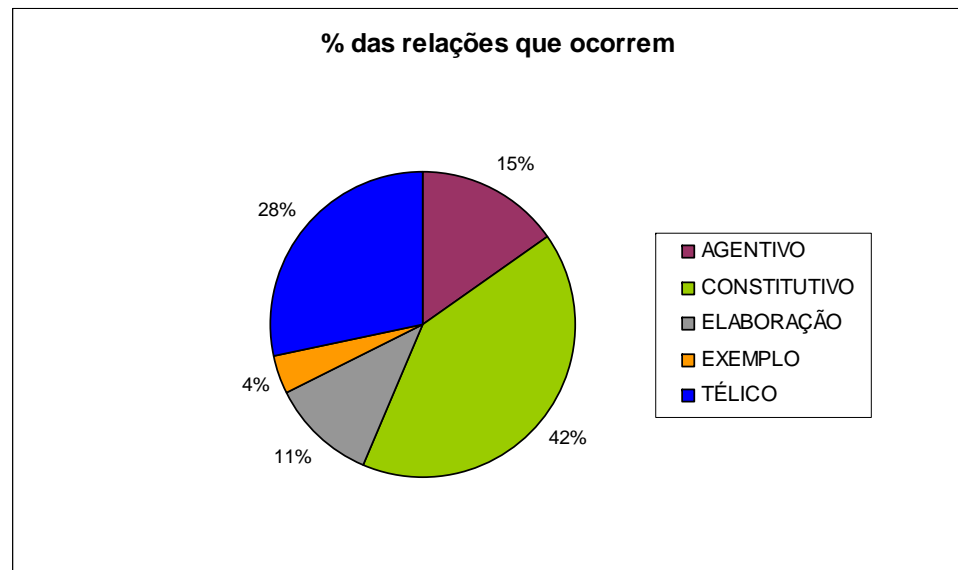
DT

Text File	OVERALL	R_PA_01	R_PA_02	R_PA_03	R_PA_04	R_PA_05	R_PA_06	R_PA_07	R_PA_08	R_PA_09	R_PA_10	R_PA_11	R_PA_12	R_PA_13
Tokens	837	80	71	89	73	127	25	28	67	34	65	47	75	56
Types	109	60	56	65	56	90	19	23	44	20	46	33	63	43
Type/Token Ratio	13,02	75	78,87	73,03	76,71	70,87	76	82,14	65,67	58,82	70,77	70,21	84	76,79
Sentences	36	2	3	3	3	4	2	2	2	1	4	3	5	2

DE

Text File	OVERALL	R_PA_01	R_PA_02	R_PA_03	R_PA_04	R_PA_05	R_PA_06	R_PA_07	R_PA_08	R_PA_09	R_PA_10	R_PA_11	R_PA_12	R_PA_13
Tokens	782	78	60	86	71	125	23	26	57	31	61	42	73	49
Types	101	58	49	62	55	88	17	21	39	17	42	29	63	40





	DE1	DE2	DE3	DE4	DE5
R_PA_01	<constitutivo>	<télico>	<elaboração>		
R_PA_02	<constitutivo>	<constitutivo>	<exemplo>	<télico>	
R_PA_03	<constitutivo>	<télico>	<constitutivo>	<agentivo>	
R_PA_04	<constitutivo>	<télico>	<elaboração>	<télico>	
R_PA_05	<exemplo>	<télico>	<constitutivo>	<agentivo>	
R_PA_06	<télico>	<constitutivo>			
R_PA_07	<constitutivo>	<télico>			
R_PA_08	<agentivo>	<télico>	<constitutivo>		
R_PA_09	<constitutivo>	<télico>			
R_PA_10	<constitutivo>	<agentivo>	<constitutivo>	<télico>	<constitutivo>
R_PA_11	<constitutivo>	<agentivo>	<constitutivo>	<télico>	<constitutivo>
R_PA_12	<constitutivo>	<télico>	<agentivo>	<elaboração>	<elaboração>
R_PA_13	<agentivo>	<constitutivo>	<elaboração>		

campo PROPRIEDADE (PR)

DT

Text File	OVERALL	R_PR_01	R_PR_02	R_PR_03	R_PR_04	R_PR_05	R_PR_06	R_PR_07	R_PR_08	R_PR_09	R_PR_10	R_PR_11	R_PR_12	R_PR_13	R_PR_14	R_PR_15	R_PR_16	R_PR_17
Tokens	1.365	38	44	28	31	26	31	25	19	17	62	76	40	27	25	27	32	66
Types	168	29	36	22	26	26	28	23	19	16	45	59	37	25	22	26	31	45
Type/Token Ratio	12,31	76,32	81,82	78,57	83,87	100	90,32	92	100	94,12	72,58	77,63	92,5	92,59	88	96,3	96,88	68,18
Sentences	63	2	3	2	2	1	2	1	1	1	2	2	1	2	1	1	2	2

DE

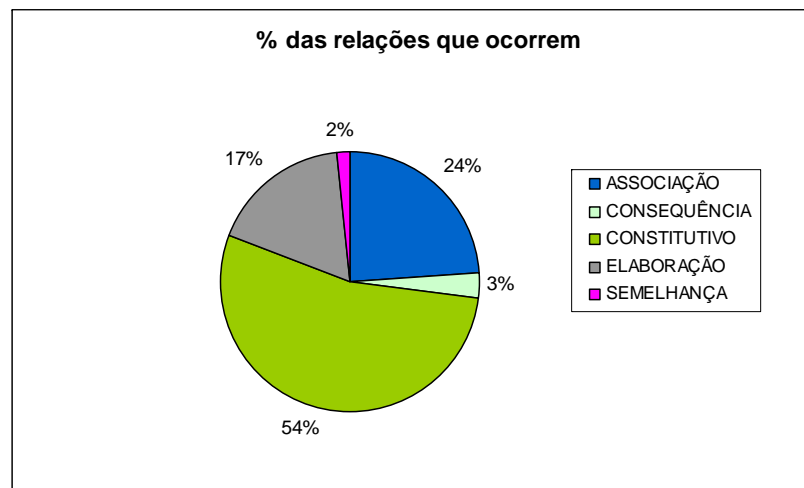
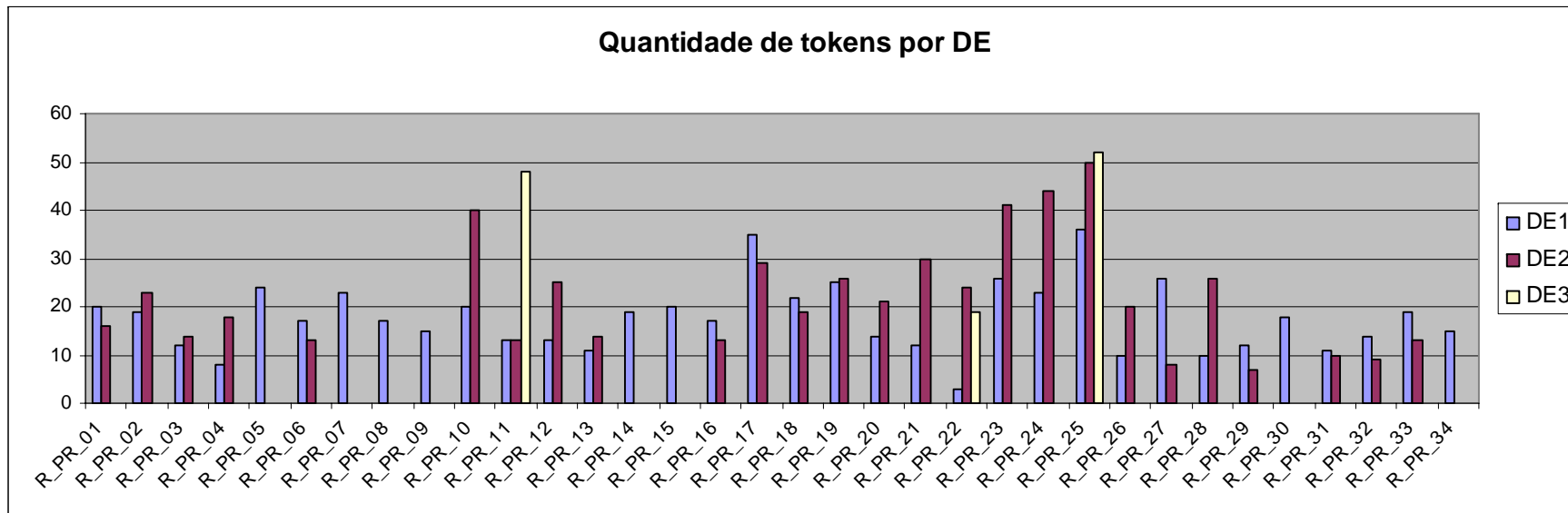
Text File	OVERALL	R_PR_01	R_PR_02	R_PR_03	R_PR_04	R_PR_05	R_PR_06	R_PR_07	R_PR_08	R_PR_09	R_PR_10	R_PR_11	R_PR_12	R_PR_13	R_PR_14	R_PR_15	R_PR_16	R_PR_17
Tokens	1.284	36	42	26	26	24	30	23	17	15	60	74	38	25	19	20	30	64
Types	161	28	34	20	24	24	27	21	17	14	44	57	35	24	17	19	29	43

DT

Text File	R_PR_18	R_PR_19	R_PR_20	R_PR_21	R_PR_22	R_PR_23	R_PR_24	R_PR_25	R_PR_26	R_PR_27	R_PR_28	R_PR_29	R_PR_30	R_PR_31	R_PR_32	R_PR_33	R_PR_34
Tokens	43	52	38	45	52	68	68	142	31	35	37	22	19	23	25	34	17
Types	36	40	36	34	40	55	51	96	27	32	27	20	18	18	25	29	14
Type/Token Ratio	83,72	76,92	94,74	75,56	76,92	80,88	75	67,61	87,1	91,43	72,97	90,91	94,74	78,26	100	85,29	82,35
Sentences	2	2	2	3	3	2	2	3	2	2	2	2	1	2	2	2	1

DE

Text File	R_PR_18	R_PR_19	R_PR_20	R_PR_21	R_PR_22	R_PR_23	R_PR_24	R_PR_25	R_PR_26	R_PR_27	R_PR_28	R_PR_29	R_PR_30	R_PR_31	R_PR_32	R_PR_33	R_PR_34
Tokens	41	51	35	42	46	67	67	138	30	34	36	19	18	21	23	32	15
Types	34	39	33	33	37	54	50	94	26	31	27	18	17	16	23	27	12



	DE1	DE2	DE3
R_PR_01	<constitutivo>	<elaboração>	
R_PR_02	<constitutivo>	<associação>	
R_PR_03	<constitutivo>	<associação>	
R_PR_04	<constitutivo>	<associação>	
R_PR_05	<constitutivo>		
R_PR_06	<constitutivo>	<associação>	
R_PR_07	<constitutivo>		
R_PR_08	<constitutivo>		
R_PR_09	<constitutivo>		
R_PR_10	<constitutivo>	<associação>	
R_PR_11	<constitutivo>	<associação>	<elaboração>
R_PR_12	<constitutivo>	<elaboração>	
R_PR_13	<constitutivo>	<elaboração>	
R_PR_14	<constitutivo>		
R_PR_15	<constitutivo>		
R_PR_16	<constitutivo>	<elaboração>	
R_PR_17	<constitutivo>	<associação>	
R_PR_18	<constitutivo>	<associação>	
R_PR_19	<constitutivo>	<associação>	
R_PR_20	<constitutivo>	<elaboração>	
R_PR_21	<constitutivo>	<consequência>	
R_PR_22	<constitutivo>	<consequência>	<associação>
R_PR_23	<constitutivo>	<associação>	
R_PR_24	<constitutivo>	<associação>	
R_PR_25	<constitutivo>	<elaboração>	<associação>
R_PR_26	<constitutivo>	<associação>	
R_PR_27	<constitutivo>	<elaboração>	
R_PR_28	<constitutivo>	<associação>	
R_PR_29	<constitutivo>	<semelhança>	
R_PR_30	<constitutivo>		
R_PR_31	<constitutivo>	<elaboração>	
R_PR_32	<constitutivo>	<elaboração>	
R_PR_33	<constitutivo>	<elaboração>	
R_PR_34	<constitutivo>		