

**UNIVERSIDADE DE BRASÍLIA**  
**FACULDADE DE ECONOMIA, ADMINISTRAÇÃO, CONTABILIDADE**  
**E CIÊNCIA DA INFORMAÇÃO E DOCUMENTAÇÃO**  
**DEPARTAMENTO DE CIÊNCIA DA INFORMAÇÃO E DOCUMENTAÇÃO**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO**

**UM ARCABOUÇO TEÓRICO PARA AUTORIA DE DOCUMENTOS VISANDO**  
**ATENUAR O SURGIMENTO DO FENÔMENO DA AMBIGÜIDADE**

**Daniel Parente Lemos dos Santos**

**Orientador: Prof. Dr. Mamede Lima-Marques**

**BRASÍLIA**

**2006**

**DANIEL PARENTE LEMOS DOS SANTOS**

**UM ARCABOUÇO TEÓRICO PARA AUTORIA DE DOCUMENTOS VISANDO  
ATENUAR O SURGIMENTO DO FENÔMENO DA AMBIGÜIDADE**

Dissertação apresentada ao Departamento de Ciência da Informação e Documentação da Universidade de Brasília, como parte dos requisitos para obtenção do título de mestre em Ciência da Informação.

**Orientador: Prof. Dr. Mamede Lima-Marques**

**BRASÍLIA**

**2006**

## FOLHA DE APROVAÇÃO

**Título:** Um arcabouço teórico para autoria de documentos visando atenuar o surgimento do fenômeno da ambigüidade

**Área de Concentração:** Transferência de Informação

**Linha de Pesquisa:** Arquitetura da Informação

Dissertação submetida à Comissão Examinadora designada pelo Colegiado do Programa de Pós-Graduação em Ciência da Informação do Departamento de Ciência da Informação e Documentação da Universidade de Brasília como requisito parcial para a obtenção do título de **Mestre em Ciência da Informação**.

Dissertação aprovada em:

Aprovado por:

---

Prof. Dr. Mamede Lima–Marques

Orientador (UnB/PPGCINF)

---

Profa. Dra. Marisa Bräscher

Membro Interno – (UnB/PPGCINF)

---

Profa. Dra. Lígia Café

Membro Externo – (UFSC)

**À minha querida vovó Didi (1924-2006),  
exemplo de ternura e amor.**

## *Agradecimentos*

Ao Prof. Dr. Mamede Lima-Marques, aquele que, com sabedoria, humildade e simpatia, inspira a todos que o cercam.

Aos colegas do departamento de Ciência da Informação e Documentação da Unb, especialmente aqueles participantes do grupo de pesquisas em Arquitetura da Informação, pelas conversas construtivas, pelas dicas e pela motivação.

A meus pais, pelo amor e pelos princípios passados durante minha criação. Pelo apoio e motivação que sempre me deram em todos os momentos. Por me mostrarem a importância e a nobreza da vida acadêmica.

À minha amada esposa Janaína, pelo amor, paciência, compreensão e suporte dados durante estes anos. Sua presença e seu carinho foram indispensáveis para o sucesso desta empreitada.

Ao meu amado filho João Pedro. Seu sorriso é o combustível que me move. Do alto de seus dois anos de idade, ministrou-me as aulas de lingüística mais impressionantes que já tive, o que, de fato, ajudou-me a compreender alguns dos tópicos deste trabalho.

Às minhas irmãs, avós, tios, primos, amigos e colegas de trabalho. Em cada linha desta dissertação também tem o dedo de vocês.

E, por fim, a Deus... pela beleza da vida.

## *Resumo*

As ambigüidades e seus efeitos na recuperação e interpretação da informação constituem-se no problema tratado por esta dissertação. Este trabalho aborda o problema no momento da autoria dos documentos, ao invés da abordagem convencional dos sistemas de Processamento Automático de Linguagem Natural, que lidam com as ambigüidades *a posteriori*. Parte-se do pressuposto que o autor do documento sabe exatamente o que quer comunicar e que, portanto, o mesmo está apto a auxiliar o receptor no processo de comunicação, sanando ou minimizando a ocorrência de ambigüidades na recuperação e interpretação da informação. Assim, o principal objetivo do trabalho é criar um arcabouço teórico que permita alçar o autor do documento a uma posição mais participativa no processo de comunicação, auxiliando o leitor através de informações de cunho morfológico, sintático e semântico sobre o conteúdo do documento. Para tal, a obra de Bräscher relativa ao tratamento automático de ambigüidades na recuperação da informação, bem como os ambientes de autoria baseados em ontologias, de Oliveira, surgem como principais ferramentas teóricas para viabilizar o arcabouço proposto. Ao final, constata-se que o arcabouço teórico mostra-se eficaz para atenuar o surgimento de ambigüidades na recuperação e interpretação da informação, principalmente no campo léxico-morfológico, e também nas esferas sintática e predicativa. Entretanto, percebeu-se que o arcabouço mostrou-se ineficaz com relação às ambigüidades de cunho pragmático e semântico.

### **Palavras-chave**

Ambigüidade, ontologia, gramática de valências, ambiente de autoria.

## *Abstract*

The ambiguities and its effects on information retrieve and interpretation are the problem treated by this dissertation. This essay approaches the problem at the moment of the authorship of documents, instead of the conventional approach of Natural Language Automatic Processing Systems, which deals with ambiguities at a moment after-authorship. The author of a document knows exactly what wants to communicate and, therefore, he is able to assist the reader in the communication process, attenuating the occurrence of ambiguities in the information recovery and interpretation. Thus, the main objective of this work is to create a framework that raises the author of a document to a more participative position in the communication process, helping the reader with morphologic, syntactic and semantic information about the document. To achieve it, the thesis of Bräscher, which treats the problem of ambiguities automatic processing in information retrieve, and the thesis of Oliveira, which considers an authoring environment based on ontologies, appear as the main methodological tools to make possible the framework. Finally, the conclusion is that the framework is effective to solve morphologic, and lexical ambiguities. The theory is also promissory to attenuate syntactic e predicative ambiguities. However, the framework did not solve the semantic and pragmatic ambiguities.

### **Keywords**

Ambiguity, ontology, valence grammar, authoring environment.

## *Lista de Ilustrações*

1	Classificação do fenômeno da Homonímia .....	p. 17
2	Compilador .....	p. 32
3	Árvore Gramatical .....	p. 33
4	Fases de um compilador .....	p. 34
5	Árvore Sintática .....	p. 36
6	Restrições Semântico-Selecionais e a Gramática de Borba .....	p. 48
7	Modelo de Comunicação de Shannon & Weaver .....	p. 61
8	Modelo de Comunicação de Tubbs & Moss .....	p. 61
9	Exemplo de um gráfico conceitual .....	p. 65
10	Camadas da Web Semântica .....	p. 71
11	Ambiente de Autoria na Web baseado em ontologia .....	p. 75
12	Tela do Ambiente de Autoria para a Web Semântica .....	p. 76
13	XML e DTD .....	p. 79
14	Tripla RDF .....	p. 81
15	Componentes de uma ontologia OWL .....	p. 84
16	Exemplo de uma ontologia OWL .....	p. 89
17	Tela inicial do plugin OWL do Protégé .....	p. 91
18	Interface gráfica do software fictício Valente .....	p. 94
19	Interface gráfica da ferramenta de busca fictícia Sherlock .....	p. 96
20	Arcabouço teórico para a Autoria de Documentos Isentos de Ambigüidades .....	p. 98
21	Encaixe dos níveis do arcabouço teórico no modelo de Lima-Marques & Oliveira .....	p. 99
22	Taxonomia das classes da ontologia de valências .....	p. 101
23	Editando uma forma canônica no Protégé .....	p.114
24	Várias formas em linguagem natural para uma única forma em MLV .....	p.118

## *Lista de Quadros*

1	Tipos de ambigüidades da Língua Portuguesa .....	p. 25
2	Hierarquia de Gramáticas de Chomsky .....	p. 31
3	Casos ou Papéis Temáticos .....	p. 44
4	Frame de uma das formas canônicas associadas à forma gráfica “comer” .....	p.119

## *Lista de Abreviaturas e Siglas*

Adj: Adjetivo  
Adv: Advérbio  
Ag : Agentivo  
Anim: Animado  
Ativ: Atividade  
B: Beneficiário  
C: Complemento  
Ca: Causativo  
Co: Comitativo  
Cont: Contável  
Det: Determinante  
Dr: Direcional  
Ex: Experimentador  
Hum: Humano  
I: Instrumental  
L: Locativo  
M: Meta  
MLV: Meta-Linguagem de Valências  
N: Nome  
Num: Numeral  
O: Oração  
Ob: Objetivo  
Or: Origem  
P: Predicado  
Pr: Percurso  
Prep: Preposição  
R: Resultativo  
SN: Sintagma Nominal  
SO: Sintagma Oracional  
Sprep: Sintagma Preposicional  
SV: Sintagma Verbal  
Tp: Temporal  
V: Verbo  
Vaux: Verbo Auxiliar  
W3C: World Wide Web Consortium

# Sumário

<b>1 – Introdução</b>	<b>3</b>
<b>2 – Requisitos Pré-pesquisa</b>	<b>6</b>
<b>2.1 – Objetivos</b>	<b>6</b>
2.1.1 - Objetivo Geral	6
2.1.2 - Objetivos Específicos	6
<b>2.2 - Justificativa</b>	<b>6</b>
<b>2.3 - Metodologia</b>	<b>9</b>
2.3.1 - Método de Pesquisa Utilizado	9
2.3.2 – Passos Metodológicos Adotados	10
<b>3 - Revisão da Literatura e Fundamentos</b>	<b>13</b>
<b>3.1 – Ambigüidades como Problema</b>	<b>13</b>
3.1.1 – Alguns Fatores Causadores de Ambigüidades	15
3.1.2 – Outros fenômenos lingüísticos relacionados à Ambigüidade	20
3.1.3 - Tipos de Ambigüidades	22
<b>3.2 – Noções da Teoria Gramatical</b>	<b>28</b>
3.2.1 – Princípios de Compiladores	31
<b>3.3 – Conceitos Gramaticais do Português Contemporâneo do Brasil</b>	<b>35</b>
<b>3.4 - Gramática de Valências para o Português</b>	<b>38</b>
3.4.1 – Teoria da Valência	42
3.4.2 – Gramática de Casos	44
3.4.3 – Restrições Semântico-Selecionais	46
3.4.3– Valência Verbal	48
3.4.4– Valência Nominal	53
3.4.5– Valência Adjetival	57
3.4.6 – Valência Adverbial	60

<b>3.5 – Teoria da Comunicação</b>	<b>61</b>
<b>3.6 - Ontologias</b>	<b>63</b>
<b>3.7 - Tratamento automático de ambigüidades na recuperação da informação</b>	<b>65</b>
<b>3.8 – A Web Semântica e os Ambientes de Autoria baseados em ontologias</b>	<b>71</b>
<b>3.9 – Fundamentação Teórica das Ferramentas</b>	<b>77</b>
3.9.1 – XML (Extensible Markup Language)	77
3.9.2 – RDF (Resource Description Framework)	80
3.9.3 – OWL (Ontology Web Language)	82
3.9.4 – Protégé	90
<b>4 – Resultados da Pesquisa</b>	<b>93</b>
<b>4.1 – Proposta de Arcabouço Teórico</b>	<b>93</b>
<b>4.2 - Ontologia das valências dos verbos, nomes, adjetivos e advérbios</b>	<b>100</b>
4.2.1 – Classes da ontologia	101
4.2.2 – Propriedades da ontologia de Valências	103
<b>4.3 - Gramática da Meta-linguagem de Valências</b>	<b>115</b>
4.3.1 – Demonstração da geração de uma sentença a partir das regras de produção da MLV	126
<b>4.4 - Desambiguação Aplicando o Referencial Proposto</b>	<b>128</b>
4.4.1 - Ambigüidade Lexical	128
4.4.2 - Ambigüidade Morfológica	132
4.4.3 - Ambigüidade Sintática	134
4.4.4 - Ambigüidade Predicativa	136
4.4.5 – Ambigüidade Semântica	138
4.4.6 - Ambigüidade Pragmática	142
<b>5 - Conclusão</b>	<b>144</b>
<b>Referências</b>	<b>150</b>
<b>ANEXO 1</b>	<b>155</b>

# 1 – Introdução

As ambigüidades e seus efeitos na recuperação e interpretação da informação constituem-se no problema a ser tratado por esta dissertação.

Na Antigüidade clássica, Aristóteles já tratava a questão das ambigüidades. Segundo Faria (1995), Aristóteles afirmava que os Sofistas, ao desenvolverem ao máximo a retórica e a dialética, não assumiam qualquer compromisso com a verdade, seus argumentos visavam antes de tudo confundir o adversário e seduzir o auditório, utilizando e explorando a natural ambigüidade da palavra.

Nos dias atuais, a questão das ambigüidades continua sendo um problema abordado de maneira recorrente por diversas áreas do conhecimento, tais como a Filosofia, a Lógica, a Lingüística, a Ciência da Informação e a Ciência da Computação. Atualmente, o problema também é tratado dentro do escopo do Processamento de Linguagens Naturais, área interdisciplinar que reúne teorias e técnicas de diversas vertentes científicas, principalmente as supracitadas, visando tratar de maneira automática textos em linguagem natural.

Este trabalho pretende abordar o problema no momento da autoria dos documentos, ao invés da abordagem convencional dos sistemas de Processamento Automático de Linguagem Natural, que lidam com as ambigüidades *a posteriori*<sup>1</sup>, ou seja, no momento da Recuperação da Informação, utilizando-se de técnicas complexas de Processamento de Linguagem Natural. Bräscher (2002) discorre sobre a complexidade e as limitações de tais sistemas:

O processo de desambiguação<sup>2</sup> automática é mais complexo que o de solução de ambigüidades realizado por um receptor humano. O recurso ao contexto em sistemas de recuperação de informação em linguagem natural é restrito, uma vez que o contexto constitui-se no conjunto de conhecimentos que o sistema possui num determinado momento da análise. Nem todo tipo de informação contextual pode

---

<sup>1</sup> Ver Hobbs (1996); Martins *et al* (1998a); Bräscher (1999).

<sup>2</sup> Desambiguar consiste no ato de remover, ou resolver, uma ambigüidade. Segundo Bräscher (2002) este termo é um neologismo sugerido pela Prof. Dra. Enilde Faulstich.

ser representado formalmente e, portanto, nem todo tipo de ambigüidade pode ser resolvido nesses sistemas.

“Um documento é todo artefato que representa ou expressa um objeto, uma idéia ou uma informação por meio de signos gráficos e icônicos (palavras, imagens, diagramas, mapas, figuras, símbolos), sonoros e visuais (gravados em suporte de papel ou eletrônico).” (Le COADIC, 1996). Dentro do escopo desta dissertação, entretanto, far-se-á referência ao termo “documento” como artefato que representa ou expressa um objeto, uma idéia ou uma informação por meio de palavras, redigidas em linguagem natural, mais especificamente na variação brasileira da língua portuguesa, gravadas em suporte de papel ou em meio eletrônico.

Em determinadas ocasiões, a ambigüidade é causada intencionalmente pelo falante ou autor do documento (MONNERAT, 2004; ZAVAGLIA, 2003). Isto ocorre geralmente em textos publicitários, literários, poéticos, provérbios, expressões idiomáticas e na música. Nas demais ocasiões, o fenômeno não é intencional, sendo, de fato, considerado como um obstáculo à comunicação. Para esses casos, Zavaglia (2003) afirma que “o fenômeno da ambigüidade existe somente para o receptor da mensagem, ao passo que para seu emissor a mensagem não é de forma alguma ambígua”. O emissor da mensagem tem, portanto, a exata consciência do significado do conteúdo da mensagem que está emitindo, enquanto que o receptor está sujeito a não ser capaz de determinar o significado pretendido pelo emissor.

É a partir deste pressuposto, ou seja, que o autor do documento sabe exatamente o que quer comunicar, que se constrói a abordagem de resolução do problema neste trabalho. Acredita-se que o autor do documento está apto a auxiliar o receptor, provendo-o com informações morfológicas, sintáticas e semânticas sobre o do conteúdo da mensagem, sanando ou minimizando a ocorrência de ambigüidades na recuperação e interpretação da informação.

Esta estratégia pode ser entendida ao se situar a mesma dentro do esquema que representa o ciclo de vida da informação, tal qual proposto em Le Coadic (1996). Segundo tal esquema, o ciclo de vida da informação se dá em três passos: Construção, Comunicação e Uso. A abordagem proposta nesta dissertação atua no âmbito da Construção do conhecimento, enquanto que as abordagens tradicionais de desambiguação atuam no escopo do uso, ou em um momento imediatamente anterior ao uso.

Para a concepção desta nova estratégia de desambiguação, a tese de Bräscher (1999) e a tese de Oliveira (2006), sobre ambientes de autoria baseados em ontologias, foram as principais fontes teóricas inspiradoras. Apesar da solução de Bräscher (1999) tratar da ambigüidade no momento da recuperação da informação, abordagem oposta à proposta aqui, a autora introduz a utilização de ferramentas como os gráficos conceituais, as ontologias, e a gramática de valências de Borba, para a desambiguação de sentenças em língua portuguesa. Tais ferramentas, especificamente as ontologias e a gramática de valências de Borba, tornam-se os componentes principais do arcabouço teórico proposto nesta dissertação.

## ***2 – Requisitos Pré-pesquisa***

### **2.1 – Objetivos**

#### **2.1.1 - Objetivo Geral**

Criar um arcabouço teórico que permita, no momento da autoria, o registro de informações morfológicas, sintáticas e semânticas sobre o conteúdo do documento, visando atenuar o surgimento de ambigüidades na recuperação e interpretação da informação.

#### **2.1.2 - Objetivos Específicos**

- a) Definir e descrever os componentes necessários à criação do arcabouço teórico.
- b) Definir a estrutura conceitual necessária a uma ontologia das valências dos verbos, nomes, adjetivos e advérbios da variação brasileira da língua portuguesa.
- c) Propor uma gramática para a meta-linguagem de valências, componente do arcabouço teórico proposto.

### **2.2 - Justificativa**

Para a Ciência da Informação, a solução da ambigüidade é extremamente relevante em função da necessidade de se obter maior relevância<sup>3</sup> na recuperação da informação. Bräscher (2002) afirma que a ambigüidade causa ruído na recuperação da informação, pois, sob um mesmo termo, o usuário (ou receptor) encontrará informação relevante e irrelevante.

---

<sup>3</sup> Entende-se por relevância a capacidade de um motor ou de uma função de busca de recuperar dados e informações apropriados para as necessidades dos usuários (ROBREDO, 2003).

Bräscher (1999) efetua algumas recomendações com relação ao desenvolvimento e aperfeiçoamento de ferramentas utilizadas para o tratamento e para recuperação de conteúdo.

São elas:

- a) Desenvolvimento de ferramentas automáticas de tratamento de conteúdo que possibilitem maior precisão na recuperação da informação (...);
- b) Preocupação com a qualidade do tratamento da informação, pois a qualidade da recuperação depende deste tratamento. (...);
- c) Desenvolvimento de metodologias e ferramentas de tratamento da informação que levem em conta os resultados obtidos em pesquisas da área da Ciência da Informação. (...);
- d) Desenvolvimento de pesquisas multidisciplinares, em que se apliquem necessidades e métodos desenvolvidos no âmbito de outras disciplinas às necessidades práticas da Ciência da Informação;
- e) Aplicações dos recursos tecnológicos na melhoria das ferramentas de tratamento e recuperação de informação, de maneira a fornecer, aos usuários, interfaces em linguagem natural que garantam a qualidade da recuperação.

Em linhas gerais, esta dissertação pretende contribuir para o atendimento de todas as recomendações tecidas por Bräscher (1999).

Em alguns tipos de documentos, tais como trabalhos científicos, especificações de regras de negócios para construção de sistemas de informação, especificações de requisitos para projetos de arquitetura, engenharia e construção, e, principalmente, em leis e contratos, lidar com a presença de ambigüidades é um dos principais problemas e desafios para a interpretação do significado contido na peça documental. Pode-se imaginar as conseqüências negativas de um requisito ambíguo em um projeto de construção de uma usina nuclear, bem como são notórias as infundáveis discussões judiciais em função de ambigüidades presentes em leis e contratos. É para estes tipos de documentos, ou seja, aqueles que não admitem a presença de ambigüidades, que o conteúdo desta dissertação se dirige.

Os resultados deste trabalho também terão valia para o dia-a-dia das organizações e da sociedade. Exemplificando, MacComb (2004) faz uma análise interessante sobre advogados e contratos, sugerindo que a complexidade semântica dos contratos ocorre propositamente para sustentar um filão bilionário da economia americana. Segundo o autor, contratos não são só complexos para um computador interpretá-los, são também complexos para juristas interpretá-los. Existem 880.000 advogados nos Estados Unidos da América, representando

uma indústria de, pelo menos, US\$ 100 bilhões de dólares. Uma das coisas mais lucrativas que eles fazem é redigir, interpretar e litigar contratos. Se os contratos fossem fáceis para ser redigidos e fáceis para interpretar, seriam também fáceis para se litigar nos poucos casos em que chegarem ao tribunal.

Problema semelhante ocorre com as leis. Em recente entrevista à revista *Consultor Jurídico* (SUPREMO TRIBUNAL FEDERAL, 2006), o presidente do Supremo Tribunal Federal, Nelson Jobim, expressou claramente sua preocupação com as ambigüidades nas leis. “Para o ministro, o preço que se paga, no Congresso, para fazer a maioria aprovar projetos de leis é a ambigüidade, pois quanto mais ambíguo um texto, maior a chance de aprovação. Essa ambigüidade transfere, aos juízes, o poder de decidir o que os parlamentares não decidiram.”.

Este trabalho também visa contribuir com o movimento científico que busca soluções para a autoria de documentos na Web Semântica. Este movimento é capitaneado pelo *World Wide Web Consortium* (W3C), e tem recebido importantes colaborações<sup>4</sup> de pesquisadores da linha de Arquitetura de Informação do departamento de Ciência de Informação da Universidade de Brasília. Oliveira, protagonista de tais publicações, expõe em sua tese de doutorado (Oliveira, 2006) que o ambiente de autoria proposto naquele trabalho deve auxiliar os autores a “escrever textos em qualquer ambiente institucional, técnico ou científico, onde a ocorrência de ambigüidade deve ser reduzida.” Entretanto, dada a vastidão teórica abordada naquele trabalho, não houve detalhamento dos procedimentos a serem adotados para a evitar as ambigüidades. Pretende-se que os resultados desta dissertação possam ajudar a complementar a obra de Oliveira.

Em resumo, o estudo do tema proposto tem importância para a melhoria dos sistemas de recuperação da informação, tornando a busca por informações mais precisa e relevante para os anseios do usuário. Esta pesquisa também contribui para enriquecer as soluções de autoria de documentos no ambiente da Web Semântica, possibilitando que o conhecimento contido nas peças documentais possa ser representado e formalizado.

---

<sup>4</sup> ver Oliveira & Lima-Marques (2006), Oliveira (2006) e Oliveira (2004).

## **2.3 - Metodologia**

### **2.3.1 - Método de Pesquisa Utilizado**

Segundo Gil (1999) é possível agrupar as mais diversas pesquisas científicas em um certo número de agrupamentos amplos, a saber: pesquisas exploratórias, descritivas e explicativas. As pesquisas descritivas têm por objetivo a descrição das características de determinada população ou fenômeno, ou o estabelecimento de relações entre variáveis. As explicativas são aquelas que têm como preocupação central identificar os fatores que determinam ou que contribuem para a ocorrência dos fenômenos. Por fim, as pesquisas exploratórias têm como principal finalidade desenvolver, esclarecer e modificar conceitos e idéias, tendo em vista, a formulação de problemas mais precisos ou hipóteses pesquisáveis para estudos posteriores. É neste último nível de pesquisa, o exploratório, que reside esta dissertação. Apesar de o tema objeto de estudo, no caso a ambigüidade, ser estudado desde a antiguidade clássica, a solução proposta nesta dissertação pode ser considerada de vanguarda. Por esse motivo não há pesquisas disponíveis que se prestam a solucionar o problema proposto com as ferramentas propostas, o que faz deste trabalho um instrumento exploratório que irá auxiliar na edificação de problemas mais esclarecidos, passíveis de investigação mediante procedimentos mais sistematizados.

O método de pesquisa escolhido para o desenvolvimento desta dissertação é a pesquisa bibliográfica. Segundo Gil (1999), a pesquisa bibliográfica é feita a partir de fontes secundárias, que são classificadas em: livros da leitura corrente, livros de referência e consulta, e publicações periódicas. As principais fontes consultadas para a edificação desta pesquisa foram:

- Annual Review of Information Science and Techonolgy;
- Communications of the ACM;
- Information Processing and Management;
- Journal of the American Society for Information Science;
- Revista Ciência da Informação;
- Datagramazero;

- D.E.L.T.A (Documentação de Estudos em Lingüística Teórica e Aplicada)
- Scielo;
- ACM Digital Library;
- ISI Web of Knowledge;
- LISA;
- Sítio do World Wide Web Consortium (W3C);
- livros;
- teses e dissertações;
- Google Acadêmico.

A pesquisa bibliográfica se deu em duas fases: uma inicial, onde foi feita uma exploração ampla da literatura sobre o tema; e uma fase posterior, onde se procurou delimitar a bibliografia, priorizando os trabalhos com relevância científica que pudessem delimitar o objeto de estudo.

Na pesquisa bibliográfica atentou-se principalmente para a questão da ambigüidade como problema lingüístico e suas conseqüências negativas para a comunicação humana. Além disso prestou-se especial atenção às soluções de outros pesquisadores para o problema, bem como às diversas classificações de ambigüidades disponíveis na literatura.

Pode-se verificar que este tema consiste em um problema multidisciplinar, sendo abordado na Filosofia, Ciência da Computação, Ciência da Informação e, notadamente, na Lingüística. Dessa maneira, procurou-se compilar, à luz dos preceitos da Ciência da Informação, os estudos oriundos dessas diversas disciplinas, a fim de se criar um construto teórico que situasse e delimitasse o objeto de estudo.

### **2.3.2 – Passos Metodológicos Adotados**

Foram quatro os passos metodológicos utilizados para atingir os objetivos propostos nesta dissertação. São eles: definição de um arcabouço teórico para solução do problema, definição da estrutura de uma ontologia para suportar o referencial proposto, proposta de uma gramática para a meta-linguagem de valências e desambiguação utilizando o arcabouço teórico proposto. Segue a descrição de cada um dos passos.

### **2.3.2.1 – Definição de um arcabouço teórico para solução do problema**

O primeiro passo metodológico consistiu na composição de um arcabouço teórico que pudesse ser aplicado na autoria de documentos para eliminar ou ao menos minimizar a susceptibilidade dos mesmos às ambigüidades.

A construção deste referencial foi norteadada pela idéia de que o autor do documento, de posse das informações morfológicas, sintáticas e semânticas sobre o conteúdo da mensagem, pode assumir um papel mais participativo na comunicação, especialmente na recuperação da informação. O autor do documento deixaria explícitas as informações morfossintáticas e semânticas, de maneira que o usuário receptor da mensagem possa, sem inferências, determinar o exato significado do conteúdo comunicado. Para isso, o autor do documento faria uso de um ambiente de autoria, baseado nos moldes propostos por Oliveira (2006). Tal ambiente serviria como meio de processamento semi-automático de linguagem natural para que o usuário redija o documento em linguagem natural. O processamento é considerado semi-automático pois, além de algumas inferências automáticas, haverá predominância de interferência humana nas escolhas morfológicas, sintáticas e semânticas do texto.

A obra de Bräscher (1999) e a gramática de valências para o português (BORBA, 1996) são as principais ferramentas para a viabilização deste arcabouço teórico. É através desta gramática, auxiliada por outras teorias importantes presentes em Bräscher (1999), que o autor do documento irá explicitar as informações morfológicas, sintáticas e semânticas sobre o conteúdo da mensagem. A linguagem que possibilitará tal feito será chamada de meta-linguagem de valências (MLV).

Dessa maneira, o arcabouço teórico será composto por três níveis, a saber: (1) O nível da linguagem natural; (2) O nível da meta-linguagem de valências e; (3) O nível ontológico. O arcabouço teórico e seus três níveis serão descritos com detalhes após a revisão de literatura.

### **2.3.2.2 – Definição da estrutura de uma ontologia para suportar o referencial proposto**

O segundo passo metodológico consiste em definir uma ontologia que possa fundamentar o referencial proposto. Basicamente uma ontologia é composta por classes, instâncias, regras e relações. Estes componentes ontológicos deverão ser criados a luz da gramática de valências de Borba.

Para tal, os dois dicionários de Borba, o Dicionário de Usos do Português do Brasil (Borba (2002)) e o Dicionário Gramatical de Verbos do Português Contemporâneo do Brasil (Borba (1990)), terão papel fundamental, já que o primeiro contempla a valência lógica e sintática dos verbos, advérbios, adjetivos e nomes, e o segundo contém informações sobre as matrizes valenciais dos verbos, incluindo as valências lógica e semântica.

A linguagem de ontologias para a Web (OWL) e o aplicativo Protégé serão os ferramentais tecnológicos que viabilizarão a concretização deste passo.

### **2.3.2.3 – Proposta de uma gramática para a meta-linguagem de valências**

O terceiro passo metodológico versará sobre uma proposta de gramática para a Meta-Linguagem de valências. A meta-linguagem de valências corresponde ao segundo nível do arcabouço teórico proposto. Como toda linguagem, esta precisa de regras gramaticais que determinarão sentenças validas (gramaticais) e sentenças não válidas (agramaticais).

Para atender a este passo metodológico o principal fundamento será extraído dos trabalhos de Noam Chomsky, principal expoente da lingüística contemporânea. A teoria gramatical do autor embasará a gramática da meta-linguagem de valências.

### **2.3.2.4 – Desambiguação utilizando o arcabouço teórico proposto**

O último passo metodológico consiste em efetuar uma validação inicial do arcabouço teórico proposto na dissertação. Neste momento metodológico o trabalho de Silva (2006) teve extrema importância pois foram utilizados alguns dos exemplos de frases ambíguas presentes naquela dissertação.

Considera-se validação inicial pois o corpus a ser submetido ao arcabouço é muito limitado. Além do corpus limitado, as frases que são submetidas ao arcabouço foram escolhidas propositalmente de maneira a demonstrar características importantes do arcabouço e evitando limitações impostas pela gramática da meta-linguagem de valências, que se mostrou ainda imatura para a representação de sentenças complexas.

## ***3 - Revisão da Literatura e Fundamentos***

A organização da fundamentação teórica se dá da seguinte forma: O item 3.1 trata das ambigüidades, problema e objeto de estudo desta dissertação. Logo após, abordar-se o tema da gramática, desde sua teoria fundamentadora, seção 3.2, passando por conceitos utilizados usualmente pelas gramáticas normativas do português contemporâneo do Brasil, seção 3.3, até chegar à gramática de valências para o português, seção 3.4, principal ferramenta para a proposta de resolução do problema. A seção seguinte, 3.5, versa sobre a teoria da comunicação, que é um importante tópico, já que a ambigüidade é fenômeno que se dá no processo de comunicação. As ontologias, seção 3.6, são tratadas em seguida. As ontologias consistem em um dos níveis do referencial proposto nesta dissertação. A seção seguinte, 3.7, intitulada Tratamento automático de ambigüidades na recuperação da informação, advém da obra da Prof. Dra. Marisa Bräscher, obra esta que inspirou o presente trabalho. Nesta seção demonstra-se como se dá uma solução para resolução de ambigüidades no momento da recuperação da informação e como a autora trabalhou com algumas das ferramentas utilizadas aqui. Logo após, na seção 3.8, o tópico tratado é a Web semântica e os ambientes de autoria baseados em ontologias. É no contexto abarcado por estes dois conceitos que futuras aplicações do arcabouço teórico desta dissertação serão ambientadas. Por fim, na seção 3.9, a revisão de literatura é concluída com os fundamentos relacionados às ferramentas utilizadas para a validação dos resultados desta pesquisa, a saber: XML – seção 3.9.1, RDF – seção 3.9.2, OWL – seção 3.9.3 e Protégé – seção 3.9.4.

### **3.1 – Ambigüidades como Problema**

Em termos gerais, os autores atribuem este fenômeno à possibilidade de se denotar<sup>5</sup> mais de um significado a partir de uma expressão da língua.

Segundo FUCHS (1996 *apud* BRÄSCHER, 2002), entende-se por ambigüidade uma expressão da língua (palavra ou frase) que possui vários significados distintos, podendo, conseqüentemente, ser compreendida de diferentes maneiras por um receptor. Já segundo WASOW *et al.* (2005), “Ambigüidade é uma propriedade semântica. (...) Uma expressão é ambígua se tem duas ou mais denotações distintas – isto é, se está associada com mais de uma região do espaço de significados. As instâncias mais óbvias da ambigüidade envolvem expressões com (pelo menos) duas denotações que são disjuntas uma da outra.”. Para Borba (2002), a ambigüidade consiste na “possibilidade de mais de uma interpretação”. Em Houaiss (2004) a descrição do termo, para a rubrica de lingüística, é a seguinte: “propriedade que possuem diversas unidades lingüísticas (morfemas, palavras, locuções, frases) de significar coisas diferentes, de admitir mais de uma leitura”. Para Wikipedia (2005), a ambigüidade refere-se a uma propriedade de palavras, termos e conceitos, (dentro de um contexto particular) que têm um significado não claro, indefinido, indefinível ou, de outra maneira, vago<sup>6</sup>.

“Significante e significado são termos utilizados em uma área da lingüística e da crítica literária para descrever os componentes de um signo: o significante, em termos simples, é a palavra, e o significado é a coisa ou idéia que ela representa.” (ROBREDO, 2003). A ambigüidade, portanto, pode ser entendida quando um único significante remete a mais de um significado.

Um exemplo de uma expressão ambígua segue abaixo:

a) Eu vi o banco.

Na frase acima, é possível denotar várias interpretações possíveis. A palavra “banco” gera o fenômeno da homonímia, e tem vários significados possíveis. Dentre as várias acepções, um

---

<sup>5</sup> Entende-se por denotação o uso do signo em seu sentido real.

<sup>6</sup> Nesta definição de ambigüidade, Wikipedia (2005) mistura os conceitos de ambigüidade, indeterminação e vagueza. Ver seção 3.1.2, para distinguir os conceitos.

“banco” pode ser um assento, uma instituição financeira, um local onde se armazena sangue doado (“banco de sangue”), ou um acúmulo de areia no fundo de um rio (“banco de areia”). Esta frase gera uma ambigüidade lexical.

O fenômeno da ambigüidade ocorre na esfera do receptor da mensagem. Conforme Zavaglia (2003), o emissor da mensagem sabe exatamente qual é o seu referente, o seu conteúdo e o seu propósito; a interpretação do item lexical (ou da frase emitida) pelo destinatário é que ocasiona a ambigüidade.

Por comprometer a recepção da mensagem por parte do receptor da mesma, Bräscher (2002) afirma que a ambigüidade causa ruído na recuperação da informação, pois, sob um mesmo termo, o usuário (ou receptor) encontrará informação relevante e irrelevante.

Embora seja evitada na maioria das vezes, em alguns casos como na literatura, poesia, música e publicidade, a ambigüidade é utilizada propositadamente como um recurso útil. Conforme Zavaglia (2003), nesses casos a polissemia e a homonímia são fenômenos produtivos de jogos de palavras que enriquecem o conteúdo textual, revelando-se uma riqueza da expressão lingüística.

Para Bear & Hobbs (1988), a ambigüidade é um problema para qualquer sistema de processamento de linguagem natural. Segundo os autores, gramáticas extensas tendem a produzir um grande número de alternativas de análises para, inclusive, sentenças muito simples.

### **3.1.1 – Alguns Fatores Causadores de Ambigüidades**

Segundo Bräscher (2002), a ambigüidade pode ser gerada pelos seguintes fatores: polissemia, homonímia, policategorização, relação contextual e estrutura sintática das frases. Silva (2006) adiciona aos fatores arrolados por Bräscher (2002) os fenômenos da anáfora e metáfora.

#### **3.1.1.1 – Polissemia e Homonímia**

Segundo Martins *et al* (1999), “A falta de correspondência biunívoca entre o significante e o significado das palavras da língua, embora não constitua problema (*grave*) para o falante humano, revela-se impedimento dos mais sérios para o processamento automático das línguas naturais.”. Os fenômenos da polissemia e homonímia são os dois tipos especiais de fenômenos decorrentes da falta de correspondência biunívoca entre o significante e o significado das palavras da língua.

Existe uma tênue fronteira entre os conceitos de polissemia e homonímia. Após uma extensa revisão da literatura sobre o assunto Zavaglia (2003) propõe um framework como critérios para distinguir os conceitos. Segundo a autora, o fenômeno da polissemia é unicategorial, ou seja, só se poder considerar um significante polissêmico quando os possíveis significados referem-se a significantes homógrafos de uma mesma categoria morfológica. Além disso, na polissemia, os significados são classificados como conjuntivos, ou seja, não se perde de vista o significado antigo e original do significante primário, existindo uma coexistência entre os novos significados gerados com significado primário. Para ilustrar os dois critérios propostos por Zavaglia (2003), analisemos algumas palavras polissêmicas, adaptadas de Borba (2002):

- a) boca1 – Substantivo – cavidade da parte inferior da face, entre as maxilas, delimitada pelos lábios, constituindo a primeira parte do tubo digestivo.
- b) boca2 – Substantivo – abertura ou fenda na parte anterior: *é a boca do canhão.*
- c) cópia1 – Substantivo – reprodução textual de um escrito.
- d) cópia2 – Substantivo – pessoa ou coisa muito semelhante à outra: *você é a cópia do Michael Jackson.*

Diferentemente da polissemia, na homonímia não existe restrição morfológica categorial. Conforme Zavaglia (2003), existem, portanto, formas homógrafas policategoriais e unicategoriais. Zavaglia amplia esta categorização das formas homógrafas ao detalhar que essas podem ser polissêmicas, formas homógrafas em que pelo menos uma delas possui mais de uma acepção, ou monossêmicas, formas homógrafas sendo que cada uma delas possui apenas uma acepção. Um critério importante que distingue a homonímia da polissemia diz respeito à questão de que os significados na homonímia são disjuntivos, ou seja, conforme Fuchs (1996), citada logo acima, existe uma “colisão acidental entre as formas de dois signos lingüísticos distintos”. Na homonímia, não há relação entre os significado. Ullmann (1964 *apud* Zavaglia, 2003), cita os três únicos processos pelos quais podem surgir as homonímias:

- (i) Convergência fonética, ou seja, o desenvolvimento de sons convergentes, isto é, quando dois ou mais itens lexicais tiveram, no passado, formas diferentes que coincidem na linguagem falada e escrita;
- (ii) Divergência semântica, ou seja, quando há o desenvolvimento de sentidos divergentes, isto é, "quando dois ou mais significados da mesma palavra se separam de tal modo que não haja nenhuma conexão evidente entre eles, a polissemia dará lugar à homonímia e a unidade da palavra será destruída (...)" (Ullmann, 1964:368). É o caso de palavras como canal1 (abertura, passagem de água, cavidade) e canal2 (meio de transmissão de sinais); criação1 (obra, invenção) e criação2 (animais domésticos criados conjuntamente), dentre muitos outros exemplos do português contemporâneo;
- (iii) Influência estrangeira, ou seja, quando palavras estrangeiras se introduzem em uma língua, adaptando-se ao sistema fonético no qual se introduziram, e coincidem com outras palavras já existentes; conseqüentemente, surgem pares de homônimos. Segundo Ullmann (1964), além de ser um processo raro, "este tipo de influência estrangeira não é, pois, uma fonte separada de homonímia, mas apenas uma forma especial de desenvolvimentos fonéticos convergentes" (p.373).

No esquema abaixo, retirado de Zavaglia (2003), estão organizados metodologicamente as homonímias, suas variações e alguns exemplos.

<b>HOMONÍMIA: significados disjuntivos</b>					
<b>Homonímia Categorial</b>		<b>Homonímia Semântica</b>		<b>Homonímia Heterófona</b>	
<b>Formas Homógrafas Pluricategoriais Monossêmicas (FHPM)</b>	<b>Formas Homógrafas Pluricategoriais Polissêmicas (FHPP)</b>	<b>Formas Homógrafas Unicategoriais Monossêmicas (FHUM)</b>	<b>Formas Homógrafas Unicategoriais Polissêmicas (FHUP)</b>	<b>Formas Heterófonas Pluricategoriais Monossêmicas (FHrPM)</b>	<b>Formas Heterófonas Pluricategoriais Polissêmicas (FHrPP)</b>
<b>EXEMPLOS:</b>					
<castigo <sub>1</sub> (v.); castigo <sub>2</sub> (subs.)>	<falta <sub>1</sub> (v.); falta <sub>2a</sub> , falta <sub>2b</sub> , falta <sub>2c</sub> (subs.)>	<banco <sub>1</sub> (subs.); banco <sub>2</sub> (subs.)>	<ato <sub>1</sub> (subs.); ato <sub>2a</sub> , ato <sub>2b</sub> , ato <sub>2c</sub> (sub.); ato <sub>3a</sub> , ato <sub>3b</sub> (subs.)>	<começo <sub>1</sub> (v.); começo <sub>2</sub> (subs.)>	<acordo <sub>1</sub> (v.); acordo <sub>2a</sub> , acordo <sub>2b</sub> , acordo <sub>2c</sub> , acordo <sub>2d</sub> (subs.)>

Figura 1 – Classificação do fenômeno da Homonímia

FONTE: ZAVAGLIA (2003)

### **3.1.1.2 – Policategorização**

A policategorização consiste no fato de se poder atribuir a uma mesma palavra mais de uma categoria morfológica, como verbo, nome ou adjetivo. “Desta forma, a frase que comporta essa palavra pode não ter seu sentido definido da melhor forma, ou ainda este pode ser exatamente o efeito procurado pelo autor.” (SILVA, 2006).

### **3.1.1.3 – Relação Contextual**

A relação contextual é um fator causador de ambigüidades que atua no nível pragmático. Muitas vezes, especialmente quando as informações contidas no co-texto oracional não são suficientes para se resolver a ambigüidade, seriam necessárias coletar informações pragmáticas sobre o contexto no qual está inserida a oração ambígua. Um exemplo de ambigüidade pragmática dado por Bräscher (2002) pode esclarecer este tópico. Na oração “Paulo vai à escola”, só o contexto pode resolver a ambigüidade pragmática, pois só com informações do tipo “Paulo é um estudante”, ou “Paulo vai buscar a irmã na escola”, é que será possível determinar com exatidão se Paulo está indo à escola por que é estudante, ou por outro motivo qualquer.

Cabe, neste momento, diferenciar os conceitos de co-texto e contexto. Segundo Welker (2003), “co-texto é o texto ao redor de determinada palavra, geralmente no interior de uma oração; contexto poderia restringir-se ao texto inteiro no qual um item lingüístico a ser estudado ou traduzido é inserido, mas pode também referir-se à situação de comunicação (portanto, a fatos extra-lingüísticos), sendo às vezes chamado, nesse caso, de `contexto de situação`.”

### **3.1.1.4 – Estrutura Sintática das Frases**

Quando uma frase pode ser analisada sintaticamente por mais de uma árvore gramatical válida, então é possível encontrar ambigüidade em tal frase. Conforme expõe Bräscher

(2002), “As frases preposicionais são uma das fontes mais frequentes de ambigüidade sintática.”.

### 3.1.1.4 – Anáforas e Metáforas

As anáforas são consideradas uma figura de estilo. As anáforas consistem na repetição sucessiva e sistemática de uma palavra no início de diferentes frase ou de membros de um mesma frase. Para Leffa (2001 *apud* SILVA, 2006), “uma anáfora pode ser descrita como um processo que acarreta a volta no texto. O processo começa quando o anaforizante é conhecido (um pronome, por exemplo) e concluído quando o anaforizado é encontrado (a palavra a qual o pronome se refere)”. Um exemplo de anáfora pode ser verificado abaixo.

- a) **Tão** simples, **tão** franco, **tão** honesto. Daniel era um exemplo a ser seguido.

Segundo Wikipedia (2006), “metáfora é uma figura de estilo lingüística, designada pelo uso de uma palavra ou expressão num sentido que não é o próprio, baseado numa relação de semelhança. O processo de produção da metáfora requer a comparação entre entes diversos retendo o que se considera semelhante, para estabelecer um novo significado. Sendo assim, a metáfora é mais do que uma figura ou esquema da língua, é do pensamento, é cognitiva.”. Vários são os exemplos da utilização de metáforas:

- a) A Amazônia é o pulmão do mundo.
- b) Meu coração é um balde despejado.
- c) Marcos é um furacão em Medicina.
- d) Felícia foi sempre o abismo que separou a família.

### 3.1.2 – Outros fenômenos lingüísticos relacionados à Ambigüidade

Vários fenômenos da língua são costumeiramente confundidos com o fenômeno da ambigüidade. Dentre esses podemos citar: vagueza, duplo sentido e a indeterminação semântica. Embora todos esses fenômenos possam também contribuir para a geração de

ambigüidades, cabe diferenciá-los, de maneira a possibilitar a definição das fronteiras do problema foco de estudo nesta dissertação.

### **3.1.2.1 – Indeterminação**

A linguagem é dinâmica e tem por característica se renovar e rearranjar indefinidamente. Sobre a linguagem Strogenski (2005) discorre: “a sua inegável característica de sempre poder significar coisas novas num engendramento ”criativo” que toma o já existente e o reelabora, devolvendo ao mundo, através de si própria, novas significações que se juntarão às já existentes, ampliando a base sobre a qual o sujeito se debruçará (ou se apoiará) para “fabricar” novos sentidos indefinidamente.”. Neste contexto dinâmico aparece o conceito de indeterminação semântica. A indeterminação ocorre justamente quando não é possível “produzir um discurso com uma forma e interpretação precisas em um sistema lingüístico indeterminado e, considerando a possibilidade de mais de uma interpretação, como a interpretação desejada (ou interpretações possíveis) é determinada.” (COUTINHO, 1996). É por isso que muitos lingüistas apontam a indeterminação semântica como um conceito que abarca outros como a própria ambigüidade, a vagueza e o duplo sentido.

### **3.1.2.2 – Vagueza**

O fenômeno da vagueza ou vaguidade diz respeito à falta de informação ou especificação sobre uma palavra ou uma oração da língua. Para Alves e Aluísio (2006), “A fonte do fenômeno da vagueza ou vaguidade encontra-se no limite entre o léxico e o contexto. Pode-se considerar que uma palavra é vaga quando não é possível – apenas com as informações fornecidas por ela, ou seja, no léxico – recuperar seu referente ou sentido com precisão.”. Segundo Wasow *et al* (2005), expressões são vagas se as regiões que elas denotam não tem fronteiras bem definidas.

### **3.1.2.3 – Duplo Sentido**

O duplo sentido diferencia-se da ambigüidade apenas pelo fato de que o duplo sentido manifesta-se através da fala. Esta concepção é proposta por Grésillon (1988 *apud* SILVA, 2006). Segundo este autor, a ambigüidade existe por si só em textos e enunciados. Já o duplo sentido, é definido como um fenômeno que não existe sem a presença do sujeito falante.

### 3.1.3 - Tipos de Ambigüidades

Existem várias propostas de tipificação de ambigüidades na literatura. Este aspecto do estudo sobre ambigüidades não é uma unanimidade no meio científico. A seguir serão descritas algumas das propostas.

Uma das classificações mais comuns de ambigüidades pode ser encontrada em Wasow *et al* (2005) e em Wikipedia (2005). Esta classificação prega a divisão de ambigüidades em Léxicas e Sintáticas. Segundo esta proposta, ambigüidades léxicas surgem quando o contexto é insuficiente para determinar um único significado para uma palavra. Um exemplo típico deste tipo de ambigüidade é a palavra “banco”, que pode ter várias interpretações, tais como: instituição financeira, assento, local de depósito, dentre outras. De fato, é possível encontrar em dicionários genéricos mais de quinze denotações para a palavra “banco” em língua portuguesa.

A ambigüidade sintática surge quando uma sentença pode ser validada sintaticamente em mais de uma forma. Um exemplo deste tipo de ambigüidade é a frase: “Ele comeu biscoitos no sofá”. Desta frase podem-se inferir duas interpretações: (1) O sujeito da frase pode ter comido os biscoitos que estavam sobre o sofá, ou (2) o sujeito poderia estar sentado no sofá comendo os biscoitos. Existe uma distinção clara entre a ambigüidade léxica e a sintática. Na frase citada a pouco, sob o olhar do senso comum, é possível afirmar que todas as palavras têm um único significado, o que implica que a ambigüidade causada pela expressão não pode ser do tipo léxica.

De acordo com Zavaglia (2003), segundo a ótica lingüística, a ambigüidade define-se em função das regras gramaticais. E assim, quando uma frase possibilita duas interpretações semânticas ou sintáticas, ela considerada ambígua. Desse modo temos a ambigüidade léxica gerada pela homografia que pode ser morfo-semântico-categorial: “A casa é bonita” / “Ela

casa amanhã”, exclusivamente semântica: “Sentei-me no banco do jardim” / “Fui ao banco sacar dinheiro” ou sintática: “Entrei no carro andando”.

Zavaglia (2003) também cita a proposta de Ullmann (1964). Segundo Ullmann (1964 *apud* ZAVAGLIA, 2003) lingüisticamente, existem três tipos principais de ambigüidades: Fonética, Gramatical e Lexical.

A ambigüidade fonética diz respeito à interpretação do som gerado pela pronúncia de uma unidade acústica da linguagem. Uma unidade acústica da linguagem pode ser composta pelo som de mais de uma palavra (ou partes de palavras). Exemplos em português deste tipo de ambigüidade são citados por Zavaglia (2003): “agosto”, o oitavo mês do ano / “a gosto”, locução adverbial que significa “à vontade”; “enquanto”, conjunção / “em quanto”, locução interrogativa.

Já a ambigüidade gramatical pode ser gerada de dois modos: (1) pela ambigüidade de formas gramaticais e/ou (2) pela ambigüidade da estrutura da frase. Em (1) temos o caso de prefixos e sufixos que possuem mais de um significado e, por isso, são ambíguos, ou aqueles que são, por natureza homônimos (Zavaglia, 2003). Exemplos em língua inglesa são citados por Ullmann (1964 *apud* ZAVAGLIA, 2003): o sufixo “*able*” não significa a mesma coisa em “*desiderable*” (desejável) ou “*readable*” (Legível) que em “*eatable*” (Comestível), “*Knowable*” (Conhecível), “*debatable*” (Debatível), etc. Já em (2) temos a possibilidade da ambigüidade por meio da “frase equívoca (anfibia)” como nomeia Ullmann. Neste caso, a combinação de palavras que não são ambíguas possibilita a interpretação de dois ou mais modos diversos da frase Zavaglia (2003). A frase “Eu vi a menina com o telescópio” é um exemplo típico deste tipo de ambigüidade.

Para Ullmann (1964, *apud* ZAVAGLIA, 2003), a ambigüidade lexical é um dos fatores mais importantes dentre os fatores de ambigüidade presentes em uma língua. Com efeito, a “polivalência das palavras”, tomando emprestado o termo de Ullmann (1964), assume duas formas diferentes: a polissemia e a homonímia.

Segundo Fuchs (1996 *apud* BRÄSCHER, 1999) a ambigüidade pode ser classificada em seis tipos: Ambigüidade Morfológica, Ambigüidade Lexical, Ambigüidade Sintática, Ambigüidade Predicativa, Ambigüidade Semântica e Ambigüidade Pragmática. Esta

classificação é determinada pelos fatores causadores de ambigüidades: polissemia, homografia, policategorização, relação contextual e estrutura sintática das frases. Bräscher (2002) descreve cada um dos subtipos de ambigüidade citados acima:

I) Ambigüidade morfológica: ocorre quando não é possível classificar determinada forma quanto à categoria gramatical. Este tipo de ambigüidade é ocasionado pela policategorização – em que palavras pertencem a mais de uma categoria gramatical, como proposta, que pode ser ou substantivo, ou adjetivo ou verbo.

II) Ambigüidade lexical: ocorre quando há mais de uma interpretação possível do significado de uma unidade lexical. Este tipo de ambigüidade é provocado por :

homografia : ocorre por meio da “colisão acidental entre as formas de dois signos lingüísticos distintos”. (FUCHS, 1996, p.9). Ex.: cobre (metal) ; cobre (do verbo cobrir).

polissemia : ocorre quando uma só e mesma expressão envolve significados distintos, sendo um único signo lingüístico; é a própria expressão que é ambígua, à medida que possui uma forma à qual corresponde uma pluralidade de significados. (FUCHS, 1996). Ex.: arquivo (móvel, instituição, conjunto de documentos).

III) Ambigüidade sintática : ocorre na estruturação da frase em constituintes hierarquizados, quando se definem as ligações que se estabelecem entre os sintagmas. As frases preposicionais são uma das fontes mais freqüentes de ambigüidade sintática. Exemplos:

- Eu li a notícia sobre a greve na universidade. (ou eu li a notícia e eu estava na universidade, ou a greve ocorre na universidade)

- A professora de dança espanhola. (ou a professora é espanhola, ou a dança é espanhola)

IV) Ambigüidade predicativa : ocorre na interpretação das relações temáticas que articulam predicado, argumentos e participantes. Exemplos :

A crítica deste autor. (autor = ou objeto da crítica, ou agente da crítica)

Eu a deixei feliz. (feliz = ou atributo do sujeito ou atributo do objeto)

V) Ambigüidade semântica : ocorre quando há mais de uma interpretação possível para o relacionamento dos termos na frase, como, por exemplo, no cálculo dos operadores de negação e de quantificação :

- Ela não chora mais porque ele partiu. (ou ela chorava porque ele havia partido, ou ela parou de chorar uma vez que ele já foi embora)

- Um rio corre através de cada país europeu. (ou um único rio corre através de todos os países, ou diferentes rios correm através de diferentes países)

VI) Ambigüidade pragmática : relaciona-se ao cálculo dos valores enunciativos, à reconstrução destes valores, que estão ligados à situação do falante no momento da enunciação, como por exemplo :

- Os pássaros voam. (referência geral ou específica?)

- Paulo vai à escola. (ele é estudante ou ele está indo à escola neste momento?)

Silva (2006) faz uma ampla revisão da literatura sobre o assunto e enumera uma série de tipos de ambigüidades encontrados na literatura. Após tratamento terminológico dos termos utilizados para representar tipos de ambigüidades, Silva (2006) descreve cada um deles conforme segue:

Quadro 1 – Tipos de ambigüidades da Língua Portuguesa

Tipo de Ambigüidade	Descrição
Âmbito	<p>Âmbito significa campo de ação, esfera ou contexto relacionado com uma área de conhecimento. É o resultado da co-ocorrência de mais de um determinante quantificacional (<i>sic</i>) na mesma frase. Ocorre uma ambigüidade de âmbito em uma afirmação quando há mais de uma maneira de interpretar o âmbito das suas conectividades.</p> <p>Exemplo:            Todas as mulheres têm um sonho.            1a interpretação: Elas compartilham o mesmo sonho.            2a interpretação: Cada uma tem seu respectivo sonho.</p>
Sintática	<p>Refere-se à posição de um sintagma, que é o conjunto de duas ou mais palavras que possuem um significado, mas que por si só não podem formar uma frase completa, no contexto expresso por uma frase. Observa-se que essa ambigüidade também é conhecida como estrutural.</p>
Sistemática	<p>Relacionada com os critérios estruturais ou funcionais, de acordo com os quais as instâncias lingüísticas são classificáveis como pertencentes a este ou aquele tipo lingüístico.</p> <p>Exemplo:            O fazendeiro viu o campeiro no cavalo.            1a interpretação: O fazendeiro estava próximo do cavalo.            2a interpretação: O fazendeiro estava montado no cavalo.            3a interpretação: O campeiro estava montado no cavalo.</p>
Lexical	<p>Ocorre quando é possível aplicar mais de uma interpretação para uma unidade lexical. A unidade lexical está relacionada com os vocábulos de uma língua. Um vocábulo é uma palavra que pode ter seu sentido considerado quanto ao som ou à forma.</p> <p>Exemplo:            A divisão já foi encerrada pela direção.            1a interpretação: Partilha, distribuição ou repartição de bens.            2a interpretação: Operação pela qual se conhece quantas vezes uma quantidade está contida noutra.</p>
Interlingual	<p>Está relacionada com a comunicação entre línguas. Essa ambigüidade tem se tornado cada vez mais freqüente por causa do grande número de informação que transita entre vários países.</p>
Lingüística	<p>Gerada apenas por questões lingüísticas e é detectada quando determinados enunciados em condições já previstas apresentam problemas de escolha lingüística ao receptor, gerando uma flutuação entre duas ou mais condições aceitáveis. Não se contamina com aspectos não lingüísticos.</p>
Morfológica	<p>Ocorre em decorrência do efeito do evento da policategorização. Esse fenômeno lingüístico permite que uma palavra seja classificada em mais de uma categoria gramatical. É possível afirmar que essa ambigüidade ocorre quando não é possível classificar determinada forma quanto à categoria gramatical, podendo esta ser um substantivo, adjetivo ou verbo.</p>
Categorial	<p>Ocorre quando há falta de correspondência entre os conjuntos de significado; também conhecida como a falta de correspondência biunívoca onde se tem o significante, que é a forma, não localizando um significado correspondente no conjunto de significados, que é o conteúdo das palavras da língua.</p>

Gramatical	<p>Ocorre quando se tem uma palavra que pode pertencer a várias categorias gramaticais.</p> <p>Exemplo: A vara era muito grande. 1a interpretação: Vara é objeto comprido e fino de madeira. 2a interpretação: Vara é substantivo coletivo de porcos.</p>
Poética	<p>Entende-se que essa ambigüidade é vista como um elemento que constitui a poesia e que, não pode ser resolvida, afim de preservar seu objetivo de permitir a geração de múltiplos sentidos, caracterizando sua linguagem e a mudança de perceptiva que ela provoca, sendo a exceção concedida à língua para a sua subversão, e se distingue suficientemente da ambigüidade manifestada na linguagem natural.</p> <p>Exemplo: Desce<sup>7</sup></p> <p>Desce do trono, rainha Desce do seu pedestal De que te vale a riqueza sozinha Enquanto é carnaval (...)</p>
Pragmática	<p>Pragmática expressa conjunto de regras ou fórmula, resultado do conflito entre a significação expressa na sentença e a intenção do enunciador no momento da enunciação. Aponta para o uso e referência de palavras e expressões da língua, estabelece relação com os pronomes pessoais e advérbios de lugar e tempo, onde a referência ultrapassa o contexto lingüístico, a significação de “eu”, “aqui” e “agora” não depende do contexto atual e excede o texto criando relações externas, buscando elementos para a sua resolução.</p> <p>Exemplo: Eles cometeram um crime. 1a interpretação: Cometeram o mesmo crime. 2a interpretação: Cometeram crimes separados.</p>
Predicativa	<p>Um predicado é definido como um verbo de significação indefinida, que qualifica o sujeito ou o complemento direto. A ambigüidade predicativa ocorre quando há problemas relacionados a interpretação das relações temáticas que articulam o predicado.</p> <p>Exemplo: O homem parou atento. 1a interpretação: O homem é atento. 2a interpretação: O homem ficou atento.</p>
Semântica	<p>Ocorre quando uma palavra possui multiplicidade de conceitos relacionados com sua aplicabilidade. A polissemia pode ser apontada como um fenômeno responsável por esse tipo de ambigüidade.</p> <p>Exemplo: Comprei uma lancha e dois barcos velozes. 1a interpretação: A lancha e os barcos são velozes. 2a interpretação: Os barcos são velozes.</p>

<sup>7</sup> Trecho da canção “Desce”, de Arnaldo Antunes.

Referencial	<p>Está relacionada com enunciados potencialmente ou efetivamente incompletos, possui a característica de ser uma ambigüidade elíptica. O efeito elíptico determina que um enunciado não pode conter todas as informações necessárias para o seu completo entendimento. Logo, é sempre possível acrescentar um elemento que possa, necessariamente, contribuir para o seu entendimento, causando um efeito de precisão, modificação ou também causando mais contradição ao enunciado em questão.</p> <p>Exmplo: E espero seja a última. 1a questão: O que é a última?</p>
Transfrástica	<p>Conhecido como ambigüidade por disjunção, é aquela configuração lingüística cuja significação se constrói pela separação de dois termos mutuamente excludentes. O conceito da disjunção implica na escolha de uma alternativa entre duas, o que irá provocar a exclusão automática da outra.</p> <p>Exemplo: Os professores do colegiado atenderam os alunos, eles estão felizes. 1a interpretação: Os alunos estão felizes. 2a interpretação: Os professores estão felizes.</p>
Virtual	<p>A aceitação do duplo, várias interpretações em um único contexto, só pode ser aceita no mundo abstrato. Essa ambigüidade tem seus moldes baseados na disjunção relativa, onde não é obrigatória a escolha de apenas uma interpretação possível. A possibilidade de múltipla escolha, mas de forma controlada e previsível.</p> <p>Exemplo: Certos alimentos me deixam doente. 1a interpretação: alguns alimentos, não todos. 2a interpretação: um certo tipo, frutos do mar. 3a interpretação: indeterminado: entre quantidade e qualidade.</p>

FONTE: SILVA, 2006.

O trabalho de Silva (2006) é importante, pois reúne em uma única obra diversas classificações de ambigüidades presentes na literatura. Entretanto, a taxonomia proposta pelo autor demonstra intersecções entre algumas classificações de mesmo nível hierárquico, fato que não pode acontecer em uma taxonomia. Exemplificando, uma única oração ambígua como “A vara era muito grande” pode ser classificada em mais de um tipo (gramatical e lexical) na taxonomia de Silva (2006). Por esse motivo, adota-se a classificação de Fuchs (1996 *apud* BRÄSCHER, 1999). A classificação proposta pela autora francesa mostra-se mais estável, principalmente por permitir que uma única palavra ou oração ambígua seja classificada em um único tipo de ambigüidade na taxonomia.

### 3.2 – Noções da Teoria Gramatical

Uma linguagem é um conjunto (finito ou infinito) de sentenças, cada uma delas finitas em tamanho e construídas a partir de um conjunto finito de elementos (Chomsky, 1957). A variante brasileira da língua portuguesa é uma instância materializada do conceito de linguagem proposto por Chomsky já que, a partir de um alfabeto finito e de um conjunto de regras de formação de seqüências válidas, infinitas sentenças gramaticais na variante brasileira da língua portuguesa podem ser elaboradas.

Chomsky (1957) prossegue explicando o conceito de gramática conforme segue:

O objetivo fundamental da análise lingüística de uma linguagem L é separar seqüências gramaticais, que são as sentenças de L, das seqüências agramaticais, que não são sentenças de L, bem como estudar a estrutura das seqüências gramaticais. A gramática de L será, portanto, um dispositivo que gera todas as seqüências gramaticais de L e nenhuma seqüência agramatical.

É didaticamente interessante observar o comportamento de algumas linguagens, bem como o relacionamento com suas respectivas gramáticas. O alfabeto de tais linguagens contém apenas as letras  $x$  e  $y$ , conforme exemplos adaptados de Chomsky (1957):

a) Linguagem 1 (L1)

- i. Sentenças de L1:  $xy, xxxy, xxxyyy, xxxxyyyy, etc.$
- ii. Gramática de L1: Um número “ $n$ ”, maior ou igual a zero, de qualquer da letra  $x$ , seguido do mesmo número “ $n$ ” de letras  $y$ .

b) Linguagem 2 (L2)

- i. Sentenças de L1:  $xx, yy, xyyx, xxxx, yyyy, xxyyxx, xyyyx, etc.$
- ii. Gramática de L1: Uma seqüência qualquer de letras seguida da mesma seqüência invertida, provocando um efeito de “espelho”.

c) Linguagem 3 (L3)

- i. Sentenças de L1:  $xx, yy, xyxy, yxyx, xxxx, yyyy, xxyxyy, xyxyyy, etc.$
- ii. Gramática de L1: Uma seqüência qualquer de letras seguida da mesma seqüência repetida.

Algumas inferências básicas podem ser feitas analisando as linguagens acima. A seqüência “xyx” é agramatical em L1; a seqüência “xxx” é agramatical em L2; e a seqüência “y” é agramatical em L3, pois as três seqüências ferem a gramática de suas respectivas linguagens.

A gramática pode ser subdividida em três partes: A sintaxe, a semântica e a fonética. Para Chomsky (1957), a sintaxe é o estudo dos princípios e processos pelos quais sentenças são construídas em uma linguagem particular. Para Wikipedia (2006a), um componente semântico corresponde a um sistema das regras que definem a interpretação das frases geradas pelo componente sintático, e, um componente fonético consiste em um sistema de regras que realizam em uma seqüência de sons as frases geradas pelo componente sintático.

Conforme D’eharbe (2003), formalmente, uma gramática G pode ser representada por uma quádrupla  $G = (V, T, P, S)$  onde:

- V é um conjunto finito de símbolos não terminais, ou variáveis;
- T é um conjunto finito de símbolos terminais ( $V \cap T = \{\}$ );
- P é um conjunto finito de pares  $(\alpha, \beta)$ , chamadas regras de produção, tal que  $l$  é uma palavra não vazia de  $V \cup T$ , e  $r$  é uma palavra possivelmente vazia de  $V \cup T$ ;
- S é o símbolo inicial, tal que  $S \in V$ .

D’eharbe (2003) prossegue com o exemplo:

a) A gramática  $G = (V, T, P, N)$  é tal que:

- a.  $V = \{N, S, A, D\}$ ,
- b.  $T = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, +, -\}$ ,
- c.  $P = \{N \rightarrow SA, S \rightarrow + | - | \epsilon, A \rightarrow DA | D, D \rightarrow 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9\}$ ,
- d. N é o símbolo inicial.

Aplicando-se as regras de produção da gramática G exemplificada acima, podemos ter:

$N \rightarrow SA \rightarrow SDA \rightarrow SDD \rightarrow +DD \rightarrow +1D \rightarrow +13$ . Logo,  $N \rightarrow +13$ .

Noam Chomsky (1956) também propõe uma hierarquia de gramáticas, hierarquia esta que é composta por 4 níveis, a saber: gramáticas tipo 0, ou gramáticas com estrutura de frase; gramáticas tipo 1, ou gramáticas sensíveis ao contexto; gramáticas tipo 2, ou gramáticas livres de contexto e; gramática tipo 3, também referenciada por gramáticas regulares. Chamam-se linguagens do tipo  $n$  todas aquelas que podem ser geradas por alguma gramática do tipo  $n$ . A gramática de nível 0 é a mais poderosa por não ter restrições impostas. De fato, conforme Nunes (1999 *et al.*), “todo o universo das linguagens que se pode definir através dos mecanismos generativos definidos pelas gramáticas corresponde exatamente ao conjunto das linguagens que esta classe de gramáticas é capaz de gerar.”. A cada nível galgado na hierarquia, as restrições vão aumentando até que se chegue à gramática mais restrita, a do tipo 3. Cada nível é um super conjunto do próximo. Portanto, uma gramática de tipo  $n$  é conseqüentemente uma gramática de tipo  $n - 1$ . Vieira & Lima (2001) fazem uma análise sobre a hierarquia de Chomsky:

As gramáticas do tipo 3, ou regulares, são as mais restritas, e por isso são as mais fáceis de serem reconhecidas. São, no entanto, insuficientes para expressar as regras de formação da linguagem natural. Gramáticas do tipo 2, livres de contexto, mais poderosas, permitem a representação de linguagens com um grau maior de complexidade; estas ainda apresentam problemas para expressar dependências, como é o caso da concordância verbal. O próximo nível de gramáticas, sensíveis ao contexto, resolve o problema das dependências, mas apresenta problemas de complexidade no reconhecimento. Decidir se uma sentença pertence a uma gramática sensível ao contexto é uma função exponencial sobre o tamanho da sentença, o que torna a implementação do procedimento de verificação uma questão complexa, do ponto de vista computacional.

Conforme Nunes (1999 *et al.*), “como alternativa para a definição de uma linguagem, é possível a utilização de dispositivos aceitadores, denominados reconhecedores da linguagem. Através dos reconhecedores é possível submeter uma cadeia de símbolos a um teste de aceitação capaz de determinar se tal cadeia pertence ou não à linguagem em questão.”. O quadro abaixo demonstra as regras de produção de cada um dos níveis da hierarquia de Chomsky, bem como os tipos de reconhecedores necessários a cada um dos níveis arrolados acima:

Quadro 2 – Hierarquia de Gramáticas de Chomsky

Nível	Nome	Reconhecedor	Regras de Produção
Tipo 0	Estrutura de Frase	Máquinas de Turing <sup>8</sup>	$\alpha \rightarrow \beta$ , onde $\alpha \in (V_n \cup V_t)^+$ , $\beta \in (V_n \cup V_t)^*$ . <sup>9</sup>
Tipo 1	Sensíveis ao Contexto	Máquinas de Turing com memória limitada	$\alpha \rightarrow \beta$ , com $ \alpha  \leq  \beta $ , onde $\alpha, \beta \in (V_n \cup V_t)^+$ ou $\beta A \gamma \rightarrow \beta \alpha \gamma$ , onde $A \in V_n$ , $\alpha \in (V_n \cup V_t)^+$ , $\beta, \gamma \in (V_n \cup V_t)^*$  Obs: A segunda regra demonstra a sensibilidade ao contexto quando que especifica que A só pode ser transformado em $\alpha$ se A estiver precedido por $\beta$ e seguido por $\gamma$ .
Tipo 2	Livres de Contexto	Autômatos à pilha	$\alpha \rightarrow \beta$ , onde $A \in V_n$ , $\alpha \in (V_n \cup V_t)^*$ .
Tipo 3	Regulares	Autômatos finitos	$\alpha \rightarrow \beta$ ou $A \rightarrow b$ , onde $A, B \in V_n$ , $a \in V_t$ , $b \in (V_t \cup \{\lambda\})$ .

As gramáticas do tipo 2 e tipo 3 são amplamente utilizadas para na descrição de linguagens de programação e na implementação de interpretadores e compiladores. Entretanto, é importante notar que, de fato, a utilização das gramáticas dos tipos 0 e 1 é limitada para os fins descritos acima justamente pela amplidão formal permitida por suas regras de produção. Analisemos a seguir alguns princípios que envolvem o conceito de compilador, intimamente relacionado ao conceito de gramática, que nos será útil para a proposta de resolução do problema proposto.

### 3.2.1 – Princípios de Compiladores

Um compilador pode ser entendido como um tradutor de linguagens. Segundo Aho, Sethi & Ullman (1995), “um compilador é um programa que lê um programa escrito em uma

<sup>8</sup> Conhecida também por máquina universal, a máquina de Turing foi concebida por Alan Turing (1912-1954). É um modelo abstrato de um computador, que se restringe apenas aos aspectos lógicos do seu funcionamento (memória, estados e transições) e não à sua implementação física. Numa máquina de Turing pode-se modelar qualquer computador digital.

<sup>9</sup> Neste caso, o símbolo de adição “+” pode ser entendido como “uma ou mais instâncias de”. O símbolo “\*” pode ser entendido como conjunto das cadeias de qualquer comprimento sobre um determinado alfabeto.

linguagem – a linguagem fonte – e o traduz num programa equivalente em outra linguagem.”. O esquema abaixo demonstra conceitualmente o funcionamento de um compilador.

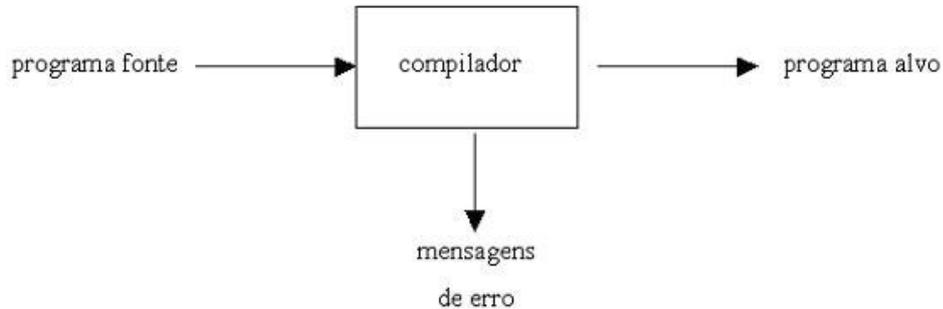


Figura 2: Compilador

Adaptada de: Aho, Sethi & Ullman (1995)

De acordo com Aho, Sethi & Ullman (1995), “existem duas partes na compilação: a análise e a síntese. A parte da análise divide o programa fonte nas partes constituintes e cria uma representação intermediária do mesmo. A relativa à síntese constrói o programa alvo desejado, a partir da representação intermediária.”. A análise pode ser dividida em três partes, a saber: Análise linear, análise hierárquica e análise semântica.

A análise linear, também conhecida por análise léxica ou escrutinador léxico, pode ser definido, conforme Nunes (1999 *et al.*), como “a parte do computador responsável por ler cada um dos elementos da sentença de entrada e transformá-los em uma representação conveniente para o analisador sintático.”. O exemplo abaixo, extraído de Aho, Sethi & Ullman (1995), descreve a atuação do analisador léxico:

- a) Sentença:  $\text{montante} := \text{depósito\_inicial} + \text{taxa\_de\_juros} * 60$ 
  - a. Identificadores: “montante”, “depósito\_inicial”, “taxa\_de\_juros”.
  - b. Símbolo de Atribuição: “:=”.
  - c. Símbolo de Adição: “+”.
  - d. Símbolo de Multiplicação: “\*”.
  - e. Número: “60”.

A análise hierárquica, também conhecida por análise gramatical ou análise sintática, consiste no “agrupamento dos tokens do programa fonte em frases gramaticais, que são utilizadas pelo compilador, a fim de sintetizar uma saída.” (Aho, Sethi & Ullman, 1995). Frequentemente, é

utilizado um tipo especial de árvore, chamado de árvore sintática, na qual cada nó representa uma operação e o filho de um nó representa o argumento da operação. O exemplo abaixo, adaptado de Aho, Sethi & Ullman (1995), demonstra a árvore gramatical da sentença utilizada no exemplo logo acima.

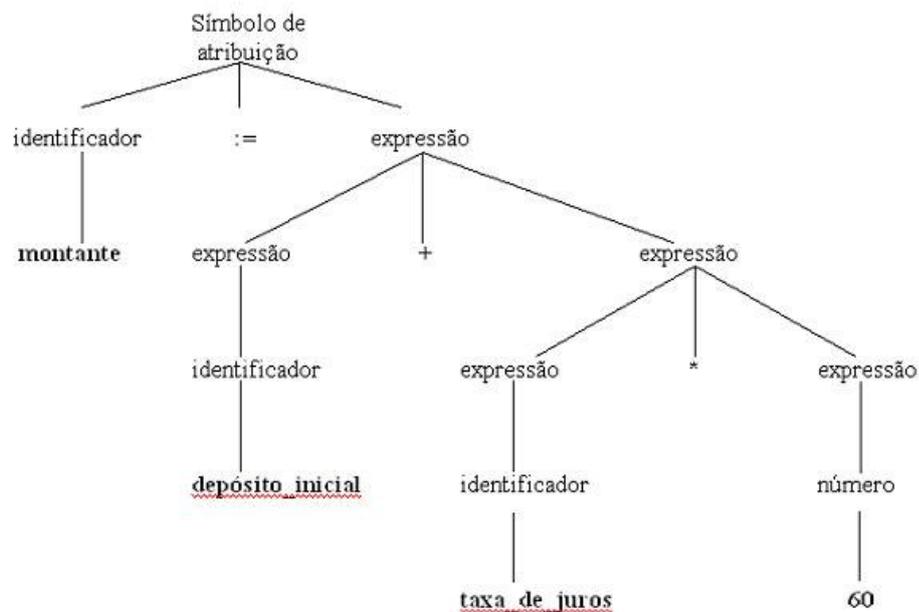


Figura 3: Árvore Gramatical

Adaptada de: Aho, Sethi & Ullman (1995)

A fase de análise semântica utiliza-se da estrutura hierárquica provida pelo analisador sintático a fim de identificar operandos e operadores. Nesta fase são verificados os erros semânticos no programa fonte e são capturadas informações de tipo para a fase subsequente de geração de código. Segundo Aho, Sethi & Ullman (1995), um importante componente da análise semântica é a verificação de tipos. Nela o compilador checa se cada operador recebe os operandos que são permitidos pela especificação de linguagem fonte. Num exemplo fictício a gramática implementada no compilador determina que os operandos do operador de adição (+) só podem ser do tipo inteiro. Quando o programa fonte submete ao compilador uma adição que envolve operandos do tipo decimal, o compilador retornará um erro, uma vez que o tipo de operando deveria ser inteiro.

Uma outra função importante exercida por um compilador é o registro de identificadores<sup>10</sup> usados no programa fonte e a coleta de informações sobre seus respectivos atributos. Este registro e coleta são efetuados na tabela de símbolos. Quando o analisador léxico detecta um identificador, instala-o na tabela de símbolos. Na tabela de símbolos são armazenadas informações como o tipo do identificador (inteiro, real, ponto flutuante, etc.), escopo (onde o identificador é válido no programa), nomes de procedimentos e funções, número e tipo dos argumentos dos procedimentos e tipo de retorno do procedimento.

As fases de um compilador são completadas com o gerador de código intermediário, o otimizador de código e o gerador de código alvo. Estas últimas três fases visam basicamente performance computacional e geração de código de máquina. Por isso as últimas três fases têm menor importância para os fins pretendidos por esta dissertação.

O arranjo completo de um compilador pode ser verificado na figura abaixo:

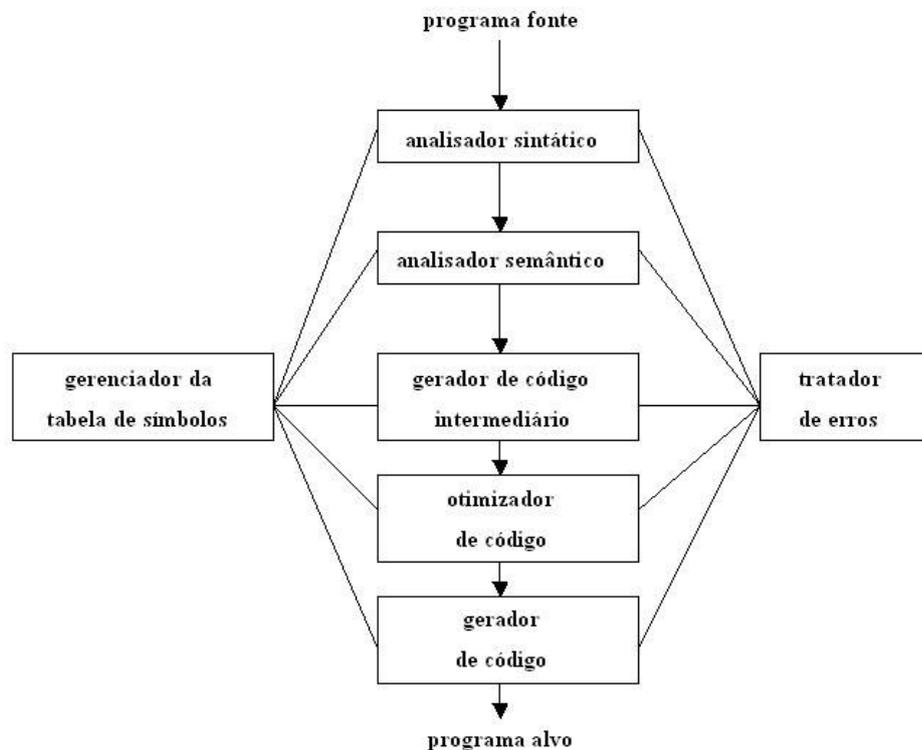


Figura 4: Fases de um compilador  
Adaptada de: Aho, Sethi & Ullman (1995)

<sup>10</sup> Os identificadores são comumente referenciados por programadores e analistas de sistemas de informação por “variáveis”.

### 3.3 – Conceitos Gramaticais do Português Contemporâneo do Brasil

Começamos na morfologia com os conceitos de letra e alfabeto. Conforme propõem Cunha & Cintra (2001), “Para reproduzirmos na escrita as palavras de nossa língua, empregamos um certo número de sinais gráficos chamados letras. O conjunto ordenado de letras de que nos servimos para transcrever os sons da linguagem falada denomina-se alfabeto.”.

A língua portuguesa é constituída de um conjunto infinito de frases. “Uma frase, por sua vez, pode ser dividida em unidades menores de som e significado – as PALAVRAS – e em unidades ainda menores, que apresentam apenas a face significativa – os FONEMAS.” (Cunha & Cintra, 2001). Não se pode confundir fonema com letra pois, enquanto o primeiro é um elemento acústico, o outro, como já observado, consiste em um sinal gráfico.

Às unidades mínimas de significação que compõem uma palavra, dá-se o nome de MORFEMA. Na palavra “ruas”, por exemplo, podem ser identificados dois morfemas. O primeiro, “rua”, forma por si só um vocábulo, realizando-se semanticamente como uma palavra isolada. O segundo, o sufixo “s”, que indica plural, nunca poderá realizar-se semanticamente como uma palavra autônoma. Os morfemas podem ser lexicais, também chamados de lexemas ou de radical, ou seja, aqueles que têm significação externa, fazendo referência intencional a fatos e objetos do mundo extralingüístico. Os morfemas também podem ser classificados como gramaticais, pois derivam das relações e categorias levadas em conta pela língua. Exemplificando, são morfemas gramaticais as preposições, artigos, e as marcas de feminino (“a”) e plural (“s”), pois esses têm função gramatical, não fazendo referência a objetos ou fatos exteriores ao mundo extralingüístico.

As classes das palavras podem ser estabelecidas através da classificação dos morfemas efetuada acima. “São morfemas lexicais os substantivos, os adjetivos, os verbos, e os advérbios de modo. São fonemas gramaticais os artigos, os pronomes, os numerais, as preposições, as conjunções e os demais advérbios, bem como as formas indicadoras de número, gênero, tempo, modo ou aspecto verbal.” (Cunha & Cintra, 2001).

Uma frase pode conter uma ou mais orações. Segundo Cunha & Cintra (2001), contém apenas uma oração quando apresenta uma só forma verbal, clara ou oculta, ou duas ou mais formas verbais, integrantes de uma única locução verbal. Uma locução verbal consiste no conjunto

formado por um verbo auxiliar e um verbo principal. Contém mais de uma oração quando há na frase mais de um verbo, claro ou oculto, seja na forma simples, seja na locução verbal. Já um período consiste na frase organizada em orações. O período pode ser simples, quando constituído de uma só oração, ou composto, quando formado por duas ou mais orações. O período sempre termina com uma pausa definida, que se marca na escrita com ponto, ponto de exclamação, ponto de interrogação, reticências e, algumas vezes, com dois pontos.

Uma oração pode ainda ser dividida em sintagmas. Tomando o exemplo fornecido por Cunha & Cintra (2001), na oração: “Este aluno obteve uma boa nota”, pode-se distinguir duas unidades maiores: o sujeito, “Este aluno”, e o predicado “obteve uma boa nota”. Nota-se que o sujeito é formado por duas palavras, um determinante<sup>11</sup> e um nome. O núcleo desta unidade é o nome “aluno”. Toda unidade que tem por núcleo um substantivo recebe o nome de sintagma nominal<sup>12</sup>. Vários sintagmas nominais podem ocorrer em uma única oração. O outro sintagma nominal da oração acima, por exemplo, é “uma boa nota”. Já o sintagma verbal constitui o predicado. Nele há sempre um verbo que, quando significativo, corresponde ao núcleo do sintagma verbal. Os sintagmas verbais podem ser complementados por sintagmas nominais e modificados por advérbios ou expressões adverbiais (MOD). O exemplo abaixo descreve a frase que serviu de exemplo logo acima, através de uma árvore sintática.

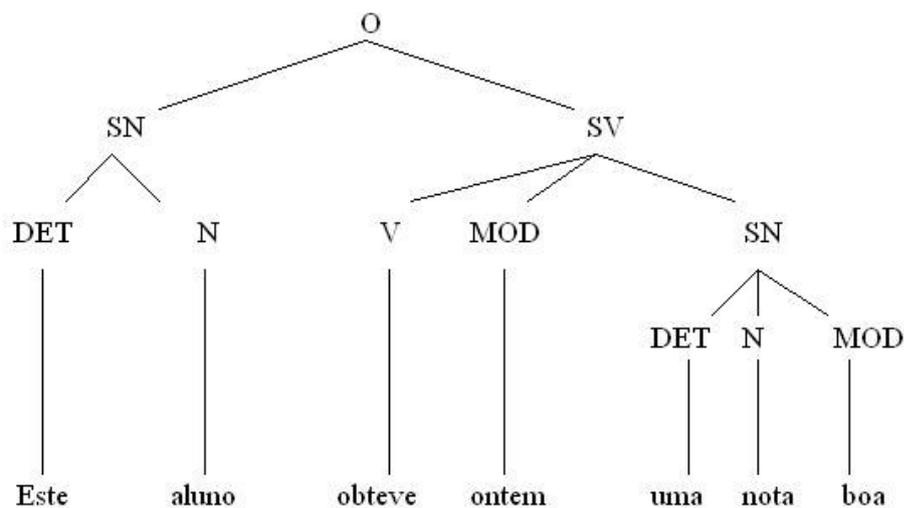


Figura 5: Árvore Sintática

Adaptada de: Cunha & Cintra (2001)

<sup>11</sup> Os artigos, os numerais e os pronomes adjetivos são considerados determinantes (DET).

<sup>12</sup> Além dos determinantes, o nome, núcleo de um sintagma nominal, pode vir precedido de um modificador (MOD), que são os adjetivos ou expressões adjetivas.

Outros tipos de sintagmas também podem ser encontrados na literatura. Os sintagmas preposicionais são formados de uma preposição e um sintagma nominal. Conforme Manosso (2002), o sintagma adjetivo, por sua vez, é formado por um sintagma adjetivo simples (aquele com apenas um adjetivo) ou então, pela concatenação de dois ou mais sintagmas adjetivos simples que se relacionam dois a dois por sintagma conectivo. Já o sintagma adverbial é formado por advérbio ou então, pela concatenação de dois ou mais advérbios que se relacionam dois a dois opcionalmente por sintagma conectivo. Por fim, o sintagma conectivo medeia a relação entre duas frases ou entre dois itens enumerados de um mesmo sintagma. O sintagma conectivo pode ser formado por conectivo (geralmente “e”) ou então, pelo morfema pausa (vírgula).

### 3.4 - Gramática de Valências para o Português

Na Química, valência é a capacidade que um átomo de um elemento tem de se combinar com outros átomos, capacidade essa que é medida pelo número de elétrons que um átomo pode dar, receber, ou compartilhar de forma a constituir uma ligação química (WIKIPEDIA, 2006b). Na molécula da água ( $H_2O$ ), por exemplo, são necessários dois átomos de Hidrogênio (H), que têm valência igual a um, para constituir uma ligação química com o Oxigênio (O), que tem valência igual a dois. Alguns elementos químicos, entretanto, tem valência igual a zero, ou seja, não se combinam com outros elementos para constituir uma molécula. Esse é o caso dos gases nobres, como o Hélio (He).

O exemplo da Química serve como uma analogia para introduzirmos a gramática de valências para o português, proposta pelo lingüista e dicionarista Francisco Borba em Borba (1996). Na referida obra, o autor resume sua intenção com as palavras abaixo:

Na língua, há palavras lexicais (substantivos, adjetivos, verbos) que são semanticamente auto-suficientes e, por isso, não precisam de nenhuma outra para se realizar plenamente. É o caso de tigre, bom, pedra, nevar, por exemplo. Há outras que são semanticamente incompletas e, por essa razão, precisam ligar-se a outras para adquirir um valor semântico completo, como, por exemplo, doação, transpor, obediente. As palavras que dão a outras uma completude de sentido recebem o nome de argumentos.

Valência é, então, o número de argumentos implicados por uma determinada palavra lexical. Uma gramática de valências descreve o número de argumentos que cada uma das palavras de uma classe pode ter, que classes de palavras (substantivo, adjetivo, etc.) podem preencher cada um dos argumentos, que argumentos devem ser ou não introduzidos por preposição, que categorias devem ter as palavras que se combinam (por exemplo, humano, abstrato, etc.).

Borba (1996, p. 16-17) diferencia uma gramática de constituintes da sua gramática de valências. Segundo o autor, uma gramática de constituintes se ocupa com a análise de estruturas tentando descobrir como um constituinte se encaixa noutro ou pertence a outro. Além disso, se ocupa também com a decomposição do enunciado por meio de regras (de reescrita) que mostram como os constituintes se aproximam e se hierarquizam. Já a gramática de valências procura detectar relações de dependência entre categorias básicas que (co) ocorrem num determinado contexto (ex: A depende de B, se a presença de A pressupõe B).

Além disso, a gramática de valências toma como nuclear um elemento oracional (o verbo) e demonstra como os demais se dispõem em torno dele através de relações de dependência.

Segundo Borba (1990) e Borba (1996), toda oração pode ser subdivida em dois níveis: o esquema profundo e o esquema superficial. O esquema profundo é subjacente e descreve as relações básicas como conexão e número de argumentos, bem como a ordem, restrições de coocorrência e relações semânticas fundamentais. No esquema superficial se decide a combinatória dos morfemas, levando-se em consideração a concordância e a regências entre eles. O verbo roubar, por exemplo, é descrito no esquema profundo como verbo de três lugares. Na frase “Daniel roubou o carro de Patrícia”, “Daniel” tem o papel semântico de agente, “o carro” tem o papel semântico de objetivo, e “de Patrícia” tem o papel semântico de origem. O esquema superficial determina o seguinte arranjo: o primeiro argumento é o sujeito, expresso por nome humano, e os outros dois na posição de complemento, ambos apagáveis, sendo que o argumento com papel semântico de origem precedido pelas preposições *a* ou *de*.

Borba (1990) explica que, ao analisar sintaticamente uma oração, o esquema superficial é o único observável diretamente pelo analista. O primeiro passo, então, é observar a estrutura superficial da oração, levando em consideração alguns expedientes que podem dar uma aparência irregular ao enunciado, são eles: alçamento, rebaixamento, apagamento, truncamento, transposição, dissimilação, coordenação e condensação. A análise desses expedientes torna-se uma ferramenta fundamental para se detectar a estrutura primitiva da oração. Parafraseando Borba (1990) e utilizando seus exemplos, seguem as descrições dos expedientes:

O alçamento consiste na promoção de um constituinte da frase, levando-se em consideração que existe uma hierarquia oracional que pode ser descrita do mais importante para o menos importante dessa maneira: sujeito, complementos e outros constituintes (adnominais, locativos, instrumentais). No exemplo abaixo o instrumento (faca) é alçado a posição de sujeito na segunda frase.

- a) Leo cortou o salame com uma faca.
- b) Faca laser corta salame em fatias bem finas.

O rebaixamento é exatamente o contrário do alçamento. O constituinte da frase é rebaixado a um nível hierárquico mais baixo.

- a) O sangue escorria da ferida.
- b) Da ferida escorria sangue.

O apagamento consiste no cancelamento de um componente da oração, quando este pode ser recuperado através de informações oriundas do contexto, ou da própria morfologia, sintaxe e semântica empregada no enunciado. Na primeira frase do exemplo logo abaixo, é possível inferir a partir do verbo compor que o complemento apagado trata-se de músicas (sinfonias). Já na segunda frase, é possível recuperar o componente apagado a partir da própria estrutura morfológica do verbo.

- a) Carlos Gomes compunha desde criança.
- b) Wanda trabalha o dia inteiro só empacotando.

O truncamento consiste na supressão de componentes, ficando a frase reduzida a seus componentes semanticamente essenciais. Borba (1990) deixa claro que o truncamento é semelhante ao apagamento, porém ele prefere diferenciar um do outro dando ao apagamento do complemento oracional o nome de truncamento. Com os exemplos abaixo, Borba (1990) demonstra que os verbos “ousar” e “proibir” exigem um complemento oracional. Nesse caso, a terceira frase corresponde ao truncamento da primeira, e a quarta frase corresponde ao truncamento da segunda.

- a) Djalma ousou externar uma opinião contrária.
- b) Jânio proibiu que as garotas usassem biquínis nas praias.
- c) Djalma ousou uma opinião contrária.
- d) Jânio proibiu biquíni nas praias.

A transposição consiste na transferência de traços semânticos exigidos para um determinado argumento na matriz valencial de um item léxico, para a palavra que ocupa este argumento. No primeiro exemplo, o verbo vender exige um complemento não-animado. De fato o complemento Sócrates é submetido a uma espacialização do expediente da transposição, a desanimação, já que o que está se vendendo é o passe de Sócrates e não o próprio. Já no

segundo exemplo, o verbo mentir exige um sujeito humano, o que faz com “jornais” seja submetido à humanização, outra espacialização do expediente da transposição.

- a) O Corinthians vendeu Sócrates.
- b) Há jornais que mentem muito.

A dissimilação diz respeito às orações onde o verbo tem o mesmo radical do complemento. Neste caso, o verbo é alterado para que os radicais passem a ficar diferentes.

- a) Beber uma bebida.
- b) Tomar uma bebida.
- c) Pensar um pensamento.
- d) Ter um pensamento.

A coordenação e a condensação ocorrem com os verbos que tem complementos simétricos, ou seja, aqueles que tem argumentos de mesma subclasse e nenhuma exigência quanto à ordem deles. Os dois primeiros exemplos descrevem uma coordenação, que consiste na troca de ordem dos argumentos. Os dois últimos exemplos indicam uma condensação, onde “leite” e “farinha” são condensados em “ingredientes”.

- a) Toni divorciou de Leia em maio.
- b) Leia divorciou de Toni em maio.
- c) Misture o leite e a farinha.
- d) Misture os ingredientes.

Após a análise dos expedientes oracionais, é importante salientar que a gramática de valências de Borba é baseada na Gramática da Dependência do francês Lucien Tesnière (TESNIÈRE, 1959), na teoria das Restrições Semântico-Selecionais, e na gramática de casos de Charles J. Fillmore (FILLMORE, 1968), as quais detalha-se a seguir.

### 3.4.1 – Teoria da Valência

Esta teoria foi inicialmente proposta por Tesnière (1959), que considera a frase como um conjunto organizado formado por palavras e pelas conexões que se estabelecem entre elas (BRÄSCHER, 1999).

A teoria de Tesnière parte do verbo como núcleo oracional tomando-o como uma espécie de pólo imantado, capaz de atrair um determinado número de actantes<sup>13</sup>, comportando um número variável de pontos de atração capazes de manter esses actantes sob sua dependência (BORBA, 1996, p. 19). É o verbo que determina quais elementos a frase tem que conter (WELKER, 2005).

Para Tesnière (1959 *apud* BRÄSCHER, 1999), valência é restrita ao verbo sendo o “número de actantes que o verbo é capaz de reger”. Esta definição é mais restrita do que a proposta por Borba (1996, p. 20), que amplia o conceito ao propor que valência é uma “(...) propriedade que uma classe de elementos, de poder ligar-se com classes específicas de outros elementos sendo que esta mesma propriedade faz que esta classe se distinga de outras de mesmo nível sintagmático.”

Por isso, para Borba (1996), a valência não é privilégio apenas dos verbos. Determinados adjetivos (Ex: obediente) e substantivos (Ex: sogra) também exigem a presença de actantes para se tornarem semanticamente plenos. As demais classes gramaticais, como determinantes, pronomes, e relacionantes (preposições) não têm peso semântico absoluto ou relativo. São, apenas, instrumentos de realização dos lexicais.

As palavras semanticamente plenas, ou seja, aquelas que não necessitam de complementos, como *tamanduá*, *pedra* e *chover*, têm valência zero ( $V_0$ ). As palavras, como *sogra*, que

---

<sup>13</sup> Entende-se por Actante o argumento que complementa semanticamente um verbo, adjetivo ou substantivo. Ou ainda, conforme Tesnière (1959 *apud* BRÄSCHER, 1999), “pessoas ou coisas que participam de alguma maneira no processo”.

exigem um argumento, têm valência um ( $V_1$ ). Os itens lexicais que exigem dois argumentos, como os verbos que indicam movimento (*ir, vir, chegar, subir, descer*), têm valência dois ( $V_2$ ), e assim por diante até a valência quatro ( $V_3, V_4$ ). Dessa maneira, existem itens lexicais aivalentes, aqueles de valência zero, e itens lexicais valenciais, ou seja, aqueles que exigem argumentos.

Borba (1996) sugere utilizar o termo “valência” em três níveis, são eles:

a) Valência Lógica – Também chamada por Borba (1996) de valência quantitativa, ou lógico-semântica. Diz respeito apenas ao número de argumentos que um predicado pode ter. Um item lexical pode ter até quatro argumentos, podendo ser, portanto, monovalente, bivalente, trivalente e tetravalente. Segundo Welker (2005), Bondzio (1971) introduziu a o conceito de valência lógica ao notar, por exemplo, que o verbo *vigiar* implica alguém que vigia e algo ou alguém que está sendo vigiado. Bondzio (1971 *apud* WELKER, 2005) esclarece:

Essas relações são representadas na lógica dos predicados mediante os conceitos ‘predicado’, ou ‘functor’, ‘lugares vazios’ e ‘argumentos’ do predicado. Assim, o verbo ‘vigiar’ pode ser descrito como predicado (functor) com dois lugares vazios que podem ser ocupados pelos respectivos argumentos.

b) Valência Sintática – “Esta é a valência ‘tradicional’ – isto é, idealizada por Tesnière (...)” (WELKER, 2005). De acordo com Borba (1996, p. 20), a valência sintática diz respeito às “características dos actantes, do preenchimento das casas vazias por determinadas classes com determinadas propriedades morfológicas” (Ex: um determinado predicado exige um argumento que seja um nome, e um segundo argumento que seja um sintagma preposicional). Borba (1996, p. 20) continua dizendo que este nível trata da função sintática (Substantivo, Complemento, Especificador, Adjunto e Circunstante); da propriedade sintática (pronominalização, ordem, etc.), das classes que preenchem os argumentos (Nome, Adjetivo, Advérbio, Oração, etc.), e das propriedades morfossintáticas dos itens que preenchem os argumentos.

c) Valência Semântica – Para Borba (1996, p. 21), a valência semântica é deduzida da observação das regularidades ou compatibilidades das unidades que operam nas

seqüências. Este nível se preocupa com as características categoriais (ex: +anim, +hum, +cont, etc.); com as funções temáticas (papéis) que cada actante tem em um determinado predicado (agente, causativo, beneficiário, etc.); e com as restrições seletivas que determinam quais classes/subclasses de itens que, funcionando como predicado, concorrem ou se excluem com as classes/subclasses dos itens que preenchem os argumentos (Ex: Em “Pedro vai à igreja” – O predicado é um verbo de ação, que se relaciona com dois argumentos: um sujeito agente, e um complemento com traço “direção”).

### 3.4.2 – Gramática de Casos

A idéia de valência semântica introduz a gramática de casos proposta por Fillmore (1968). Fillmore (1968) desenvolveu um modelo de gramática cujo traço característico é o fato de considerar que a sentença, em sua estrutura básica, é formada por um verbo e um ou mais sintagmas nominais, cada um deles associado ao verbo por uma relação de caso particular. (TEIXEIRA, 1999 *apud* BRÄSCHER).

De acordo com Borba (1996, p. 27), a análise valencial deve, também, determinar as relações sintático-semânticas ou temáticas (funções ou papéis temáticos) que fazem parte da estrutura conceitual dos itens léxicos. Ainda segundo o autor, as relações temáticas se traduzem ou se fazem representar por um sistema de casos ou gramática de casos. “Da aproximação entre as estruturas conceituais resultam os papéis temáticos.” (BORBA, 1996 p. 52)

Borba (1996) segue Fillmore (1968), quando propõe que os casos, ou papéis temáticos, são:

Quadro 3 – Casos ou Papéis Temáticos

Caso	Descrição
Agentivo (Ag)	O que desencadeia uma atividade, sendo origem dela ou controlador.
Experimentador (Ex)	O que experimenta um evento psicológico.
Beneficiário (B)	O que se beneficia. Marca o destinatário de posse (perda ou ganho).
Locativo (L)	Marca o lugar onde.
Direcional (Dr)	Especialização de Locativo. Movimento para.

Percurso (Pr)	Especialização de Locativo. Movimento através de.
Objetivo (Ob)	É o afetado por aquilo que o verbo indica.
Instrumental (I)	Exprime uma causa indireta tendo como traços básicos a atividade e o fato a ser controlado
Causativo (Ca)	Provoca um efeito ou que desencadeia algo.
Meta (M)	O ponto de chegada.
Origem (Or)	O ponto de partida.
Resultativo (R)	É um efetuado, algo que passa a existir.
Temporal (Tp)	Localização no tempo.
Comitativo (Co)	Associação (Romeu morreu <i>com Julieta</i> )

FONTE: BORBA, 1996.

As frases abaixo, demonstram os casos que surgem a partir da aproximação dos argumentos com o predicado:

a) Paulo espancou o gato com um porrete.

Ag P Ob I

b) João ganhou um carro de seu pai.

B P Ob Or

c) O menino sentiu a falta do pai em Brasília.

Ex P Ob L

“O nome de cada caso refere-se ao valor semântico que expressa (...) esse valor semântico não está ligado a um item léxico em si, senão a uma relação, ou seja, o semântico é uma consequência do sintático.” (BORBA, 1996, p. 30)

Por exemplo, o item lexical *cachorro* só será agentivo ao se relacionar com outro item com o traço *+ativ*, que funcione como núcleo do predicado, tal como *late*: O cachorro [Ag] late [P, *+ativ*].

Borba (1996, p. 35) define, ainda, a diferença entre casos profundos e casos superficiais. Um caso é profundo quando faz parte da valência do verbo, enquanto que o caso superficial não o faz. Para determinar quem faz parte da estrutura valencial de um verbo, diferenciando casos

profundos de casos superficiais, é necessária uma rigorosa identificação dos traços semânticos dos componentes da oração. Dessa maneira, Borba efetua a seguinte orientação:

1º - Casos exclusivamente profundos são aqueles que descrevem relações causais subjacentes, regulares e implícitas, podendo realizar-se ou não efetivamente: Ag, Ex, Ca. Se algum deles aparecer no esquema superficial, é realização do correspondente no caso profundo.

2º - Casos exclusivamente superficiais são realizações semânticas resultantes de combinatória léxica e atuando nas orações. São eles: M e I.

3º - Existem casos que tanto podem ser profundos como superficiais e, quando aparecem no esquema superficial, podem ou não fazer parte da valência verbal. São eles: B, Ob, Or, R, L, Tp.

O núcleo do predicado “*sair*”, exposto na frase abaixo, tem uma matriz valencial que exige um Ag no primeiro argumento e um Or no segundo argumento. O caso Tp, apenas complementa a oração, não fazendo parte da valência do verbo.

a) João Pedro saiu da loja na segunda-feira.

Ag            P        Or        Tp

### 3.4.3 – Restrições Semântico-Selecionais

De acordo com Androutopoulos & Dale (2005), o termo restrição seletional refere-se ao conjunto de restrições impostas aos participantes de construções linguísticas. Este conceito pode ser esclarecido ao analisarmos os seguintes exemplos:

a) Daniel bebeu o carro.

b) A menina comeu a televisão.

c) Magda bebeu café.

Enquanto a terceira frase não causa espanto, evidentemente, o senso comum já indica que carro e televisão, nas primeiras frases, não são objetos bebíveis e comestíveis respectivamente. Existe uma restrição Semântico-Selecionais (semântica ou pragmática)

imposta pelo verbo comer que impõe que o objeto afetado pelo verbo tenha a característica de ser comestível. Da mesma maneira, para o verbo beber, entende-se a necessidade de um objeto bebível. É importante verificar que, sintaticamente, as duas primeiras frases do exemplo imediatamente acima são perfeitas, ou seja, não existe relação entre a violação de uma restrição seletional com o fato da sentença ser gramatical ou não.

De acordo com Sailer (2004), desde Chomsky (1965), o status das restrições semântico-selecionais vem sendo tema de debates, especificamente se estas restrições estão no campo da pragmática ou no campo da semântica. Esta discussão é justificável e, ao mesmo tempo, expõe uma das principais limitações da utilização das restrições semântico-selecionais. O contexto é quem determina as restrições válidas. Por exemplo, supondo que uma rede de lanches rápidos crie uma refeição chamada "televisão", então o exemplo "b" passa a ser válido. Outra limitação exposta por Androutsopoulos & Dale (2005) é a questão das metáforas e de sentenças que contenham negações. Quando diz-se "Daniel não bebeu o carro", então a sentença, que na forma afirmativa causava estranheza, passa a ser válida. Conforme Amaral (2003), "Por mais que descrevam os diferentes sentidos de um mesmo verbo, incluindo usos metafóricos, sempre haverá um novo uso sendo criado ou ainda não descrito."

Apesar das limitações, conforme Androutsopoulos & Dale (2005), a utilização de restrições semântico-selecionais tem sido de muita utilidade em aplicações práticas relacionadas ao processamento automático de linguagem natural. Segundo este autor, a utilização das restrições semântico-selecionais tem papel importante na resolução de ambigüidades.

A figura abaixo, extraída de Amaral (2003), demonstra como a teoria das restrições semântico-selecionais pode ser integrada com a gramática de valências de Borba (1996). Na primeira coluna encontra-se a classe sintático-semântica do verbo comer. Na segunda as definições das acepções. A terceira e quinta coluna trazem o papel temático que surge da relação entre o argumento e o predicado. A quarta e a sexta coluna indicam a restrição semântico-selecional para o argumento em questão. O sinal de negativo na frente da restrição, indica negação (Ex.: -animado → não animado).

Classe	Valor (Definição)	SUJEITO		COMPLEMENTO		
		Caso	Descrição Semântica	Caso	Descrição Semântica	Desc. Sint.
Ação	Engolir para se alimentar; ingerir.	Ag	+animado	Obj	Alimento sólido	SN
Ação-processo	Corroer; destruir	C	-animado elemento corrosivo; doença	Obj	-animado	SN
Ação-processo	Gastar; dilapidar	Ag	nome	Obj	Tempo; bens; acontecimento	SN
Ação-processo	Suprimir; omitir	Ag	+animado	Obj	Elementos lingüísticos	SN
Ação-processo	Eliminar	Ag	+animado	Obj	Peças de jogo de tabuleiro	SN
Ação	Provar; experimentar	Ag	+animado	Obj	Alimento sólido	SP

Figura 6: Restrições Semântico-Selecionais e a Gramática de Borba

Adaptada de: Amaral (2003)

### 3.4.3– Valência Verbal

“A valência verbal tem por objetivo básico observar e determinar o comportamento do verbo no âmbito da frase.” (BORBA, 1996, p.46)

A análise da valência dos verbos começa, naturalmente, pela valência lógica ou quantitativa, que trata apenas do número de argumentos exigidos pelo verbo. Dessa maneira, existem verbos que são avalentes ( $V_0$ ), ou seja, não exigem argumentos. Os verbos meteorológicos são exemplos típicos de verbos avalentes: *chover*, *nevar*, *trovejar*. Os verbos de valência igual a um ( $V_1$ ), tais como os verbos que indicam atividade fisiológica (Ex: *escarrar*, *espirrar*, *urinar*, etc.) e verbos que indicam atividades dos seres (Ex: *sonhar*, *dormir*, *andar*, *galopar*, etc.) exigem apenas um argumento. Os verbos de valência igual a dois ( $V_2$ ), exigem dois argumentos. Alguns exemplos de verbos bivalentes são: verbos que indicam movimento direcional (Ex: *ir*, *vir*, *subir*, etc.); verbos que indicam afetividade (Ex: *amar*, *odiar*, *namorar*, *desprezar*, etc.). Os verbos trivalentes ( $V_3$ ) são aqueles que exigem três argumentos, tais como

verbos de atribuição (Ex: *dar, doar, emprestar*, etc.). Os verbos tetravalentes (V<sub>4</sub>) exigem quatro argumentos. Os verbos que indicam transferência de localização (Ex: *transferir, passar, levar, arrastar*, etc.), bem como os que indicam transposição (Ex: *traduzir, verter, transpor*, etc.), são exemplos típicos de verbos de valência igual a quatro.

Seguindo os níveis da análise valencial, passamos à análise da valência sintática dos verbos. “A primeira tarefa da valência sintática é identificar as classes que preenchem os argumentos.” (BORBA, 1996, p. 47).

Geralmente os nomes preenchem as casas vazias dos verbos, entretanto, eventualmente, alguns advérbios modais (Ex: SN + *proceder* + “bem” ou “mal”) e também orações (Ex: SN + *proibir* alguém de –O) podem funcionar como actante do verbo.

Borba (1996, p. 47-48) propõe uma série de regras sintáticas gerais de estruturação dos verbos, à luz de sua valência quantitativa. Dessa maneira:

- Verbos monovalentes: Estes verbos têm a forma SN + V, ou seja, o actante está localizado fora do SV. Por isso, o argumento nunca é precedido por preposição (Ex: *morrer* – Ele morreu, *urinar* – O cão urinou).
- Verbos Bivalentes: A estrutura dos bivalentes é: SN<sub>1</sub> + V + SN<sub>2</sub>. Um dos argumentos dos verbos bivalentes faz parte da estrutura interna do SV e, por isso, pode ser introduzido opcionalmente por preposição. O outro argumento é externo ao SV, fazendo parte do SN (Ex: *conspirar* – A oposição *conspira contra* o governo).
- Verbos Trivalentes: A estrutura dos verbos de três actantes é a seguinte: SN<sub>1</sub> + V + SN<sub>2</sub> + Sprep. Nestes verbos, dois argumentos fazem parte da estrutura interna do SV, sendo que um deles exige obrigatoriamente ser introduzido por preposição (Ex: *enviar* – Daniel *enviou* condolências *a* Jorge).
- Verbos Tetravalentes: Os verbos de quatro actantes terão três argumentos internos ao SV sendo que dois destes precedidos por preposição (Ex: *traduzir* – O tradutor *traduziu* a carta do português para o mandarim).

O próximo passo na análise da valência dos verbos remete à valência semântica. “A valência semântica diz respeito, em primeiro lugar, às propriedades semânticas dos verbos, ou seja, sua

subcategorização em traços, o que está diretamente implicado nas restrições seletivas”. (BORBA, 1996, p. 49)

Borba (1996, p. 51) continua sua exploração sobre a valência semântica dos verbos:

As propriedades semânticas do verbo importam na medida em que permitem estabelecer regras de seleção. São elas que controlam toda essa diversificação. Por exemplo, para ‘subir’, podemos estabelecer regras do tipo:

Rni: Com sujeito +anim e com complemento equivale a ‘trepar’, ‘galgar’, ‘grimpar’.

Rnj: Com sujeito +concreto e designativo de líquido, equivale a elevar-se, atingir nível (mais alt, etc., etc., etc. ...)

A valência semântica tem estreita ligação com a gramática de casos. “Nos anos setenta, vários autores, percebendo semelhanças entre as idéias de Fillmore e de Tesnière – pois ambos viam no verbo aquele elemento que determina a estrutura básica da frase – incluíram os casos profundos na teoria da valência, denominando essa parte de valência semântica.” (WELKER, 2005)

Ou seja, os papéis temáticos surgem da aproximação ou associação entre o verbo, núcleo do predicado, e um nome. Por exemplo, da relação entre o verbo *subir*, que tem o traço semântico +mov, com o nome *macaco* (+anim), resulta a função de agente (Ag) para o nome e a classe de ação para o verbo. Da mesma maneira, para *cortar*, outro verbo da classe ação, surgem os seguintes casos: alguém que corta (Ag), algo que cortado (Ob), e um instrumento (I).

“Isso significa que, pela identificação das funções temáticas, se consegue estabelecer um sistema de casos para os nomes e uma taxonomia sintático-semântica para os verbos. De posse de um sistema de casos montam-se esquemas como: entrar (Ag, L); amar (Ex, Ob); chorar (Ag); dar (Ag, Ob, B); morrer (Ob); saber (Ex, Ob); horrorizar (Ca, Ob); vacilar (Ex); herdar (B, Ob, Or).” (BORBA, 1996, p. 52)

“Como acontece com a valência lógica, não há unanimidade quanto ao estabelecimento dos quadros de casos dos diversos verbos. Borba (1996a, p.31) reconhece que a ‘identificação dos casos por seus traços constituintes às vezes se torna sutil ou difícil’.” (WELKER, 2005)

Como já dito, um dos resultados da aproximação entre um verbo e um nome diz respeito à classificação sintático-semântica dos verbos. Existem quatro classes, a saber: verbos de ação, verbos de processo, verbos de ação-processo e verbos de estado. Conforme Borba (2002, p. VII):

- *Os verbos de Ação expressam “uma atividade associada a um sujeito agente, ou seja, aquele que, por si mesmo, desencadeia uma atividade, física ou não, sendo origem dela ou seu controlador (o galo canta, o pássaro voa).”*
- *Os verbos de processo expressam “um evento ou sucessão de eventos cujo suporte está num sujeito paciente ou afetado por aquilo que o predicado indica (A planta cresce; O gato morreu); experimentador ou que expressa uma experiência ligada a uma disposição mental (Marina lembra-se com saudade da infância), uma sensação (O cururu viu a cobra), uma emoção (Lina sente a morte da avó); beneficiário ou que é sede de transferência de posse ou destinatário de um benefício (Lúcia herdou da avó um gato siamês).”*
- *Os verbos de ação-processo expressam “uma mudança de estado ou condição levada a efeito por um sujeito agente, causativo ou instrumental, atingindo um complemento que é, então, um afetado ou efetuado: Ana abriu a porta; O raio partiu a árvore; Uma velha tesoura cortou o umbigo do nenê; Maria tricou uma blusa; O medo afugentou o rapaz”.*
- *Os verbos de estado são aqueles “cujo sujeito é mero suporte de propriedades (estado > Uma estrada margeia o lago; ou condição > Meu vizinho tem fazendas em Goiás) ou, então, é experimentador delas (João amava Maria; Nuno sabe tocar oboé).”*

De posse das características arroladas acima, é possível explicitar a valência quantitativa, sintática e semântica dos verbos através das matrizes valenciais. Por exemplo, a matriz valencial do verbo *assobiar*, verbo de ação que exige apenas um actante, será descrita por “*assobiar* (Ag)”, ou, de maneira mais genérica, pela classe do verbo com seus respectivos argumentos: Assobiar – Ação (Ag). É importante salientar que as diferentes acepções de um mesmo verbo podem ter matrizes valenciais diferentes. Lista-se abaixo alguns exemplos de matrizes valenciais <sup>14</sup>:

- a) Assobiar – Ação (Ag)
- b) Abdicar – Ação (Ag, Ob)
- c) Amar – Estado (Ex, Ob)

---

<sup>14</sup> Adota-se neste trabalho a vírgula para separar os argumentos de um verbo, ao invés do sinal de adição, conforme utilizado por Borba (1996). Esta escolha se dá propositadamente para aproximarmos a notação utilizada na expressão das matrizes valenciais da lógica de predicados.

- d) Chorar<sub>1</sub> – Ação (Ag) – (Acepção: derramar lágrimas)
- e) Chorar<sub>2</sub> – Ação (Ag, Ob) – (Acepção: lamentar, prantear)
- f) Abençoar<sub>1</sub> – Ação-Processo (Ag, Ob) – (Acepção: benzer, proteger)
- g) Abençoar<sub>2</sub> – Ação-Processo (Ob, L) – (Acepção: constituir proteção para)

As matrizes valenciais representam os casos profundos, ou seja, aqueles casos que fazem parte da valência verbal. Entretanto, Borba (1996, p.68) expõe uma preocupação com os diferentes usos de um verbo, principalmente com o “jogo metafórico e metonímio” que por vezes é utilizado (o bebê *chora* / o tronco da seringueira *chora* (verte) um leite grosso e pegajoso). Nestes casos as matrizes valenciais têm que ser revistas para cada utilização.

Os verbos funcionais (auxiliares e verbos-suporte), modais e substitutos não são abarcados pela valência verbal, uma vez que tais verbos não são plenos, ou seja, semanticamente não tem significação lexical e sintaticamente não ocupam o núcleo do predicado.

Um verbo é funcional quando se relaciona com outro verbo pleno, núcleo do predicado, ocupando uma posição periférica no interior do SV, e, quando tiver uma significação gramatical ou for suporte de categorias gramaticais. (BORBA, 1996, p.75).

Nas frases abaixo, retiradas de Borba (1996, p.75), a predicação é concentrada em *fugir*, *sair* e *saudar* sendo que *tinha* expressa tempo; *ia*, aspecto progressivo e *foi*, voz passiva. Os últimos três exercem a função verbo auxiliar. Um verbo auxiliar é reconhecido por ter função semântica restrita, apenas indicando categorias de tempo, voz e aspecto, e, por ligarem-se a verbos plenos por preposição ou não, sendo que o verbo principal estará sempre no participípio, gerúndio ou no infinitivo impessoal. De acordo com Cunha e Cintra (2001), “Os auxiliares mais freqüentes são *ter*, *haver* e *estar*.”

- a) Teresa tinha fugido de casa.
- b) O trem ia saindo.
- c) O papa foi saudado pela multidão.

Os verbos-suporte, ou verbalizadores, são verbos que, tornando-se vazios de sua significação léxica, compõem sintagmas verbais complexos introduzindo predicados nominais cujo núcleo é um nome ou adjetivo (BORBA, 1996, p.78). Seguem alguns exemplos de verbos-suporte:

- a) ter medo (temer)
- b) abrir falência (falir)
- c) ficar triste (entristecer)

Os modais constituem-se um predicado que rege um outro predicado para expressar modificação da relação entre o sujeito e o enunciado, traduzindo assim uma atitude subjetiva do falante em relação ao que ele comunica (BORBA, 2002, p. VII). Seguem alguns exemplos:

- a) João *deve* ser bom.
- b) O vendedor *podia* trapacear.
- c) *Mandei* que o encanador arrumasse a pia.
- d) A casa *está* por varrer.
- e) A brisa *fez* cair uma folha seca perto da cobra.

Os verbos substitutos são verbos semanticamente esvaziados que remetem a outra oração ou parte dela. Em português, o substituto por excelência é *ser*, seguido por *fazer* e eventualmente por *acontecer*. Sem alguns exemplos extraídos de Borba (1996, p.82-83):

- a) Você está falando de casamento? *É*.
- b) Bebê não é de fazer força. *É*, mas um dia ele vai ser dar mal.
- c) O preto só *fez* foi rir.
- d) Há muito tempo eu não entrava ali, e agora o *fazia* com certa curiosidade.
- e) Ele não pestaneja como *acontece* com todos os entrevistados.

#### 3.4.4– Valência Nominal

Por uma questão metodológica, Borba (1996) propõe que a análise da valência nominal se subdivida em duas partes: a análise dos nomes abstratos e a análise dos nomes concretos. Seguiremos a metodologia proposta pelo autor.

### 3.4.4.1 – Valência dos Nomes Abstratos

Baseado em uma longa análise das propostas de outros autores, Borba (1996, p. 95-96) adotar, como nos verbos, uma descrição em dois níveis, o esquema profundo e o esquema superficial, admitindo que todos os nomes abstratos podem ser descritos no esquema profundo pela estrutura Vsup + Na. Segundo o autor, todo sintagma nominal cujo núcleo é um nome abstrato resulta de nominalização<sup>15</sup>. Exemplos:

- a) A cabrita está no cio > O cio da cabrita.
- b) Pedro teve vontade de gritar > A vontade de gritar por (parte de) Pedro.
- c) Ana é bela > Ana tem beleza > A beleza de Ana.
- d) Os vampiros dançam > Os vampiros fazem dança > A dança dos vampiros.

Borba (1996) admite o paralelismo no esquema profundo entre os verbos, adjetivos e nomes abstratos, sendo que a interpretação dos últimos pode ser previsível a partir dos traços lexicais dos primeiros. Exemplos:

- a) Paulo educou Pedro.
- b) A educação de Pedro por Paulo.
- c) Pedro é educável por Paulo.

Borba (1996, p. 96) chama a atenção que a análise do paralelismo pode ser importante para a eliminação de ambigüidades. Em “a educação de Pedro” ocorre ambigüidade, por que se pode interpretar que Pedro é educado por alguém ou que, simplesmente, Pedro é uma pessoa educada.

Quantitativamente, os nomes abstratos são todos valenciiais, podendo-se realizar com um, dois, três ou quatro actantes. Seguem exemplos retirados de Borba (1996, p. 99):

- a) Nomes Abstratos com 1 Argumento: greve, cio, saúde, sofrimento, apetite.

---

<sup>15</sup> Segundo McCawley (1968) *apud* Borba (1996), uma nominalização é uma representação semântica derivada transformacionalmente de uma representação semântica subjacente. A palavra “educação”, por exemplo, é uma nominalização do verbo educar.

- b) Nomes Abstratos com 2 Argumentos: obediência, amor, ódio, diálogo, divórcio.
- c) Nomes Abstratos com 3 Argumentos: doação, atribuição, narração, afirmação.
- d) Nomes Abstratos com 4 Argumentos: tradução, transferência, mudança.

Com relação às matrizes sintáticas dos nomes abstratos, Borba (1996, p. 102) propõe:

- Nomes Abstratos com 1 Argumento: de N/O. Exemplo: Possibilidade de Chuva.
- Nomes Abstratos com 2 Argumentos: Prep N/O + Prep N. Exemplo: A confiança nos pais pelas crianças.
- Nomes Abstratos com 3 Argumentos: de N + Prep N + por N. Exemplo: A narração do assalto à polícia pelo garoto.
- Nomes Abstratos com 4 Argumentos: de N + de N + para N + por N. Exemplo: A tradução do poema do francês para o português por Denise.

Os nomes abstratos seguem os verbos também no que diz respeito às classes sintático-semânticas. Borba (1996) justifica tal constatação em função de que os nomes abstratos se comportam ora de um modo estático, ora de modo dinâmico. Dessa maneira temos nomes abstratos que podem indicar estado, ação ou processo. Os exemplos abaixo demonstram exemplares de cada uma das classes arroladas:

- a) Abstratos de Estado: negrura, amargura, bravura, braveza, baronato, condado, irritabilidade, tranquilidade.
- b) Abstratos de Ação: tosse, fuga, fala, gemido, entrada, saída, ferroadada.
- c) Abstratos de Processo: gozo, queda, sonho, sono, lembrança.

Borba (1996, p. 115) nota que há uma vinculação entre as classes sintático-semânticas dos nomes abstratos com suas respectivas matrizes de valência semântica. Quando um nome abstrato pertence a mais de uma classe, é a ligação do caso que decide a classificação. Nomes como chegada, corrida e alfinetada, por exemplo, serão classificados como ação ou processo de acordo com o caso a que pertence o argumento com o qual se relacionam. Para esses nomes, se o argumento se tratar de um caso agentivo, então o nome é um abstrato de ação, caso contrário, indica processo.

- a) A chegada do ministro. / A chegada do inverno.

- ação Ag processo Ob
- b) A corrida dos atletas. / A corrida dos rios para o mar.
- ação Ag processo Ob
- c) A velha deu duas alfinetadas na nora. / O velho sentia umas alfinetadas no rim.
- Ag ação Ex processo

Já os abstratos de estado geralmente se ligam a actantes do tipo experimentador, conforme se pode notar nos exemplos abaixo:

- a) O aborrecimento dos velhos.  
Ex
- b) A agitação da noiva.  
Ex
- c) A tranqüilidade de Daniel.  
Ex
- d) A antipatia do avô pelo neto  
Ex Ob

#### 3.4.4.2 – Valência dos Nomes Concretos

Borba (1996) adota a concepção de que um nome concreto se refere a um objeto físico, localizável no tempo e no espaço, com propriedades perceptuais diretamente observáveis. Ainda segundo o autor, a imaginação também pode criar objetos físicos materializados, tais como fantasma, sereia, lobisomem e centauro, sendo esses, portanto, também caracterizados como nomes concretos.

A valência lógica dos nomes concretos prevê nomes avalentes e nomes que exigem um, dois ou três actantes. Os nomes concretos de valência zero são aqueles “que se definem por si mesmo, sendo, portanto, capazes de formar, sozinhos, um sintagma nominal” (Borba, 1996, p. 125), tais como pedra, árvore, tamanduá, etc. Os nomes concretos valenciados, ou seja, aqueles que exigem argumentos, não são capazes de formar um sintagma nominal sozinhos, pois precisam de outros actantes para se realizar plenamente. São exemplos de nomes concretos os nomes que indicam relações sociais ou de parentesco, tais como pai, sogra, vizinho e

vendedor; bem como aqueles que derivam de radicais verbais, tais como urina, jorro, comprador e copista.

De acordo com Borba (1996, p.128), sintaticamente, os nomes concretos valenciais podem se arranjar da seguinte forma:

- Nomes Concretos com 1 Argumento: Prep N – Exemplos: amante de mamãe. crente em Deus.
- Nomes Concretos com 2 Argumentos: de N + a N – Exemplo: pedinte de esmolas aos abastados.
- Nomes Concretos com 3 Argumentos: de N + de N + para N – tradutor de poemas do árabe para o português.

A valência semântica dos nomes concretos segue as valências semânticas dos verbos e dos nomes abstratos. O primeiro passo, portanto, é determinar as possíveis matrizes que especificam as funções temáticas. Para os nomes concretos de um argumento surge o papel temático Ob. Para os de dois argumentos, no primeiro argumento podem surgir os papéis temáticos Ob ou Ag, e o segundo argumento pode aparecer com os casos B, Or ou Dr. Finalmente, para os nomes concretos que exigem três argumentos, surgem no primeiro deles o caso Ob, no segundo actante Or e no terceiro M. Os exemplos abaixo, retirados de Borba (1996, p.129), demonstram as matrizes semânticas:

- a) O assassino do industrial; Esguicho d'água.

Ob                      Ob

- b) Doador de sangue aos aidéticos; enviado do presidente ao senado.

Ob            B                      Ag            Dr

- c) Uma bela tradução do poema do inglês para o português.

Ob            Or            M

### 3.4.5– Valência Adjetival

“Um adjetivo é essencialmente um modificador do substantivo.” (Cunha & Cintra, 2001).

O primeiro desafio para a determinação da valência dos adjetivos é saber diferenciá-los dos substantivos. Segundo Cunha & Cintra (2001), é muito estreita a relação entre o substantivo (termo determinado) e o adjetivo (termo determinante). O exemplo abaixo demonstra que, na primeira frase, “preta” tem a função de substantivo e “velha” de adjetivo. Já na segunda frase do exemplo, as funções se invertem.

- a) Uma preta velha vendia laranjas.
- b) Uma velha preta vendia laranjas.

Uma saída para este problema, segundo Borba (1996) é que os dicionaristas registram nos dicionários em entradas diferentes, ora como adjetivos, ora substantivos, palavras como caçador, amante, impostor, transformista, sectários, etc.

Um outro ponto a ser observado é que existe uma relação entre o verbo, o nome e o adjetivo. O verbo é o primitivo, sendo que os demais seguem o primeiro no esquema profundo. O verbo emprestar, por exemplo, é derivado em empréstimo (nome) e emprestável (adjetivo). Todos os três têm três argumentos, todos argumentos com as mesmas exigências semânticas, modificando-se apenas no esquema superficial, com arranjos diferentes expressos nas preposições:

- emprestar (Sujeito, Complemento 1 (nome), para/a Complemento 2 (nome humano))  
Paulo emprestou o carro para Maria.
- Empréstimo (de Sujeito, de Complemento 1 (nome), para/a Complemento 2 (nome humano))  
Empréstimo do carro de Paulo para Maria.
- Emprestável (de Sujeito, Complemento 1 (nome), para/a Complemento 2 (nome humano))  
Carro emprestável de Paulo para Maria.

Segundo Felippo & Dias-Da-Silva (2005), “na conceituação da subclasse valencial, pressupõe-se que os adjetivos em posição predicativa (p.ex.: (4) O rapaz era descendente de portugueses.) são verdadeiros predicadores, expressando, assim, um “estado-de-coisas” (predicação).”.



### 3.4.6 – Valência Adverbial

No final de sua obra, Borba (1996) deixa como problema residual a questão da valência adverbial. De fato o autor sugere que alguns advérbios parecem ser valenciais como os de localização relativa (longe, perto, além, aquém, dentro, fora). Entretanto, para determinar se tais advérbios são valenciais, é necessário decidir primeiramente se, na análise sintática, o item corresponde a um advérbio ou a uma locução prepositiva (sintagma preposicional). O exemplo abaixo demonstra esta preocupação.

- a) “Longe” (adv) “de casa” (Sprep).
- b) “Longe de casa” (Sprep).

Ao escolher a primeira opção de análise do exemplo acima, o advérbio tem um argumento que é caracterizado por ser um sintagma nominal, ficando assim a matriz valencial: Adv (Sprep). Borba (1996) segue só até este ponto e não dá mais pistas sobre a valência dos advérbios.

### 3.5 – Teoria da Comunicação

Muitos são as propostas de autores para uma teoria de comunicação. A teoria mais conhecida é a de Shannon e Weaver, de 1949, onde a comunicação é descrita como linear, em uma única via. Apesar das inúmeras críticas sofridas, muitas pelo fato de que os pesquisadores tinham preocupação específica com os problemas de comunicação de dispositivos eletro-eletrônicos, este modelo serviu como estímulo para avanços posteriores na área. A figura abaixo descreve o modelo proposto por Shannon e Weaver:

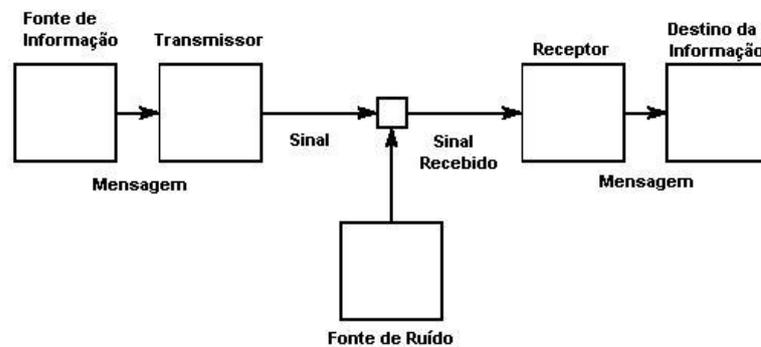


Figura 7: Modelo de Comunicação de Shannon & Weaver

Adaptada de: Shannon & Weaver (1949)

No presente trabalho, é adotado o modelo de comunicação de Tubbs & Moss (2003) por contemplar elementos concernentes ao processo de comunicação humana. Mesmo assim é notória a inspiração de Tubbs & Moss (2003) em Shannon e Weaver.

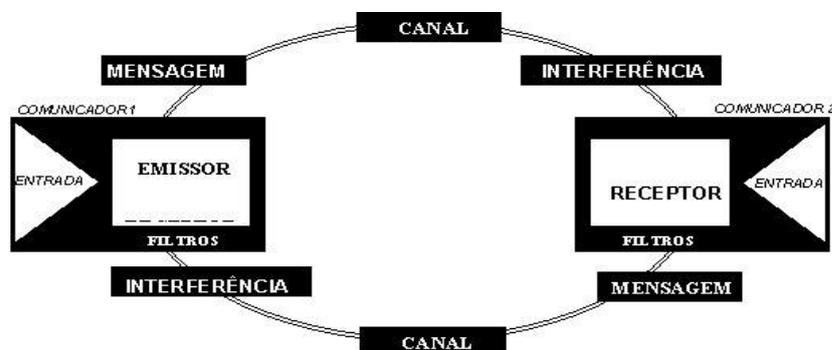


Figura 8: Modelo de Comunicação de Tubbs & Moss

Adaptada de: Tubbs & Moss (2003)

A figura acima representa o modelo de comunicação de Tubbs & Moss (2003). Comunicador 1 e Comunicador 2 enviam e recebem mensagens através de canais de comunicação. As mensagens são susceptíveis a interferências, ou ruídos, que podem alterar seu conteúdo ou sua compreensão. Cada comunicador possui necessidades, construtores de modelos mentais, experiências pessoais, acadêmicas e profissionais singulares. Este conjunto de fatores provoca a criação de filtros que envolvem os comunicadores. Os filtros afetam o envio e a recepção de mensagens. Comunicador 1 e 2 se comportam ora como emissor, ora como receptor, num ciclo *ad infinitum* de comunicação.

O modelo de comunicação de Tubbs & Moss (2003) pode ser empregado em qualquer tipo de comunicação, inclusive na comunicação humana e organizacional. Barichello et al (2003) alerta que é importante salientar a diferença entre os conceitos de comunicação e informação, eventualmente tratados quase como sinônimos em publicações especializadas em comunicação organizacional. A comunicação é um processo de construção de sentidos, enquanto a informação refere-se ao conteúdo das mensagens postas em circulação.

### 3.6 - Ontologias

Conforme propõe Lima-Marques (2006), o primeiro passo para entender este tema consiste em estabelecer uma distinção entre a Ontologia, disciplina filosófica, e as ontologias, formas de representação do conhecimento.

A primeira, simbolizada com a letra inicial maiúscula, consiste em uma disciplina filosófica, a qual pode ser chamada de Ciência do “ente” ou do “ser”. Esta definição tem sua gênese na antigüidade clássica, mais especificamente em Aristóteles, que definia “Ontologia como o estudo do ser das coisas (...)” (LIMA-MARQUES, 2006).

Historicamente, a definição de Ontologia se entrelaça com a definição de Metafísica. Quanto a esta afirmação, a definição encontrada em Wikipedia (2006) é sintomática: “Metafísica é uma divisão da filosofia que se ocupa do estudo da realidade, dos primeiros princípios (filosofia primeira) e do ser (Ontologia). Os problemas centrais da metafísica são Ontológicos, principalmente o ser enquanto ser.”. Segundo Lima-Marques (2006), o termo Metafísica “foi usado desde a antigüidade clássica pelos filósofos para designar o que Aristóteles chamou de Filosofia Primeira, ou estudo do ser enquanto ser.” Já o termo Ontologia foi introduzido no século XVIII por Christian Wolff “para denominar suas idéias filosóficas, equivalentes à Metafísica no sentido geral aristotélico (...)”.

Já o termo empregado com a letra inicial minúscula diz respeito a estruturas de representação do conhecimento de um determinado contexto, “que buscam a essência do objeto representado.”. (LIMA-MARQUES, 2006). Atualmente as ontologias são aplicadas nas pesquisas em inteligência artificial, representação do conhecimento e arquitetura da informação e, mais recentemente, entraram em voga em função de sua importância para a viabilização da Web Semântica.

Uma das definições clássicas de ontologia é dada por Gruber (2001), que a descreve como uma especificação explícita de uma conceituação, ou seja, como uma especificação dos conceitos de um determinado domínio e suas relações, restrições e axiomas, definidos de forma declarativa. “Uma conceituação é uma visão abstrata e simplificada do mundo por esta representado para algum propósito.” (Lima-Marques, 2006).

As ontologias servem como ferramenta para organização, reuso e disseminação de conhecimento já especificado. Dessa forma, elas servem também de vocabulário de comunicação entre os agentes de software. Pode-se dizer que um agente se compromete com uma ontologia se suas ações observáveis são consistentes com as definições na ontologia.

Segundo Almeida & Bax (2003), “os componentes básicos de uma ontologia são classes (organizadas em uma taxonomia), relações (representam o tipo de interação entre os conceitos de um domínio), axiomas (usados para modelar sentenças sempre verdadeiras) e instâncias (utilizadas para representar elementos específicos, ou seja, os próprios dados)”.

Os conceitos em uma ontologia são organizados em quadros (*frames*). Segundo Lustosa & Alvarenga (2004), “O modelo de frames para a representação do conhecimento foi introduzido inicialmente em 1975 por Marvin Minsky. Em geral, um frame é uma coleção de atributos, chamados de slots, e valores, que descrevem alguma entidade do mundo. Os frames integram conhecimento declarativo sobre objetos e eventos e conhecimento procedimental sobre como recuperar informações ou calcular valores.”.

As ontologias podem ser tipificadas quanto à sua função em três tipos: ontologias gerais, ontologias de domínio e ontologias de tarefa. Almeida & Bax (2003) reúnem na literatura a descrição de cada um desses tipos: As ontologias gerais descrevem um vocabulário independente de um problema ou domínio específico, referindo-se a coisas, eventos, tempo, espaço, causalidade, comportamento, funções, etc. As ontologias de domínio fornecem um vocabulário sobre os conceitos, seus relacionamentos, atividade e regras que governam um determinado domínio de discurso ou problema a ser resolvido, tais como a medicina e o direito. Por fim, as ontologias de tarefas fornecem um vocabulário sistematizado de termos, especificando tarefas, como, por exemplo, diagnósticos e compras, que podem ou não estar em um mesmo domínio de discurso.

No Anexo 1 desta dissertação, está disponibilizado um exemplo em linguagem OWL da ontologia proposta nesta dissertação.

### 3.7 - Tratamento automático de ambigüidades na recuperação da informação

Conforme já citado, esta seção dedica-se à obra da professora Dra. Marisa Bräscher, especialmente sua tese de doutorado (BRÄSCHER, 1999) e posterior artigo publicado na revista DataGramZero (BRÄSCHER, 2002), ambos tratando da questão do tratamento de ambigüidades na recuperação da informação.

Em sua tese de doutorado, Bräscher (1999) busca resolver casos de ambigüidades morfológicas e lexicais, mais especificamente da policategorização, da homografia e da polissemia, em textos científicos e técnicos da variação brasileira da língua portuguesa.

Para tal, Bräscher (1999) demonstra que a adição de informações de cunho sintático-semântico a um ambiente de processamento de linguagem natural, organizadas com base na gramática de valências de Borba<sup>16</sup> e na teoria dos gráficos conceituais de Sowa (1994), pode vir a sanar diversos casos de ambigüidades.

Segundo Bräscher (1999), os gráficos conceituais formam uma linguagem de representação do conhecimento e são constituídos por gráficos que possuem dois tipos de nós, a saber: os conceitos e as relações. Os conceitos correspondem a conteúdos de pensamento, representando entidades, ações ou estados. As relações simbolizam as ligações existentes entre os conceitos e demonstram os papéis desempenhados por cada entidade. Bräscher (2002) explicita um exemplo de uma sentença representada através de um gráfico conceitual:

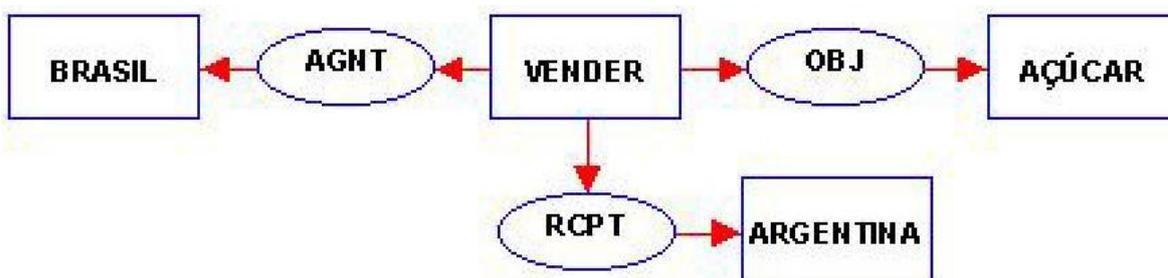


Figura 9: Exemplo de um gráfico conceitual

Adaptada de: Bräscher (2002)

<sup>16</sup> Para maiores informações sobre a Gramática de Borba, ver seção 3.4.

Para o gráfico conceitual acima, segundo Bräscher (2002), deve-se fazer a seguinte interpretação: “VENDER tem por agente BRASIL, por objeto AÇÚCAR e por receptor ARGENTINA.”.

Embasada teoricamente pela gramática de valências de Borba (1996) e pela teoria dos gráficos conceituais de Sowa (1984), a autora faz uso de um sistema de tratamento automático de linguagem natural, o Zstation, que foi desenvolvido na Universidade de Nice por Zinglé (1994). “O Zstation constitui-se num sistema de tratamento automático de linguagem natural, cuja idéia básica é que, para desempenhar uma tarefa, como analisar uma sentença, faz-se necessário coletar toda informação sobre esta sentença, quanto a propriedades semânticas e morfológicas das palavras, possíveis grupos de palavras e frases, e conexões possíveis entre eles, até que o conhecimento coletado permita propor uma ou várias interpretações.” (BRÄSCHER, 2002).

O Zstation é composto por módulos especialistas em determinadas funções, a saber: Geração morfossintática, Análise Morfossintática, Análise sintagmática e Análise semântica. Na geração morfossintática são construídas “formas corretas a partir de lemas de acordo com variáveis morfossintáticas (número, tempo, etc).” (BRÄSCHER, 1999). A análise morfossintática encontra o lema morfológico correspondente para cada forma no texto, bem como sua categoria morfológica. Já função da análise sintagmática consiste em extrair todos os tipos de grupos necessários para a análise sintática da sentença ou de unidades de texto maiores, atuando como um analisador sintático (*parser*). Por fim, a análise semântica se ocupa de procurar todos os conceitos que podem ser associados a um lema morfológico, para, posteriormente, obter as informações necessárias à análise semântica.

O Zstation também faz uso de um dicionário específico, denominado dicionário de base. Esta estrutura armazena um conjunto de lemas e de dados lingüísticos referentes a eles. Bräscher (1999) exemplifica uma entrada de um lema no dicionário de base do sistema:

**brasileiro**

{CPT=brasileiro() }

MOD=amigo

VSM=

VGR=

```

APD=
}
{CPT=brasileiro()
MOD=belo
VSM=
VGR=
APD=$qual
      Arg(0, rel=CHRC, cat=adj, fonct=modSub, conds=[])
}

```

Bräscher (1999) descreve cada um dos dados lingüísticos encontrados na entrada de um lema:

- a) Identificador do conceito (CPT): conjunto de caracteres que simbolizam o conceito representado pelo lema. (...)
- b) Modelo Morfológico (MOD): lema escolhido para representar uma classe de lemas que, pertencendo a uma mesma categoria, sofre a mesma flexão quanto ao tempo, ao modo e à pessoa, para verbos, e quanto ao gênero e ao número para as demais categorias. (...)
- c) Argumentos (APD): dados sintático-semânticos relacionados ao lema de entrada. (...). Os argumentos são definidos com base na valência sintática e semântica do lema. Cada argumento é estruturado da seguinte forma:

```
(Code, rel=R, cat=C, fonct= F, conds=[r(...)],
```

Em que:

Code = código da prioridade que pode ser 0 para um argumento facultativo; 1 para argumento obrigatório representado em termos de relação conceitual; 2 para argumento obrigatório que não é passível de representação em relação conceitual e 3 para argumento proibido (regra de bloqueio).

rel = Relação Temática estabelecida com o conceito do lema na ontologia.

Cat = Categoria morfossintática do argumento.

Conds= condições de validação intralingüísticas, que não podem ser deduzidas na ontologia utilizada. São definidas na forma de relações conceituais.

Além do dicionário de base, o Zstation faz uso de uma gramática morfológica. Esta gramática reúne “o conjunto de lemas selecionados como modelos morfológicos para os demais lemas incluídos no dicionário de base”. (BRÄSCHER, 2002). Constam em cada entrada da

gramática, conforme exemplo abaixo extraído de Bräscher (2002), o modelo morfológico do lema, a categoria gramatical (CAT), as variáveis flexionais (VARs) e a regra morfológica a ser aplicada (REGs). Dessa maneira, é possível gerar formas derivadas das formas canônicas a partir das regras disponíveis na gramática morfológica.

**amigo**

CAT = sub VARs=[masc, sing] REGs=[]

CAT = sub VARs=[masc, plur] REGs=[+s]

CAT = sub VARs=[fem, sing] REGs=[-o, +a]

CAT = sub VARs=[fem, plur] REGs=[-o, +as]

Associada à gramática morfológica, a autora propõe a adoção de uma gramática de argumentos. “As regras da gramática de argumentos possibilitam que o sistema identifique e analise, nos enunciados do corpus, as seqüências que devem ser interpretadas segundo os parâmetros estabelecidos nos argumentos.” (BRÄSCHER, 2002).

O último componente do arcabouço proposto por Bräscher (1999) é a ontologia. Segundo a autora, a relação temática definida num argumento é especificada na ontologia, que, de fato, equivale-se a uma representação linear dos gráficos conceituais. “A cada relação podem ser especificadas, se necessário, as características ou categorias conceituais que delimitam os tipos de conceitos com os quais esta relação pode ser estabelecida.” (BRÄSCHER, 2002). As relações são formadas pela seguinte tripla:

- a) um indicador se uma declaração é obrigatória ou facultativa;
- b) um identificador da relação temática; e
- c) um conceito, característica ou classe de conceito com o qual se estabelece a relação temática. O símbolo “+” indica os que são aceitos, e o símbolo “-” informa os que não são aceitos.

No exemplo abaixo, adaptado de Bräscher (2002), vê-se como o conceito “Título de Crédito”, relacionado à forma canônica “ação”, é registrado na ontologia:

**titulocredito0**

r(0, isa, +produtoComercializável0)

r(0, ORIG, +instituicaoFinanceira0)

r(0, POSS, +animado0)

Na primeira linha é possível perceber a relação taxonômica (isa) entre os conceitos “titulocredito0” e “produtoComercializável0”. Na segunda linha percebe-se a relação temática que caracteriza o conceito “instituicaoFinanceira0” como origem (ORIG) de “titulocredito0”. Finalizando o exemplo, a última linha indica que a entidade a possuir um “titulocredito0” deve ter a restrição seletional “animado0”.

Utilizando-se de um corpus constituído de documentos oficiais do Mercosul, Bräscher (1999) efetua testes de desambiguação, submetendo as sentenças do corpus à teoria implementada no sistema Zstation. Vejamos algumas das conclusões da autora.

Na desambiguação de formas policategoriais, a autora chegou aos seguintes resultados: Na classe Adjetivos/Substantivos (hm1), em 3480 ocorrências de formas policategoriais da classe hm1, o sistema interpretou corretamente 1309 (37,61 %) como substantivos e 1667 (47,90 %) como adjetivos. Em 14,48 % das ocorrências (504), o sistema não foi capaz de resolver este tipo de ambigüidade morfológica. Na classe Adjetivos/Verbos (hm2), em 1262 ocorrências de formas policategoriais da classe hm2, o sistema interpretou corretamente 63,7 % das ocorrências, sendo 411 como verbos e 393 como adjetivos. Em 36,3 % das ocorrências, o sistema não foi capaz de resolver a ambigüidade morfológica. Na classe Substantivos/Verbos (hm3), em 1919 ocorrências de formas policategoriais da classe hm3, o sistema interpretou corretamente 77 % das ocorrências. Na classe Adjetivos/Substantivos/Verbos (hm4), em 899 ocorrências de formas policategoriais da classe hm4, o sistema interpretou corretamente 291 (32,36 %) formas como adjetivos, 295 (32,81 %) como substantivos e 112 (12,45 %) como verbos.

No enunciado “regras de preferência para os casos de venda de ações e aumento do capital social”, a ambigüidade presente em “ações”, que é causada pela homonímia (ou homografia como prefere a autora), é resolvida. A forma canônica do verbo “vender” exige um objeto que tenha o traço “produto comercializável”. Portanto, o sistema é capaz de interpretar corretamente que a palavra “ações” corresponde a papéis comercializáveis pela bolsa de valores, evitando a escolha de outras acepções de “ações”, tais como “praticar ação” (Ex: “os

países-membros realizarão ações necessárias”) ou “efeito” (Ex: “substância de ação hormonal”).

Em ambigüidades causadas pela polissemia, como em “notificação” (ação ou resultado de notificar?), “pedido” (Ação ou resultado de pedir?) e “administração” (Ação ou agente de administrar?) é possível efetuar a desambiguação a partir dos traços semânticos exigidos para o argumento ou a valência de outro elemento do contexto. No enunciado “A Presidência Pro Tempore da Comissão remeterá aos demais Estados-Parte cópia das notificações referidas no ...”, o sistema consegue inferir que “notificações” corresponde ao resultado de notificar pois, a valência sintática de “copiar” (cópia) exige um argumento precedido por preposição “de” sendo exigido o traço semântico +documento para o mesmo.

Nomes abstratos de ação que indicam ação-processo e que possuem argumentos objeto e agente precedidos pela preposição “de” podem apresentar ambigüidades predicativas, ou seja, aquelas em que ocorre mais de um tipo de relação temática (caso ou papel temático) entre predicado e argumentos. Segundo Bräscher (2002), nestes casos, a ambigüidade ocorre apenas quando o argumento apresenta o traço +animado, pois este pode ser interpretado como agente ou objeto. É o caso de: “aprovação da comissão”, “designação do diretor” e “regulamentação do estado”. Bräscher (2002) cita Borba (1996) para explicar que este tipo de ambigüidade ocorre em função do apagamento do sintagma preposicional em relação subjetiva. Bräscher (1999) afirma que “a ambigüidade predicativa pode ser solucionada, por meio de tratamento sintático-semântico, quando outros elementos do contexto oferecem restrições que orientem a escolha da interpretação correta.”. A autora cita como exemplo o enunciado: “todo local alfandegado no qual podem ingressar mercadorias com autorização e sob controle da autoridade aduaneira”. A análise das relações que se estabelecem entre mercadoria, controle e autoridade indicam que “autoridade” exerce o papel temático de agente.

Por fim, a autora conclui que é possível introduzir procedimentos automáticos de solução de ambigüidades através da inserção de informações de cunho morfológico, sintático e semântico em sistemas de tratamento da linguagem natural.

### 3.8 – A Web Semântica e os Ambientes de Autoria baseados em ontologias

De acordo com Berners-Lee (2001), a Web Semântica se propõe a estruturar o conteúdo das páginas Web, de maneira a criar um ambiente onde agentes de software possam responder a tarefas sofisticadas solicitadas pelos usuários. Neste ambiente, os agentes de software operariam de maneira autônoma, percorrendo o ambiente informacional e selecionando informação relevante, deduzindo ou aprendendo quais as informações úteis para o objetivo pelo qual foi construído.

A Web semântica é uma evolução em relação à Web atual. Os sítios baseados na tecnologia atual, notadamente os fundamentados na linguagem HTML, são voltados para a mera exibição da informação, em detrimento da semântica do conteúdo informacional. Os problemas inerentes a esta abordagem são conhecidos, principalmente os relacionados à recuperação da informação em sítios de busca. A nova web “visa incorporar semântica às informações. Isso proporcionará não somente aos usuários entenderem as informações como também as máquinas. Ela pretende fornecer estruturas e dar significado semântico ao conteúdo das páginas web, criando um ambiente onde agentes de software e usuários possam trabalhar de forma cooperativa.” (Dziekaniak & Kirinus, 2004). A figura abaixo demonstra como a Web Semântica é organizada em camadas.

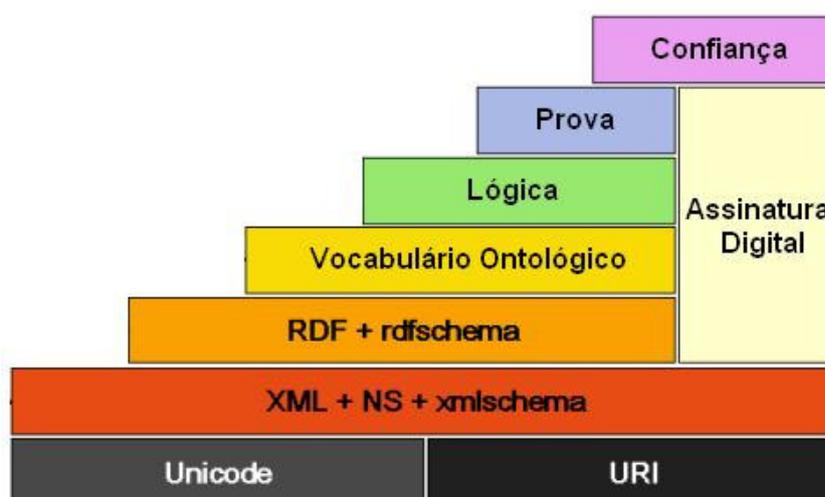


Figura 10: Camadas da Web Semântica  
Adaptada de: Koivunen & Miller (2001)

Segundo Koivunen & Miller (2001), a camada de Unicode e URI serve para certificar que estarão sendo utilizados padrões de caracteres aceitos internacionalmente (Unicode) e para prover meios para identificar cada objeto da Web Semântica (URI). A camada que agrega XML, NS (namespace) e XML Schema serve para estruturar os dados e certificar que tais dados possam ser integrados com outros recursos baseados em XML. Através de camada que engloba as tecnologias RDF e RDF Shema é possível efetuar sentenças sobre os objetos que são identificados univocamente por URI's. Esta é uma camada tipicamente de metadados onde é possível tipificar os recursos e as ligações entre eles. A camada de vocabulário ontológico permite definir conceitos e relações semânticas entre os conceitos. A função da camada de assinatura digital é assegurar que o conteúdo dos documentos não foi alterado indevidamente. As três camadas superiores: Lógica, Prova e Confiança, ainda não estão plenamente desenvolvidas e, por isso, são alvo atualmente de pesquisas por parte da comunidade científica. A camada de Lógica permite o registro de regras de produção enquanto que a camada de prova executa tais regras e avalia, juntamente com a camada de confiança, o grau de confiança das inferências realizadas.

Dentro do contexto da Web Semântica, destaca-se para os fins deste trabalho a Iniciativa de Acessibilidade da Web (W3C, 2006a). Esta iniciativa do W3C visa desenvolver estratégias, guias, e conteúdos para fazer a Web mais acessível a pessoas com pouca ou nenhuma habilidade para acesso e criação de conteúdo na Web. Dentro desta iniciativa, destaca-se o grupo de trabalho que pesquisa métodos e tecnologias para aumentar a acessibilidade das ferramentas de autoria de conteúdo para a Web. W3C (2005), define uma ferramenta de autoria como um *software*, ou coleção de componentes de *software*, que os autores utilizam para criar ou modificar conteúdo da Web para publicação. Ainda segundo W3C (2005), uma coleção de componentes de software consiste em quaisquer produtos de software utilizados em conjunto ou em separado, tais como editores de marcação, editores de imagem e ferramentas de validação, independentemente de haver qualquer tipo de colaboração formal entre os desenvolvedores dos produtos.

Contribuindo com esta frente de pesquisas, Oliveira (2004) propõe uma série de requisitos que devem ser considerados no desenvolvimento de um novo ambiente para autoria na Web Semântica, de modo que este novo ambiente possa fazer uso integrado de tecnologias com potencial. Esta proposta consiste em um ambiente de produção de conhecimento baseado em ontologias, integrado pelas tecnologias disponíveis atualmente dentro de uma plataforma que

permita que os usuários, incluindo autores, revisores e editores, executem as atividades necessárias para produção de conhecimento. Os requisitos, extraídos de Oliveira (2004), são resumidos abaixo:

- Autoria – integração de tecnologias de processamento de textos a ontologias e bases de conhecimento, de maneira a disponibilizar para os autores um ambiente hipertextual onde estes possam navegar facilmente e executar tarefas de autoria.
- Anotação e Marcação Semântica - anotação e marcação Semântica de termos relevantes que podem ser ligados a outros documentos ou ontologias, provendo suporte para navegação conceitual, descoberta de conhecimento e geração de metadados.
- Navegação Semântica – visualização das conexões semânticas entre os conceitos e seus respectivos recursos.
- Geração de Metadados – atribuição de elementos de metadados a termos ou sentenças do documento.
- Suporte a multi-ontologias – o uso de uma ou mais ontologias simultaneamente. Em áreas multidisciplinares, é fundamental que as ferramentas de autoria permitam a ligação do documento a mais de uma ontologia no mesmo ambiente.
- Visualização de ontologias – navegação semântica através de ontologias visando aquisição de conhecimento (por máquinas e humanos) bem como permitir a atribuição ou localização de um conceito dentro de uma estrutura hierárquica a seus respectivos conceitos relativos (pais ou filhos) visando busca e indexação.
- Escolha de conceitos baseada em thesaurus – suporte ao reconhecimento de sinônimos e antônimos dentro de uma estrutura conceitual, sugerindo o uso de termos preferidos, evitando variações e ambigüidade
- Extração de Informação – reconhecimento de informação presente em diversas fontes de informação para fins de indexação e recuperação.
- Engenharia de Texto – identificação de elementos de semântica em textos para atividades de processamento de linguagem natural, tais como troca de informação, tradução, operação com sistemas diferentes, bem como visualização de árvores gramaticais que indicam as estruturas lingüísticas das sentenças.
- Ambiente de descoberta de recursos – busca e navegação em fontes de referência, citação, etc., através de marcação automática de fontes relacionadas e criação de *hyperlinks* dos documentos para as páginas de destino.

- Argumentação e revisão pelos pares – uma estrutura para documentar revisão de pares com o intuito de criticar um artigo e julgar sua aceitabilidade pela comunidade científica de pares, baseada em uma ontologia compartilhada pelos pares.
- Publicação – submissão de documentos a editores e conferências para formalização da publicação. Moldes (*templates*), linhas guia e recomendações orientarão os autores a formatar a estrutura do documento e a registrar os metadados de acordo com as demandas editoriais do veículo de comunicação científica onde se pretende publicar.

Após revisar a literatura e avaliar as ferramentas disponíveis atualmente para a autoria na Web, Oliveira & Lima-Marques (2006) propõem um modelo de arquitetura para ambientes de autoria baseados em ontologias que visa vencer os desafios propostos pelo W3C bem como atender os requisitos propostos em Oliveira (2004). Conforme Oliveira & Lima-Marques (2006), o principal pré-requisito para este ambiente é que as ontologias de contexto estejam bem desenvolvidas e suficientemente maduras para prover os autores com bases de conhecimento consistentes, deixando o mínimo de esforço possível para o autor na edição e população da ontologia.

A filosofia do ambiente proposto consiste em anotar automaticamente os termos no documento e relacioná-los às ontologias. Dessa maneira, os documentos já serão ligados às ontologias no momento de sua geração, procedimento este que viabiliza geração automática de metadados, indexação e navegação conceitual em fontes associadas às ontologias, bem como facilitará a publicação e, principalmente, a recuperação da informação, uma vez que as buscas estarão baseadas em semântica precisa que foi adicionada no momento da autoria.

Através de uma interface homem/máquina os documentos serão criados e ligados a ontologias e bases de conhecimento. Para tal é necessário um ambiente integrado onde conviverão os editores atuais e as ferramentas de autoria. O processo de autoria será facilitado por um módulo que permitirá navegar semanticamente e visualmente pelas ontologias, o que facilitará sobremaneira a localização, escolha e validação de conceitos. Ontologias poderão ser editadas e carregadas com novos conceitos. Um mesmo documento pode se utilizar de várias ontologias e metadados poderão ser gerados em padrões pré-definidos, tais como o Dublin Core. O módulo

de extração de informação permite extração automática de documentos e bases de conhecimento, além de viabilizar a anotação e marcação nos documentos.

A descrição do modelo termina com um módulo de descoberta de conhecimento, onde usuários e máquinas (agentes de software, por exemplo) estarão aptos a navegar pelas estruturas conceituais, inferindo novos conhecimentos, não só a partir do conhecimento relativo ao documento em processo de autoria, mas também a partir das representações disponíveis em banco de dados locais, na própria Web Semântica e em outras fontes de conhecimento. Na figura abaixo é possível apreciar o modelo proposto.

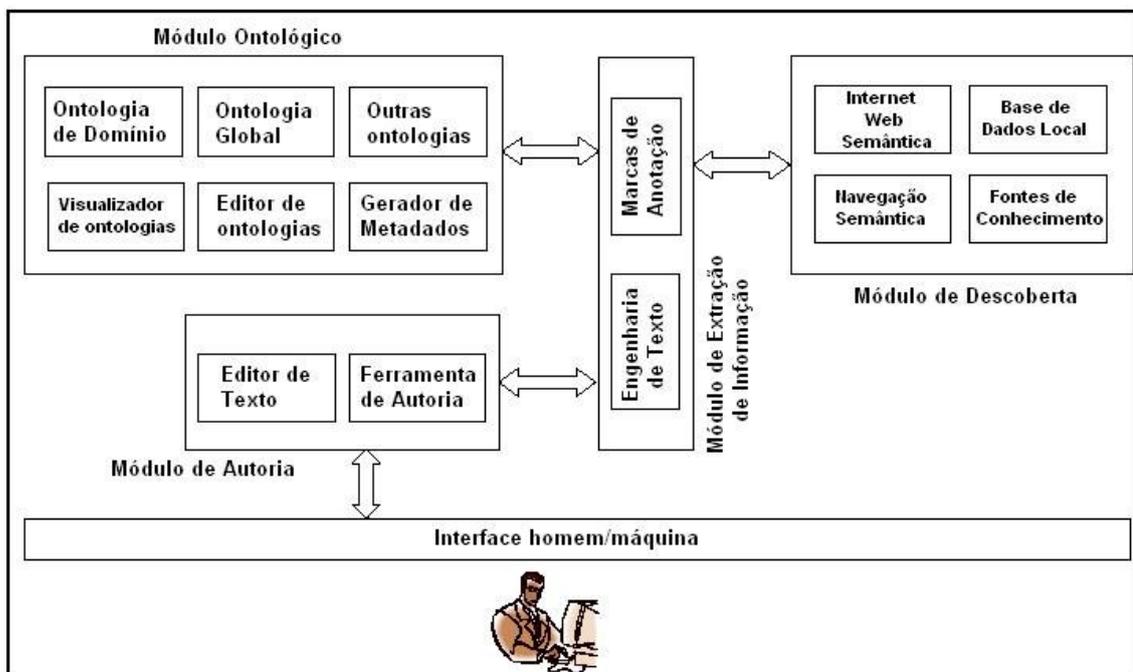


Figura 11: Ambiente de Autoria na Web baseado em ontologia

Adaptada de: Oliveira & Lima-Marques (2006)

Em sua tese de doutorado, Oliveira (2006) dá prosseguimento à sua solução para ambientes de autoria e propõe um protótipo não funcional para a interface homem/máquina que servirá para a redação de documentos. Oliveira (2006) também disponibiliza a especificação de alto nível para a implementação da referida interface, incluindo diagramas UML (*Uniform Modelling Language*), tais como casos de uso, diagramas de seqüência e diagramas de atividades. A figura que demonstra o protótipo de Oliveira (2006) é demonstrada abaixo:

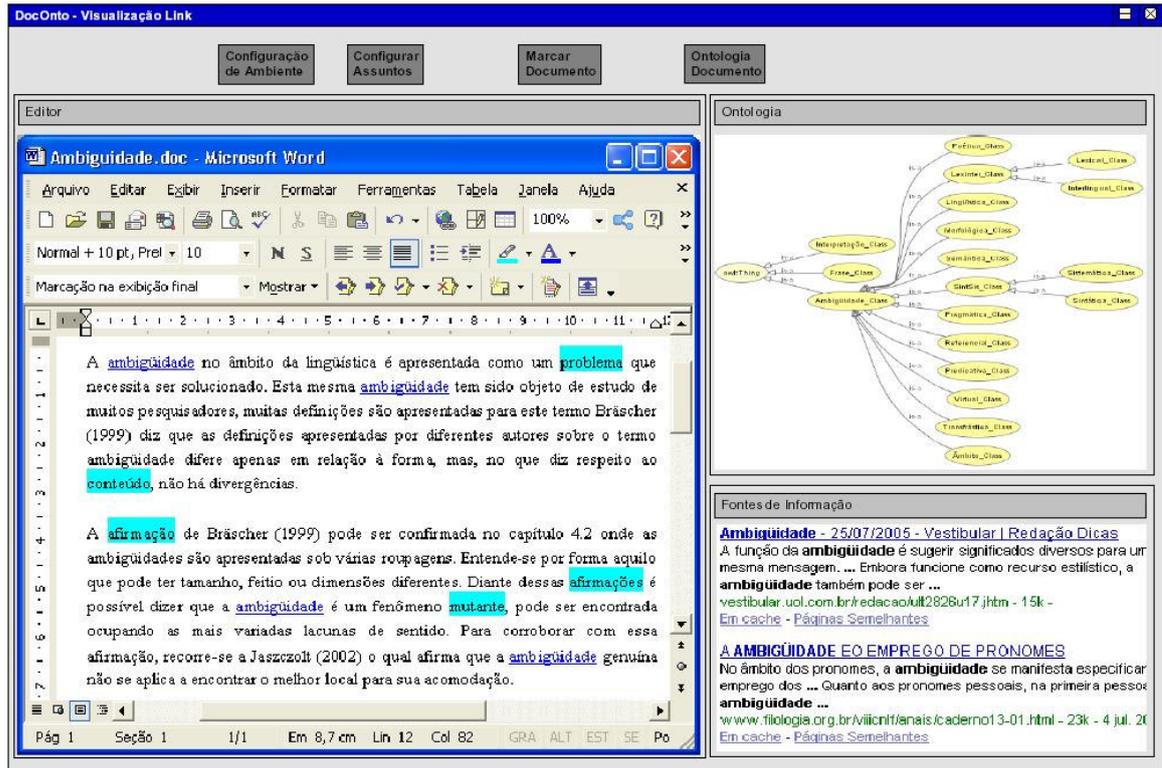


Figura 12: Tela do Ambiente de Autoria para a Web Semântica

Adaptada de: Oliveira (2006)

A interface acima está subdividida em três partes, a saber: Editor de texto (à esquerda), Ambiente de ontologia (na parte superior direita) e Navegador WEB (na parte inferior direita). Conforme Oliveira (2006), a integração destas três telas em um único ambiente permite auxiliar o autor nas seguintes tarefas:

- fazer link entre os termos do documento e ontologias de domínio;
- oferecer um ambiente de descoberta interligado via ontologias para pesquisa e seleção de outros termos semanticamente relacionados;
- exibir as variações que os termos podem apresentar de modo a suscitar consciência dos usuários de possíveis relações de sinonímia e outras equivalências, de modo a ajudá-los na padronização da representação do texto de acordo com padrões terminológicos de domínios de conhecimento específicos, representados pelas ontologias;
- fazer links semânticos entre os documentos e referências com recursos em linha por meio das ontologias;
- criar automaticamente os metadados do documento a partir de modelos previamente apresentados e definidos;
- escrever textos em qualquer ambiente institucional, técnico ou científico, onde a ocorrência de ambigüidade deve ser reduzida e uma relação de univocidade entre os termos e conceitos representados seja necessária.

### 3.9 – Fundamentação Teórica das Ferramentas

A seguir, são arrolados os fundamentos relacionados às ferramentas utilizadas para a validação dos resultados desta pesquisa, a saber: XML – seção 3.9.1, RDF – seção 3.9.2, OWL – seção 3.9.3 e Protégé – seção 3.9.4.

#### 3.9.1 – XML (Extensible Markup Language)

Segundo W3C (2006b), a linguagem extensível de marcação (*Extensible Markup Language - XML*) consiste em uma linguagem simples e muito flexível derivada da SGML<sup>17</sup> (*Standard Generalized Markup Language - ISO-8879*). Originalmente desenvolvido para atender aos desafios da publicação eletrônica em larga escala, o XML vem atualmente ocupando um espaço muito importante no intercâmbio de dados na Internet e em outros tipos de aplicação. Segundo o W3C, alguns dos principais objetivos do XML são:

- ser diretamente utilizável na Internet;
- ser legível por humanos;
- possibilitar um meio independente para publicação eletrônica;
- permitir a definição de protocolos para troca de dados pelas empresas (independentemente da plataforma de hardware e software);
- facilitar às pessoas o processamento de dados.

Tanto o XML quanto o HTML (HyperText Markup Language), são derivados do SGML. O HTML é notoriamente a linguagem mais utilizada na Internet atualmente. Entretanto, existe uma grande diferença entre o XML e o HTML. Esta diferença é notada por Almeida (2002): enquanto as marcas são pré-definidas em HTML, o XML “permite ao autor do documento a definição de suas próprias marcas. Esta característica confere à linguagem XML “habilidades” semânticas, que possibilitam melhorias significativas em processos de recuperação e disseminação da informação.”. A autora continua: “O XML é uma arquitetura

---

<sup>17</sup> SGML é um padrão internacional para definição de métodos para a representação de textos em formato eletrônico.

que não possui elementos e marcas predefinidas. Não especifica como os autores vão utilizar metadados, sendo que existe total liberdade para utilizar qualquer método disponível, desde simples atributos, até a implementação de padrões mais complexos”, tais como o Dublin Core e o RDF. Entretanto, é importante entender o XML como um complemento do HTML, e não como um substituto. XML é utilizado para descrever dados, enquanto o HTML é bem empregado quando utilizado para formatar e mostrar tais dados.

Segundo Furnival (2002), o XML representa documentos a partir de elementos de informação. Esses elementos podem ser, por sua vez, compostos de outros elementos. O XML opera sobre eles supondo que cumprem uma estrutura rigorosamente definida em DTDs (*Data Type Definitions*). O exemplo abaixo<sup>18</sup> esclarece a estrutura de um arquivo XML:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<nota>
<para>João</ para >
<de>Maria</de>
<assunto>Lembrete</assunto>
<corpo>Não esqueça de nosso encontro!</corpo>
</nota>
```

A primeira linha no documento – a declaração XML – define a versão do XML e a codificação de caracteres utilizada no arquivo. Neste exemplo, a versão utilizada é a 1.0 e a codificação corresponde ao padrão ISO Latino-1/Europa Ocidental. A segunda linha indica o elemento raiz do documento (<nota>), enquanto que as quatro linhas subsequentes correspondem aos elementos contidos no elemento raiz (<para>,<de>,<assunto> e <corpo>). As marcas contendo uma barra (/) antes do nome, correspondem a marcas de fim, enquanto que as marcas sem barra, correspondem a marcas de início. Portanto, a última linha do documento contém uma marca fechando o elemento raiz. Todas os elementos em XML devem ter uma marca de fim.

Um documento XML pode se utilizar de uma gramática DTD. Segundo Almeida (2002), “Um DTD é uma gramática para o documento XML, e sua importância está relacionada à

---

<sup>18</sup> Adaptado de <[http://www.w3schools.com/xml/xml\\_syntax.asp](http://www.w3schools.com/xml/xml_syntax.asp)>.

possibilidade que o próprio usuário defina suas marcações. Assim, é necessária uma gramática que apresente o significado da marcas criadas.”. A autora esclarece o assunto com o exemplo abaixo:

#### Fragmento de uma página XML

---

```
<bd>
  <peessoa>
    <nome> João </nome>
    <idade> 30 </idade>
    <email> joao@ufmg.br </email>
  </peessoa>
  <peessoa> ... </peessoa>
  ...
</bd>
```

---

#### Parte de um DTD para o fragmento XML acima

---

```
<DOCTYPE bd [
  <ELEMENT BD (peessoa*)>
  <ELEMENT peessoa (nome, idade, email)>
  <ELEMENT nome (#PCDATA)>
  <ELEMENT idade (#PCDATA)>
  <ELEMENT email (#PCDATA)>
]>
```

---

Figura 13: XML e DTD.

Adaptada de: Almeida (2002)

A autora explica o exemplo, relacionando o DTD com XML: “A primeira linha (*do DTD*) diz que o elemento raiz (aquele que está situado no topo da árvore) é <bd>. As próximas cinco linhas são declarações de marcações, que mostram que <bd> pode conter um número arbitrário (representado pelo asterisco) de elementos <peessoa>, cada um contendo os elementos <nome>, <idade> e <email>, os quais contêm apenas caracteres “data” (não possuem mais elementos). A expressão “peessoa\*” é uma expressão regular, significando qualquer número de elementos peessoa.”.

### 3.9.2 – RDF (Resource Description Framework)

De acordo com W3C (2004b), o RDF é um framework para representação da informação na Web, sendo que o mesmo foi desenvolvido para representar informação de maneira flexível e com um mínimo de restrições.

Segundo Powers (2003) o propósito do RDF consiste em prover meios para registrar dados em um formato entendível por máquina, possibilitando, de maneira eficiente e sofisticada, intercâmbio de dados, busca, catalogação, navegação, classificação e assim por diante.

Dentro da visão do W3C para a Web Semântica, o RDF tem papel muito importante. O RDF, segundo W3C (2004b), foi desenvolvido visando possibilitar o processamento automático de informação por agentes de software. Além disso seu desenvolvimento também visa possibilitar a interoperabilidade entre aplicações distintas, bem como servir como framework para registro de metadados sobre recursos na Web e sobre as aplicações que os utilizam.

Segundo Dziekaniak & Kirinus (2004), o RDF recebeu a influência de várias fontes diferentes. As principais influências vieram das comunidades de padronização da web (HTML, XML e SGML), da Biblioteconomia (metadados de catalogação), da representação do conhecimento (ontologias), da programação orientada a objetos, da linguagem de modelagem, entre outras.

A estrutura conceitual fundamental que suporta qualquer expressão RDF consiste em uma coleção de triplas, cada uma delas composta por um Sujeito, um Predicado (também chamado de propriedade) e um Objeto. Segundo Dziekaniak & Kirinus (2004), o significado de uma tripla pode ser resumido como: “o recurso (sujeito) que possui a propriedade (predicado) com determinado valor (objeto)”. Os autores prosseguem com o exemplo: “a tripla <“http://www.urcamp.tche.br/josiane/metadados”, “criador”, “Josiane”> teria o significado: Josiane é a criadora da página <http://www.urcamp.tche.br/josiane/metadados>”. A figura abaixo descreve uma tripla RDF:

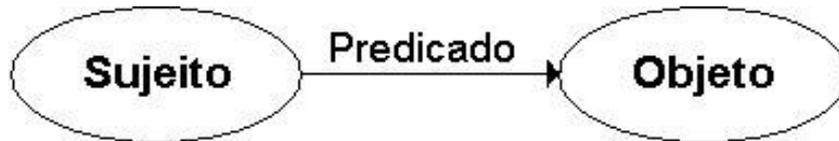


Figura 14: Tripla RDF.

Adaptada de: W3C (2004b)

O sujeito e o objeto são também chamados de nós. Um nó pode ser representado por um URI (*Uniform Resource Identifier*), uma literal ou, simplesmente, ser nulo. Um predicado também é identificado através de um URI. Um URI é uma cadeia de caracteres que identifica univocamente um recurso na web, tais como: documentos, imagens, arquivos, serviços, etc. O exemplo abaixo demonstra um trecho de um arquivo RDF que descreve o discos compactos (CD) do estopque de uma loja:

```
<?xml version="1.0"?>

<rdf:RDF
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:cd="http://www.recshop.fake/cd#">

<rdf:Description
rdf:about="http://www.recshop.fake/cd/Empire Burlesque">
<cd:artista>Bob Dylan</cd:artista>
<cd:pais>USA</cd:pais>
<cd:gravadora>Columbia</cd:gravadora>
<cd:preco>10.90</cd:preco>
<cd:ano>1985</cd:ano>
</rdf:Description>

<rdf:Description
rdf:about="http://www.recshop.fake/cd/Hide your heart">
<cd:artista>Bonnie Tyler</cd:artista>
<cd:pais>UK</cd:pais>
<cd:gravadora>CBS Records</cd:gravadora>
<cd:preco>9.90</cd:preco>
<cd:ano>1988</cd:ano>
</rdf:Description>
```

O framework RDF é estendido semanticamente através do RDFS (*RDF Schema*). RDFS provê mecanismos para descrever grupos de recursos e relacionamentos entre estes recursos. Através do RDFS é possível descrever classes e propriedades de aplicações específicas. As classes definidas em RDFS são similares ao conceito utilizado em linguagens de programação orientadas a objeto. Isso permite que recursos sejam definidos como instâncias de classes, e

subclasses de classes. O exemplo abaixo<sup>19</sup> consiste em um trecho de um documento RDFS, que define a classe “cavalo” como subclasse da classe “animal”.

```
<rdfs:Class rdf:ID="animal" /><rdfs:Class rdf:ID="horse">  
  <rdfs:subClassOf rdf:resource="#animal"/>  
</rdfs:Class>
```

### 3.9.3 – OWL (Ontology Web Language)

Diversas linguagens podem ser empregadas para se implementar uma ontologia. Dentre elas, destacam-se atualmente as linguagens OIL<sup>20</sup>, DAML+OIL<sup>21</sup> e OWL, todas baseadas nos fundamentos da tríade XML, RDF e RDFS.

Para os fins deste trabalho, abordaremos com detalhes a linguagem OWL, que será utilizada mais adiante para se implementar um protótipo de ontologia que suporta o arcabouço teórico proposto por esta dissertação. Em primeiro lugar, a escolha desta linguagem foi motivada por esta ser um padrão de linguagem para criação de ontologias na Web, conforme determinado pelo *World Wide Web Consortium* (W3C). Em segundo lugar, pela facilidade de implementação através da ferramenta Protégé, que possui uma interface específica para a criação de ontologias em OWL.

A OWL foi projetada justamente para servir de suporte para a Web Semântica. Segundo *World Wide Web Consortium* (2004), a OWL foi desenvolvida para uso em aplicações que tenham a necessidade de processador o conteúdo informacional dos documentos, ao invés de apenas exibi-los.

A OWL é dividida em três linguagens, cada qual com um nível de expressividade voltado para públicos específicos de usuários e desenvolvedores. As três linguagens componentes são: OWL Lite, OWL DL e OWL Full. Segue a descrição de cada uma delas, conforme *World Wide Web Consortium* (2004):

---

<sup>19</sup> Adaptado de [http://www.w3schools.com/rdf/rdf\\_schema.asp](http://www.w3schools.com/rdf/rdf_schema.asp).

<sup>20</sup> Maiores informações em <http://www.ontoknowledge.org/oil/>.

<sup>21</sup> Maiores informações em <http://www.daml.org/>.

- OWL Lite – Suporta aqueles usuários que necessitam basicamente de classificações hierárquicas e restrições simples. Por exemplo, é possível definir uma restrição de cardinalidade apenas com valores 0 ou 1. A OWL Lite provê uma migração rápida para thesaurus e taxonomias.
- OWL DL – Suporta aqueles usuários que precisam de expressividade máxima, mantendo completeza computacional (é garantido que todas as conclusões são computáveis) e decidibilidade (é garantido que todas as computações finalizam em um tempo finito). OWL DL inclui todas os construtos da linguagem OWL, porém tais construtos podem ser utilizados sob certas restrições (por exemplo, enquanto uma classe pode ser subclasse de várias classes, uma classe não pode ser uma instância de outra classe). OWL DL tem seu nome oriundo da lógica descritiva, campo de pesquisa que vem estudando a lógica que forma a fundamentação formal do OWL.
- OWL Full – É indicada para usuários que precisam de expressividade máxima e liberdade sintática, porém, sem garantias computacionais (computabilidade e decidibilidade). Por exemplo, em OWL Full uma classe pode ser tratada simultaneamente como uma coleção de indivíduos e como um indivíduo propriamente dito. OWL Full permite a uma ontologia aumentar o significado do vocabulário pré-definido. É improvável que qualquer software esteja hábil a suportar raciocínio completo para todas as capacidades da OWL Full.

Toda ontologia desenvolvida em linguagem OWL Lite, é uma ontologia válida para as linguagens OWL DL e OWL Full. Toda ontologia desenvolvida em linguagem OWL DL, é uma ontologia válida para a linguagem OWL Full. O caminho inverso não é válido.

Uma ontologia em OWL é composta por Indivíduos, Classes e Propriedades. Seguem considerações sobre cada um desses componentes, conforme Horridge *et al.* (2004):

- Indivíduos - Os indivíduos, também conhecidos por instâncias, representam os objetos do domínio de interesse coberto pela ontologia.
- Classes - Já as classes podem ser interpretadas como um conjunto de indivíduos que tenham um mesmo grupo de requisitos exigidos para pertencerem à classe. As classes podem ser organizadas em taxonomias. Por exemplo, a classe Animal pode

conter uma sub-classe Gatos que por sua vez pode conter todos os indivíduos caracterizados como gatos.

- Propriedades – A linguagem OWL provê dois tipos de propriedades: as propriedades de tipo de dados, e as propriedades de objetos. As propriedades de objetos podem ser entendidas como relações binárias entre indivíduos. Por exemplo, a propriedade hipotética temFilho pode ligar o indivíduo Kirk Douglas ao indivíduo Michael Douglas. As propriedades de tipo de dados ligam um indivíduo a um tipo de dado definido na linguagem XMLS. Por exemplo, a propriedade nomeProfessor pode ser associada ao tipo de dado XMLS string.

A figura abaixo demonstra como indivíduos, classes e propriedades se relacionam em uma ontologia. Os círculos indicam classes. Os losangos indicam indivíduos. As setas indicam propriedades.

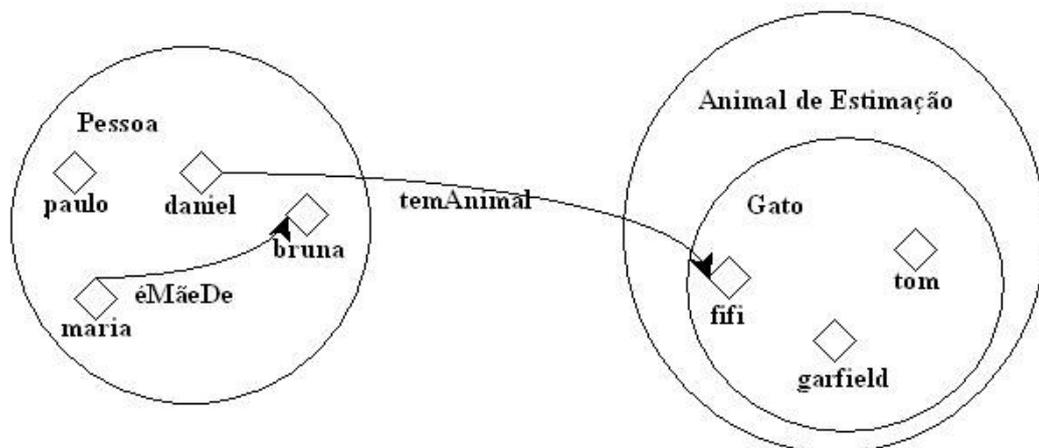


Figura 15: Componentes de uma ontologia OWL

Passamos agora à descrição das propriedades da linguagem OWL. Começamos pelas funcionalidades herdadas da linguagem RDF. São elas, conforme World Wide Web Consortium (2004):

- Class – Uma classe define um grupo de indivíduos que estão unidos sob uma mesma classe por possuírem propriedades em comum.
- rdfs:subClassOf – Hierarquias de classes podem ser criadas através da criação de sentenças que determinam se uma classe é subclasse de outra classe. Por exemplo, a classe Pessoa pode ser determinada a ser uma subclasse da classe Mamífero.

Através desta hierarquia, um software de raciocínio (*reasoner*) pode determinar que se um indivíduo é uma Pessoa, então também é um Mamífero.

- `rdf:Property` – Propriedades podem ser usadas para estabelecer relacionamentos entre indivíduos, ou de indivíduos para valores de tipos de dados.
- `rdfs:subPropertyOf` – Hierarquia de propriedades podem ser criadas através do estabelecimento de sentenças que estabelecem que uma propriedade é subtipo de outra. Por exemplo, `temFilho` pode ser designada como uma subpropriedade de `temParente`.
- `rdfs:domain` – Um domínio de uma propriedade limita os indivíduos cuja propriedade pode ser aplicada. Se uma propriedade relaciona um indivíduo a outro indivíduo, e esta propriedade tem uma classe como um de seus domínios, então o primeiro indivíduo precisa pertencer àquela classe de domínio.
- `rdfs:range` – O alcance de uma propriedade limita os indivíduos cuja propriedade pode ter como valor. Se uma propriedade relaciona um indivíduo a outro indivíduo, e esta propriedade tem uma classe como um de seus alcances, então o segundo indivíduo precisa pertencer à classe de alcance.
- `Individual` – Indivíduos são instâncias de classes.

As próximas funcionalidades estão relacionadas à definição de igualdade ou desigualdade entre classes ou indivíduos.

- `equivalentClass` – Duas classes podem ser designadas como equivalentes. Classes equivalentes têm as mesmas instâncias. Por exemplo, a classe `Carro` pode ser designada como classe equivalente de `Automóvel`.
- `equivalentProperty` - Duas instâncias podem ser designadas como equivalentes. Instâncias equivalentes relacionam um indivíduo ao mesmo conjunto de outros indivíduos.
- `sameAs` – Dois indivíduos podem ser designados como o mesmo indivíduo. Esta funcionalidade pode ser utilizada para criar diferentes nomes para um mesmo indivíduo.
- `differentFrom` – Um indivíduo pode ser designado como diferente de outro indivíduo.
- `AllDifferent` – Um número qualquer de indivíduos pode ser designado como mutuamente distintos.

As próximas funcionalidades são características especiais que uma propriedade OWL pode ter. São elas:

- `inverseOf` – Uma propriedade pode ser designada como propriedade inversa de outra propriedade. Se uma propriedade P1 é designada como propriedade inversa de P2, então se X é um indivíduo relacionado a Y pela propriedade P2, então Y é relacionado a X pela propriedade P1.
- `TransitiveProperty` – Propriedades podem ser transitivas. Se uma propriedade é transitiva, então se o par (x,y) é uma instância da propriedade transitiva P, e o par (y,z) é uma instância de P, então o par (x,z) é também uma instância de P.
- `SymmetricProperty` – Propriedades podem ser transitivas. Se caracterizada como tal, então se o par (x,y) for uma instância da propriedade P, então o par (y,x) é também uma instância de P.
- `FunctionalProperty` – As propriedades funcionais podem referenciar um valor único em seu alcance.
- `InverseFunctionalProperty` – Propriedades podem ser inversamente funcionais. Se a propriedade é caracterizada como tal, então o inverso desta propriedade é funcional.

As próximas funcionalidades são restrições especiais que envolvem classes e propriedades OWL. São elas:

- `allValuesFrom` – Esta restrição significa que uma determinada propriedade, para uma classe em particular, tem uma restrição de alcance local associada a ela. Por exemplo, a classe Pessoa pode ter uma propriedade chamada `temFilha` restrita a ter todos os valores oriundos da classe mulher.
- `someValuesFrom` – Uma classe particular pode ter uma restrição em uma propriedade em que, pelo menos um valor desta propriedade, é de um certo tipo. Por exemplo, a classe `JornalDaWebSemantica` pode ter uma restrição `someValuesFrom` na propriedade `temPalavraChave` que designa que alguns valores de `temPalavraChave` devem ser uma instância da classe `AssuntoWebSemantica`.

As próximas funcionalidades são restrições relativas à cardinalidade de uma classe OWL. São elas:

- `minCardinality` – Se a cardinalidade mínima é igual a 1 para uma propriedade relacionada a uma classe, então todas as instâncias desta classe estarão relacionadas a pelo menos um indivíduo por aquela propriedade.
- `maxCardinality` – Se a cardinalidade máxima é igual a 1 para uma propriedade relacionada a uma classe, então qualquer instância desta classe estará relacionada a no máximo um indivíduo por aquela propriedade.
- `cardinality` – A cardinalidade é provida como uma conveniência quando é útil designar que uma propriedade em uma determinada classe tem tanto cardinalidade mínima quanto cardinalidade máxima igual a 0 ou tanto cardinalidade mínima quanto cardinalidade máxima igual a 1.

A propriedade a seguir diz respeito à possibilidade de designar intersecções entre classes.

- `intersectionOf` – É possível estabelecer intersecções entre classes em OWL. Por exemplo, a classe `PessoaEmpregada` pode ser descrita como uma intersecção ente a classe `Pessoa` e a classe `CoisasEmpregáveis`.

Todas as funcionalidades descritas acima são parte do vocabulário da linguagem OWL Lite, e portanto, também pertencem às linguagens OWL DL e OWL Full. Não há diferença de vocabulários entre OWL DL e OWL Full. O que existem são restrições impostas à linguagem OWL DL que a OWL Full não tem. Por exemplo, em OWL DL não é possível que uma classe possa ser considerada ao mesmo tempo como um indivíduo e uma propriedade, procedimento que é possível na versão Full. As funcionalidades abaixo são extensões à linguagem OWL Lite, que estão disponíveis no vocabulário de OWL DL e OWL Full.

- `oneOf` – Classes podem ser descritas por enumeração dos indivíduos que fazem parte da classe. Os membros da classe são exatamente o conjunto de indivíduos enumerados. Por exemplo, a classe `DiasDaSemana` pode ser descrita pela simples enumeração dos indivíduos domingo, segunda-feira, terça-feira, quarta-feira, quinta-feira, sexta-feira e sábado.

- `hasValue` – Uma propriedade pode se requisitada a ter um certo indivíduo como um valor. Por exemplo, instâncias da classe de `CidadãosHolandeses` podem ser caracterizadas como pessoas que tenham Holanda como valor de sua nacionalidade.
- `disjointWith` – Classes podem ser designadas como disjuntas umas das outras. Por exemplo, `Homem` e `Mulher` podem ser classes caracterizadas como disjuntas.
- `unionOf`, `complementOf`, `intersectionOf` – Utilizando `unionOf`, podemos estabelecer que a classe contém coisas que são pertencentes às classes `CidadãosAmericanos` e `CidadãosHolandeses`. Utilizando `complementOf`, pode-se estabelecer que crianças não são pertencentes à classe `CidadãosIdosos`.
- `minCardinality`, `maxCardinality`, `cardinality` – Em OWL Lite, cardinalidades eram restritas a pelo menos, no máximo o extamente 1 ou 0. OWL DL e OWL Full permitem cardinalidades que envolvem números inteiros não-negativos.

A imagem abaixo demonstra um exemplo de uma ontologia redigida na linguagem OWL, com o auxílio do software Protégé. Este exemplo é equivalente à ontologia descrita na figura de título “Componentes de uma ontologia OWL”, disponível nesta seção, imediatamente acima.

```

<?xml version="1.0"?>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns="http://www.owl-ontologies.com/unnamed.owl#"
  xml:base="http://www.owl-ontologies.com/unnamed.owl">
  <owl:Ontology rdf:about=""/>
  <owl:Class rdf:ID="CLASSE_GATO">
    <rdfs:subClassOf>
      <owl:Class rdf:ID="CLASSE_ANIMAL_ESTIMACAO"/>
    </rdfs:subClassOf>
  </owl:Class>
  <owl:Class rdf:about="#CLASSE_ANIMAL_ESTIMACAO">
    <owl:disjointWith>
      <owl:Class rdf:ID="CLASSE_SER_HUMANO"/>
    </owl:disjointWith>
    <rdfs:subClassOf>
      <owl:Class rdf:ID="CLASSE_ANIMAL"/>
    </rdfs:subClassOf>
  </owl:Class>
  <owl:Class rdf:about="#CLASSE_SER_HUMANO">
    <rdfs:subClassOf rdf:resource="#CLASSE_ANIMAL"/>
    <owl:disjointWith rdf:resource="#CLASSE_ANIMAL_ESTIMACAO"/>
  </owl:Class>
  <owl:ObjectProperty rdf:ID="eMaeDe">
    <rdfs:range rdf:resource="#CLASSE_SER_HUMANO"/>
    <rdfs:domain rdf:resource="#CLASSE_SER_HUMANO"/>
  </owl:ObjectProperty>
  <owl:ObjectProperty rdf:ID="temAnimalEstimacao">
    <rdfs:range rdf:resource="#CLASSE_ANIMAL_ESTIMACAO"/>
    <rdfs:domain rdf:resource="#CLASSE_SER_HUMANO"/>
  </owl:ObjectProperty>
  <CLASSE_GATO rdf:ID="fifi"/>
  <CLASSE_SER_HUMANO rdf:ID="maria">
    <eMaeDe>
      <CLASSE_SER_HUMANO rdf:ID="bruna"/>
    </eMaeDe>
  </CLASSE_SER_HUMANO>
  <CLASSE_SER_HUMANO rdf:ID="paulo"/>
  <CLASSE_GATO rdf:ID="tom"/>
  <CLASSE_GATO rdf:ID="garfield"/>
  <CLASSE_SER_HUMANO rdf:ID="daniel">
    <temAnimalEstimacao rdf:resource="#fifi"/>
  </CLASSE_SER_HUMANO>
</rdf:RDF>

```

Figura 16: Exemplo de uma ontologia OWL.

### 3.9.4 – Protégé

Segundo Gennari *et al.* (2002), Protégé é um ambiente para o desenvolvimento de sistemas baseados em conhecimento. O Protégé começou como uma pequena aplicação desenvolvida para o domínio da medicina, evoluindo em uma década para um conjunto de ferramentas de propósito genérico, utilizado amplamente pela comunidade científica e para fins comerciais. Segundo o mesmo autor, os principais objetivos do Protégé podem ser resumidos em:

- a) servir para aquisição de conhecimento diretamente de especialistas de domínios com menos dependência de engenheiros de conhecimento;
- b) permitir diversos formalismos e estratégias de inferência;
- c) integrar tarefas de aquisição de ontologias e instâncias, ambiente de teste com inferência em um mesmo ambiente;
- d) criar automaticamente formulários para entrada de conhecimento, acessar ontologias e também relacionar as ontologias.

Basicamente, a arquitetura do Protégé é separada em duas partes: modelo e visão. A parte do modelo corresponde ao mecanismo interno de representação de conhecimento para ontologias e bases de conhecimento. No contexto da visão são disponibilizados componentes que provêm uma interface homem/máquina para manipular o modelo de representação do conhecimento.

Segundo Knublauch *et al* (2004), o Protégé é uma ferramenta de código aberto que possui uma comunidade de milhares de usuários. Esta característica permite que a ferramenta evolua de maneira muito rápida, incorporando sugestões e melhorias efetuadas pela comunidade científica. O número de plugins, extensões e aplicativos disponíveis para a ferramenta é um sintoma claro da amplidão obtida pelas contribuições da comunidade de usuários. Algumas das aplicações e plugins mais populares podem ser citados como: Jambalaya e Ontoviz – extensões que permitem a visualização gráfica de ontologias; o RACER - software de raciocínio que permite, entre outras coisas, checar a consistência das ontologias; e o plugin OWL – plugin utilizado nesta dissertação que permite a criação de ontologias em OWL. Ainda segundo Knublauch *et al* (2004), o Protégé permite atualmente que se desenvolva

ontologias em vários formatos, a saber: CLIPS, RDF, XML, UML, OWL e bases de dados relacionais.

Dar-se-á ênfase ao plugin OWL da ferramenta, pois este foi o escolhido para o desenvolvimento da ontologia de valências, um dos resultados desta dissertação. Para Knublauch *et al* (2004), o plugin OWL é uma extensão complexa do Protégé que pode ser utilizada para criar arquivos OWL e bases de dados. Este plugin inclui uma coleção de abas específicas para criação e edição de ontologias OWL, permitindo inclusive ações como classificação, checagem de consistência e teste da ontologia. A figura abaixo demonstra a tela padrão do plugin OWL do Protégé, quando selecionada a aba “OWL Classes”:

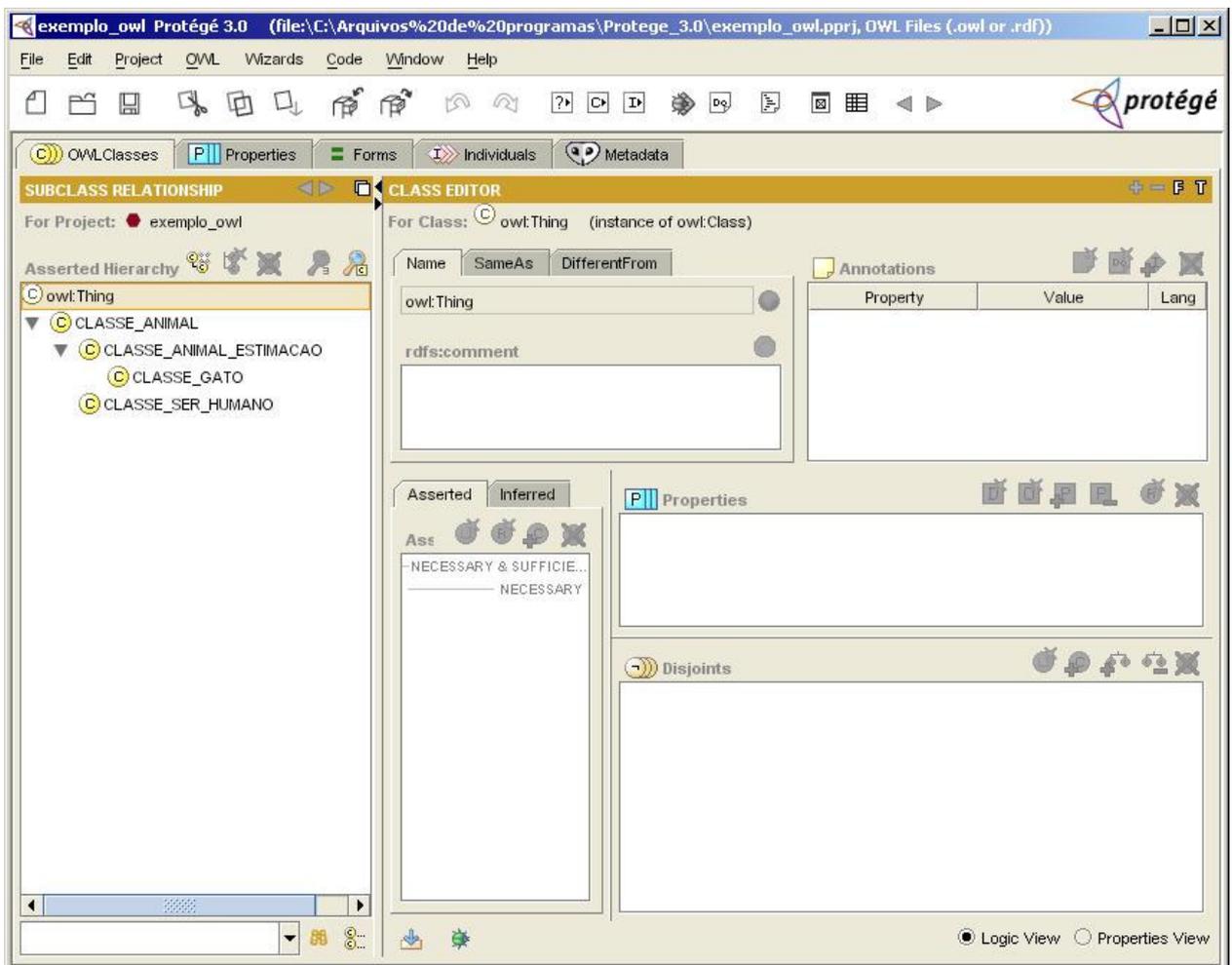


Figura 17: Tela inicial do plugin OWL do Protégé

Na figura acima se pode perceber, à esquerda da tela, a seção onde se visualiza e edita a taxonomia da ontologia OWL. Toda taxonomia OWL começa pela classe “owl: Thing”.

Conforme explica Horridge *et al.* (2004), a classe “owl: Thing” representa o conjunto que contém todos os indivíduos ou instâncias. Por isso, todas as classes de uma ontologia OWL são subclasses da classe “owl: Thing”. À direita é possível perceber uma série de seções onde se pode verificar características da classe selecionada, tais como: suas propriedades (*Properties*), as classes que são disjuntas à classe em questão (*Disjoints*), anotações (*Annotations*) e comentários (*rdf: comment*) genéricos e uma seção onde se pode estabelecer axiomas relativos à classe (*Asserted / Inferred*).

Na parte superior é possível verificar as abas disponíveis para a edição da ontologia. Na aba *Properties* é possível criar e gerenciar propriedades. Na aba *Forms* é possível criar formulários de entrada de dados na ontologia. Na aba *Individuals* são cadastradas e mantidas as instâncias das classes. Na aba *Metadata* é possível registrar metadados genéricos relativos à ontologia. Outras abas podem ser adicionadas, dependendo da configuração do projeto da ontologia. Por exemplo, selecionando-se o plugin do Jambalaya na configuração do projeto da ontologia, uma aba adicional, referente a este plugin, aparecerá na tela do Protégé.

## *4 – Resultados da Pesquisa*

### **4.1 – Proposta de Arcabouço Teórico**

A fim de determinar o arcabouço teórico para a autoria de documentos isentos de ambigüidades, inicialmente far-se-á uso de um exercício criativo. Imagina-se um autor pronto a iniciar a edição de um documento. Este autor liga seu computador pessoal e aciona o aplicativo Valente<sup>22</sup>. O Valente é um editor de texto fictício, com opções de formação de texto e interface gráfica similares ao Microsoft Word, tal qual proposto por Oliveira (2006) em seu ambiente de autoria baseado em ontologias. O Valente utiliza-se de uma ou mais estruturas de representação de conhecimento para auxiliar o usuário no processo de autoria. A cada palavra digitada pelo usuário o Valente acessa estas estruturas de representação de conhecimento e verifica qual é a exata acepção da palavra, bem como quais são as exigências sintáticas para aquela palavra dentro de uma determinada oração. Caso haja mais de uma acepção para a palavra digitada, casos de polissemia e homonímia, o Valente solicitará ao usuário que escolha a exata acepção pretendida. Durante o processo de autoria do documento, o Valente vai conduzindo o autor, orientando-o e inquirindo-o sobre questões morfológicas, sintáticas e semânticas referentes ao texto. Paralelamente ao arquivo contendo a redação do contrato em linguagem natural, o Valente vai registrando em um arquivo anexo os metadados referentes ao documento em linguagem natural. Estes metadados registrarão informações morfológicas, sintáticas e semânticas, que facilitarão posteriormente a recuperação da informação contida no documento.

A figura abaixo demonstra a interface fictícia do Valente. Na imagem pode-se verificar uma interação simples entre o software e o usuário quando o usuário digita uma palavra que tem mais de uma acepção possível na estrutura de representação de conhecimento. Neste momento, o software abre uma caixa de possibilidades para que o usuário escolha qual é a exata acepção pretendida. A escolher a acepção pretendida, o Valente registra formalmente tal escolha no arquivo de metadados anexo ao documento em linguagem natural.

---

<sup>22</sup> Nome inspirado na gramática de Valências.

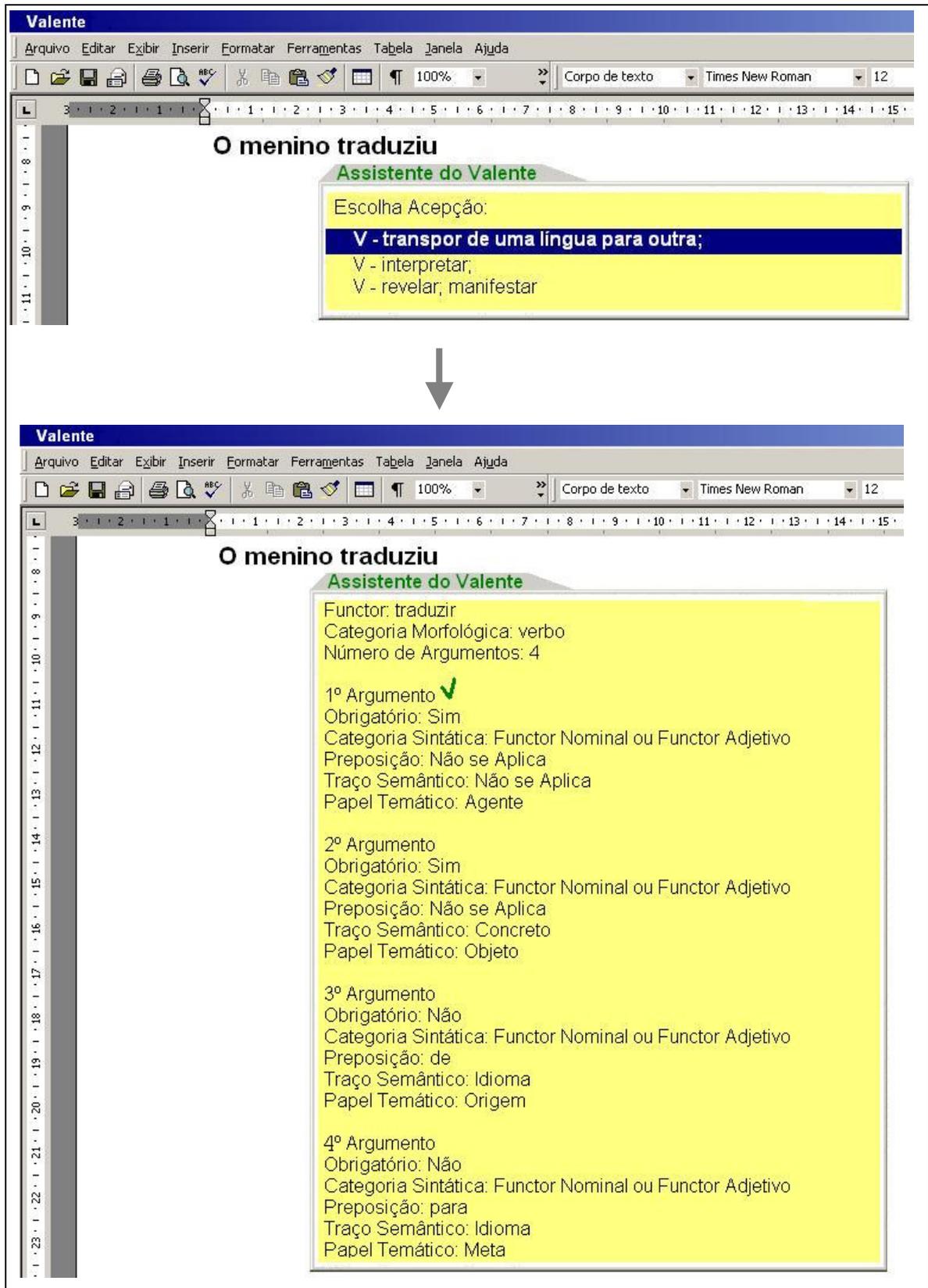


Figura 18: Interface gráfica do software fictício Valente

Na seqüência de figuras apresentada acima, é possível verificar o autor digitando a palavra “traduziu” e, logo após constatar que esta forma gráfica é derivada de “n” formas canônicas referentes às diferentes acepções do verbo traduzir, o software solicita que o usuário designe a exata acepção pretendida. Em seguida, o ambiente de autoria verifica que a palavra digitada trata-se de um predicado verbal e, imediatamente, demonstra ao usuário as exigências sintáticas e semânticas para preenchimento dos argumentos do referido predicado.

Outra faceta importante que o referencial deve contemplar é a questão da recuperação da informação. Um outro exercício de criatividade mostra como se imagina a recuperação da informação utilizando-se da teoria contida neste trabalho. Imagina-se uma ferramenta fictícia de busca na Web. A esta ferramenta daremos o nome de Sherlock<sup>23</sup>. Através desta ferramenta, o usuário poderia fazer buscas simples digitando uma única palavra. Ao digitar a palavra para busca, a ferramenta de busca verificaria na estrutura de representação de conhecimento se a forma gráfica digitada apresenta mais de uma acepção. Em caso positivo, a ferramenta de busca solicitaria ao usuário, de maneira semelhante ao Valente, que escolhesse a acepção exata para busca. Outra possibilidade seria a de fazer pesquisas mais complexas, utilizando-se dos construtos disponíveis pela gramática de valências na estrutura de representação de conhecimento. Por exemplo, pesquisas como “Recupere todos os documentos em que Arnaldo Jabor aparece como agente (papel temático) de uma crítica.”, “Recupere todos os documentos em que uma ação qualquer foi executada inconseqüentemente”, “Recupere todos os documentos em que um nome com traço semântico “corrosivo” foi utilizado para queimar um nome com traço semântico “humano””, “Recupere todos os documentos em que “os detentos” forma transferidos da “penitenciária de Presidente Bernardes” para a “penitenciária da Papuda””. A figura abaixo demonstra a interface fictícia do Sherlock.

---

<sup>23</sup> Nome inspirado no famoso detetive Sherlock Holmes.

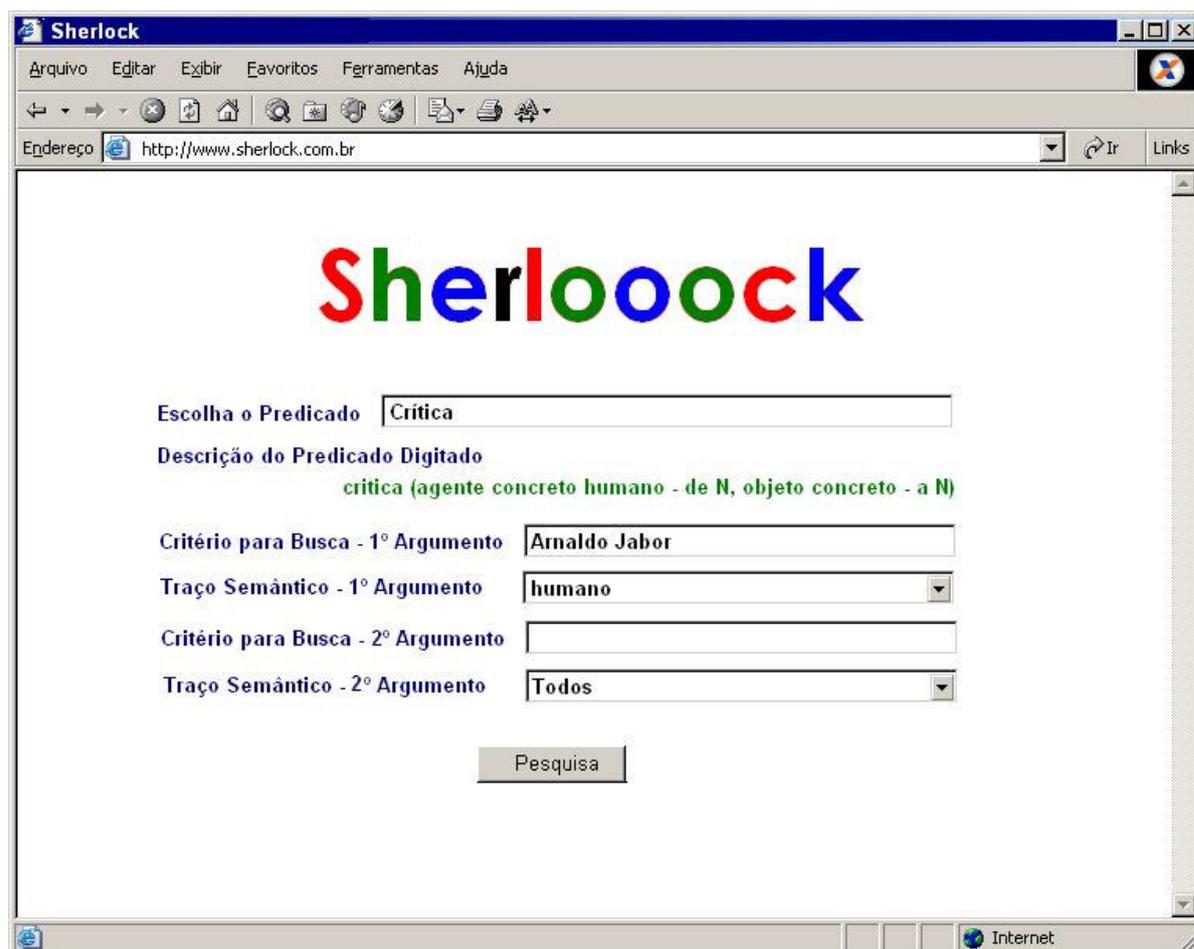


Figura 19: Interface gráfica da ferramenta de busca fictícia Sherlock

Os cenários expostos acima foram propositadamente preparados para introduzir a idéia de resolução do problema das ambigüidades através do processo de autoria dos documentos. Através do cenário fictício exposto acima, portanto, é possível extrair e organizar os componentes conceituais que comporão o arcabouço teórico.

O primeiro componente do referencial que salta aos olhos no cenário acima é a própria linguagem natural. A linguagem natural, seja qual for a língua utilizada, é o meio mais comum para a apresentação do conteúdo informacional dos documentos. Entretanto, como já visto, a linguagem natural é susceptível a diversos fenômenos lingüísticos, tais como a ambigüidade e a vagueza. Visto por um lado, esta susceptibilidade pode ser encarada como uma fragilidade. Por outro lado, a flexibilidade dada pela linguagem natural compensa a falta de um formalismo mais rígido, ampliando os horizontes da comunicação humana. A este nível do arcabouço teórico daremos o nome de nível da linguagem natural, ou, opcionalmente, nível do documento.

O segundo componente está relacionado aos metadados registrados pelo ambiente de autoria em um documento anexo ao documento da linguagem natural. A este segundo componente daremos o nome de Nível da Meta-linguagem de Valências, ou, opcionalmente, Nível do Meta-documento. Através da interação entre o ambiente de autoria e o usuário, especialmente através das inquirições efetuadas a todo tempo pelo software ao usuário em busca de informações morfológicas, sintáticas e semânticas, será edificado um documento paralelo<sup>24</sup> ao documento em linguagem natural. A linguagem através da qual o documento paralelo será redigido (meta-linguagem de valências) será detalhada em uma seção específica. Esta meta-linguagem é inspirada diretamente na gramática de valências de Borba (1996) e na lógica de predicados. Neste meta-documento, as sentenças escritas em linguagem natural serão reescritas em uma linguagem predicativa, que organizará formalmente os componentes da sentença de acordo com a gramática proposta por Borba. Os princípios teóricos dos compiladores serão fundamentais nesta reescrita, que necessariamente deverá ser efetuada através da passagem pelas diversas fases de um compilador, tais como a análise léxica, sintática e semântica. O nível do meta-documento estará intimamente ligada ao próximo nível do referencial, o nível ontológico. Cada palavra registrada no meta-documento será relacionada a sua respectiva entrada na ontologia.

Como já adiantado, o terceiro nível corresponde ao nível ontológico. Este nível corresponde à estrutura de representação de conhecimento citada no cenário de exemplo. Resumidamente, pois adiante detalharemos este nível, na ontologia estarão registradas informações sobre a valência lógica, sintática e semântica dos verbos, nomes, adjetivos e advérbios. Além das informações sobre a valência, cada acepção de um item lexical será descrita na ontologia. A função da ontologia é, primeiramente, auxiliar o usuário na composição do documento e, posteriormente, em conjunto com o meta-documento, otimizar a recuperação da informação provendo informações de caráter morfológico, sintático e semântico sobre o conteúdo do documento em linguagem natural. A figura abaixo esquematiza os três níveis propostos:

---

<sup>24</sup> Para registrar os metadados em Meta-linguagem de Valências, propõe-se um documento paralelo, ou anexo, ao documento referente à linguagem natural apenas por uma questão didática e conceitual. Caso haja implementação de um software baseado nas concepções contidas nesta dissertação, ficará a cargo dos cientistas da computação determinar se é melhor registrar os metadados em um arquivo em separado, ou no próprio arquivo que contém a linguagem natural.



Figura 20: Arcabouço teórico para a Autoria de Documentos Isentos de Ambigüidades

É possível verificar como o arcabouço teórico proposto acima pode ser encaixado no modelo de autoria na Web baseado em ontologias proposto por Oliveira & Lima-Marques (2006). Propõe-se que o módulo denominado interface homem/máquina daquele ambiente fique entre os níveis do documento e do meta-documento do referencial proposto nesta dissertação, gerenciando-os. O módulo ontológico proposto por Oliveira & Lima-Marques (2006), especificamente as ontologias de domínio, corresponderiam ao nível ontológico do framework proposto neste trabalho. A figura abaixo demonstra o encaixe:

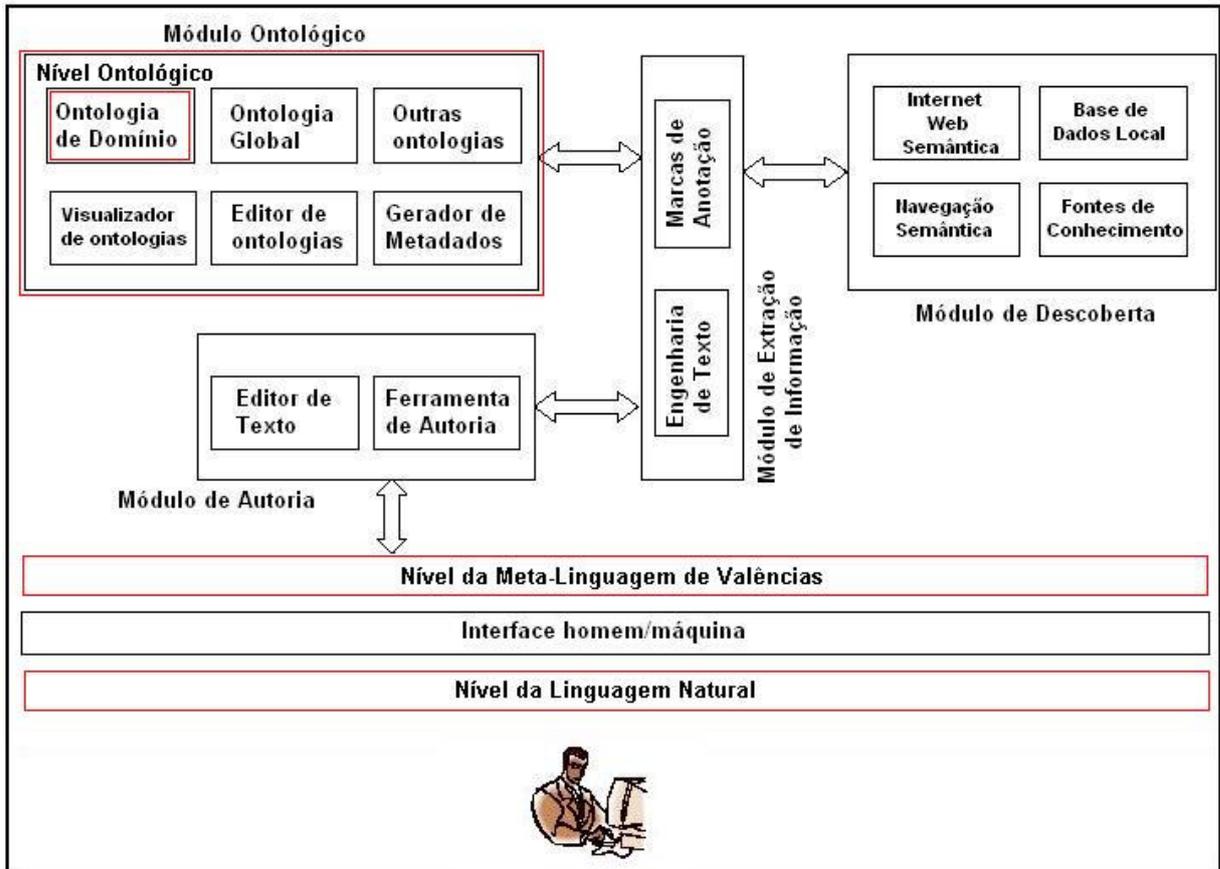


Figura 21: Encaixe dos níveis do arcabouço teórico no modelo de Lima-Marques & Oliveira

Adaptada de: Oliveira & Lima-Marques (2006)

Na seção a seguir descreve-se em detalhes a estrutura proposta para a ontologia de valências. Logo após, define-se a gramática para a meta-linguagem de valências. Por fim, verifica-se a aplicabilidade do referencial através da desambiguação de várias orações de exemplo.

## 4.2 - Ontologia das valências dos verbos, nomes, adjetivos e advérbios

Lembrando Almeida & Bax (2003), “os componentes básicos de uma ontologia são classes (organizadas em uma taxonomia), relações (representam o tipo de interação entre os conceitos de um domínio), axiomas (usados para modelar sentenças sempre verdadeiras) e instâncias (utilizadas para representar elementos específicos, ou seja, os próprios dados)”. Descreve-se abaixo cada uma das classes da ontologia. Logo após, passamos à descrição das propriedades, incluindo aí as propriedades de tipo de dados (*datatype properties*), aquelas que assumem um determinado valor, e as propriedades de objetos (*object propertires*), que são aquelas propriedades que possibilitam relacionar instâncias na linguagem OWL, sem assumir um valor específico.

Primeiramente é importante definir algumas convenções. Todos os nomes que identificam univocamente uma classe (ID), subclasse, propriedade ou instância não terão acentuação, cedilha (este será substituído por um c), e estarão escritos no singular. Para a linguagem OWL, estes nomes (IDs) identificarão de maneira única um item dentro da ontologia, não podendo, portanto serem repetidos. O nome das classes e subclasses será descrito em letras maiúsculas com o prefixo “CLASSE\_”. O nome (ID) das propriedades será escrito em letras minúsculas. Quando a propriedade se tratar de uma propriedade de tipo de dados, será identificada com prefixo “dprop\_”. Quando a propriedade se tratar de uma propriedade de objetos, será identificada com prefixo “oprop\_”. O nome (ID) das instâncias será escrito em letras minúsculas com o prefixo “ins\_”. Quando duas ou mais instâncias tiverem o mesmo ID, será adicionado um posfixo com numeração seqüencial (1,2,3, etc.) para diferenciá-las.

Para cada acepção de uma mesma forma gráfica, será adicionada uma instância. Por exemplo, no dicionário de usos de Borba (2002), a forma gráfica “persa” tem dez acepções. As informações do dicionário de usos de Borba (2002) serão complementadas pelas informações contidas no dicionário gramatical de verbos de Borba (1990).

A ontologia aqui proposta assume uma função mais genérica do que a ontologia proposta em Bräscher (1999). A presente ontologia reúne as funções desempenhadas pelo dicionário, gramática morfológica, gramática de argumentos e da própria ontologia proposta por Bräscher (1999).

No Anexo 1 desta dissertação, está disponibilizado um exemplo da ontologia aqui proposta em linguagem OWL.

#### 4.2.1 – Classes da ontologia

A figura abaixo descreve a taxonomia das classes utilizadas na ontologia. Esta figura foi gerada através do Protege, utilizando o *plugin* Jambalaya. Logo após a apresentação da figura, passa-se à descrição de cada uma das classes.

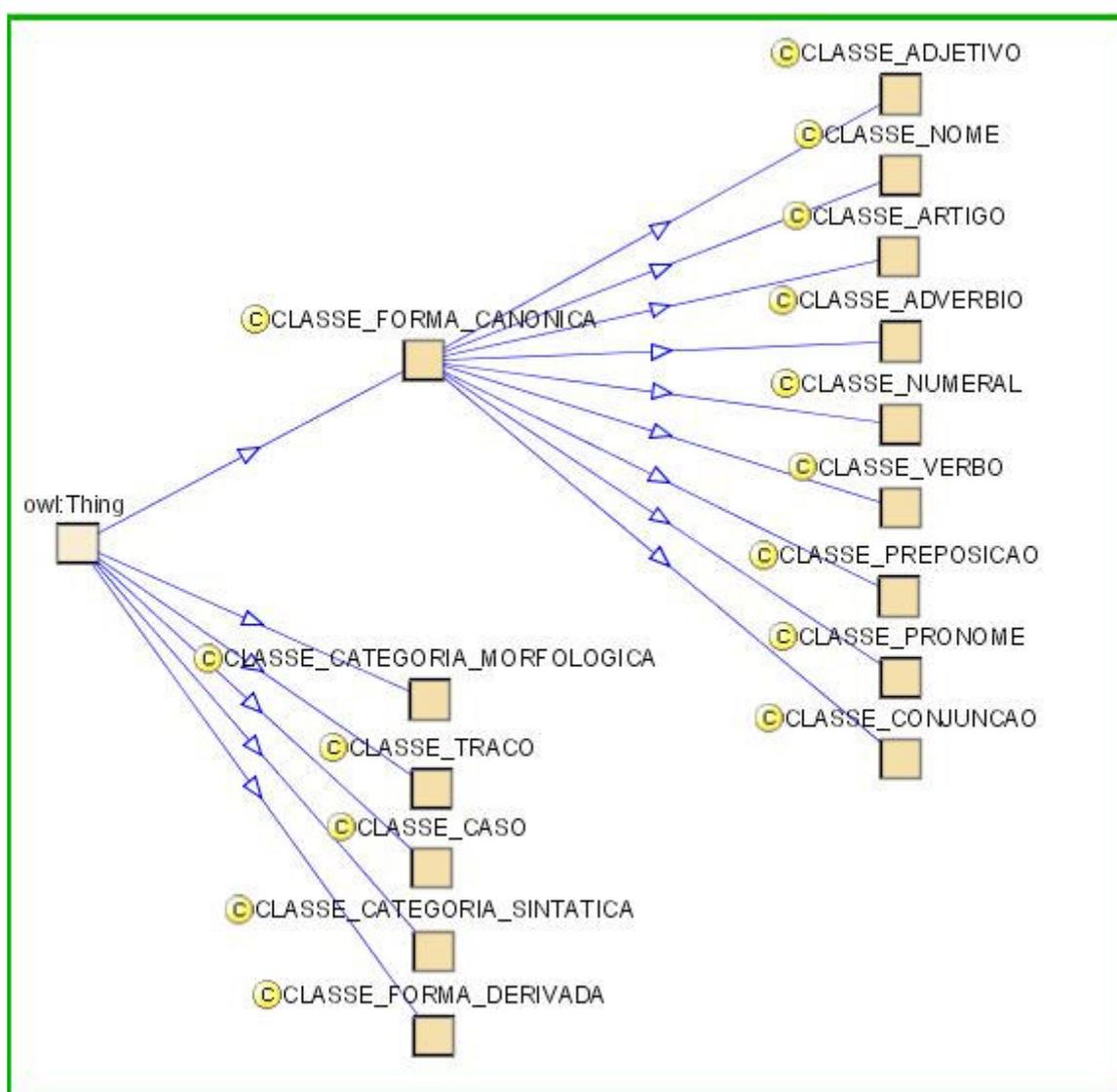


Figura 22: Taxonomia das classes da ontologia de valências

#### **4.2.1.1 - CLASSE\_FORMA\_CANONICA**

A primeira classe a ser descrita é a classe denominada “CLASSE\_FORMA\_CANONICA”. Uma forma canônica corresponde a um lema, ou seja, “unidade de coleta na sua forma gramatical neutra, i.e., sem conjugação, sem flexão, etc.” (BRÄSCHER, 1999). Esta classe equivale a um dicionário de base do Zstation conforme proposto por Bräscher (1999). Dentro desta classe residem nove subclasses, cada uma correspondente a uma das classes morfológicas determinadas na gramática de Cunha & Cintra (2001), a saber: Substantivos (aos quais chamaremos de Nomes, buscando integração com o texto de Borba (1996)), Adjetivos, Advérbios, Verbos, Pronomes, Artigos, Conjunções, Numerais e Preposições. O nome das subclasses fica portanto: CLASSE\_NOME, CLASSE\_ADJETIVO, CLASSE\_ADVERBIO, CLASSE\_VERBO, CLASSE\_PRONOME, CLASSE\_ARTIGO, CLASSE\_CONJUNCAO, CLASSE\_NUMERAL e CLASSE\_PREPOSICAO.

#### **4.2.1.2 – CLASSE\_FORMA\_DERIVADA**

Nesta classe estarão registradas, além das formas canônicas propriamente ditas (apenas por uma questão de performance computacional), todas as formas derivadas das formas canônicas presentes na CLASSE\_FORMA\_CANONICA. Entende-se por forma derivada aquelas que derivam dos lemas, ou seja, aquelas que são flexionadas em conjugação, gênero, diminutivo, aumentativo e número. Por exemplo, para a forma canônica “menino”, teremos as seguintes formas derivadas: “menino”, “meninos”, “meninão”, “meninões”, “menina”, “meninas”, “menininha”, “meninhas”. Esta classe pode ser encarada como “porta de entrada” da ontologia no momento da autoria do documento. Quando o usuário digita qualquer palavra no ambiente de autoria, esta palavra será buscada na CLASSE\_FORMA\_DERIVADA. Caso a palavra digitada seja forma derivada de mais de uma forma canônica, então o ambiente de autoria irá inquirir o usuário para que escolha a forma canônica sobre qual refere-se a forma derivada em questão.

#### **4.2.1.3 – CLASSE\_CATEGORA\_MORFOLOGICA**

As instâncias desta classe dizem respeito às classes morfológicas, a saber: Verbo, nome, adjetivo, advérbio, artigo, conjunção, numeral, preposição e pronome.

#### **4.2.1.4 – CLASSE\_CATEGORIA\_SINTATICA**

Nesta classe estarão cadastradas as categorias sintáticas conforme proposto na gramática da meta-linguagem de valências, disponível na próxima seção desta dissertação. Como pode-se perceber as instâncias desta classe estão organizadas em uma taxonomia de dois níveis. Para possibilitar esta taxonomia, abre-se mão de se utilizar as classes do OWL, partindo-se para uma solução que se utiliza de uma propriedade de objetos, denomina `oprop_e_um`, que faz a ligação taxonômica entre os níveis. Mais detalhes sobre esta hierarquia de categorias sintáticas podem ser obtidos na próxima seção.

#### **4.2.1.5 – CLASSE\_TRACO**

Nesta classe estarão registrados todas restrições semântico-selecionais, às quais chamaremos de traços semânticos, que podem estar associados a um determinado substantivo. Como verificado na revisão de literatura, a escolha do conjunto de restrições seletivas depende eminentemente do contexto para o qual a aplicação se dirige. Para os fins didáticos deste trabalho, as seguintes restrições serão utilizadas: abstrato, alimento, animado, concreto, contável, humano, não animado, não contável, não humano, sólido e líquido.

#### **4.2.1.6 – CLASSE\_CASO**

A função desta classe é registrar todos os caso, ou papéis temáticos, que podem surgir da relação entre um predicado e um argumento, conforme a teoria de Borba (1996). Os casos, conforme a teoria de Tesnière (1959) utilizada por Borba (1996), são: Agentivo, Experimentador, Beneficiário, Locativo, Direcional, Percurso, Objetivo, Instrumental, Causativo, Meta, Origem, Resultativo, Temporal e Comitativo

### **4.2.2 – Propriedades da ontologia de Valências**

Como se sabe, na linguagem OWL tem-se dois tipos de propriedades, as de tipo de dados e as de objetos. Descreve-se inicialmente as propriedades de tipo de dados e, logo após, as propriedades de objetos.

#### **4.2.2.1 – Propriedades de Tipo de Dados da ontologia de Valências**

##### **4.2.2.1.1 – ID**

A propriedade ID corresponde ao nome, código ou número, dado para um determinado item da ontologia, através do qual o item será identificado univocamente dentro da estrutura de representação de conhecimento. Um ID, portanto, não pode ser repetido em hipótese alguma, mesmo quando se tratar de componentes ontológicos de categorias diferentes, como classes, instâncias e propriedades.

Domínio: owl: Thing

Tipo de Dado: String

Funcional: Sim

##### **4.2.2.1.2 - dprop\_nome**

Esta propriedade corresponde ao nome de uma instância desta classe morfológica, em sua forma canônica. É importante diferenciar esta propriedade do nome que identifica univocamente um item na ontologia (ID). Dessa forma, por exemplo, o valor da propriedade dprop\_nome para o verbo correr será “correr”, enquanto que seu ID será “ins\_correr”.

Domínio: CLASSE\_TRACO, CLASSE\_CASO, CLASSE\_CATEGORIA\_SINTATICA,  
CLASSE\_FORMA\_CANONICA, CLASSE\_FORMA\_DERIVADA,  
CLASSE\_CATEGORIA MORFOLOGICA.

Tipo de Dado: String

Funcional: Sim

#### 4.2.2.1.3 - dprop\_descricao

Esta propriedade conterá a descrição de cada acepção cadastrada na ontologia. Ou seja, para a primeira acepção do verbo correr, teremos, conforme Borba (2002): “fazer deslizar”.

Domínio: CLASSE\_FORMA\_CANONICA.

Tipo de Dado: String

Funcional: Sim

#### 4.2.2.1.4 - dprop\_numero\_argumento

Esta propriedade está relacionada à valência quantitativa da teoria de Borba (1996), ou seja, quantos argumentos o verbo exige para se tornar auto-suficiente semanticamente. Conforme reza a teoria, pode variar de zero a quatro.

Domínio: CLASSE\_ADJETIVO, CLASSE\_ADVERBIO, CLASSE\_VERBO, CLASSE\_NOME.

Tipo de Dado: Int

Funcional: Sim

#### 4.2.2.1.5 - dprop\_arg1\_obrigatorio

Esta propriedade está relacionada à valência sintática da teoria de Borba (1996). Esta propriedade indica se o primeiro argumento exigido pelo verbo é de preenchimento obrigatório.

Domínio: CLASSE\_ADJETIVO, CLASSE\_ADVERBIO, CLASSE\_VERBO, CLASSE\_NOME.

Tipo de Dado: Boolean.

Funcional: Sim

#### 4.2.2.1.6 - dprop\_arg2\_obrigatorio

Esta propriedade está relacionada à valência sintática da teoria de Borba (1996). Esta propriedade indica se o segundo argumento exigido pelo verbo é de preenchimento obrigatório.

Domínio: CLASSE\_ADJETIVO, CLASSE\_ADVERBIO, CLASSE\_VERBO, CLASSE\_NOME.

Tipo de Dado: Boolean.

Funcional: Sim

#### 4.2.2.1.7 - dprop\_arg3\_obrigatorio

Esta propriedade está relacionada à valência sintática da teoria de Borba (1996). Esta propriedade indica se o terceiro argumento exigido pelo verbo é de preenchimento obrigatório.

Domínio: CLASSE\_ADJETIVO, CLASSE\_ADVERBIO, CLASSE\_VERBO, CLASSE\_NOME.

Tipo de Dado: Boolean.

Funcional: Sim

#### 4.2.2.1.8 - dprop\_arg4\_obrigatorio

Esta propriedade está relacionada à valência sintática da teoria de Borba (1996). Esta propriedade indica se o quarto argumento exigido pelo verbo é de preenchimento obrigatório.

Domínio: CLASSE\_ADJETIVO, CLASSE\_ADVERBIO, CLASSE\_VERBO, CLASSE\_NOME.

Tipo de Dado: Boolean.

Funcional: Sim

#### **4.2.2.1.9 - dprop\_abreviatura**

Nesta propriedade estarão redigidas as abreviaturas, redução da forma de uma palavra, das instâncias das classes de domínio relacionadas abaixo.

Domínio: CLASSE\_CASO, CLASSE\_CATEGORIA\_SINTATICA,  
CLASSE\_CATEGORIA\_MORFOLOGICA.

Tipo de Dado: String

Funcional: Sim

#### **4.2.2.2 - Propriedades de Objetos da ontologia de Valências**

##### **4.2.2.2.1 - oprop\_e\_forma\_derivada\_de**

Permite que uma única forma derivada seja associada a zero ou várias formas canônicas diferentes. A forma derivada “banco”, por exemplo, pode ser relacionada a diversas formas canônicas, a saber: “banco” (substantivo – instituição financeira), “banco” (substantivo – assento) e “bançar” (verbo, cuja conjugação na primeira pessoa do singular no presente do indicativo é “banco”. Ex.: “Eu banco.”).

Domínio: CLASSE\_FORMA\_DERIVADA.

Alcance: CLASSE\_FORMA\_CANONICA.

Funcional: Não.

##### **4.2.2.2.2 - oprop\_tem\_cat\_sint\_arg1**

Esta propriedade está relacionada à valência sintática da teoria de Borba (1996). Através desta propriedade será especificada a categoria sintática, conforme estabelecido pela gramática da meta-linguagem de valências, do primeiro argumento exigido pelo verbo.

Domínio: CLASSE\_ADJETIVO, CLASSE\_ADVERBIO, CLASSE\_VERBO,  
CLASSE\_NOME.

Alcance: CLASSE\_CATEGORIA\_SINTATICA

Funcional: Não.

#### **4.2.2.2.3 - oprop\_tem\_cat\_sint\_arg2**

Esta propriedade está relacionada à valência sintática da teoria de Borba (1996). Através desta propriedade será especificada a categoria sintática, conforme estabelecido pela gramática da meta-linguagem de valências, do segundo argumento exigido pelo verbo.

Domínio: CLASSE\_ADJETIVO, CLASSE\_ADVERBIO, CLASSE\_VERBO,  
CLASSE\_NOME.

Alcance: CLASSE\_CATEGORIA\_SINTATICA

Funcional: Não.

#### **4.2.2.2.4 - oprop\_tem\_cat\_sint\_arg3**

Esta propriedade está relacionada à valência sintática da teoria de Borba (1996). Através desta propriedade será especificada a categoria sintática, conforme estabelecido pela gramática da meta-linguagem de valências, do terceiro argumento exigido pelo verbo.

Domínio: CLASSE\_ADJETIVO, CLASSE\_ADVERBIO, CLASSE\_VERBO,  
CLASSE\_NOME.

Alcance: CLASSE\_CATEGORIA\_SINTATICA

Funcional: Não.

#### **4.2.2.2.5 - oprop\_tem\_cat\_sint\_arg4**

Esta propriedade está relacionada à valência sintática da teoria de Borba (1996). Através desta propriedade será especificada a categoria sintática, conforme estabelecido pela gramática da meta-linguagem de valências, do quarto argumento exigido pelo verbo.

Domínio: CLASSE\_ADJETIVO, CLASSE\_ADVERBIO, CLASSE\_VERBO,  
CLASSE\_NOME.

Alcance: CLASSE\_CATEGORIA\_SINTATICA

Funcional: Não.

#### **4.2.2.2.6 - oprop\_tem\_prep\_arg1**

Esta propriedade está relacionada à valência sintática da teoria de Borba (1996). Caso o primeiro argumento deva ser precedido de uma preposição, esta propriedade possibilitará listar todas as possíveis preposições que podem preceder o argumento.

Domínio: CLASSE\_ADJETIVO, CLASSE\_ADVERBIO, CLASSE\_VERBO,  
CLASSE\_NOME.

Alcance: CLASSE\_PREPOSICAO.

Funcional: Não.

#### **4.2.2.2.7 - oprop\_tem\_prep\_arg2**

Esta propriedade está relacionada à valência sintática da teoria de Borba (1996). Caso o segundo argumento deva ser precedido de uma preposição, esta propriedade possibilitará listar todas as possíveis preposições que podem preceder o argumento.

Domínio: CLASSE\_ADJETIVO, CLASSE\_ADVERBIO, CLASSE\_VERBO,  
CLASSE\_NOME.

Alcance: CLASSE\_PREPOSICAO.

Funcional: Não.

#### **4.2.2.2.8 - oprop\_tem\_prep\_arg3**

Esta propriedade está relacionada à valência sintática da teoria de Borba (1996). Caso o terceiro argumento deva ser precedido de uma preposição, esta propriedade possibilitará listar todas as possíveis preposições que podem preceder o argumento.

Domínio: CLASSE\_ADJETIVO, CLASSE\_ADVERBIO, CLASSE\_VERBO, CLASSE\_NOME.

Alcance: CLASSE\_PREPOSICAO.

Funcional: Não.

#### **4.2.2.2.9 - oprop\_tem\_prep\_arg4**

Esta propriedade está relacionada à valência sintática da teoria de Borba (1996). Caso o quarto argumento deva ser precedido de uma preposição, esta propriedade possibilitará listar todas as possíveis preposições que podem preceder o argumento.

Domínio: CLASSE\_ADJETIVO, CLASSE\_ADVERBIO, CLASSE\_VERBO, CLASSE\_NOME.

Alcance: CLASSE\_PREPOSICAO.

Funcional: Não.

#### **4.2.2.2.10 - oprop\_tem\_caso1**

Esta propriedade está relacionada à valência semântica da teoria de Borba (1996). Esta propriedade indica qual é o caso, ou papel temático, que surge da relação entre o primeiro argumento com o núcleo do predicado verbal.

Domínio: CLASSE\_ADJETIVO, CLASSE\_VERBO, CLASSE\_NOME.

Alcance: CLASSE\_PREPOSICAO.

Funcional: Sim.

#### **4.2.2.2.11 - oprop\_tem\_caso2**

Esta propriedade está relacionada à valência semântica da teoria de Borba (1996). Esta propriedade indica qual é o caso, ou papel temático, que surge da relação entre o segundo argumento com o núcleo do predicado verbal.

Domínio: CLASSE\_ADJETIVO, CLASSE\_VERBO, CLASSE\_NOME.

Alcance: CLASSE\_PREPOSICAO.

Funcional: Sim.

#### **4.2.2.2.12 - oprop\_tem\_caso3**

Esta propriedade está relacionada à valência semântica da teoria de Borba (1996). Esta propriedade indica qual é o caso, ou papel temático, que surge da relação entre o terceiro argumento com o núcleo do predicado verbal.

Domínio: CLASSE\_ADJETIVO, CLASSE\_VERBO, CLASSE\_NOME.

Alcance: CLASSE\_PREPOSICAO.

Funcional: Sim.

#### **4.2.2.2.13 - oprop\_tem\_caso4**

Esta propriedade está relacionada à valência semântica da teoria de Borba (1996). Esta propriedade indica qual é o caso, ou papel temático, que surge da relação entre o quarto argumento com o núcleo do predicado verbal.

Domínio: CLASSE\_ADJETIVO, CLASSE\_VERBO, CLASSE\_NOME.

Alcance: CLASSE\_PREPOSICAO.

Funcional: Sim.

#### **4.2.2.2.14 - oprop\_tem\_traco\_semantico\_arg1**

Esta propriedade diz respeito à valência semântica da teoria de Borba (1996). Caso seja exigido que o primeiro argumento do verbo tenha um traço semântico em especial, nesta propriedade estarão registradas estas restrições seletivas.

Domínio: CLASSE\_ADJETIVO, CLASSE\_VERBO, CLASSE\_NOME.

Alcance: CLASSE\_TRACO.

Funcional: Não.

#### **4.2.2.2.15 - oprop\_tem\_traco\_semantico\_arg2**

Esta propriedade diz respeito à valência semântica da teoria de Borba (1996). Caso seja exigido que o segundo argumento do verbo tenha um traço semântico em especial, nesta propriedade estarão registradas estas restrições seletivas.

Domínio: CLASSE\_ADJETIVO, CLASSE\_VERBO, CLASSE\_NOME.

Alcance: CLASSE\_TRACO.

Funcional: Não.

#### **4.2.2.2.16 - oprop\_tem\_traco\_semantico\_arg3**

Esta propriedade diz respeito à valência semântica da teoria de Borba (1996). Caso seja exigido que o terceiro argumento do verbo tenha um traço semântico em especial, nesta propriedade estarão registradas estas restrições seletivas.

Domínio: CLASSE\_ADJETIVO, CLASSE\_VERBO, CLASSE\_NOME.

Alcance: CLASSE\_TRACO.

Funcional: Não.

#### **4.2.2.2.17 - oprop\_tem\_traco\_semantico\_arg4**

Esta propriedade diz respeito à valência semântica da teoria de Borba (1996). Caso seja exigido que o quarto argumento do verbo tenha um traço semântico em especial, nesta propriedade estarão registradas estas restrições seletivas.

Domínio: CLASSE\_ADJETIVO, CLASSE\_VERBO, CLASSE\_NOME.

Alcance: CLASSE\_TRACO.

Funcional: Não.

#### **4.2.2.2.18 - oprop\_tem\_traco\_semantico**

Esta propriedade atende à valência semântica da teoria de Borba (1996). Através desta propriedade serão indicados os traços semânticos de um substantivo.

Domínio: CLASSE\_NOME.

Alcance: CLASSE\_TRACO.

Funcional: Não.

#### **4.2.2.2.19 - oprop\_e\_um**

A função desta propriedade é possibilitar uma hierarquia taxonômica sem a necessidade da criação de classes. Esta propriedade constitui-se em um subterfúgio para driblar uma limitação do owl, relativa às propriedades de objetos. As propriedades de objetos ligam instâncias de uma classe à instâncias de outra classe. Não é possível ligar uma instância a uma classe inteira.

Domínio: CLASSE\_CATEGORIA\_SINTATICA.

Alcance: CLASSE\_CATEGORIA\_SINTATICA.

Funcional: Sim.

### 4.2.2.3 – Visualizando a edição de uma forma canônica no Protègè

A figura abaixo demonstra, a título de exemplo, a edição da forma canônica do verbo “traduzir” na interface da ferramenta Protègè. Na figura é possível observar que o verbo em questão tem quatro argumentos, sendo que apenas o primeiro é obrigatório. Pode-se constatar também as exigências sintáticas e semânticas para o preenchimento dos quatro argumentos do verbo.

<b>Dprop Nome</b>		<b>Dprop Numero Argumento</b>	
Traduzir		4	
<b>Dprop Descricao</b>			
Transpor de um idioma para outra			
<input checked="" type="checkbox"/> Dprop Arg1 Obrigatorio <input type="checkbox"/> Dprop Arg2 Obrigatorio <input type="checkbox"/> Dprop Arg3 Obrigatorio <input type="checkbox"/> Dprop Arg4 Obrigatorio			
<b>Oprop Tem Cat Sint Arg1</b>		<b>Oprop Tem Cat Sint Arg2</b>	
<ul style="list-style-type: none"> <li>◆ Functor Nominal</li> <li>◆ Functor Adjetivo</li> </ul>		<ul style="list-style-type: none"> <li>◆ Functor Nominal</li> <li>◆ Functor Adjetivo</li> </ul>	
<b>Oprop Tem Cat Sint Arg3</b>		<b>Oprop Tem Cat Sint Arg4</b>	
<ul style="list-style-type: none"> <li>◆ Functor Nominal</li> <li>◆ Functor Adjetivo</li> </ul>		<ul style="list-style-type: none"> <li>◆ Functor Nominal</li> <li>◆ Functor Adjetivo</li> </ul>	
<b>Oprop Tem Prep Arg1</b>		<b>Oprop Tem Prep Arg2</b>	
<b>Oprop Tem Prep Arg3</b>		<b>Oprop Tem Prep Arg4</b>	
◆ de		◆ para	
<b>Oprop Tem Traco Semantico Arg1</b>		<b>Oprop Tem Traco Semantico Arg2</b>	
		◆ Concreto	
<b>Oprop Tem Traco Semantico Arg3</b>		<b>Oprop Tem Traco Semantico Arg4</b>	
◆ Idioma		◆ Idioma	
<b>Oprop Tem Caso1</b>		<b>Oprop Tem Caso2</b>	
◆ Agentivo		◆ Objetivo	
<b>Oprop Tem Caso3</b>		<b>Oprop Tem Caso4</b>	
◆ Origem		◆ Meta	

Figura 23: Editando uma forma canônica no Protègè

### 4.3 - Gramática da Meta-linguagem de Valências

Toda linguagem pressupõe uma gramática. Porém, antes de partir para a formalização da gramática da meta-linguagem de valências, resolveremos um problema inacabado na revisão de literatura, relativo à sintaxe dos advérbios na gramática de valências. Logo após, descreve-se, de maneira informal, a forma pretendida para as sentenças da meta-linguagem, para só depois definirmos a gramática formal desta linguagem.

Como verificado, as informações sobre a valência dos advérbios são mínimas. A teoria da gramática de valências, em função desta limitação, é uma teoria inacabada. A única pista que Borba (1996) deixa é que alguns advérbios parecem ser valenciais (advérbios de lugar, especificamente, os de localização relativa). Esta lacuna teórica causa um problema a esta dissertação, já que a utilização de advérbios, locuções e adjuntos adverbiais, consistem em uma importante causa de ambigüidades sintáticas (veja o caso de “Eu vi a menina com o telescópio”, onde “com o telescópio” pode ser um adjunto adverbial de instrumento ou complemento nominal de “a menina”).

Mesmo sem exaurir teoricamente o problema da valência dos advérbios, pois isto seria assunto para uma nova dissertação, é necessário tomar decisões sobre como abordar este tema na solução das ambigüidades através do referencial proposto. Acataremos a sugestão de Borba (1996), com relação aos advérbios de lugar, que estes necessitam de um argumento, um sintagma nominal, para se concretizarem semanticamente (Ex.: embaixo da cama, sobre o cavalo, adiante da escola). Os demais advérbios seriam avalentes.

Será necessário também controlar o escopo de ação dos advérbios, ou seja, determinar a área de ação que um advérbio tem sobre um outro advérbio, adjetivo ou sintagma nominal. A necessidade deste controle está baseada na afirmação de Cunha & Cintra (2001): "O advérbio é, fundamentalmente, um modificador do verbo". Ainda segundo Cunha & Cintra (2001), o advérbio modifica também um outro advérbio ou um adjetivo. A sentença abaixo, escrita em meta-linguagem de valências, demonstra nossa intenção sobre a valência e o controle de escopo dos advérbios.

- a) embaixo (da cama)                      (pediu (a menina, que comprem (eles, um carro))).<sup>25</sup>  
 Advérbio Valencial                              Escopo de Ação do Advérbio “embaixo”

Outro problema relacionado aos advérbios é a questão das locuções adverbiais. “Uma locução adverbial é um conjunto de duas ou mais palavras que funciona como advérbio. De regra, formam-se da associação de uma preposição com um substantivo, com um adjetivo ou com um advérbio”.(Cunha & Cintra, 2001). Como o ambiente de autoria poderia detectar que uma locução se trata de uma locução adverbial? Uma possível solução para este problema seria fazer com que o ambiente de autoria perguntasse ao usuário, toda vez que este digitasse uma forma gráfica relacionada a um verbo, quais são as possíveis modificações que aquela ação poderia sofrer. A resposta a estas perguntas seria limitada por um universo de opções dado pelo ambiente de autoria, que conduziria o usuário sintaticamente. Exemplificando, quando um usuário digitasse a forma gráfica “veja”, relacionada ao verbo “ver”, o ambiente de autoria faria as seguintes perguntas: Como (Advérbio de Modo)? Quando (Advérbio de Tempo)? Onde (Advérbio de Lugar)? Com que instrumento (Advérbio de Instrumento)? A resposta a estas perguntas seria opcional, pois, de fato, não é obrigatório que a ação seja sempre alvo de um advérbio. Mas, caso o usuário optasse por responder uma das perguntas, a relativa ao advérbio de instrumento, por exemplo, o ambiente de autoria guiaria a sintaxe da resposta, que seria regida por uma regra como: PREP + Nome (Ex.: “com o telescópio”, “com o machado”).

Passemos agora a descrever de maneira informal, como seriam sentenças válidas na meta-linguagem de valências. Começa-se com as sentenças mais simples, aquelas vazias ou com uma única palavra de valência zero, ou seja, uma palavra plenamente concretizada semanticamente. Para as sentenças vazias, utilizaremos o símbolo  $\epsilon$ . Para as sentenças com uma única palavra, apenas a própria palavra e o ponto final, indicando o fim da sentença.

- a) Pedra.
- b) Tigre.
- c) Choveu.

---

<sup>25</sup> Sentença equivalente à frase redigida em linguagem natural “A menina pediu embaixo da cama que eles comprem um carro.”.

Passa-se então a descrever informalmente como seria o comportamento dos itens lexicais valenciais. De acordo com a gramática de valências de Borba (1996), a sentença em meta-linguagem de valência será escrita da seguinte forma: (1) o núcleo do predicado, seja ele verbal, nominal, adjetivo ou adverbial, fará a função de um functor<sup>26</sup>; (2) Os argumentos deste functor serão organizados na ordem estabelecida nos dicionários de Borba (Borba (2002) e Borba (1990)), que estão formalizadas na ontologia de valências. Percebe-se esta organização nos exemplos abaixo:

- a) traduziu (a menina, o livro, do português, para o inglês).
- b) censurou (o general, o estudante).
- c) censura (pelo general, do estudante).
- d) embaixo (da cama) (traduziu (a menina, o livro, do português, para o inglês)).
- e) muito (vasta (uma cabeleira)).
- f) havíamos (comprado (nós, os livros)).

Algumas características desta linguagem podem ser evidenciadas através dos exemplos acima:

- a) A ordem dos argumentos obedece ao paralelismo entre verbo, nome e adjetivo, ou seja, por mais que na linguagem natural se diga usualmente “a censura do estudante pelo general”, a ordem dos argumentos obedece à ordem da classe gramatical primária (o verbo);
- b) Percebe-se o comportamento de um advérbio, que, no exemplo “d”, além de ter um argumento direto, “da cama”, engloba o functor verbal, pois modifica-o. No exemplo “e” percebe-se um advérbio modificando um adjetivo, que modifica um sintagma nominal.
- c) A utilização de preposições nos argumentos também obedece às regras estabelecidas por Borba;
- d) Percebe-se no exemplo “f” a relação entre um verbo auxiliar e um verbo núcleo do predicado verbal, na forma nominal (particípio);
- e) Uma mesma sentença em meta-linguagem de valências pode ser traduzida (compilada) em diversas formas diferentes em linguagem natural. Isto é possível

---

<sup>26</sup> Functor é um termo utilizado na programação em lógica. Consiste uma relação estabelecida entre argumentos.

através da utilização dos expedientes de alçamento, rebaixamento, apagamento, truncamento, transposição, dissimilação, coordenação e condensação. Um exemplo simples, mostra como uma mesma sentença em meta-linguagem de valências pode ser reescrita em diferentes formas em linguagem natural, utilizando-se destes expedientes..

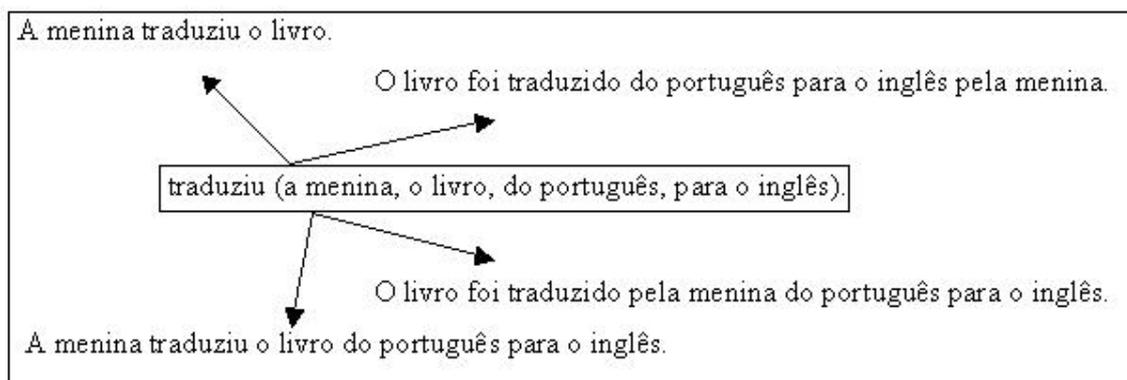


Figura 24: Várias formas em linguagem natural para uma única forma em MLV

Por fim, antes de adentrarmos na formalização da gramática, é preciso ligar cada palavra escrita em meta-linguagem de valências com a ontologia de valências. A ligação do meta-documento com a ontologia é fundamental para recuperação da informação e para desambiguação. Isto seria feito através da inserção do ID da forma canônica da ontologia, referente àquela palavra escrita no meta-documento. Desta forma teríamos sentenças descritas da forma como segue:

- a) traduziu [ins\_traduzir] (a [ins\_o] menina [ins\_menino], o [ins\_o] livro [ins\_livro], do [ins\_de] português [ins\_português], para [ins\_para] o [ins\_o] inglês [ins\_inglês]).
- b) muito [ins\_muito] (vasta [ins\_vasto] (uma [ins\_um] cabeleira [ins\_cabeleira])).

Com a noção das sentenças válidas para a meta-linguagem de valências, passemos à formalização da gramática. É importante lembrar que, de acordo com as regras estabelecidas, o nível de restrição de uma gramática pode ser ajustado. Para os fins didáticos desta dissertação, as poucas regras estabelecidas permitirão uma flexibilidade grande para a geração de sentenças, entretanto, uma susceptibilidade a gerar sentenças com pouco ou nenhum

sentido. Melhorar a gramática proposta aqui é um dos trabalhos futuros a serem desenvolvidos.

De fato, a gramática da meta-linguagem de valências pode ser entendida pela soma das micro-gramáticas de cada forma canônica dos itens lexicais na ontologia (Nomes, verbos, advérbios e adjetivos). Ou seja, nos quadros (*frames*) de cada item lexical estará contida a gramática para este item lexical. Vejamos o exemplo de uma das formas canônicas associadas à forma gráfica “comer”:

Quadro 4 – Frame de uma das formas canônicas associadas à forma gráfica “comer”

Propriedade	Valor
ID	ins_comer1
dprop_nome	Comer
dprop_descricao	Engolir para se alimentar, depois de ter mastigado, ingerir.
oprop_pertence_a_categoria_morfologica	Verbo
dprop_numero_argumento	2
dprop_arg1_obrigatorio	Sim
dprop_arg2_obrigatorio	Sim
dprop_arg3_obrigatorio	Não se aplica.
dprop_arg4_obrigatorio	Não se aplica.
oprop_tem_categoria_sintatica_arg1	FN ou FADJ.
oprop_tem_categoria_sintatica_arg2	FN ou FADJ.
oprop_tem_categoria_sintatica_arg3	Não se aplica.
oprop_tem_categoria_sintatica_arg4	Não se aplica.
oprop_tem_prep_arg1	Não se aplica.
oprop_tem_prep_arg2	Não se aplica.
oprop_tem_prep_arg3	Não se aplica.
oprop_tem_prep_arg4	Não se aplica.
oprop_tem_caso1	Agentivo
oprop_tem_caso2	Objetivo
oprop_tem_caso3	Não se aplica.
oprop_tem_caso4	Não se aplica.
oprop_tem_traco_semantico_arg1	Animado
oprop_tem_traco_semantico_arg2	Alimento Sólido Concreto
oprop_tem_traco_semantico_arg3	Não se aplica.
oprop_tem_traco_semantico_arg4	Não se aplica.

Podemos determinar se as sentenças abaixo são gramaticais ou agramaticais para a meta-linguagem de valências com o auxílio do quadro acima:

- a) comeu (bela (mulher), o estrogonofe)

Validação da Sentença: Sentença Gramatical. “comeu” é uma forma derivada de “comer”; “bela (mulher)” corresponde a um functor adjetivo (FADJ) que tem a função de sujeito agente da ação de comer e também tem o traço semântico exigido (Animado); “o estrogonofe” é um functor nominal, que tem os traços semânticos exigidos: “Alimento”, “Sólido”, “Concreto”.

b) comeria (bela (mulher), o amor)

Validação da Sentença: Sentença Agramatical. “comeria” é uma forma derivada de “comer”; “bela (mulher)” corresponde a um functor adjetivo (FADJ) que tem a função de sujeito agente da ação de comer e também tem o traço semântico exigido (Animado); “o amor” é um functor nominal, porém, não atende a exigência dos traços-semânticos, pois é “Abstrato”.

c) comerá (mulher, do estrogonofe)

Validação da Sentença: Sentença Agramatical. “comerá” é uma forma derivada de “comer”; “mulher” corresponde a um functor nominal (FN) que tem a função de sujeito agente da ação de comer e também tem o traço semântico exigido (Animado); “estrogonofe” é um nome, que tem os traços semânticos exigidos: “Alimento”, “Sólido”, “Concreto”. Entretanto, “estrogonofe” vem precedido de uma preposição, o que fere a gramática do item lexical.

Apesar de constatarmos que a gramática da meta-linguagem de valências está descrita através dos *frames* de cada item lexical, podemos estabelecer um conjunto de regras gramaticais gerais, que abarcam as micro-gramáticas de cada item lexical. Para tal é preciso lembrar que uma gramática  $G$  pode ser representada por uma quádrupla  $G = (V, T, P, S)$  onde:

- a.  $V$  é um conjunto finito de símbolos não terminais, ou variáveis;
- b.  $T$  é um conjunto finito de símbolos terminais ( $V \cap T = \{\}$ );
- c.  $P$  é um conjunto finito de pares  $(\alpha, \beta)$ , chamadas regras de produção, tal que  $l$  é uma palavra não vazia de  $V \cup T$ , e  $r$  é uma palavra possivelmente vazia de  $V \cup T$ ;
- d.  $S$  é o símbolo inicial, tal que  $S \in V$ .

Começaremos pela definição dos símbolos não terminais (V), todos redigidos em letras maiúsculas:

S – Símbolo inicial.

FADV - Functor Adverbial Genérico

FADV0 – Functor Adverbial com zero argumento.

FADV1 – Functor Adverbial com um argumento.

FADJ - Functor Adjetivo Genérico

FADJ1 - Functor Adverbial com um argumento.

FADJ2 - Functor Adverbial com dois argumentos.

FADJ3 - Functor Adverbial com três argumentos.

FADJ4 - Functor Adverbial com quatro argumentos.

FN - Functor Nominal Genérico

FN0 - Functor Nominal com zero argumento.

FN1 - Functor Nominal com um argumento.

FN2 - Functor Nominal com dois argumentos.

FN3 - Functor Nominal com três argumentos.

FN4 - Functor Nominal com quatro argumentos.

FV - Functor Verbal Genérico

FV0 – Functor Verbal com zero argumento.

FV1 – Functor Verbal com um argumento.

FV2 – Functor Verbal com dois argumentos.

FV3 – Functor Verbal com três argumentos.

FV4 – Functor Verbal com quatro argumentos.

ADV - Advérbio

ADJ - Adjetivo

N - Nome

V - Verbo

DET – Determinante (Pronomes demonstrativos e artigos)

PREP – Preposição

NUM – Numeral

VAUX – Verbo Auxiliar

Os símbolos terminais (T) desta gramática são todas instâncias da CLASSE\_FORMA\_DERIVADA da ontologia de valências. Estes são escritos em letras minúsculas. Portanto temos:  $\epsilon$ , abóbora, carro, amor, fazer, feito, de, para, com, o, a, zoeira, zebra, etc.

Antes de definir as regras de produção (P) da gramática, vamos a algumas convenções: (1) Utiliza-se a seta para a direita ( $\rightarrow$ ) com o sentido de “reescreva” ou “substitua” (Ex.:  $A \rightarrow B$ , quer dizer substitua A por B); (2) Utiliza-se o símbolo barra vertical (“|”) para indicar a operação lógica “ou exclusivo”, ou seja, para que “A|B” seja verdadeiro é necessário que ou A seja verdadeiro, ou B seja verdadeiro, mas não ambos ao mesmo tempo; (3) O símbolo  $\epsilon$  equivale a uma sentença vazia; (4) As chaves indicam a utilização opcional de uma variável; (5) Os parênteses indicam a área onde estarão dispostos os argumentos de um functor; (6) As vírgulas separam os argumentos; (7) O ponto final indica o final da sentença.

Passamos agora às regras de produção propriamente ditas. A partir do item “b”, serão dados exemplos sobre a aplicação da regra. Os exemplos serão compostos por uma frase em linguagem natural, e pela mesma frase redigida em meta-linguagem de valências, conforme as regras gramaticais determinadas.

a) Regras gerais:

$$1: S \rightarrow \epsilon.$$

$$2: S \rightarrow FV.$$

$$3: S \rightarrow FN.$$

$$4: S \rightarrow FADJ.$$

$$5: S \rightarrow FADV.$$

$$6: FV \rightarrow FV0.$$

$$7: FV \rightarrow FV1.$$

$$8: FV \rightarrow FV2.$$

$$9: FV \rightarrow FV3.$$

$$10: FV \rightarrow FV4.$$

$$11: FN \rightarrow FN0.$$

$$12: FN \rightarrow FN1.$$

- 13: FN → FN2.  
 14: FN → FN3.  
 15: FN → FN4  
 16: FADJ → FADJ1.  
 17: FADJ → FADJ2.  
 18: FADJ → FADJ3.  
 19: FADJ → FADJ4.  
 20: FADV → FADV0.  
 21: FADV → FADV1.  
 22: V → correr [ID], chover [ID], transferir [ID], traduzir [ID], e demais verbos.  
 23: N → carro [ID], bola [ID], tamanduá [ID], pedra [ID], e demais nomes.  
 24: ADJ → belo [ID], bonito [ID], transferível [ID], rápido [ID], e demais adjetivos.  
 25: ADV → freqüentemente [ID], aqui [ID], agora [ID], já [ID], e demais advérbios.  
 26: DET → este [ID], esse [ID], aquele, o [ID], e demais determinantes.  
 27: PREP → a [ID], ante [ID], após [ID], perante [ID], e demais preposições.  
 28: NUM → um [ID], dois [ID], três [ID], e demais numerais.  
 29: VAUX → ter [ID], haver [ID], ser [ID], estar [ID], e demais verbos auxiliares.

Observação: A sigla “ID” será sempre substituída pela seqüência de caracteres que identifica univocamente a forma canônica em questão na ontologia de valências.

#### b) Regras para Verbos

- 30: FV0 → {VAUX} V | FADV.  
 Ex.: Choveu.  
       Choveu.  
 Ex.: Nevou muito.  
       muito (nevou).  
 31: FV1 → {VAUX} V (FN | FADJ).  
 Ex.: João Pedro dormiu.  
       dormiu (João Pedro).  
 32: FV2 → {VAUX} V (FN | FADJ, {PREP} FN | FADJ).

Ex.: Daniel poderia morar em Brasília.

poderia morar (Daniel, em Brasília).

33: FV3 → {VAUX} V (FN | FADJ, FN | FADJ, PREP FN | FADJ).

Ex.: Daniel deu um carro rápido para João Pedro.

deu( Daniel, Rápido(um carro), para João Pedro).

34: FV4 → {VAUX} V (FN | FADJ, FN | FADJ, de FN | FADJ, para FN | FADJ).

Ex.: A diretoria teria traduzido o poema do árabe para o português arcaico.

teria traduzido (a diretoria, o poema, do árabe, para o arcaico (português)).

### c) Regras para Nomes Concretos

35: FN0 → {DET} {NUM} N.

Ex.: Uns oito carros.

Uns oito carros.

Ex.: Árvore.

Árvore.

36: FN1 → {DET} {NUM} N (PREP FN | FADJ).

Ex.: Amante da bela mãe.

amante (da bela (mãe)).

37: FN2 → {DET} {NUM} N (de FN | FADJ, a FN | FADJ).

Ex.: Pedinte de esmolas aos abastados.

pedintes (de esmolas, aos abastados).

38: FN3 → {DET} {NUM} N (de FN | FADJ, de FN | FADJ, para FN | FADJ).

Ex.: O tradutor de poemas do árabe para o português.

o tradutor (de poemas, do árabe, para o português).

### d) Regras para Nomes Abstratos

39: FN1 → {DET} {NUM} N (PREP FN | FADJ).

Ex.: Possibilidade de neve.

possibilidade (de neve).

40: FN1 → {DET} {NUM} N (PREP FV | FADV).

Ex.: Possibilidade de que ele dispute as eleições

possibilidade (de que dispute (ele, as eleições)).

41: FN2 → {DET} {NUM} N (PREP FN | FADJ, PREP FN | FADJ).

Ex.: A confiança nos pais pelas crianças.

a confiança (pelas crianças, nos pais).

42: FN2 → {DET} {NUM} N (PREP FN | FADJ, PREP FV | FADV).

Ex.: A crença em Deus de que tudo correrá bem.

a crença (em Deus, de que correrá (tudo, bem)).

43: FN3 → {DET} {NUM} N (por FN | FADJ, de FN | FADJ, PREP FN | FADJ).

Ex.: A doação de agasalhos aos flagelados pelas assistentes sociais.

a doação (pelas sociais (assistentes), de agasalhos, aos flagelados).

44: FN4 → {DET} {NUM} N (por FN | FADJ, de FN | FADJ, de FN | FADJ, para FN | FADJ).

Ex.: A tradução do poema do árabe para o português pelo tradutor.

a tradução (pelo tradutor, do poema, do árabe, para o português).

#### e) Regras para Adjetivos

45: FADJ1 → ADJ (FN | FADJ).

Ex: Uma bela mulher.

bela (uma mulher).

46: FADJ2 → ADJ (FN | FADJ, PREP FN | FADJ).

Ex.: Operações geradoras de divisas.

geradoras (operações, de divisas).

47: FADJ2 → ADJ (FN | FADJ, PREP FV | FADV).

Ex.: Quadro demonstrativo de que as despesas são inúteis.

demonstrativo (quadro, de que são (despesas, inúteis)).

48: FADJ3 → ADJ (FN | FADJ, de FN | FADJ, PREP FN | FADJ).

Ex.: Capataz contador de lorotas aos companheiros.

contador (capataz, de lorotas, aos companheiros).

49: FADJ3 → ADJ (por FN | FADJ, FN | FADJ, PREP FN | FADJ).

Ex.: Réu violento condenável à pena máxima pelo juiz.

condenável (pelo juiz, violento (réu), à máxima (pena)).

50: FADJ3 → ADJ (FN | FADJ, a FN | FADJ, PREP FV | FADV).

Ex.: Textos comprobatórios às comissões de que está tudo correto.  
 comprobatórios (textos, às comissões, de que está (tudo, correto)).

51: FADJ4 → ADJ (por FN | FADJ, FN | FADJ, de FN | FADJ, para FN | FADJ).

Ex.: Livros transferíveis da diretoria para biblioteca pelo funcionário.  
 transferíveis (pelo funcionário, livros, da diretoria, para biblioteca).

#### f) Regras para Advérbios

52: FADV0 → ADV (FV | FADJ | FADV).

Ex.: Primeiramente, inconseqüentemente, você roubou o carro.

Primeiramente (inconseqüentemente (roubou (você, o carro))).

Ex.: Uma cabeleira muito vasta.

muito (vasta (uma cabeleira)).

53: FADV1 → ADV (PREP FN | FADJ) (FV | FADV).

Ex.: embaixo (da velha (cama)) (gritou (ela)).

### **4.3.1 – Demonstração da geração de uma sentença a partir das regras de produção da MLV**

Visando demonstrar a aplicabilidade da gramática descrita logo acima, descreve-se abaixo a utilização das regras de produção da gramática para a geração de uma sentença gramaticalmente válida em MLV. A frase em linguagem natural a ser utilizada como exemplo será “O menino já comprou uma moto rápida”. Para se gerar qualquer sentença na gramática em questão, começa-se pelo símbolo inicial S e, a partir dele, escolhe-se as regras de produção para se chegar à sentença pretendida. Quando uma regra de produção for utilizada, ela será referenciada pelo seu respectivo número seqüencial, tal qual disposto na seção 4.3 logo acima.

S	Símbolo Inicial
FADV.	Regra 5
FADV0.	Regra 20
ADV (FV   FADJ   FADV).	Regra 52
já [ins_ja] (FV   FADJ   FADV).	Regra 25
já [ins_ja] (FV2).	Regra 8
já [ins_ja] ((VAUX) V (FN   FADJ, {PREP} FN   FADJ)).	Regra 32
já [ins_ja] (comprou [ins_comprar1] (FN   FADJ, {PREP} FN   FADJ)).	Regra 22
já [ins_ja] (comprou [ins_comprar1] (FN0, {PREP} FN   FADJ)).	Regra 11
já [ins_ja] (comprou [ins_comprar1] ({DET} {NUM} N, {PREP} FN   FADJ)).	Regra 35
já [ins_ja] (comprou [ins_comprar1] (o [ins_o1] N, {PREP} FN   FADJ)).	Regra 26
já [ins_ja] (comprou [ins_comprar1] (o [ins_o1] menino [ins_menino1], {PREP} FN   FADJ)).	Regra 23
já [ins_ja] (comprou [ins_comprar1] (o [ins_o1] menino [ins_menino1], FADJ1)).	Regra 16
já [ins_ja] (comprou [ins_comprar1] (o [ins_o1] menino [ins_menino1], ADJ (FN   FADJ))).	Regra 45
já [ins_ja] (comprou [ins_comprar1] (o [ins_o1] menino [ins_menino1], ADJ (FN0))).	Regra 11
já [ins_ja] (comprou [ins_comprar1] (o [ins_o1] menino [ins_menino1], rápida [ins_rapido1] (FN0))).	Regra 24
já [ins_ja] (comprou [ins_comprar1] (o [ins_o1] menino [ins_menino1], rápida [ins_rapido1] ({DET} {NUM} N))). Regra 35	
já [ins_ja] (comprou [ins_comprar1] (o [ins_o1] menino [ins_menino1], rápida [ins_rapido1] ({NUM} moto [ins_moto1])). Regra 23	
já [ins_ja] (comprou [ins_comprar1] (o [ins_o1] menino [ins_menino1], rápida [ins_rapido1] (uma [ins_um1] moto [ins_moto1])). Regra 28	

#### **4.4 - Desambiguação Aplicando o Referencial Proposto**

Nesta seção aplicaremos o referencial proposto a uma série de sentenças redigidas em linguagem natural, todas elas como algum tipo de ambigüidade. Nos aproveitaremos do trabalho de Silva (2006), utilizando muitos dos exemplos contidos naquela dissertação. Alguns exemplos foram trazidos dos trabalhos de Bräscher (1999) e a minoria restante foi elaborada propositamente para demonstrar o comportamento da MLV. Agruparemos as sentenças por tipo de ambigüidade, tal qual taxonomia proposta por Fuchs (1996) e, logo após cada sentença, descreveremos as possíveis (e mais prováveis) interpretações da frase ambígua. Cada interpretação possível será reescrita em meta-linguagem de valências. Por fim, para cada sentença ambígua julgaremos se a ambigüidade foi resolvida pelo referencial proposto.

Todas as palavras contidas nas sentenças redigidas em meta-linguagem de valências farão referência à ontologia de valências, através dos IDs dispostos entre colchetes. Entretanto, para não poluir visualmente a dissertação, prejudicando seu entendimento, só deixaremos explícitos nos exemplos abaixo os IDs que atuarem diretamente na resolução da ambigüidade. Por exemplo, nas ambigüidades lexicais, só estarão aparentes os IDs das palavras ambíguas. Os demais IDs serão escondidos propositamente. O símbolo de interrogação (?) será utilizado para indicar um argumento apagado.

##### **4.4.1 - Ambigüidade Lexical**

De acordo com a proposta contida nesta dissertação, a resolução deste tipo de ambigüidade é dada através da escolha formal, feita pelo autor do documento, do significado pretendido para um determinado significante representado por uma forma gráfica. Quando o autor conclui a digitação de uma forma gráfica, o ambiente de autoria busca na classe CLASSE\_FORMA\_DERIVADA a instância que condiz exatamente com a grafia da palavra digitada. A instância encontrada é remetida para uma ou várias instâncias de formas canônicas, CLASSE\_FORMA\_CANONICA, através da propriedade “oprop\_e\_forma\_derivada\_de”. Cada forma canônica representa um significado diferente para uma determinada forma gráfica. O ambiente de autoria solicita, então, que o autor escolha qual das formas canônicas encontradas é a desejada para a palavra digitada. Desta maneira, é

criada uma relação biunívoca entre significado e significante, resolvendo, assim a ambigüidade léxica.

a) A divisão já foi encerrada pela direção.

1a interpretação: Partilha, distribuição ou repartição de bens.

Sentença em MLV: já (encerrou (a direção, a divisão [ins\_divisao1])).

2a interpretação: Operação pela qual se conhece quantas vezes uma quantidade está contida noutra.

Sentença em MLV: já (encerrou (a direção, a divisão [ins\_divisao2])).

3a interpretação: Parte de um exército formada de brigadas.

Sentença em MLV: já (encerrou (a direção, a divisão [ins\_divisao3])).

Diagnóstico: Ambigüidade resolvida. Ambigüidade causada por polissemia. Como, no momento da autoria a acepção desejada é escolhida pelo autor, não existe mais ambigüidade.

b) Maria escolheu o canto.

1a interpretação: Série de sons musicais cadenciados, formados pela voz.

Sentença em MLV: escolheu (Maria, o canto [ins\_canto1]).

2a interpretação: Ângulo saliente ou reentrante formado pelo encontro de linhas ou superfícies.

Sentença em MLV: escolheu (Maria, o canto [ins\_canto2]).

Diagnóstico: Ambigüidade resolvida. Ambigüidade causada por homonímia. Como, no momento da autoria a acepção desejada é escolhida pelo autor, não existe mais ambigüidade.

c) Carlos indicou o banco.

1a interpretação: Instituição financeira.

Sentença em MLV: indicou (Carlos, o banco [ins\_banco1]).

2a interpretação: Estrutura utilizada como apoio.

Sentença em MLV: indicou (Carlos, o banco [ins\_banco2]).

3a interpretação: Local onde doa-se sangue.

Sentença em MLV: indicou (Carlos, o banco [ins\_banco3]).

4a interpretação: Estrutura utilizada para o armazenamento de dados.

Sentença em MLV: indicou (Carlos, o banco [ins\_banco4]).

Diagnóstico: Ambigüidade resolvida. Ambigüidade causada por homonímia. Como, no momento da autoria a acepção desejada é escolhida pelo autor, não existe mais ambigüidade.

d) Fernanda escolheu o salto.

1a interpretação: Taco que sustenta calçados femininos.

Sentença em MLV: escolheu (Fernanda, o salto [ins\_salto1]).

2a interpretação: Pulo que é executado em provas atléticas.

Sentença em MLV: escolheu (Fernanda, o salto [ins\_salto2]).

3a interpretação: Queda de grande porção de água corrente.

Sentença em MLV: escolheu (Fernanda, o salto [ins\_salto3]).

Diagnóstico: Ambigüidade resolvida. Ambigüidade causada por polissemia. Como, no momento da autoria a acepção desejada é escolhida pelo autor, não existe mais ambigüidade.

e) O assalto terminou.

1a interpretação: Combate simulado em esgrima.

Sentença em MLV: terminou (o assalto[ins\_assalto1]).

2a interpretação: Combate simulado em lutas de boxe.

Sentença em MLV: terminou (o assalto[ins\_assalto2]).

3a interpretação: Ato de subtrair violentamente.

Sentença em MLV: terminou (o assalto[ins\_assalto3]).

4a interpretação: Ato de atacar ao uma pessoa.

Sentença em MLV: terminou (o assalto[ins\_assalto4]).

Diagnóstico: Ambigüidade resolvida. Ambigüidade causada por polissemia. Como, no momento da autoria a acepção desejada é escolhida pelo autor, não existe mais ambigüidade.

f) Aquele assento não existe mais.

1a interpretação: Sítio onde esteve ou está fundada alguma povoação.

Sentença em MLV: não (mais (existe( aquele assento [ins\_assento1])))).

2a interpretação: Estrutura utilizada como apoio.

Sentença em MLV: não (mais (existe( aquele assento [ins\_assento2])))).

3a interpretação: Registro de acórdão do Supremo Tribunal com força de lei.

Sentença em MLV: não (mais (existe( aquele assento [ins\_assento3])))).

Diagnóstico: Ambigüidade resolvida. Ambigüidade causada por polissemia. Como, no momento da autoria a acepção desejada é escolhida pelo autor, não existe mais ambigüidade.

g) O reitor comentou as cadeiras da universidade.

1a interpretação: Assento com encosto para costas.

Sentença em MLV: comentou (o reitor, as cadeiras [ins\_cadeira1] da universidade).

2a interpretação: Ramo dos conhecimentos humanos considerados como objeto do ensino.

Sentença em MLV: comentou (o reitor, as cadeiras [ins\_cadeira2] da universidade).

3a interpretação: Árvore da Guiné que libera uma espécie de borracha.

Sentença em MLV: comentou (o reitor, as cadeiras [ins\_cadeira3] da universidade).

4a interpretação: parte do corpo humano também conhecido como quadris ou nádegas.

Sentença em MLV: comentou (o reitor, as cadeiras [ins\_cadeira4] da universidade).

Diagnóstico: Ambigüidade resolvida. Ambigüidade causada por polissemia. Como, no momento da autoria a acepção desejada é escolhida pelo autor, não existe mais ambigüidade.

h) Mariana conheceu outra língua durante a aula.

1a interpretação: Idioma utilizando ou originado de um país ou civilização.

Sentença em MLV: durante a aula (conheceu (Mariana, outra língua [ins\_lingua1])).

2a interpretação: Corpo carnudo alongado, móvel, situado dentro da boca.

Sentença em MLV: durante a aula (conheceu (Mariana, outra língua [ins\_lingua2])).

Diagnóstico: Ambigüidade resolvida. Ambigüidade causada por polissemia. Como, no momento da autoria a acepção desejada é escolhida pelo autor, não existe mais ambigüidade.

i) A vara era muito grande.

1a interpretação: Vara é objeto comprido e fino de madeira.

Sentença em MLV: muito ((era (a vara [ins\_vara1], grande)).

2a interpretação: Vara é substantivo coletivo de porcos.

Sentença em MLV: muito ((era (a vara [ins\_vara2], grande)).

Diagnóstico: Ambigüidade resolvida. Ambigüidade causada por homonímia. Como, no momento da autoria a acepção desejada é escolhida pelo autor, não existe mais ambigüidade.

j) Ela preparou um molho.

1a interpretação: Molho é substantivo coletivo de chaves ou objetos pequenos.

Sentença em MLV: preparou (ela, um molho [ins\_molho1]).

2a interpretação: Molho é um tipo de alimento ensopado.

Sentença em MLV: preparou (ela, um molho [ins\_molho2]).

Diagnóstico: Ambigüidade resolvida. Ambigüidade causada por homonímia. Como, no momento da autoria a acepção desejada é escolhida pelo autor, não existe mais ambigüidade.

k) Os pássaros cantam.

1a interpretação: Os pássaros estão cantando no momento.

Sentença em MLV: cantam (os pássaros [ins\_passaro1]).

2a interpretação: Os pássaros têm habilidade de cantar.

Sentença em MLV: cantam (os pássaros [ins\_passaro2]).

Diagnóstico: Ambigüidade resolvida. Apesar de Bräscher (1999) caracterizar esta ambigüidade como pragmática, esta tem características de ambigüidade lexical, pois pode ser revolvida através da manipulação do léxico. Adicionando uma entrada na ontologia para pássaro, referência específica (traço semântico contável), e outra para pássaro, referência geral (traço semântico não contável), é possível resolver esta ambigüidade. Tal procedimento já vem sendo adotado por alguns dicionaristas, tais como Borba (2002).

l) Maria esqueceu João.

1a interpretação: Maria esqueceu João em algum lugar.

Sentença em MLV: esqueceu [ins\_esquecer1] (Maria, João, ?).

2a interpretação: Maria esqueceu da pessoa de João.

Sentença em MLV: esqueceu [ins\_esquecer2] (Maria, João).

Diagnóstico: Ambigüidade resolvida. Da mesma maneira que a sentença anterior, apesar de Silva (2006) caracterizar esta como uma ambigüidade semântica, de fato parece esta ser um exemplo de ambigüidade lexical. O verbo esquecer tem acepções e matrizes valenciais distintas. Na primeira acepção, “esquecer por inadvertência” (Borba, 2002), temos um predicado em três argumentos, sendo que o último compreende o lugar onde o objeto foi esquecido. A segunda acepção, “deixar sair da memória; não lembrar” (Borba, 2002), só tem dois argumentos.

#### **4.4.2 - Ambigüidade Morfológica**

A resolução da ambigüidade morfológica é dada de maneira idêntica à resolução da ambigüidade lexical. Portanto, da mesma maneira, quando o autor conclui a digitação de uma forma gráfica, o ambiente de autoria busca na classe CLASSE\_FORMA\_DERIVADA a instância que condiz exatamente com a grafia da palavra digitada. A instância encontrada é remetida para uma ou várias instâncias de formas canônicas,

CLASSE\_FORMA\_CANONICA, através da propriedade “oprop\_e\_forma\_derivada\_de”. Cada forma canônica pode representar uma classe morfológica diferente para uma mesma forma gráfica. O ambiente de autoria solicita, então, que o autor escolha qual das formas canônicas encontradas é a desejada para a palavra digitada..

a) O jovem trabalhador começou suas funções.

1a interpretação: O trabalhador é jovem.

Sentença em MLV: começou (jovem [ins\_jovem2] (o trabalhador [ins\_trabalhador1]), suas funções).

2a interpretação: O jovem é trabalhador.

Sentença em MLV: começou (trabalhador [ins\_trabalhador2] (o jovem [ins\_jovem1]), suas funções).

Diagnóstico: Ambigüidade resolvida. Jovem e trabalhador terão entradas na ontologia, tanto para a função de substantivo, quanto para a função de adjetivo.

b) Ela escolhia calma e vagorosamente os presentes

1a interpretação: Ela estava calma (adjetivo) enquanto escolhia os presentes.

Sentença em MLV: vagorosamente (escolhia (calma (ela), os presentes)).

2a interpretação: Ela escolheu os presentes de forma calma (advérbio).

Sentença em MLV: calmamente (vagosamente (escolhia (ela, os presentes))).

Diagnóstico: Ambigüidade resolvida. Segundo Cunha & Cintra (2001), quando em uma mesma frase dois advérbios modificam o mesmo verbo, então, para tornar o enunciado mais leve, o sufixo “mente” só é adicionado ao segundo advérbio. Entretanto, “calma” também pode ter função de adjetivo na sentença acima, modificando “ela”.

c) Ela casa amanhã.

1a interpretação: Casa, conjugação do verbo casar.

Sentença em MLV: amanhã (casa [ins\_casar1] (Ela, ?)).

2a interpretação: Casa, substantivo.

Sentença em MLV: ela casa [ins\_casa2] amanhã.

Diagnóstico: Ambigüidade resolvida. A acepção pretendida na primeira interpretação, logo acima, corresponde ao verbo casar com dois argumentos (alguém casa com alguém). Na segunda interpretação tem-se “casa” como substantivo.

d) Castigo.

1a interpretação: Verbo castigar.

Sentença em MLV: castigo [ins\_castigar1] (?, ?).

2a interpretação: Nome que indica pena ou punição.

Sentença em MLV: castigo [ins\_castigo].

Diagnóstico: Ambigüidade resolvida. Uma sentença com uma única palavra, onde o co-texto não pode auxiliar na desambiguação. Na primeira interpretação temos o verbo castigar, com dois argumentos apagados. Na segunda temos “castigo” como nome.

#### 4.4.3 - Ambigüidade Sintática

Seguindo a gramática de valências de Borba, cada instância da CLASSE\_FORMA\_CANONICA tem, de maneira explícita, a função sintática de cada um dos argumentos da palavra lexical. Dessa maneira, como o ambiente de autoria auxilia o autor no preenchimento dos argumentos, cada um dos actantes terá atrelado a si, através da ontologia, uma função sintática específica, sanando a ambigüidade sintática. Outro foco importante deste tipo de ambigüidade é o emprego palavras ou locuções adverbiais que podem ser confundidas com complementos nominais, e vice-versa. O exemplo “a”, logo abaixo, demonstra este tipo de fenômeno. Para resolução da ambigüidade, o ambiente de autoria guia o autor de maneira que a locução “no cavalo” seja formalmente definida como um adjunto adverbial, ou como um complemento de “campeiro”.

a) O fazendeiro viu o campeão no cavalo.

1a interpretação: O fazendeiro estava montado no cavalo.

Sentença em MLV: no cavalo (viu (o fazendeiro, o campeão)).

2a interpretação: O campeão estava montado no cavalo.

Sentença em MLV: viu (o fazendeiro, o campeão no cavalo).

Diagnóstico: Ambigüidade resolvida. No primeiro caso, “no cavalo” trata-se de um adjunto adverbial. No segundo, trata-se de um adjunto adnominal.

b) O guarda viu o rapaz com os binóculos.

1a interpretação: O guarda usava os binóculos.

Sentença em MLV: com os binóculos (viu (o guarda, o rapaz)).

2a interpretação: O rapaz usava os binóculos.

Sentença em MLV: viu (o guarda, o rapaz com os binóculos).

Diagnóstico: Ambigüidade resolvida. Mesmo caso da anterior. No primeiro caso, “com os binóculos” trata-se de um adjunto adverbial. No segundo, trata-se de um adjunto adnominal.

c) Maria disse que choveu ontem.

1a interpretação: Maria disse que ontem choveu.

Sentença em MLV: disse (Maria, que ontem (choveu ())).

2a interpretação: Maria disse ontem que choveu.

Sentença em MLV: ontem (disse (Maria, que choveu ())).

3a interpretação: Maria disse ontem que choveu ontem.

Sentença em MLV: ontem (disse (Maria, que ontem (choveu ())))).

Diagnóstico: Ambigüidade resolvida.

d) Comprei uma revista sobre um acidente no trem.

1a interpretação: A revista é sobre o acidente que ocorreu no trem.

Sentença em MLV: no trem (comprei (eu, uma revista sobre um acidente)).

2a interpretação: A revista foi comprada no trem e fala sobre um acidente.

Sentença em MLV: comprei (eu, uma revista sobre um acidente no trem).

Diagnóstico: Ambigüidade resolvida.

e) O carro da menina que vimos ontem.

1a interpretação: O carro foi visto ontem.

Sentença em MLV: ontem (vimos (nós, o carro da menina)).

2a interpretação: A menina foi vista ontem.

Sentença em MLV: O carro da ontem (vimos (nós, a menina)).

Diagnóstico: Ambigüidade resolvida. Esta é uma questão relativa ao escopo do advérbio e do verbo. No primeiro caso o verbo modificado age sobre “o carro da menina”. No segundo, o verbo modificado pelo advérbio age só sobre a menina.

f) A lata perto da panela com a colher.

1a interpretação: A colher está na lata.

Sentença em MLV: perto (da panela) (está (a lata com a colher)).

2a interpretação: A colher está na panela.

Sentença em MLV: perto (da panela com a colher) (está (a lata)).

Diagnóstico: Ambigüidade resolvida.

g) Comprei uma lancha e dois barcos velozes.

1a interpretação: A lancha e os barcos são velozes.

Sentença em MLV: comprei (eu, velozes (uma lancha e dois barcos)).

2a interpretação: Os barcos são velozes.

Sentença em MLV: comprei (eu, uma lancha e velozes (dois barcos)).

Diagnóstico: Ambigüidade resolvida. Na primeira interpretação, temos o adjetivo englobando tanto “uma lancha” quanto “dois barcos”. Na segunda, temos o adjetivo modificando apenas “dois barcos”.

h) Cachorros e gatos felizes vivem na fazenda.

1a interpretação: Os cachorros e os gatos são felizes.

Sentença em MLV: vivem (felizes (cachorros e gatos), na fazenda).

2a interpretação: Os gatos são felizes.

Sentença em MLV: vivem (cachorros e felizes (gatos), na fazenda).

Diagnóstico: Ambigüidade resolvida.

#### **4.4.4 - Ambigüidade Predicativa**

A resolução deste tipo de ambigüidade é dada pela atuação das propriedades `oprop_tem_caso1`, `oprop_tem_caso2`, `oprop_tem_caso3` e `oprop_tem_caso4`. Estas propriedades ligam cada um dos actantes das palavras lexicais à classe `CLASSE_CASO`, definindo assim, formalmente, qual é o papel temático exercido pelo argumento em um determinado predicado, tal qual propõe a teoria de Tesnière (1959).

a) A crítica deste autor.

1a interpretação: Autor alvo da crítica.

Sentença em MLV: crítica (? , ao autor).

2a interpretação: Autor criador da crítica.

Sentença em MLV: crítica (do autor , a ?).

Diagnóstico: Ambigüidade resolvida. O primeiro argumento do nome abstrato “crítica”, derivado do verbo criticar, é o sujeito com papel temático agente. O segundo argumento é o complemento com papel temático de objeto. A disposição “do autor” no primeiro ou segundo argumento resolve a ambigüidade.

b) Arranquei a folha do caderno inútil.

1a interpretação: O caderno é inútil.

Sentença em MLV: arranquei (eu, a folha do inútil (caderno)).

2a interpretação: A folha é inútil.

Sentença em MLV: arranquei (eu, inútil (a folha do caderno)).

Diagnóstico: Ambigüidade resolvida.

c) O leão avançou sobre o caçador violento.

1a interpretação: O Leão é violento.

Sentença em MLV: avançou (violento (o leão), sobre o caçador).

2a interpretação: O Caçador é violento.

Sentença em MLV: avançou (o leão, sobre violento (o caçador)).

Diagnóstico: Ambigüidade resolvida.

d) O caçador defendeu-se do animal amedrontado.

1a interpretação: O caçador estava amedrontado.

Sentença em MLV: defendeu-se (amedrontado (o caçador), do animal).

2a interpretação: O animal estava amedrontado.

Sentença em MLV: defendeu-se (o caçador, do amedrontado (animal)).

Diagnóstico: Ambigüidade resolvida.

e) A platéia aplaudiu a bailarina entusiasmada.

1a interpretação: A platéia estava entusiasmada.

Sentença em MLV: aplaudiu (entusiasmada (a platéia), a bailarina).

2a interpretação: A bailarina estava entusiasmada.

Sentença em MLV: aplaudiu (entusiasmada (a platéia), a bailarina).

Diagnóstico: Ambigüidade resolvida.

#### 4.4.5 – Ambigüidade Semântica

Os exemplos de ambigüidades semânticas não foram resolvidos. Nas orações com este tipo de ambigüidade é possível definir exatamente qual é o significado, classe morfológica, classe sintática e papel temático de todos os componentes. Ainda assim, a ambigüidade semântica surge, pois a combinação de tais componentes faz com que não se possa determinar o significado da oração como um todo. De fato, este tipo de ambigüidade se aproxima bastante da questão pragmática, pois apenas o contexto ou, melhor dizendo, o co-texto oracional, pode resolvê-la.

a) Carla não chora mais porque Antônio partiu.

1a interpretação: Carla chorava porque Antônio havia partido.

Sentença em MLV: não (mais (chora (Carla, porque partiu (Antônio)))).

2a interpretação: Carla parou de chorar uma vez que Antônio já foi embora.

Sentença em MLV: não (mais (chora (Carla, porque partiu (Antônio)))).

Diagnóstico: Ambigüidade não resolvida. Não é possível formalizar a relação causa consequência expressa na sentença acima.

b) As garrafas valem 30 reais.

1a interpretação: Cada garrafa vale 30 reais.

Sentença em MLV: valem (as garrafas, 30 reais).

2a interpretação: Todas as garrafas valem 30 reais.

Sentença em MLV: valem (as garrafas, 30 reais).

Diagnóstico: Ambigüidade não resolvida.

c) Um rio corre através de cada país latino.

1a interpretação: Um único rio corre através de todos os países.

Sentença em MLV: corre (um rio, através de cada latino (país)).

2a interpretação: Diferentes rios correm através de diferentes países.

Sentença em MLV: corre (um rio, através de cada latino (país)).

Diagnóstico: Ambigüidade não resolvida.

d) Carlos não comprou mais daquele produto.

1a interpretação: Carlos ainda tem o produto em estoque.

Sentença em MLV: não (mais (comprou, Carlos, daquele produto)).

2a interpretação: Carlos acredita que o produto é ruim.

Sentença em MLV: não (mais (comprou, Carlos, daquele produto)).

Diagnóstico: Ambigüidade não resolvida.

e) Onde moro não há mais energia elétrica.

1a interpretação: A quantidade para consumo é pequena.

Sentença em MLV: não (mais (há (onde moro, energia))).

2a interpretação: O fornecimento foi interrompido.

Sentença em MLV: não (mais (há (onde moro, energia))).

Diagnóstico: Ambigüidade não resolvida.

f) Ela usa blusas de linha e seda.

1a interpretação: As blusas são confeccionadas com matéria mista.

Sentença em MLV: usa (ela, blusas de linha e seda).

2a interpretação: As blusas são confeccionadas com matérias diferentes.

Sentença em MLV: usa (ela, blusas de linha e seda).

Diagnóstico: Ambigüidade não resolvida.

g) Todas as mulheres têm um sonho.

1a interpretação: Elas compartilham o mesmo sonho.

Sentença em MLV: sonham (todas as mulheres).

2a interpretação: Cada uma tem seu respectivo sonho.

Sentença em MLV: sonham (todas as mulheres).

Diagnóstico: Ambigüidade não resolvida. Segundo Borba (1990), neste caso o verbo ter tem função de verbalizador (ter medo = temer, ter um sonho = sonhar). Por isso o predicado corresponde ao verbo sonhar.

h) Algumas pessoas foram enganadas pelos políticos.

1a interpretação: As pessoas foram enganadas todas ao mesmo tempo.

Sentença em MLV: enganaram (os políticos, algumas pessoas).

2a interpretação: Cada pessoa foi enganada por vez.

Sentença em MLV: enganaram (os políticos, algumas pessoas).

Diagnóstico: Ambigüidade não resolvida. Esta ambigüidade tem um fundo pragmático. Apenas o contexto pode fornecer informações para sanar esta ambigüidade.

i) Todas as pessoas são amadas por alguém.

1a interpretação: Todas as pessoas são amadas por uma única pessoa.

Sentença em MLV: ama (uma única pessoa, todas as pessoas).

2a interpretação: Cada pessoa é amada pelo seu par respectivo.

Sentença em MLV: ama (uma única pessoa, todas as pessoas).

Diagnóstico: Ambigüidade não resolvida.

j) Os animais domésticos possuem casa.

1a interpretação: Todos os animais moram na mesma casa.

Sentença em MLV: possuir (os domésticos (animais), casa).

2a interpretação: Cada animal mora em sua casa respectiva.

Sentença em MLV: possuir (os domésticos (animais), casa).

Diagnóstico: Ambigüidade não resolvida.

k) Ambos os estudantes visitaram Carlos.

1a interpretação: Juntos visitaram Carlos.

Sentença em MLV: visitaram (ambos os estudantes, Carlos).

2a interpretação: Separados visitaram Carlos.

Sentença em MLV: visitaram (ambos os estudantes, Carlos).

Diagnóstico: Ambigüidade não resolvida.

l) O médico visitou o filho da enfermeira que se machucou.

1a interpretação: A enfermeira que se machucou.

Sentença em MLV: visitou (o médico, o filho da enfermeira que se machucou ()).

2a interpretação: O filho que se machucou.

Sentença em MLV: visitou (o médico, o filho da enfermeira que se machucou ()).

Diagnóstico: Ambigüidade não resolvida. Não é possível distinguir se o trecho “que se machucou” refere-se ao “filho” ou à “enfermeira”.

m) Maria já emprestou todos os livros a José.

1a interpretação: Emprestou todos os livros ao mesmo tempo.

Sentença em MLV: já (emprestou (Maria, todos os livros, a José)).

2a interpretação: Emprestou um livro após o outro.

Sentença em MLV: já (emprestou (Maria, todos os livros, a José)).

Diagnóstico: Ambigüidade não resolvida.

n) Aqueles homens compraram todos os computadores.

1a interpretação: Todos os homens compraram os computadores todos.

Sentença em MLV: compraram (aqueles homens, todos os computadores).

2a interpretação: Cada homem comprou um computador.

Sentença em MLV: compraram (aqueles homens, todos os computadores).

Diagnóstico: Ambigüidade não resolvida.

o) Os alunos riscaram todas as folhas.

1a interpretação: Todos os alunos riscaram todas as folhas.

Sentença em MLV: riscaram (os alunos, todas as folhas).

2a interpretação: Cada aluno riscou sua folha.

Sentença em MLV: riscaram (os alunos, todas as folhas).

Diagnóstico: Ambigüidade não resolvida.

#### 4.4.6 - Ambigüidade Pragmática

Como nas semânticas, não foi possível solucionar os exemplos de ambigüidades pragmáticas. O arcabouço teórico mostra-se eficaz nas questões morfológicas, sintáticas, lexicais e predicativas, mas não consegue captar o *pragma*, ou seja, não é capaz de registrar formalmente o contexto interno e externo ao documento. É nestes contextos (interno e externo ao documento) que reside a solução das ambigüidades pragmáticas.

a) Eles cometeram um crime.

1a interpretação: Cometeram o mesmo crime.

Sentença em MLV: cometeram (eles, um crime).

2a interpretação: Cometeram crimes separados.

Sentença em MLV: cometeram (eles, um crime).

Diagnóstico: Ambigüidade não resolvida. Só o contexto pode resolvê-la.

b) Carlos reviu alguns de seus amigos.

1a interpretação: Os amigos estavam juntos.

Sentença em MLV: reviu (Carlos, alguns de seus amigos).

2a interpretação: Os amigos estavam separados.

Sentença em MLV: reviu (Carlos, alguns de seus amigos).

Diagnóstico: Ambigüidade não resolvida.

c) Carlos vai à faculdade.

1a interpretação: Carlos é estudante.

Sentença em MLV: vai (Carlos, à faculdade).

2a interpretação: Carlos está indo à faculdade neste momento para buscar a namorada.

Sentença em MLV: vai (Carlos, à faculdade).

Diagnóstico: Ambigüidade não resolvida. Só o contexto pode resolvê-la.

d) Carlos veio ver Antônio, o inocente estava contente.

1a interpretação: Carlos é o inocente.

Sentença em MLV: veio ver (Carlos, Antônio). estava (o inocente (?), contente).

2a interpretação: Antônio é o inocente.

Diagnóstico: Ambigüidade não resolvida. Frase composta por duas orações. A segunda oração remete à primeira. Não é possível saber que é o inocente na segunda oração.

e) Os quadros expostos foram comprados.

1a interpretação: Os quadros foram comprados pela mesma pessoa.

Sentença em MLV: comprou ( ? , os expostos (quadros)).

2a interpretação: Os quadros foram comprados por pessoas diferentes.

Sentença em MLV: comprou ( ? , os expostos (quadros)).

Diagnóstico: Ambigüidade não resolvida. Esta ambigüidade é causada pelo apagamento do sujeito. Só o contexto pode resolvê-la.

## ***5 - Conclusão***

Definitivamente não é fácil abordar os problemas da linguagem natural e da comunicação humana. Esta dificuldade pode ser constatada na literatura em textos como o de Martins (2002), que chega a afirmar que “Neste panorama de valorização da indeterminação, teóricos chegam a afirmar que a (possibilidade de) compreensão mútua é quase um ‘milagre’, o que torna ‘bastante irracionais’ as teorias que defendem a comunicação em termos de codificação e decodificação de idéias.”. Frente a uma afirmação contraproducente como esta, tem-se duas opções: deixar os problemas relativos à comunicação e conhecimento de lado ou; buscar alternativas para atenuar tais problemas. Fica-se com a segunda opção.

Foi constatado que a ambigüidade é geralmente tratada na recuperação da informação ou em um momento imediatamente antes dela, ou seja, sempre após a autoria do documento. Esta abordagem é plenamente justificável, tendo em vista o vasto legado de documentos disponíveis, principalmente em função da explosão informacional ocorrida após o advento da Internet. Entretanto, a eliminação de ambigüidades através desta abordagem é extremamente complicada, exigindo soluções complexas que não conseguem abarcar todo o problema. Além disso, toda desambiguação feita através desta abordagem convencional, por mais sofisticada que seja, não passa de mera inferência do receptor sobre o significado da mensagem, inferência esta que pode ser equivocada. Quem realmente conhece o conteúdo a ser comunicado é o emissor, o autor do documento, e este deve ter uma posição mais compromissada no ato da comunicação.

Alçar o autor do documento a uma posição mais compromissada significa tornar o processo de autoria mais trabalhoso. É possível imaginar o quão enfadonho seria produzir um documento com um software inquirindo o autor a todo o momento sobre informações morfológicas, sintáticas e semânticas sobre o conteúdo do texto. Este é o preço a ser pago pela adoção de um ambiente de autoria que siga os preceitos propostos nesta dissertação. Este também é o motivo que limita a utilização do arcabouço teórico proposto neste trabalho a apenas determinados tipos de documentos, tais como leis e contratos. Não é justificável, por exemplo, a utilização do framework proposto aqui para a criação de documentos de cunho literário poético.

O objetivo geral proposto foi atingido. O arcabouço teórico proposto neste trabalho mostrou-se suficientemente sólido para atenuar o surgimento de ambigüidades na recuperação e interpretação da informação. A noção de que o autor do documento deve assumir uma posição mais comprometida na comunicação é o ponto mais forte do arcabouço proposto. De fato, esta noção mostrou-se eficaz para atenuar alguns fenômenos lingüísticos indesejáveis na comunicação, como as ambigüidades e a vagueza. Como visto, esta abordagem é uma alternativa aos complexos procedimentos de processamento automático de linguagem natural.

O primeiro objetivo específico, “definir e descrever os componentes necessários à criação do arcabouço teórico”, foi atingido. Tal feito foi conseguido a partir de um exercício de criatividade e de embasamento teórico que nos guiou à definição dos níveis necessários ao arcabouço.

O segundo objetivo específico, “definir a estrutura conceitual necessária a uma ontologia das valências dos verbos, nomes, adjetivos e advérbios da variação brasileira da língua portuguesa”, também foi atingido. Entretanto a ontologia proposta carece de detalhamento e expansão. Na ontologia não foram contemplados, por exemplo, diversos aspectos morfológicos, sintáticos e semânticos que poderiam contribuir para otimizar o ambiente de autoria para o qual a ontologia se propõe. Informações morfológicas tais como regras para formação de relativas a número e gênero poderiam ser adicionadas. No campo da sintaxe, regras relativas à concordância também poderiam auxiliar os autores no momento da autoria dos documentos. No contexto semântico, poderiam ser adicionadas relações do tipo parte-todo (Ex: carburador é parte de motor que é parte de carro) e até mesmo taxonômicas (Ex: gato é um mamífero) entre as formas canônicas.

O terceiro objetivo específico também foi atingido. A gramática da meta-linguagem de valência foi definida de maneira formal. É interessante observar como a meta-linguagem de valências é capaz de registrar o enunciado primitivo de uma oração. É neste enunciado primitivo que reside a arquitetura informacional da sentença. Um único enunciado primitivo pode gerar “n” formas diferentes em linguagem natural. Isso faz com que a linguagem natural passe a ter um caráter estético, deixando a representação do conhecimento para uma linguagem mais adequada para tal.

Dos quarenta e nove exemplos de sentenças ambíguas, em vinte e nove a ambigüidade foi resolvida (59,18 %). Todas as doze sentenças com ambigüidade lexical foram desambiguadas. As quatro sentenças apresentadas com ambigüidades morfológicas foram desambiguadas. As oito sentenças disponíveis com ambigüidade sintática tiveram sua ambigüidade resolvida. Todas as cinco sentenças com ambigüidade predicativa foram desambiguadas. Nenhuma das quinze sentenças com ambigüidade semântica, bem como nenhuma das cinco sentenças com ambigüidade pragmática, teve sua ambigüidade resolvida.

Primeiramente, é necessário cautela na análise dos resultados descritos acima. O corpus submetido à MLV é muito pequeno, sendo necessário submeter a referida linguagem a um corpus muito maior para que se tenha maior segurança sobre os resultados. Entretanto, os testes realizados permitem algumas interpretações:

1 – Em teoria, bem como nos testes realizados, o arcabouço teórico mostrou-se sólido o bastante para afirmarmos que o mesmo é capaz de sanar a totalidade das ambigüidades lexicais e morfológicas, dado que o ambiente de autoria exige que o autor formalize suas escolhas lexicais e morfológicas no momento da redação do documento;

2 – Sobre as ambigüidades predicativas, não é possível ter certeza de que o arcabouço proposto pode resolvê-las em cem por cento dos casos. Este sentimento está baseado no fato de que alguns autores não se contentam com os papéis temáticos oriundos da teoria de Tesnière (1959), que são utilizados no presente trabalho. Tais autores adicionam novos papéis temáticos em suas soluções de processamento de linguagem natural para atingir os objetivos pretendidos.

3 – Quanto às ambigüidades sintáticas, também não se pode ter certeza de que a gramática proposta pode resolver todos os casos de tal tipo de ambigüidade. Tal gramática necessita ser melhorada, estando ainda longe de uma operacionalização razoável. É preciso melhorá-la para permitir a redação de sentenças complexas. É preciso formalizar várias lacunas não cobertas pela gramática, tais como: a questão das frases compostas por mais de uma oração (Ex: Paulo Pediu a Pedro que pedisse a Maria que falasse ao público); a questão dos conectivos (Ex: um carro, uma vaca, uma bola e uma galinha); e a questão referente às locuções adverbiais e adjetivas.

4 – O arcabouço mostrou-se ineficaz com relação às ambigüidades pragmáticas e semânticas. O grande dificultador para tratar tais tipos de ambigüidades é falta de uma

estrutura de representação de conhecimento no arcabouço aqui proposto capaz de formalizar o contexto de dentro do documento, bem como o contexto externo ao documento. Muitas das ambigüidades semânticas apresentadas aqui são causadas por anáforas que remetem a outras sentenças do documento, fora do co-texto oracional. Já as pragmáticas podem remeter, inclusive, ao contexto fora do documento, o que torna a tarefa de resolvê-las ainda mais desafiadora.

Com relação à obra de Oliveira (2006), entende-se que este trabalho contribuiu para detalhar os procedimentos que devem ser tomados por um ambiente de autoria para atenuar o surgimento de ambigüidades. Este detalhamento era necessário, dado que o autor, apesar de salientar a problemática das ambigüidades, não trata o tema em detalhes, e que um dos requisitos que seu ambiente de autoria deve atender diz respeito ao auxílio que o software deve dar ao autor para “escrever textos em qualquer ambiente institucional, técnico ou científico, onde a ocorrência de ambigüidade deve ser reduzida.”.

Como obra inspiradora desta dissertação, e natural referencial para comparação de resultados, cabem algumas considerações sobre Bräscher (1999; 2002). Primeiramente, é necessário salientar a qualidade metodológica e didática da obra da autora, bem como os resultados consideráveis obtidos com sua proposta. Para tal constatação basta verificar o percentual de acertos obtidos na desambiguação de formas policategoriais, chegando a mais de 85 % na distinção entre adjetivos e substantivos. A autora utilizou em seu trabalho conceitos vanguardistas, como as ontologias, que viriam a ser alardeadas na comunidade científica alguns anos depois. Almeja-se a honra de que esta dissertação seja considerada como uma extensão e um complemento considerável à bela obra de Bräscher.

Entende-se que este trabalho atendeu às recomendações de trabalhos futuros efetuadas por Bräscher (1999) no final de seu trabalho<sup>27</sup>. O arcabouço aqui proposto pode subsidiar o desenvolvimento “de ferramentas automáticas de tratamento de conteúdo que possibilitem maior precisão na recuperação da informação”, especialmente no âmbito dos ambientes de autoria na Web Semântica baseados em ontologias. Compreende-se que atuar no momento da autoria do documento demonstra uma compreensão holística do ciclo de vida da informação, demonstrando-se assim a “preocupação com a qualidade do tratamento da informação, pois a

---

<sup>27</sup> Vide seção 2.2 para uma descrição das recomendações de Bräscher (1999).

qualidade da recuperação depende deste tratamento”. Foram levados em consideração “os resultados obtidos em pesquisas da área da Ciência da Informação” para a concepção desta dissertação, uma vez que a própria obra de Bräscher (1999), utilizada à exaustão no presente trabalho, é um resultado importante de pesquisas na área da Ciência da Informação. A presente pesquisa também se utilizou de resultados obtidos em outras áreas do conhecimento, tal qual a Lingüística e a Ciência da Computação, para atender “às necessidades práticas da Ciência da Informação”. Nesta dissertação também foram aplicados “recursos tecnológicos na melhoria das ferramentas de tratamento e recuperação de informação”. A utilização de conceitos como ontologias e linguagens como XML, RDF e OWL, demonstram a aplicação de ferramentas tecnológicas de vanguarda.

Como uma solução de desambiguação que atua na fase de pós-autoria do documento, Bräscher (1999) necessita de pistas sintático-semânticas dadas pelo co-texto para solucionar as ambigüidades. Alias, é importante esclarecer que a solução de Bräscher (1999) não faz uso do contexto externo à oração onde se encontra a ambigüidade, nem ao contexto externo ao documento. Quando o co-texto não é capaz de subsidiar o sistema de processamento automático de linguagem natural utilizado pela autora (Zstation) não é possível efetuar a desambiguação. Esta constatação pode ser efetuada através do exemplo dado pela própria autora, quando a ambigüidade presente em “ações” é resolvida pelo co-texto, através da valência semântica do verbo vender, que exige como complemento algo comercializável: “regras de preferência para os casos de venda de ações e aumento do capital social”. Em uma hipotética sentença tal como: “Ações e aumento do capital social”, a desambiguação já não seria possível, pois o verbo “vender” não está presente para auxiliar na resolução da palavra ambígua.

Diferentemente de Bräscher (1999), que atua na pós-autoria, a proposta da presente nesta dissertação se beneficia do fato de atuar no momento da autoria do documento, fazendo com que o próprio autor forneça informações morfológicas, sintáticas e semânticas sobre o conteúdo do documento. A mesma ontologia que é utilizada na autoria será utilizada para a recuperação da informação, fazendo com que, ao menos no campo léxico-morfológico, tenha-se 100 % de certeza de que não haverá ambigüidade. Não há dependência do co-texto para resolução da ambigüidade morfológica ou lexical, como existe em Bräscher (1999). Já com relação às ambigüidades predicativas, também se atinge um resultado considerável, uma vez

que o autor, auxiliado pelo ambiente de autoria, posiciona os argumentos explicitamente em seus respectivos papéis temáticos.

Outro ponto importante a ser tocado é referente à automatização de um possível ambiente de autoria baseado no referencial aqui proposto. Como dito anteriormente, a idéia por traz deste referencial exige uma maior participação do autor no processo de comunicação, especialmente na recuperação da informação. Esta filosofia exige que autor interaja fortemente com o ambiente. Por um lado, esta predominância de interferência humana pode ser considerada um retrocesso tecnológico. Por outro, apesar de a ontologia proposta aqui ter informações suficientes para inferências automáticas, seguiu-se a filosofia que norteou este trabalho: Apenas o autor do documento tem a exata consciência do quer transmitir. Os problemas relacionados à comunicação humana estão longe de serem resolvidos por ferramentas tecnológicas. De fato, acredita-se que nunca serão, especialmente quando se trata da questão pragmática.

Por fim, a constatação da eficácia léxica-morfológica do arcabouço aqui proposto aponta para trabalhos futuros e imediatos que visem a construção de aplicações práticas, baseadas na filosofia de desambiguação na autoria de documentos, para solucionar ambigüidades morfológicas e, especialmente, as léxicas. De fato, chama a atenção o movimento científico mundo afora que se preocupa unicamente com a solução de ambigüidades lexicais (*Word Sense Disambiguation*, termo difundido mundialmente em língua inglesa). Entretanto, toda literatura científica para solucionar o referido problema é voltada para soluções que não levam em conta o momento da autoria do documento. Por isso, a contribuição desta dissertação seria considerável e importante.

## *Referências*

ALVES, I. M. R.; ALUÍSIO, S. M. Web como um corpus versus corpus tradicionais: uma avaliação da aplicabilidade na construção da Wordnet.Br. Instituto de Ciências Matemáticas e de Computação - USP. Relatórios Técnicos do Icmc, nº 268. 2006. Disponível em <[http://www.icmc.usp.br/~biblio/download/RT\\_268.pdf](http://www.icmc.usp.br/~biblio/download/RT_268.pdf)>. Acesso em: 25 maio 2006.

AMARAL, Luiz. Um Modelo de Restrições Semântico-Selecionais para Sistemas de Processamento de Linguagem Natural. *Jornal Veredas*, v.9, p 31 - 45. Juiz de Fora - MG. 2003.

ANDROUTSOPOULOS, I.; DALE, R. Selectional Restrictions in HPSG. In *Proceedings of COLING 2000*, p. 15-20. 2000.

BARRICHELLO, E.; POZZOBON, C.; RIBEIRO, M. Comunicação Informal e Cultura Organizacional. *Revista Comunicação Organizacional*. 2003. Disponível em <<http://www.pucrs.br/famecos/geacor/texto4-03.html>>. Acesso em: 12 de abr. 2006.

BEAR, J.; HOBBS, J. R. Localizing Expression of Ambiguity. *Proceedings of the Second Conference on Applied Natural Language Processing*, Austin, Texas, pp. 235-241. 1988

BORBA, Francisco da Silva (Coord.). *Dicionário gramatical de verbos do Português Contemporâneo do Brasil*. São Paulo: UNESP, 1990.

BORBA, Francisco da Silva (Coord.). *Dicionário de usos do Português do Brasil*. São Paulo: Ática, 2002.

BORBA, Francisco da Silva. *Uma gramática de valências para o português*. São Paulo : Ática, 1996.

BRÄSCHER, M. Tratamento automático de ambigüidades na recuperação da informação. 286 p. Tese (Tese de Doutorado) — Universidade de Brasília, Brasília, 1999.

BRÄSCHER, M. A ambigüidade na recuperação da informação. *DataGramaZero - Revista de Ciência da Informação*, v. 3, n. 1, p. 40, fev. 2002.

CHOMSKY, N. Three models for the description of language. *IRE Transactions PGIT*, 2. (pp. 113-124), 1956.

CHOMSKY, N. *Syntactic Structures*. Haia: Mouton, 1957.

COUTINHO, M.. Ecologia Polissêmica. *IMAGINARIO - USP (REVISTA DO NUCLEO INTERDISCIPLINAR DO IMAGINARIO E MEMORIA*, Universidade de São Paulo, v. 3, p. 121-149, 1996.

CUNHA, C.; CINTRA, L. F. L. Nova gramática do português contemporâneo. 3. ed. – Rio de Janeiro: Nova Fronteira, 2001.

DENICULI, Camila Dias. Panorama nacional sobre a utilização de MARC e Protocolo Z39.50, através das Bibliotecas de Universidades Federais Brasileiras até 2004. Dissertação (Mestrado em Ciência da Informação) - Escola de Ciência da Informação, UFMG, Belo Horizonte. 2004.

D'EHARBE, David. Gramáticas. 2003. Disponível em <<http://www.consiste.dimap.ufrn.br/~david/ENSEIGNEMENT/SUPPORT/330-gramaticas.pdf>>. Acesso em: 17 maio 2006.

DZIEKANIAK, Gisele Vasconcelos; KIRINUS, Josiane Boeira. Web semântica. Encontros Bibli, Florianópolis, v. 18, p. 20-39, 2. sem. 2004. Disponível em: <[http://www.encontros-bibli.ufsc.br/Edicao\\_18/2\\_Web\\_Semantica.pdf](http://www.encontros-bibli.ufsc.br/Edicao_18/2_Web_Semantica.pdf)>. Acesso em: 02 mar. 2006.

FARIA, Maria Do Carmo B. Aristóteles Plenitude Com Horizonte Do Ser (C.Logos). MODERNA, 1995.

FELIPPO, Ariani Di ; DIAS-DA-SILVA, B.C. . Modelo lingüístico-computacional da estrutura valencial de adjetivos do português do Brasil. Estudos Lingüísticos, São Paulo, v. 34, n. 1, p. 857-862, 2005.

FILLMORE, Charles J. The case for case. In: BACH, E. & HARMS, R.T. (Ed.) *Universals in Linguistic Theory*, pp.1- 88. New York: Holt, Rinehart and Winston, 1968.

FURNIVAL, A. C. Os fundamentos da lógica aplicada à recuperação da informação. Editora da Universidade Federal de São Carlos. São Carlos. 2002.

FUCHS, C. Les ambiguïtés du français. Paris : Orphys, 1996. 183p.

GENNARI, J.; MUSEN, M. A.; FERGERSON, R. W.; GROSSO, W. E.; CRUBÉZY, M.; ERIKSSON, H.; NOY, N. F.; S.; TU, W. The Evolution of Protégé: An Environment for Knowledge-Based Systems Development. 2002. Disponível em <<http://protege.stanford.edu/doc/users.html>>. Acesso em: 02 jun. 2006.

GIL, A. C. Métodos e Técnicas da Pesquisa Social. 5. ed. São Paulo: Editora Atlas, 1999.

HOBBS, J. Overview of the TACITUS Project, CL, Vol. 12, No. 3. 1986.

HOUAISS, A. Dicionario Houaiss da língua Portuguesa. Objetiva, 1ª Edição. 2004.

KNUBLAUCH, H.; FERGERSON, R. W.; NOY, N. F.; MUSEN, M. A. The Protégé OWL Plugin: An Open Development Environment for Semantic Web Applications. Third International Semantic Web Conference. 2004.

Le COADIC, Yves-François. A ciência da informação. Brasília: Briquet de Lemos Livros, 1996.

LIMA-MARQUES, M. Ontologias: da filosofia à representação do conhecimento. Brasília: Thesaurus, 2006.

LUSTOSA, V. G.; ALVARENGA, R. O Estado da Arte em Inteligência Artificial. Colabor@ - Revista Digital da CVA-Ricesu. ISSN 1519-8529. 2004. Disponível em: <[http://www.ricesu.com.br/colabora/n8/artigos/n\\_8/pdf/id\\_03.pdf](http://www.ricesu.com.br/colabora/n8/artigos/n_8/pdf/id_03.pdf)>. Acesso em: 01 maio 2006.

MACCOMB, D. Semantics in business systems: the savvy manager's guide: the discipline underlying web-services, business rules, and the semantic web. Morgan Kaufman, 2004.

MANOSSO, R. Gramática descritiva. 2002. Disponível em <http://www.radames.manosso.nom.br>. Acesso em: 10 maio 2006.

Martins, R.T.M.; HASEGAWA, R.; NUNES, M.D.G.V.; MONTILHA, G.; OLIVEIRA JR, O. N. Linguistic issues in the development of ReGra: a grammar checker for Brazilian Portuguese. *Natural Language Engineering* 4 (4): 287-307. 1998a.

MARTINS, R.T.M. Linguistic issues in the development of ReGra: a grammar checker for Brazilian Portuguese. *Natural Language Engineering* 4 (4): 287-307. 1998.

MARTINS, R.T.; RINO, L.H.M.; MONTILHA, G.; NUNES, M.G.V. Dos modelos de resolução da ambigüidade categorial: O problema do SE. IV Encontro para o Processamento Computacional da Língua Portuguesa Escrita e Falada (PROPOR'99). Universidade de Évora, Portugal. 1999.

MARTINS, C. A indeterminação do significado nos estudos sócio pragmáticos: Divergências teórico-metodológicas. *D.E.L.T.A.*, 18:1, 2002.

MOURA, A. M. de C. A web semântica: fundamentos e tecnologias. Rio de Janeiro: IME, 2002. Disponível em: <http://ipanema.ime.eb.br/~anamoura/publicacoes.html>. Acesso em: 02 maio 2006.

NUNES, M. G. V.; MACEDO, A. A.; DOSUALDO, D. G.; BARBOSA, T. SCE-123: Introdução a Compilação. 1999. Disponível em <<http://www.icmc.usp.br/~gracan/download/sce126/>>. Acesso em: 18 maio 2006.

OLIVEIRA, E. C. Towards a new authoring environment: overview of some ontology-based systems. In: Jan Engelen (Ed.), *Proceedings of the 8th Iccc International Conference on Electronic Publishing*, p. 121-130. 2004.

OLIVEIRA, E. C.; LIMA-MARQUES, M. An Architecture of Authoring Environments for the Semantic Web. *ELPUB 2006 Proceedings*. 2006. Sujeito a alterações.

OLIVEIRA, E. C. Autoria de documentos para a Web Semântica: um ambiente de produção de conhecimento baseado em ontologias. Universidade de Brasília, UnB. Tese de doutorado. 2006.

POWERS, S. *Practical RDF*. O'Reilly, 2003.

ROBREDO, Jaime. Da Ciência da Informação Revisitada aos Sistemas Humanos de Informação. Brasília: Thesaurus, 2003.

SAILER, M. Local Semantics in Head-Driven Phrase Structure Grammar. In: Olivier Bonami and Patricia Cabredo Hofherr: Empirical Issues in Formal Syntax and Semantics 5. p. 197-214. 2004.

SHANNON, C.; WEAVER, W. The Mathematical Theory of Communication. University of Illinois Press, Urbana, 1949.

SILVA, Lúcio Buzon. Ambigüidades da língua portuguesa: recorte classificatório para a elaboração de um modelo ontológico. Dissertação (Mestrado em Ciência da Informação) – Departamento de Ciência da Informação e Documentação, UNB, Brasília. 2006.

SOWA, J. F. Conceptual Structures: information processing in mind and machine. Massachusetts: Addison-Wesley, 1984. 435 p.

STROGENSKI, Paulo J.R. Discurso e expressão nos estudos da linguagem. Revista de Letras 7. UTFPR: Curitiba, 2005. Disponível em : <<http://www.cefetpr.br/deptos/dacex/paulo7.htm>>. Acesso em: 26 maio 2006.

SUPREMO TRIBUNAL FEDERAL. Jobim: a Constituição precisa de uma "lipoaspiração". 2006. Disponível em: <<http://www.stf.gov.br/noticias/imprensa/ultimas/ler.asp?CODIGO=177347&tip=UN&param=>>>. Acesso em: 12 jul. 2006.

TESNIÈRE, Lucien. Eléments de syntaxe structurale. Paris: Klincksieck, 1959.

THE LIBRARY OF CONGRESS. What is a marc record, and why is it important?. 2003. Disponível em <http://www.loc.gov/marc/umb/um01to06.html>. Acesso em: 12 mai. 2006.

TUBBS, S. L.; MOSS, S. Human communication: principles and contexts. 9. Ed. Boston: McGraw Hill, 2003.

VIEIRA, R. e LIMA, V. L. S. Lingüística Computacional: Princípios e Aplicações. In: As Tecnologias da Informação e a questão social: anais. Carlos Eduardo Ferreira (Ed.) Fortaleza, SBC. ISBN 85-88442-03-5 (v.2). p. 47-88. 2001.

WASOW, T.; PERFORS, A.; BEAVER, D. The Puzzle of Ambiguity. In O. Orgun and P. Sells (eds) Morphology and The Web of Grammar: Essays in Memory of Steven G. Lapointe. CSLI Publications. 2005.

WELKER, H. A. A valência verbal em três dicionários brasileiros. Linguagem & Ensino, Vol. 8, no. 1, 2005.

WELKER, H. A. Traduzir frases isoladas na aula de língua estrangeira - por que não? Horizontes de Lingüística Aplicada, Brasília, 2.3, 2003, pp. 149-163.

WIKIPEDIA. Ambiguity. In: Wikipedia, the free encyclopedia. [s.n.], 2005. Disponível em<<http://en.wikipedia.org/wiki/Ambiguity>>. Acesso em: 26 maio 2006.

WIKIPEDIA. A Gramática gerativa. In: Wikipedia, a enciclopédia livre. [s.n.], 2006a. Disponível em <[http://pt.wikipedia.org/wiki/Gram%C3%A1tica\\_gerativa](http://pt.wikipedia.org/wiki/Gram%C3%A1tica_gerativa)>. Acesso em: 10 maio 2006.

WIKIPEDIA. Valência (química). In: Wikipedia, a enciclopédia livre. [s.n.], 2006b. Disponível em <<http://pt.wikipedia.org/wiki/Val%C3%Aancia>>. Acesso em: 02 maio 2006.

World Wide Web Consortium (W3C). OWL Web Ontology Language Overview. 2004. Disponível em <<http://www.w3.org/TR/owl-features/>>. Acesso em: 31 maio 2006.

World Wide Web Consortium (W3C). Resource Description Framework (RDF): Concepts and Abstract Syntax. 2004b. Disponível em <<http://www.w3.org/TR/rdf-concepts/>>. Acesso em: 01 jun. 2006.

World Wide Web Consortium (W3C). Authoring Tool Accessibility Guidelines 2.0. 2005. Disponível em <<http://www.w3.org/TR/2005/WD-ATAG20-20051123/>>. Acesso em: 07 maio 2006.

World Wide Web Consortium (W3C). Web Accessibility Initiative. 2006a. Disponível em <<http://www.w3.org/WAI/>>. Acesso em: 07 maio 2006.

World Wide Web Consortium (W3C). Extensible Markup Language (XML). 2006b. Disponível em <<http://www.w3.org/XML/>>. Acesso em: 07 maio 2006.

ZAVAGLIA, C. Ambigüidade gerada pela homonímia: revisão teórica, linhas limítrofes com a polissemia e proposta de critérios distintivos. São Paulo: DELTA, 2003.

ZINGLÉ, H. The Zstation workbench and the modelling of linguistic knowledge. In: CURRENT issues in mathematical linguistics. North Holland: Elsevier, 1994.

# ANEXO 1

## EXEMPLO ILUSTRATIVO DA ONTOLOGIA DAS VALÊNCIAS DOS VERBOS, NOMES, ADJETIVOS E ADVÉRBIOS EM OWL

### A.1 – Introdução

O intuito deste anexo é demonstrar na prática a ontologia de valências dos verbos, nomes, adjetivos e advérbios em linguagem OWL. O arquivo OWL descrito abaixo foi editado na ferramenta Protege e contém a estrutura completa de classes e propriedades descritas na seção 4.2 desta dissertação. Entretanto, em função dos fins meramente ilustrativos deste anexo, o número de instâncias das classes `CLASSE_FORMA_CANONICA` e `CLASSE_FORMA_DERIVADA` é bastante reduzido.

### A.2 – Estrutura de Classes e Instâncias do Exemplo

TAXONOMIA DE CLASSES		INSTÂNCIAS	
Owl: Thing	CLASSE_FORMA_CANONICA	CLASSE_PRONOME	Aquele Esse Eu Isto
		CLASSE_CONJUNCAO	Quando Que
		CLASSE_NUMERAL	Um Dois Trezentos

		<b>CLASSE_PREPOSICAO</b>	a ante após até com contra de desde em entre para perante por sem sob sobre trás
		<b>CLASSE_ADJETIVO</b>	Belo
		<b>CLASSE_NOME</b>	Banco Banco Telescópio Vendedor
		<b>CLASSE_VERBO</b>	Amar Bancar Chover Espirrar Traduzir Ver Ver
		<b>CLASSE_ADVERBIO</b>	Abaixo Agora Felizmente
		<b>CLASSE_ARTIGO</b>	O Um
	<b>CLASSE_CATEGORIA_MORFOLOGICA</b>		Adjetivo Advérbio Artigo Conjunção Nome Numeral Preposição Pronome Verbo

	<b>CLASSE_CATEGORIA_SINTATICA</b>	Functor Adjetivo Functor Adjetivo Aridade 1 Functor Adjetivo Aridade 2 Functor Adjetivo Aridade 3 Functor Adjetivo Aridade 4 Functor Adverbial Functor Adverbial Aridade 1 Functor Nominal Functor Nominal Aridade 0 Functor Nominal Aridade 1 Functor Nominal Aridade 2 Functor Nominal Aridade 3 Functor Nominal Aridade 4 Functor Verbal Functor Verbal Aridade 0 Functor Verbal Aridade 1 Functor Verbal Aridade 2 Functor Verbal Aridade 3 Functor Verbal Aridade 4
	<b>CLASSE_TRACO</b>	Abstrato Alienável Alimento Animado Concreto Contável Humano Idioma Não Alienável Não Animado Não Contável Não Humano Sólido
	<b>CLASSE_FORMA_DERIVADA</b>	banco bancos um uma uns vê

	<b>CLASSE_CASO</b>	Agentivo Beneficiário Causativo Comitativo Direcional Experimentador Locativo Meta Objetivo Origem Percurso Resultativo Temporal
--	--------------------	--

### A.3 – Código-fonte OWL do exemplo

```

<?xml version="1.0"?>
<rdf:RDF
  xmlns:protege="http://protege.stanford.edu/plugins/owl/protege#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns="http://www.owl-ontologies.com/unnamed.owl#"
  xml:base="http://www.owl-ontologies.com/unnamed.owl">
  <owl:Ontology rdf:about="">
    <owl:imports rdf:resource="http://protege.stanford.edu/plugins/owl/protege"/>
  </owl:Ontology>
  <owl:Class rdf:ID="CLASSE_PRONOME">
    <rdfs:subClassOf>
      <owl:Class rdf:ID="CLASSE_FORMA_CANONICA"/>
    </rdfs:subClassOf>
  </owl:Class>
  <owl:Class rdf:ID="CLASSE_PREPOSICAO">
    <rdfs:subClassOf rdf:resource="#CLASSE_FORMA_CANONICA"/>
  </owl:Class>
  <owl:Class rdf:ID="CLASSE_TRACO"/>
  <owl:Class rdf:ID="CLASSE_CASO"/>
  <owl:Class rdf:ID="CLASSE_NUMERAL">
    <rdfs:subClassOf rdf:resource="#CLASSE_FORMA_CANONICA"/>
  </owl:Class>
  <owl:Class rdf:ID="CLASSE_NOME">
    <rdfs:subClassOf rdf:resource="#CLASSE_FORMA_CANONICA"/>
  </owl:Class>
  <owl:Class rdf:ID="CLASSE_ADJETIVO">
    <rdfs:subClassOf rdf:resource="#CLASSE_FORMA_CANONICA"/>

```

```

</owl:Class>
<owl:Class rdf:ID="CLASSE_FORMA_DERIVADA"/>
<owl:Class rdf:ID="CLASSE_VERBO">
  <rdfs:subClassOf rdf:resource="#CLASSE_FORMA_CANONICA"/>
</owl:Class>
<owl:Class rdf:ID="CLASSE_CATEGORIA_MORFOLOGICA"/>
<owl:Class rdf:ID="CLASSE_ADVERBIO">
  <rdfs:subClassOf rdf:resource="#CLASSE_FORMA_CANONICA"/>
</owl:Class>
<owl:Class rdf:ID="CLASSE_CONJUNCAO">
  <rdfs:subClassOf rdf:resource="#CLASSE_FORMA_CANONICA"/>
</owl:Class>
<owl:Class rdf:ID="CLASSE_CATEGORIA_SINTATICA"/>
<owl:Class rdf:ID="CLASSE_ARTIGO">
  <rdfs:subClassOf rdf:resource="#CLASSE_FORMA_CANONICA"/>
</owl:Class>
<owl:ObjectProperty rdf:ID="oprop_tem_traco_semantico_arg1">
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#CLASSE_ADJETIVO"/>
        <owl:Class rdf:about="#CLASSE_NOME"/>
        <owl:Class rdf:about="#CLASSE_VERBO"/>
        <owl:Class rdf:about="#CLASSE_ADVERBIO"/>
      </owl:unionOf>
    </owl:Class>
  </rdfs:domain>
  <rdfs:range rdf:resource="#CLASSE_TRACO"/>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="oprop_tem_cat_sint_arg3">
  <rdfs:range rdf:resource="#CLASSE_CATEGORIA_SINTATICA"/>
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#CLASSE_ADJETIVO"/>
        <owl:Class rdf:about="#CLASSE_NOME"/>
        <owl:Class rdf:about="#CLASSE_VERBO"/>
      </owl:unionOf>
    </owl:Class>
  </rdfs:domain>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="oprop_e_forma_derivada_de">
  <rdfs:domain rdf:resource="#CLASSE_FORMA_DERIVADA"/>
  <rdfs:range rdf:resource="#CLASSE_FORMA_CANONICA"/>
</owl:ObjectProperty>

```

```

<owl:ObjectProperty rdf:ID="oprop_tem_prep_arg4">
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#CLASSE_ADJETIVO"/>
        <owl:Class rdf:about="#CLASSE_NOME"/>
        <owl:Class rdf:about="#CLASSE_VERBO"/>
      </owl:unionOf>
    </owl:Class>
  </rdfs:domain>
  <rdfs:range rdf:resource="#CLASSE_PREPOSICAO"/>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="oprop_e_um">
  <rdfs:range rdf:resource="#CLASSE_CATEGORIA_SINTATICA"/>
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/>
  <rdfs:domain rdf:resource="#CLASSE_CATEGORIA_SINTATICA"/>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="oprop_tem_cat_sint_arg4">
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#CLASSE_ADJETIVO"/>
        <owl:Class rdf:about="#CLASSE_NOME"/>
        <owl:Class rdf:about="#CLASSE_VERBO"/>
      </owl:unionOf>
    </owl:Class>
  </rdfs:domain>
  <rdfs:range rdf:resource="#CLASSE_CATEGORIA_SINTATICA"/>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="oprop_tem_prep_arg3">
  <rdfs:range rdf:resource="#CLASSE_PREPOSICAO"/>
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#CLASSE_ADJETIVO"/>
        <owl:Class rdf:about="#CLASSE_NOME"/>
        <owl:Class rdf:about="#CLASSE_VERBO"/>
      </owl:unionOf>
    </owl:Class>
  </rdfs:domain>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="oprop_tem_caso4">
  <rdfs:range rdf:resource="#CLASSE_CASO"/>
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/>
  <rdfs:domain>

```

```

<owl:Class>
  <owl:unionOf rdf:parseType="Collection">
    <owl:Class rdf:about="#CLASSE_ADJETIVO"/>
    <owl:Class rdf:about="#CLASSE_NOME"/>
    <owl:Class rdf:about="#CLASSE_VERBO"/>
  </owl:unionOf>
</owl:Class>
</rdfs:domain>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="oprop_tem_traco_semantico_arg4">
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#CLASSE_ADJETIVO"/>
        <owl:Class rdf:about="#CLASSE_NOME"/>
        <owl:Class rdf:about="#CLASSE_VERBO"/>
      </owl:unionOf>
    </owl:Class>
  </rdfs:domain>
  <rdfs:range rdf:resource="#CLASSE_TRACO"/>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="oprop_tem_caso1">
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#CLASSE_ADJETIVO"/>
        <owl:Class rdf:about="#CLASSE_NOME"/>
        <owl:Class rdf:about="#CLASSE_VERBO"/>
      </owl:unionOf>
    </owl:Class>
  </rdfs:domain>
  <rdfs:type rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/>
  <rdfs:range rdf:resource="#CLASSE_CASO"/>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="oprop_tem_prep_arg2">
  <rdfs:range rdf:resource="#CLASSE_PREPOSICAO"/>
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#CLASSE_ADJETIVO"/>
        <owl:Class rdf:about="#CLASSE_NOME"/>
        <owl:Class rdf:about="#CLASSE_VERBO"/>
      </owl:unionOf>
    </owl:Class>
  </rdfs:domain>

```

```

</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="oprop_tem_traco_semantico_arg3">
  <rdfs:range rdf:resource="#CLASSE_TRACO"/>
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#CLASSE_ADJETIVO"/>
        <owl:Class rdf:about="#CLASSE_NOME"/>
        <owl:Class rdf:about="#CLASSE_VERBO"/>
      </owl:unionOf>
    </owl:Class>
  </rdfs:domain>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="oprop_tem_cat_sint_arg1">
  <rdfs:range rdf:resource="#CLASSE_CATEGORIA_SINTATICA"/>
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#CLASSE_ADJETIVO"/>
        <owl:Class rdf:about="#CLASSE_NOME"/>
        <owl:Class rdf:about="#CLASSE_VERBO"/>
        <owl:Class rdf:about="#CLASSE_ADVERBIO"/>
      </owl:unionOf>
    </owl:Class>
  </rdfs:domain>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="oprop_tem_traco_semantico">
  <rdfs:domain rdf:resource="#CLASSE_NOME"/>
  <rdfs:range rdf:resource="#CLASSE_TRACO"/>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="oprop_tem_prep_arg1">
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#CLASSE_ADJETIVO"/>
        <owl:Class rdf:about="#CLASSE_NOME"/>
        <owl:Class rdf:about="#CLASSE_VERBO"/>
        <owl:Class rdf:about="#CLASSE_ADVERBIO"/>
      </owl:unionOf>
    </owl:Class>
  </rdfs:domain>
  <rdfs:range rdf:resource="#CLASSE_PREPOSICAO"/>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="oprop_tem_traco_semantico_arg2">
  <rdfs:domain>

```

```

<owl:Class>
  <owl:unionOf rdf:parseType="Collection">
    <owl:Class rdf:about="#CLASSE_ADJETIVO"/>
    <owl:Class rdf:about="#CLASSE_NOME"/>
    <owl:Class rdf:about="#CLASSE_VERBO"/>
  </owl:unionOf>
</owl:Class>
</rdfs:domain>
<rdfs:range rdf:resource="#CLASSE_TRACO"/>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="oprop_tem_cat_sint_arg2">
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#CLASSE_ADJETIVO"/>
        <owl:Class rdf:about="#CLASSE_NOME"/>
        <owl:Class rdf:about="#CLASSE_VERBO"/>
      </owl:unionOf>
    </owl:Class>
  </rdfs:domain>
  <rdfs:range rdf:resource="#CLASSE_CATEGORIA_SINTATICA"/>
</owl:ObjectProperty>
<owl:DatatypeProperty rdf:ID="dprop_nome">
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#CLASSE_FORMA_CANONICA"/>
        <owl:Class rdf:about="#CLASSE_CATEGORIA_MORFOLOGICA"/>
        <owl:Class rdf:about="#CLASSE_CATEGORIA_SINTATICA"/>
        <owl:Class rdf:about="#CLASSE_TRACO"/>
        <owl:Class rdf:about="#CLASSE_FORMA_DERIVADA"/>
        <owl:Class rdf:about="#CLASSE_CASO"/>
      </owl:unionOf>
    </owl:Class>
  </rdfs:domain>
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
</owl:DatatypeProperty>
<owl:DatatypeProperty rdf:ID="dprop_abbreviatura">
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#CLASSE_CATEGORIA_MORFOLOGICA"/>

```

```

    <owl:Class rdf:about="#CLASSE_CATEGORIA_SINTATICA"/>
    <owl:Class rdf:about="#CLASSE_CASO"/>
  </owl:unionOf>
</owl:Class>
</rdfs:domain>
</owl:DatatypeProperty>
<owl:DatatypeProperty rdf:ID="dprop_arg2_obrigatorio">
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#boolean"/>
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#CLASSE_ADJETIVO"/>
        <owl:Class rdf:about="#CLASSE_NOME"/>
        <owl:Class rdf:about="#CLASSE_VERBO"/>
        <owl:Class rdf:about="#CLASSE_ADVERBIO"/>
      </owl:unionOf>
    </owl:Class>
  </rdfs:domain>
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/>
</owl:DatatypeProperty>
<owl:DatatypeProperty rdf:ID="dprop_descricao">
  <rdfs:domain rdf:resource="#CLASSE_FORMA_CANONICA"/>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/>
</owl:DatatypeProperty>
<owl:FunctionalProperty rdf:ID="oprop_tem_caso3">
  <rdfs:range rdf:resource="#CLASSE_CASO"/>
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#CLASSE_ADJETIVO"/>
        <owl:Class rdf:about="#CLASSE_NOME"/>
        <owl:Class rdf:about="#CLASSE_VERBO"/>
      </owl:unionOf>
    </owl:Class>
  </rdfs:domain>
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#ObjectProperty"/>
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="dprop_arg4_obrigatorio">
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#boolean"/>
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#DatatypeProperty"/>
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#CLASSE_ADJETIVO"/>

```

```

    <owl:Class rdf:about="#CLASSE_NOME"/>
    <owl:Class rdf:about="#CLASSE_VERBO"/>
    <owl:Class rdf:about="#CLASSE_ADVERBIO"/>
  </owl:unionOf>
</owl:Class>
</rdfs:domain>
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="dprop_arg1_obrigatorio">
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#DatatypeProperty"/>
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#CLASSE_ADJETIVO"/>
        <owl:Class rdf:about="#CLASSE_NOME"/>
        <owl:Class rdf:about="#CLASSE_VERBO"/>
        <owl:Class rdf:about="#CLASSE_ADVERBIO"/>
      </owl:unionOf>
    </owl:Class>
  </rdfs:domain>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#boolean"/>
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="oprop_tem_caso2">
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#CLASSE_ADJETIVO"/>
        <owl:Class rdf:about="#CLASSE_NOME"/>
        <owl:Class rdf:about="#CLASSE_VERBO"/>
      </owl:unionOf>
    </owl:Class>
  </rdfs:domain>
  <rdfs:range rdf:resource="#CLASSE_CASO"/>
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#ObjectProperty"/>
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="dprop_numero_argumento">
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#CLASSE_ADJETIVO"/>
        <owl:Class rdf:about="#CLASSE_NOME"/>
        <owl:Class rdf:about="#CLASSE_VERBO"/>
        <owl:Class rdf:about="#CLASSE_ADVERBIO"/>
      </owl:unionOf>
    </owl:Class>
  </rdfs:domain>

```

```

<rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#int"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#DatatypeProperty"/>
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="dprop_arg3_obrigatorio">
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#boolean"/>
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#CLASSE_ADJETIVO"/>
        <owl:Class rdf:about="#CLASSE_NOME"/>
        <owl:Class rdf:about="#CLASSE_VERBO"/>
        <owl:Class rdf:about="#CLASSE_ADVERBIO"/>
      </owl:unionOf>
    </owl:Class>
  </rdfs:domain>
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#DatatypeProperty"/>
</owl:FunctionalProperty>
<CLASSE_TRACO rdf:ID="ontoval_dissertacao_Instance_12">
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Idioma</dprop_nome>
</CLASSE_TRACO>
<CLASSE_CASO rdf:ID="CASO_Beneficiario">
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Beneficiário</dprop_nome>
  <dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >B</dprop_abreviatura>
</CLASSE_CASO>
<CLASSE_CASO rdf:ID="CASO_Experimentador">
  <dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Ex</dprop_abreviatura>
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Experimentador</dprop_nome>
</CLASSE_CASO>
<CLASSE_CATEGORIA_SINTATICA rdf:ID="SINT_FADJ2">
  <oprop_e_um>
    <CLASSE_CATEGORIA_SINTATICA rdf:ID="SINT_FADJ">
      <dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
      >FADJ</dprop_abreviatura>
      <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
      >Functor Adjetivo</dprop_nome>
    </CLASSE_CATEGORIA_SINTATICA>
  </oprop_e_um>
  <dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >FADJ2</dprop_abreviatura>
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"

```

```

    >Functor Adjetivo Aridade 2</dprop_nome>
</CLASSE_CATEGORIA_SINTATICA>
<CLASSE_TRACO rdf:ID="TRACO_Contavel">
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Contável</dprop_nome>
</CLASSE_TRACO>
<CLASSE_CATEGORIA_MORFOLOGICA rdf:ID="M_CAT_Preposicao">
  <dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >PREP</dprop_abreviatura>
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Preposição</dprop_nome>
</CLASSE_CATEGORIA_MORFOLOGICA>
<CLASSE_CASO rdf:ID="CASO_Meta">
  <dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >M</dprop_abreviatura>
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Meta</dprop_nome>
</CLASSE_CASO>
<CLASSE_CATEGORIA_SINTATICA rdf:ID="SINT_FADJ1">
  <oprop_e_um rdf:resource="#SINT_FADJ"/>
  <dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >FADJ1</dprop_abreviatura>
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Functor Adjetivo Aridade 1</dprop_nome>
</CLASSE_CATEGORIA_SINTATICA>
<CLASSE_CATEGORIA_SINTATICA rdf:ID="SINT_FN3">
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Functor Nominal Aridade 3</dprop_nome>
  <dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >FN3</dprop_abreviatura>
  <oprop_e_um>
  <CLASSE_CATEGORIA_SINTATICA rdf:ID="SINT_FN">
    <dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
    >FN</dprop_abreviatura>
    <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
    >Functor Nominal</dprop_nome>
  </CLASSE_CATEGORIA_SINTATICA>
  </oprop_e_um>
</CLASSE_CATEGORIA_SINTATICA>
<CLASSE_FORMA_DERIVADA rdf:ID="FD_VE">
  <oprop_e_forma_derivada_de>
  <CLASSE_VERBO rdf:ID="FC_V_Ver2">
    <oprop_tem_traco_semantico_arg1>
    <CLASSE_TRACO rdf:ID="TRACO_Humano">
      <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"

```

```

    >Humano</dprop_nome>
  </CLASSE_TRACO>
</oprop_tem_traco_semantico_arg1>
<dprop_descricao rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Conhecer.</dprop_descricao>
<oprop_tem_caso2>
  <CLASSE_CASO rdf:ID="CASO_Objetivo">
    <dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
    >Ob</dprop_abreviatura>
    <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
    >Objetivo</dprop_nome>
  </CLASSE_CASO>
</oprop_tem_caso2>
<dprop_numero_argumento rdf:datatype="http://www.w3.org/2001/XMLSchema#int"
>2</dprop_numero_argumento>
<dprop_arg2_obrigatorio rdf:datatype="http://www.w3.org/2001/XMLSchema#boolean"
>true</dprop_arg2_obrigatorio>
<oprop_tem_traco_semantico_arg2>
  <CLASSE_TRACO rdf:ID="TRACO_Animado">
    <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
    >Animado</dprop_nome>
  </CLASSE_TRACO>
</oprop_tem_traco_semantico_arg2>
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Ver</dprop_nome>
<dprop_arg1_obrigatorio rdf:datatype="http://www.w3.org/2001/XMLSchema#boolean"
>true</dprop_arg1_obrigatorio>
<oprop_tem_caso1 rdf:resource="#CASO_Experimentador"/>
</CLASSE_VERBO>
</oprop_e_forma_derivada_de>
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>vê</dprop_nome>
<oprop_e_forma_derivada_de>
  <CLASSE_VERBO rdf:ID="FC_V_Ver1">
    <oprop_tem_cat_sint_arg2 rdf:resource="#SINT_FN"/>
    <oprop_tem_traco_semantico_arg1 rdf:resource="#TRACO_Animado"/>
    <oprop_tem_caso1 rdf:resource="#CASO_Experimentador"/>
  </oprop_tem_traco_semantico_arg2>
  <CLASSE_TRACO rdf:ID="TRACO_Concreto">
    <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
    >Concreto</dprop_nome>
  </CLASSE_TRACO>
</oprop_tem_traco_semantico_arg2>
<oprop_tem_caso2 rdf:resource="#CASO_Objetivo"/>
<oprop_tem_cat_sint_arg2 rdf:resource="#SINT_FADJ"/>

```

```

<dprop_descricao rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Perceber pela visão; enxergar.</dprop_descricao>
<oprop_tem_cat_sint_arg1 rdf:resource="#SINT_FN"/>
<oprop_tem_cat_sint_arg1 rdf:resource="#SINT_FADJ"/>
<dprop_numero_argumento rdf:datatype="http://www.w3.org/2001/XMLSchema#int"
>2</dprop_numero_argumento>
<dprop_arg1_obrigatorio rdf:datatype="http://www.w3.org/2001/XMLSchema#boolean"
>true</dprop_arg1_obrigatorio>
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Ver</dprop_nome>
</CLASSE_VERBO>
</oprop_e_forma_derivada_de>
</CLASSE_FORMA_DERIVADA>
<CLASSE_PREPOSICAO rdf:ID="PREP_tras">
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>trás</dprop_nome>
</CLASSE_PREPOSICAO>
<CLASSE_CASO rdf:ID="CASO_Causativo">
<dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Ca</dprop_abreviatura>
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Causativo</dprop_nome>
</CLASSE_CASO>
<CLASSE_CATEGORIA_MORFOLOGICA rdf:ID="ins_m_cat_numeral">
<dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>NUM</dprop_abreviatura>
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Numeral</dprop_nome>
</CLASSE_CATEGORIA_MORFOLOGICA>
<CLASSE_TRACO rdf:ID="ins_traco_nao_alienavel">
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Não Alienável</dprop_nome>
</CLASSE_TRACO>
<CLASSE_PREPOSICAO rdf:ID="PREP_entre">
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>entre</dprop_nome>
</CLASSE_PREPOSICAO>
<CLASSE_CATEGORIA_SINTATICA rdf:ID="SINT_FN2">
<dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>FN2</dprop_abreviatura>
<oprop_e_um rdf:resource="#SINT_FN"/>
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Functor Nominal Aridade 2</dprop_nome>
</CLASSE_CATEGORIA_SINTATICA>
<CLASSE_TRACO rdf:ID="TRACO_Inanimado">

```

```

<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Não Animado</dprop_nome>
</CLASSE_TRACO>
<CLASSE_PREPOSICAO rdf:ID="PREP_para">
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>para</dprop_nome>
</CLASSE_PREPOSICAO>
<CLASSE_CASO rdf:ID="CASO_Origem">
<dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Or</dprop_abreviatura>
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Origem</dprop_nome>
</CLASSE_CASO>
<CLASSE_PRONOME rdf:ID="FC_PRON_Eu">
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Eu</dprop_nome>
<dprop_descricao rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Refere-se a pessoa que fala e funciona como sujeito.</dprop_descricao>
</CLASSE_PRONOME>
<CLASSE_CASO rdf:ID="CASO_Resultativo">
<dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>R</dprop_abreviatura>
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Resultativo</dprop_nome>
</CLASSE_CASO>
<CLASSE_VERBO rdf:ID="ins_amar">
<oprop_tem_cat_sint_arg2 rdf:resource="#SINT_FADJ"/>
<oprop_tem_caso2 rdf:resource="#CASO_Objetivo"/>
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Amar</dprop_nome>
<dprop_arg1_obrigatorio rdf:datatype="http://www.w3.org/2001/XMLSchema#boolean"
>true</dprop_arg1_obrigatorio>
<dprop_numero_argumento rdf:datatype="http://www.w3.org/2001/XMLSchema#int"
>2</dprop_numero_argumento>
<oprop_tem_cat_sint_arg1 rdf:resource="#SINT_FN"/>
<dprop_descricao rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>ter amor, afeição a, sentir ternura ou paixão por.</dprop_descricao>
<oprop_tem_cat_sint_arg2 rdf:resource="#SINT_FN"/>
<dprop_arg2_obrigatorio rdf:datatype="http://www.w3.org/2001/XMLSchema#boolean"
>true</dprop_arg2_obrigatorio>
<oprop_tem_caso1 rdf:resource="#CASO_Experimentador"/>
<oprop_tem_cat_sint_arg1 rdf:resource="#SINT_FADJ"/>
</CLASSE_VERBO>
<CLASSE_CATEGORIA_SINTATICA rdf:ID="SINT_FN4">
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"

```

```

>Functor Nominal Aridade 4</dprop_nome>
<dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>FN4</dprop_abreviatura>
<oprop_e_um rdf:resource="#SINT_FN"/>
</CLASSE_CATEGORIA_SINTATICA>
<CLASSE_VERBO rdf:ID="ontoval_dissertacao_Instance_10">
<oprop_tem_cat_sint_arg4 rdf:resource="#SINT_FN"/>
<oprop_tem_cat_sint_arg2 rdf:resource="#SINT_FADJ"/>
<oprop_tem_traco_semantico_arg4 rdf:resource="#ontoval_dissertacao_Instance_12"/>
<oprop_tem_traco_semantico_arg3 rdf:resource="#ontoval_dissertacao_Instance_12"/>
<oprop_tem_caso1>
<CLASSE_CASO rdf:ID="CASO_Agentivo">
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Agentivo</dprop_nome>
<dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Ag</dprop_abreviatura>
</CLASSE_CASO>
</oprop_tem_caso1>
<oprop_tem_caso2 rdf:resource="#CASO_Objetivo"/>
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Traduzir</dprop_nome>
<oprop_tem_traco_semantico_arg2 rdf:resource="#TRACO_Concreto"/>
<oprop_tem_cat_sint_arg4 rdf:resource="#SINT_FADJ"/>
<oprop_tem_prep_arg3>
<CLASSE_PREPOSICAO rdf:ID="PREP_de">
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>de</dprop_nome>
</CLASSE_PREPOSICAO>
</oprop_tem_prep_arg3>
<oprop_tem_cat_sint_arg1 rdf:resource="#SINT_FADJ"/>
<dprop_descricao rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Transpor de um idioma para outra</dprop_descricao>
<oprop_tem_caso4 rdf:resource="#CASO_Meta"/>
<oprop_tem_cat_sint_arg3 rdf:resource="#SINT_FN"/>
<oprop_tem_cat_sint_arg3 rdf:resource="#SINT_FADJ"/>
<dprop_numero_argumento rdf:datatype="http://www.w3.org/2001/XMLSchema#int"
>4</dprop_numero_argumento>
<oprop_tem_caso3 rdf:resource="#CASO_Origem"/>
<oprop_tem_cat_sint_arg1 rdf:resource="#SINT_FN"/>
<dprop_arg1_obrigatorio rdf:datatype="http://www.w3.org/2001/XMLSchema#boolean"
>true</dprop_arg1_obrigatorio>
<oprop_tem_prep_arg4 rdf:resource="#PREP_para"/>
<oprop_tem_cat_sint_arg2 rdf:resource="#SINT_FN"/>
</CLASSE_VERBO>
<CLASSE_CATEGORIA_SINTATICA rdf:ID="SINT_FADJ4">

```

```

<oprop_e_um rdf:resource="#SINT_FADJ"/>
<dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>FADJ4</dprop_abreviatura>
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Functor Adjetivo Aridade 4</dprop_nome>
</CLASSE_CATEGORIA_SINTATICA>
<CLASSE_PRONOME rdf:ID="FC_DET_esse">
<dprop_descricao rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Aponta para o que está próximo do ouvinte.</dprop_descricao>
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Esse</dprop_nome>
</CLASSE_PRONOME>
<CLASSE_CONJUNCAO rdf:ID="ontoval_dissertacao_Instance_0">
<dprop_descricao rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Na ocasião em que.</dprop_descricao>
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>quando</dprop_nome>
</CLASSE_CONJUNCAO>
<CLASSE_CATEGORIA_SINTATICA rdf:ID="SINT_FADJ3">
<oprop_e_um rdf:resource="#SINT_FADJ"/>
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Functor Adjetivo Aridade 3</dprop_nome>
<dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>FADJ3</dprop_abreviatura>
</CLASSE_CATEGORIA_SINTATICA>
<CLASSE_CATEGORIA_SINTATICA rdf:ID="SINT_FV4">
<dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>FV4</dprop_abreviatura>
<oprop_e_um>
<CLASSE_CATEGORIA_SINTATICA rdf:ID="SINT_FV">
<dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>FV</dprop_abreviatura>
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Functor Verbal</dprop_nome>
</CLASSE_CATEGORIA_SINTATICA>
</oprop_e_um>
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Functor Verbal Aridade 4</dprop_nome>
</CLASSE_CATEGORIA_SINTATICA>
<CLASSE_CATEGORIA_MORFOLOGICA rdf:ID="N_CAT_Pronome">
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Pronome</dprop_nome>
<dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>PRON</dprop_abreviatura>
</CLASSE_CATEGORIA_MORFOLOGICA>

```

```

<CLASSE_CATEGORIA_MORFOLOGICA rdf:ID="ins_m_cat_artigo">
  <dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >ART</dprop_abreviatura>
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Artigo</dprop_nome>
</CLASSE_CATEGORIA_MORFOLOGICA>
<CLASSE_CATEGORIA_MORFOLOGICA rdf:ID="ins_m_cat_conjuncao">
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Conjunção</dprop_nome>
  <dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >CONJ</dprop_abreviatura>
</CLASSE_CATEGORIA_MORFOLOGICA>
<CLASSE_PREPOSICAO rdf:ID="PREP_sem">
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >sem</dprop_nome>
</CLASSE_PREPOSICAO>
<CLASSE_ARTIGO rdf:ID="FC_DET_o">
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >o</dprop_nome>
</CLASSE_ARTIGO>
<CLASSE_TRACO rdf:ID="TRACO_Abstrato">
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Abstrato</dprop_nome>
</CLASSE_TRACO>
<CLASSE_PREPOSICAO rdf:ID="PREP_por">
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >por</dprop_nome>
</CLASSE_PREPOSICAO>
<CLASSE_FORMA_DERIVADA rdf:ID="FD_BANCOS">
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >bancos</dprop_nome>
  <oprop_e_forma_derivada_de>
    <CLASSE_NOME rdf:ID="ontoval_Individual_3">
      <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
      >Banco</dprop_nome>
      <dprop_descricao rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
      >Estabelecimento de Crédito</dprop_descricao>
      <oprop_tem_traco_semantico rdf:resource="#TRACO_Concreto"/>
      <dprop_numero_argumento rdf:datatype="http://www.w3.org/2001/XMLSchema#int"
      >0</dprop_numero_argumento>
    </CLASSE_NOME>
  </oprop_e_forma_derivada_de>
  <oprop_e_forma_derivada_de>
    <CLASSE_NOME rdf:ID="ontoval_Individual_2">
      <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"

```

```

>Banco</dprop_nome>
<oprop_tem_traco_semantico rdf:resource="#TRACO_Contavel"/>
<oprop_tem_traco_semantico>
  <CLASSE_TRACO rdf:ID="TRACO_Solido">
    <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
    >Sólido</dprop_nome>
  </CLASSE_TRACO>
</oprop_tem_traco_semantico>
<dprop_descricao rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Móvel usado como assento</dprop_descricao>
<oprop_tem_traco_semantico rdf:resource="#TRACO_Concreto"/>
<dprop_numero_argumento rdf:datatype="http://www.w3.org/2001/XMLSchema#int"
>0</dprop_numero_argumento>
</CLASSE_NOME>
</oprop_e_forma_derivada_de>
</CLASSE_FORMA_DERIVADA>
<CLASSE_CASO rdf:ID="CASO_Comitativo">
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Comitativo</dprop_nome>
  <dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Co</dprop_abreviatura>
</CLASSE_CASO>
<CLASSE_CATEGORIA_SINTATICA rdf:ID="SINT_FV3">
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Functor Verbal Aridade 3</dprop_nome>
  <oprop_e_um rdf:resource="#SINT_FV"/>
  <dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >FV3</dprop_abreviatura>
</CLASSE_CATEGORIA_SINTATICA>
<CLASSE_NUMERAL rdf:ID="ontoval_dissertacao_Instance_2">
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Um</dprop_nome>
</CLASSE_NUMERAL>
<CLASSE_CASO rdf:ID="CASO_Direcional">
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Direcional</dprop_nome>
  <dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Dr</dprop_abreviatura>
</CLASSE_CASO>
<CLASSE_PREPOSICAO rdf:ID="PREP_ante">
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >ante</dprop_nome>
</CLASSE_PREPOSICAO>
<CLASSE_NUMERAL rdf:ID="ontoval_dissertacao_Instance_3">
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"

```

```

    >Dois</dprop_nome>
  </CLASSE_NUMERAL>
  <CLASSE_CONJUNCAO rdf:ID="ontoval_dissertacao_Instance_1">
    <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
    >que</dprop_nome>
    <dprop_descricao rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
    >Introduz uma oração que expressa um fato certo ou real e funciona como sujeito ou complemento da
    principal.</dprop_descricao>
  </CLASSE_CONJUNCAO>
  <CLASSE_PRONOME rdf:ID="FC_DET_aquele">
    <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
    >Aquele</dprop_nome>
    <dprop_descricao rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
    >Aponta para ser ou coisa afastada do falante.</dprop_descricao>
  </CLASSE_PRONOME>
  <CLASSE_FORMA_DERIVADA rdf:ID="FD_UNE">
    <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
    >uns</dprop_nome>
    <oprop_e_forma_derivada_de>
      <CLASSE_ARTIGO rdf:ID="FC_DET_um">
        <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
        >Um</dprop_nome>
      </CLASSE_ARTIGO>
    </oprop_e_forma_derivada_de>
  </CLASSE_FORMA_DERIVADA>
  <CLASSE_FORMA_DERIVADA rdf:ID="FD_BANCO">
    <oprop_e_forma_derivada_de>
      <CLASSE_VERBO rdf:ID="ontoval_dissertacao_Instance_7">
        <dprop_arg1_obrigatorio rdf:datatype="http://www.w3.org/2001/XMLSchema#boolean"
        >true</dprop_arg1_obrigatorio>
        <oprop_tem_caso1 rdf:resource="#CASO_Agentivo"/>
        <oprop_tem_cat_sint_arg1 rdf:resource="#SINT_FADJ"/>
        <dprop_descricao rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
        >Sustentar.</dprop_descricao>
        <oprop_tem_traco_semantico_arg1 rdf:resource="#TRACO_Humano"/>
        <oprop_tem_cat_sint_arg1 rdf:resource="#SINT_FN"/>
        <oprop_tem_caso2 rdf:resource="#CASO_Beneficiario"/>
        <dprop_arg2_obrigatorio rdf:datatype="http://www.w3.org/2001/XMLSchema#boolean"
        >true</dprop_arg2_obrigatorio>
        <dprop_numero_argumento rdf:datatype="http://www.w3.org/2001/XMLSchema#int"
        >2</dprop_numero_argumento>
        <oprop_tem_cat_sint_arg2 rdf:resource="#SINT_FN"/>
        <oprop_tem_traco_semantico_arg2 rdf:resource="#TRACO_Humano"/>
        <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
        >Bancar</dprop_nome>
      </CLASSE_VERBO>
    </oprop_e_forma_derivada_de>
  </CLASSE_FORMA_DERIVADA>

```

```

    <oprop_tem_cat_sint_arg2 rdf:resource="#SINT_FADJ"/>
  </CLASSE_VERBO>
</oprop_e_forma_derivada_de>
<oprop_e_forma_derivada_de rdf:resource="#ontoval_Individual_2"/>
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>banco</dprop_nome>
<oprop_e_forma_derivada_de rdf:resource="#ontoval_Individual_3"/>
</CLASSE_FORMA_DERIVADA>
<CLASSE_CASO rdf:ID="CASO_Percurso">
  <dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Pr</dprop_abreviatura>
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Percurso</dprop_nome>
</CLASSE_CASO>
<CLASSE_CATEGORIA_SINTATICA rdf:ID="SINT_FV2">
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Functor Verbal Aridade 2</dprop_nome>
  <oprop_e_um rdf:resource="#SINT_FV"/>
  <dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >FV2</dprop_abreviatura>
</CLASSE_CATEGORIA_SINTATICA>
<CLASSE_TRACO rdf:ID="ins_traco_alienavel">
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Alienável</dprop_nome>
</CLASSE_TRACO>
<CLASSE_ADVERBIO rdf:ID="FC_ADV_Abaixo1">
  <oprop_tem_cat_sint_arg1 rdf:resource="#SINT_FN"/>
  <dprop_arg1_obrigatorio rdf:datatype="http://www.w3.org/2001/XMLSchema#boolean"
  >true</dprop_arg1_obrigatorio>
  <dprop_numero_argumento rdf:datatype="http://www.w3.org/2001/XMLSchema#int"
  >1</dprop_numero_argumento>
  <oprop_tem_traco_semantico_arg1 rdf:resource="#TRACO_Concreto"/>
  <dprop_descricao rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Em direção a um ponto inferior.</dprop_descricao>
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Abaixo</dprop_nome>
  <oprop_tem_prep_arg1 rdf:resource="#PREP_de"/>
  <oprop_tem_cat_sint_arg1 rdf:resource="#SINT_FADJ"/>
</CLASSE_ADVERBIO>
<CLASSE_ADVERBIO rdf:ID="FC_ADV_Felizmente1">
  <dprop_descricao rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Expressa uma avaliação positiva do falante em relação ao que está sendo dito.</dprop_descricao>
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Felizmente</dprop_nome>
  <dprop_numero_argumento rdf:datatype="http://www.w3.org/2001/XMLSchema#int"

```

```

>0</dprop_numero_argumento>
</CLASSE_ADVERBIO>
<CLASSE_CATEGORIA_SINTATICA rdf:ID="SINT_FV1">
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Functor Verbal Aridade 1</dprop_nome>
  <dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >FV1</dprop_abreviatura>
  <oprop_e_um rdf:resource="#SINT_FV"/>
</CLASSE_CATEGORIA_SINTATICA>
<CLASSE_CATEGORIA_MORFOLOGICA rdf:ID="M_CAT_Verbo">
  <dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >V</dprop_abreviatura>
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Verbo</dprop_nome>
</CLASSE_CATEGORIA_MORFOLOGICA>
<CLASSE_VERBO rdf:ID="ontoval_dissertacao_Instance_8">
  <dprop_arg1_obrigatorio rdf:datatype="http://www.w3.org/2001/XMLSchema#boolean"
  >true</dprop_arg1_obrigatorio>
  <oprop_tem_caso1 rdf:resource="#CASO_Agentivo"/>
  <oprop_tem_cat_sint_arg1 rdf:resource="#SINT_FADJ"/>
  <dprop_numero_argumento rdf:datatype="http://www.w3.org/2001/XMLSchema#int"
  >1</dprop_numero_argumento>
  <oprop_tem_traco_semantico_arg1 rdf:resource="#TRACO_Animado"/>
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Espirrar</dprop_nome>
  <dprop_descricao rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Dar espirros.</dprop_descricao>
  <oprop_tem_cat_sint_arg1 rdf:resource="#SINT_FN"/>
</CLASSE_VERBO>
<CLASSE_FORMA_DERIVADA rdf:ID="FD_UM">
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >um</dprop_nome>
  <oprop_e_forma_derivada_de rdf:resource="#FC_DET_um"/>
</CLASSE_FORMA_DERIVADA>
<CLASSE_NOME rdf:ID="FC_N_Vendedor1">
  <oprop_tem_cat_sint_arg1 rdf:resource="#SINT_FADJ"/>
  <oprop_tem_prep_arg1 rdf:resource="#PREP_de"/>
  <dprop_descricao rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Profissional, cujo trabalho é fazer vendas.</dprop_descricao>
  <oprop_tem_traco_semantico rdf:resource="#TRACO_Concreto"/>
  <dprop_arg2_obrigatorio rdf:datatype="http://www.w3.org/2001/XMLSchema#boolean"
  >false</dprop_arg2_obrigatorio>
  <oprop_tem_traco_semantico rdf:resource="#TRACO_Humano"/>
  <oprop_tem_cat_sint_arg1 rdf:resource="#SINT_FN"/>
  <dprop_numero_argumento rdf:datatype="http://www.w3.org/2001/XMLSchema#int"

```

```

>1</dprop_numero_argumento>
<oprop_tem_caso1 rdf:resource="#CASO_Objetivo"/>
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Vendedor</dprop_nome>
<dprop_arg1_obrigatorio rdf:datatype="http://www.w3.org/2001/XMLSchema#boolean"
>>true</dprop_arg1_obrigatorio>
</CLASSE_NOME>
<CLASSE_PREPOSICAO rdf:ID="PREP_sob">
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >sob</dprop_nome>
</CLASSE_PREPOSICAO>
<CLASSE_PREPOSICAO rdf:ID="PREP_contra">
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >contra</dprop_nome>
</CLASSE_PREPOSICAO>
<CLASSE_TRACO rdf:ID="TRACO_Nao_Humano">
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Não Humano</dprop_nome>
</CLASSE_TRACO>
<CLASSE_ADVERBIO rdf:ID="FC_ADV_Agora">
  <dprop_descricao rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Neste instante.</dprop_descricao>
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Agora</dprop_nome>
  <dprop_numero_argumento rdf:datatype="http://www.w3.org/2001/XMLSchema#int"
  >0</dprop_numero_argumento>
</CLASSE_ADVERBIO>
<CLASSE_FORMA_DERIVADA rdf:ID="FD_UMA">
  <oprop_e_forma_derivada_de rdf:resource="#FC_DET_um"/>
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >uma</dprop_nome>
</CLASSE_FORMA_DERIVADA>
<CLASSE_PREPOSICAO rdf:ID="PREP_com">
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >com</dprop_nome>
</CLASSE_PREPOSICAO>
<CLASSE_PREPOSICAO rdf:ID="PREP_desde">
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >desde</dprop_nome>
</CLASSE_PREPOSICAO>
<CLASSE_PREPOSICAO rdf:ID="PREP_sobre">
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >sobre</dprop_nome>
</CLASSE_PREPOSICAO>
<CLASSE_CASO rdf:ID="CASO_Temporal">

```

```

<dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Tp</dprop_abreviatura>
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Temporal</dprop_nome>
</CLASSE_CASO>
<CLASSE_PRONOME rdf:ID="FC_DET_isto">
<dprop_descricao rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Objeto próximo do falante.</dprop_descricao>
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Isto</dprop_nome>
</CLASSE_PRONOME>
<CLASSE_CATEGORIA_SINTATICA rdf:ID="SINT_FV0">
<oprop_e_um rdf:resource="#SINT_FV"/>
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Functor Verbal Aridade 0</dprop_nome>
<dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>FV0</dprop_abreviatura>
</CLASSE_CATEGORIA_SINTATICA>
<CLASSE_TRACO rdf:ID="TRACO_ALimento">
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Alimento</dprop_nome>
</CLASSE_TRACO>
<CLASSE_ADJETIVO rdf:ID="FC_ADJ_Belo">
<dprop_descricao rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>que tem forma perfeita e proporções harmonicas</dprop_descricao>
<oprop_tem_cat_sint_arg1 rdf:resource="#SINT_FADJ"/>
<dprop_numero_argumento rdf:datatype="http://www.w3.org/2001/XMLSchema#int"
>1</dprop_numero_argumento>
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Belo</dprop_nome>
<oprop_tem_cat_sint_arg1 rdf:resource="#SINT_FN"/>
</CLASSE_ADJETIVO>
<CLASSE_PREPOSICAO rdf:ID="PREP_perante">
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>perante</dprop_nome>
</CLASSE_PREPOSICAO>
<CLASSE_CATEGORIA_MORFOLOGICA rdf:ID="N_CAT_Nome">
<dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>N</dprop_abreviatura>
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Nome</dprop_nome>
</CLASSE_CATEGORIA_MORFOLOGICA>
<CLASSE_CASO rdf:ID="CASO_Locativo">
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Locativo</dprop_nome>

```

```

<dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>L</dprop_abreviatura>
</CLASSE_CASO>
<CLASSE_PREPOSICAO rdf:ID="PREP_em">
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >em</dprop_nome>
</CLASSE_PREPOSICAO>
<CLASSE_PREPOSICAO rdf:ID="PREP_a">
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >a</dprop_nome>
</CLASSE_PREPOSICAO>
<CLASSE_NOME rdf:ID="ontoval_Individual_4">
  <oprop_tem_traco_semantico rdf:resource="#TRACO_Contavel"/>
  <oprop_tem_traco_semantico rdf:resource="#TRACO_Concreto"/>
  <dprop_descricao rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Instrumento óptico destinado a aumentar a imagem</dprop_descricao>
  <oprop_tem_traco_semantico rdf:resource="#TRACO_Solido"/>
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Telescópio</dprop_nome>
</CLASSE_NOME>
<CLASSE_NUMERAL rdf:ID="ontoval_dissertacao_Instance_4">
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Trezentos</dprop_nome>
</CLASSE_NUMERAL>
<CLASSE_PREPOSICAO rdf:ID="PREP_apos">
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >após</dprop_nome>
</CLASSE_PREPOSICAO>
<CLASSE_VERBO rdf:ID="ontoval_dissertacao_Instance_6">
  <dprop_numero_argumento rdf:datatype="http://www.w3.org/2001/XMLSchema#int"
  >0</dprop_numero_argumento>
  <dprop_descricao rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Cair água em gotas da atmosfera.</dprop_descricao>
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Chover</dprop_nome>
</CLASSE_VERBO>
<CLASSE_CATEGORIA_MORFOLOGICA rdf:ID="M_CAT_Adverbio">
  <dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >ADV</dprop_abreviatura>
  <dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Advérbio</dprop_nome>
</CLASSE_CATEGORIA_MORFOLOGICA>
<CLASSE_CATEGORIA_SINTATICA rdf:ID="SINT_FN0">
  <dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >FN0</dprop_abreviatura>

```

```

<oprop_e_um rdf:resource="#SINT_FN"/>
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Functor Nominal Aridade 0</dprop_nome>
</CLASSE_CATEGORIA_SINTATICA>
<CLASSE_CATEGORIA_SINTATICA rdf:ID="SINT_FADV">
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Functor Adverbial</dprop_nome>
<dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>FADV</dprop_abreviatura>
</CLASSE_CATEGORIA_SINTATICA>
<CLASSE_CATEGORIA_SINTATICA rdf:ID="SINT_FADV1">
<dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>FADV1</dprop_abreviatura>
<oprop_e_um rdf:resource="#SINT_FADV"/>
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Functor Adverbial Aridade 1</dprop_nome>
</CLASSE_CATEGORIA_SINTATICA>
<CLASSE_TRACO rdf:ID="TRACO_Nao_Contavel">
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Não Contável</dprop_nome>
</CLASSE_TRACO>
<CLASSE_PREPOSICAO rdf:ID="PREP_ate">
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>até</dprop_nome>
</CLASSE_PREPOSICAO>
<CLASSE_CATEGORIA_MORFOLOGICA rdf:ID="M_CAT_Adjetivo">
<dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>ADJ</dprop_abreviatura>
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Adjetivo</dprop_nome>
</CLASSE_CATEGORIA_MORFOLOGICA>
<CLASSE_CATEGORIA_SINTATICA rdf:ID="SINT_FN1">
<dprop_abreviatura rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>FN1</dprop_abreviatura>
<dprop_nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Functor Nominal Aridade 1</dprop_nome>
<oprop_e_um rdf:resource="#SINT_FN"/>
</CLASSE_CATEGORIA_SINTATICA>
</rdf:RDF>

```

<!-- Created with Protege (with OWL Plugin 1.3, Build 225.4) <http://protege.stanford.edu> -->