

UNIVERSIDADE PRESBITERIANA MACKENZIE

CELSO VETTORAZZO

**Desafios para a Construção de uma Ferramenta de Apoio à
Redação e Estilo**

São Paulo

2009

CELSO VETTORAZZO

DESAFIOS PARA A CONSTRUÇÃO DE UMA FERRAMENTA DE APOIO À
REDAÇÃO E ESTILO

Monografia apresentada ao Programa de Pós-Graduação em Tecnologia da Universidade Presbiteriana Mackenzie, como requisito para a obtenção do título de Especialista em Análise de Sistemas.

ORIENTADORA: Prof^a.Dr^a.Elida Jacomini Nunes

São Paulo
2009

V592p Vettorazzo, Celso.

Desafios para a construção de uma ferramenta de apoio à redação e estilo / Celso Vettorazzo – 2009.

90 f. : il. ; 30 cm.

Trabalho de Conclusão de Curso (Especialização em Análise de Sistemas) – Universidade Presbiteriana Mackenzie, São Paulo, 2009.

Referências Bibliográficas: f. 90.

1. Verificadores gramaticais. 2. Redação. 3. Estilo. 4. Colocações de palavras. 5. Orientação a objetos. I. Título.

CDD 003.3

AGRADECIMENTOS

Ao Banco Itaú SA, cuja política de incentivo à formação acadêmica de seus funcionários contribuiu decisivamente para viabilizar a minha graduação e pós nesta Universidade.

RESUMO

A principal finalidade deste estudo é explorar as necessidades e conceituar uma aplicação virtual com recursos que permitam incorporar manuais de redação e estilo automáticos configurados no editor de textos do consulente, o que virá a contribuir no desenvolvimento das suas habilidades de escrita e adicionar bônus na sua eficiência. Fazem parte do argumento desta monografia: análise de um dos verificadores ortográficos já existentes, pesquisa a respeito do material complementar disponível sobre assuntos relacionados e proposta de extensão destes recursos por meio da implementação de conceitos de colocabilidade inspirados no Oxford *Collocations Dictionary* (OCD). Para atender a este objetivo, este trabalho faz algumas considerações sobre o processo de desenvolvimento de um produto de Processamento de Linguagem Natural que incorpore uma solução baseada no modelo de Orientação a Objetos. Este modelo foi escolhido por ser suficientemente flexível para permitir análise em camadas, e suportar soluções para algumas das situações apresentadas no Manual de Redação e Estilo (MRE) do Estado de São Paulo de Eduardo Martins (*in memoriam*). A opção por esta obra como referência justifica-se pela sua ampla aceitação entre os especialistas, a popularidade que angariou com o seu formato “guia prático”, além da preocupação constante do autor com a precisão e a objetividade no uso da palavra. Esta preocupação pode ser observada tanto no cuidado com que ele recomenda utilizações das palavras dentro dos contextos mais adequados, quanto no sentido mais próximo possível da intenção que se pretendeu expressar. O autor justifica suas escolhas por esta ou aquela construção, lastreado num sólido conhecimento da nossa gramática. Ele provê, ainda, uma rica variabilidade de exemplos colhidos durante a sua vivência no exercício da profissão, o que proporciona material para a construção de Corpora estáveis, abrangentes e robustos. O estudo de um caso real sobre o desenvolvimento e colocação no mercado de um verificador ortográfico, levou a algumas considerações importantes a respeito da performance, portabilidade, limitações tecnológicas e estratégias para atrair eventuais investidores. Trata-se do projeto **ReGra**, que desenvolveu ferramentas de revisão de texto para o nosso idioma incorporados ao pacote Microsoft Office. Este projeto foi resultado de uma parceria bem sucedida entre a Itautec Philco S.A. com diversos centros de pesquisas de universidades. Além de permitir o desenvolvimento de um revisor gramatical para o português, este projeto proporcionou ainda, como benefício indireto, uma série de estudos que culminaram na produção de recursos lingüísticos de apoio, e em projetos completamente independentes de um revisor gramatical. No formato do trabalho, a fim de facilitar a leitura por não-especialistas, haverá uma nítida distinção entre as partes meramente informativas e as considerações de natureza mais teóricas. Ou seja; farão parte do escopo deste estudo tanto as estratégias para atrair financiadores e levantamentos de perfis de eventuais consulentes – orientando o seu desenvolvimento a algum nicho de mercado para o qual o uso da linguagem seja essencial –, quanto detalhes físicos de implementação desta aplicação e o estudo sobre materiais complementares – o que dará a esta solução um direcionamento mais concreto.

Palavras-chave: verificadores gramaticais; redação e estilo; colocações de palavras; orientação a objetos.

ABSTRACT

*The main purpose of this study is to explore the needs and elaborate a virtual application with resources that permit the incorporation of automatic writing and style manuals configured in text editors. This will contribute towards the development of the user's writing skills and further increase efficiency. The argument on behalf of this monograph includes: analysis by one of the already existing spell checkers, research on complementary material available about related subjects and the proposal to extend these resources by implementing collocation concepts inspired by the Oxford Collocations Dictionary (OCD). In order to meet this objective, this study makes a few considerations about the development process for a Natural Language Processing product that incorporates a solution based on Object-Oriented Programming. This model was chosen because it is sufficiently flexible to permit analysis in layers and to support solutions for some situations presented in the Manual de Redação e Estilo (MRE) do Estado de São Paulo (Writing and Style Manual of the Estado de São Paulo newspaper) by Eduardo Martins (in memoriam). This work was chosen as a reference due to its broad acceptance among specialists, the popularity it secured with its "practical guide" format, as well as the author's constant concern with precision and objectivity in the use of words. This concern can be observed in the care with which he recommends using words within the most appropriate contexts as well as in the closest possible sense in terms of what was intended to be expressed. The author justifies his choices for this or that construction, founded on solid knowledge of our grammar. He also provides a rich variety of examples gathered throughout his professional experience, resulting in material for the construction of stable, comprehensive and robust Corpora. A real case study about developing and placing a spell checker in the market led to some important considerations regarding performance, portability, technological limitations and strategies to attract eventual investors. It is the **ReGra** project, which developed text revision tools for our language incorporated in the Microsoft Office package. This project was the result of a successful partnership between Itautec Philco S.A. and several university research centers. Besides permitting development of a grammar reviser for Portuguese, this project also led to a series of studies that culminated in the production of support linguistic resources, as an indirect benefit, and in completely independent projects for a grammar reviser. In order to facilitate reading by non-specialists, there will be a clear distinction between the merely informative parts and the more theoretical considerations. In other words, strategies to attract investors and studies of possible consultant profiles will be part of this study – guiding its development to a market niche for which language use is essential. It will also include physical details for implementing the application and a study of complementary materials, which will provide this solution with a more solid direction.*

Key words: grammar checkers; writing and style; word collocation; object-oriented programming.

SUMÁRIO

I.	INTRODUÇÃO.....	7
	1. <i>Workflow</i> esboçando aplicação da solução na construção do modelo de classe.....	8
	2. Premissas.....	9
	3. Estratégia para mitigar resistências.....	12
II.	REVISORES DE TEXTO PARA O PORTUGUÊS.....	15
	1. Análise do desempenho do ReGra.....	18
	2. A Questão das Locuções.....	23
III.	DICIONÁRIO DE COLOCAÇÕES DAS CLASSES DE PALAVRAS.....	30
	1. Grupos de Colocações.....	31
	2. Incorporação de um Tesouro à solução.....	31
	3. Termos que exigem complemento.....	32
	4. Regras sintáticas para uma sentença.....	33
	5. Omissões do OCD.....	34
IV.	APLICAÇÃO DO MODELO DE OO PARA A ANÁLISE SEMÂNTICA.....	38
	1. Herança.....	39
V.	CONCLUSÃO.....	41
APÊNDICE 1.	EXEMPLOS DE APLICAÇÃO DO WORKFLOW.....	49
APÊNDICE 2.	GLOSSÁRIO.....	80
	REFERÊNCIAS BIBLIOGRÁFICAS.....	90

I. INTRODUÇÃO

Sob a pressão da necessidade de aumentar a produtividade, as formas tradicionais de escrita evoluíram dos manuscritos primitivos para a invenção da imprensa com tipos móveis, datilografia, linotipos, fotolitos, edição e diagramação eletrônica de textos; até chegar, nos dias de hoje, aos revisores ortográficos e gramaticais. A evolução natural deste ciclo aponta para aperfeiçoamentos nos revisores gramaticais existentes, de maneira que eles passem a incorporar recursos automáticos dos materiais de apoio à redação e estilo.

Para que esta proposta se realize numa solução concreta, o passo inicial teria que ser dado na direção de mapear – em ordem decrescente de abrangência – toda a diversidade e complexidade das variáveis envolvidas nesta abordagem; partindo das mais amplas, até atingir metas mais restritas, pragmáticas e, portanto, realistas; conforme representação na Figura 1.

1. *Workflow* esboçando aplicação da solução na construção do modelo de classe

Esta proposta apresenta um esboço do *WorkFlow* para atingir este objetivo, utilizando como exemplos alguns dos Conceitos colhidos no MRE e apresentando uma solução na forma de uma Modelo de Classes aplicável, por meio da qual uma versão traduzida para o Português do Brasil do OCD – o OCD-PT-BR – seria compilada como um Catálogo de Colocações. Assim, na sua primeira versão, uma base de conhecimento neste formato exigiria um esforço prévio de tradução da obra original em Inglês que lhe serviu de inspiração.

Neste estudo, foi composta uma amostra com verbetes escolhidos apenas dentre aqueles que aparentavam ser de mais fácil aplicação do modelo proposto. Assim,

esses exemplos funcionam meramente como Casos de Uso ilustrativos da estratégia

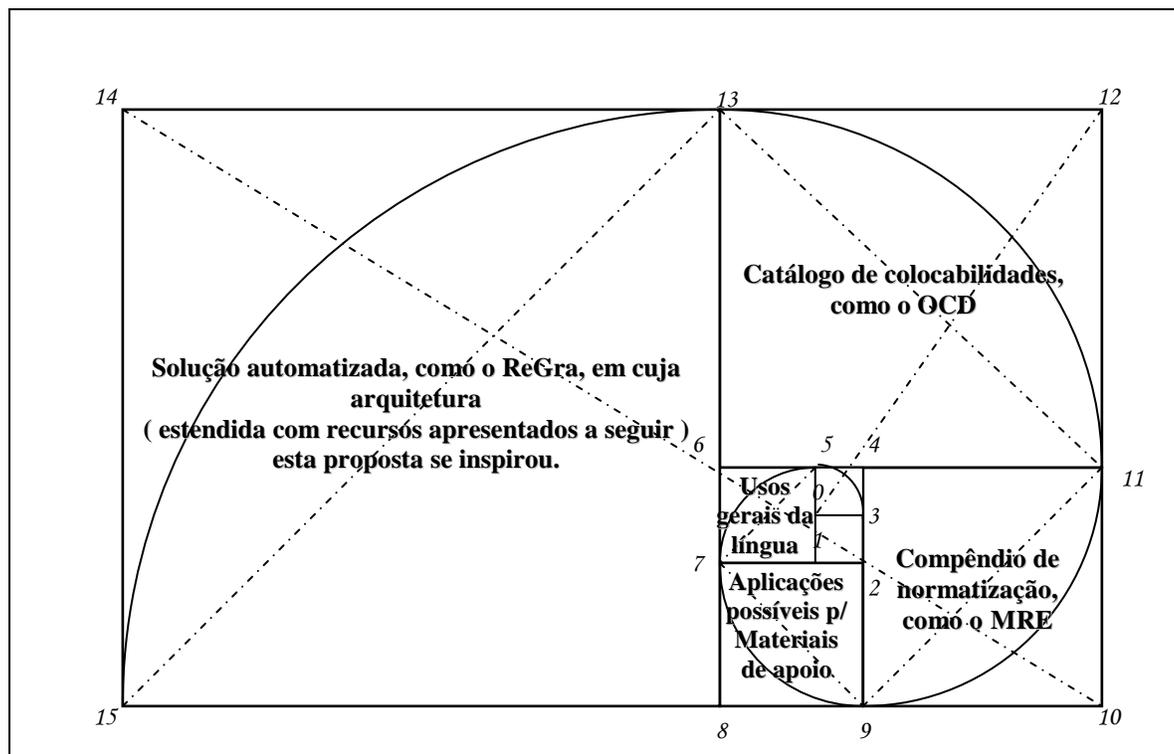


Figura 1 – Perspectiva artística representando a linha da evolução da solução

de funcionamento proposta nesta abordagem. A aplicação prática do *Workflow*, representado na figura 2, está descrita no **Apêndice 1 – Exemplos de aplicação do *Workflow*** composto por exemplos colhidos no MRE^① que foram resolvidos a partir do registro do termo no Catálogo de Colocações traduzido do OCD para o Português do Brasil^②, compilado no formato de Modelo de Classes no paradigma de Orientação a Objetos^③.

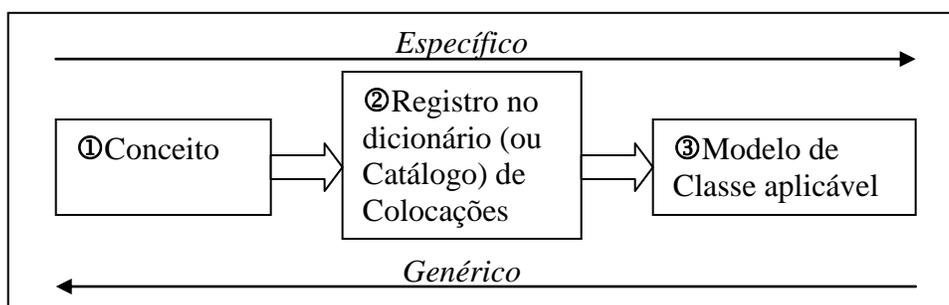


Figura 2 - Workflow esboçando a aplicação da solução até a construção modelo de classe.

OBSERVAÇÃO: Mesmo compondo-se esta amostra por exemplos em que a aplicação desse modelo transcorreu de forma tão direta quanto possível, a experiência com alguns casos práticos de aplicação deste *Workflow* trouxe algumas considerações adicionais que, por constituírem parte fundamental para a proposta deste trabalho, não podem ser relevadas. Estas considerações são apresentadas no decorrer deste trabalho; contendo comentários sobre a necessidade de ajustes mais finos e – prevendo aperfeiçoamentos posteriores – detalhes a respeito das situações observadas para as quais a solução não dá suporte.

2. Premissas

No esboço de um cronograma representando a linha da evolução da solução, a primeira etapa teria que ser, conforme as considerações anteriores, a carga inicial de uma base de conhecimento que a suporte. Em paralelo, as interfaces da solução poderiam ser prototipadas juntamente com o planejamento da colocação da ferramenta no mercado. Em seguida viriam: o projeto da arquitetura, o desenvolvimento da solução, e finalmente o lançamento no uma versão beta. Opcionalmente, o formato inicial da solução poderia começar pela automação parcial, evoluindo até realizar plenamente a sua automação.

OBSERVAÇÃO: Para esta solução realizar-se completamente, neste ponto caberia um detalhamento maior de cada uma destas etapas abrangendo todas as fases do seu projeto de implantação. Justifica-se esta omissão pelos limites do escopo deste trabalho, cujo aspecto principal é o estudo exploratório, objetivando adquirir conhecimentos sobre os principais desafios a serem confrontados para a realização de um projeto com estas características.

O formato das interfaces teria que levar em conta as implicações apresentadas a seguir:

O seu manejo deverá ser aberto a fim de dinamizar a flexibilização das regras, preferindo antes sugeri-las que impô-las. A opção por este formato prevê também a possibilidade do consulente colaborar com sugestões, pois ninguém melhor que ele para auxiliar na construção, propor melhorias e incrementar a base de conhecimento lingüístico original. A solução deverá permitir ao usuário, ainda, customizar o produto instalado no seu *desktop* por meio de alterações nas diretivas de um *PROFILE*. Posteriormente, a leitura e análise destas particularizações viriam também a contribuir no aperfeiçoamento do produto final;

A interface, nas suas intervenções, teria que estar preparada para analisar e fazer sugestões nos diversos níveis de criticidade, como: *incorreta, imprópria, inadequada e há-forma-preferível-a-esta*, seguida de exemplos práticos;

Aproximar-se tanto quanto possível da organização gráfica do OCD, em um formato padronizado e amigável para quem se dispuser a efetuar consultas e modificações, conforme a estrutura simplificada dos registros apresentada na Figura 3, utilizando os mesmos termos da obra.

Estruturas possíveis no OCD	<i>noun</i>	<i>verb</i>	PHRASAL VERBS	<i>result in algo refer to sb/algo</i>	adj.
	'n' significado 'n'	'n' significado 'n'			'n' significado 'n'
					<u>VERBOS</u>
	<u>ADJ.</u>				<u>ADV.</u>
		<u>ADV.</u>			
	<u>QUANT.</u>				
	<u>VERB + termo</u>	<u>VERB + termo</u>			
	<u>termo + VERB</u>				
	<u>termo + NOUN</u>				
	<u>PREP.</u>	<u>PREP.</u>			<u>PREP.</u>
	<u>FRASES</u>	<u>FRASES</u>			
NOTE					

Figura 3 - organização gráfica do OCD

Nas intervenções da interface, a função das FRASES da estrutura acima auxiliaria no entendimento das colocações por meio de exemplos práticos.

A arquitetura deverá prever expansões futuras para:

Recursos do OCD ainda não implementados neste modelo, conf. será visto mais adiante neste trabalho nas considerações sobre as experiências adquiridas na aplicação da solução;

Suportar um formato colaborativo da solução utilizando arquitetura WEB 2.0, em que o produto em si não ficaria residente no desktop do consulente (exceto o *PROFILE*), mas centralizado em servidores próprios. Isto proporcionaria o agendamento de *uploads* periódicos das particularizações feitas pelo usuário durante o uso, mas que não foram formalizadas, para análise e aperfeiçoamento do produto final;

Incorporar um Módulo Administrativo, prevendo a dinamização da intervenção de lexicógrafos qualificados atuando sobre as sugestões de aperfeiçoamentos feitas pelos usuários antes de incorporá-las à base centralizada, sem que haja a necessidade de incrementos na versão do produto a cada pacote de melhorias implementado;

A implementação física da solução deve ser capaz de compilar este formato de interface (semelhante a um verbete num dicionário) em um modelo de Orientação por Objetos; de maneira que os parâmetros fornecidos derivem naturalmente para um diagrama de classes. A opção pelo modelo de Orientação a Objetos proporcionaria recursos para estabelecer: Heranças, Vínculos com outras classes de palavras (a função *NOTE* da estrutura apresentada na Figura 3), Especializações, Métodos representando comportamentos, etc. Ao mesmo tempo, este modelo proporcionaria a flexibilidade necessária para inclusões dinâmicas de novas colocações, bem como o estabelecimento de novos vínculos entre os termos. Como um subproduto independente, o diagrama de classes poderia ser ofertado para outros usos em plataforma aberta, podendo ser incorporado diretamente aos mais diversos tipos de soluções e formatos.

3. Estratégia para mitigar resistências

Uma boa estratégia para mitigar resistências por parte dos eventuais consulentes seria fundamental para o sucesso do produto. Uma ferramenta desta natureza provocaria um incômodo natural por conta de aspectos subjetivos, como a vaidade dos consulentes, por exemplo, indo de encontro à presunção da tutela imposta. Raros seriam os usuários que se não se julgariam proficientes a ponto de prescindir de uma ajuda deste tipo.

Uma argumentação convincente deveria expor aspectos do dinamismo das línguas, como bem disse REGULA (1957): "A língua, expressão consciente de impressões exteriores e interiores, está sujeita a uma perpétua transformação. Os vocábulos mudam de sentido ou porque as coisas se modificam ou porque a 'constelação psíquica' sob cuja influência nasce o sentido do objeto, se altera graças a causas diversas" (REGULA¹ apud BECHARA, 1975, p. 340). De fato, a vigência a partir de 2009 do Acordo Ortográfico da Língua Portuguesa confirma esta necessidade, pois prevê, entre outras mudanças: a inclusão das letras "w", "k" e "y" no alfabeto, diversas alterações na escrita das palavras, o fim do trema, a supressão de algumas consoantes mudas e vários acentos, além de simplificar o uso do hífen. Um material de apoio, parcial ou totalmente automatizado, ajudaria a reduzir o período de adaptação a estas novas regras; uma vez que o decreto assinado que normatizou a adesão do Brasil ao Acordo já impôs um cronograma de implantação.

BECHARA (1975) estabelece uma clara distinção entre **traço estilístico** e **erro gramatical**: "(...) – Não se há de entender que o estilo seja sempre uma deformação da norma lingüística. Isto nos leva à distinção entre *traço estilístico* e *erro gramatical*. O traço estilístico pode ser um desvio ocasional de norma gramatical vigente, mas se impõe pela sua intenção estética." (BECHARA, 1975, p. 350). Seria conveniente ao consulente dispor de um recurso automático para apontar correções indevidas por parte de revisões excessivamente formais, razão pela qual estes revisores relevariam aspectos mais polêmicos de interpretação vernácula. Apenas para

¹ Moritz Regula na sua *Grammaire Française Explicative*, 1957, 26.

ilustrar, em outro trecho desta mesma obra BECHARA (1975) discorre sobre a questão da colocação dos pronomes átonos: “**Colocação e clareza** Levadas em conta as construções fundamentais de que a linguagem natural e espontânea não costuma afastar-se, é certo que para a estrutura oracional temos em português bastante liberdade. (...) Colocação dos pronomes pessoais átonos e do demonstrativo o é questão de fonética sintática – Durante muito tempo viu-se o problema apenas pelo aspecto sintático, criando-se a falsa teoria da ‘atração’ vocabular do *não*, do *quê*, de certas conjunções e tantos outros vocábulos. Graças a notáveis pesquisadores, e principalmente Said Ali, passou-se a considerar o assunto pelo aspecto fonético^①. Abriram-se com isso os horizontes, estudou-se a questão dos vocábulos átonos e tônicos, e chegou-se à conclusão de que muitas das regras estabelecidas pelos puristas ou estavam erradas, ou se aplicavam em especial atenção ao falar lusitano^②. A Gramática, alicerçada na tradição literária^③, ainda não se dispôs a fazer concessões a algumas tendências do falar de brasileiros cultos^④, e não leva em conta as possibilidades estilísticas que os escritores conseguem extrair da colocação de pronomes átonos.” (BECHARA, 1975, p. 325 e 326).

Apenas neste trecho transcrito da obra, o autor cita quatro correntes conflitantes:

- ① pelo aspecto fonético;
- ② puristas, com atenção especial ao falar lusitano;
- ③ alicerçada na tradição literária;
- ④ tendências do falar de brasileiros cultos;

Destas, o autor manifesta a sua opção pessoal pela ^④, com as seguintes ressalvas “Daremos aqui apenas aquelas normas que, sem exagero, são observadas na linguagem escrita e falada das pessoas cultas. Não se infringindo os critérios expostos, o problema é questão pessoal de escolha, atendendo-se às exigências da eufonia.” (BECHARA, 1975, p. 326).

Utilizar o material de apoio como uma espécie de “mediador isento de paixões” neste tipo de debate, adicionaria outro bom argumento na estratégia de convencimento dos usuários; pois ele passaria a ter acesso a uma contra-argumentação consistente

para defesa das particularidades do estilo que ele pretende fixar como a sua marca pessoal.

Porém, nada motivaria mais um consulente a persistir no uso do material de apoio do que incentivar a sua participação ativa no processo de aperfeiçoamento contínuo do produto. Creditando-lhe a autoria sobre sugestões que vierem a ser incorporadas à solução, o status adquirido diante da comunidade composta pelos seus consulentes/colaboradores dentro de um ambiente participativo, viria a incentivar uma saudável rivalidade entre eles. Haveria, inclusive, a possibilidade de um escritor consagrado enriquecer este legado ao incorporar sutilezas características do seu estilo particular neste repositório. Isto faria este material de apoio funcionar também como um formato alternativo ao formalismo das oficinas de escrita tradicionais. Neste aspecto, o produto viria a contribuir também como um poderoso recurso pedagógico para a população em geral. Este seria outro argumento importante de apelo à responsabilidade social (o que ajudaria na colocação do produto no mercado), pois; conforme IBOPE (2004) “Apenas 25% da população adulta dominam (sic) habilidades de escrita e leitura”. Por outro lado, esta exigência levaria a estratificar as versões do produto para atender aos diferentes perfis dos consulentes, diferenciando os especialistas dos usuários comuns; em que somente os primeiros teriam os privilégios mencionados de contribuir com aperfeiçoamentos no produto, que viriam a beneficiar diretamente também os segundos.

II. REVISORES DE TEXTO PARA O PORTUGUÊS

As Ferramentas de revisão de texto do Microsoft Office — foco principal deste estudo — são um pacote suplementar que contém as ferramentas de revisão de texto da Microsoft para vários idiomas, fontes, verificador ortográfico e gramatical, listas de Auto-Correção, regras de Auto-Resumo (Microsoft Word somente), dicionários e, ferramentas voltadas para os idiomas asiáticos. A criação destas ferramentas no nosso idioma, ficou a cargo do projeto ReGra; uma parceria bem sucedida entre a Itautec Philco SA e alguns centros de pesquisas de universidades. Além de permitir o desenvolvimento de um revisor gramatical para o português, cujo desempenho é o melhor do mercado; propiciou ainda, como benefício indireto, uma série de estudos que culminaram na produção de recursos lingüísticos de apoio e em projetos completamente independentes de um revisor gramatical.

O *parser* desenvolvido neste projeto foi baseado na frase como estrutura gramatical, compreendendo regras heurísticas sensíveis ao contexto para detecção de erros comuns. Foi necessário um relaxamento destas regras a fim de reconhecer erros a partir de determinados padrões estabelecidos para inferir sentenças gramaticalmente impróprias.

Um processo de revisão humana atua sobre cinco níveis básicos da estrutura de um texto: léxico, sintático, semântico, pragmático e análise de discurso. Entretanto, diante das limitações tecnológicas, o **ReGra** atua somente sobre alguns destes níveis; sendo constituído por três módulos principais: o módulo mecânico, o verificador ortográfico e o módulo, cuja arquitetura será o propósito desta leitura; o módulo lingüístico; representado na Figura 4.

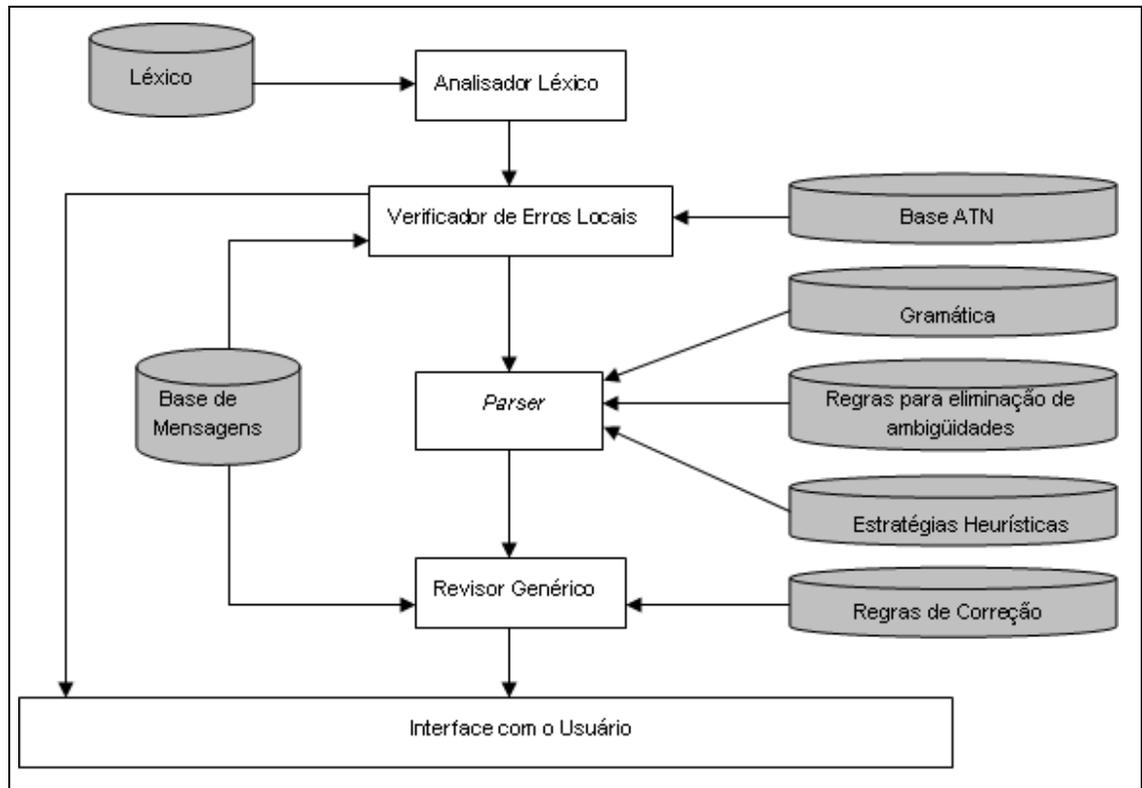


Figura 4: Arquitetura do Módulo Lingüístico do ReGra

Fonte: MARTINS et al (1998, p.8)

Léxico / Analisador Léxico Um léxico contém cerca de 1,5 milhões de palavras, contendo inflexões (por exemplo: gênero, tempo verbal), derivações, nomes próprios, acrônimos largamente utilizados e abreviaturas. Foi desenvolvido para dar suporte a revisões ortográficas e gramaticais de uma forma geral, mas não abrange o domínio da categorização semântica.

Verificador de erros locais Cada unidade passa pelo revisor gramatical que aponta erros locais de acordo com um conjunto de padrões extraídos de mais de dois mil textos não revisados produzidos por escritores comuns, como: estudantes, secretárias, etc.

Base ATN As regras corretas neste sub-módulo foram implementadas no formato de ATNs (vide glossário).

Parser Categoriza sintaticamente as palavras

Gramática Analisa a unidade tentando reconhecer alguns padrões, tanto para construções adequadas, como para erros comuns obtidos numa gramática híbrida para o Português do Brasil. Estas construções gramaticalmente impróprias foram obtidas por meio de aprendizado adquirido em um grande corpus compilado especialmente para este projeto.

Regras para eliminação de ambigüidades Quando houver mais de uma categoria gramatical para uma palavra no dicionário, a mais freqüente dentre elas é a escolhida. Isto pressupõe que os itens léxicos estejam categorizados hierarquicamente, de acordo com a freqüência do seu uso. Caso estes itens tenham freqüências de uso semelhantes, regras de eliminação de ambigüidades baseadas na análise de palavras adjacentes auxiliam a inferir qual a sua classificação mais provável.

Estratégias Heurísticas Havendo categorizado todas as palavras, o *parser* utiliza uma tabela extensiva de possibilidades com o objetivo de inferir a estrutura sintática da sentença. Dificuldades nesta fase da análise são resolvidas por intermédio de estratégias heurísticas para busca e identificação dos padrões de erros mais comuns para o Português do Brasil.

Revisor Genérico / Regras de Correção Uma vez que os desvios elementares da norma tenham sido corrigidos, a próxima fase do processo tratará da correção dos erros que dependam da aplicação de regras sintáticas mais complexas, como aqueles relacionados à concordância e regência, por exemplo. A referência para as estruturas lingüísticas adequadas à norma culta são os gramáticos de maior prestígio.

Base de Mensagens / Interface com o usuário Diante das freqüentes divergências de opinião entre os gramáticos, as mensagens do **ReGra** são cuidadosamente redigidas no sentido de tentar evitar que alguma convenção adotada pelo revisor seja imposta ao usuário como se fosse uma regra gramatical definitiva; considerando que

em alguns casos é impossível inferir se uma construção lingüística particular está errada ou não. Ainda que a ferramenta não seja dirigida para o ensino de gramática, a mensagem, ao descrever a regra, provê informações adicionais sobre os erros detectados. (MARTINS et al, 1998).

1. Análise do desempenho do ReGra

Alguns limites do escopo desta solução:

Pressupor uma idéia conservadora – sempre estacionária – de linguagem que rejeita muitos itens léxicos que ainda não estão nos dicionários, como neologismos ou gírias;

Atuar de maneira segura somente sobre as dez classes gramaticais tradicionalmente admitidas: substantivo, adjetivo, advérbio, preposição, interjeição, conjunção, numeral, pronome, verbo, artigo e algumas subclassificações (nomes próprios, nomes comuns, etc.);

Desconsiderar construções sofisticadas ou idiossincráticas, pois o revisor gramatical é direcionado para o uso padrão do Português Brasileiro por parte de escritores não proficientes.

Eduardo Martins – autor de um dos materiais de apoio à escrita sobre o qual esta argumentação se fundamentou – convidado a opinar sobre o desempenho dos corretores ortográficos e gramaticais automáticos existentes fez as seguintes considerações: “os corretores ortográficos são excelentes instrumentos para evitar os erros de grafia, falta de acento nas palavras, para fugir de formas verbais erradas (p. ex., o corretor informa que o certo é ‘interveio’ e não ‘interview’). Acho todos recomendáveis, por mais incompletos que sejam. Quanto aos corretores gramaticais, existe um no Word 2000¹, que funciona em alguns casos e em outros, não. Acho que

¹ Trata-se do ReGra, conf. visto anteriormente, foco principal deste estudo.

estes ainda precisam ser aperfeiçoados.”¹; como ilustra a análise apresentada adiante, avaliando o desempenho deste corretor ortográfico diante dos “CEM ERROS MAIS COMUNS” e dos “DEZ ERROS MAIS GRAVES” apontados por MARTINS (1997, p. 321-328).

OS CEM ERROS MAIS COMUNS			
	Erro	Comportamento do ReGra	Natureza do erro
1	<u>Mal</u> cheiro, <u>mau</u> -humorado	Parcialmente detectado	Gramática
2	Fazem cinco anos	Corretamente detectado	Gramática
3	Houveram muitos acidentes	Corretamente detectado	Gramática
4	Existe muitas esperanças	Corretamente detectado	Gramática
5	Para mim fazer	Corretamente detectado	Gramática
6	Entre eu e você	Não detectado	Gramática
7	Há dez anos atrás	Não detectado	Retórica ou Semântica
8	Entrar dentro	Não detectado	Gramática
9	Venda à prazo	Corretamente detectado	Gramática
10	Porque você foi?	Não detectado	Retórica ou Semântica
11	Vai assistir o jogo hoje.	Corretamente detectado	Gramática
12	Preferia ir do que ficar.	Corretamente detectado	Gramática
13	O resultado do jogo, não o abateu.	Não detectado	Gramática
14	Não há regra sem <u>excessão</u>	Corretamente detectado	Ortografia
15	Quebrou o óculos	Corretamente detectado	Gramática
16	Comprei ele para você	Corretamente detectado	Gramática
17	Nunca lhe vi	Não detectado	Gramática
18	Aluga-se casas	Corretamente detectado	Gramática
19	Tratam-se de	Não detectado	Gramática
20	Chegou em São Paulo	Não detectado	Retórica ou Semântica
21	Atraso implicará em punição	Não detectado	Retórica ou Semântica
22	Vive às custas do pai.	Não detectado	Retórica ou Semântica
23	Todos somos <u>cidadões</u>	Detectado. Sugestão errada	Ortografia
24	O ingresso é <u>gratuito</u>	Corretamente detectado	Ortografia
25	A última seção de cinema	Não detectado	Retórica ou Semântica
26	Vendeu uma grama de ouro	Corretamente detectado	Gramática

¹ Trecho do bate-papo com o autor organizado pelo site **uni>ersia** disponível em: <<http://www.universia.com.br/materia/materia.jsp?id=1616>>. Publicado em 23/05/2003.

27	<u>Porisso</u>	Detectado. Sugestão errada	Ortografia
28	Não viu qualquer risco	Não detectado	Retórica ou Semântica
29	A feira inicia amanhã	Não detectado	Gramática
30	Soube que os homens feriram- se	Não detectado	Gramática
31	O peixe tem muito espinho	Não detectado	Retórica ou Semântica
32	Não sabiam aonde ele estava	Corretamente detectado	Gramática
33	Obrigado , disse a moça.	Não detectado	Retórica ou Semântica
34	O governo <u>entrevistou</u>	Não detectado	Ortografia
35	Ela era meia louca	Corretamente detectado	Gramática
36	Fica você comigo	Não detectado	Gramática
37	A questão não tem nada haver com você	Não detectado	Retórica ou Semântica
38	A corrida custa cinco real	Corretamente detectado	Gramática
39	Vou emprestar dele	Não detectado	Retórica ou Semântica
40	Foi taxado de ladrão	Não detectado	Retórica ou Semântica
41	Ele foi um dos que chegou antes	Não detectado	Gramática
42	Cerca de 18 pessoas o saudaram	Corretamente detectado	Retórica ou Semântica
43	Ministro nega que é negligente	Não detectado	Gramática
44	Tinha chego atrasado	Não detectado	Gramática
45	Tons pastéis predominam	Não detectado	Retórica ou Semântica
46	Lute pelo meio-ambiente	Não detectado	Ortografia
47	Querida namorar com o colega	Corretamente detectado	Gramática
48	O processo deu entrada junto ao STF	Não detectado	Retórica ou Semântica
49	As pessoas <u>esperavam-o</u>	Corretamente detectado	Ortografia
50	Vocês fariam-lhe um favor?	Corretamente detectado	Gramática
51	Chegou a duas horas e partirá daqui há cinco minutos	Corretamente detectado	Gramática
52	Blusa em seda	Não detectado	Retórica ou Semântica
53	A artista deu à luz a gêmeos.	Não detectado	Retórica ou Semântica
54	Estávamos em quatro à mesa	Não detectado	Retórica ou Semântica
55	Sentou na mesa para comer	Corretamente detectado	Gramática
56	Ficou contente por causa que ninguém se feriu	Corretamente detectado	Gramática
57	O time empatou em 2 a 2	Não detectado	Retórica ou Semântica

58	À medida em que a epidemia se espalhava	Corretamente detectado	Gramática
59	Não queria que <u>receiassem</u> a sua companhia	Corretamente detectado	Ortografia
60	Eles tem razão	Corretamente detectado	Gramática
61	A moça estava ali há muito tempo	Não detectado	Gramática
62	Não se o diz	Não detectado	Gramática
63	Acordos <u>políticos-partidários</u>	Detectado. Sem sugestão para correção	Ortografia
64	Fique <u>tranquilo</u>	Corretamente detectado	Ortografia
65	Andou por todo país	Não detectado	Retórica ou Semântica
66	Todos amigos o elogiavam	Não detectado	Gramática
67	Favoreceu ao time da casa	Não detectado	Gramática
68	Ela mesmo arrumou a sala.	Corretamente detectado	Gramática
69	Chamei-o e o mesmo não atendeu	Não detectado	Retórica ou Semântica
70	Vou sair essa noite	Não detectado	Retórica ou Semântica
71	A temperatura chegou a 0 graus	Corretamente detectado	Gramática
72	A promoção veio de encontro aos seus desejos	Não detectado	Retórica ou Semântica
73	Comeu frango ao invés de peixe	Não detectado	Retórica ou Semântica
74	Se eu ver você por aí	Corretamente detectado	Gramática
75	Ele <u>intermedia</u> a negociação	Detectado. Sugestão errada	Ortografia
76	Ninguém se <u>adequa</u>	Corretamente detectado	Ortografia
77	Evite que a bomba <u>expluda</u>	Detectado. Sugestão errada	Ortografia
78	Governo <u>reavê</u> confiança	Detectado. Sugestão errada	Ortografia
79	Disse o que <u>quiz</u>	Corretamente detectado	Ortografia
80	O homem <u>possue</u> muitos bens	Detectado. Sugestão errada	Ortografia
81	A tese onde	Não detectado	Retórica ou Semântica
82	Já foi comunicado da decisão	Não detectado	Retórica ou Semântica
83	Venha por a roupa	Não detectado	Ortografia
84	<u>Inflingiu</u> o regulamento	Corretamente detectado	Ortografia
85	A modelo pousou o dia todo	Não detectado	Retórica ou Semântica
86	Espero que viagem hoje	Não detectado	Ortografia
87	O pai sequer foi avisado	Não detectado	Retórica ou Semântica
88	Comprou uma TV a cores	Não detectado	Gramática
89	Causou -me estranheza as palavras	Corretamente detectado	Gramática
90	A realidade das pessoas podem mudar	Corretamente detectado	Gramática

91	O fato passou desapercebido	Não detectado	Retórica ou Semântica
92	Haja visto seu empenho	Corretamente detectado	Gramática
93	A moça que ele gosta	Não detectado	Retórica ou Semântica
94	É hora dele chegar	Não detectado	Retórica ou Semântica
95	Vou consigo	Não detectado	Retórica ou Semântica
96	Já é 8 horas	Corretamente detectado	Gramática
97	A festa começa às 8 <u>hrs</u>	Detectado. Sugestão errada	Ortografia
98	Dado os índices das pesquisas	Não detectado	Gramática
99	Ficou sobre a mira do assaltante	Não detectado	Retórica ou Semântica
100	Ao meu ver	Não detectado	Retórica ou Semântica
Índice de intervenções corretas			≈ 40,5%

OS DEZ ERROS MAIS GRAVES			
	Erro	Comportamento do ReGra	Natureza do erro
1	Quando estiver voltado da Europa	Não detectado	Gramática
2	Que <u>seje</u> feliz	Erro detectado. Sugestão errada	Ortografia
3	Ele é de menor	Não detectado	Gramática
4	A gente fomos embora	Corretamente detectado	Gramática
5	De formas que	Corretamente detectado	Gramática
6	Fiquei fora de si	Corretamente detectado	Gramática
7	Acredito de que	Não detectado	Gramática
8	Fale alto porque ele houve mal	Não detectado	Gramática
9	Ela veio, mais você, não.	Não detectado	Gramática
10	Fale sem <u>exitar</u>	Corretamente detectado	Ortografia
Índice de intervenções corretas			≈ 45%

OBSERVAÇÃO: Por exigir competências lexicográficas que vão além da formação exigida para um trabalho com esta proposta, optou-se por não discernir as diferenças sutis entre os erros de natureza retórica ou semântica nesta análise (vide Glossário).

Buscando a fidelidade desta análise com o cenário sobre o qual se fundamentou o comentário citado anteriormente, utilizou-se a mesma versão do corretor gramatical do *Word* existente no pacote do *Office* 2000, ignorando-se os eventuais aperfeiçoamentos das versões posteriores desta ferramenta. Originalmente, entretanto, regras que escapam da alçada estritamente gramatical estão mesmo fora do escopo do **ReGra**, como reconhecem os responsáveis pelo projeto numa análise própria: “*The main reason for this high rate of omission is that a number of grammatical mistakes may be semantic in nature and/or depend on the context, which is clearly outside ReGra’s scope.*” (MARTINS et al, 1998, p. 25). Isto leva a crer que este índice pouco mudaria, caso esta análise utilizasse alguma das suas versões mais recentes. Também em razão do seu enfoque generalista, o desempenho do **ReGra** talvez fosse melhor com um corpus não tão especializado em termos jornalísticos.

2. A Questão das Locuções

Ao pretender alargar a usabilidade desta proposta enriquecendo o seu conteúdo morfológico e semântico; impõe-se como um dos principais desafios, pela sua importância, a questão das locuções. Qualquer solução para funcionar adequadamente deve reconhecer locuções em grupos de palavras: suas regras de variação; o fato de que as palavras que a compõem podem aparecer juntas na frase ou separadas por outras classes de termos; e analisá-la numa primeira camada, antes de qualquer outra.

Para ilustrar esta necessidade como mandatória para nortear o desenvolvimento do projeto já nas suas pesquisas exploratórias, a figura 4, que representa a arquitetura do módulo lingüístico do **ReGra**, foi aqui redesenhada na figura 5 como uma proposta de aperfeiçoamento, destacando a importância de incorporar esta subfunção no seu analisador léxico, como uma extensão dos recursos já existentes que resolvem a verificação de Inflexões (gênero, tempo verbal, etc.), pronomes átonos (enclíticos ou mesoclíticos), regências e regras de concordância.

Em uma avaliação superficial da obra de BECHARA (1975), destacam-se cinco grupos principais de locuções: **verbais**, **adverbiais**, **conjuntivas**, **interjetivas** e **prepositivas**; sendo que para estas últimas, o autor adverte para algumas situações em que o acúmulo de preposições pode ser confundido com uma locução prepositiva. Também, o autor faz algumas considerações sobre o uso impróprio do hífen unindo seus elementos e cita algumas unidades fonéticas originadas a partir delas:

Locução verbal “verbos auxiliares. Chama-se locução verbal a combinação das diversas formas de um verbo auxiliar com o infinitivo, gerúndio ou particípio de outro verbo que se chama principal: *hei de estudar, estou estudando, tenho estudado. (...) haveremos de fazer, estavam por sair, iam trabalhando, tinham visto. (...) Era chegada a ocasião da fuga. São passados três meses. (...)é amado, está prejudicado, ficaram rodeados(...). começar a escrever, pôr-se a escrever, (...)estar para escrever(...), estar a escrever, andar escrevendo, vir escrevendo, ir escrevendo(...), tornar a escrever, costumar escrever(...), acabar de escrever, cessar de escrever, deixar de escrever, parar de escrever(...), haver de escrever, ter de escrever, dever escrever, precisar(de) escrever(...), tenho que estudar(...), poder escrever(...), querer escrever, desejar escrever, odiar escrever, abominar escrever(...), buscar escrever, pretender escrever, tentar escrever, ousar escrever, atrever-se a escrever(...), conseguir escrever, lograr escrever(...), parecer escrever(...), ir escrever(...), vir a escrever, chegar a escrever(...).*

(...) *Vir a + infinitivo* de certos verbos tem quase o mesmo sentido do verbo principal empregado sozinho: *Isto vem a traduzir a mesma idéia (=isto por fim traduz a mesma idéia). Vir a ser* pode ainda ser sinônimo de *tornar-se*: *Ele veio a ser famoso.*” (BECHARA, 1975, p. 110).

Também têm o mesmo sentido: *levar em consideração* e *considerar*.

“NOTA FINAL – Nem sempre a aproximação de dois ou mais verbos constitui uma locução verbal; a intenção da pessoa que fala ou escreve é que determinará a existência da locução. ‘Por exemplo, na frase: *queríamos colher rosas*, os verbos *queríamos colher* constituirão expressão verbal se pretendo dizer que *queríamos colher rosas e não outra flor*, sendo *rosas* o objeto da declaração. Se, porém,

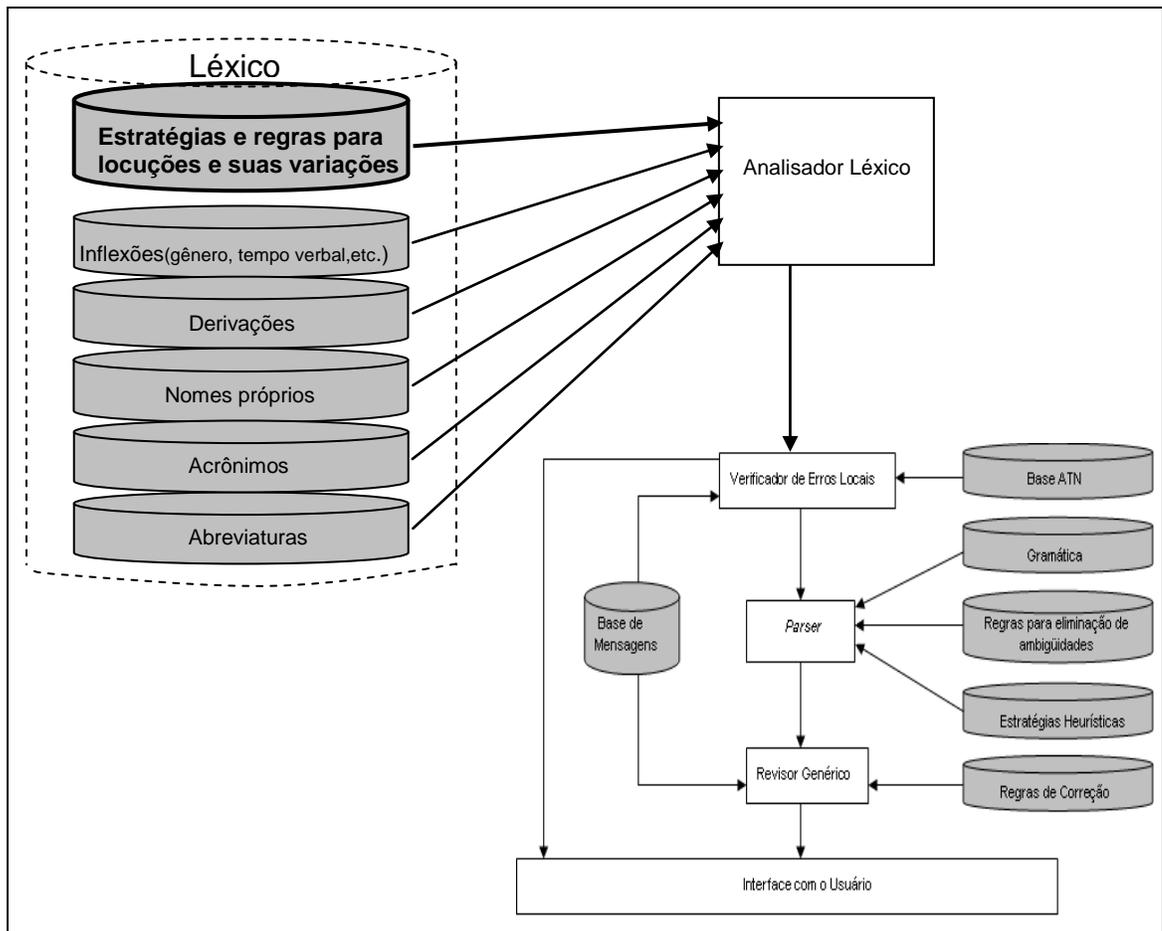


Figura 5: Meramente ilustrativa, representando a Arquitetura do seu Módulo lingüístico do ReGra e uma proposta para passar a tratar locuções no seu Analisador Léxico

pretendo dizer que *o que nós queremos era colher rosas e não fazer outra coisa*, o objeto da declaração é *colher rosas* e a declaração principal se contém incompletamente em *queríamos* (OITICICA apud BECHARA, 1975, p. 110¹).” (BECHARA, 1975, p. 110).

Também é locução verbal: *entrar e sair*.

Locução adverbial “É o grupo geralmente constituído de preposição + substantivo que tem o valor e o emprego de advérbio: *Com efeito, de graça, às vezes, em silêncio, por prazer, sem dúvida (...)*.”

¹ José Oiticica, Manual de Análise, 202-203

Outras vezes o substantivo vem com acompanhante e pode ocorrer até a omissão do substantivo: *Na verdade, de nenhum modo, em breve* (subentende-se tempo), *à direita (ao lado de, à mão direita)*, *à francesa* (subentende-se à moda).

(...) Frequentemente se omite a preposição nas locuções adverbiais de tempo e modo:

*Esta semana (por nesta semana) teremos prova. Espingarda ao ombro (por **de** espingarda ao ombro), juntou-se ao grupo de pessoas*” (BECHARA, 1975, p. 152).

Também pertencem ao grupo das locuções adverbiais: *a forma como, a ponto de, às vezes, ao mesmo tempo, como um todo, da mesma forma que, de forma que, de maneira que, de vez em quando, em função de, em via de, frente a frente, grosso modo, pelo andar da carruagem, pelo contrário, por ora, por outro lado, por sua vez, quanto possível, a favor ou contra, que nem.*

Locução conjuntiva “É um grupo de palavras com valor e empregos de uma conjunção: *para que, a fim de que, tanto que, por isso, por isso que, etc.*” (BECHARA, 1975, p. 160).

Outras locuções conjuntivas: *à vista de, acaba de, além do mais, antes (,durante) e depois, apesar disso, assim por diante, assim sendo, como se não bastasse, devido ao fato, eis que, enquanto isso, fazer frente a, forma pela qual, isto é, mas até, mas também, nada mais é, não obstante, não só, não somente, ou seja, pois então, uma vez que.*

Locução interjetiva “É um grupo de palavras com valor de interjeição: *ai de mim; ora bolas; com todos os diabos (...)*” (BECHARA, 1975, p. 167).

Locução prepositiva “É o grupo de palavras com valor e empregos de uma preposição. Em geral a locução prepositiva é constituída de advérbio ou locução adverbial seguida da preposição *de, a* ou *com*: *O garoto escondeu-se **atrás do** móvel, Não saímos **por causa da** chuva, O colégio ficava **em frente a** casa, O ofício foi redigido **de acordo com** o modelo.*

Às vezes a locução prepositiva se forma de duas preposições, como de *per* (na locução *per si*), *até a* e *para com*: *Foi até ao colégio, Mostrava-se bom para com todos.*

OBSERVAÇÃO: Acúmulo de preposições – Não raro duas preposições se juntam para dar maior efeito expressivo às idéias, guardando cada uma o seu sentido primitivo: *Andou por sobre o mar.*

Estes acúmulos de preposições não constituem uma locução prepositiva porque valem por duas preposições distintas. Combinam-se com mais freqüência as preposições: *de*, *para* e *por* com *entre*, *sob* e *sobre*. ‘*De uma vez olhou por entre duas portadas mal fechadas para o interior de outra sala..*’ (CAMILO apud BECHARA, 1975, p. 156¹), ‘*Os deputados oposicionistas conjuravam-no a não levantar a mão de sobre os projetos depredadores*’(Id., *ibid.*).” (BECHARA, 1975, p. 156).

Também pertencem ao grupo das locuções prepositivas: *a partir de*, *em frente à*, *em frente de*, *na frente de*, *o quanto antes*, *o que quer que seja*.

Considerações adicionais “Como as locuções não têm unidade de sentido, os seus elementos não devem ser unidos por hífen, seja qual for a categoria gramatical a que elas pertençam. Assim, escreve-se, *v.g.*, *vós outros* (locução pronominal), *a desoras* (locução adverbial), *a fim de* (locução prepositiva), *contanto que* (locução conjuntiva), porque essas combinações vocabulares não são verdadeiros compostos, não formam perfeitas unidades semânticas. Quando, porém, as locuções se tornam unidades fonéticas, devem ser escritas numa só palavra: *acerca* (adv.), *afinal*, *apesar*, *debaixo*, *decerto*, *defronte*, *depressa*, *devagar*, *deveras*, *resvés*, etc.” (BECHARA, 1975, p. 65).

Há, ainda, algumas locuções que não se enquadram em nenhum dos grupos anteriormente citados, e que pertencem à categoria das expressões idiomáticas: *longo trabalho pela frente*, *pano de fundo*, *pôr em xeque*, *prova dos nove*, *tomar parte em*, *um e outro*, *valer a pena*, *ver-se às voltas*, *via de regra*, etc.

¹ Camilo, A queda dum Anjo, 175

Em MARTINS, 1997, há várias instruções orientando a evitar a utilização inadequada de alguns grupos de locuções: *regra geral, ter lugar, trabalhar com a hipótese, troca de farpas, tudo a ver, último adeus, de vez que, em vias de e às vistas de*. A seguir, as que devem ser evitadas por se tratar de lugares-comuns: *requintes de crueldade, rindo à toa e rota de colisão*. Também, há aquelas que podem ser suprimidas por serem irrelevantes ao sentido da frase: *que é, que era, que foi e que forem*.

Ainda sobre este assunto, em MARTINS, 1997 o autor faz as seguintes considerações sobre o emprego correto das seguintes locuções: "**Quer ... quer**. É a forma correta: *Quer queira quer não queira... / Quer sejam filhos quer apenas amigos...* Nunca: *quer queira 'ou' não queira, quer sejam filhos 'ou' apenas amigos*, etc. Pode-se usar também: *ou queira ou não queira*. Neste caso, o primeiro **ou** pode ser omitido: *queira ou não queira...*" (MARTINS, 1997, p. 247).

"**'Que tem direito'**. O certo é **a que tem** (tinha, teve, terá) **direito**: *Exige tudo a que tem direito (e não 'tudo que tem direito')*. / *Reivindicaram as vagas a que tinham direito*. / *Imaginou as regalias a que terá direito*. / *Recebeu a herança a que teve direito*)." (MARTINS, 1997, p. 247).

Para algumas delas, a recomendação do autor é para que sejam substituídas; ou que não sejam utilizadas, pois são consideradas vícios lingüísticos sob o aspecto do estilo: "2 - Considere também **muletas** (e, portanto, fuja delas) locuções utilizadas para ligar fatos ou declarações diferentes na mesma matéria, como: *por outro lado, pelo contrário, ao mesmo tempo, não obstante, enquanto isso, por sua vez, apesar disso* e outras já desgastadas pelo uso excessivo. Para cumprir funções semelhantes você tem palavras ou expressões como *ainda também, mas, porém, no entanto, pois, como, porque, portanto*, etc." (MARTINS, 1997, p. 183). "29 - Por encadeamento de parágrafos não se entenda o cômodo uso de **vícios lingüísticos**, como *por outro lado, enquanto isso, ao mesmo tempo, não obstante* e outros do gênero. Busque formas menos batidas ou simplesmente as dispense: se a seqüência do texto estiver correta, esses recursos se tornarão absolutamente desnecessários." (MARTINS, 1997, p. 18).

OBSERVAÇÃO: Na teoria da Análise do Discurso, muitos Marcadores Discursivos (vide glossário) são constituídos por estas locuções. Suprimi-los pode comprometer o desempenho na separação de sentenças de um *parser* desenvolvido para aplicações com esta finalidade.

A inconveniência em se analisar separadamente cada palavra que compõe uma locução também aparece registrada como uma preocupação recorrente no processo de construção do OCD: “A informação complementar sobre o registro disponível serve também para quando qualquer par de palavras em combinação assumam um sentido diverso ao destas duas palavras separadamente.” (LEA, 2002, p. x). Alguns dos exemplos colhidos nesta obra foram detalhados no APÊNDICE 1 como propostas de aplicação do *Workflow*:

cirurgião *substantivo* VERB+CIRURGIÃO **trabalhar como, qualificar-se como**, apresentado no exemplo III;

registrar *verb* VERB+REGISTRAR **falhar ao não, ter que, ser autorizado a**, apresentado no exemplo X;

preço *substantivo* VERB+PREÇO **jogar para o alto, ir de ... a ...**, apresentado no exemplo IV;

proferir *verb* VERB+PROFERIR **ser difícil para**, apresentado no exemplo V;

taxa *substantivo* ADJ de mercado, apresentado no exemplo IV;

vítima *substantivo* ADJ **propenso a se tornar**, apresentado no exemplo XI.

III. DICIONÁRIO DE COLOCAÇÕES DAS CLASSES DE PALAVRAS

Conforme visto no CAPÍTULO I, na sua primeira versão, a base de conhecimento no formato de um Catálogo de Colocabilidades exigiria um esforço prévio de tradução da obra original em Inglês que lhe serviu de inspiração. Na realização desta proposta, a adaptação desta obra ao Português falado no Brasil, mesmo levando-se em conta as diferenças ortográficas e sintáticas entre os dois idiomas, permitiria um bom índice de aproveitamento; pois, como observou BECHARA (1975, p. 340): “A significação dos vocábulos está intimamente relacionada com o mundo das idéias e dos sentimentos”. Da mesma forma que a significação dos vocábulos, muitas das categorizações entre eles (ou colocações, conf. será visto mais adiante) também estão no mundo – ou domínio – das idéias, independentemente do idioma: “De todo modo, parece evidente que, como mostravam os filósofos clássicos (exatamente os que montaram o suporte do pensamento ocidental sobre a linguagem), substantivo e verbo comandam o espetáculo da linguagem, o primeiro representando as categorizações do mundo que nos permitem falar dele, e o segundo representando os processos ativados para, a partir dos estados de coisas, expressar julgamentos (duas matérias de reflexão de Aristóteles, no *Das categorias* e no *Da interpretação*), compondo a interação.” (NEVES, 2007). O esforço no desenvolvimento deste subproduto, totalmente independente do restante da solução, viria a agregar uma participação importante a outras pesquisas desenvolvidas nessa linha de estudos; o *Dicionário de Usos de Português Contemporâneo*, citado em NEVES (2007): “(...)com orientação eminentemente construcional elaborado por uma equipe da UNESP de Araraquara, sob coordenação de F. S. Borba”. Ainda sobre este artigo, a autora aponta para uma lacuna nas obras de apoio ao nosso léxico, pois não há dicionário algum “(...)centrado na colocabilidade das diversas palavras lexicais da língua”.

Seguem algumas considerações a respeito da compilação desta obra no formato de uma base de conhecimento:

1. Grupos de Colocações

Há diferentes definições para o termo **Categorização**, conforme considerações apresentadas no Glossário. Em LEA (2002, p. ix), o termo é citado no seguinte contexto: “(...) outra área da colocação pode ser chamada '*category collocation*', em que uma palavra pode se combinar com qualquer outra palavra de um conjunto bem definido. Este pode ser um conjunto extenso, mas os seus membros são previsíveis por seguirem um padrão determinado, como por exemplo: todas as palavras usadas para designar nacionalidades, ou medidas de tempo. (...). Isto também ocorre com certos grupos de palavras que compartilham todas ou muitas das suas colocações, característica de alguns conjuntos bem determinados, como: os dias da semanas/meses e pontos cardeais; mas que é igualmente comum entre outros grupos de palavras não tão coesos, embora tão restritos quanto os primeiros, como: os conjuntos de diferentes tipos de moedas, pesos, medidas e refeições.”.

No modelo proposto neste estudo, optou-se por representar tais conjuntos como referências a objetos:

Exemplos do APÊNDICE 1:

imigrante *substantivo* ADJ. **irlandês, italiano, etc.** `nacionalidades()`, no exemplo I;

juro *substantivo* ADJ. **anual, mensal, etc.** `períodos()`, no exemplo IV;

quantia *substantivo* ADJ. **anual, mensal, etc.** `períodos()`, no exemplo VII;

fruta *substantivo* QUANT. **dúzia, dezena, etc** `medidas_de_quantidade()`, | **quilo, grama, etc** `medidas_de_peso()` | **caixa, saco, pacote, etc**, `tipos_de_embalagem()`, no exemplo VII.

2. Incorporação de um Tesouro à solução

Para situações em que uma palavra pode se combinar com uma lista de sinônimos cabe uma interconexão com um tesouro (vide glossário) para relacioná-los, respeitando algumas restrições, como a observada a seguir: “(...) palavras determinadas que combinam, apenas e tão somente, com alguma outra específica: *small fortune* não pode ser alterada para *little fortune*, ainda que: *small* e *little*

aparentem ser sinônimos”(LEA, 2002, p. ix). Em BECHARA (1975, p. 345) este fenômeno lingüístico é denominado **gradação semântica**: “**3) Sinonímia**. – É o fato de haver mais de um vocábulo com a mesma ou quase a mesma significação: *casa, lar, moradia, residência, mansão*. Um exame detido nos mostrará que a identidade dos sinônimos é muito relativa; no uso (quer literário, quer popular) eles assumem sentidos ‘ocasionais’ que no contexto um não pode ser empregado pelo outro sem que se quebre um pouco o matiz da expressão. Uma série sinonímica apresenta-se nos com pequenas gradações semânticas quanto a diversos domínios: o sentido abstrato ou concreto; o valor literário ou popular (fenecer / morrer); a maior ou menor intensidade de significação (chamar / clamar / bradar / berrar); o aspecto cultural (escutar / auscultar) e tantas outras.”. Portanto, a interconexão ao Tesouro mencionada teria que ser estendida com capacidades adicionais, que permitissem discernir sobre situações apropriadas ao emprego do termo em um determinado contexto, levando em conta as identidades dos sinônimos.

Exemplos do APÊNDICE 1:

cachorro *substantivo* ADJ. **bom, bem-comportado, bem-treinado, etc.**
sinônimos(“bom”, termo), no exemplo II, para uma lista composta por termos com sentido positivo aplicáveis ao cachorro.

3. Termos que exigem complemento

A solução teria que estar preparada para atuar apropriadamente sobre termos que exijam complemento — representados em LEA, 2002 por **~a algo**, por exemplo — acusando a omissão ou identificando a sua presença e particularidades para o tipo de complemento que melhor se aplica ao termo.

Exemplos do apêndice 1:

referir *verb* PREP. **a ~(algo)** locução(termo, a(algo)), no exemplo X;
paciente *substantivo* PACIENTE+VERB **desenvolver (algo), ter (algo)** termo(algo), no exemplo III.

4. Regras sintáticas para uma sentença

Neste formato simplificado do modelo, não foram detalhados os algoritmos para validar a ordem correta dos termos na sentença, ou a disposição relativa aos demais termos num dado contexto. A solução para esta necessidade impõe uma verificação mais complexa em que a seqüência, ou a posição, dos termos dentro da sentença obedeça a múltiplos fatores, conforme comenta BECHARA (1975, p. 322) sob um aspecto para o conceito de Colocação totalmente diverso do apresentado por LEA (2002): “7— COLOCAÇÃO *Sintaxe de colocação* ou *de ordem* — é aquela que trata da maneira de dispor os termos dentro da oração e as orações dentro do período. A *colocação*, dentro de um idioma, obedece a tendências variadas, quer de ordem estritamente gramatical, quer de ordem rítmica psicológica e estilística, que se coordenam e completam. O maior responsável pela ordem favorita numa língua ou grupo de línguas parece ser a entonação oracional.”.

Aplicações existentes na área de linguística computacional utilizam-se de uma metalinguagem própria chamada BNF (Backus-Naur *form*, vide glossário) para representar e resolver regras sintáticas, conforme demonstração de um trecho de código na figura 6.

```

SENTENCE ::= COMPOSED_CLAUSE | SIMPLE_CLAUSE
COMPOSED_CLAUSE ::= SIMPLE_CLAUSE COMPOSED_CLAUSE |
SIMPLE_CLAUSE SIMPLE_CLAUSE
SIMPLE_CLAUSE ::= BOUNDARY [COORDINATIVE CONJUNCTION]
[ADVERBIAL ADJUNCT] SUBJECT [ADVERBIAL ADJUNCT] BOUNDARY |
[COORDINATIVE CONJUNCTION] [ADVERBIAL ADJUNCT] PREDICATE
[ADVERBIAL ADJUNCT] SUBJECT [ADVERBIAL ADJUNCT] |
[COORDINATIVE CONJUNCTION] [ADVERBIAL ADJUNCT] [SUBJECT]
[ADVERBIAL ADJUNCT] PREDICATE [ADVERBIAL ADJUNCT]

```

Figura 6 Algumas regras sintáticas expressas em BNF para uma sentença determinada
Fonte: MARTINS et al, 1998, p.14

Exemplos do apêndice 1:

venda *substantivo* PREP. **de** **~(algo)** **~a (subs)** *Venda de trigo ao Brasil*

locução (termo, de (algo), a (subs)), no exemplo IX.

5. Omissões do OCD

No início deste estudo não se esperava que o formato apresentado por LEA (2002) convergisse para a estrutura da obra de MARTINS (1997) de forma natural e absoluta. Porém, é conveniente destacar algumas situações importantes para as quais este dicionário é omissivo; mesmo levando-se em conta que esta obra foi escrita originalmente para outro idioma, e que venha a sofrer adaptações ao Português falado no Brasil. Prevendo impactos negativos no funcionamento desta solução, cabem algumas considerações sobre implementações adicionais ou extensões que viriam a mitigá-los:

Colocações Pronominais O OCD não contempla colocações pronominais. Principalmente para os verbos, as verificações das colocações pronominais são fundamentais para, como no caso do termo *proliferar* no exemplo VI do APÊNDICE 1, identificar os casos em que os pronomes não são admitidos:

“**Proliferar**. E não ‘proliferar-se’.” (MARTINS, 1997, p. 238).

proliferar *verbo* PRON¹, no exemplo VI;

Colocações de Substantivos para Verbos e Adjetivos Pelas justificativas apresentadas a seguir, LEA (2002) omite também as colocações de substantivos para verbos e adjetivos: “A terceira questão (Eles encontrarão com facilidade este registro quando precisarem?) levou a omitir substantivos colocados para verbos e adjetivos. Ao conceber suas idéias, uma pessoa geralmente parte de um substantivo. Ao pensar na chuva, pensamos, em seguida, quais os adjetivos que melhor descrevem uma grande precipitação em um curto espaço de tempo. Seria improvável que partíssemos do adjetivo *heavy*, e surpreendente a variedade de idéias que poderíamos associar a este termo (chuva, brisa, danos, fogos de artilharia?). Da mesma maneira, podemos procurar o verbo mais adequado para utilizar numa

¹ A notação em tachado duplo representa uma colocação não registrada para o termo.

situação em que precisemos saber o que fazer em resposta a um *challenge*. Mas não escolheríamos *meet* e sim *what to meet* (*a challenge, an acquaintance, your death, the expense*).” (LEA, 2002, p. ix).

Este ponto de vista não é compartilhado por NEVES (2007). A autora prefere destacar o papel das ações no processo de concepção das idéias: “É fácil perceber que, ao compormos nossas frases, ao estruturarmos as orações, selecionamos predicados/verbos e, para junto deles, levamos os participantes implicados naquele tipo de evento: se alguém falar em ‘latir’, vai aparecer um cachorro (ou, por exemplo, uma pessoa que resolveu imitar um cachorro); se alguém falar em ‘comprar’, vai aparecer uma pessoa que compra algo assim como vai aparecer aquilo que é o objeto da compra; se alguém falar em ‘entregar’, vai existir uma pessoa que pratica a ação, vai existir uma coisa ou animal que é objeto dessa ação (ou até uma pessoa, se ‘entregar’ tiver um sentido metafórico), e vai existir uma outra pessoa que recebe aquilo que é entregue. Montam-se, assim, pela linguagem, os eventos, que são estados de coisas do mundo do qual se fala. Até aí, parti do verbo, coloquei-o no centro de montagem das cenas e invoquei substantivos que colocam como objetos-de-discurso os participantes desses eventos.”.

Apesar disto, considerando-se as limitações de espaço impostas pelo seu formato impresso e a abrangência da sua proposta (“Um dicionário de colocações não faz generalizações desta natureza: ele abrange, por princípio, a linguagem completa – ou grande parte dela! –; palavra por palavra, colocação por colocação.” [LEA, 2002, p. viii]); esta obra, de fato, não comportaria uma solução que viesse a preencher todas essas lacunas. As justificativas apresentadas procedem, portanto; e, dentro do seu escopo como material de apoio à escrita, o OCD atinge os seus objetivos. Já para uma solução automatizada, os volumes para o armazenamento eletrônico dos dados são virtualmente ilimitados. Considerando apenas os impactos sobre as exigências de desempenho, um projeto fundamentado numa Base de Conhecimento no formato de um Catálogo de Colocabilidades, da forma como foi apresentado no CAPÍTULO 1, comportaria plenamente tais extensões. A figura que representa a organização gráfica do OCD (figura 3) foi redesenhada na figura 7, de forma a representar estas considerações.

Estruturas possíveis no OCD	<i>noun</i>	<i>verb</i>	<i>PHRASAL VERBS</i>	<i>result in algo refer to sb/algo</i>	<i>adj.</i>
	'n' significado 'n'	'n' significado 'n'			'n' significado 'n'
					<u>VERBOS</u>
	<u>ADJ.</u>				
		<u>ADV.</u>			<u>ADV.</u>
	<u>QUANT.</u>				
	<u>VERB + termo</u>	<u>VERB + termo</u>			
	<u>termo + VERB</u>				
	<u>termo + NOUN</u>	<u>termo + NOUN</u>			<u>termo + NOUN</u>
	<u>PREP.</u>	<u>PREP.</u>			<u>PREP.</u>
	<u>PRON.</u>				
<u>FRASES</u>	<u>FRASES</u>				
NOTE					

Figura 7 - meramente ilustrativa representando a organização gráfica do OCD alterada para passar a tratar pronomes e substantivos para verbos e substantivos para adjetivos.

Segue exemplo do APÊNDICE 1 apresentando colocações de substantivo para verbo:

“**Rapar, raspar.** Para cabelo, use apenas **rapar**, que significa cortar rente: *Rapou a barba. / Rapou o cabelo do calouro. Raspar* significa lixar, desbastar, tocar ou ferir de raspão: *O marceneiro raspou a madeira.*” (MARTINS, 1997, p. 248).

rapar verbo SUBSTANTIVO **cabelo, barba**, no exemplo VIII;

raspar verbo SUBSTANTIVO **madeira**, no exemplo VIII. Segue exemplo do APÊNDICE 1 apresentando colocações de substantivos para adjetivos:

“**Vítima fatal**’. Fatal significa mortífero, que causa a morte, que traz ruína ou desgraça. Por isso, não existe a expressão ‘vítima fatal’: a vítima **recebe** a morte, e não **a produz**. Fatal é um golpe, um tiro, um acidente, uma pancada, um choque, uma batida, e nunca a vítima.” (MARTINS, 1997, p. 312).

fatal adjetivo SUBSTANTIVO **golpe, tiro, acidente, pancada, choque, batida**, no exemplo XI.

OBSERVAÇÕES: Haveria ainda outros aspectos detectados, mas não abordados, para que esta análise não viesse a se tornar muito extensa; como: a verificação de

colocabilidades para alguns dos verbos mais comuns: “Algumas das palavras mais comuns – como os verbos *make* e *do* – não justificam, por si, registros próprios. Isto porque estes verbos, individualmente, não têm colocações reais. Eles são as colocações de muitos substantivos, e aparecem nos registros como coadjuvantes de muitos deles.” (LEA, 2002, p. x); o enfoque da obra sobre a linguagem ordinária, ignorando as formas de expressão mais sofisticadas de usuários especializados: “A atenção principal voltou-se para o que poderia ser chamada ‘linguagem moderadamente formal’ – a linguagem utilizada por não-especialistas instruídos para redigirem seus trabalhos” (LEA, 2002, p. viii e ix). Também, este dicionário omite colocações para: artigos, conjunções, etc. O **ReGra**, porém, já provê recursos que resolvem satisfatoriamente alguns destes itens.

Finalmente, apesar de não ser o objetivo desta proposta, um Dicionário de Colocações das Classes de Palavras viria agregar uma participação importante no desenvolvimento de um subproduto totalmente independente do conjunto da solução: novos conjuntos de regras para eliminação de ambigüidades para produtos já existentes no mercado (vide **Regras para eliminação de ambigüidades** no **CAPÍTULO II**), como nos exemplos do APÊNDICE 1 em que, conforme o contexto, os seguintes termos podem ser adjetivos ou substantivos:

barba *substantivo* ADJ. **(tipo)cavanhaque** , no exemplo VIII;

BARBA + SUBSTANTIVO **--de-dois-dias** *Ele tinha uma barba-de-dois-dias crescida sobre o rosto*, no exemplo VIII;

quantia *substantivo*, no exemplo VII.

IV. APLICAÇÃO DO MODELO DE OO PARA A ANÁLISE SEMÂNTICA

Como forma de ilustrar a aplicação prática do *Workflow*, este trabalho apresenta no **APÊNDICE 1** alguns Casos de Uso utilizando exemplos colhidos no MRE que foram resolvidos a partir do registro do termo no Catálogo de Colocações traduzido do OCD para o Português do Brasil. Em seguida, o registro deste termo foi compilado no formato de Modelo de Classes no paradigma de Orientação a Objetos.

Conforme a representação da organização gráfica do OCD, vista anteriormente na figura 3, seu formato é compatível com um projeto de um diagrama de classes convencional. Para facilitar o entendimento das demonstrações desta solução, ela foi codificada na linguagem de programação Java, com a opção pelos seus comandos mais básicos. Nesta, que é uma das linguagens mais populares para OO, os atributos dos objetos inicializados com os conteúdos das colocações foram representados na forma de Vetores de Strings, mas poderiam ser pilhas, *collections*, bancos de dados¹, ou quaisquer outras soluções que suportem operação com tabelas, listas ou pilhas. Necessidades mais específicas foram representadas como referências a outros objetos (como *locução()*, por exemplo) cuja implementação não foi detalhada. Como pressuposto, os termos da sentença deverão estar previamente rotulados por meio de algum *tagger* morfossintático de alto desempenho, como o MXPOST e URML² (vide glossário). A classe então viria a atuar no sentido de acusar a tentativa de uso indevido do termo ao não encontrá-lo entre as colocações

¹ A utilização de bancos de dados é questionável, pois os dados das classes são estáveis e o desempenho será melhor se estas listas forem incorporadas diretamente ao código. Somente atualizações realizadas no *PROFILE* do usuário (Vide CAPÍTULO 1), poderiam ser feitas desta forma.

² Além das classes de palavras mais comuns, como: substantivos, adjetivos, verbos, etc.; estes *taggers* deveriam estar capacitados a reconhecer e rotular as classes da forma como o OCD organiza as colocações dos termos, inclusive identificando os substantivos coletivos (QUANT. ou *quantifiers*).

catalogadas para ele¹. Neste formato simplificado do modelo, não foram detalhados os algoritmos para validações gramaticais mais complexas, que envolvam o contexto do emprego do termo na sentença ou na frase. Tais sofisticções caberiam num modelo pronto mais elaborado e próximo da implementação física, pois extensões deste tipo seriam plenamente suportadas pelos muitos recursos a serem explorados no paradigma da OO. Isto, porém, estaria fora do objetivo principal deste trabalho, de apenas abstrair conceitos do modelo proposto para exemplificar um direcionamento mais prático da solução e mapear os desafios para a sua implementação.

1. Herança

Diante dos muitos recursos a serem explorados suportados no paradigma da OO, um especialmente útil é o da Herança. Este recurso permite a formação de uma hierarquia de classes, em que cada nível é a especialização do nível anterior, que é mais genérico, aproveitando a sua implementação de estruturas de dados e métodos.

Foram estudados dois Casos de Uso em que a aplicação deste recurso viria a contribuir para simplificar o código e garantir uma organização maior no projeto:

O mais representativo é a situação descrita por MARTINS (1997) em que um conjunto de verbos que compartilham um sentido comum de **dizer** (verbos *dicendi* no exemplo X do APÊNDICE 1) não admitem a preposição **que** (MARTINS, 1997, p. 307). O recurso mencionado representaria todos os verbos que compartilham esta característica herdada de uma classe genérica de nível superior.

Outro caso mais específico, diz respeito a particularidades zoológicas das espécies, dividindo-as entre animais que **mordem** ou **picam** (MARTINS, 1997, p. 181), conforme registro no exemplo II do APÊNDICE 1. Sob este aspecto, haveria dois grandes grupos de animais em que os descendentes de cada um deles

¹ Com capacidades, também, para cruzar as análises dos diferentes termos que compõem a frase; a fim de atender necessidades de verificações mais complexas, que impliquem em interrelacioná-los.

compartilhariam uma destas características mutuamente exclusivas: ser dotado de dentes ou algum outro apêndice incisivo para a sua defesa, conforme a representação na figura 8.

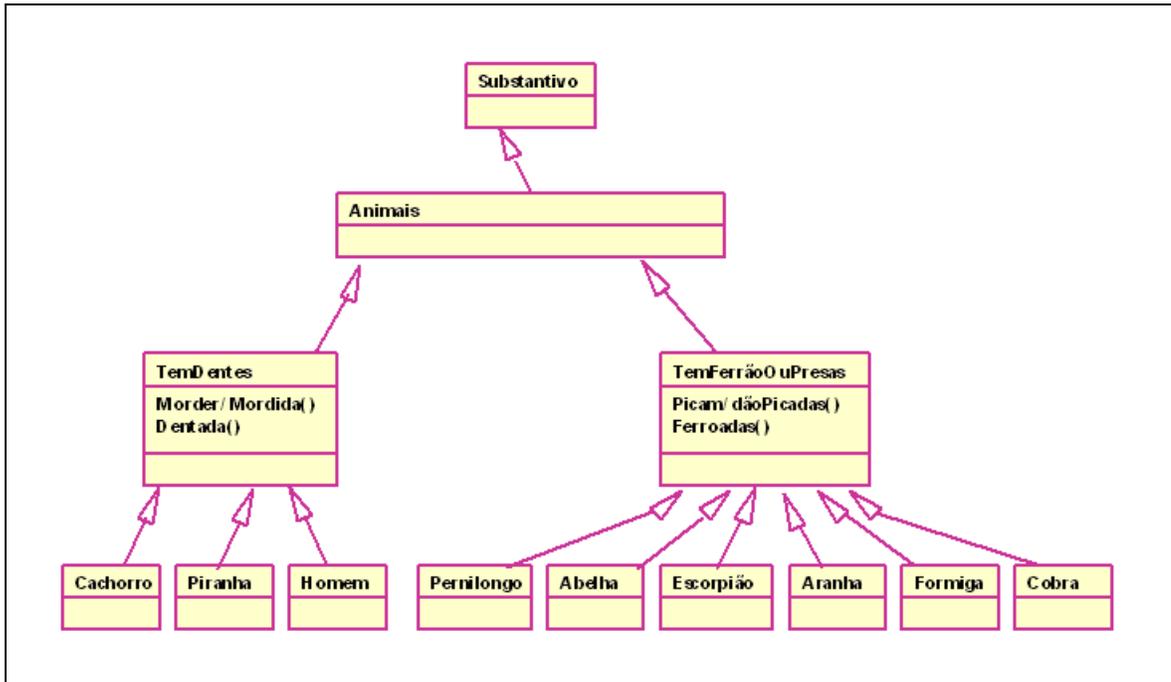


Figura 8 - Modelo de classes representando a hierarquia dos objetos com as características herdadas e suas especializações

V. CONCLUSÃO

Conforme foi visto no decorrer desta argumentação, esta proposta fundamenta sua abordagem na utilização de dois materiais principais; cuja boa aceitação entre os mais diversos tipos de públicos pode ser confirmada pela impressão de múltiplas edições destas obras (o MRE chegou à 4ª edição). No desenvolvimento do argumento deste estudo, estes dois materiais convergiram para uma solução integrada, apesar das diferenças importantes observadas entre eles. A destacar:

Em MARTINS (1997), um único autor assina a obra, o que garante a uniformidade do seu conteúdo. Já em LEA (2002), contribuíram: “*over the years, a large number of lexicographers and editors have been involved in this Project*” (LEA, 2002, p. v). Em decorrência disto, ao se confrontar as anotações nos registros dos diversos termos, é possível notar-se diferentes níveis de abstração entre eles;

O conteúdo apresentado em MARTINS (1997) é dirigido a erros comuns do idioma e as formas para evitá-los. Mais abrangente, a proposta em LEA (2002) é “*covers the entire language (or a large parte of it!) on a word by word, collocation by collocation*” (LEA, 2002, p. viii);

As análises em MARTINS (1997) são mais aprofundadas, pois foram originalmente dirigidas a um público especializado, com capacidade de discernimento suficiente para a reflexão sobre a aplicação de algumas regras necessárias para garantir a uniformidade da edição de um jornal, levando em conta as normas editoriais e de estilo adotadas. Já o enfoque principal em LEA (2002), é o uso produtivo da linguagem por parte de um público não-especialista, que encontra dificuldades específicas em questões envolvendo colocabilidades entre os termos; o que dirigiu o seu desenvolvimento para um formato “*as comprehensive and acessible as possible*” (LEA, 2002, p. x);

Os critérios utilizados no processo de construção do OCD foram essencialmente mecânicos, baseados em um *corpus* consolidado: “*The main source used was the 100 million words British National Corpus. From this, compilers of the dictionary were able to check how frequently any given combination occurred, in how many (and what kind of) sources, and in what particular contexts.*” (LEA, 2002, p. viii). Diferentemente desta obra, o MRE é composto de amostras obtidas baseadas unicamente no julgamento do autor, reunidas por ele ao longo das décadas da sua rica vivência profissional.

Assim, a leitura de MARTINS (1997) com uma visão analítica sob o aspecto da teoria lingüística computacional, levou às considerações que foram elaboradas no formato de aplicações de Casos de Uso (para algumas situações específicas, que pareceram adaptar-se mais facilmente ao modelo); e realizadas, no APÊNDICE 1, em uma proposta de modelo de solução.

Na maior parte dos casos, porém, a leitura revelou que esta análise iria muito além da simples aplicação de algumas poucas regras claras e bem definidas. Nestas situações, pelo que postula o Modelo Essencial (vide glossário) em Análise de Sistemas, coube apenas abstrair conceitos a fim de entendê-los melhor.

Dentre os casos mais representativos, por conveniência e facilidade de interpretação, optou-se por subdividi-los em grupos em vez de estudá-los caso a caso. Estes grupos foram formados de acordo com a presunção do nível de dificuldade desta análise e algumas características comuns compartilhadas por eles:

Análise do Contexto Para os casos estudados neste grupo seriam necessárias etapas preliminares de análise em camadas, envolvendo o contexto e termos adjacentes. A dificuldade estaria em interpretar sentidos e/ou intenções, de forma a garantir a precisão da análise. Espera-se que determinadas palavras forneçam algum indício lingüístico – forte, fraco ou menos evidente – que permita presumir estas intenções; como, por exemplo, o emprego de determinados substantivos numa frase: “**Grama** (gênero). Palavra masculina quando significa peso: *um grama*,

trezentos gramas, oitocentos gramas. Como equivalente a relva é que tem o gênero feminino: *A grama do jardim. / Não pise na grama.*” (MARTINS, 1997, p. 136). Neste caso, um numeral precedendo o termo pode apontar para o sentido de medida de peso. Após esta análise, o termo poderá, então, ser classificado e a sua colocabilidade analisada por suas próprias regras, permitindo inferir sobre o seu emprego e posição corretos entre os outros termos da frase.

Há exemplos em que, para discernir sobre construções apropriadas em um dado contexto, o autor se refere à necessidade de inferir sentidos positivos, como: desejo e expectativa favorável: “**Ganhar.** O verbo tem sentido **positivo**. Por isso, um time não pode ‘*ganhar*’ *mais baixas*. Da mesma forma, ninguém ‘*ganha*’ uma *cicatriz no rosto*, um *processo*, uma *punição*, uma *repreensão*, uma *advertência*, uma *multa* ou uma *descompostura*. Pode, isso sim, *receber uma punição, uma repreensão, etc.*” (MARTINS, 1997, p. 133); ou negativos, como: oposição, choque e confronto: “‘**Enfrentar boas condições**’. **Enfrentar** indica confronto. Por isso, é errado empregar o verbo como no exemplo seguinte: *O motorista no momento vai ‘enfrentar’ (o certo: encontrar) boas condições de trânsito.*” (MARTINS, 1997, p. 107).

Para outras situações analisadas, há a necessidade de identificar intenções como: *em segredo, na intimidade, em confiança, etc.*: “**Confidenciar.** O verbo significa dizer em segredo, em confiança, na intimidade, e está sendo muito usado erradamente no lugar de revelar ou informar. Veja como empregá-lo corretamente: *O diretor da organização confidenciou a amigos que não vai tomar a iniciativa de deixar o cargo. / O deputado confidenciou à família que não será candidato à reeleição.* Ou seja, disseram em segredo, em confiança, na intimidade. Nos casos seguintes, no entanto, **confidenciar** não tem sentido: ‘*Realmente, a negociação está adiantada*’, *confidenciou* (disse, garantiu, informou, adiantou) *o diretor do time no vestiário. / A atriz confidenciou* (revelou) *à revista Notícias que está esperando um filho. / O líder confidenciou* (comunicou, revelou, informou) *aos 400 convencionais qual a posição do partido na questão.*” (MARTINS, 1997, p. 79 e 80).

Em outros exemplos a análise passa pela substituição de um termo na frase e a capacidade para inferir se o sentido manteve-se idêntico ao original: “**Tachar, taxar.** Com o significado de acusar, censurar, pôr defeito em, o correto é **tachar**. O *deputado tachou o adversário de corrupto. / Eles o tacharam de leviano. Taxar* quer dizer impor tributo a: *Os governos taxavam o país a mais não poder.*” (MARTINS, 1997, p. 277).

Análise de Discurso A interpretação de alguns outros exemplos constitui um desafio ainda maior, pois implica em estender as capacidades da solução com recursos adicionais que somente encontram suporte na teoria da Análise de Discurso (vide glossário). Para outras, ainda, esta análise precisaria ser dotada com recursos avançados que não se encontram suficientemente amadurecidos, nem mesmo nas pesquisas de vanguarda nesta área de estudos. Nos exemplos analisados, há relações – retóricas e/ou semânticas – sutis a tal ponto que uma ferramenta automática teria que ser dotada de uma capacidade de discernimento próxima à intuição humana para poder resolvê-las sintaticamente.

Avaliado superficialmente sob o enfoque da *Rhetorical Structure Theory* – RST apresentada em PARDO (2005, p. 12, 18 e 149), o seguinte exemplo extraído de MARTINS (1997, p. 175) confronta o que seria uma relação semântica de Jordan *After Simultaneous* contra uma relação retórica de CAUSE (revelada pelo Marcador Textual **na medida em que** Conf. Figura 9): “**Medida.** Veja a diferença entre as expressões: 1 - **À medida que** (e não ‘à medida em que’) equivale a à proporção que, ao mesmo tempo que, conforme: *As mortes iam aumentando à medida que a epidemia se alastrava. / À medida que conquistava cargos, tornava-se mais autoritário.* 2 - **Na medida em que** corresponde a tendo em vista que: *Na medida em que não tenham ficado claras as acusações, todos estão sob suspeita. / É preciso cumprir as leis, na medida em que elas existem.* ” .

	um bem alheio às escondidas: <i>furto de carros, livros furtados da estante.</i>	Efeito: Use Furto e furtar
	2 – Roubo e roubar pressupõem violência ou ameaça: <i>Os assaltantes roubaram o carro e feriram o motorista. / O roubo ao banco ocorreu na hora de maior movimento”</i> (MARTINS, 1997, p.132)	Causa: apoderar de um bem alheio usando de violência ou ameaça Efeito: Use Roubo e roubar
Ir a, ir para	“Ir a indica curta permanência, <i>Vai a Paris</i> (vai e volta logo) <i>este mês; vai a Campos do Jordão esta semana.</i>	Causa: Curta permanência Efeito: Use Ir a
	ir para dá idéia de destinação, demora, <i>vai para Paris no fim do ano</i> (vai e fica, pelo menos algum tempo). <i>Vai para Campos do Jordão nas férias;”</i> (MARTINS, 1997, p. 150)	Causa: Destinação, demora Efeito: Use Ir para
Visar, visar a	“1 - No sentido de mirar ou dar o visto a, use a regência direta: <i>Visou o alvo. / Visou o ladrão. / Visou o cheque.</i>	Causa: Mirar, dar o visto a Efeito: Regência direta; Use Visar
	2 - Como equivalente a ter em vista ou ter por objetivo, exige a preposição a <i>Visou exclusivamente aos seus interesses. / Nunca visaram à riqueza. / Tudo a que visavam eram melhores condições de vida. / Visava a atender aos pedidos dos amigos. / Visou a proporcionar conforto à família.”</i> (MARTINS, 1997, p. 312)	Causa: Ter em vista, ter por objetivo Efeito: Exige a preposição a ; Use Visar a

Figura 10: Amostra contendo exemplos do MRE estruturados em pares “causa-efeito”.

Domínios de Conhecimento Há exemplos em que o sentido de um termo é relativo ao domínio de conhecimento sobre o qual a argumentação do autor se fundamentou; e há até mesmo situações em que esta análise serviria para resolver ambigüidades, pois palavras semelhantes podem denotar sentidos diversos de acordo com o jargão em que ela estiver sendo empregada. Assim como a questão das locuções, esta necessidade se impõe no desenvolvimento do projeto já nas suas pesquisas exploratórias, e aparece registrada como uma preocupação recorrente na construção do OCD: “(...) tratando os diversos assuntos - negócios, ciência, história, esporte, etc. (esta lista poderia ocupar metade da página). Adicionalmente, o dicionário inclui algumas das mais importantes colocações utilizadas nas diferentes áreas de especialização, particularmente as mais corriqueiras nos jargões jurídico e médico; expressões populares, mais comuns ao tratar intimidades como sentimentos e relacionamentos; colocações informais, de uso corrente na linguagem coloquial; e

umas poucas, embora freqüentes, colocações características do jornalismo Britânico. As características do uso – técnico, informal ou jornalístico – foram devidamente discriminadas como informação suplementar.” (LEA, 2002, p. ix).

Um esboço do que seria esta divisão do universo dos termos nos seus diferentes domínios do conhecimento – subdivididos em jargões e dialetos específicos – é apresentado na figura 11 num formato meramente ilustrativo, aplicando rudimentos da teoria dos conjuntos.

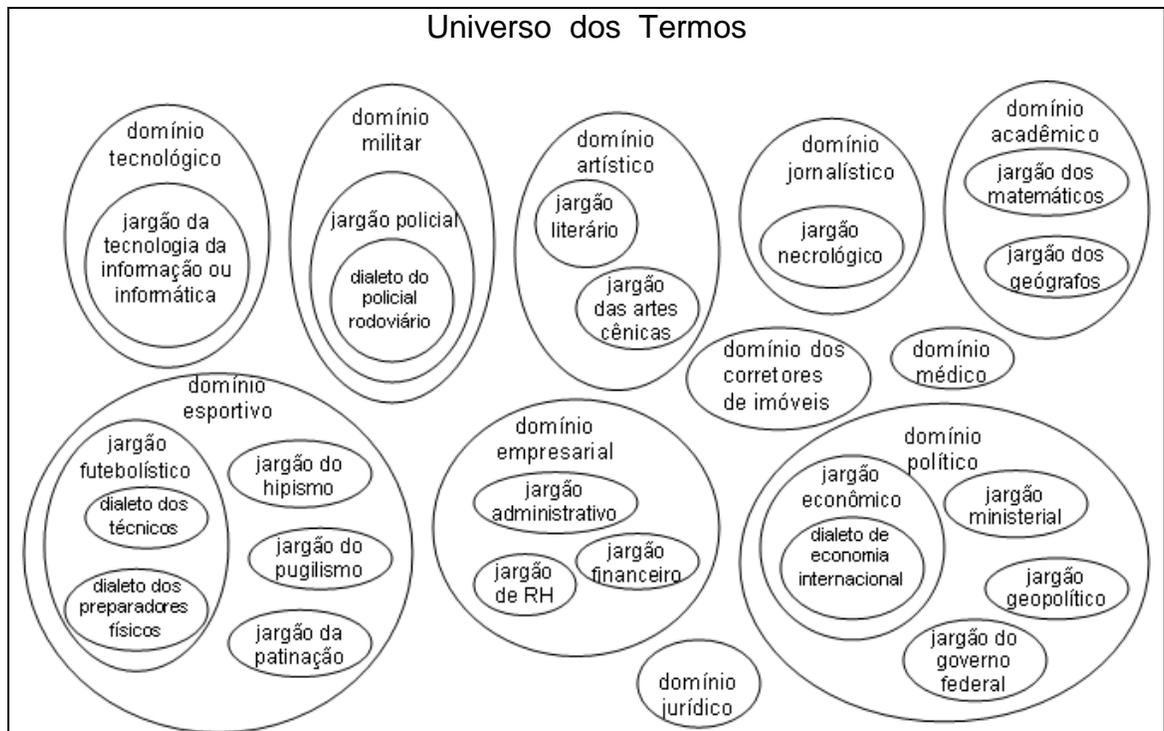


Figura 11: Universo dos Termos e os Domínios de Conhecimento.

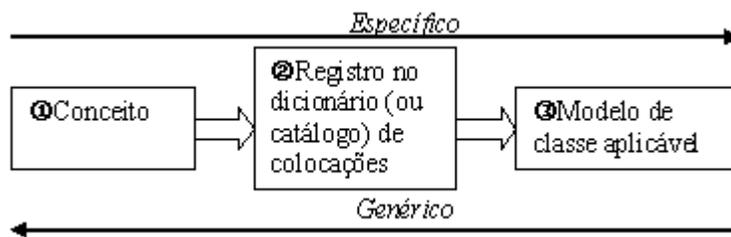
O aspecto, citado anteriormente, sobre a pluralidade dos sentidos de um mesmo termo de acordo com as especificidades da área em que ele é utilizado, poderia aparecer aqui representado como intersecções entre estes domínios. Segue uma demonstração de como isto ocorre na prática: “**Implantar**. 1 - Use **implantar**, livremente, no sentido médico: *O cirurgião implantou duas pontes de safena no deputado. / O dr. Hélio de Almeida implantou um novo rim no doente.* 2 - Por se tratar de modismo incontrolável, porém, **não use** implantar nos demais casos. Ele pode ser

substituído, sempre, com vantagem, por um dos verbos seguintes: adotar, introduzir, estabelecer, efetivar, instituir, criar, construir, levantar, edificar, formar, constituir, instalar, iniciar, apresentar, aplicar, firmar, consolidar, fixar, montar, aprontar, impor, imprimir, inculcar, desenvolver, instaurar, erguer, lançar, compor, organizar, consumir, assentar e ajustar. 3 - Escolha os substantivos correspondentes para substituir ***implantação***, outra palavra desgastada pelo uso abusivo e impreciso.” (MARTINS, 1997, p. 142 e 143).

APÊNDICE 1. EXEMPLOS DE APLICAÇÃO DO WORKFLOW

OBSERVAÇÃO: A notação adotada para os diagramas de classe foi a da ferramenta Rational da IBM.

I.



① **“Illegal.** É a situação de alguém e não a pessoa. Assim, não existem ‘imigrantes ilegais’” (MARTINS, 1997, p. 141 e 142).

ou

“Irregular. Irregular é a situação da pessoa e não ela própria. Assim: *Imigrantes em situação irregular serão deportados* (em vez de: *Imigrantes ‘irregulares’ serão deportados*).” (MARTINS, 1997, p. 151).

② **“imigrante substantivo**

ADJ. estrangeiro | Irlandês, Italiano, etc. | recente | primeira-geração, segunda-geração, etc. *A Primeira-geração de imigrantes sonhava retornar para o ‘lar’; seus filhos consideram a Inglaterra o seu lar.*

QUANT. influxo, onda

VERB + IMIGRANTE aceitar | deportar, retornar *barcos carregados com imigrantes que eram forçados a retornar*

IMIGRANTE + VERB chegar, entrar *imigrantes procurando entrar no país* | **chegar**
| **estabelecer** *Muitos imigrantes europeus se estabeleceram na Austrália.*

IMIGRANTE + SUBSTANTIVO **comunidade, família, grupo, população** |
trabalhador | **situação**

PREP. ~ **de** *Ela era a filha de imigrantes.*" (LEA, 2002, p. 395, tradução minha).

③



```

class Imigrante_Substantivo {
String adj[] = {nacionalidades(), "recente", "primeira-geração", "segunda-geração" };
String quant[] = {"influxo", "onda" };
String verb_imigrante[] = {"aceitar", "deportar", "retornar"};
String imigrante_verb[] = {"chegar", "entrar", "estabelecer"};
String imigrante_substantivo [] = {"comunidade", "família", "grupo", "população",
"trabalhador", "situação"};
String prep[] = {"de" };

private boolean ValidarAdj(String termo) {
    while(i < adj.length) {
        if adj[i].equal(termo)
            return true;
        }
    return false;
}

private boolean ValidarQuant(String termo) {
    while(i < quant.length) {
        if quant[i].equal(termo)
            return true;
        }
    return false;
}

private boolean ValidarVerb_Imigrante(String termo) {

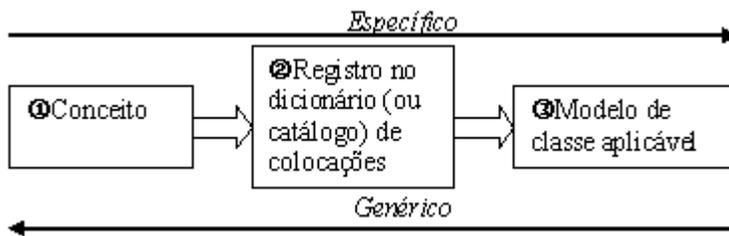
```

```

        while(i < verb_imigrante.length) {
        if verb_imigrante[i].equal(termo)
            return true;
        }
        return false;
    }
    private boolean ValidarImigrante_Verb(String termo) {
        while(i < imigrante_verb.length) {
        if imigrante_verb[i].equal(termo)
            return true;
        }
        return false;
    }
    private boolean ValidarImigrante_Substantivo(String termo) {
        while(i < imigrante_substantivo.length) {
        if imigrante_substantivo[i].equal(termo)
            return true;
        }
        return false;
    }
    private boolean ValidarPrep(String termo) {
        while(i < prep.length) {
        if prep[i].equal(termo)
            return true;
        }
        return false;
    }
}
}

```

II.



① **"Morder, mordida.** O termo aplica-se apenas a animais que têm dentes: *O cachorro mordeu o carteiro. / A mordida da piranha lhe tirou pedaços.* No caso, pode-se ainda usar **dentada**: *O homem deu uma dentada no inimigo.* Pernilongos, abelhas, escorpiões, aranhas, formigas, etc., **picam** ou **dão picadas e ferroadas**. Cobras, embora tenham presas, igualmente picam ou dão picadas: *A picada da urutu quase sempre é fatal. / A jararaca picou o menino.*" (MARTINS, 1997, p. 181).

②

"cachorro substantivo

ADJ. doméstico, familiar, estimação | selvagem | pedigree | pequeno *A madame estava beijando um cachorro pequeno. | amigável | bom, bem-comportado, bem-*

treinado, etc. | **mau** *Cachorro mau! O que você está fazendo aqui?* | **perigoso, assassino, selvagem** | **louco, raivoso** | **lutador, caçador, trabalhador** | **fazendeiro, guarda, guia, policial, pastor, farejador, rastreador** *Cães farejadores foram usados para encontrar as drogas.*

QUANT. matilha

VERB + CACHORRO **ter, guardar** *Os donos dos cachorros foram multados em £500 e proibidos de ter cachorros por cinco anos.* | **criar** *Aqueles cachorros foram criados para trabalhar como cães-guia para os cegos.* | **treinar** *Ele havia treinado o seu cachorro para sentar no bagageiro da sua bicicleta.* | **levar para passear, passear** *Eu estou levando o cachorro para passear.* | **castrar** *Nós não queremos filhotes, por isso castramos o cachorro.*

CACHORRO + VERB **latir, ganir** *O cachorro latiu furiosamente para o estranho. Os cachorros pequenos estavam ganindo nos meus tornozelos.* | **saltar, vagar, correr, caminhar** *O cachorro saltou para cima de mim e começou a lamber a minha mão. Cachorros abandonados vagam nas ruas a noite.* | **chegar** | **atacar, morder** | **roer** *O cachorro roeu um de meus sapatos.* | **farejar** *Um cachorro estava farejando em volta dos meus calcanhares.* | **arranhar** *O cachorro estava arranhando a porta para que o deixassem entrar.* | **sujar** *Donos que deixarem seus cachorros sujar a calçada serão multados.* | **atacar** *Um cachorro que atacou uma criança de cinco anos foi morto. A polícia está investigando.*

CACHORRO + SUBSTANTIVO **comida** | **coleira** | **sujeira, excremento, fezes** | **dono, treinador, guarda** | **luta** | **corrida** | **rastro**" (LEA, 2002, p. 235, tradução minha).

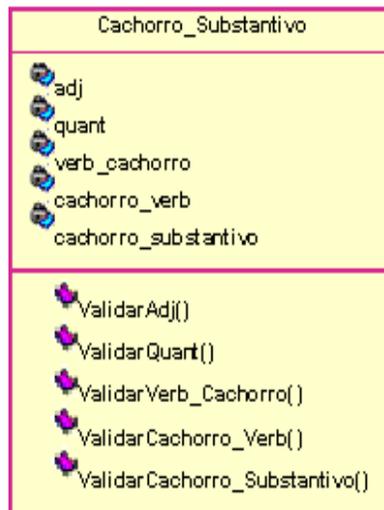
"**cobra** *substantivo*

ADJ. **letal** *Uma das cobras mais letais do mundo* | **venenosa** | **marítima, arborícola**

COBRA + VERB picar, chacoalhar *Ela havia sido picada por uma cobra enquanto estava caminhando sobre a relva. A cobra ergueu sua cabeça antes chacoalhar sua cauda.* | **deslizar, rastejar** *Uma pequena cobra verde rastejava sobre a estrada.* | **trocar a pele** | **enrolar-se** *A cobra enrolou-se.*

COBRA + SUBSTANTIVO picada | veneno (LEA, 2002, p. 724, tradução minha)

③



```

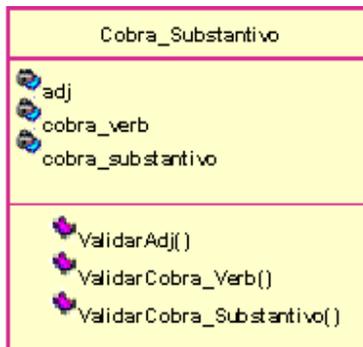
class Cachorro_Substantivo {
String adj[] = {"doméstico", "familiar", "estimação", "selvagem", "pedigree", "pequeno",
"amigável", sinônimos("bom", termo), "mau", "perigoso", "assassino", "selvagem", "louco",
"raivoso", "lutador", "caçador", "trabalhador", "fazendeiro", "guarda", "guia", "policia"
"pastor", "farejador", "rastreador" };
String quant[] = {"matilha"};
String verb_cachorro[] = {"ter", "guardar", "criar", "treinar", locução("levar", item, "para",
"passear"), "passear", "castrar" };
String cachorro_verb[] = {"latir", "ganir", "saltar", "vagar", "correr", "caminhar", "chegar",
"atacar", "morder", "roer", "farejar", "arranhar", "sujar", "atacar" };
String cachorro_substantivo [] = {"comida", "coleira", "sujeira", "excremento", "fezes",
"dono", "treinador", "guarda", "luta", "corrida", "rastros"};

private boolean ValidarAdj(String termo) {
while(i < adj.length) {
if adj[i].equal(termo)
return true;
}
return false;
}
private boolean ValidarQuant(String termo) {
while(i < quant.length) {
if quant[i].equal(termo)
return true;
}
return false;
}
private boolean ValidarVerb_Cachorro(String termo) {
  
```

```

        while(i < verb_cachorro.length) {
            if verb_cachorro[i].equal(termo)
                return true;
        }
        return false;
    }
    private boolean ValidarCachorro_Verb(String termo) {
        while(i < cachorro_verb.length) {
            if cachorro_verb[i].equal(termo)
                return true;
        }
        return false;
    }
    private boolean ValidarCachorro_Substantivo(String termo) {
        while(i < cachorro_substantivo.length) {
            if cachorro_substantivo[i].equal(termo)
                return true;
        }
        return false;
    }
}
}

```



```

class Cobra_Substantivo {
    String adj[] = {"letal", "venenosa", "marítima", "arboricola" };
    String cobra_verb[] = {"picar", "chacoalhar", "deslizar", "rastejar", locução("trocar", "a",
    "pele"), enrolar(pronome)};
    String cobra_substantivo [] = {"picada", "veneno"};

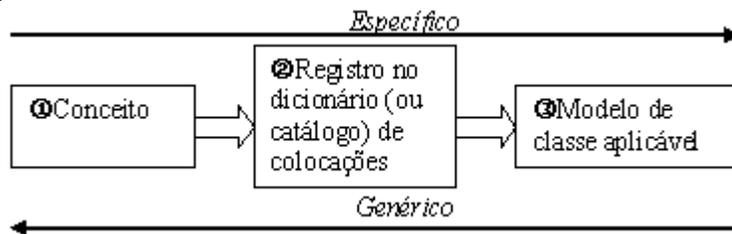
    private boolean ValidarAdj(String termo) {
        while(i < adj.length) {
            if adj[i].equal(termo)
                return true;
        }
        return false;
    }

    private boolean ValidarCobra_Verb(String termo) {
        while(i < cobra_verb.length) {
            if cobra_verb[i].equal(termo)
                return true;
        }
        return false;
    }

    private boolean ValidarCobra_Substantivo(String termo) {
        while(i < cobra_substantivo.length) {
            if cobra_substantivo[i].equal(termo)
                return true;
        }
        return false;
    }
}
}

```

III.



① “Operar. Quem opera é o cirurgião, e nunca o paciente. Assim: *Jogador viaja amanhã para ser operado na terça* (e não: *Jogador viaja amanhã e ‘opera’ na terça*). Igualmente: *General foi operado do coração em Cleveland* (e não: *General “operou” o coração em Cleveland*.” (MARTINS, 1997, p. 204).

②

"**cirurgião** substantivo

ADJ. **distinto, eminente | brilhante | pioneiro | chefe, consultor, sênior | qualificado | geral | cérebro, dentista, ocular, coração, ortopédico, plástico, transplantes, veterinário** *Um cirurgião plástico reconstruiu o seu nariz.*

VERB + CIRURGIÃO **ser, trabalhar como | tornar-se, qualificar-se como** *Ele irá se qualificar como um cirurgião.* | **consultar**

CIRURGIÃO + VERB **operar, remover algo** *o cirurgião que irá operá-lo. Cirurgiões removeram sua perna direita acima do joelho.*

FRASES **bisturi do cirurgião** *Muitas pessoas ficam apavoradas ao imaginar o bisturi do cirurgião.*" (LEA, 2002, p. 774, tradução minha)

"**paciente** substantivo

ADJ. aidético, canceroso, diabético, cardíaco, etc. | **mental, psicótico** | **terminal** | **particular** *Ele somente atende pacientes particulares.* | **velho** | **adulto, criança**

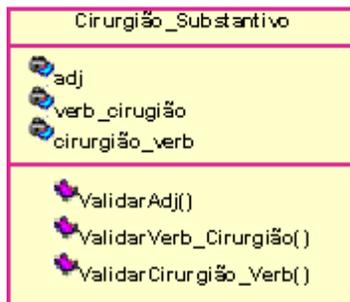
VERB + PACIENTE ter alta *pacientes aguardando para terem alta*

PACIENTE + VERB desenvolver (algo), ter (algo) *O paciente tem uma condição mental severa.* | **responder (a algo)** *Estes pacientes estão respondendo bem à nova droga.* | **melhorar**

PACIENTE + SUBSTANTIVO cuidados

PREP. ~ **com** *pacientes com doenças do fígado*" (LEA, 2002, p. 557, tradução minha)

③



```

class Cirurgião_Substantivo {
String adj[] = {"distinto", "eminente", "brilhante", "pioneiro", "chefe", "consultor", "sênior",
"qualificado", "geral", "cérebro", "dentista", "ocular", "coração", "ortopédico", "plástico",
"transplantes", "veterinário" };
String verb_cirurgião [] = {"ser", trabalhar("como"), tornar(pron),
Locução(Qualificar(pronome), ("como"))};
String cirurgião_verb[] = {"operar", remover(algo)};

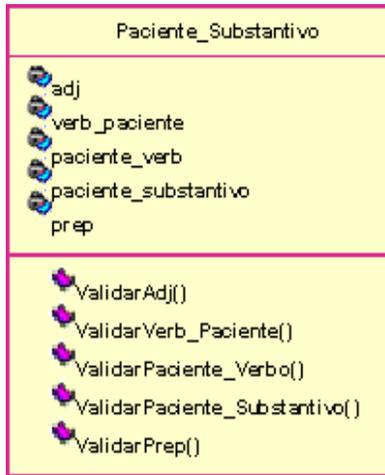
private boolean ValidarAdj(String termo) {
while(i < adj.length) {
if adj[i].equal(termo)
return true;
}
return false;
}

private boolean ValidarVerb_Cirurgião(String termo) {
while(i < verb_cirurgiao.length) {
if verb_cirurgiao[i].equal(termo)
return true;
}
return false;
}

private boolean ValidarCirurgião_Verb(String termo) {
while(i < cirurgiao_verb.length) {
if cirurgiao_verb[i].equal(termo)
return true;
}
return false;
}
}

```

```
}
}
```



```
class Paciente_Substantivo {
String adj[] = {doenças(adj), "mental" , "psicótico" , "terminal" , "particular" , "velho" ,
"adulto" , "criança" };
String verb_paciente [] = {locução("ter","alta")};
String paciente_verb[] = {desenvolver(algo), ter(algo), responder("a", algo), "melhorar"};
String paciente_substantivo[] = {"cuidados"};
String prep[] = {"com"};

private boolean ValidarAdj(String termo) {
while(i < adj.length) {
if adj[i].equal(termo)
return true;
}
return false;
}

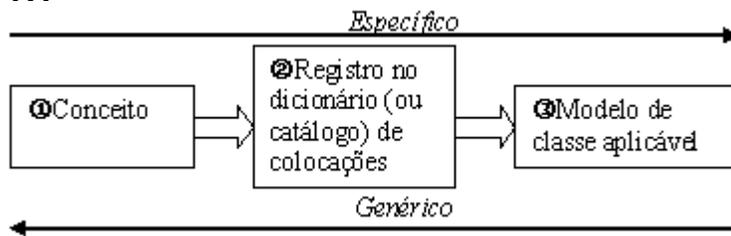
private boolean ValidarVerb_Paciente(String termo) {
while(i < verb_paciente.length) {
if verb_paciente[i].equal(termo)
return true;
}
return false;
}

private boolean ValidarPaciente_Verbo(String termo) {
while(i < paciente_verb.length) {
if paciente_verb[i].equal(termo)
return true;
}
return false;
}

private boolean ValidarPaciente_Substantivo(String termo) {
while(i < paciente_substantivo.length) {
if paciente_substantivo[i].equal(termo)
return true;
}
return false;
}

private boolean ValidarPrep(String termo) {
while(i < prep.length) {
if prep[i].equal(termo)
return true;
}
return false;
}
}
```

IV.



① “**Praticar**’. Fuja ao modismo e à impropriedade: não se ‘praticam’ juros, taxas, preços ou alíquotas, que podem, isso sim, ser cobrados, estabelecidos, fixados ou determinados.” (MARTINS, 1997, p. 230).

②

"**juro** *substantivo*

ADJ. **anual, mensal, etc. | composto, simples**

VERB + JURO **ganhar, fazer, receber | pagar | cobrar | estabelecer, fixar, determinar**

JURO + VERB **resultar, sustentar**

JURO + SUBSTANTIVO **taxa | pagamento | cobrança**

PREP. ~ **sobre** *para pagar juros sobre um empréstimo*

FRASES **uma taxa de juros** *hipoteca com uma taxa de juros fixa/flexível*" (LEA, 2002, p. 426, tradução minha).

"**taxa** *substantivo*

ADJ. **competitiva, menor, baixa, moderada, razoável** *Nós temos um grande estoque de veículos disponíveis para locação a taxas competitivas. As chamadas telefônicas estão com taxas menores após as 18 horas. | extorsiva, alta | excelente, boa | reduzida* *Neste orçamento, os juros foram calculados a uma taxa reduzida. |*

fixa, pós-fixada *Você pode optar por uma taxa pós-fixada para ter acesso ilimitado à Internet.* | **comum** | **futura** *Você pode preferir pagar pela taxa futura (= a taxa presente é a mais comum para pagamentos).* | **variável** | **anual, horária, semanal** | **base, básica, padrão** | **máxima** *pagando a taxa máxima de impostos* | **média** | **mercado** *Empréstimo a taxas de mercado* | **grupo, diferenciada** *Pergunte sobre as taxas diferenciadas para a sua categoria profissional.* | **bancária, hipotecária, tributária**

VERB + TAXA **determinar, fixar, estabelecer, estipular** | **aumentar, subir, elevar** | **cortar, abaixar, reduzir** | **segurar** *Nós iremos segurar estas taxas até 1 de Abril.* | **cobrar** *Eles cobram a taxa de juros de mercado.* | **pagar** | **dar, oferecer**

TAXA + VERB **ir para as alturas, aumentar, elevar** *Suas taxas horárias foram elevadas.* | **chegar, abaixar** | **flutuar** *As taxas de câmbio estão flutuando ferozmente.* | **aplicar a algo** *Taxa de juros padrão aplicadas a estes empréstimos.*

PREP. **a um, a** ~ *emprestando dinheiro a uma alta taxa de juros* | ~ **para** *a taxa média para um trabalhador não qualificado* | ~ **de** *um aumento na taxa de tributação*

FRASES **uma queda nas taxas de juros, um aumento nas taxas sobre as hipotecas, etc.** *um ponto percentual de aumento nas básicas taxas sobre empréstimos* | *investimentos com uma taxa de retorno segura dão uma boa taxa de retorno* " (LEA, 2002, p. 620, tradução minha).

"**preço** *substantivo*

ADJ. **exorbitante, alto, inflacionado, proibitivo, exagerado** | **baixo** *desenhar roupas por um preço baixo* | **atrativo, razoável, justo** *Nós vendemos ferramentas de qualidade por um preço justo.* | **bom** *Eu comprei este carro usado por um bom preço.* | **médio** | **anunciado, compra, pago** *Qual é o preço anunciado para esta casa? Você precisa pagar um depósito de dez por cento sobre o preço pago pela propriedade.* | **aluguel, venda** | **custo** *Eles estão liquidando o estoque de sapatos de*

verão pelo preço de custo. | inteiro, metade Crianças até a idade de dez anos viajam pela metade do preço. | mercado Este website tem a tabela dos preços de mercado para carros usados de todos os modelos. | aceitável preços aceitáveis para os imóveis | consumidor | commodity, alimento, casa, terra, propriedade, parcelado | eletricidade, energia, combustível, óleo, petróleo

VERB + PREÇO **comandar, alcançar, ir para** *As propriedades naquela área alcançaram preços extraordinariamente altos. | dar, cotar Eu pedi ao corretor para cotar os melhores preços para mim. | cobrar | estabelecer, fixar, determinar | aumentar, subir, crescer | trazer para baixo, cortar, abaixar, reajustar, empurrar para baixo, reduzir | subir em, aumentar em, crescer em Reajustar o preço do óleo. | abaixar em | diminuir*

PREÇO + VERB **subir, crescer, reajustar, disparar** *Os preços das casas foram reajustados em 5 por cento no último ano. Os preços dispararam durante a guerra. | cair, descer Se os preços caírem demais, os fazendeiros irão morrer à míngua. | ir de ... a ... , variar de ... para... , começar por Os preços vão de \$30 para o modelo standard a \$150 por uma versão de luxo. | variar de Os preços destes computadores variam de £1,300 a £2,000.*

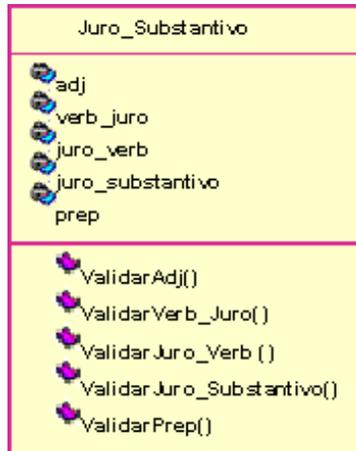
PREÇO + SUBSTANTIVO **nível, limites | aumento, crescimento | corte | mudança, movimento | guerra | etiqueta** *Eu tive um choque quando eu olhei para o preço na etiqueta. | lista | índice o índice geral de preços*

PREP. **por um/oleste** ~ *Estes alimentos ainda estão disponíveis, mas por um preço muito alto (= por um preço alto). Eu não posso pagar isto por este preço. | de* ~ *A estabilidade de preços tende a se manter por algum tempo.*

FRASES **um pingo/uma queda/uma redução no preço, um aumento /um crescimento no preço, pagar um preço alto (por algo)** *A equipe pagou um preço alto por estas deficiências na preparação. | localizado/pôr um preço em algo Você não pode pôr um preço na felicidade. | o preço da liberdade /sucesso, etc. (= as coisas desagradáveis a que você deve se sujeitar para ter liberdade, sucesso, etc.),*

um preço baixo a pagar (por algo) *O custo do prêmio da apólice é um preço baixo a pagar pela segurança que o seguro proporciona.* (LEA, 2002, p. 594, tradução minha).

③



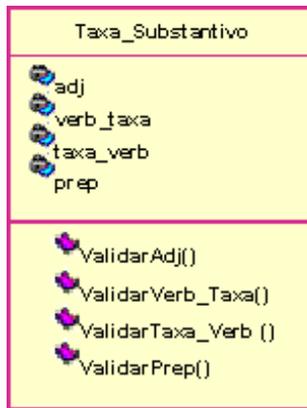
```

class Juro_Substantivo {
String adj[] = {períodos(), "composto", "simples" };
String verb_juro [] = {"ganhar", "fazer", "receber", "pagar", "cobrar", "estabelecer",
"fixar", "determinar"};
String juro_verb [] = {"resultar", "sustentar"};
String juro_substantivo [] = {"taxa", "pagamento", "cobrança"};
String prep [] = {"sobre"};
private boolean ValidarAdj(String termo) {
    while(i < adj.length) {
        if adj[i].equal(termo)
            return true;
    }
    return false;
}
private boolean ValidarVerb_Juro(String termo) {
    while(i < verb_juro.length) {
        if verb_juro[i].equal(termo)
            return true;
    }
    return false;
}
private boolean ValidarJuro_Verb(String termo) {
    while(i < juro_verb.length) {
        if juro_verb[i].equal(termo)
            return true;
    }
    return false;
}
private boolean ValidarJuro_Substantivo(String termo) {
    while(i < juro_substantivo.length) {
        if juro_substantivo[i].equal(termo)
            return true;
    }
    return false;
}
private boolean ValidarPrep(String termo) {
    while(i < prep.length) {
        if prep[i].equal(termo)
            return true;
    }
}
}
  
```

```

    }
    return false;
}
}

```



```

class Taxa_Substantivo {
String adj[] = {"competitiva", "menor", "baixa", "moderada", "razoável", "extorsiva",
"alta", "excelente", "boa", "reduzida", "fixa", "pós-fixada", "comum", "futura",
"variável", "anual", "horária", "semanal", "base", "básica", "padrão", "máxima",
"média", locução("de", "mercado"), "diferenciada", "bancária", "hipotecária",
"tributária" };
String verb_taxa [] = {"determinar", "fixar", "estabelecer", "estipular", "aumentar", "subir",
"elevar", "cortar", "abaixar", "reduzir", "segurar", "cobrar", "pagar", "dar", "oferecer"};
String taxa_verb [] = {locução("ir", "para", "as", "alturas"), "aumentar", "elevar", "chegar",
"abaixar", "flutuar", aplicar("a", algo) };
String prep [] = {locução("a", "uma"), "a", "para", "de"};

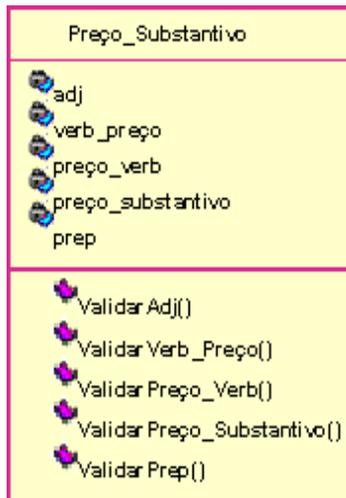
private boolean ValidarAdj(String termo) {
    while(i < adj.length) {
        if adj[i].equal(termo)
            return true;
        }
    return false;
}

private boolean ValidarVerb_Taxa(String termo) {
    while(i < verb_taxa.length) {
        if verb_taxa[i].equal(termo)
            return true;
        }
    return false;
}

private boolean ValidarTaxa_Verb(String termo) {
    while(i < taxa_verb.length) {
        if taxa_verb[i].equal(termo)
            return true;
        }
    return false;
}

private boolean ValidarPrep(String termo) {
    while(i < prep.length) {
        if prep[i].equal(termo)
            return true;
        }
    return false;
}
}
}

```



```

class Preço_Substantivo {
String adj[] = {"exorbitante", "alto", "inflacionado", "proibitivo", "exagerado", "baixo",
"atrativo", "razoável", "justo", "bom", "médio", "anunciado", "compra", "pago", "aluguel",
"venda", "custo", "inteiro", "metade", "mercado", "aceitável", "consumidor", "commodity",
"alimento", "casa", "terra", "propriedade", "parcelado", "eletricidade", "energia",
"combustível", "óleo", "petróleo"};
String verb_preço [] = {"comandar", "alcançar", ir(para), "dar", "cotar", "cobrar",
"estabelecer", "fixar", "determinar", "aumentar", "subir", "crescer", locução("trazer",
"para", "baixo"), "cortar", "abaixar", "reajustar", locução("empurrar", "para", "baixo")
, "reduzir", locução("subir", "em"), locução("aumentar", "em"), locução("crescer", "em")
, locução("abaixar", "em"), "diminuir"};
String preço_verb [] = {"subir", "crescer", "reajustar", "disparar", "cair", "descer",
ir("de", algo, "a", algo), variar("de", algo, "para", algo), começar("por", algo),
variari("de", algo)};
String preço_substantivo [] = {"nível", "limites", "aumento", "crescimento", "corte",
"mudança", "movimento", "guerra", "etiqueta", "lista", "índice"};
String prep [] = {locução("por", "um"), locução("por", "este"), "pelo", "de"};

private boolean ValidarAdj(String termo) {
while(i < adj.length) {
if adj[i].equal(termo)
return true;
}
return false;
}
private boolean ValidarVerb_Preço(String termo) {
while(i < verb_preço.length) {
if verb_preço[i].equal(termo)
return true;
}
return false;
}
private boolean ValidarPreço_Verb(String termo) {
while(i < preço_verb.length) {
if preço_verb[i].equal(termo)
return true;
}
return false;
}
private boolean ValidarPreço_Substantivo(String termo) {
while(i < preço_substantivo.length) {
if preço_substantivo[i].equal(termo)
return true;
}
return false;
}
}

```

```

}
private boolean ValidarPrep(String termo) {
    while(i < prep.length) {
        if prep[i].equal(termo)
            return true;
        }
    return false;
}
}
}

```

V.



① "Proferir que'. Alguém *profere* alguma coisa, mas não *profere que*." (MARTINS, 1997, p. 238).

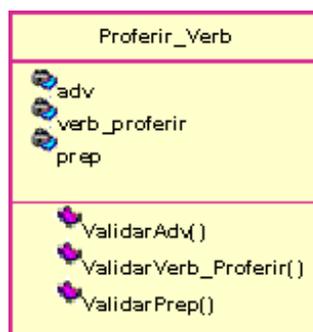
② "proferir verb

ADV. claramente, distintamente | corretamente, devidamente

VERB + PROFERIR poder, saber *Eu não sei como proferir o nome da cidade.* | ser difícil para

PREP. como 'Vamos embora' às vezes é proferido como 'vam'bora'. | como um *Ela proferia o 'l' como um 'r'.*" (LEA, 2002, p. 602, tradução minha).

③



```

class Proferir_Verb {
String adv[] = {"claramente", "distintamente","corretamente","devidamente"};
String verb_proferir [] = {"poder", "saber", locução("ser", "difícil", "para")};
String prep [] = {"como", locução("como", "um")};
private boolean ValidarAdv(String termo) {
    while(i < adv.length) {
        if adv[i].equal(termo)
            return true;
        }
    return false;
}
private boolean ValidarVerb_Proferir(String termo) {
    while(i < verb_proferir.length) {
        if verb_proferir[i].equal(termo)
            return true;
        }
    return false;
}
private boolean ValidarPrep(String termo) {
    while(i < prep.length) {
        if prep[i].equal(termo)
            return true;
        }
    return false;
}
}
}

```

VI.



① "Proliferar. E não 'proliferar-se'." (MARTINS, 1997, p. 238)

② "proliferar verb

ADV. **assexuadamente, sexuadamente**

VERB + PROLIFERAR **ser capaz de, poder, dever** *O cupim é capaz de proliferar na alvenaria. | falhar ao*

PREP. **por** *Muitos organismos unicelulares proliferam por divisão.* (LEA, 2002, p. 649, tradução minha).

③



```

class Proliferar_Verb {
String adv[] = {"assexuadamente", "sexuadamente"};
String verb_proliferar [] = {locução("ser", "capaz", "de"), "poder", "dever", falhar("ao")};
String prep [] = {"por"};
private boolean ValidarAdv(String termo) {
    while(i < adv.length) {
        if adv[i].equal(termo)
            return true;
    }
    return false;
}
private boolean ValidarVerb_Proliferar(String termo) {
    while(i < verb_proliferar.length) {
        if verb_proliferar[i].equal(termo)
            return true;
    }
    return false;
}
private boolean ValidarPrep(String termo) {
    while(i < prep.length) {
        if prep[i].equal(termo)
            return true;
    }
    return false;
}
}
}
  
```

VII.



① “**Quantia de dinheiro**’. Redundância: modernamente só se usa quantia para dinheiro. Por isso, não diga: *Havia ali grande ‘quantia’ de pessoas, de frutas, de objetos, mas quantidade.*” (MARTINS, 1997, p. 244).

②

"**quantia substantivo**

ADJ. **considerável, generosa, boa, bonita, robusta, alta, grande, desprezível, suficiente, apreciável, considerável, magnífica (ironia), significativa, substancial, elevada (informal)** *Isto lhe parecia ser uma quantia absurdamente alta a pagar pela unidade. Por seu primeiro livro ele recebeu a magnífica quantia de \$400. A equipe destinou uma quantia substancial para a caridade local. | astronômica, colossal, enorme, prodigiosa, exorbitante, larga, magnificante, pesada, recorde, surpreendente, vasta* *£200 era uma quantia astronômica em 1547. Dois anos atrás eles associaram-se ao clube por uma quantia recorde. | pequena, modesta, simbólica, razoável, simbólica, insignificante, diminuta* *A caridade paga uma quantia simbólica para arrendar o local. | derrisória, irrisória | máxima, mínima | média | absoluta | arredondada* *a quantia foi arredondada para \$10,000. | completa, global, total | crescente | acordada, fixada | declarada, especificada | anual, mensal, etc. | regular | garantida | capital, em espécie | livre de impostos | líquida | bruta* *Em caso de morte do cônjuge, a viúva receberá uma quantia bruta de \$900.000.*

VERB + QUANTIA **emprestar, ganhar, alcançar, crescer, recuperar** *Algumas das pinturas devem alcançar uma quantia considerável no leilão de hoje. Você teria que ir à corte para recuperar esta quantia. | pegar, receber | conceder, contribuir, dar, autorizar, oferecer, pagar, repassar* *O juiz concedeu-lhes uma quantia não revelada como ressarcimento pelos prejuízos. | investir, gastar | custar um projeto que custou uma quantia enorme aos cofres públicos | orçar, multar | chegar a um acordo (sobre)* *chegamos a um acordo sobre a quantia e eu lhe dei um cheque. | dever, lucrar* *O senhorio lucrou a quantia de \$100 sobre este arrendamento | igualar, equivaler* *O que os gangsters ofereceram a ele equivalia à quantia integral dos seus ganhos anuais. | exceder*

PREP. ~ **de** *A quantia de \$100" (LEA, 2002, p. 768, tradução minha).*

"fruta *substantivo*

ADJ. fresca | madura, verde | podre | confeitada, cristalizada, seca | enlatado | exótica, tropical | cítrica *frutas cítricas como limões e limões- galegos*

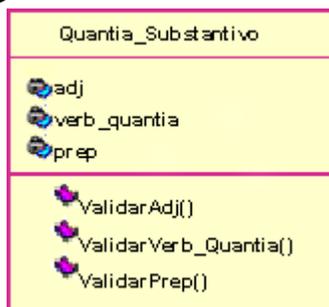
QUANT. quantidade *Havia ali grande quantidade de frutas| penca | cacho | dúzia, dezena, etc | quilo, grama, etc | caixa, saco, pacote, etc*

VERB + FRUTA **comer, ter | trazer** *O pássaro trouxe uma pequena fruta em seu bico para o filhote. | escolher | descaroçar, descascar, preparar | deleitar* *Após deleitarem-se com a fruta proibida foram expulsos do Éden. | colher* *O dono do pomar colhia as frutas maduras.*

FRUTA + SUBSTANTIVO **árvore | suco, salada | pedaço** *Termine a refeição com um pedaço de fruta fresca.*

FRASES **as primeiras frutas de algo** *As primeiras frutas da estação chegaram nesta semana às feiras livres.* "(LEA, 2002, p. 331, tradução minha).

③



```

class Quantia_Substantivo {
String adj[] = {"considerável", "generosa", "boa", "bonita", "robusta", "alta", "grande",
"desprezível", "suficiente", "apreciável", "considerável", "magnífica", "significativa",
"substancial", "elevada", "astronômica", "colossal", "enorme", "prodigiosa", "exorbitante",
"larga", "magnificente", "pesada", "recorde", "surpreendente", "vasta", "pequena", "modesta",
"simbólica", "razoável", "simbólica", "insignificante", "diminuta", "derrisória", "irrisória",
"máxima", "mínima", "média", "absoluta", "arredondada", "completa", "global", "total",
"crecente", "acordada", "fixada", "declarada", "especificada", periodos(), "regular",
"garantida", "capital", locução("em", "espécie"), locução("livre", "de", "impostos"),
"líquida", "bruta" };
String verb_quantia [] = {"emprestar", "ganhar", "alcançar", "crescer", "recuperar", "pegar",
"receber", "conceder", "contribuir", "dar", "autorizar", "oferecer", "pagar", "repassar",
"investir", "gastar", "custar", "orçar", "multar", locução("chegar", "a", "um", "acordo",
sobre(algo)), "dever", "lucrar", "igualar", "equivaler", "exceder" };
String prep [] = {"de"};
private boolean ValidarAdj(String termo) {
while(i < adj.length) {
if adj[i].equal(termo)

```

```

        return true;
    }
    return false;
}
private boolean ValidarVerb_Quantia(String termo) {
    while(i < verb_quantia.length) {
        if verb_quantia[i].equal(termo)
            return true;
        }
    return false;
}
private boolean ValidarPrep(String termo) {
    while(i < prep.length) {
        if prep[i].equal(termo)
            return true;
        }
    return false;
}
}
}

```



```

class Fruta_Substantivo {
    String adj[] = {"fresca", "madura", "verde", "podre", "confeitada", "cristalizada", "seca",
    "enlatado", "exótica", "tropical", "cítrica"};
    String quant[] = {"quantidade", "penca", "cacho", medidas_de_quantidade(),
    medidas_de_peso(), tipos_de_embalagem()};
    String verb_fruta [] = {"comer", "ter", "trazer", "escolher", "descaroçar", "descascar",
    "preparar", "deleitar", "colher"};
    String fruta_substantivo [] = {"árvore", "suco", "salada", "pedaço"};
    private boolean ValidarAdj(String termo) {
        while(i < adj.length) {
            if adj[i].equal(termo)
                return true;
            }
        return false;
    }
}
private boolean ValidarQuant(String termo) {
    while(i < quant.length) {
        if quant[i].equal(termo)
            return true;
        }
    return false;
}
private boolean ValidarVerb_Fruta(String termo) {
    while(i < verb_fruta.length) {
        if verb_fruta[i].equal(termo)
            return true;
        }
    return false;
}
private boolean ValidarFruta_Substantivo(String termo) {

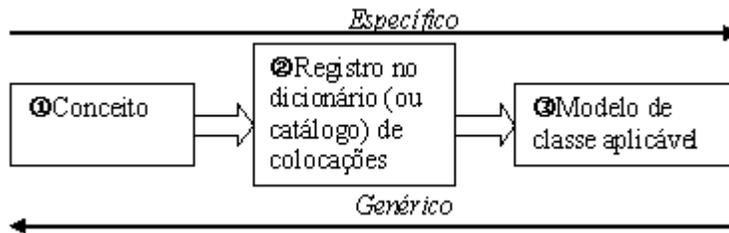
```

```

while(i < fruta_substantivo.length) {
  if fruta_substantivo[i].equal(termo)
    return true;
  }
return false;
}
}

```

VIII.



① “**Rapar, raspar.** Para cabelo, use apenas **rapar**, que significa cortar rente: *Rapou a barba. / Rapou o cabelo do calouro.* **Raspar** significa lixar, desbastar, tocar ou ferir de raspão: *O marceneiro raspou a madeira. / O carro raspou o portão. / Raspou a perna no chão.*” (MARTINS, 1997, p. 248).

②

rapar verb

ADV. **cuidadosamente, gentilmente** *Ela rapou cuidadosamente a barba do enteado.*

PREP. **de** *Sua barba foi rapada de seu rosto..*

RAPAR + SUBSTANTIVO **cabelo, barba**

FRASES **rapar algo** *Rapou a barba. / Rapou o cabelo do calouro*

raspar verb

ADV. **severamente** *A mesa havia sido vigorosamente raspada.*

PREP. **algo ~em** *Eu raspei meu braço em um espinho de rosa.* | **com** *Ela raspou seu rosto com suas unhas.*

RASPAR + SUBSTANTIVO **madeira**

FRASES **raspar algo** *O marceneiro raspou a madeira. / O carro raspou o portão. / Raspou a perna no chão.*

"**barba** *substantivo*

ADJ. **erçada, cerrada, cheia | bem-feita | com falhas, rala | tipo cavanhaque | ruiva, grisalha, branca, etc. | postiça, falsa**

VERB + BARBA **crescer | ter, ostentar, usar** *Ele ostentava uma barba bem-feita do tipo cavanhaque.* | **cortar, aparar, escanhoar, rapar | barbear | remover, afagar, puxar** *Rapou a barba. Jim afagava sua barba pensativo.*

BARBA + VERB **crescer** *Minha barba cresceu muito.* | **ter mechas/fios** *Sua barba tinha mechas/fios cinza.*

PREP. **com** ~ *Ela tinha medo de homens com barba.*

FRASES **uma barba de três-dias /sete-dias, etc., crescimento da barba"** (LEA, 2002, p. 60, tradução minha).

③

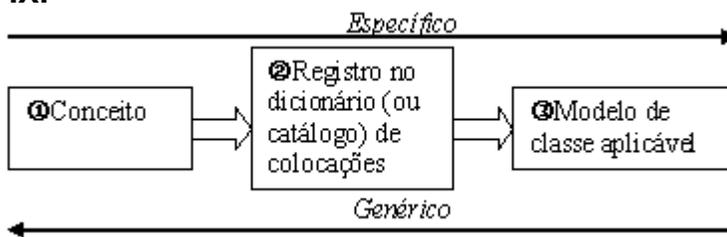
Barba_Substantivo
<ul style="list-style-type: none">  adj  verb_barba  barba_verb  prep
<ul style="list-style-type: none">  Validar Adj()  Validar Verb_Barba()  Validar Barba_Verb()  Validar Prep()

```

class Barba_Substantivo {
String adj[] = {"eriçada", "cerrada", "cheia", "bem-feita", locução("com", "falhas"), "rala",
locução("tipo", "cavanhaque"), tonalidades("cor", termo), "postiça", "falsa", "crescida",
locução("por", "fazer")};
String verb_barba [] = {"crescer", "ter", "ostentar", "usar", "cortar", "aparar", "escanhoar",
"rapar", "barbear", "remover", "afagar", "puxar"};
String barba_verb [] = {"crescer", locução("ter", "mechas"), locução("ter", "fios")};
String prep [] = {"com"};
private boolean ValidarAdj(String termo) {
    boolean done = false;
    while adj[i].equal(termo)
        return true;
}
private boolean ValidarVerb_Barba(String termo) {
    while(i < verb_barba.length) {
        if verb_barba[i].equal(termo)
            return true;
    }
    return false;
}
private boolean ValidarBarba_Verb(String termo) {
    while(i < barba_verb.length) {
        if barba_verb[i].equal(termo)
            return true;
    }
    return false;
}
private boolean ValidarPrep(String termo) {
    while(i < prep.length) {
        if prep[i].equal(termo)
            return true;
    }
    return false;
}
}
}

```

IX.



① **"Venda a.** E não venda **para:** *Venda de trigo ao Brasil. / A venda de mercadorias ao supermercado.*" (MARTINS, 1997, p. 297).

② **"venda substantivo**

ADJ. rápida *O preço está baixo para garantir uma venda rápida. | maioria absoluta a maioria absoluta das vendas de pinturas coube a investidores estrangeiros | ilegal*

a venda ilegal de álcool | de leilão, de feira de trocas, de banca, de garagem | em domicílio, compartilhada, etc.

VERB + VENDA **suspender, interditar, pausar, impedir, proibir, parar | fechar, fazer** *Fechar a venda significa forçar o comprador a se decidir. | marcar* *Marcou uma venda de pinturas sábado no Town Hall.*

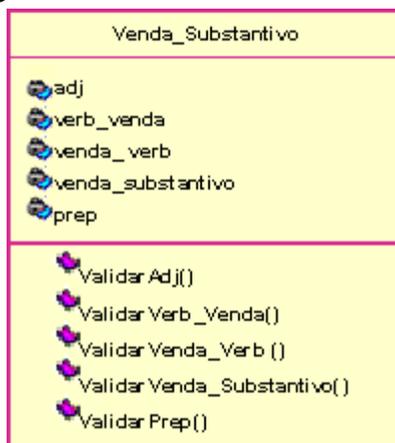
VENDA + VERB **fazer para algo, realizar por algo, totalizar** *A venda do último sábado fez £358 para as pesquisas do câncer. A venda de pinturas totalizou £250,000. | prosseguir, proceder | fracassar* *A venda da casa fracassou quando o comprador desistiu.*

VENDA + SUBSTANTIVO **acordo, contrato | preço | trâmite** *para maximizar o preço da venda*

PREP. **a ~** *Os novos selos estão agora a venda nos principais postos de correio. | de ~(algo) ~a (subs)* *Venda de trigo ao Brasil. / A venda de mercadorias ao supermercado.*

FRASES **condições de venda** *As condições de venda estão publicadas em torno da sala de leilão. | um contrato de venda, estar com as vendas liberadas* *Drogas estão com as vendas liberadas no clube. | lucros de uma venda* *Todos os lucros da venda do livro irão para caridade." (LEA, 2002, p. 676, tradução minha).*

③

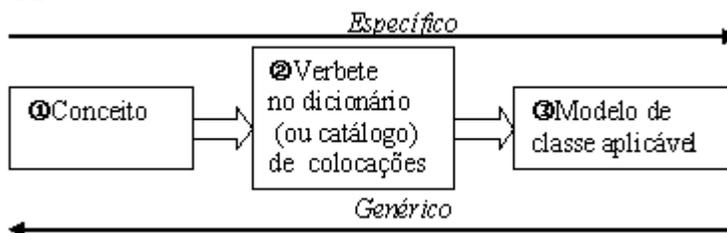


```

class Venda_Substantivo {
String adj[] = {"rápida", locução("maioria", "absoluta"), "ilegal", locução("de", "leilão"),
locução("de", "feira", "de", "trocas"), locução("de", "banca"), locução("de", "garagem"),
modos_de_negócio()};
String verb_venda [] = {"suspender", "interditar", "pausar", "impedir", "proibir", "parar",
"fechar", "fazer", "marcar"};
String venda_verb [] = {fazer(locução("por", algo)), realizar(locução("por", algo)),
"totalizar", "prosseguir", "proceder", "fracassar" };
String venda_substantivo [] = {"acordo", "contrato", "preço", "trâmite"};
String prep [] = {"a", locução(termo, de(algo), a(subs))};
private boolean ValidarAdj(String termo) {
while(i < adj.length) {
if adj[i].equal(termo)
return true;
}
return false;
}
private boolean ValidarVerb_Venda(String termo) {
while(i < verb_venda.length) {
if verb_venda[i].equal(termo)
return true;
}
return false;
}
private boolean ValidarVenda_Verb(String termo) {
while(i < venda_verb.length) {
if venda_verb[i].equal(termo)
return true;
}
return false;
}
private boolean ValidarVenda_Substantivo(String termo) {
while(i < venda_substantivo.length) {
if venda_substantivo[i].equal(termo)
return true;
}
return false;
}
private boolean ValidarPrep(String termo) {
while(i < prep.length) {
if prep[i].equal(termo)
return true;
}
return false;
}
}
}

```

X.



① **Verbos mais que errados.** 1 - Há verbos que, por questão de significado, não podem ser acompanhados de **que**. Em geral, eles equivalem aos diversos sentidos de **dizer** (verbos **dicendi**). Veja um exemplo: alguém defende uma idéia, uma

posição, mas nunca **defende que** alguma coisa se realize ou concretize. Assim, são erradas ou no mínimo impróprias as formas: acusar que, alertar que, antecipar que, apontar que, aprovar que, assumir que, citar que, comentar que, continuar que, defender que, definir que, denunciar que, desmentir que, difundir que, divulgar que, enfatizar que, indicar que, justificar que, mencionar que, narrar que, proferir que, prosseguir que, referir que, registrar que e relatar que.” (MARTINS, 1997, p. 307).

②

"referir verb

ADV. brevemente, de passagem *Ele referiu-se brevemente ao relatório.* | **especificamente | habitualmente** *Referia-se habitualmente à doença como ‘a epidemia verde’.* | **freqüentemente | geralmente, usualmente | algumas vezes | sempre | nunca | ironicamente**

VERB + REFERIR ser usado para *O termo ‘alexia’ é usado para se referir a qualquer desordem marcada pela perda da habilidade de leitura.*

PREP. ~a algo como *Ela sempre se referia a Ben como ‘aquele belo homem’.*" (LEA, 2002, p. 632, tradução minha).

"registrar verb

ADV. abertamente, dificilmente, raramente *Ela havia registrado abertamente a sua presença.* | **displicentemente** *Ele registrou displicentemente a chegada da mulher.* | **devidamente | legalmente | oficialmente, formalmente** *Registrei formalmente o meu compromisso de apresentar-me à junta militar quando completasse 18 anos.* | **automaticamente**

VERB + REGISTRAR falhar ao não *O almoxarife falhou ao não registrar a necessidade urgente da reposição daquela peça no estoque.* | **ser convocado a, ter**

que, dever *Você deve registrar o óbito dentro de três dias.* | **ser cogitado, ser autorizado a | falhar ao | recusar-se a**

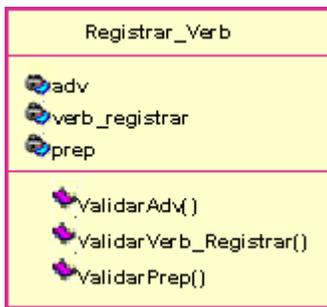
PREP. **como** *o número de pessoas que oficialmente se registraram como desempregadas* | **na** *A reclamação foi registrada na Ouvidoria.* | **com** *Deixe registrado com o médico a sua intolerância ao princípio ativo de algum dos remédios desta lista.* (LEA, 2002, p. 636, tradução minha).

③



```

class Referir_Verb {
String adv[] = {"brevemente", locução("de", "passagem"), "especificamente", "habitualmente",
"freqüentemente", "geralmente", "usualmente", locução("algumas", "vezes"), "sempre", "nunca",
"ironicamente"};
String verb_referir[] = {locução("ser", "usado", "para")};
String prep [] = {locução(termo, "a", algo, "como")};
private boolean ValidarAdv(String termo) {
while(i < adv.length) {
if adv[i].equal(termo)
return true;
}
return false;
}
private boolean ValidarVerb_Referir(String termo) {
while(i < verb_referir.length) {
if verb_referir[i].equal(termo)
return true;
}
return false;
}
private boolean ValidarPrep(String termo) {
while(i < prep.length) {
if prep[i].equal(termo)
return true;
}
return false;
}
}
}
  
```

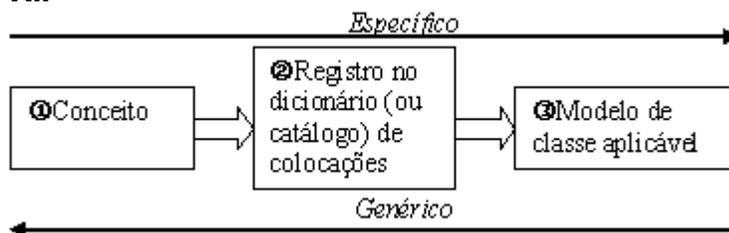


```

class Registrar_Verb {
String adv[] = {"abertamente", "dificilmente", "raramente", "displicentemente", "devidamente",
"legalmente", "oficialmente", "formalmente", "automaticamente"};
String verb_registrar[] = {locução("falhar", "ao", "não"), locução("ser", "convocado", "a"),
locução("ter", "que"), "dever", locução("ser", "cogitado"), locução("ser", "autorizado", "a"),
locução("falhar", "ao"), locução(recusar(pron), "a")};
String prep [] = {"como", "na", "com" };
private boolean ValidarAdv(String termo) {
while(i < prep.length) {
if prep[i].equal(termo)
return true;
}
return false;
}
private boolean ValidarVerb_Registrar(String termo) {
while(i < verb_registrar.length) {
if verb_registrar[i].equal(termo)
return true;
}
return false;
}
private boolean ValidarPrep(String termo) {
while(i < prep.length) {
if prep[i].equal(termo)
return true;
}
return false;
}
}
}

```

XI.



① "Vítima fatal". Fatal significa mortífero, que causa a morte, que traz ruína ou desgraça. Por isso, não existe a expressão 'vítima fatal': a vítima **recebe** a morte, e

não **a produz**. Fatal é um golpe, um tiro, um acidente, uma pancada, um choque, uma batida, e nunca a vítima.” (MARTINS, 1997, p. 312).

② "**vítima substantivo**

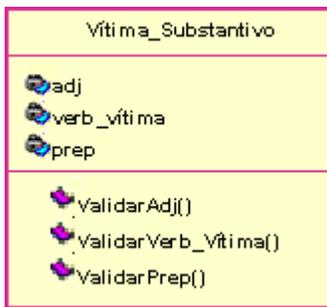
ADJ. infeliz, desassistida, inocente, pobre, desafortunada, insuspeita, teimosa.
Ele lesou muitas vítimas inocentes em milhões de libras. | fácil Turistas são vítimas fáceis de punquistas. | passiva | propenso a se tornar Em suas fantasias, ela estava sempre propensa a se tornar sua vítima | em potencial A vítima em potencial era escolhida por aparentar vulnerabilidade. | jovem, de idade avançada as jovens vítimas da guerra | de Aids, de câncer, de ataque de coração, da peste, de trauma | de acidente, da bomba, da quebra, do desastre, do terremoto, da penúria, da inundação O governo está enviando ajuda para as vítimas da inundação. | de seqüestro, de roubo | de sacrifício | da moda Ela é uma vítima da moda (= cede facilmente aos modismos, mesmo que a peça não lhe caia bem).

VERB + VÍTIMA retratar sb como *Durante o trajeto, ele tentava retratar a si próprio como vítima de uma sociedade egoísta. | fazer A quebra do trem fez sua décima vítima ontem quando o condutor morreu no hospital. | compensar uma conta para compensar as vítimas dos efeitos da poluição do ar | penalizar O corte nos benefícios dos desempregados vem penalizando suas vítimas ainda mais.*

PREP. de *Eles foram vítimas de uma brincadeira cruel.*

FRASES cair vítima de sb/algo *Lamentavelmente, ela caiu vítima de um senhorio inescrupuloso. | uma vítima de seu próprio sucesso A pequena firma tornou-se uma vítima de seu próprio sucesso quando não pôde atender a todos os seus pedidos em dia." (LEA, 2002, p. 852, tradução minha).*

③



```

class Vítima_Substantivo {
String adj[] = {"infeliz", "desassistida", "inocente", "pobre", "desafortunada", "insuspeita",
"teimosa", "fácil", "passiva", locução("propenso", "a", "se", "tornar"), locução("em",
"potencial"), "jovem", locução("de", "idade", "avançada"), locução("de", "Aids"),
locução("de", "câncer"), locução("de", "ataque", "do", "coração"), locução("da", "peste"),
locução("de", "trauma"), locução("de", "acidente"), locução("da", "bomba"), locução("da",
"quebra"), locução("do", "desastre"), locução("do", "terremoto"), locução("da", "penúria"),
locução("da", "inundação"), locução("de", "seqüestro"), locução("de", "roubo"), locução("de",
"sacrifício"), locução("da", "moda")};
String verb_vítima[] = {locução(retratar(alguém), "como"), "fazer", "compensar",
"penalizar"};
String prep [] = {"de"};
private boolean ValidarAdj(String termo) {
while(i < adj.length) {
if adj[i].equal(termo)
return true;
}
return false;
}
private boolean ValidarVerb_Vítima(String termo) {
while(i < verb_vítima.length) {
if verb_vítima[i].equal(termo)
return true;
}
return false;
}
private boolean ValidarPrep(String termo) {
while(i < prep.length) {
if prep[i].equal(termo)
return true;
}
return false;
}
}
}

```

APÊNDICE 2. GLOSSÁRIO

Análise do Discurso (teoria da) – “Um texto possui uma estrutura subjacente altamente elaborada que relaciona todo o seu conteúdo, atribuindo-lhe coerência. A essa estrutura dá-se o nome de estrutura discursiva, sendo ela objeto de estudo da área de pesquisa conhecida como Análise de Discurso. Diante da grande utilidade desse conhecimento para diversas aplicações de Processamento de Línguas Naturais, (...) a análise discursiva automática tem recebido muita atenção.” (PARDO, 2005, p. xiii).

ArgBank – repositório das estruturas argumentais dos 1.500 verbos mais freqüentes do inglês. Esta base de conhecimento foi produzida automaticamente aplicando-se técnicas de aprendizado de máquina, fundamentadas em modelos estatísticos de aprendizado de relações discursivas. Deverá, no futuro, ser aplicada ao português falado no Brasil. (PARDO, 2005, p. 108).

Árvore de estrutura retórica – técnica muito popular na literatura da teoria lingüística computacional. Representa uma generalização das técnicas de aprendizado automático de análise gramatical que necessitam representar palavras como *bitstrings* derivadas de modelos estatísticos de *clustering* de palavras.

ATN – *Augmented Transition Networks*, ou redes de transição estendidas - campo de estudos no domínio da inteligência artificial cuja meta é adicionar bônus na eficiência da análise gramatical das linguagens naturais. Esta teoria postula uma análise recursiva formando uma rede das transições de estados finitos de máquina.

BNF – *Backus–Naur form* - uma metalinguagem utilizada para a definição da sintaxe de linguagens formais, legível tanto ao desenvolvedor quanto ao usuário. Esta linguagem é definida em termos de símbolos reais (chamados terminais) e outras metavariáveis, escritas entre os sinais "<" e ">". Além destas, há os símbolos " ::= "

representando "é definido como" e "|", separando alternativas; seguidas por um número na forma Backus-Naur.

Categorização, Colocação ou Colocabilidade? – os termos Categorização e Colocação, na visão da teoria lingüística, muitas vezes se confundem; ou, dependendo do ponto de vista do autor que os emprega, assumem significados distintos ou até mesmo conflitantes.

Confrontando as diversas definições para o termo **Categorização**; em NEVES (2007), por exemplo, a autora se refere à orientação do OCD como “eminente categorizadora” no trecho transcrito a seguir: “**Categorização** Vejamos, por outro lado, um dicionário centrado na colocabilidade das diversas palavras lexicais da língua (tipo em que não se inspirou nenhuma obra elaborada em português), a *Oxford Collocations Dictionary* (2002), que, com orientação eminentemente **categorizadora**, põe as classes de palavras lexicais como selecionadoras de combinatórias.”. Porém, na obra propriamente dita, esta orientação não se confirma, pois a única referência à categorização é feita na forma do termo “*category collocation*” num sentido totalmente diverso a este, conforme a minha tradução para o trecho desta obra a seguir: “Muitas das colocações existentes no dicionário podem ser chamadas de 'colocações de palavras', isto é, aquelas palavras determinadas que combinem, apenas e tão somente, com alguma outra específica: *small fortune* não pode ser alterada para *little fortune*, ainda que: *small* e *little* aparentem ser sinônimos. Essa é outra área da colocação que pode ser chamada '**category collocation**', em que uma palavra pode se combinar com qualquer outra palavra de um conjunto bem definido. Este pode ser um conjunto extenso, mas os seus membros são previsíveis por seguirem um padrão determinado, como por exemplo: todas as palavras usadas para designar nacionalidades, ou medidas de tempo. Para o registro *walk*, um dos grupos de colocações apresentado é '*three-minute, five-minutes, etc.*': em que o 'etc.' representa qualquer figura que possa ser substituída por '*three*' ou '*five*' nestas expressões.”(LEA, 2002, pag ix). Uma terceira definição para um termo semelhante aparece expressa em PARDO (2005, p. 103), ao se referir aos: **dicionários de subcategorização** “dicionários que especificam os

comportamentos dos verbos, isto é, os argumentos que exigem e como estes se realizam sintaticamente em uma sentença”. Entretanto, os sentidos que melhor traduzem este significado complementam-se nas seguintes definições vistas na Wikipédia: “A categorização é um mecanismo fundamental para a razão e a comunicação humana, estabelecendo bases para muitos dos mais importantes processos mentais, tais como a percepção, a representação, a linguagem, a lógica e a aprendizagem.”. Ou ainda: “**Categorização** é o processo pelo qual idéias e objetos são reconhecidos, diferenciados e classificados. Em linhas gerais, a categorização consiste em organizar os objetos de um dado universo em grupos ou categorias, com um propósito específico.”. Portanto, para o “propósito específico” deste estudo, admitiu-se para o termo categorizar o sentido mais abrangente de catalogar colocações, de uma maneira geral.

Os sentidos para o termo **colocação**, por sua vez, são convergentes nas definições expressas por LEA (2002) e BECHARA (1975), conf. as seguintes transcrições destas obras: “O que é colocação? Colocação é o modo como as palavras são combinadas dentro de uma linguagem para que a pronúncia e a escrita soem de forma natural.” (LEA, 2002, p. vii). “7— COLOCAÇÃO *Sintaxe de colocação* ou *de ordem* — é aquela que trata da maneira de dispor os termos dentro da oração e as orações dentro do período. A *colocação*, dentro de um idioma, obedece a tendências variadas, quer de ordem estritamente gramatical, quer de ordem rítmica psicológica e estilística, que se coordenam e completam. O maior responsável pela ordem favorita numa língua ou grupo de línguas parece ser a entonação oracional.” – (BECHARA, 1975, p. 322). Enfim, o sentido que melhor traduz o significado do termo colocação parece ser o de representar a ordem natural da pronúncia ou entonação dos termos dentro da oração.

Diagramas de casos de uso – em UML, estudo de um caso qualquer decomposto nas suas unidades básicas.

Diferenças entre Relações retóricas e semânticas – distinguem-se “(...) as relações retóricas das semânticas pela força argumentativa das relações: a retórica

tem força argumentativa, enquanto a semântica não. (...) a força argumentativa das relações retóricas é dada pelas relações intencionais que se estabelecem entre as proposições” (PARDO, 2005, p. 25). Segue exemplo ilustrativo de uma situação em que uma análise semântica seria insuficiente para inferir sobre as relações retóricas intencionais entre as proposições expressas nas sentenças: “20 - Faça textos **imparciais e objetivos**. Não exponha opiniões, mas fatos, para que o leitor tire deles as próprias conclusões. Em nenhuma hipótese se admitem textos como: *Demonstrando mais uma vez seu caráter volúvel, o deputado Antônio de Almeida mudou novamente de partido. Seja direto: O deputado Antônio de Almeida deixou ontem o PMT e entrou para o PXN. É a terceira vez em um ano que muda de partido. O caráter volúvel do deputado ficará claro pela simples menção do que ocorreu.*” (MARTINS, 1997, p. 17).

Estruturas argumentais dos verbos – indica quantos e quais são os possíveis argumentos que o verbo exige. É um dos fundamentos dos modelos para análise discursiva (PARDO, 2005, p. 25).

Itens léxicos – padrões que indicam superficialmente a estruturação textual e que podem, ou não, ser indícios da presença de uma marca de segmento.

Marca de segmento – delimitação das Proposições textuais simples.

Marcador Discursivo – são elementos coesivos formados de uma ou mais palavras que explicitam o relacionamento que existe entre as partes de um texto. Por exemplo, ao se encontrar o marcador “entretanto”, “contudo” ou “porém” conectando dois segmentos textuais, há grandes chances de haver uma relação retórica de oposição.

Marcadores discursivos, em oposição às palavras e frases indicativas, são, normalmente, utilizados uniforme e consistentemente em qualquer gênero ou domínio de textos, independente de interpretações. Constituem o principal mecanismo lingüístico para a detecção de relações retóricas.

Marcadores textuais – delimitam orações dentro de uma sentença. Marcadores textuais “fortes” são aqueles que claramente sinalizam relações retóricas; como, por exemplo, sinais de vírgula, ponto e vírgula e dois pontos, entre outros.

Modelo Essencial – “apresenta o sistema num grau de abstração completamente independente de restrições tecnológicas. Antes que um sistema seja implementado, é necessário conhecer a sua verdadeira essência, não importando saber se sua implementação vai ser manual ou automatizada, e nem mesmo que tipo de ‘hardware’ ou ‘software’ vai ser usado. É o que os autores da análise estruturada chamam de modelo lógico; no caso, o modelo essencial corresponde ao modelo lógico proposto, que é a denominação que a Análise Estruturada dá ao modelo lógico desejada para o sistema que ainda vai ser implementado. A Análise Estruturada recomenda também a elaboração do modelo lógico atual, que é a denominação por ela dada ao modelo lógico do sistema existente, o qual vai ser substituído pelo sistema em desenvolvimento. A Análise Essencial não recomenda a elaboração do modelo lógico atual a não ser em casos muito particulares e, assim mesmo, parcialmente, apenas quando for necessário para auxiliar o entendimento do sistema a ser desenvolvido”. (POMPILHO, 2002, p.148).

Morfossintaxe – especificação das classes gramaticais das palavras.

MXPOST – assinalamento morfossintático automático com *Part-Of-Speech-Tags* treinado a partir de grandes corpora anotados manualmente, utilizando o modelo estatístico fundamentado no Princípio da Máxima Entropia para solução de ambigüidades nas linguagens naturais. Para este treinamento, utilizam-se estes corpora anotados no aprendizado de cada probabilidade de distribuição ou regras. RATNAPARKHI (1996) discorre sobre testes realizados com este modelo sobre um corpus do *Wall Street Journal*. Transpondo este conhecimento adquirido para outro texto não-revisado, registrou-se um grau de precisão no assinalamento que, em seu estado-da-arte, chegou a atingir 96,6%. Qualquer outro grande corpus anotado poderia ter sido utilizado para treinar este modelo com graus equivalentes de precisão, inclusive algum que contivesse palavras desconhecidas. As exigências de

desempenho de muitas tarefas de Linguagem Natural não podem prescindir da precisão de assinalamento deste modelo. Combinando as diferentes formas de informação contextual, esta técnica mostra a flexibilidade exigida para a modelagem lingüística.

Objeto – base da OO (*Object Oriented*) na Tecnologia da Informação. Consistem de modelos (abstrações) de objetos reais. Eles preservam as características essenciais de um objeto real: o seu estado e o seu comportamento. Qualquer objeto real pode ser abstraído para estudar o seu estado e comportamento, como no exemplo da Figura 12.

SUPERCLASSE	OBJETO (substantivo)	COMPORTAMENTO (verbo na 3ª pessoa do Presente do Indicativo)	ESTADO (adjetivo)
Zoologia	macaco	Dorme, come	Faminto
Botânica	banana	Cresce, amadurece	madura

Figura 12: Representação de entidades reais como classes de palavras no paradigma da Orientação a Objetos.

Parser – analisador gramatical, cujo desempenho melhora significativamente ao ser dotado de algum conhecimento semântico. As tecnologias deverão evoluir naturalmente para adquirir estas capacidades.

Part-of-speech Tagging – (*POS Tagging* ou *POST*) – "Também chamada de 'etiquetagem' gramatical, consiste no processo de rotular cada palavra em um texto de acordo com a sua classe ou categoria gramatical particular; levando em conta também o contexto na qual ela está inserida; isto é, o relacionamento dela com termos adjacentes e palavras relacionadas numa frase, sentença ou parágrafo. Uma forma simplificada de entender esta teoria, é por meio da analogia com o processo de aprendizado de crianças em idade escolar; e como elas aprendem a identificar as palavras como: substantivos, verbos, adjetivos, advérbios, pronomes, preposições, conjunções, interjeições e artigos. *Part-Of-Speech Tagging* constitui parte indissociável do domínio do Processamento de Linguagem Natural." (WIKIPEDIA -

tradução minha). A figura 13 ilustra o relacionamento entre alguns destes rótulos e as classes de palavras correspondentes na língua inglesa.

Number	Tag	Description
1.	CC	Coordinating conjunction
2.	CD	Cardinal number
3.	DT	Determiner
4.	EX	Existential there
5.	FW	Foreign word
6.	IN	Preposition or subordinating conjunction
7.	JJ	Adjective
8.	JJR	Adjective, comparative
9.	JJS	Adjective, superlative
10.	LS	List item marker
11.	MD	Modal
12.	NN	Noun, singular or mass
13.	NNS	Noun, plural
14.	NNP	Proper noun, singular
15.	NNPS	Proper noun, plural
16.	PDT	Predeterminer
17.	POS	Possessive ending
18.	PRP	Personal pronoun
19.	PRP\$	Possessive pronoun
20.	RB	Adverb
21.	RBR	Adverb, comparative
22.	RBS	Adverb, superlative
23.	RP	Particle
24.	SYM	Symbol
25.	TO	to
26.	UH	Interjection
27.	VB	Verb, base form
28.	VBD	Verb, past tense
29.	VBG	Verb, gerund or present participle
30.	VBN	Verb, past participle
31.	VBP	Verb, non-3rd person singular present
32.	VBZ	Verb, 3rd person singular present
33.	WDT	Wh-determiner
34.	WP	Wh-pronoun
35.	WP\$	Possessive wh-pronoun
36.	WRB	Wh-adverb

Figura 13: Lista alfabética das *part-of-speech tags* utilizadas em um Projeto chamado *Penn Treebank*.

Disponível em: <http://www.ling.upenn.edu/courses/Fall_2003/ling001/penn_treebank_pos.html>.

PLN – Processamento de Linguagem Natural: como é chamada a área da computação que se dedica a estudar linguagens de seres humanos, em oposição às linguagens artificiais, ou linguagens de programação.

Princípio da máxima entropia – de aplicação corrente no PLN, atua na solução de ambigüidades: "O princípio da máxima entropia é um método de análise da informação qualitativa disponível, com o propósito de determinar uma única distribuição de probabilidades epistêmica (origem, natureza e limite do conhecimento). Isto implica em que a distribuição minimamente preconcebida que codifica certas informações disponíveis é o que maximiza a entropia da informação ('entropia (...) na teoria da informação e em ciência de computador, é uma medida do conteúdo de informação de uma mensagem, avaliada por sua incerteza. ' (Michaelis). O princípio da máxima entropia faz uso explícito da informação anteriormente obtida. Esta é uma alternativa para os métodos de inferência clássicos baseados em modelos estatísticos. Sendo uma técnica genérica de aprendizado de máquina para resolver problemas de ambigüidades em processamento do *framework* de linguagem natural, o modelo probabilístico da máxima entropia oferece uma forma limpa de combinar diversas partes de evidências contextuais, com o propósito de estimar a probabilidade de certas classes lingüísticas ocorrerem num determinado contexto. Uma vantagem adicional do *framework* da máxima entropia é que a sua implementação de software é altamente re-usável. A teoria do *framework* da máxima entropia é independente de qualquer tarefa de linguagem natural. Na prática, desde que não tenha sido construído para atender a alguma necessidade específica, o software desenvolvido para uma tarefa pode ser reutilizado para outras tarefas que são implementadas sobre este framework. Uma simples implementação de software (C++ , Java, etc.) pode treinar todos demais modelos de máxima entropia para aumentar a portabilidade em outras plataformas." (WIKIPEDIA - tradução minha).

PROFILE – termo usado para se referir às particularizações feitas pelo consultante no decorrer da utilização de um material de apoio.

Retórica (relação) – forma de expressão das intenções no discurso.

Semântica – “é o estudo da significação dos vocábulos e das transformações de sentido por que estes mesmos vocábulos passam. No decorrer de sua história nem sempre o vocábulo guarda seu sentido etimológico, isto é, originário. Por motivos variadíssimos o sentido ultrapassa os limites de sua primitiva ‘esfera semântica’ e assume valores novos.” (BECHARA, 1975, p.340).

Tagger – etiquetador morfossintático.

Tesouro (Thesaurus) – "Também conhecido como dicionário de idéias afins, é uma lista de palavras com significados semelhantes, dentro de um domínio específico de conhecimento. Por definição, um tesouro é restrito. Não deve ser encarado simplesmente como uma lista de sinônimos, pois o objetivo do tesouro é justamente mostrar as diferenças mínimas entre as palavras e ajudar o escritor a escolher a palavra exata. Tesouros não incluem definições, pelo menos muito detalhadas, acerca de vocábulos, uma vez que essa tarefa é da competência de dicionários. Um thesaurus eletrônico tem a função principal de sugerir a um usuário conjuntos de sinônimos e antônimos de palavras da língua. Com essa funcionalidade, disponibiliza-se para o usuário uma ferramenta de auxílio à escrita que lhe permite encontrar com eficiência a palavra que procura enquanto escreve seu texto." (WIKIPEDIA, Disponível em: <<http://pt.wikipedia.org/wiki/Tesouro>>).

UML – *Unified Modeling Language* – Linguagem de modelagem por intermédio de diagramas. Para o escopo deste trabalho, restringimos a sua aplicação aos Modelos de Classes e alguns Casos de Uso.

URML – derivada da XML (linguagem padrão nas diversas plataformas tecnológicas em que é necessária a representação de informações dentro das suas próprias estruturas de dados), oferece recursos no nível da anotação discursiva que suportam notação das árvores de estrutura retórica. Corpora anotados com esta linguagem aproximam a análise humana e automática. Isto é especialmente útil no caso da anotação discursiva voltada para a representação da estrutura retórica, pois ela é

freqüentemente muito complexa e os algoritmos de análise gramatical que ela representa deverão ser legíveis para humanos, prevendo situações em que estas anotações precisem ser revistas e analisadas. Outro benefício considerável é prover anotações não somente no nível da sentença, mas também ao nível do discurso. Isto dá o suporte necessário a uma representação que integre várias camadas de anotação, em que a primeira camada represente a análise automática, necessária para diferenciar os tipos de informação; e as camadas superiores se especializem na decisão sobre qual o relacionamento retórico será interpretado. URML provê ainda a capacidade de subespecificação, formato em que os tópicos da análise são representados de forma aberta para extensões e especializações, prevendo que componentes adicionais farão escolhas e aperfeiçoamentos posteriores com base no conhecimento adicional adquirido (REITTER, 2001).

REFERÊNCIAS BIBLIOGRÁFICAS

BECHARA, Evanildo. *Moderna Gramática Portuguesa*. 19. ed. São Paulo: Companhia Editora Nacional, 1975¹.

MARTINS, Eduardo. *Manual de Redação e Estilo de O Estado de São Paulo*. 3. ed. São Paulo: O Estado de S. Paulo, 1997.

LEA, Diana; CROWTHER, Jonathan; DIGNEN, Sheila. *Oxford Collocations Dictionary*. 1. ed. Oxford: Oxford University Press, 2002.

PARDO, T. A. S. *Métodos para Análise Discursiva Automática*. Tese (Doutorado em Ciências Matemáticas e da Computação) – Universidade de São Paulo, São Carlos, 2005.

NEVES, Maria Helena de Moura. Quem comanda o espetáculo da linguagem? *Revista Língua Portuguesa*, 2007. Disponível em <<http://revistalingua.uol.com.br/textos.asp?codigo=11476>>. Acesso em: 20 fev. 2008.

IBOPE. Tema: *Apenas 25% da população adulta dominam (sic) habilidades de escrita e leitura*. Pesquisa publicada em 26/05/2004. Disponível em: <<http://www.ibope.com.br>> seção: Biblioteca Área: *Papers*. Acesso em: 6 mar. 2008

MARTINS, R. T.; HASEGAWA, R.; NUNES, M. G. V.; MONTILHA, G.; OLIVEIRA JR, O. N. *Linguistic issues in the development of ReGra: a Grammar Checker for Brazilian Portuguese*, 1998.

RATNAPARKHI, A. *A Maximum Entropy Model for Part-Of-Speech Tagging – Univ. of Pennsylvania – Dept. of Computer and Information Science, Pennsylvania*, 1996.

POMPILHO, S. *Análise Essencial*. 1. ed. Rio de Janeiro: Editora Ciência Moderna, 2002.

REITTER, D.; STEDE, M. *Step by step: underspecified markup in incremental rhetorical analysis – Media Lab Europe – Adaptive Speech Interfaces Manfred Stede – University of Potsdam – Dept. of Linguistics, Potsdam*, 2001.

WIKIPEDIA, a enciclopédia livre. Disponível em: <<http://wikipedia.org>>.

¹ Há edições mais atualizadas. Porém, graças à estabilidade nas normas gramaticais vigentes desde então, esta edição atendeu aos objetivos deste estudo.