



UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, LETRAS E CIÊNCIAS HUMANAS
DEPARTAMENTO DE LETRAS MODERNAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS LINGÜÍSTICOS
E LITERÁRIOS EM INGLÊS

GUILHERME FROMM

**VoTec: a construção de vocabulários eletrônicos para
aprendizes de tradução**

SÃO PAULO

2007

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, LETRAS E CIÊNCIAS HUMANAS
DEPARTAMENTO DE LETRAS MODERNAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS LINGÜÍSTICOS E
LITERÁRIOS EM INGLÊS

**VoTec: a construção de vocabulários eletrônicos para
aprendizes de tradução**

Guilherme Fromm

Tese apresentada ao Programa de Pós-Graduação em Estudos Lingüísticos e Literários em Inglês do Departamento de Letras Modernas da Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo, para a obtenção do título de Doutor em Letras.

Orientador: Profa. Dra. Stella Esther Ortweiller Tagnin

SÃO PAULO

2007

FOLHA DE APROVAÇÃO

Guilherme Fromm

VoTec: a construção de vocabulários eletrônicos para aprendizes de tradução

Tese apresentada à Faculdade de Filosofia,
Letras e Ciências Humanas da Universidade
de São Paulo para obtenção do título de
Doutor.

Área de Concentração: Estudos
Linguísticos e Literários em Inglês

Aprovado em: _____

Banca Examinadora

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Ao meu pai, Oswaldo, que sempre me incentivou a estudar.

AGRADECIMENTOS

Gostaria de agradecer, em primeiro lugar, a minha orientadora, profa. Dra. Stella E. O. Tagnin pela enorme paciência e por aceitar meu jeito “não muito ortodoxo” de trabalho.

Um agradecimento especial ao amigo Daniel Henrique Passerini por todo tipo de ajuda, direta ou indireta, que vem me prestando nos últimos três anos.

Ao amigo Roberto Böker, sem o qual a primeira pesquisa, com os tradutores, não teria acontecido.

Aos amigos e professores universitários que me ajudaram sobremaneira com a segunda pesquisa: Paulo Stella, Luciana Ginezi, Leila Darin, Glória R. L. Sampaio e Cristina A. E. Kindermann.

Aos amigos da Uniban, Denilson Schäfer e Sidnei Aparecido dos Reis, essenciais para ajudar na reflexão da construção do campo da computação e na ajuda com questões de formatação do trabalho.

Às professoras Maria Aparecida Barbosa e Sandra M. Aluísio, pelas valiosas sugestões oferecidas na banca de qualificação.

RESUMO

FROMM, G. **VoTec: a construção de vocabulários eletrônicos para aprendizes de tradução**. 2007. 210 f. Tese (Doutorado) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2007.

O presente trabalho visa descrever a construção de um Vocabulário Técnico Online, denominado VoTec e disponibilizado em uma página da Internet, para aprendizes de Tradução. São discutidos vários modelos de obras similares que serviram de base para o desenvolvimento do mesmo. O VoTec é uma ferramenta que se vale de *corpora* técnicos para a construção de seus verbetes e de um banco de dados (ambos exaustivamente descritos) para o seu funcionamento. São apresentados, também, o resultado de duas pesquisas realizadas com tradutores: uma com profissionais da área, para colher suas opiniões sobre obras terminográficas em geral, e outra com aprendizes, para colher suas opiniões após o uso da ferramenta. A inovação proposta pela ferramenta é a possibilidade de customização de seus modos de apresentação na tela e as formas de busca no banco de dados. O **modo de apresentação** possibilita duas formas de visualização da microestrutura: a **normal**, formatada como os dicionários impressos, e a **descritiva**, uma forma própria dos dicionários eletrônicos, em que a microestrutura é descrita de forma hierárquica. A **consulta ao banco** de dados pode ser feita de três modos diferentes: total, tradutor e modular. A consulta **total** disponibiliza praticamente todos os campos existentes no banco de dados; a do **tradutor** apresenta os campos que se destacaram na pesquisa realizada com os profissionais da área; a **modular** é montada de acordo com as necessidades tradutórias do consulente, isto é, ele pode visualizar ou não cada campo disponível no banco de dados. Além dessas consultas ligadas ao banco, a página do VoTec disponibiliza, também, links para pesquisa externa em páginas pertinentes, como dicionários e enciclopédias online.

Palavras-chave: Terminologia, Terminografia, Lingüística de Corpus, Ensino da Tradução, Tecnologia da Informação.

ABSTRACT

FROMM, G. **VoTec: the construction of electronic vocabularies for translation learners**. 2007. 210 f. Thesis (Doctoral) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2007.

The present work aims at describing the construction of an Online Technical Vocabulary, to be known as VoTec and available on an Internet site for Translation learners. Many similar models are discussed and taken into account in its development. The VoTec is a tool that relies on a technical corpora and a data bank (both exhaustively described) to build up the entries available at the site. The results of two researches with translators are also shown: one with the professionals in the area, to collect their opinions about terminographical reference works in general, and another one with the learners, to elicit their opinions on the tool after they had their first contact with it. The innovation proposed by the tool is the possibility of customizing the display of its screens and the searches the data bank make available. The **screen display** offers two ways to view the microstructure: the **normal** one, similar to a paperback dictionary model, and the **descriptive** one, typical of electronic dictionaries, where the microstructure is described in a hierarchical way. The **data bank** can be **searched** in three different ways: total, translator and modulated. The **total** search shows almost all the fields available in the data bank; the **translator** one shows the fields that were highlighted by the professional translators in one of the researches; the **modulated** one is built up according to the user's translation necessities, that is, he/she can choose which field or fields of the data bank will be shown. Besides these searches which access the data bank, the VoTec page also offers links to external searches in related webpages, such as online dictionaries and encyclopedias.

Keywords: Terminology, Terminography, Corpus Linguistics, Translation Training, Information Technology.

Sumário

1. Introdução	12
2. Sobre o aprendiz de tradução: algumas considerações.....	16
2.1 O aprendiz de tradução	16
2.2 O Ensino da Tradução no Brasil e no mundo	17
2.3 O aprendizado em tradução	20
3. Obras de consulta disponíveis	23
3.1 Obras tradicionais	25
3.2 Novas tecnologias em consulta.....	28
3.3 Ferramentas de Tradução.....	31
3.4 Tradução e Terminologia: a necessidade da interface	33
4. A construção dos <i>corpora</i> de estudo.....	36
4.1 O que é um <i>corpus</i> ?	36
4.2 Os <i>corpora</i> desta tese	37
4.3 Ontologia/Taxonomia: a árvore de campo	38
4.4 A coleta dos textos.....	41
4.5 A coleta dos <i>corpora</i>	43
4.6 A análise dos <i>corpora</i>	44
4.6 Mais <i>corpora</i>	56
4.7 Em suma	57
5. A opinião de quem já chegou lá: os tradutores profissionais	59
6. Abordagem de modelos, a Ficha Terminológica e o banco de dados	77
6.1 A abordagem lexicográfica para a construção do banco de dados	77
6.2 Uma abordagem híbrida para a construção da página de consulta.....	78
6.3 Modelos de fichas terminológicas e bancos de dados	80
6.4 A inserção de dados pelo pesquisador	87
6.5 O Paradigma Pragmático (PP).....	90
6.6 A construção do Paradigma Definicional (PD)	92
6.7 Paradigma Informacional (PI)	102
6.8 Paradigma Semântico (PS)	103
6.9 Paradigma de Forma Equivalente (PFE)	104
6.10 Remissivas	105
6.11 Paradigma Enciclopédico (PE).....	105
6.12 O Enunciado Terminográfico resultante.....	106
7. O aprendiz e o VoTec	108

7.1 O aprendiz iniciante: usando a página de consulta.....	108
7.1.1 Apresentação da página e modo de consultar.....	108
7.1.2 Modos de exibição: normal e descritivo.....	110
7.1.3 Consultas internas: total, tradutor e modular.....	115
7.1.4 Consultas externas	118
7.1.5 A ajuda online.....	120
7.2 O aprendiz avançado: usando o banco de dados	127
8. Avaliação da ferramenta.....	130
8.1 As perguntas, suas motivações e a relevância para a tese	131
8.2 A mini-pesquisa com os professores	151
9. Considerações finais	158
Referências bibliográficas.....	164
APÊNDICES	170
APÊNDICE A. Questionário para os tradutores profissionais, em inglês.	170
APÊNDICE B. Pesquisa com os tradutores. Perguntas e gráficos de respostas.	174
APÊNDICE C. Relação dos dicionários aventados pelos tradutores	180
APÊNDICE D. Exemplo de Ficha Terminológica Não-computadorizada.	183
APÊNDICE E. Documento de requisitos para a construção do banco e da página.	184
APÊNDICE F. Ajuda online em inglês.	189
APÊNDICE G. Questionário com o público-alvo/professores. Versão completa. ..	197
APÊNDICE H. Resultados completos da pesquisa com o público-alvo.	201
ANEXO.....	214
ANEXO A. Organização das possíveis entradas de um dicionário bilíngüe.....	214

1. Introdução

A ferramenta básica de um tradutor, há muito tempo é, e acredito que continuará sendo, o dicionário. Como professor dos cursos de Letras e Tradução, no entanto, percebo que, cada dia mais, os professores trabalham menos o funcionamento desse tipo de obra. Consideram que os alunos, a priori, já sabem trabalhar, e bem, com um dicionário ou vocabulário e, portanto, esta tarefa de aprendizagem pode ser desconsiderada nesses cursos. Nada mais errôneo! O ensino fundamental ou médio não instrumentaliza, já há tempos, o aluno a trabalhar com obras de referência: o máximo que os professores comentam em sala de aula é a questão da ordem alfabética (e, talvez, semasiológica) presente nessas obras. Seu processo de escolha de palavras, sua disposição gráfica interna, a construção da microestrutura de cada verbete, tudo isso é passado em branco. Tem-se, como resultado para o aprendizado do futuro profissional de Tradução, enfoque primeiro desta tese, a parcial ignorância dos recursos que obras de referência, em especial de caráter terminológico, oferecem. O mundo globalizado e os clientes requerem, cada vez mais, tradutores conectados com as novas tecnologias. Esse profissional deve, portanto, aprender o uso de novas ferramentas numa velocidade muito rápida. O que acontece quando esse aprendizado está calcado numa base nada sólida em relação à primeira ferramenta com que ele deveria ter travado contato? Será que esse profissional consegue traduzir com a mesma rapidez que aqueles que tiveram um bom aprendizado (na escola ou por experiência própria) do uso de ferramentas eletrônicas de busca ou voltadas para a tradução? Ele conseguirá se destacar num mercado cada vez mais competitivo? Acredito que não.

Por essa razão, é importante proporcionar aos aprendizes de Tradução um primeiro contato com uma ferramenta terminológica informatizada que servirá para treiná-los no uso de dicionários (ou outras obras de referência) . Para tanto, esta tese apresenta o VoTec (Vocabulário Técnico Online), que pretende auxiliar não só esse aluno nos seus estudos iniciais sobre o uso de obras terminográficas, mas também os tradutores profissionais que já atuam no mercado.

O próximo capítulo pretende descrever quem é esse aprendiz no Brasil e seu perfil, discutindo o ensino de Tradução no país e seus possíveis currículos. São analisados modelos de ensino de tradução, partindo de uma abordagem tradicional para uma tendência mais comum nos textos atuais, que é o desenvolvimento de uma Competência Tradutória.

Um panorama das obras de consulta disponíveis para os tradutores é o cerne do terceiro capítulo. Parte-se do dicionário como fonte de consulta para o aluno que aprende uma língua estrangeira até chegar ao foco desta tese, que é o apoio essencial que uma obra terminográfica oferece aos tradutores. Analisam-se obras tradicionais, em papel, seguindo com as novas tecnologias disponíveis para os dicionários e vocabulários eletrônicos e citando outros tipos de programas que, ultrapassando os limites do dicionários, são bastante utilizados pelos tradutores na atualidade. Discuto a idéia que há a necessidade de uma maior interação entre as disciplinas Tradução e Terminologia porque, cada vez mais, o tradutor também desempenha o papel de terminólogo e, ocasionalmente, de terminógrafo.

Foram necessárias várias etapas para a construção do VoTec. A primeira, descrita no capítulo quatro, é a construção de um *corpus* de estudo. Diferente de muitas outras ferramentas eletrônicas disponíveis no mercado, que também se valem de obras já publicadas para a construção de seu banco de dados, esta baseia-se, para a elaboração

de sua macro- e microestruturas, apenas em *corpora* técnicos especialmente constituídos para esse fim. Analiso o significado atual do que é um *corpus*, a ciência ligada ao estudo e à construção de *corpora* – a Linguística de *Corpus*. Descrevo como se dá o planejamento para a construção dos dois *corpora* bilíngües que servem de base para a ferramenta, a construção de uma árvore de campo (onde discuto, também, questões sobre taxonomia/ontologia), as fontes textuais, o processo de levantamento dessas fontes, como os *corpora* são analisados e como podem ser reutilizados em novas empreitadas.

Antes de construir uma ferramenta que atenda bem aos aprendizes da área, decidi por uma pesquisa com os profissionais que já estão atuando no mercado. Através dessa pesquisa, desenvolvida no capítulo cinco, averiguo quais suas preferências na composição da macroestrutura e, especialmente, da microestrutura de cada verbete.

Construir uma ferramenta informatizada, nos dias de hoje, pressupõe construir, também, um banco de dados. Antes de partir para um modelo qualquer de banco, discuto, no capítulo seis, uma abordagem lexicográfica para a construção desse banco, ou seja, as possibilidades de campos que farão parte de sua estrutura e que se inter-relacionarão. Há diversas maneiras, também, de se apresentar uma página de consulta. Proponho, visto a grande quantidade de modelos disponíveis e as respostas obtidas pelo questionário do capítulo cinco, uma abordagem híbrida, em que o banco oferece, através da página de consulta proposta, mais de uma possibilidade de consulta para o usuário. Ainda neste capítulo, discuto a base para a construção de um banco de dados na área, que é a Ficha Terminológica. Na sua composição são indicados os campos que constam do banco. Descrevo como os dados, partindo do conceito de Ficha Terminológica, são inseridos no banco e como a microestrutura é construída, passo a passo, por meio do desenvolvimento de seus Paradigmas: Pragmático, Definicional, Informacional,

Semântico, Forma Equivalente, Enciclopédico. Também apresento o desenvolvimento do sistema de remissivas, bem como a construção do Enunciado Terminográfico resultante.

O VoTec é apresentado, formalmente, no capítulo sete. Lá estão descritos os modos de consulta - normal e descritivo - da página disponível na Internet e os tipos de consulta de que os aprendizes dispõem, uma das inovações propostas pelo trabalho. Essas consultas pretendem representar diferentes tipos de contatos com uma obra terminográfica pelos quais o aluno iniciante na área pode passar: a *total* copia o modelo dos dicionários tradicionais, a consulta *tradutor* apresenta uma microestrutura baseada nas opiniões dos profissionais levantadas no capítulo cinco e a consulta *modular* permite a manipulação da página, onde o consulente constrói a microestrutura de acordo com suas necessidades. Descreverei, também, com a Ajuda Online, o modo de usar a página. Mas como o projeto não se restringe aos iniciantes na área de Tradução, discuto, também, de que modo o aluno de nível mais avançado pode se beneficiar de um aprendizado complementar em Terminografia através do manuseio, sob supervisão docente, do banco de dados.

Com a intenção de avaliar o uso da página na Internet, apresento, no capítulo oito, uma pesquisa realizada com os alunos. Discuto, como no capítulo cinco, minhas motivações para as perguntas apresentadas no questionário a eles apresentado e como os resultados obtidos servem para corroborar o que foi produzido ou sugerir alterações em futuras versões da página. Analiso, também, a opinião dos professores que aplicaram a pesquisa e se acreditam que a ferramenta será útil para o aprendizado de tradução.

O último capítulo conta com minhas considerações finais, relacionando a construção da ferramenta com seu uso e sugestões de possíveis mudanças para melhorar o seu desempenho. Seguem-se as referências bibliográficas, apêndices e um anexo.

2. Sobre o aprendiz de tradução: algumas considerações

O objetivo deste capítulo é a apresentar uma revisão bibliográfica a respeito do público-alvo desta tese, o aprendiz de tradução, e discutir questões como: Quem é esse público? Quais são seus objetivos com o curso de tradução? Quais habilidades eles já dominam para fazer uma tradução? Quais novas habilidades precisam ser trabalhadas? Eles têm conhecimento de ferramentas (como dicionários eletrônicos, vocabulários online, memórias de tradução) disponíveis para os profissionais?

2.1 O aprendiz de tradução

Milton (2000) levanta uma série de questões referentes aos cursos de Tradução no Brasil e seus respectivos alunos. Muitas dessas questões foram por mim vivenciadas quando atuei como professor na área. O curso de Tradução é, quase que unanimemente, ligado à área de Letras Modernas (como as línguas alemã, francesa, italiana, espanhola e inglesa), e isso já indica um perfil do aluno egresso: muitos alunos parecem escolher o curso de Tradução como uma forma de melhorar suas habilidades lingüísticas na língua desejada para poderem dar melhores aulas; outros buscam uma nova fonte de renda, já que o curso os habilita para uma nova profissão e as escolhas nas grandes áreas de Letras e Lingüística se resumem, cada vez mais, ao ensino.

O ensino superior, no Brasil, também tende, nos últimos anos, num processo contínuo, a se deslocar do eixo Público para o Privado. A faculdade/universidade

privada, porém, não tem o mesmo compromisso que a pública com a pesquisa. Isso se reflete, também, nos cursos de Tradução e seus currículos. Pouca pesquisa se dá sobre como organizar um bom currículo de Tradução. Essa preocupação é comum aos pesquisadores da área, tanto no Brasil quanto no exterior, como bem demonstram alguns textos de Pym (1998, 2002). O que o aluno de Tradução, então, deveria ter como base de seu currículo?

Antes de mais nada, no Brasil, devemos atentar ao panorama que Milton (2000, p. 83-105) traçou das possibilidades de formação de tradutores no país. Existem aqueles que fazem uma graduação:

- a. em uma área qualquer e depois partem para um curso de especialização em Tradução;
- b. em uma área qualquer e depois participam de vários cursos de curta duração em Tradução;
- c. em Tradução.

2.2 O Ensino da Tradução no Brasil e no mundo

Alguns pesquisadores, entre eles Gonçalves e Machado (2006), vem se preocupando com a questão do ensino da Tradução no Brasil. Schäffner e Adab¹ (2000, *apud* GONÇALVES; MACHADO, 2006), refletindo sobre o ensino de Tradução na Europa, indicam que

[...] diferentes tipos de instituições acadêmicas oferecem cursos para a formação profissional. Dependendo de restrições sócio-culturais, seus currículos e programas podem priorizar a teoria de tradução, habilidades práticas e, mais freqüentemente, uma combinação de ambas. (p.46; tradução de GONÇALVES e MACHADO)

¹ SCHÄFFNER, C.; ADAB, B. **Developing translation competence**. Amsterdam: John Benjamins Publishing, 2000.

Isso acaba polarizando os cursos, segundo Gonçalves e Machado, entre o treinamento prático (automatização/desenvolvimento de conhecimentos procedimentais – saber como) e a reflexão teórica (conscientização/desenvolvimento de conhecimentos declarativos – saber o quê), quando da discussão sobre a construção daquilo que denominam Competência do Tradutor (CT). Essa discussão começa a ser levantada no Brasil também. Os autores identificaram, entre graduação e especialização, cento e dezoito instituições que oferecem cursos de Tradução no Brasil.

Apesar do grande número de cursos, uma questão pode ser levantada: será que esses futuros profissionais estão preparados para lidar com as necessidades do mercado de Tradução? Os currículos dos cursos os preparam de modo satisfatório para exercerem sua profissão? Pym (1993, 2002, 2006), corretamente, dá uma resposta negativa para ambas as perguntas. O aprendiz, a princípio, está ali para burilar seus conhecimentos na língua de partida (o inglês, por exemplo) e melhorar sua redação na língua de chegada (português, no caso do Brasil). Os cursos, então, tentam atender a essa demanda e investem no aperfeiçoamento dos pares de línguas escolhidas. Porém, quanto é investido no aprendizado das ferramentas ligadas ao ato de traduzir? Quanto o mercado de trabalho é analisado em sala de aula? Terão os professores vivência suficiente na área de Tradução e no uso de ferramentas para repassar suas experiências aos alunos?

Como exemplo de cursos desenvolvidos no exterior que investem nessa vertente de aprendizado de ferramentas informatizadas, posso citar as experiências coletadas por Jaatinen e Jääskeläinen (2006) durante o projeto Colc (Computing for Language Careers), que pretendia inserir a tecnologia da informação no treinamento de tradutores. Entre as disciplinas do currículo básico, estavam inclusas: introdução à localização, habilidades básicas de computação para tradutores, interfaces de uso gráfico, introdução

à engenharia de software, documentação de software, ferramentas de memória de tradução e bancos de dados de termos, lingüística de *corpus* e tradução assistida por *corpus*.

Araújo (2001), por outro lado, realizou um dos melhores estudos sobre cursos de Tradução no Brasil, inclusive com a análise de grades curriculares e ementas de disciplinas. Dentre as dez instituições analisadas que ministravam cursos de graduação ou especialização em Tradução, seis não contavam com disciplinas de Terminologia, base para o aprendizado de uso e criação de novas ferramentas para tradutores. Numa segunda fase de sua pesquisa, trabalhando apenas com alunos que cursaram a disciplina de Terminologia, a autora destaca (p. 61), ao analisar a familiaridade desses alunos com recursos terminológicos, uma conclusão que remete ao pensamento de Pym sobre problemas no ensino de ferramentas informatizadas para aprendizes:

Neste grupo de respondentes, apenas 25% mencionaram conhecer algum tipo de ferramenta terminológica para a tradução. Destes, 33% apontaram conhecer bancos de dados, 27% mencionaram estar familiarizados com glossários e 13% mencionaram conhecem dicionários especializados. Apenas 27% desses alunos mencionaram conhecer um programa específico, o Autolex.

Essas e outras conclusões levantadas pela pesquisa da autora apenas corroboram a idéia de que há muito a ser discutido sobre o ensino de Tradução no Brasil, onde pouco espaço é dado ao aprendizado de ferramentas informatizadas de apoio ao profissional, enquanto os tradutores experientes, na sua maioria, trabalham com textos técnicos que exigem o manuseio dessas ferramentas e o conhecimento da Terminologia das áreas a serem traduzidas².

² Uma outra pesquisa realizada por Araújo (2001, p. 46), com firmas de tradução, revela que 82% da fatia do mercado de trabalho dos tradutores está nas áreas técnicas. Isso vai ao encontro da pesquisa realizada

2.3 O aprendizado em tradução

Diversos autores trabalham com a questão do ensino de Tradução. Começo a discussão com aqueles que tratam de questões de caráter geral sobre o ensino e específico, mais adiante, outros que destacam questões pertinentes a esta tese.

Arrojo (2003), por exemplo, propõe que aprender a traduzir é aprender a ler: “aprender a produzir significados, a partir de um determinado texto, que sejam ‘aceitáveis’ para a comunidade cultural da qual participa o leitor.” (p. 76). O bom tradutor seria, continua a autora, aquele que aprende então a ler, a escrever (com o mesmo cuidado daqueles que se preparam para serem escritores) e a se manter informado a respeito das teorias e dos estudos sobre tradução.

Já Stolze (1997), analisando indicadores de como avaliar uma tradução, repara que o ensino de tradução está frequentemente aliado ao ensino de língua. Segundo a autora, contudo, “[...] a proficiência lingüística e o conhecimento do assunto são condições para a tradução” (p. 165-166). Entre as categorias lingüísticas descritas pela autora, destaca-se a questão do léxico e da terminologia (comunicação em áreas especializadas). A autora salienta que a didática da tradução deve destacar diferentes modos de terminologização a fim de sensibilizar o tradutor para as diferenças terminológicas e de constituição textual entre a área científica e a de humanidades (p. 162).

A grande vertente atual de discussão entre os estudiosos, no entanto, caminha em direção à chamada Competência Tradutória. Gonçalves e Machado (2006, p. 53-57) levantam dezessete categorias de Competência defendidas por vários teóricos e

com tradutores profissionais para essa tese, cujo gráfico final, apresentado no apêndice B, revela fatos semelhantes.

apontam o quanto elas são adotadas, em questão de porcentagem, pelas instituições de ensino. Essas categorias de competência foram agrupadas num gráfico (figura 1) e três, particularmente, são bastante relevantes para esta tese: a questão da Terminologia (8), das Tecnologias que podem ser aplicadas à tradução (12) e os conhecimentos operativos/procedimentais sobre a tradução (13).



Figura 1. Competências Tradutórias trabalhadas nos diversos cursos de Tradução no Brasil (GONÇALVES; MACHADO, 2006).

A questão da competência, já levantada, também é trabalhada por Alves (1997), que divide a questão da aprendizagem entre competência lingüística e competência tradutória. O autor indica que a competência lingüística é algo que deveria ser inerente ao aprendiz (em consonância com a opinião de STOLZE, 1997, quando fala de proficiência lingüística); o que deve ser desenvolvido é a sua competência tradutória. Alves defende, na busca da conscientização do tradutor nos passos do ofício, a existência de duas etapas da competência tradutória: o Bloco Adhoc (composto por automatismos) e o Bloco Restante (dividido em seis passos: bloqueio processual; apoio interno; apoio externo; a combinação de apoios interno e externo; priorização e omissão de informações e burilamento final do texto de chegada). Partindo desse Bloco Restante e tendo o apoio externo como base, já que vai ao encontro do tema desta tese, examino, no próximo capítulo, como esse autor e outros trabalham o assunto.

3. Obras de consulta disponíveis

Antes de entrar na questão da Tradução propriamente dita, apresento um autor que trabalha com a questão do uso de dicionários como um fator de aprendizado para a competência lingüística (como o aprendizado de outra língua, por exemplo). Schmitz (2000), discutindo o uso de dicionários por parte de aprendizes de língua estrangeira, questiona não só a utilidade dessas obras, mas também a vontade desses alunos em decifrar códigos gramaticais impostos pelas editoras em seus dicionários.

Transitando entre o enfoque de aprendizado de língua (vernácula ou estrangeira) e o aprendizado de tradução, Darin (1997, p. 118), comenta sobre a competência lingüística que deveria ser inerente ao aluno de Tradução que já teve contato prévio com essas ferramentas:

[...] o trabalho [de tradução] pode ser bem feito se houver uma pesquisa adequada. Entretanto, surpreendeu-nos a dificuldade às vezes apresentada na utilização do dicionário, principalmente tendo em vista que disciplinas do primeiro ano estão incumbidas de preparar os alunos para esta atividade tão vital ao estudante de Letras.

Voltando à questão do apoio externo proposto por Alves (1997, p. 25-26), é importante destacar a ênfase dada às ferramentas a serem utilizadas por aqueles que estão desenvolvendo a competência tradutória:

[...] sem conseguir solucionar, através de conhecimentos prévios, o problema gerado no decorrer do processo tradutório, o tradutor passa a utilizar-se de mecanismos de apoio externo para tomar sua decisão de tradução. Ao utilizar-se de fontes de apoio externo, o tradutor depende de um tipo de conhecimento diferente dos conhecimentos declarativos utilizados como apoio interno. Ele se vale agora de um outro tipo de conhecimento: o conhecimento procedural. Em outras palavras, o tradutor deve adquirir a capacidade de operar instrumentos que lhe auxiliem na tarefa de traduzir. Tanto dicionários quanto obras de consulta e referência desempenham um papel fundamental nessa fase.

Além desses recursos, cabe ao tradutor saber buscar fontes de consulta alternativas através do contato com informantes e agências de informação. Enfim, desenvolve-se aqui, junto ao tradutor em formação, um trabalho de conscientização de como implementar as estratégias e técnicas de tradução discutidas na literatura da área.

No entanto, Pagano (2003) aponta que, “embora seja o recurso mais comumente associado à tradução e ao tradutor, o dicionário está longe de ser o recurso mais bem utilizado” (p. 41). A autora critica a crença, bastante difundida entre os aprendizes, de que basta um dicionário bilíngüe para dar conta de qualquer tradução. Para ela, um dicionário bilíngüe “é apenas um dos recursos existentes, cuja utilização requer a verificação ou checagem das informações em outros tipos de dicionários, como é o caso dos dicionários monolíngües, que oferecem uma descrição ou explicação do termo procurado” (p. 41). Segundo Pagano, os dicionários usados pelos tradutores devem ser atualizados constantemente ou deve-se usar versões mais modernas dos mesmos, construídas por meio de recursos computacionais e bancos de dados calcados em diversos tipos de texto. As definições seriam básicas para o profissional da tradução, mas, além delas, a autora destaca a necessidade de a obra dispor de sinônimos, antônimos, hipônimos e palavras afins. Defende, ainda, que o tradutor, para poder trabalhar com expressões de uma área técnica, deve: consultar os especialistas na área; pesquisar em glossários e dicionários especializados; ler textos paralelos e buscar por palavras-chave na Internet (colocados nessa ordem pela autora).

Percebe-se, com as opiniões dos vários autores apresentados, que o ensino da ferramenta “dicionário”, constantemente negligenciada, é de vital importância para aqueles que estudam uma língua estrangeira e para o aprendiz de Tradução. É importante, também, que o aluno tenha um conhecimento das estruturas que compõem essas obras para poder melhor escolher os vários tipos de obras de referência quando da necessidade de consulta.

3.1 Obras tradicionais

Além dos grandes dicionários gerais de língua, uma vasta quantidade de vocabulários de áreas de especialidade conta, cada dia mais, com publicações em papel. Há, no caso, dois tipos de público para essas obras: o público-alvo representado pelos profissionais de cada área, que necessitam dominar a terminologia para um melhor desempenho profissional, e os tradutores, que consultam esses tipos de obra como fonte terminológica em várias áreas. Já há trabalhos que analisam essas obras terminológicas. Fromm (2002), por exemplo, realizou uma análise contrastando cinco vocabulários técnicos na área de computação e apresentou as diferentes construções para a microestrutura e as incongruências internas presentes em todas as obras, especialmente aquelas referentes à inconsistência na construção da microestrutura.

Existem vários modelos possíveis de formatação da macro e microestruturas dessas obras. Duas, a título de ilustração, são apresentadas a seguir. No *Vocabulaire d'Internet* (Office de langue française – Canadá), obra impressa de 1995, há uma preocupação já no preparo da macroestrutura. Novas acepções de um mesmo termo são colocadas como novos verbetes (como o termo *hypertext*, figura 2). As entradas são impressas em itálico. Na seqüência da entrada é (são) apresentada(s) a(s) tradução(ções) para o francês (em negrito), a categoria gramatical, gênero, sinônimos, etc. (que compõe o Paradigma Informacional, discutido no capítulo 7). A definição, em francês, é apresentada logo em seguida e, por fim, são apresentadas notas de caráter enciclopédico. A obra não apresenta exemplos ou abonações para os termos propostos. A reflexão sobre todas essas questões de macro- e microestrutura são essenciais quando da construção de uma obra lexicográfica/terminográfica, tendo em vista, especialmente,

o público-alvo. Tanto essa obra quanto a série proposta pela editora SBS levaram em conta essa preocupação.

169. *hypertext*

hypertexte n. m.

Présentation de l'information qui permet une lecture non linéaire grâce à la présence de liens sémantiques activables dans les documents.

Notes. — 1. Le nom *hypertexte* peut aussi être adjectif qualificatif.

2. C'est Ted Nelson qui a inventé le terme *hypertexte* en 1965.

3. Dans W3, on parle beaucoup d'hyper-texte alors qu'il s'agit de plus en plus d'hypermédia puisque les données peuvent se présenter sous la forme de texte, d'image ou de son.

170. *hypertext document*

document hypertexte n. m.;

hyperdocument n. m.

Quasi-syn. **document W3** n. m.

Termes à éviter : document World Wide Web;
document WWW

Document contenant des liens hypertextes.

Notes. — 1. Il existe plusieurs systèmes qui utilisent des documents hypertextes. Dans Internet, c'est W3 qui est constitué de documents hypertextes (et de documents hypermédias). Dans ce cadre précis, on peut donc parler indifféremment de *document hypertexte* ou de *document W3*.

2. Le terme *World Wide Web* et son abréviation *WWW* sont des emprunts à l'anglais difficiles à intégrer au français sur le plan phonétique et donc à éviter.

171. *hypertext link*,

hyperlink;

link

lien hypertexte n. m.

Connexion activable à la demande dans W3, reliant des données ayant une relation

Figura 2. Vocabulaire d'Internet, termo *hypertext*, p. 29.

A coleção Mil e Um Termos, da SBS, segue parcialmente o formato de estrutura adotado pela obra canadense. Veja-se, por exemplo, o verbete *band* (figura 3): o termo base é seguido pela classe gramatical (quando da entrada em inglês, como no exemplo abaixo) ou classe gramatical/gênero (quando da entrada em português) e a tradução para a outra língua. Há um caráter enciclopédico presente na obra com a inclusão de

inúmeras ilustrações e tabelas. Os exemplos apresentados, identificados por uma figura com a letra “E”, não são baseados em textos reais, são apenas um indicativo de uso. A figura com o dedo apontando indica comentários ou sinônimos (como no exemplo abaixo) e uma figura com a letra “P” indica uma forma popular (banalizada) para o termo, enquanto as remissivas são indicadas por uma seta para a direita (esses dois últimos não disponíveis na figura abaixo).

balancing side (n.) lado de balanceio
 See also **working side**

Balter's activator (n.) bionator de Balter

band (v.) bandar (*fixar uma banda ortodôntica a um dente*)

band (n.) banda ortodôntica

- **[abutment] band** banda de sustentação
- **[adjustable] band** banda regulável (*banda ajustável ao colo dentário*)
- **[all-closing] band** banda completa
- **[anchor] band** banda de ancoragem
- **[anchor clamp] band** grampo de ancoragem
- **[Angle] band** banda de Angle
- **[apron] band** banda em avental (*banda ortodôntica com extensão vestibuloincisal ou gengival*)
- **[clamped] band** banda de pressão | banda travada
- **[contoured] band** banda contornada
- **[elastic] band** elásticos ortodônticos | anéis elásticos (*bandas elásticas utilizadas para estimular a movimentação dentária*)

Same as **orthodontic elastic band**

See illustration on page 79

- **[nickel silver] band** banda de níquel-prata

- **[preformed] band** banda pré-fabricada | banda pré-moldada
- **[Richmond] band** banda de Richmond
- **[seamless] band** banda ortodôntica sem emenda | banda pré-fabricada
- **[Hunter-Schreger] bands** bandas de Hunter-Schreger
- **[molar] bands** bandas para molar

band adapter (n.) assentador de bandas (*instrumento utilizado para adaptação de bandas ortodônticas aos dentes*)

band driver (n.) calcador para banda

Same as **band pusher**

band forming plier (n.) conformador de bandas

band pusher (n.) calcador para banda (*instrumento utilizado para adaptar e posicionar bandas ortodônticas aos dentes*)



band pusher = calcador para banda

Figura 3. Vocabulário para Ortodontia, termo *band*, p. 83.

Novas acepções, com palavras à direita do termo em questão, entram como novos verbetes (*band adapter, band driver*). Já as novas acepções formadas por uma palavra acrescentada à esquerda do termo são colocadas como diferentes acepções na

microestrutura do mesmo e são dispostas em diferentes linhas com indicação de marcadores ([abutment] band; [adjustable] band); o colocado do termo base é indicado por colchetes ([]).

3.2 Novas tecnologias em consulta

Gil e Pym (2006, p. 17) já pontuaram que a

Tecnologia não é uma opção no mundo profissional hoje em dia; é uma necessidade. Anos atrás falava-se de Tradução Assistida por Computador (CAT). Isso, agora, parece redundante. Virtualmente todas as traduções são assistidas por computador.
(minha tradução)³.

Analisando obras brasileiras, no entanto, pode-se dizer que apenas os dicionários gerais de língua em versão eletrônica, como o Houaiss (2002), aproveitam as vantagens do computador. Algumas funcionalidades, antes impensáveis, passam a existir: os vínculos em hipertexto, por exemplo, ligam todos os semas que compõem uma definição às respectivas entradas dos mesmos. As obras dispõem, também, de mais de um modo de exibição para a microestrutura dos verbetes (assunto melhor discutido no capítulo seis).

Cada vez mais, obras já disponíveis há bastante tempo estão sendo passadas da mídia tradicional (em papel) para uma mídia eletrônica (CDROM, Internet). Parece haver uma clara tendência, também para os vocabulários técnicos, em seguir a digitalização de suas obras. Muitos, no entanto, perdem a oportunidade de acrescentar novas funcionalidades e, com um visual bastante pobre, tentam apenas reproduzir a versão em papel. O Dicionário Jurídico Noronha representa esta tendência (figura 4).

³ Technology is not an option in today's professional world; it is a necessity. Years ago one talked about Computer-Aided Translation (CAT). That now seems a redundancy. Virtually all translating is aided by computers.

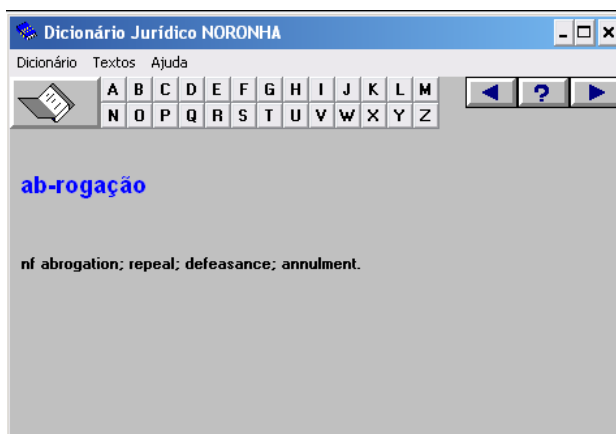


Figura 4. Dicionário Jurídico Noronha. Tela para o termo *ab-rogação* e os equivalentes em inglês.

O Dicionário Eletrônico Michaelis Técnico apresenta, em seis línguas, a equivalência para termos da área industrial (figura 5). Como no caso do Noronha, não apresenta a definição do termo e tampouco faz uso de funcionalidades disponibilizadas pelas ferramentas eletrônicas; possui, por outro lado, um esquema gráfico colorido⁴ e com vários ícones que facilitam a consulta.



Figura 5. Tela do Dicionário Eletrônico Michaelis Técnico, verbete *abelito* (em português).

⁴ Todas as telas de programas e sites aqui apresentadas são coloridas, portanto esta impressão não representa, com total fidelidade, os originais.

Os dicionários e vocabulários não só se modernizaram, mas também seguiram para a Internet e lá foram desenvolvidas novas ferramentas, como programas de tradução automática e enciclopédias *online* que, cada vez mais, auxiliam os profissionais da tradução a realizar o seu trabalho. O site Answers.com (figuras 6a e 6b), uma combinação de dicionário e enciclopédia, é um bom exemplo da convergência de obras num mesmo espaço.

The screenshot shows the Answers.com website interface. At the top, there is a navigation menu with links for Home, Business, Entertainment, Food, Health, People, Places, Reference, Shopping, Words, Careers, and More. The main header features the Answers.com logo and a search bar containing the word 'Computer'. Below the search bar, there are several sections:

- Contents:** A list of various reference topics including Dictionary, Sci-Tech Encycl., Modern Science, Computer Ency., Business, Britannica Concise, Philosophy, Columbia Ency., Abbreviations, Word Tutor, Quotes About, Wikipedia, Translations, Best of Web, and Cite this page.
- WikiAnswers:** Options to Ask Questions, Answer Questions, and Browse Questions.
- Resources:** Links for Bloggers & Webmasters, Educators, Journalists, Small Businesses, Advertisers, and Computer User Groups.
- Did you mean:** A list of related terms and pages, such as 'Computer User Group Page', 'Computer Enterprises, Inc.', 'computer', and 'Computer (magazine)'.
- computer (in computers, communication, mathematics):** The main definition section, which includes:
 - Dictionary:** A section with a 'Save Big on PCs' sponsored link and a Dell advertisement for a 'Computador Dell 530'.
 - com·put·er (kəm-pyū'tar):** The phonetic transcription and a list of two definitions:
 - A device that computes, especially a programmable electronic machine that performs high-speed mathematical or logical operations or that assembles, stores, correlates, or otherwise processes information.
 - One who computes.
 - Sci-Tech Encyclopedia:** A section with a 'Computer Parts' sponsored link from Alibaba and a 'Promoção Notebook HP' advertisement for a Core 2 Duo laptop.
 - Computer:** A detailed encyclopedic definition starting with 'A device that receives, processes, and presents information...' and discussing analog and digital computers.

Figura 6a. Answers.com. Definição do verbete *computer*. Informações lexicográficas e enciclopédicas.

Na visualização do verbete *computer* (o site é todo em inglês e só apresenta essa língua como ponto de partida), por exemplo, a página mostra, em primeiro lugar, a definição do mesmo (característica de dicionário monolíngüe; figura 6a, parte superior, *dictionary*), retirado do dicionário The American Heritage. Na sequência, várias informações, gráficos, fotos, etc. são apresentados para complementar o entendimento (característica de enciclopédia; figura 6a, parte inferior, *sci-tech encyclopedia*). Por fim, podem-se visualizar as formas equivalentes do verbete em várias línguas (característica de glossário de formas equivalentes; figura 6b).

The screenshot displays the following content for the word 'computer':

- [computer conferencing](#) (Comput) téléconférence
- [computer dating](#) rencontres sélectionnées par ordinateur
- [computer graphics](#) infographie, images de synthèse

Deutsch (German)
n. - Computer, elektronische Datenverarbeitungsanlage

idioms:

- [computer conferencing](#) an einer elektronischen Konferenz teilnehmen
- [computer dating](#) Datenbanken benutzen, um jdm. einen Partner zu finden
- [computer graphics](#) Computergrafik

Ελληνική (Greek)
n. - ηλεκτρονικός υπολογιστής

idioms:

- [computer conferencing](#) διάσκεψη μέσω υπολογιστών, τηλεδιάσκεψη
- [computer dating](#) ραντεβού μέσω υπολογιστών
- [computer graphics](#) γραφικά μέσω υπολογιστή

Italiano (Italian)
computer, calcolatore

Português (Portuguese)
n. - computador (m)

Figura 6b. Answers.com. Definição do verbete *computer*. Informações de forma equivalente.

A proposta desta tese é fornecer ao aprendiz, num único ambiente, várias das características acima apresentadas com o intuito de treiná-lo e torná-lo proficiente no uso das mesmas, a fim de otimizar seu trabalho como tradutor.

3.3 Ferramentas de Tradução

Muitos dos tradutores profissionais, conforme menciona Araújo (2001, p. 49), já desenvolvem, por iniciativa própria e para consumo individual, glossários e vocabulários para suas traduções, porém trabalham de forma precária com programas de processamento de texto ou planilhas para organizá-los. Tendo em vista essa problemática, além de dicionários e vocabulários, de uso do público em geral ou de especialistas, novas ferramentas foram e continuam sendo desenvolvidas especialmente para os tradutores. Zerfass (2002) destaca, entre várias ferramentas, os utilitários de conversão para gerenciamento de terminologia, memórias de tradução, tradução por máquina e sistemas de gerenciamento de projeto e fluxo de trabalho. Bowker e Barlow (2000) fazem uma interessante análise comparativa entre a memória de tradução os concordanceadores bilíngües (os autores indicam que esses são muito mais conhecidos entre acadêmicos do que entre tradutores profissionais), também indicados para o ensino de tradução. Mas a memória de tradução é, com larga vantagem, a ferramenta que mais se destaca (entre as acima apresentadas). Ela armazena trechos inteiros de traduções já realizadas (por meio de um processo de segmentação do texto), pelo tradutor ou por outro tradutor/firma de tradução, e os disponibiliza para comparação e/ou simples cópia do trecho. Um exemplo é o programa WordFast, que funciona como uma macro para o Microsoft Office (figura 7). Outros programas, além de conterem um módulo de memória de tradução, disponibilizam para o tradutor, também, uma série de ferramentas para o gerenciamento de terminologia; o Trados é o melhor exemplo entre esses programas (a ele são acopladas diversas ferramentas, entre elas, o Translator's Workbench, o Multiterm e o SDLX).

Esses programas estão na vanguarda das tecnologias que auxiliam o tradutor, porém pressupõem que o usuário seja um tradutor com alguma ou muita experiência na área. Programas mais avançados estariam ligados, portanto, a uma segunda fase de

aprendizado de ferramentas eletrônicas de ajuda ao profissional. Dicionários, em qualquer formato, ainda fazem parte do aprendizado básico para quem trabalha com tradução.

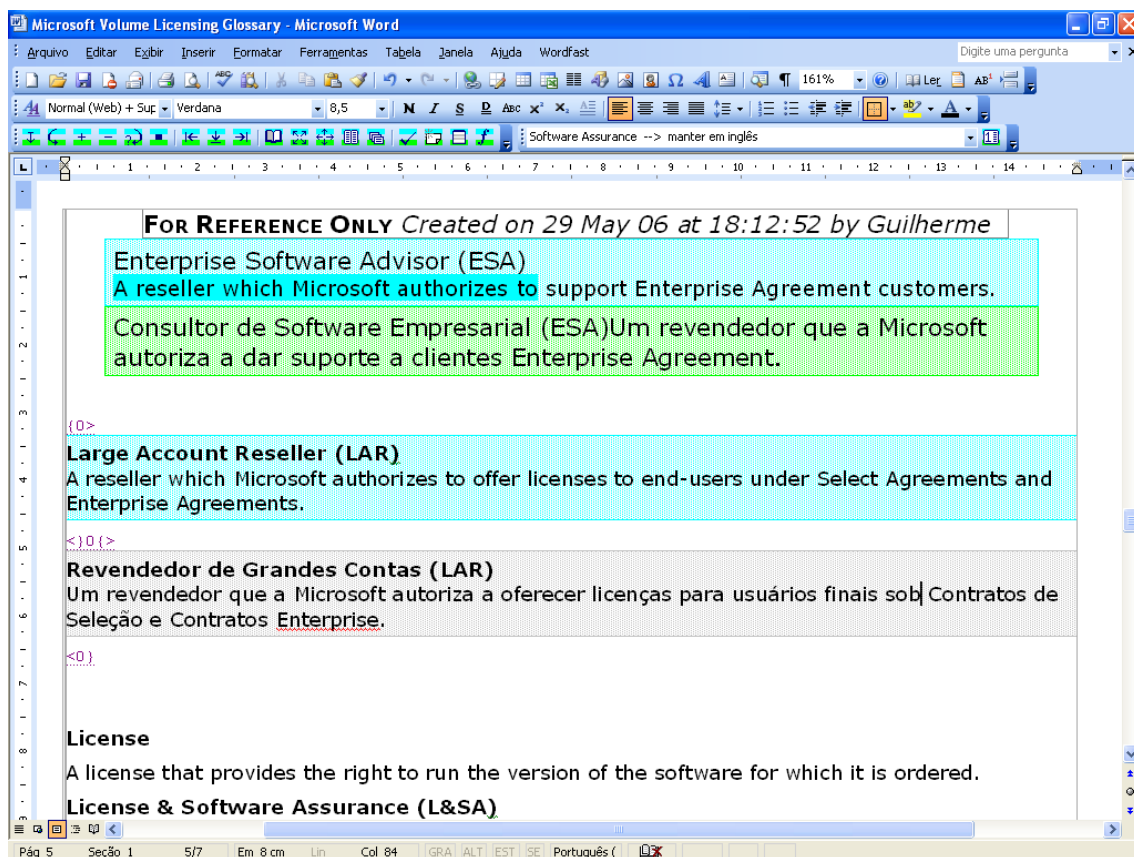


Figura 7. Exemplo de uma tradução na área de computação com recursos do WordFast. Na primeira parte (primeiro quadro) são apresentados o original e a tradução já realizada e armazenadas pelo programa. Na segunda parte o tradutor pode copiar pedaços já disponibilizados ou usá-los como parâmetro para o segmento a ser traduzido.

3.4 Tradução e Terminologia: a necessidade da interface

Além das obras de consulta externa, alguns autores sustentam o uso de outras fontes de pesquisa. Aubert (2001) advoga que a terminologia e a terminografia são ferramentas essenciais aos tradutores (p. 42). Uma situação ideal de busca por um termo, por parte dos tradutores, consistiria na consulta a:

- a. uma fonte monolíngüe na língua-fonte;
- b. vários materiais terminográficos bilíngües (bidirecionais);
- c. materiais monolíngües (dicionário de especialidade) na língua alvo, para efeito de verificação.

O ensino da Terminologia, uma das competências levantadas por Gonçalves e Machado (2006), está, portanto, intimamente ligado à Tradução. Araújo (2002), partindo dos problemas relacionados à tradução do livro *A brief history of time* (Uma breve história do tempo), do físico inglês Stephen Hawking, questiona a ligação da Tradução com a Terminologia. No caso do livro de Hawking, os físicos, um dos públicos-alvo, mencionaram vários erros de tradução de acordo com a terminologia vigente para as áreas de física e astronomia; não houve, pelo lado da tradutora, uma preocupação em usar a terminologia adequada (em parte porque o livro também era destinado ao público leigo) e isso resultou em lacunas de entendimento por parte dos profissionais que estão acostumados com essa terminologia.

A necessidade de tradução adequada de termos técnico-científicos é apontada por Krieger (2006, p. 190), ao defender o aprendizado de Terminologia para tradutores:

Os tradutores compreendem que os termos técnico-científicos, objetos centrais da disciplina terminológica, são componentes lingüísticos e cognitivos nucleares dos textos especializados; constituindo-se, conseqüentemente, em peças-chave de representação e de divulgação do saber científico e tecnológico. Daí a importância de identificá-los e traduzi-los adequadamente, embora os termos não sejam os únicos elementos que permitem que a comunicação profissional cumpra suas finalidades.

A necessidade da associação entre as duas áreas – Tradução e Terminologia - é bastante clara e já foi levantada por Araújo (2002). Embora as estatísticas de Gonçalves e Machado (2006, p. 57) indiquem que 41,3% dos estabelecimentos de ensino que

ministram o curso de Tradução já trabalhem com Terminologia, seja através de disciplinas, laboratórios e centros de pesquisa, muito, ainda, pode ser desenvolvido para que essa união seja mais profícua.

Essa união já foi levantada por Araújo (2002, p. 177) com uma certa preocupação, pois, para a autora,

[...] o que podemos constatar na prática é uma falta generalizada de produtos terminológicos no mercado nacional voltados para o auxílio à tradução, quer na forma de dicionários especializados ou, numa abordagem informatizada, bancos de dados terminológicos (BDTs), que contemplem o português brasileiro e possam ser utilizados como fontes de consulta confiáveis [...]

O objetivo primordial desta tese, portanto, é facilitar o contato dos aprendizes de tradução com a Terminologia por intermédio do uso do VoTec, uma ferramenta que disponibiliza, através de um vocabulário online de áreas técnicas, meios para que os alunos entrem em contato com obras terminográficas já finalizadas ou, no caso daqueles mais avançados, “ponham a mão na massa” e produzam uma obra terminográfica.

4. A construção dos *corpora* de estudo

Um dos pilares do VoTec é que a obra de consulta terminológica a ser desenvolvida obedecerá um princípio básico: todo o conteúdo terminográfico será totalmente construído a partir de *corpora* de textos reais, autênticos, nas duas línguas abordadas (inglês/português). Convém, antes de mais nada, explicar o que é um *corpus*.

4.1 O que é um *corpus*?

Um *corpus*, segundo Tagnin (2004), é “[...] uma coletânea de textos em formato eletrônico, compilada segundo critérios específicos, considerada representativa de uma língua (ou da parte que se pretende estudar), destinada à pesquisa”. Bidermann (2001, p. 79) coloca como *corpus* um conjunto homogêneo de amostras de língua de qualquer tipo que deve possibilitar, mediante análise lingüística, a ampliação do conhecimento das estruturas lingüísticas da língua que ele representa. A área da Lingüística que trata dos estudos sobre *corpora* (assim como de suas compilações), é a Lingüística de *Corpus*. Para Berber Sardinha,

A Lingüística de *Corpus* ocupa-se da coleta e da exploração de *corpora*, ou conjuntos de dados lingüísticos textuais coletados criteriosamente, com o propósito de servirem para a pesquisa de uma língua ou variedade lingüística. Como tal, dedica-se à exploração da linguagem por meio de evidências empíricas, extraídas por computador. (2004, p. 3).

Ainda segundo Berber Sardinha (p. 20/21), quanto à tipologia, os *corpora* podem ser de diferentes:

a. modos: falados (transcrições) ou escritos;

- b. tempos: sincrônicos ou diacrônicos; contemporâneos ou históricos;
- c. seleções: por amostragem (estático, amostra finita da linguagem como um todo), monitor (dinâmico, reciclável), balanceado (textos distribuídos em quantidades semelhantes);
- d. conteúdos: especializados (gêneros ou registros definidos), regionais ou dialetais, multilíngües;
- e. autorias: de aprendiz (falantes não-nativos) ou de língua nativa (falantes nativos);
- f. disposições internas: paralelos (original e tradução) e alinhados⁵;
- g. finalidades: de estudo (*corpus* a ser descrito), de referência (para contrastar com o *corpus* de estudo) e de treinamento (para desenvolvimento de aplicações e ferramentas de análise).

4.2 Os corpora desta tese

A construção do *corpus* para a alimentação do banco de dados passou por várias fases. A idéia original era o uso de *corpora* bilíngües comparáveis já prontos, adotando o princípio da reusabilidade. Esses *corpora* seriam buscados num dos vários projetos do COMET: o CORTEC. Segundo o site do COMET (www.fflch.usp.br/dlm/comet), o CORTEC “[...]é um *corpus* comparável de textos técnicos e/ou científicos originalmente escritos em português brasileiro e em inglês.” As áreas iniciais abrangidas pelo projeto são: Direito Contratual, Informática, Hipertensão Arterial, Culinária e Ecoturismo. O projeto prevê a inserção contínua de *corpora* em novas áreas e a complementação, também contínua, dos *corpora* já existentes.

⁵ Utiliza-se, nesta tese e nos trabalhos propostos pelo projeto COMET, a oposição entre corpora paralelos (textos originais e suas traduções) e comparáveis (textos equivalentes em línguas diferentes), diferente, portanto, dessa apresentada por Berber Sardinha.

Após conseguir os *corpora* completos do CORTEC, verificou-se que, para o presente projeto, os mesmos não apresentavam alguns aspectos essenciais:

1. não havia árvores ou mapas conceituais para todos os campos envolvidos, o que é essencial para verificar se todas as áreas foram contempladas na construção;
2. o balanceamento desses *corpora* estava bastante irregular;
3. o planejamento original dos mesmos não previa um fim lexicográfico/terminográfico, de modo que nem sempre incluía textos que permitissem a construção de definições para o banco de dados;
4. o tamanho de cada *corpus*, de aproximadamente duzentas mil palavras, também não se mostrou suficiente para selecionar uma quantidade de termos em todas as áreas e/ou a possibilidade de criar suas respectivas definições.

Verificada a necessidade de novos *corpora*, partiu-se, em primeiro lugar, para a reconstrução do *corpus* de Informática (ou Computação). Embora já houvesse um *corpus* semelhante, organizado durante o mestrado de Fromm (2002), o mesmo era monolíngüe (português). Decidiu-se, então, pela ampliação dos *corpora* desenvolvidos para o CORTEC, aproveitando o que já havia sido levantado. A estruturação final desses *corpora* ficou assim delineada: escritos, sincrônicos, de amostragem (embora exista a possibilidade de se transformarem em monitor), especializados, bilíngües, de língua nativa, comparáveis (segundo os critérios do COMET) e de estudo.

4.3 Ontologia/Taxonomia: a árvore de campo

Um dos pontos básicos para a elaboração de um banco de dados é a criação de uma estrutura para organizar a informação a ser coletada. Vários tipos de estruturas

podem ser elaboradas de acordo com o objetivo final. Segundo Almeida e Bax (2003, p. 7):

[e]struturas que se organizam a partir da utilização de termos são os *arquivos de autoridade, glossários e dicionários*. Estruturas que se organizam com a classificação e a criação de categorias são os *cabeçalhos de assunto* e os *esquemas de classificação* (ou *taxonomias*). As estruturas que se organizam a partir de conceitos e de seus relacionamentos são as *ontologias*, os *thesaurus* e as *redes semânticas*. (grifos dos autores)

Embora essa classificação dos autores pareça bastante clara, há diversos problemas em torná-la universal. Sowa (1999), por exemplo, defende uma idéia de categorização para ontologias ao colocar que

O assunto da *ontologia* é o estudo das *categorias* de coisas que existem ou podem vir a existir em algum domínio. O produto de tal estudo, chamado *ontologia*, é um catálogo de tipos de coisas que se pressupõe existirem em um domínio de interesse *D* da perspectiva de uma pessoa que usa uma língua *L* para o propósito de falar sobre *D*.⁶
(grifos do autor; minha tradução)

Tendo em vista esses diferentes conceitos para denominar o que é uma ontologia e uma taxonomia, para o presente trabalho foi escolhido o termo *taxonomia* para indicar a construção da árvore do campo pesquisada para a construção do *corpus*. No site, no entanto, optou-se pelo uso do termo *ontologia* para designar essa mesma árvore. O termo *ontologia*, cada vez mais, está associado ao uso de ferramentas computacionais para diversos tipos de análise, o que se prova pertinente para o presente caso.

O modelo tomado como base para a construção de uma árvore de campo foi aquele apresentado por Marinotto (1995) para a área de Aeronáutica e a divisão hierárquica proposta para o saber humano: campo, área, domínio, subdomínio e outros.

⁶ The subject of *ontology* is the study of the *categories* of things that exist or may exist in some domain. The product of such a study, called *an ontology*, is a catalog of the types of things that are assumed to exist in a domain of interest *D* from the perspective of a person who uses a language *L* for the purpose of talking about *D*.

A árvore do campo da computação já havia sido previamente desenvolvida por Fromm (2002) para a informática⁷ em geral; aquela, no entanto, não mais representa um estado da arte do campo em questão: o extremo dinamismo desse campo na criação de novas tecnologias e produtos requer uma atualização constante da mesma. A construção de uma taxonomia, no entanto, não é infalível: há sempre controvérsias por parte dos especialistas quanto à sua montagem.

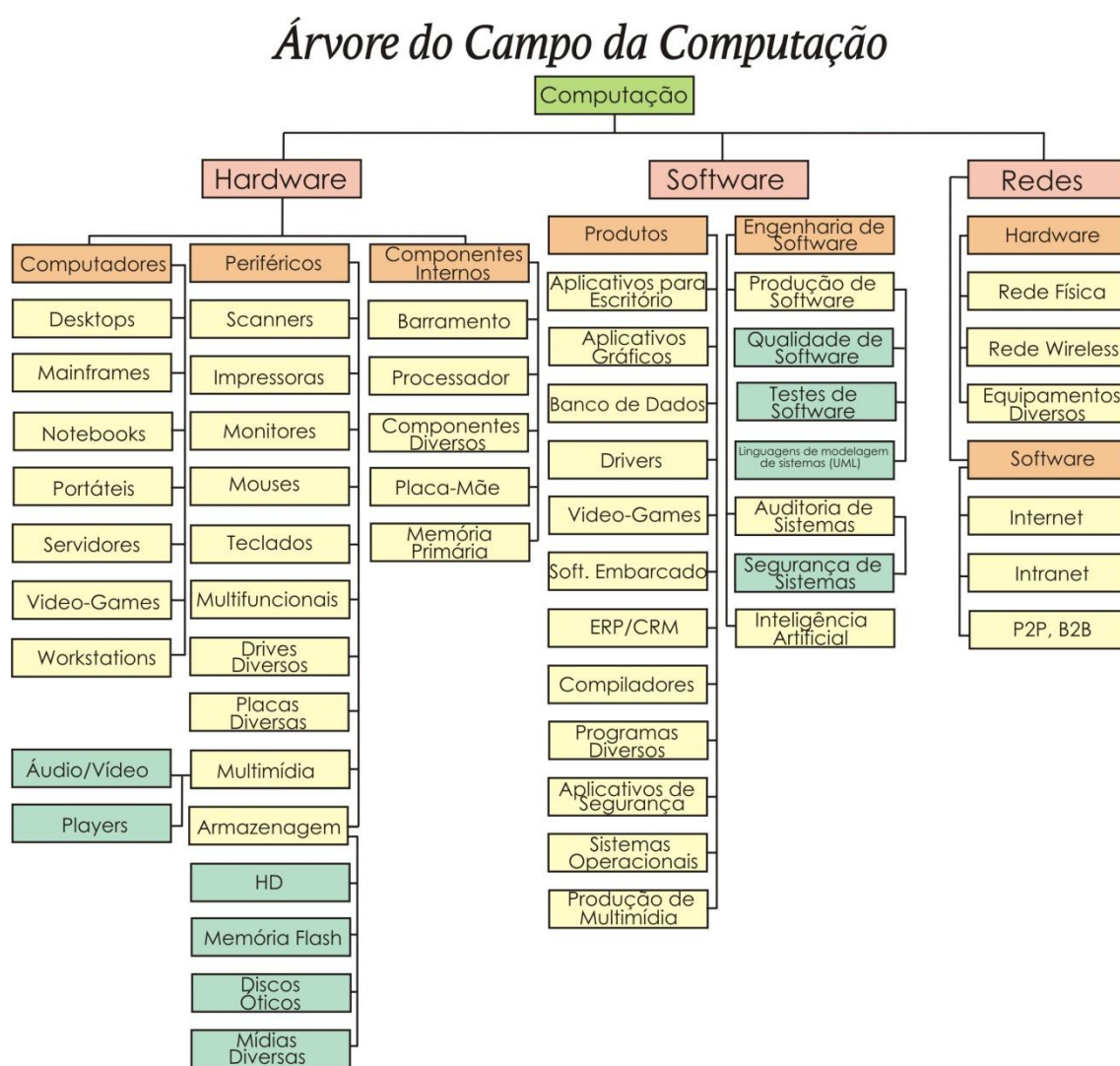


Figura 8. Árvore do Campo da Computação. Cada cor representa um novo nível.

⁷ Embora os termos *informática* e *computação* não se apresentem como sinônimos para Houaiss, eles pertencem ao mesmo campo. Tomo, aqui, esses termos como sinônimos.

A árvore acima (figura 8) representa o consenso entre a opinião de alguns professores especialistas na área, o que não quer dizer que seja unanimidade. Além disso, ainda que a Árvore de Campo (também designada Árvore de Domínio) seja uma das possíveis formas de representar uma taxonomia, uma outra forma bastante comum é aquela apresentada na figura 24, no capítulo 7, quando da inserção das áreas feita pelo administrador no banco de dados.

4.4 A coleta dos textos

Estabelecida a árvore, o passo seguinte foi a captura de uma quantidade de textos, em todas as áreas, suficiente para exibir contextos que pudessem criar definições para os termos. Já existem programas que fazem a coleta e extração de termos automaticamente, como o BootCaT, e ambientes de criação, armazenamento e análise de *corpora*, como o Corpógrafo (bem detalhados por ALMEIDA; OLIVEIRA; ALUÍSIO, 2006). Preferiu-se aqui, no entanto, não utilizar essas e outras ferramentas disponíveis, pois muitas ainda estão em fase de testes e não garantem o balanceamento de *corpus* exigido pelo trabalho; a coleta foi feita manualmente e depois os textos foram processados de acordo com as necessidades aqui propostas.

Segundo Aubert (1996), as fontes de busca para a definição de um termo podem apresentar três tipos de contextos possíveis:

O *contexto associativo* apresenta o termo como pertinente ao tema objeto da pesquisa, mas não indica os traços conceptuais específicos destes termos, [...] Já os *contextos explicativos* apresentam alguns traços conceptuais pertinentes específicos do termo sob observação, freqüentemente relativos à materialidade, finalidade, funcionamento e similares. [...] Talvez mais desejáveis, mas certamente menos contraditórios, os *contextos definitórios* proporcionam um conjunto completo dos traços conceptuais distintivos do termo. Tal distintividade, no entanto, representa freqüentemente um certo nível de abstração, sem indícios claros da gama efetiva de usos em situação do termo. (p. 66-67)

A busca por contextos associativos, no caso da presente pesquisa, pode ser automatizada através dos programas de análise lexical (como o WordSmith Tools, que será explicado adiante no item 4.6). Os contextos explicativos e definitórios, por outro lado, exigem certo conhecimento do terminógrafo sobre como localizá-los.

A necessidade de refazer os *corpora* e não apenas reutilizar os que já existiam deveu-se justamente à falta de contextos explicativos e definitórios. Verificou-se que, ao proceder à análise computadorizada dos textos previamente selecionados, havia lacunas em alguns subdomínios. Mesmo nos subdomínios com vários textos já coletados, o levantamento dos contextos foi insatisfatório.

Levando tudo isso em conta, ao começar uma nova coleta de textos, foi estabelecido um número mínimo de vinte mil palavras para cada subdomínio da árvore, quantidade que se acreditou razoável (e que se mostrou acertada, após alguns testes iniciais com um dos subdomínios disponíveis e a construção de alguns termos como teste) para o levantamento dos termos e um bom balanceamento entre esses subdomínios. Notou-se, porém, que já havia mais de um milhão de palavras, em cada língua, quando do término da primeira área da árvore (*hardware*). Decidiu-se, então, limitar os *corpora* a esse tamanho para o desenvolvimento da pesquisa. Como o objetivo da tese não é fazer um levantamento completo de um campo técnico e sim coletar alguns exemplos de termos e seus contextos para posterior análise com os aprendizes de tradução, o número obtido foi julgado suficiente, inclusive por abranger uma área completa.

4.5 A coleta dos *corpora*

Os textos coletados para os *corpora* de análise no campo da computação foram totalmente levantados pela Internet em sites especializados, muitos de caráter enciclopédico. A escolha se deveu à facilidade de encontrar textos do campo na rede (isso é uma característica marcante, já que nem todos os campos do saber estão bem representadas em termo de quantidade e qualidade na Internet) e a velocidade com que os mesmos podem ser resgatados. Embora existam muito mais sites em inglês sobre o campo da computação, não houve dificuldade para achar sites semelhantes (ou até mesmo traduzidos, como o *How Stuff Works*⁸) em português. Foi dada preferência aos sites de revistas especializadas, acadêmicos ou aqueles especializados em determinado assunto para o levantamento da pesquisa. Um site enciclopédico, no entanto, foi deixado de lado: a Wikipedia. A razão é que os termos apresentados pela mesma serão usados, na íntegra, para a elaboração do Paradigma Enciclopédico (item 7.8).

Para coletar os *corpora*, criou-se um diretório no computador que exibia pastas na mesma estrutura da árvore de campo (figuras 9 e 10). O mesmo foi subdividido entre as línguas (inglês e português) e todas as áreas, domínios e subdomínios. Novos textos coletados e aqueles remanescentes do projeto original de Informática do CORTEC, na área de *hardware*, já foram distribuídos dentro de suas respectivas pastas. Os textos remanescentes das áreas *software* e rede foram alocados, também, nas respectivas pastas. Embora somente na área de *hardware* novos textos tenham sido coletados, todos aqueles já coletados para o CORTEC foram aproveitados para o estudo.

⁸ < <http://hsw.uol.com.br>>. O endereço desta e de todas as outras páginas consultadas está disponível no banco de dados.

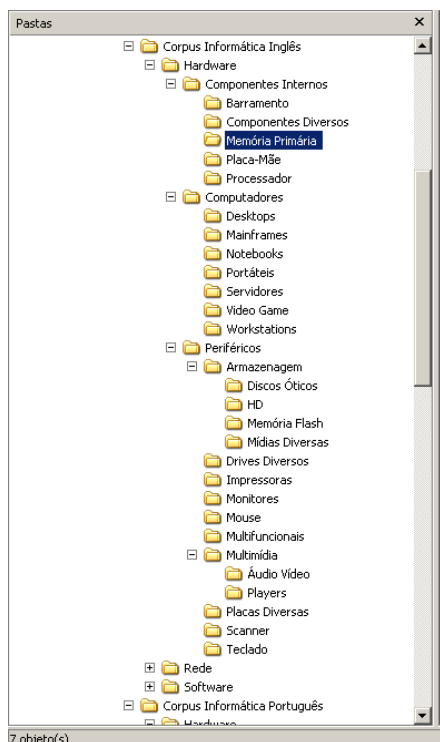


Figura 9. Diretório com pastas na forma da árvore de campo; área: *hardware*.

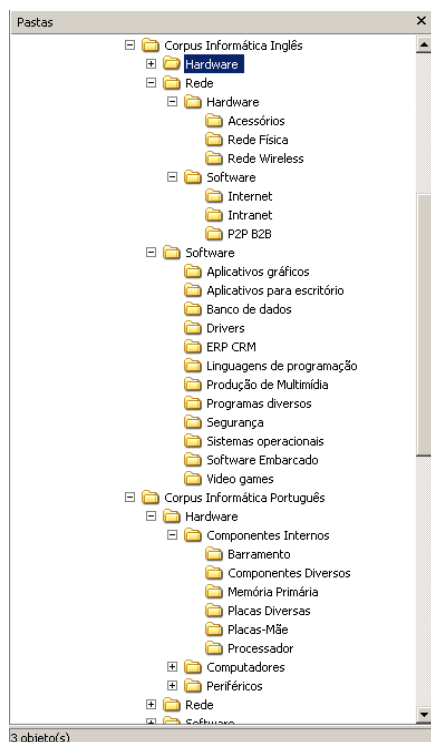


Figura 10. Idem; áreas: *software* e *rede*.

4.6 A análise dos *corpora*

Para verificar se a quantidade de textos alocados a cada pasta obedecia ao critério de vinte mil palavras por subdomínio⁹, foi usada a ferramenta Wordlist (listagem de palavras) do programa de Análise Lexical WordSmith Tools, versão 4, de Scott (2007), para fazer a contagem (veja figura 11, no destaque). Embora haja vários programas de análise computadorizada, conforme estudos anteriores (FROMM, 2004), o WordSmith Tools é o mais indicado para grande quantidade de dados e para os tipos de análise que serão demonstrados a seguir.

O volume total de palavras para o *corpus* de computação foi de 1.029.187 palavras em inglês e 1.055.375 palavras em português. Segundo Berber Sardinha (2007,

⁹ Verificada através da quantidade de *tokens* que a listagem apresenta. Os *tokens* representam a quantidade total de palavras nos textos, os *types* representam a quantidade de palavras não repetidas (distintas) nos textos.

p.26), esses *corpora* seriam classificados, de acordo com a quantidade de palavras, como médios (de 250 mil a um milhão de palavras).

N	Overall	1	2	3	4	5	6	7	8	9	10
text file	Overall	ção.txt	usb.txt	ento.txt	cpu.txt	rios.txt	ntos.txt	3gio.txt	ress.txt	ress.txt	le dvi.txt
file size	209,783	35.981	7.811	48.776	2.500	5.400	42.025	14.713	8.016	21.426	22.135
tokens (running words) in text	36,324	6.237	1.329	8.273	419	951	7.354	2.673	1.392	3.855	3.841
tokens used for word list	33.956	5.878	1.188	7.359	380	905	7.067	2.535	1.319	3.648	3.677
types (distinct words)	4.101	1.342	406	1.377	189	339	1.441	697	451	858	995
type/token ratio (TTR)	12	23	34	19	50	37	20	27	34	24	27
standardised TTR	36,55	35,30	33,60	34,97			38,66	36,65	37,10	35,00	39,70
standardised TTR std dev.	60,39	51,07		55,09			51,67	44,80		49,89	46,28
standardised TTR basis	1.000,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
mean word length (in characters)	5	5	5	5	5	5	5	5	5	5	5
word length std dev.	3,19	3,27	3,89	3,29	3,43	3,21	3,07	3,06	3,09	3,01	3,07
sentences	1.432,00	328,00	72,00	324,00	9,00	38,00	258,00	89,00	53,00	132,00	129,00
mean (in words)	24	18	17	23	42	24	27	28	25	28	29
std dev.	17,60	12,66	23,95	22,40	38,75	9,26	15,20	15,36	15,12	13,86	14,50
paragraphs	11,00	1,00	1,00	1,00	1,00	1,00	2,00	1,00	1,00	1,00	1,00
mean (in words)	3,087	5,878	1,188	7,359	380	905	3,534	2,535	1,319	3,648	3,677
std dev.	2,218,25						726,05				
headings											
mean (in words)	0	0	0	0	0	0	0	0	0	0	0
std dev.											
sections	10,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
mean (in words)	3,396	5,878	1,188	7,359	380	905	7,067	2,535	1,319	3,648	3,677
std dev.	2,597,35										
numbers removed	2.368,00	359,00	141,00	914,00	39,00	46,00	287,00	138,00	73,00	207,00	164,00
stoplist tokens removed	0	0	0	0	0	0	0	0	0	0	0
stoplist types removed	0	0	0	0	0	0	0	0	0	0	0
1-letter words	4.307	784	151	1.016	48	94	858	310	137	447	462
2-letter words	5.981	1.150	217	1.418	68	173	1.125	468	238	584	540
3-letter words	5.015	621	131	1.063	63	163	1.178	408	219	613	556
4-letter words	3.413	626	138	708	20	99	680	252	120	371	399
5-letter words	3.392	579	100	644	31	95	748	271	129	404	391
6-letter words	2.802	500	117	486	28	60	654	234	100	329	294
7-letter words	2.388	393	71	519	31	42	488	164	125	271	284

Figura 11. A subárea *componentes diversos* (em português) apresenta uma quantidade de 36.324 palavras no total (em destaque).

A cada vinte mil palavras levantadas, partia-se para uma nova subárea. Algumas subáreas, no entanto, têm um valor bastante superior a esse. O limite de vinte mil palavras, portanto, foi o mínimo a ser levantado; não houve preocupação com o volume máximo. Berber Sardinha, ao citar Sinclair¹⁰ (1997, p.27-39 *apud* BERBER SARDINHA, 2004, p.26), comenta uma entre as possíveis abordagens a respeito da extensão do *corpus* (no caso, a Impressionística):

Sinclair [...] postula que o *corpus* deva ser tão grande quanto a tecnologia permitir para a época, deixando subentender que a extensão de um *corpus* deva variar de acordo com o padrão corrente nos

¹⁰ SINCLAIR, J. Corpus evidence in language description. In: WICHMANN, A. S. et al. **Teaching and language corpora**. Londres/Nova Iorque: Longman, 1997.

grandes centros de pesquisa, que possuem equipamentos de última geração” (p. 26).

As variações de tamanho deram-se em virtude dos tipos de arquivos baixados: de algumas páginas o texto foi retirado no formato .html, copiado e colado para um arquivo formato .txt; outras páginas forneceram arquivos no formato .pdf que, sempre que possível, foram copiados para .txt também (alguns não puderam ser copiados e foram, portanto, descartados). Os arquivos em formato .pdf, normalmente estudos acadêmicos sobre a área, manuais de instrução ou propaganda dos fabricantes, têm uma quantidade maior de palavras. Essa preocupação em transformar todos os arquivos para o formato .txt dá-se por causa da velocidade de análise do programa WordSmith Tools 4. Embora ele também leia arquivos salvos em outros formatos, é no .txt que ele tem o máximo de desempenho. Os arquivos foram salvos com o título do texto (quando havia repetição dos títulos, foram acrescentados números sequenciais) e, para fins de posterior análise, depois de copiados os textos, foram incluídos o endereço do site e a data de coleta (figura 12).

Ao término da coleta e primeira análise dos *corpora* em forma de Wordlist, partiu-se para o segundo passo, que é a criação das palavras-chave (*Keywords*). Antes de iniciar a ferramenta *Keywords* do WordSmith Tools, é necessária a criação dos chamados *corpora* de referência, que são grandes *corpora* de textos gerais da língua em análise e que servem como parâmetro de comparação para a ferramenta. Em português, foi usada a versão beta do Banco de Português (BERBER SARDINHA, 2007), totalizando 689.294.592 palavras; em inglês usou-se uma combinação das listas de palavras do BNC (British National *Corpus*) e de uma versão beta do ANC (American

National *Corpus*)¹¹, totalizando 122.224.832 palavras. Em ambos os casos, os *corpora* de referência são bem maiores do que a proporção de cinco para um (o *corpus* de referência é cinco vezes maior que o *corpus* de análise) proposta por Berber-Sardinha (2004, p.102) como o tamanho recomendado.

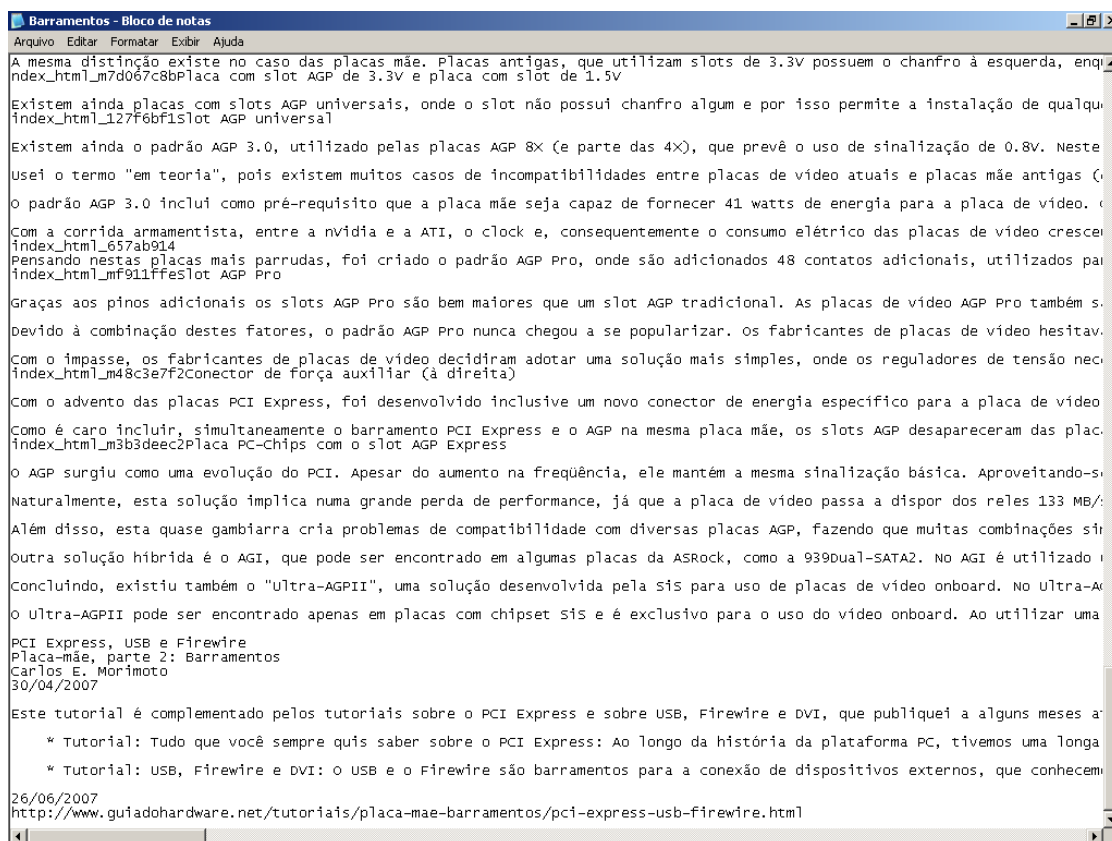


Figura 12. Arquivo .txt do *corpus*, com data de coleta e endereço na Internet (final da página).

Com a ferramenta Keywords do WordSmith Tools foram criadas, então, as listagens de palavras-chave em cada língua. Essas palavras, escolhidas por meio de análises estatísticas (*log likelihood*) entre o *corpus* de estudo e o *corpus* de referência, correspondem aos contextos associativos já citados. Na figura 13 temos a tela do programa com as palavras-chave em inglês; na planilha 1, a tela com as palavras-chave em português, agora numa listagem em Excel.

¹¹ A listagem do BNC foi obtida no site do programa Wordsmith Tools. A listagem do ANC foi elaborada tendo a segunda versão do CD como *corpus* e o programa Wordsmith Tools como ferramenta de análise.

N	Key word	Freq.	%	RC. Freq.	RC. %	Keyness	P	Lemmas	Set
1	DISK	2.322	0,23	2.887		15.117,46	0,0000000000		
2	MEMORY	2.709	0,26	8.744		13.550,35	0,0000000000		
3	DATA	3.246	0,32	31.771	0,03	9.986,71	0,0000000000		
4	CPU	1.092	0,11	316		8.958,52	0,0000000000		
5	DRIVE	2.154	0,21	11.440		8.928,05	0,0000000000		
6	ARN	882	0,09	12		8.315,24	0,0000000000		
7	DRIVES	1.233	0,12	1.735		7.802,31	0,0000000000		
8	PCI	833	0,08	84		7.413,06	0,0000000000		
9	PC	1.267	0,12	2.755		7.162,23	0,0000000000		
10	VIRTUAL	1.053	0,10	1.225		6.954,83	0,0000000000		
11	KEYBOARD	983	0,10	1.098		6.549,06	0,0000000000		
12	BIT	2.468	0,24	31.963	0,03	6.398,03	0,0000000000		
13	USB	639	0,06	37		5.829,93	0,0000000000		
14	ADDRESS	1.491	0,14	9.535		5.695,09	0,0000000000		
15	DEVICES	1.016	0,10	2.611		5.466,72	0,0000000000		
16	CONNECTOR	664	0,06	200		5.423,94	0,0000000000		
17	PARALLELS	744	0,07	559		5.350,71	0,0000000000		
18	INTERFACE	952	0,09	2.188		5.296,06	0,0000000000		
19	CD	910	0,09	1.834		5.254,41	0,0000000000		
20	DVD	707	0,07	555		5.045,30	0,0000000000		
21	MOV	527	0,05	6		4.978,49	0,0000000000		
22	FLOPPY	698	0,07	559		4.963,20	0,0000000000		
23	COMPUTER	1.644	0,16	17.184	0,01	4.868,08	0,0000000000		
24	SOFTWARE	1.425	0,14	11.505		4.861,09	0,0000000000		
25	DEVICE	987	0,10	3.395		4.829,11	0,0000000000		
26	SERVER	844	0,08	1.770		4.819,72	0,0000000000		
27	DISKS	715	0,07	829		4.725,61	0,0000000000		
28	WORKSTATION	697	0,07	759		4.668,34	0,0000000000		
29	HARDWARE	866	0,08	2.321		4.599,46	0,0000000000		
30	HARD	1.923	0,19	28.720	0,02	4.518,88	0,0000000000		
31	PROCESSOR	772	0,08	1.578		4.440,14	0,0000000000		

Figura 13. Palavras-chave na área de computação, em inglês.

As palavras na primeira coluna indicam as palavras mais freqüentes que, assim indicam os candidatos prováveis a termos naquela área; a ordem de palavras na primeira/segunda colunas leva em conta a sétima coluna, ou seja, sua chavicidade (*keyness*), que significa o quanto a palavra em destaque, na relação entre o *corpus* de análise e o *corpus* de referência, é representativa na freqüência relativa (o programa compara, estatisticamente, a freqüência desta palavra em ambos os *corpora*; se ela apresenta um uso mais [ou menos] destacado no *corpus* de análise do que no de referência, ela é incluída na lista).

WordSmith Tools 4.0 -- 27/6/2007							
N	Key word	Freq.	%	RC. Freq.	RC. %	Keyness	P
1	COMPUTADOR	3380	0,3203	28792		22302,23	7E-23
2	IMPRESSORA	1812	0,1717	2526		17610,62	1E-22
3	CLIQUE	1553	0,1472	2491		14760,42	3E-22
4	PLACA	2104	0,1994	19180		13613,9	3E-22
5	WINDOWS	1603	0,1519	11497		11086,75	6E-22
6	PROCESSADOR	1212	0,1148	2901		10738,48	7E-22
7	BITS	1183	0,1121	2547		10688,28	7E-22
8	USB	913	0,0865	326		10412,25	7E-22
9	IMPRESSÃO	1617	0,1532	19473		9616,599	9E-22
10	BARRAMENTO	815	0,0772	344		9159,807	1E-21
11	PLACAS	1514	0,1435	17867		9061,677	1E-21
12	MEMÓRIA	1964	0,1861	49114		8969,695	1E-21
13	GEFORCE	672	0,0637	2		8686,718	1E-21
14	SELECIONE	755	0,0715	496		8111,6	2E-21
15	PCI	782	0,0741	1351		7341,269	2E-21
16	VÍDEO	1379	0,1307	25009		7134,15	2E-21
17	TELA	1241	0,1176	20775		6607,95	3E-21
18	CONTROLADOR	874	0,0828	4499		6575,389	3E-21
19	DVD	755	0,0715	2026		6544,251	3E-21
20	VOCÊ	2539	0,2406	193135	0,028	6410,745	3E-21
21	RADEON	491	0,0465	0		6366,831	3E-21
22	CONSULTE	630	0,0597	701		6330,112	3E-21
23	HARDWARE	684	0,0648	1707		6012,326	4E-21
24	PALM	700	0,0663	2140		5911,728	4E-21
25	TECLADO	692	0,0656	2275		5757,233	4E-21
26	DISCO	1345	0,1274	39388		5742,442	4E-21
27	MOUSE	750	0,0711	3769		5675,047	4E-21
28	MONITOR	796	0,0754	6107		5405,326	5E-21
29	XP	558	0,0529	862		5335,347	5E-21
30	MB	781	0,074	6179		5258,846	6E-21
31	BOTÃO	693	0,0657	3595		5203,829	6E-21
32	EAX	388	0,0368	9		4945,257	7E-21
33	MEMORIA	480	0,0455	466		4914,386	7E-21

Planilha 1. Palavras-chave em português.

Identificados os candidatos a termos nas duas línguas, é preciso verificar quais deles estão presentes em ambas as listas. Nesse momento é necessário um pouco da *expertise* (conhecimento sobre a área) do pesquisador para delimitar quais termos são equivalentes nas duas línguas. Alguns são empréstimos (nessa área, em especial, são bastante numerosos; como, por exemplo, *mainframe*) ou decalques (em que os verbos são destaque: *deletar*, *chipar*, etc.), outros são acrônimos ou abreviações usados indistintamente nas duas línguas (como *AGP*), outros ainda requerem uma consulta a

obras bilíngües já existentes para verificar, num primeiro momento, se são equivalentes (Platters – Discos, componentes do disco rígido; a primeira acepção de *platter*, segundo o American Heritage Dictionary, é o equivalente, em português, a *travessa* ou *prato grande*; o termo corrente em português, neste caso, é *disco*).

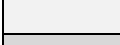




Na planilha 2, são mostradas as colunas das palavras-chave, numa planilha em Excel, indicando sua ordem pela chavicidade dos termos em cada língua.

Ordem	Português	Ordem	Inglês
1	COMPUTADOR	23	COMPUTER
2	IMPRESSORA	65	PRINTER
4	PLACA	782	BOARD
6	PROCESSADOR	31	PROCESSOR
8	USB	13	USB
10	BARRAMENTO	71	BUS
12	MEMÓRIA	2	MEMORY
15	PCI	8	PCI
18	CONTROLADOR	46	CONTROLLER
25	TECLADO	11	KEYBOARD
26	DISCO	132	PLATTERS
36	MHZ	203	MHZ
37	APLICATIVOS	88	APPLICATIONS
38	TECLA	199	KEY
45	AGP	237	AGP
47	DADOS	3	DATA
57	DISPOSITIVO	25	DEVICE
64	SERVIDOR	26	SERVER
67	INSTALAR	194	INSTALL
68	DRIVE	5	DRIVE
72	SCSI	45	SCSI
77	HTTP	82	HTTP
81	INTERFACE	18	INTERFACE
83	ROM	53	ROM
90	CHIP	142	CHIP
92	RÍGIDO	30	HARD
102	DRIVER	499	DRIVER
107	CONFIGURAÇÕES	181	SETTINGS
108	FIREWIRE	330	FIREWIRE
110	MAINFRAME	157	MAINFRAME
113	RAID	684	RAID

Planilha 2. Relação de termos equivalentes nas duas línguas.

A equivalência dos termos na listagem não garante, contudo, que todos eles apresentem contextos explicativos ou definitórios. Como o objetivo da tese é fornecer um ambiente de pesquisa que indique também a definição do termo, é necessário identificar um desses dois contextos, explicativos ou definitórios, para termos equivalentes nas duas línguas. Muitos termos, nessa comparação, não foram aprovados pela dificuldade em se achar contextos claros (já prevendo essa, foram selecionados com termos equivalentes em cada língua para haver uma margem de descarte). A planilha 3 apresenta a listagem parcial dos candidatos a termos equivalentes na área de computação. As escalas de cinza das legendas indicam os termos com contextos explicativos e/ou definitórios encontrados nas duas línguas, encontrados somente em uma língua ou não encontrados em nenhuma das duas¹². Conforme os termos eram inseridos no banco de dados, uma marca com tons de cinza ou preto também era feita ao lado. Os números, antepostos ao termo, assim como na planilha 2, indicam sua ordem de chavicidade.

	Português		Inglês
1	COMPUTADOR	23	COMPUTER
2	IMPRESSORA	65	PRINTER
4	PLACA	782	BOARD
6	PROCESSADOR	31	PROCESSOR
8	USB	13	USB
10	BARRAMENTO	71	BUS
12	MEMÓRIA	2	MEMORY
15	PCI	8	PCI
18	CONTROLADOR	46	CONTROLLER
25	TECLADO	11	KEYBOARD
26	DISCO	132	PLATTERS
36	MHZ	203	MHZ
37	APLICATIVOS	88	APPLICATIONS
38	TECLA	199	KEY
45	AGP	237	AGP
47	DADOS	3	DATA

Legenda	
	definição encontrada nas duas línguas
	definição não encontrada nas duas línguas
	definição encontrada em inglês, mas não em português
	definição encontrada em português, mas não em inglês
	adicionado ao banco de dados
	não adicionado ao banco de dados

¹² Uma possível ampliação do *corpus* de estudo, inclusive diacronicamente, deve fornecer todos os contextos necessários para o campo de definição dos termos.

57	DISPOSITIVO	25	DEVICE
64	SERVIDOR	26	SERVER
67	INSTALAR	194	INSTALL
68	DRIVE	5	DRIVE
72	SCSI	45	SCSI
77	HTTP	82	HTTP
81	INTERFACE	18	INTERFACE
83	ROM	53	ROM
90	CHIP	142	CHIP
92	RÍGIDO	30	HARD
102	DRIVER	499	DRIVER
107	CONFIGURAÇÕES	181	SETTINGS
108	FIREWIRE	330	FIREWIRE
110	MAINFRAME	157	MAINFRAME
113	RAID	684	RAID

Planilha 3. Área de computação, alguns candidatos a termos.

Para obter os contextos de cada termo, utilizamos uma terceira ferramenta do WordSmith Tools: o concordanciador (*Concordancer*). Ao selecionar o termo na listagem de palavras-chave e pedir suas concordâncias, o programa cria uma nova tela, com o termo em questão centralizado e na cor azul (tela KWIC, *key word in context*), mostrando suas ocorrências em todos os textos (figura 14). Basta clicar duas vezes na linha desejada, na coluna *File*, para que o texto seja mostrado por completo.

Para descobrir quais dessas linhas (cada uma representa a seleção de uma linha de um texto) podem nos fornecer os contextos desejados, foram usados, basicamente, dois artifícios:

1. uma busca por sinais de pontuação. Nos textos da figura 14, fiz, inicialmente, uma busca usando os parâmetros de : (dois pontos), (parênteses) ou , (vírgula). A idéia era achar esses contextos depois de pontuação (dois pontos ou parênteses) ou como aposto (entre vírgulas). Para realizar essa busca no programa, é necessário acrescentar o asterisco (*) depois da pontuação desejada. No caso do exemplo

acima, a busca seria realizada como: **computador:***, **computador** (* ou **computador,***;

N	Concordance	Set	Tag	Word #	t. #	os. %	# os.	t. #	os. %	File	%	
1	acesso a todos componentes PCI do computador Isto torna este padrão			1.649	76	0%	0	9%	0	9%	ra interação.txt	31%
2	do core ao barramento PCI do computador hospedeiro (host). Para			4.762	227	2%	0	4%	0	4%	ra interação.txt	83%
3	por Gerald Estrin [2]. Estrin propôs um "computador com estrutura fixa e			657	32	4%	0	2%	0	2%	ra interação.txt	13%
4	destas a placas de extensão de um computador. A forma encontrada para			1.486	69	0%	0	6%	0	6%	ra interação.txt	28%
5	receber os sinais PING provenientes do computador antes que este seja			898	48	4%	0	7%	0	7%	ramento usb.txt	73%
6	os dispositivos USB conectados ao computador e verifica se algum deles é o			921	49	7%	0	9%	0	9%	ramento usb.txt	75%
7	1 Número de comandos recebidos do computador byte 2 Número de			737	48	1%	0	3%	0	3%	ramento usb.txt	59%
8	de sinais Pacote de dados enviado ao computador Posição Função byte 0			763	48	5%	0	5%	0	5%	ramento usb.txt	61%
9	para a memória RAM. A partir dali, o computador está pronto para funcionar			986	82	0%	0	4%	0	4%	\barramento.txt	22%
10	1-56205-195-4, 1994. "Como funciona o computador III", Ron White, Quark, ISBN			6.979	321	2%	0	0%	0	0%	\barramento.txt	99%
11	por muitos como o primeiro computador, surgiu em 1942. Essa			484	54	9%	0	7%	0	7%	\barramento.txt	16%
12	a Xerox criou o Alto, um computador pessoal para ser usado em			677	63	8%	0	0%	0	0%	\barramento.txt	18%
13	automaticamente reconhecida pelo computador. Hoje em dia, os slots PCIs			479	18	0%	0	4%	0	4%	proprietários.txt	54%
14	de um barramento antigo). Como esse computador trabalhava a uma velocidade			215	8	5%	0	4%	0	4%	proprietários.txt	25%
15	na placa-mãe. O ISA surgiu no computador IBM PC, na versão de 8 bits			182	7	7%	0	1%	0	1%	proprietários.txt	21%
16	taxas de transferência de dados entre o computador em si e um dispositivo, por			343	14	0%	0	6%	0	6%	pci express.txt	26%
17	para a conexão de dispositivos ao computador, principalmente placas de			124	5	3%	0	0%	0	0%	pci express.txt	10%
18	interno) de 8 bits Usado no primeiro computador pessoal – Altair 8086b em			881	8	5%	0	1%	0	1%	putadores 2.txt	61%
19	a sua velocidade. 3_Componentes do Computador Prof. A. Neco Figura 3.44:			4.694	225	0%	0	5%	0	6%	mputadores.txt	76%
20	dos dados. 3_Componentes do Computador Prof. A. Neco 3.3.3.4			4.453	217	0%	0	1%	0	2%	mputadores.txt	72%
21	Para vídeo. 3_Componentes do Computador Prof. A. Neco Figura 3.46:			4.884	233	0%	0	9%	0	9%	mputadores.txt	79%
22	nao e prejudicial a performance do computador. As placas de rede tem			5.040	241	0%	0	2%	0	1%	mputadores.txt	81%
23	dos ambientes com mais de um computador sentem a necessidade de			5.001	240	0%	0	1%	0	0%	mputadores.txt	81%
24	respectivamente. 3_Componentes do Computador Prof. A. Neco Figura 3.39:			4.354	211	0%	0	9%	0	0%	mputadores.txt	70%
25	soaquete PGA 370 3_Componentes do Computador Prof. A. Neco rocessador			3.209	157	3%	0	8%	0	2%	mputadores.txt	52%
26	a 66MHz. 3_Componentes do Computador Prof. A. Neco Figura 3.28:			3.035	143	0%	0	5%	0	9%	mputadores.txt	49%
27	Figura 3.32 3_Componentes do Computador Prof. A. Neco 3.3.2.2			3.408	170	1%	0	2%	0	5%	mputadores.txt	55%
28	alem dessa marca. 3_Componentes do Computador Prof. A. Neco Figura 3.37:			4.108	199	0%	0	5%	0	6%	mputadores.txt	67%
29	dos 2 GHz. 3_Componentes do Computador Prof. A. Neco FIAura 3.35:			3.720	182	0%	0	8%	0	0%	mputadores.txt	60%
30	exemplo bem claro. Imaginemos um computador com processador, memória			5.608	268	4%	1	6%	0	0%	mputadores.txt	90%
31	aeral do micro. 3_Componentes do Computador Prof. A. Neco exemplificar			5.581	265	0%	1	2%	0	0%	mputadores.txt	90%
32	dos atuais, o que veremos sera um computador com desempenho limitado			5.633	268	3%	1	9%	0	1%	mputadores.txt	91%
33	o FAT32 e FAT16. 3_Componentes do Computador Prof. A. Neco 08/06/2007			6.208	284	0%	1	9%	0	0%	mputadores.txt	100%

Figura 14. Termo “computador”, em uma tela de concordâncias, totalizando 3.380 delas.

- uma busca pelos colocados (para o programa, colocado é a combinação de alta frequência entre o termo selecionado mais um termo a ele associado)¹³. O primeiro termo procurado como colocação foi o verbo *ser* (ou *to be*) em todas as suas formas. Veja na figura 15 as colocações para o termo “computador”: existem vinte e oito colocações com o verbo *ser* (*é*) no primeiro campo à direita de computador (coluna R1, seguinte à coluna “centre”). Ao clicar no número vinte e oito (em vermelho, no original; em destaque, aqui), a tela volta para a apresentação das concordâncias e destaca os segmentos de texto que apresentam essa combinação (figura 16). Na

¹³ Para Sardinha (2004, p. 40) é a “[...] associação entre itens lexicais, ou entre o léxico e campos semânticos”.

linha quatro dessa nova tela, por exemplo, temos um contexto definitório para o termo computador (“... podemos aprender que computador é uma máquina utilizada...”). Para ver todo o parágrafo, basta clicar duas vezes sobre a linha e o programa abre uma nova tela (figura 17).

N	Word	With elation	Total	tal Left	al Right	L5	L4	L3	L2	L1	Centre	R1	R2	R3	R4
1	COMPUTADOR	0,000	3.511	67	67	29	18	14	5	1	3.377	1	5	14	18
2	DE COMPUTADOR	0,000	2.435	696	1.739	149	175	192	128	52	0	1.260	60	78	171
3	O COMPUTADOR	0,000	1.944	1.507	437	108	116	97	73	1.113	0	25	113	71	172
4	MÃO COMPUTADOR	0,000	1.115	40	1.075	14	11	10	3	2	0	0	1.067	0	0
5	DO COMPUTADOR	0,000	942	828	114	36	46	41	42	663	0	9	25	16	27
6	A COMPUTADOR	0,000	676	320	356	90	96	113	21	0	0	56	75	102	60
7	E COMPUTADOR	0,000	613	203	410	52	48	61	39	3	0	154	28	109	47
8	PARA COMPUTADOR	0,000	513	321	192	59	67	101	81	13	0	72	11	51	22
9	NO COMPUTADOR	0,000	509	427	82	11	19	18	31	348	0	16	13	15	23
10	UM COMPUTADOR	0,000	486	342	144	40	33	37	13	219	0	5	48	21	36
11	PALM COMPUTADOR	0,000	412	21	391	3	5	7	6	0	0	0	0	362	3
12	Z22 COMPUTADOR	0,000	374	18	356	15	0	0	0	3	0	0	0	0	356
13	QUE COMPUTADOR	0,000	351	203	148	50	40	36	76	1	0	31	28	39	26
14	COM COMPUTADOR	0,000	313	145	168	30	18	21	76	0	0	39	17	70	23
15	EM COMPUTADOR	0,000	254	156	98	17	36	25	76	2	0	26	16	23	21
16	AO COMPUTADOR	0,000	236	200	36	10	9	11	12	158	0	7	7	16	3
17	OU COMPUTADOR	0,000	233	79	154	18	17	27	12	5	0	55	27	31	20
18	SE COMPUTADOR	0,000	226	143	83	22	31	16	74	0	0	20	17	20	13
19	INFORMAÇÕES COMPUTADOR	0,000	204	162	42	17	18	59	68	0	0	0	5	8	8
20	SEU COMPUTADOR	0,000	190	174	16	5	6	1	0	162	0	3	1	3	7
21	MESA COMPUTADOR	0,000	189	24	165	10	3	8	3	0	0	0	155	1	0
22	É COMPUTADOR	0,000	155	70	85	18	27	17	8	0	0	28	12	20	19
23	NÃO COMPUTADOR	0,000	152	67	85	32	14	20	1	0	0	20	10	34	12
24	DA COMPUTADOR	0,000	147	73	74	20	22	28	3	0	0	13	14	9	20
25	AS COMPUTADOR	0,000	143	67	76	16	26	22	3	0	0	8	25	12	15
26	UMA COMPUTADOR	0,000	141	40	101	18	11	10	1	0	0	5	39	10	31
27	OS COMPUTADOR	0,000	141	71	70	24	19	27	1	0	0	5	13	18	22
28	COMO COMPUTADOR	0,000	128	72	56	9	10	26	22	5	0	19	11	6	11
29	NA COMPUTADOR	0,000	119	73	46	21	21	31	0	0	0	9	3	12	10
30	QUANDO COMPUTADOR	0,000	95	67	28	3	19	11	34	0	0	12	6	3	4
31	MAIS COMPUTADOR	0,000	93	39	54	13	7	18	1	0	0	9	4	21	11
32	ESTÁ COMPUTADOR	0,000	80	27	53	3	8	15	1	0	0	26	5	13	4

Figura 15. Lista de colocações do termo *computador*.

N	Concordance	Set	Tag	Word #	t. #	os. #	os. #	os. #	os. #	File	%	
1	Ricardo Hardware e Software O termo "Computador" é utilizado hoje em dia para			429	21	9%	0	2%	0	2%	informatics.txt	13%
2	natural: se a aplicação para o qual o computador é utilizado manipula grande			714	35	4%	0	1%	0	1%	s de 32 bits.txt	60%
3	pela oferta de produtos em que o computador é usado para criar e manter			1.628	58	3%	0	5%	0	5%	omunicação.txt	25%
4	Por enquanto podemos aprender que "Computador é uma máquina utilizada			459	18	5%	0	1%	0	1%	ática básica.txt	11%
5	aparelho eletrônico? Pois desligar um computador é uma operação muito mais			29	1	5%	0	4%	0	4%	ligando o pc.txt	4%
6	é essencial. Principalmente se seu computador é um notebook que			55	1	3%	0	4%	0	4%	sete chaves.txt	15%
7	LOCAL Em teoria, escolher um computador é simples: basta verificar as			31	0	8%	0	1%	0	1%	permercado.txt	1%
8	maior que a suportada pelo seu computador, é recomendável que você			17.480	925	6%	0	5%	0	5%	itude™ d510.txt	85%
9	realmente arrasadoras, uma parada no computador é quase obrigatória. Com			2.665	144	3%	0	6%	0	6%	da imagem.txt	37%
10	assistir filmes em DVD usando um computador é preciso que ele tenha uma			1.675	104	6%	0	4%	0	4%	a informação.txt	54%
11	A principal função do teclado no computador é permitir que você possa			11	0	6%	0	1%	0	1%	na o teclado.txt	1%
12	ou o Processador O cérebro de um computador é o que chamamos de			580	33	9%	0	7%	0	7%	informatics.txt	17%
13	de software. A IHC - Interação Humano-Computador - é o campo de pesquisa			583	17	3%	0	0%	0	0%	putacionais.txt	10%
14	aparece sempre que o cabo ou o computador é ligado, desactive o			12.519	968	6%	0	8%	0	8%	do utilizador.txt	19%
15	status do dispositivo Acende quando o computador é ligado ou pisca quando ele			476	20	7%	0	2%	0	2%	itude™ d510.txt	3%
16	aparece sempre que o cabo ou o computador é ligado, desactive o			47.927	427	6%	0	8%	0	8%	do utilizador.txt	69%
17	Computador X Gabinete A rigor o computador é formado pelo gabinete e			344	16	6%	0	1%	0	1%	a informação.txt	12%
18	A conexão dos drives de CD-ROM ao computador é feita através da parte			136	6	8%	0	6%	0	2%	ve de cd-rom.txt	13%
19	troca deles -afinal, a interação com o computador é essencialmente tátil. Um			2.007	89	8%	0	0%	0	0%	permercado.txt	81%
20	Nela, os dados se perdem quando o computador é desligado. Os módulos de			447	24	2%	0	1%	0	1%	ra iniciantes.txt	22%
21	diminuir para um nível satisfatório, o computador é desligado, de forma a			89	3	4%	0	3%	0	3%	o pentium 4.txt	14%
22	artigos sobre Periféricos Introdução O computador é cheio de barramentos -			49	8	8%	0	2%	0	2%	ciona o scs.txt	2%
23	vídeo 3D, disco e memória que o computador é capaz de fornecer. Nós			2.555	81	2%	0	8%	0	8%	100+ e 3100+.txt	59%
24	fixa e nenhum código de bipe, mas o computador é bloqueado durante o			11.792	539	7%	0	9%	0	9%	tipler™ 1701.txt	47%
25	InfoWay Note M3420, da Itautec. Esse computador é baseado no chip set			14.281	910	6%	0	8%	0	8%	ncia do wi-fi.txt	98%
26	o scanner para qualquer local, o computador é apenas necessário para			4.634	302	1%	0	5%	0	5%	1700 docupen.txt	76%
27	de expandir a visualização do seu computador é adicionar um segundo			4.882	283	1%	24	4%	0	5%	computador.txt	84%
28	Outro componente fundamental do Computador é a Memória RAM (do			903	50	9%	0	6%	0	6%	informatics.txt	26%
29	SE e Me 4. Como posso proteger meu computador desse worm? Você deve			460	27	5%	0	8%	0	8%	bre o sasser.txt	36%
30	ataque? 4. Como posso proteger meu computador desse worm? 5. O que é um			40	4	8%	0	3%	0	3%	bre o sasser.txt	3%
31	se a unidade não aparecer em Meu computador, no Windows Explorer ou			206	12	1%	0	4%	0	4%	usb firewire.txt	22%
32	O ícone da unidade aparecerá em Meu computador, no Windows Explorer ou na			101	6	3%	0	2%	0	2%	usb firewire.txt	11%
33	para instalar o Palm Desktop em um computador com Windows 2000/XP ou			3.115	336	8%	0	6%	0	6%	uês\sumário.txt	10%

Figura 16. Colocações de *computador* + “é”

N	Concordance	Set	Tag	Word #	t. #	os. #	os. #	os. #	os. #	File	%
1.1.12	Placa de som/kit multimídia										
1.2	Como funciona o computador										
1.2.1	O que é um software? Um Programa ou Sistema?										
1.2.2	Ligando/desligando o computador										
1.2.3	Problemas frequentes										
1.3	Rede de computadores										
1.4	Estrutura de um Computador										
1.4.1	Unidade de disco (lógica)										
1.4.2	Diretório ou pasta										
1.4.3	Arquivo										
1.5	Localizando uma Informação										
1.6	Informações sobre o mundo da informática										
1.6.1	Vírus de computador										
1.6.2	Pirataria										
1.6.3	Segurança da informação										
1.6.4	Profissionais da área de informática										
1.7	Falando um pouco de INTERNET										
1.8	Unidades de Medidas em informática										
- 1 -											
1.	Introdução: Noções de microinformática										
1.1	Conhecendo o computador										
	A palavra informática pode ser traduzida como a ciência que trata e usa a Informação. Outro conceito utilizado é de que toda informação, ou dado, deve sofrer alguma mudança ou alteração para que possa ser repassada adiante. Po exemplo em uma soma de dois números, mesmo se feita mentalmente, nosso cérebro precisa conhecer quais os números que serão somados, além das regras de uma soma. Então nos apresentará qual é o resultado. Nisto tudo foi dado um tratamento nas informações (os dois números somados e o resultado), ou seja, houve um Processamento dos Dados para alcançar um resultado.										
	É aí que entra o computador, como sendo uma máquina inventada pelo homem para auxiliar na execução de tarefas que tratem a informação e processem dados de forma rápida segura e confiável.										
	Uma calculadora, por exemplo, usa forma de cálculos que permite com que tenhamos um resultado de uma operação matemática (soma, multiplicação, etc.) muitas vezes mais rápida do que se tivéssemos que calcular usando lápis e papel ou até mesmo de cabeça.										
	O computador funciona da mesma forma. Ele calcula o resultado rapidamente e nos apresenta a resposta. O que difere e torna-o mais completo que a calculadora do nosso exemplo, é que o computador possui formas diferentes de mostrar o resultado, seja na tela, impresso, ou até mesmo falando a resposta. Além disso o computador tem mais utilidade, executando tarefas bem mais complexas que cálculos matemáticos, por exemplo ser utilizado para substituir a máquina de escrever na edição de textos, ou para desenhar, etc.										
	O seu funcionamento dá-se através de comandos, ou seja, há a necessidade de dizermos de alguma forma como deve ser feito cada passo daquilo que queremos que seja executado. A isto dá-se o nome de comando ou instrução.										
	No final deste módulo veremos uma explicação mais completa sobre o computador. Por enquanto podemos aprender que "Computador é uma máquina utilizada para auxiliar o homem na execução de tarefas que exigem rapidez e qualidade no tratamento de informações. Porém ele não possui vontade própria, sendo necessário repassar de forma clara e completa as instruções de como deve ser feito o trabalho solicitado."										
- 2 -											
1.1.1	Tipos de computadores										
	PC Personal Computer (Computador pessoal): É o mais utilizado. Lançado inicialmente pela IBM, atualmente são vários os fabricantes deste tipo de máquina. Foi definido para uso pessoal das pessoas, sendo conhecido também como DeskTop (Em cima da mesa).										
	WorkStation (Estação de trabalho): É uma máquina de uso mais profissional (por isso mais cara) e caracteriza-se pela grande capacidade e velocidade no processamento de informações. Muito usado em atividades que usam desenhos e gráficos.										

Figura 17. Contexto da quarta linha de concordância (figura anterior).

A busca através desses mecanismos nem sempre retorna contextos definitórios, que são aqueles mais fáceis de serem incluídos no banco de dados. A busca por outras colocações pode fornecer pistas para contextos explicativos que, somados, podem criar uma definição.

4.6 Mais corpora

Terminada a fase acima, decidiu-se pela elaboração de novos *corpora*, dessa vez no campo da Lingüística, para que não houvesse a necessidade de explicar termos pertinentes desse campo na “Ajuda Online” do site em desenvolvimento. A idéia era que houvesse um sistema de metalinguagem. Cada vez que o aluno tivesse uma dúvida sobre um termo do campo da Lingüística que aparecesse na microestrutura do site, bastaria consultar esse termo no próprio site. Para isso, a construção de novos *corpora* se fez necessária.

Todos os passos descritos nos itens anteriores foram realizados novamente e uma nova árvore de campo foi criada. Dessa vez, contudo, não houve a necessidade de se desdobrar mais do que três sub-níveis da árvore, já que o objetivo desses *corpora* é diferente. Essa nova árvore ficou configurada como na figura 18.

Assim como nos *corpora* anteriores, esses contam com, no mínimo, vinte mil palavras em cada domínio. O *corpus* em português totalizou 1.309.967 palavras e o *corpus* em inglês totalizou 1.921.811 palavras.

Árvore do Campo da Lingüística

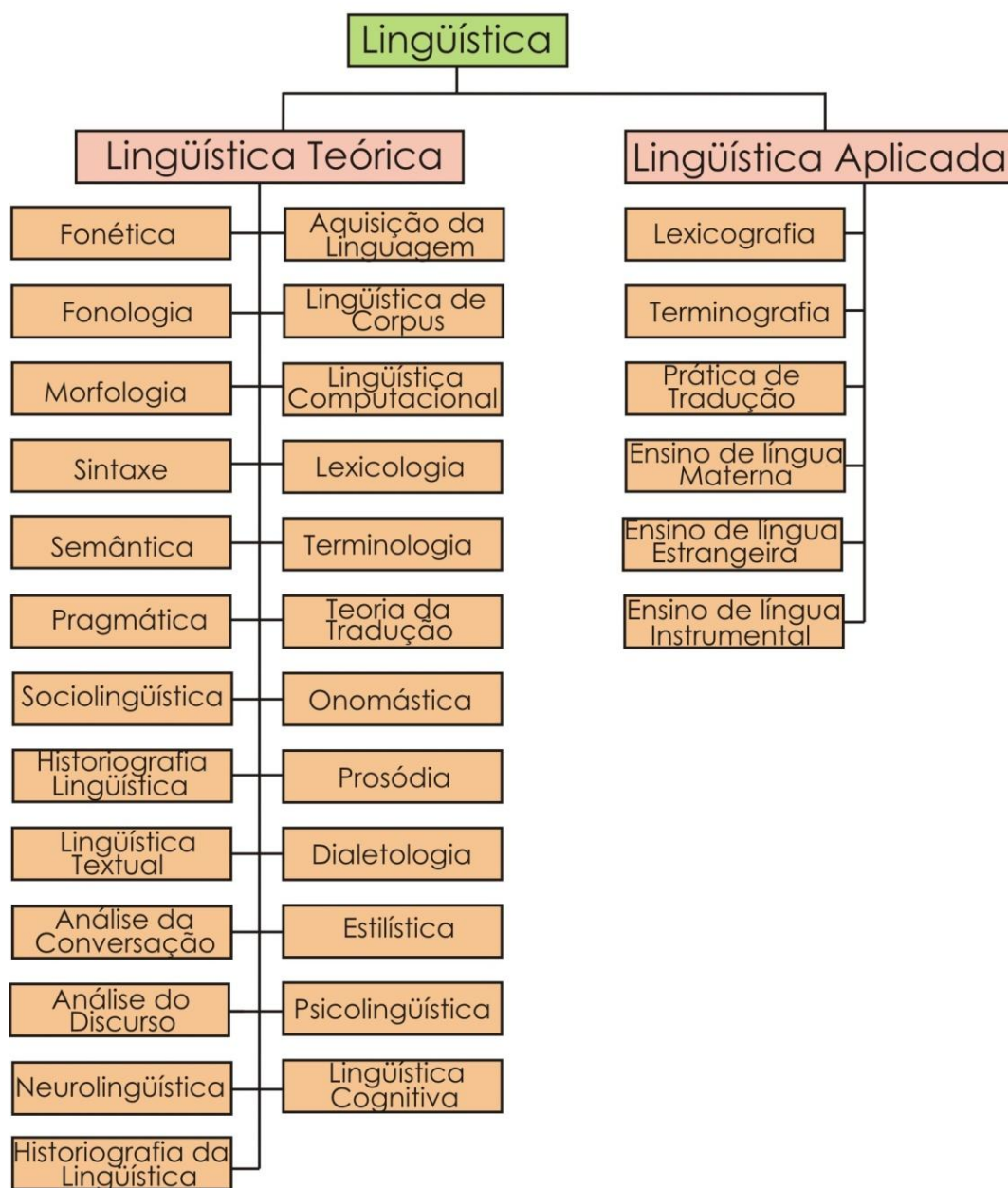


Figura 18. Árvore do Campo da Lingüística

4.7 Em suma

O projeto prevê que os *corpora* construídos para esta tese sejam dinâmicos, isto é, novos textos e áreas (com os respectivos domínios e subdomínios) poderão ser acrescentados no futuro para aumentar sua precisão e escopos de análise. O aumento do

corpus implica, porém, a atualização de dados referentes ao *corpus* para cada termo no banco de dados (frequência no *corpus* e número total de exemplos encontrados).

No momento foram incluídos somente textos escritos, já que os mesmos representam bem os campos técnicos, mas nada impede que futuramente outros tipos de texto (como os orais) sejam adicionados.

É de extrema importância notar que, diferente de algumas ferramentas disponíveis na Internet (como o Corpógrafo), a presente proposta não oferece uma solução de armazenamento do corpus. O mesmo deve ficar disponibilizado no computador do pesquisador. Somente os contextos dele extraídos é que serão armazenados no banco de dados.

5. A opinião de quem já chegou lá: os tradutores profissionais

Miranda e Farias (2006) comentam, quanto aos dicionários gerais e escolares da língua portuguesa, que os mesmos possuem um elevado número de informações não discretas, ou seja, que não têm nenhuma relevância para os consulentes, como a etimologia em dicionários escolares, e/ou não discriminantes, que não acrescentam nada ao seu conhecimento de língua, como exemplos não pertinentes. Houve, portanto, desde o começo deste trabalho, uma preocupação em não repetir erros que essas obras cometem. Fica clara, nos casos citados pelos autores, uma falta de perspectiva quanto ao público-alvo, embora comentem que o dicionário é concebido para um número amplo de falantes e a composição da microestrutura.

Antes de trabalhar com nosso público-alvo, que são os aprendizes de tradução, considerei importante, para o desenvolvimento de um produto melhor direcionado e funcional, conhecer a opinião de quem já atua na área. Para tanto, foi aplicado um questionário, disponibilizado na Internet, em formato bilíngüe (português/inglês), de março a dezembro de 2004. Esse questionário foi dirigido aos tradutores já em atividade e chegou em suas mãos por meio de contatos com cursos superiores de tradução e listas de tradução na Internet. Foram retornados 177 questionários respondidos. Por uma falha de programação, as pessoas que acessaram a pesquisa não foram “obrigadas” a responder todas as perguntas. Em virtude disso, muitas respostas não representam 100% dos 177 acessos à mesma.

O objetivo primário da pesquisa era levantar o modo pelo qual o tradutor se relaciona com sua principal ferramenta de trabalho, o dicionário¹⁴. Algumas perguntas não se propunham ao levantamento referido, mas aproveitei o meio eletrônico para apreender algumas idéias sobre o profissional. Toma-se como exemplo a pergunta seis, que não é respondida neste capítulo: quais os nomes dos três dicionários que você mais usa? Foram obtidas 516 respostas (apêndice C). A informação não foi incluída na análise da pesquisa porque muitas obras se repetem, porém com nomes parcialmente diferentes. Um novo estudo, a partir desses dados, pode ser elaborado no futuro.

Analisando, a seguir, cada pergunta, especificando os motivos (minhas pressuposições como profissional e pesquisador da área) que me levaram a formulá-la, as respostas obtidas e como contribuíram para o objetivo final da tese.

1. Quantos dicionários você costuma usar durante uma consulta?

	Respostas	%
1	5	3
2	33	19
3	68	38
4	14	8
mais de 4	52	29
Total	172	

Motivação da pergunta: parto da hipótese de que quanto mais obras o tradutor usa, menor é, provavelmente, a sua velocidade de tradução.

¹⁴ Foi usado, na pesquisa, o termo dicionário: ele é mais comum em meio não acadêmico. É importante ressaltar que o termo torna-se, então, hiperônimo de qualquer obra de consulta lexicográfica e/ou terminográfica.

Resultado: o número de profissionais que usam somente uma obra é quase irrelevante; notei que a maioria usa ‘três’ ou ‘mais de quatro’ obras.

Relevância para a tese: prova de que há a necessidade de busca em várias obras, o que diminui a velocidade de tradução. O profissional se beneficiaria bastante consultando apenas uma obra de conteúdo abrangente ou um local que reunisse vários tipos de obras.

2. Quais os dicionários que você mais costuma pesquisar? (mais de uma resposta possível).

	Resp.	%
gerais monolíngües	97	55
gerais bilíngües	115	65
técnicos monolíngües	62	35
técnicos bilíngües	134	76
técnicos multilíngües	22	12

Motivação: descobrir o quanto os tradutores usam dicionários gerais de língua e dicionários técnicos, monolíngües ou multilíngües.

Resultado: observa-se que a maioria usa tanto dicionários gerais quanto técnicos na tradução, ambos bilíngües (o que era de se esperar).

Relevância para a tese: mostra que a maior faixa de uso, por parte dos tradutores, é de dicionários técnicos bilíngües, e que é nesse viés que devo investir; a consulta a obras monolíngües, porém, não deve ser descartada.

3. Dentre os seguintes meios de busca para uma palavra, qual o grau de eficiência de cada um? Classifique: 1 = nenhum, 2 = pouco, 3 = médio, 4 = bom e 5 = alto.

	1		2		3		4		5	
Dicionários Impressos	15	8%	21	12%	44	25%	52	29%	39	22%
Dicionários em CD-ROM	17	10%	20	11%	32	18%	44	25%	50	28%
Dicionários pela Internet	10	6%	24	14%	40	23%	56	32%	38	21%
Download de glossários	18	10%	32	18%	48	27%	44	25%	16	9%
Glossários na Internet	5	3%	33	19%	53	30%	53	30%	23	13%
Listas de Discussão na Internet	14	8%	28	16%	37	21%	53	30%	32	18%
Consulta a especialistas	20	11%	11	6%	38	21%	48	27%	52	29%
<i>Corpus</i> on-line	25	14%	34	19%	40	23%	26	15%	19	11%
Busca na Internet (Google, etc.)	5	3%	10	6%	27	15%	57	32%	65	37%
Textos impressos semelhantes	15	8%	34	19%	49	28%	38	21%	22	12%

Motivação: acredito que, ao oferecer uma nova ferramenta de tradução, preciso, antes, conhecer o relacionamento dos profissionais com os produtos que já estão disponíveis. Há a necessidade de conhecer o grau de eficiência de cada um para montar a estratégia de composição do vocabulário proposto.

Resultado: foi tomado como base, aqui, um índice de aprovação entre médio (3) e alto (5)¹⁵.

- a. Dicionários impressos: apesar dos avanços da tecnologia, ainda são bastante usados; na classificação de médio a alto, o papel ainda é importante para 76% dos tradutores;
- b. dicionários em CD-ROM: por acompanharem a mesma estrutura dos dicionários impressos, também têm alto índice de aceitação (71%);
- c. dicionários pela Internet: embora seja difícil identificar a procedência ou seriedade da maioria dos dicionários disponíveis na Internet¹⁶, os profissionais

¹⁵ Essa opção se aplicará às demais perguntas formuladas no mesmo estilo, já que esses índices representam de 50 a 100% de aprovação por parte dos tradutores.

¹⁶ Esse dado refere-se à minha experiência como tradutor e à busca constante por novas fontes de consulta.

- parecem acreditar bastante na sua eficiência com 76% de aprovação, ou seja, tanto quanto os dicionários impressos e mais do que os dicionários em CD-ROM;
- d. download de glossários: como no item anterior, a maioria não mostra sua procedência ou a seriedade de quem os compôs; diferente do item anterior, o índice de aprovação cai para 61%, provavelmente revelando uma certa percepção dos profissionais para o problema de confiabilidade da fonte;
 - e. listas de discussão na Internet: com 69% de aprovação, a discussão com colegas da área parece ser um recurso bastante usado;
 - f. consulta a especialistas: requisitar ajuda à “fonte”, para 77% dos entrevistados, é ainda mais importante do que a busca por dicionários disponíveis na Internet;
 - g. *corpus* on-line: com um índice de aprovação de 49%, inferior à média, a proposta de busca por *corpora* na Internet provavelmente indica o estágio inicial no qual essa prática ainda se encontra e a dificuldade dos tradutores em achar instrumentos que transformem dados dispersos em um *corpus* eficiente de consulta;
 - h. busca na Internet: a procura de contextos definitórios ou explicativos através das ferramentas de busca parece ser, atualmente, a grande preferência por parte dos profissionais da área (84%), superando qualquer outro tipo de instrumento;
 - i. textos impressos semelhantes: embora não use computador, essa técnica ainda é bastante usada¹⁷.

Relevância para a tese: indica que os tradutores já estão preparados para trabalhar com obras em formato digital, embora não indique se essas estão formatadas de modo

¹⁷ Provavelmente porque o profissional dispõe de grandes quantidades de textos, originais e traduzidos, nas áreas em que atua.

tradicional ou já usem novas tecnologias de busca. O acesso à banda larga, na Internet, sugere que consultas a bases de dados ali disponíveis deve crescer bastante nos próximos anos.

4. Dentro do processo de busca por uma tradução em dicionários bilíngües, o que é mais comum para você?

	Respostas	%
buscar somente a tradução (forma equivalente) da palavra	18	10
buscar a definição da palavra	3	2
buscar a exemplificação apresentada para a palavra	7	4
buscar a tradução e a definição da palavra	19	11
buscar a definição e a exemplificação da palavra	11	6
buscar a tradução (forma equivalente), definição e exemplificação da palavra	115	65

Motivação: antes de montar um sistema específico para o tradutor, é preciso saber quais partes da microestrutura o profissional costuma acessar com mais frequência¹⁸: o Paradigma Definicional (PD), o Paradigma Pragmático (PP) ou o Paradigma de Forma Equivalente (PFE).

Resultado:

- a. buscar somente a tradução (forma equivalente) da palavra: embora a porcentagem seja baixa (10%), ela está em terceiro lugar na tabela, ou seja, ainda é um paradigma bastante procurado;
- b. buscar a definição da palavra: com o menor índice (2%), mostra que a leitura completa da definição, provavelmente, desacelera a velocidade de tradução e, ao mesmo tempo, não fornece a tradução que o profissional precisa;

¹⁸ Para o Paradigma Informacional (PI), elaboramos uma pergunta específica, a de número 7.

- c. buscar a exemplificação apresentada pela palavra: os exemplos são, para apenas 4% dos entrevistados, mais importantes do que os demais itens, ou seja, apenas o uso do termo no contexto não parece ser um fator para a compreensão do mesmo;
- d. buscar a tradução e a definição da palavra: em segundo lugar (11%), indica que somente associar a tradução à definição não é suficiente para a compreensão total do termo;
- e. buscar a definição e a exemplificação da palavra: a baixa porcentagem de escolha (6%) parece indicar que o tradutor está sempre, como no item 'b', em busca de uma tradução pronta, já que aqui ele necessita criar uma a partir do seu entendimento;
- f. buscar a tradução (forma equivalente), definição e exemplificação da palavra: a forma completa ainda é a mais procurada pelos tradutores (65%), o que sugere que somente a partir da composição dos três paradigmas ele pode desambigüizar possíveis traduções errôneas e precisar melhor a sua tradução.

Relevância para a tese: mostra a necessidade de trabalhar com todos os três paradigmas básicos (Informacional, Definicional e Pragmático), mas também apresenta, nos itens menos valorizados, qual pode ser a ordem dos fatores dentro da microestrutura; indica também que, embora poucos, alguns tradutores buscam somente a tradução de um determinado termo, o que está em consonância com muitas obras impressas disponíveis na praça.

5. Qual o tipo de dicionário bilíngüe ou multilíngüe que você mais utiliza?

	Respostas	%
o que apresenta somente a tradução da palavra	14	8%
o que apresenta a tradução e exemplos na língua de chegada	34	19%
o que apresenta a tradução e a definição da palavra na língua de chegada	19	11%
o que apresenta a tradução e a definição da palavra na língua de chegada e exemplificação na língua de partida	23	13%
o que apresenta tradução, definição, exemplificação e um sistema de remissivas (tipo VER:)	22	12%
o que apresenta a definição na língua de partida e o equivalente e exemplos na língua de chegada	22	12%
o que apresenta todos as possibilidades acima	41	23%

Motivação: tentar descobrir que tipo de microestrutura o consulente usa para a sua tradução. Foram apresentados, aos entrevistados, alguns tipos de composições internas mais comuns, conforme experiência própria, pesquisa prévia a dicionários e vocabulários (FROMM, 2002) e as possibilidades apresentadas por alguns teóricos (como ATKINS; FILLMORE; JOHNSON, 2003) .

Resultado:

- a. o que apresenta somente a tradução da palavra: é interessante notar que, embora existam muitas obras disponíveis no mercado, que só apresentam o PFE, esse tipo de obra é a menos consultada por tradutores, indicando que só esse paradigma é insuficiente para o profissional;
- b. o que apresenta a tradução e exemplos na língua de chegada: em segundo lugar, parece mostrar que, às vezes, o exemplo, associado à tradução, pode substituir a definição, considerada essencial pela maioria dos lexicólogos e terminólogos;

c. o que apresenta a tradução e a definição da palavra na língua de chegada; d. o que apresenta a tradução e a definição da palavra na língua de chegada e exemplificação na língua de partida; e. o que apresenta tradução, definição, exemplificação e um sistema de remissivas (tipo VER:): composições menos “clássicas” de uma microestrutura parecem não agradar aos consulentes, ainda que muitas obras (como aquelas apresentadas no capítulo 3) sigam esse tipo de construção hoje em dia;

f. o que apresenta a definição na língua de partida e o equivalente e exemplos na língua de chegada: a inversão dos paradigmas na microestrutura parece também não agradar aos consulentes;

g. o que apresenta todas as possibilidades acima: embora a combinação de todas as acima resultasse numa microestrutura bastante incomum, os tradutores parecem optar pelo “quanto mais completo, melhor”.

Relevância para a tese: os tradutores acreditam na necessidade da apresentação de uma microestrutura a mais completa possível, com destaque para a exemplificação na língua de chegada e a forma equivalente do termo. Inversões dos paradigmas e novas propostas de arranjo na microestrutura parecem causar estranheza.

7. Dentre as seguintes informações que um verbete pode apresentar, antes da definição do mesmo, quais você considera importantes para sua tarefa de tradutor? Classifique-as de pouco importantes (1) até muito importantes (5).

	1		2		3		4		5	
Abreviações, Abreviaturas, Siglas, Acrônimos	16	9%	23	13%	41	23%	33	19%	38	21%
Categoria Morfológica	29	16%	32	18%	46	26%	25	14%	25	14%
Gênero (masculino, feminino, neutro)	18	10%	30	17%	41	23%	28	16%	41	23%
Número (singular, plural)	20	11%	29	16%	45	25%	31	18%	29	16%
Pronúncia	48	27%	37	21%	34	19%	21	12%	14	8%
Tradução (ões)	13	7%	5	3%	6	3%	24	14%	112	63%
Conjugação verbal	21	12%	24	14%	45	25%	44	25%	25	14%
Classificação verbal (transitivo, intransitivo, etc.)	20	11%	23	13%	44	25%	37	21%	35	20%
Homônimos	20	11%	34	19%	40	23%	40	23%	26	15%
Sinônimos	12	7%	18	10%	35	20%	47	27%	51	29%
Etimologia	31	18%	35	20%	49	28%	27	15%	13	7%
Níveis de linguagem (popular, gíria, regionalismos, etc.)	9	5%	22	12%	43	24%	43	24%	45	25%
Diferenças Ortográficas	18	10%	30	17%	46	26%	38	21%	22	12%
Divisão Silábica	42	24%	40	23%	36	20%	26	15%	14	8%
Pragmática (uso, campo de utilização da palavra)	9	5%	14	8%	35	20%	43	24%	53	30%
Área de Especialidade	10	6%	12	7%	30	17%	42	24%	70	40%
Vida útil da palavra (neologismo, arcaísmo, etc.)	21	12%	36	20%	51	29%	28	16%	26	15%

Motivação: tentativa de descobrir, entre várias possibilidades de construção, a composição ideal do Paradigma Informacional (que congrega os dados pré-definição) para os tradutores. O levantamento dessas possibilidades foi feito a partir da bibliografia e de dicionários disponíveis no mercado (conforme análises de FROMM, 2002).

Resultado:

Abreviações, Abreviaturas, Siglas, Acrônimos	112	63%
Categoria Morfológica	96	54%
Gênero (masculino, feminino, neutro)	110	62%
Número (singular, plural)	105	59%
Pronúncia	69	39%
Tradução (ões)	142	80%
Conjugação verbal	114	64%
Classificação verbal (transitivo, intransitivo, etc.)	116	66%
Homônimos	106	61%
Sinônimos	133	76%
Etimologia	89	50%

Níveis de linguagem (popular, gíria, regionalismos, etc.)	131	73%
Diferenças Ortográficas	106	59%
Divisão Silábica	76	43%
Pragmática (uso, campo de utilização da palavra)	131	74%
Área de Especialidade	142	81%
Vida útil da palavra (neologismo, arcaísmo, etc.)	105	60%

A mesma tabela, dessa vez em ordem decrescente:

Área de Especialidade	142	81%
Tradução (ões)	142	80%
Sinônimos	133	76%
Pragmática (uso, campo de utilização da palavra)	131	74%
Níveis de linguagem (popular, gíria, regionalismos, etc.)	131	73%
Classificação verbal (transitivo, intransitivo, etc.)	116	66%
Conjugação verbal	114	64%
Abreviações, Abreviaturas, Siglas, Acrônimos	112	63%
Gênero (masculino, feminino, neutro)	110	62%
Homônimos	106	61%
Vida útil da palavra (neologismo, arcaísmo, etc.)	105	60%
Número (singular, plural)	105	59%
Diferenças Ortográficas	106	59%
Categoria Morfológica	96	54%
Etimologia	89	50%
Divisão Silábica	76	43%
Pronúncia	69	39%

Relevância para a tese: com esses resultados, pode-se montar um exemplo, numa ordem decrescente de escolha por parte dos consulentes, do que seria um Paradigma Informacional¹⁹ para os tradutores:

¹⁹ Para compor essa microestrutura, os componentes foram em ordem decrescente de frequência a partir da tabela anterior; se o item apresenta 75% ou mais (o critério de corte é nosso), item obrigatório (+); com menos de 75%, item opcional (±); com menos de 50%: item descartado.

PI: +área de especialidade +tradução +sinônimos ±pragmática ±níveis de linguagem
 ±classificação verbal ±conjugação verbal ±abreviações ±gênero ±homônimos ±vida útil
 da palavra ±número ±diferenças ortográficas ± categoria morfológica ±etimologia.

8. Dentre as seguintes informações que um verbete pode apresentar, depois da definição do mesmo, quais você considera importantes para sua tarefa de tradutor? Classifique-as de pouco importantes (1) até muito importantes (5).

	1		2		3		4		5	
Exemplificações	10	6%	8	5%	11	6%	33	19%	104	59%
Sinônimos	12	7%	14	8%	16	9%	62	35%	62	35%
Antônimos	20	11%	24	14%	40	23%	48	27%	31	18%
Expressões, ditados, citações com essa palavra	7	4%	15	8%	29	16%	54	31%	56	32%
Palavras relacionadas/referências cruzadas (remissiva)	7	4%	23	13%	40	23%	56	32%	39	22%
Outras categorias morfológicas (ex.: adjetivo, advérbio, etc.) da mesma palavra	17	10%	29	16%	48	27%	47	27%	18	10%
Verbos preposicionados	8	5%	17	10%	29	16%	47	27%	64	36%
Verbos frasais (Phrasal Verbs)	8	5%	21	12%	25	14%	55	31%	54	31%
Notas culturais	12	7%	35	20%	41	23%	42	24%	34	19%
Figuras (desenhos)	16	9%	53	30%	37	21%	34	19%	22	12%
Palavras derivadas (sem definição)	27	15%	50	28%	51	29%	29	16%	6	3%
Pragmática (uso, campo de utilização da palavra)	9	5%	30	17%	41	23%	38	21%	48	27%

Motivação: de modo semelhante à pergunta anterior, levantamos, dessa vez, a composição do Paradigma Pragmático (normalmente os dados pós-definição).

Resultado:

Exemplificações	148	84%
Sinônimos	140	79%
Antônimos	119	68%
Expressões, ditados, citações com essa palavra	139	79%
Palavras relacionadas/referências cruzadas (remissiva)	135	77%
Outras categorias morfológicas (ex.: adjetivo, advérbio, etc.) da mesma palavra	113	64%
Verbos preposicionados	140	79%
Verbos frasais (Phrasal Verbs)	134	76%
Notas culturais	117	66%
Figuras (desenhos)	93	52%
Palavras derivadas (sem definição)	86	48%
Pragmática (uso, campo de utilização da palavra)	127	71%

A mesma tabela, dessa vez em ordem decrescente:

Exemplificações	148	84%
Sinônimos	140	79%
Expressões, ditados, citações com essa palavra	139	79%
Verbos preposicionados	140	79%
Palavras relacionadas/referências cruzadas (remissiva)	135	77%
Verbos frasais (Phrasal Verbs)	134	76%
Pragmática (uso, campo de utilização da palavra)	127	71%
Antônimos	119	68%
Notas culturais	117	66%
Outras categorias morfológicas (ex.: adjetivo, advérbio, etc.) da mesma palavra	113	64%
Figuras (desenhos)	93	52%
Palavras derivadas (sem definição)	86	48%

Relevância para a tese: com esses resultados, pode-se montar um exemplo, numa ordem decrescente de escolha por parte dos consulentes, de um possível Paradigma Pragmático:

PP: +exemplificações +sinônimos +expressões, ditados, citações com essa palavra +verbos preposicionados +palavras relacionadas/referências cruzadas (remissiva) + verbos frasais (Phrasal Verbs) ±pragmática ±antônimos ±notas culturais ±outras categorias morfológicas da mesma palavra ±figuras

9. Qual o número de acepções, dentro da mesma definição, que você considera ideal?

	Respostas	%
1	2	1
2	2	1
3	5	3
4	2	1
quantas forem necessárias	163	92

Motivação: tentativa de descobrir se os tradutores, pelo motivo da necessidade de velocidade na tradução, se limitavam a consultar um número limitado de acepções de um verbete.

Resultado: os profissionais buscam tantas acepções quantas necessárias para buscar a tradução/definição de que necessitam.

Relevância para a tese: limitar o número de acepções, caso surjam (já que uma das pressuposições básicas da Terminologia é que um termo deve ter apenas uma acepção

em determinada área), não parece ser uma necessidade, já que os consulentes estão acostumados a ler tudo até achar o que procuram; por outro lado, para aumentar a eficiência da busca, pode-se organizar as acepções em ordem decrescente de frequência no *corpus*, o que provavelmente aumentaria a velocidade de trabalho (pelo menos para as acepções mais frequentes).

10. Na sua opinião, qual a forma mais prática de consultar palavras homônimas (como banco)?

	Respostas	%
agrupadas dentro do mesmo verbete	90	51
colocadas em verbetes diferentes	85	48

Motivação: já havia sido confirmado, em pesquisa de mestrado (FROMM, 2002), que obras lexicográficas e terminográficas nem sempre apresentam uma padronização quanto às entradas e acepções; ora os homônimos são colocados como acepções de um mesmo verbete, ora são colocados como verbetes diferentes (isso dentro da mesma obra).

Resultado: parece haver um equilíbrio entre as respostas (com uma pequena vantagem para o agrupamento dentro de um mesmo verbete), sugerindo indiferença por parte dos consulentes entre uma opção ou outra.

Relevância para a tese: o que parece ser uma indiferença também pode ser interpretado como uma realidade que os tradutores encontram nas obras, que é o problema da falta

de padronização. Para que isso não continue a existir, deve-se selecionar uma das duas opções e padronizá-la dentro da obra.

11. Na sua opinião, qual a forma mais prática de consultar palavras polissêmicas (com mais de um significado, como cavalo)?

	Respostas	%
agrupadas dentro do mesmo verbete	113	64
colocadas em verbetes diferentes	59	33

Motivação: assim como na pergunta anterior, tentou-se, aqui, descobrir a melhor localização para palavras, nesse caso polissêmicas, se dentro da macro- ou da microestrutura.

Resultado: a preferência pela microestrutura não surpreende, já que esse é o modelo vigente nas obras impressas.

Relevância para a tese: confirma a necessidade de não alterar o que já é, mesmo que informalmente, um padrão de consulta. Como há uma baixa probabilidade de haver uma palavra com mais de um significado dentro de uma mesma subárea, esse dado serve apenas para preparar o banco de dados para tal eventualidade.

12. Em qual(is) área(s) você costuma traduzir mais?

	Respostas	%
Técnica em geral	99	56
Outra (s)	86	49
Administração de empresas	68	38
Jurídica em geral	64	36
Tecnologia da Informação (informática)	63	36
Médica em geral	52	29
Técnica industrial	47	27
Jornalística em geral	45	25
Tecnologia da Informação (telefonía)	42	24
Turismo	39	22
Literária	34	19
Automobilística	30	17
Educação	30	17
Seguros	23	13
Linguística	20	11
Esportes	6	3
Humor	2	1

Motivação: ter um parâmetro dos campos em que o tradutor mais atua, inclusive para o levantamento dos *corpora*. A taxonomia dos campos, obviamente, pode ser contestada. O objetivo, no entanto, foi propiciar o maior número possível de campos com os quais os tradutores pudessem se identificar.

Resultado: as áreas técnicas se sobressaem, tais como administração de empresas, jurídica e tecnologia da informação. Interessante também notar a opção “outras”.

Relevância para a tese: já que a área técnica, em geral, se destaca como principal fonte de textos para o tradutor, foi escolhido um campo de destaque: a computação (ou

informática). Essa questão também pode servir para análises futuras da ordem de necessidade dos tradutores quanto aos vocabulários a serem desenvolvidos.

Todas essas respostas contribuíram para a construção da *consulta tradutor* (capítulo 7.1.3), um dos modos de montagem da microestrutura na página do VoTec. Essa microestrutura específica pretende recuperar as respostas aqui fornecidas e mostrar um padrão de busca que reflita o pensamento da maioria dos tradutores. Os aprendizes podem, então, contar com a experiência dos veteranos para melhorar sua busca.

6. Abordagem de modelos, a Ficha Terminológica e o banco de dados

Sempre há, quando do planejamento de um banco de dados e/ou de uma página de consulta para terminologia, mais de uma abordagem possível. Analiso, neste capítulo, dois tipos de abordagens e aquela adotada para o presente trabalho.

6.1 A abordagem lexicográfica para a construção do banco de dados

Quando da construção do banco de dados, pensou-se numa abordagem lexicográfica. Entenda-se, aqui, lexicográfica como a possibilidade de abarcar, dentro da ficha terminológica (detalhada adiante), todos os campos possíveis para a criação de um verbete de dicionário.

A organização de uma obra lexicográfica ou terminográfica deve ter em vista, em primeiro lugar, o protótipo do consulente. Esse trabalho não se propôs, em princípio, a abordar todas as variações diatópicas, diafásicas, diastráticas e diacrônicas, embora se reconheça a importância de algumas delas para nichos específicos de tradução, tais como intérpretes, legendadores e outros. Para a elaboração desta tese, seguiu-se a linha adotada por Barbosa (2001), que destaca as questões sincrônicas e sínfásicas como mais relevantes para a terminologia/terminografia.

A visão do que é importante para um pesquisador, no entanto, pode não coincidir com a opinião de outros pesquisadores ou com a dos próprios consulentes. A ficha terminológica foi, então, concebida com o maior número possível de campos a serem preenchidos, desde que não representassem variações (diatópicas, diastráticas, diacrônicas) do termo. O que está ali presente é um recorte das possibilidades de campos levantadas por Atkins, Fillmore e Johnson (2003, anexo A), dentre outros pesquisadores (apresentados em FROMM, 2002), para a construção de uma entrada em

um dicionário bilíngüe. A constituição do banco de dados permite que, no futuro, outros campos venham a ser adicionados ao mesmo. A disponibilização da consulta pela Internet, meio dinâmico por natureza, sugere que esses futuros campos devam estar associados, também, à questão da multimídia. Assim, o termo poderia apresentar, entre os campos de consulta, por exemplo, informações enciclopédicas interativas (vídeo, som, animações, etc.).

6.2 Uma abordagem híbrida para a construção da página de consulta

Ao mesmo tempo em que a ficha terminológica teve como base uma visão lexicográfica, com o intuito de criar diferentes produtos no futuro, a página de consulta teve como base uma visão híbrida entre lexicografia e terminografia.

Muitas obras disponíveis no mercado, rotuladas como dicionários, como, por exemplo, a de Rost (2003; figura 19), não passam de listagens de nomenclaturas de áreas técnicas em duas ou mais línguas. Snell-Hornby²⁰ (1990 *apud* SCHMITZ, 2000), já havia levantado o problema: uma listagem de quase-sinônimos, sem orientação para o consulente e sem indicação que pode haver variações dos diferentes significados, pode induzi-lo a erros na língua alvo.

²⁰ SNELL-HORNBY, M. Dynamics in meaning as a problem for bilingual lexicography. In: TOMASZCZYK, J.; LEWANDOWSKA-TOMASZCZYK, B., **Meaning and Lexicography**. Amsterdam: John Benjamins Publishing Company, 1990.

Dictionary of Insurance Terms		
ENGLISH	PORTUGUESE	ESPAÑOL
average clause	cláusula de rateio	regla proporcional
average condition	cláusula de rateio	regla proporcional
average daily value	valor médio diário	valor promedio diario
average guarantee	carta fiança em avaria grossa	fianza de avería gruesa
average premium	prêmio médio	prima media
average rate	taxa média	tasa media
aviation insurance	seguro aeronáutico	seguro de aviación
to avoid	anular	anular
avoidance	contestação	contestación
award	adjudicação	adjudicación
B		
backdating	datação retroativa	retroacción
bad faith	má fé	maña fe
bad workmanship	defeito de fabricação	defecto de fabricación
bailed property	bens sob custódia	bienes bajo custodia
bail bond	fiança	fianza
bailee	depositário	depositario
bale	fardo	fardo
bank burglary and robbery	seguro de roubo de bancos	seguro de atraco de bancos
bankers' blanket bond	seguro global de bancos	seguro global de bancos
bankruptcy	falência	insolvencia
bareboat charter	afretamento a casco nu	fletamento por translación de dominio
barge	barcaça	gabarra
barred	prescrito	prescrito
basic premium	prêmio básico	prima basica
barratry	barataria	baratería
batch clause	sinistros em série	sinistros en serie
battery	lesões corporais	lesiones corporales
bench error	erro de fabricação	error en la fabricación
benchmark	parâmetro, referencial	cota de referencia
beneficiary	beneficiário	beneficiario, derechohabiente
benefit	beneficio	beneficio
berth	ancoradouro	fondeadero

Figura 19. Dicionário Técnico de Seguros (ROST; 2003)

Pode-se criticar a postura de Rost na feitura de sua obra ou concordar com Schmitz/Snell-Hornby, mas uma questão se coloca: será que o tradutor não precisa somente disso? Embora já se saiba, com a pesquisa relatada no capítulo anterior, que o profissional da tradução prefere consultar não só o Paradigma de Forma Equivalente

(PFE), mas também os Paradigmas Definicional (PD), Informacional (PI) e Pragmático (PP), algumas perguntas práticas se fizeram presentes: qual o primeiro paradigma a ser consultado? Na necessidade de consulta de mais de um paradigma, qual a ordem de consulta aos mesmos?

Quando da elaboração da página de consulta, que é um dos objetivos básicos desta tese, um questionamento primordial sobre possíveis abordagens dessa consulta foi levantado: o aprendiz deve ter um primeiro contato com a ferramenta somente adotando-se um modelo de consulta ou deve ser oferecido mais de um? Decidiu-se, então, vista a multiplicidade de modelos (melhor detalhado no capítulo 3), oferecer mais de uma abordagem. Dessa decisão nasceram três possibilidades: uma de caráter mais lexicográfico (consulta total), outra de caráter mais terminológico (consulta tradutor) e outra híbrida (consulta modular), de acordo com as necessidades do tradutor naquele momento. Antes de trabalhar com essas possibilidades, alguns modelos são analisados.

6.3 Modelos de fichas terminológicas e bancos de dados

O projeto COMET (*Corpora Multilíngüe para Ensino e Tradução*), da FFLCH/USP, coleta, já há alguns anos, vários *corpora* em diferentes áreas. Os alunos de mestrado e doutorado ligados ao projeto, além de contribuírem para a construção desses *corpora*, também vêm desenvolvendo trabalhos de pós-graduação para expor diferentes maneiras de trabalhar com os mesmos. Uma das vertentes de estudo é a construção de vocabulários baseados em *corpora* de áreas de especialidade. Foram tomados como modelos dicionários, monolíngües ou bilíngües, baseados em grandes *corpora* gerais de língua (como os das editoras inglesas Longman e Oxford, que trabalharam com *corpora* próprios ou com o *British National Corpus*). O direcionamento de alguns doutorandos é a construção de ferramentas e modelos que

funcionem como alicerce para a futura organização de obras terminológicas baseadas exclusivamente em *corpora*.

O objetivo da construção desse banco, além de prover o VoTec com as informações necessárias para a modelagem de suas estruturas, é também prover o Projeto COMET de uma ferramenta informatizada semi-automatizada que auxilie no desenvolvimento de obras terminológicas criadas a partir da grande base de *corpora* bilíngües já levantados e disponibilizados pelo mesmo. Isso implica que os vocabulários a serem desenvolvidos terão suas definições construídas unicamente através com exemplos tirados desses *corpora* (outros exemplos de fichas terminológicas podem, também ou exclusivamente, buscar suas definições em obras lexicográficas ou terminográficas já desenvolvidas e disponíveis no mercado).

O objetivo secundário é fornecer uma base para o desenvolvimento de futuras ferramentas, a serem associadas ao banco, ligadas à extração de termos a partir de *corpora* de áreas de especialidade e à construção de novas ferramentas de visualização do produto final (vocabulários técnicos bilíngües – inglês/português, em uma primeira fase) para diferentes usuários.

Embora existam vários programas de gerenciamento de terminologia disponíveis no mercado internacional (como o Multiterm, Term-PC e outros, muito bem analisados por GAVENSKI, 2001) e vários bancos de dados terminológicos já desenvolvidos no país, como os pequenos bancos usados pelo CETRAD/USP no ensino de Terminologia para a área de tradução (desenvolvidos pelo Prof. Dr. Francis H. Aubert, baseados em um modelo construído no banco de dados Access, da Microsoft) ou os grandes bancos, como o TERMISUL (MACIEL, 2001), pensou-se na construção de um banco personalizado para as necessidades do COMET. A vantagem, além do baixo custo de

desenvolvimento e implementação²¹, é a possibilidade de agregação de futuros novos módulos (como a automatização de certas partes do processo) e/ou funcionalidades (como ferramentas de análise lexical), associados às pesquisas de mestrado e doutorado desenvolvidas no âmbito do projeto COMET ou por seus parceiros.

Qualquer trabalho terminológico pressupõe várias etapas para a construção do produto final, que seria um vocabulário de uma determinada área ou um glossário (usando as concepções de dicionário, vocabulário e glossário, propostas por BARBOSA, 2001). Entre essas etapas, uma das mais importantes é a organização dos dados recolhidos através de uma ficha, comumente chamada de ficha terminológica. Cabré (1993) explica o que vem a ser essa ficha:

[a]s fichas terminológicas são materiais estruturados que devem conter toda a informação relevante sobre cada termo. As informações apresentadas são extraídas das fichas preenchimento ou da documentação de referência e são representadas segundo critérios fixados previamente.

Há muitos modelos de fichas terminológicas, de acordo com o objetivo de cada trabalho e as necessidades de cada organização. Para iniciar, podemos distinguir entre fichas monolíngües, fichas monolíngües com equivalência e fichas bilíngües ou plurilíngües.²²

(tradução minha).

A criação de uma ficha terminológica é essencial para o desenvolvimento de um vocabulário técnico. Vários modelos já foram propostos e, entre eles, pode-se citar boas descrições como aquelas fornecidas por Aubert (1996), Krieger e Finatto (2004), Gavenski (2001) e Bacellar (2002) em seus trabalhos. A ficha terminológica foi, durante muito tempo, elaborada e preenchida manualmente. O advento dos computadores permitiu não só o desenvolvimento da Linguística de *Corpus*, como também a

²¹ A serviço da Empresa Jr., do ICMC/USP São Carlos.

²² “Las fichas terminológicas son materiales estructurados que deben contener toda la información relevante sobre cada término. Las informaciones que presentan se extraen de las fichas de vaciado o de la documentación de referencia, y se representan siguiendo unos criterios fijados previamente. Hay muchos modelos de fichas terminológicas, de acuerdo con los objetivos de cada trabajo y las necesidades de cada organismo. De entrada, podemos distinguir entre fichas monolingües, fichas monolingües con equivalência y fichas bilingües o plurilingües.”

informatização dessas fichas e a construção de bancos de dados. É apresentada abaixo, como exemplificação, a tela de uma versão em formato eletrônico dessas fichas (figura 20). Esse banco de dados, bilíngüe inglês/português, é voltado para a área de comércio eletrônico (Dicionário de termos do comércio eletrônico) e foi desenvolvido pelo ILTEC²³ (Instituto de Lingüística Teórica e Computacional, vinculado à Universidade Nova de Lisboa), apresentando as informações (como explica Cabré, acima citada, cada ficha é pensada de acordo com as necessidades e objetivos de quem a organiza) que devem ser cadastradas pelos pesquisadores.

Termo em Português	<input type="text"/> *
Categoria Morfossintática	Seleccionar Opção <input type="button" value="v"/> *
Índice de Fiabilidade	Seleccionar Opção <input type="button" value="v"/> *
Sinónimos	<input type="text"/>
Termo em Inglês	<input type="text"/> *
Subdomínio	<input type="text"/> *
Definição	<input type="text"/> *
Definição (Fonte)	<input type="text"/> *
Contexto	<input type="text"/>
Contexto (Fonte)	<input type="text"/>
Observações	<input type="text"/>
Utilizador	<input type="text"/>
Data Revisão	<input type="text"/>
Data Revisão Técnica	<input type="text"/>
Data Revisão Linguística	<input type="text"/>
Data Última Revisão	<input type="text"/>
<input type="button" value="Inserir"/>	

Figura 20. Ficha Terminológica Informatizada do ILTEC.

Almeida, Souza e Pino (2007) explicitam o modelo de ficha construído para o projeto GETerm (Grupo de Estudos e Pesquisas em Terminologia, associado ao Departamento de Letras da UFSCAR; figura 21), construído num banco de dados Access, da Microsoft, e bem mais complexo que a versão do ILTEC:

²³ <http://www.iltec.pt/projectos/em_curso/dtce.html>

Figura 21. Exemplo de ficha terminológica elaborada pelo GETerm para seu Dicionário de Revestimento Cerâmico (DiRC). (ALMEIDA; SOUZA; PINO, 2007, p. 12).

O modelo tomado como ponto de partida para o banco do COMET, no entanto, é baseado em Braga (2000) e uma variante do mesmo, proposta por Fromm (2002; um exemplo de ficha terminológica ali exibida está disponibilizado nesta tese no apêndice D) que visava elaborar um modelo para a construção da microestrutura de um vocabulário especializado de informática para tradutores. Tendo como ponto de partida esse modelo de ficha terminológica monolíngüe não-informatizada, foi desenvolvida uma nova proposta para uma ficha monolíngüe com equivalência, que servirá de base para a construção do banco de dados para esta tese.

Essa ficha com equivalência pressupõe que todas as telas da ficha terminológica a serem preenchidas são monolíngües (inglês ou português), porém com um campo para a entrada do termo equivalente na outra língua. Cada termo, portanto, requer o preenchimento de duas fichas, uma para o termo em português e outra para o termo em inglês. A página de consulta é o local onde essas fichas monolíngües podem ser contrastadas, obtendo-se o efeito bilíngüe desejado obtido.

Em discussões com o representante da Empresa Jr., decidiu-se pela construção de um banco de dados padrão SQL, com várias telas para a inserção de dados. Devido à complexidade de trabalho num banco de dados desse padrão, foi criado um ambiente WEB para que futuros pesquisadores (profissionais ligados à tradução, terminologia ou áreas técnicas) possam preencher as fichas. Em virtude dos custos de elaboração do projeto, somente um administrador (a figura 22 mostra a tela dese administrador) terá acesso ao controle do banco numa primeira fase.

http://www.rodrigopedra.mat.br - Dicionário Eletrônico - Mozilla Firefox

English

Bem Vindo, Administrador

Seu último acesso foi em: 14/05/2007 12:26

[Trocar Senha](#)
[Cadastro de Pesquisadores](#)
[Novo Termo](#)
[Cadastro de Áreas](#)
[Cadastro de Fontes](#)
[Sair](#)

Fichas ainda não aprovadas

Termo	Língua	Coletado em	Cadastrada por	Ações
memória		07/05/2007	Administrador	editar - revisar - reprovar - aprovar

Total de Termos: 1

Fichas aguardando revisão

Termo	Língua	Coletado em	Cadastrada por	Ações
Nenhum termo para revisar				

Total de Termos: 0

Buscar fichas aprovadas

Termo:

14/05/2007 16:51 © 2007 ICMC Jr.

Concluído

Figura 22. Exemplo da tela do administrador do banco de dados²⁴. O termo *memória* aguarda aprovação para ser disponibilizado no VoTec.

Ao administrador caberá o cadastro de pesquisadores (para que esses possam alimentar as fichas) e as fichas por eles preenchidas só serão disponibilizadas para

²⁴ Essa mesma tela, assim como todas as outras, pode ser visualizada em inglês também.

consulta no sistema (página do VoTec) após aprovação pelo mesmo. Ao administrador caberá, também, a inserção de novas fichas terminológicas, atualização e remoção de fichas existentes no sistema (uma reprodução do documento inicial de requisitos para a construção do sistema, elaborada pelo representante da empresa no decorrer de várias reuniões presenciais e contatos eletrônicos, está disponível no apêndice E).

Há várias opções para cadastramento na tela inicial do administrador (troca de senha, cadastro de novos pesquisadores, cadastro de novas áreas e fontes). Na figura 23, disponibiliza-se um exemplo da tela para o cadastro de fontes:

The screenshot shows a web interface for managing sources. At the top right, it says 'Tela Cheia | English'. The main heading is 'Fontes'. Below it is a button 'Voltar ao painel'. The section 'Nova Fonte' has three input fields: 'Nome*', 'Abreviação*', and 'Link', with a 'Salvar' button below them. The 'Fontes Cadastradas' section contains a table with the following data:

Nome	Abreviação	Ações
Angelfire	AF	editar
APC Magazine	APC	editar
Clube do Hardware	CDH	editar

Figura 23. Tela para cadastro de fontes. Na parte inferior, verificam-se as fontes já cadastradas.

Essencial para qualquer levantamento terminográfico, também, é a construção da árvore de campo da área a ser estudada. A inserção dessa árvore no banco, já elaborada quando da construção do(s) *corpus (corpora)*, é imprescindível para seu correto funcionamento. A interrelação entre os termos em diferentes línguas (Paradigma de Forma Equivalente), por exemplo, só se dará se os mesmos pertencerem às mesmas

subáreas de uma mesma área técnica. O banco possibilita que o pesquisador insira dados e os visualize; a apresentação, no entanto, não é mais em formato de árvore e sim de um sistema conceptual (ainda que a palavra que aparece como título, *árvore*, remeta a esse tipo de construção). Embora sistemas desse tipo costumem numerar seus itens, aqui eles são disponibilizados sem numeração. Esse sistema pode ser observado na figura 24 abaixo:

The screenshot shows a web application interface with the following elements:

- Page Header:** "Tela Cheia | English"
- Breadcrumbs:** "Computação > Hardware > Componentes Internos"
- Navigation:** "Voltar ao painel" and "Voltar um nível" buttons.
- Árvore (Tree View):**
 - Grandes Áreas
 - Computação
 - Hardware
 - Componentes Internos
 - Barramento
 - Componentes Diversos
 - Memória Primária
 - Placa-mãe
 - Processador
 - Computadores
 - Periféricos
 - Rede
 - Software
 - Direito Contratual
 - Lingüística

- Nova Sub-Área (Form):**

Informações sobre a área

Nome em Português*:

Nome em Inglês*:
- Sub-Áreas Cadastradas (Table):**

Nome em Português	Nome em Inglês	Ações
Nenhuma sub-área cadastrada		

Sub-Áreas Cadastradas: 0
- Page Footer:** "14/05/2007 17:18 © 2007 ICMC Jr." and "Concluído"

Figura 24. Sistema conceptual e cadastro de áreas e subáreas.

6.4 A inserção de dados pelo pesquisador

Após a inserção de seu nome e senha, o pesquisador depara com uma primeira tela para iniciar a sessão e o cadastro de novos termos. Essa tela (figura 25) é bastante semelhante à tela de cadastro do administrador (figura 22), porém sem as possibilidades de acrescentar novos pesquisadores. O endereço provisório para acesso ao banco é: <<http://jr.icmc.sc.usp.br/~comet/dic/admin>>²⁵. Para acessar o banco, como pesquisador, digite o usuário *teste* e a senha *teste*.



Figura 25. Tela inicial do pesquisador. Neste caso, cadastrei-me como pesquisador e não administrador. A faixa superior indica o nome do pesquisador cadastrado que inicia a sessão.

A tela de cadastro de fontes segue a mesma orientação apresentada anteriormente. A tela “novo termo” servirá para a inserção de contextos (previamente selecionados, conforme capítulo 4) retirados de um *corpus* de especialidade de uma área escolhida (os exemplos aqui apresentados são da área de computação). Na primeira tela o pesquisador insere o termo em si, escolhe a língua em que ele vai ser cadastrado e a

²⁵ Preve-se a mudança da hospedagem, do servidor da Empresa Jr. para o servidor do projeto COMET.

parte da Ontologia em que se insere (conforme explicado no item 4.3; figura 26). Na divisão hierárquica dessa Ontologia, ao invés de campo, área, domínio, subdomínio e outros, optou-se pelos termos *grande área*, *sub-área 1*, *sub-área_n* (há cinco níveis disponíveis), mais acessíveis para não-lingüistas.

Figura 26. Primeira tela, na área do pesquisador, para inserção de termo.

A partir dessa tela, podemos dividir a inserção de dados naquilo que denomino Paradigmas Terminográficos. Eles representam, na microestrutura de um termo, os campos que a compõem. Haensch (1982, p. 470-501) já fazia uma divisão dos componentes básicos da microestrutura de um verbete lexicográfico e os denominava Paradigmas: Informacional (informações ortográficas, fonéticas, etimológicas, cronológicas, gramaticais, etc.), Definicional (definição) e Pragmático (exemplos e abonações). Segundo Barbosa (1994), a organização dos paradigmas que compõem a microestrutura está relacionada com a obra lexicográfica (ou terminográfica, no caso)

em que aparecem. Não há uma fórmula pronta para a composição de uma obra lexicográfica ou terminográfica; o que deve ser observada, desde a fase de elaboração, é a coerência interna da mesma. Ainda, para a autora, (1989, p. 571), novos paradigmas podem ser acrescentados para o enriquecimento da obra (não há, portanto, um número “correto” de paradigmas), desde que o principal deles, o Definicional, não seja deixado de lado. São apresentados, a seguir, os Paradigmas que compõem a microestrutura proposta por esta tese para o VoTec.

6.5 O Paradigma Pragmático (PP)

A segunda tela solicita, para cada termo, tantos contextos quanto o pesquisador achar necessários para a alimentação dos demais campos e construção da definição; esses contextos são extraídos do *corpus*. Cada contexto exige o preenchimento de quatro campos: exemplo (o contexto em si), fonte, data de coleta e conceito (figura 27a).

http://www.rodrigopedra.mat.br - Dicionário Eletrônico - Mozilla Firefox

English

Cadastro de Contextos para memória

Voltar ao Painel Próximo Passo

Novo Contexto

Passo 2

Dados do Contexto

Exemplo*:

Conceito*:

Fonte*: APC Magazine

Data de Coleta*: (dia/mês/ano ex.: 18/03/2007)

Figura 27-a. Tela de cadastro de contextos, parte superior.

A partir de cada contexto, o pesquisador deve, também, extrair um primeiro conceito formado pelos traços distintivos explicitados pelo exemplo (e que serão retomados na construção do Paradigma Definicional). Veja o exemplo, na figura 27-B, com o termo “memória” (o termo é bastante polissemêmico, não só nessa área como em outras; a especificação do mesmo na área *hardware*, nesse caso, demandou uma grande quantidade de contextos retirados do *corpus*):

Contextos Cadastrados			
Exemplo	Conceito	Fonte	Ações
Quando o micro é ligado o processador carrega o sistema operacional, executa programas e manipula os dados conforme solicitado pelo usuário. Sabendo que, o processador não tem uma área interna de armazenamento muito grande de dados, essas informações são armazenadas em uma área chamada memória.	área para armazenamento de dados	CDH 08/06/2007	editar - excluir
Um chip de memória é um exército de clones, formado por um brutal número de células idênticas, organizadas na forma de linhas e colunas, de uma forma similar a uma planilha eletrônica. O chip de memória em si serve apenas para armazenar dados, não realiza nenhum tipo de processamento.	chip formado por células idênticas, na forma de linhas e colunas; apenas armazena dados, não realiza nenhum tipo de processamento	GHD 08/06/2007	editar - excluir
A maneira mais rápida de analisar o desempenho da memória é compararmos a sua taxa máxima de transferência, que é calculada com a fórmula $\text{clock} \times \text{bits por clock} / 8$. Os chips voltados ao mercado de alto desempenho acessam a memória a 256 bits, contra 128 bits ou mesmo 64 bits dos chips voltados aos mercados intermediário e iniciante.	desempenho medido pela taxa de transferência (256, 128 ou 64 bits).	CDH 01/02/2005	editar - excluir
O ambiente de edição é o que mais exige recursos do computador. Isso pode ser notado com o uso de um micro com chip Celeron 2,8 GHz e 256 Mbytes de memória RAM.	medida em unidades de Mbytes	FSP 01/02/2005	editar - excluir
Quando a memória RAM está ocupada, o Windows cria a chamada memória virtual, ou seja, usa o disco rígido, que é muito mais	pode ser virtual, criada pelo windows	FSP 01/02/2005	editar - excluir

Figura 27 - B. Tela de cadastro de conceitos, parte inferior. Aqui são apresentados os contextos e respectivos conceitos já cadastrados para o termo *memória*.

Esses exemplos são os mesmos que serão disponibilizados ao consulente na página do VoTec. No modo padrão, somente um exemplo é mostrado; mas o usuário pode solicitar a visualização de todos.

A terceira tela, após a inserção dos contextos, disponibiliza uma segunda página de inserção de dados sobre o termo (as figuras seguintes apresentam as várias abas que constituem essa página) e servirá para a construção de todos os demais paradigmas para a construção da microestrutura do termo.

6.6 A construção do Paradigma Definicional (PD)

A elaboração da definição (“tipo específico de texto bidimensionado: com suas dimensões semântica e pragmática”; [FINATTO, 1994]), como em qualquer obra lexicográfica ou terminográfica, representou um desafio para a construção do banco. Em primeiro lugar, há de se apresentar uma distinção (LARA, 2000, p. 94) entre a definição lexicográfica e terminográfica tradicional:

A diferença entre a definição terminológica e a definição lexicográfica remete, portanto, a perspectivas distintas de abordagem, muito embora elas não sejam isentas: a lexicografia parte do signo para chegar à determinação do conceito (procedimento semasiológico); a terminografia parte da noção (ou conceito) e pesquisa os termos que lhe correspondem (procedimento onomasiológico).

Vários outros autores trabalham com a questão da terminografia moderna e suas implicações. Finatto (1998) faz um bom resumo do que é a criação da definição terminológica:

No âmbito do dicionário terminológico, tal como verificamos, via de regra, na definição contida na microestrutura, definir pode ser o mesmo que discriminar características, explicar, situar ou arrolar diferentes modalidades ou variedades do *definiendum*, corporificado pelo termo, o que se dá pela expressão lingüística de um conceito de estatuto técnico ou científico. Entretanto, a formulação da definição estará igualmente mediada e condicionada por vários elementos tais como a natureza da obra, pelo tipo de discurso em que originalmente se inscrevam o conceito e o termo, pelo tipo de público-alvo e perfil do produto terminográfico em que se insere, entre outros fatores.

A premissa básica para a construção do banco de dados para esta tese foi que todos os campos da Ficha Terminológica deveriam ser preenchidos a partir de, e unicamente a partir de, contextos retirados de *corpora* reais de língua. Para tanto, houve uma grande preocupação com o balanceamento da quantidade de palavras dentro de

cada subárea para que o processo de construção das definições fosse o mais acurado possível, já que “...a microestrutura de dicionário, em sua porção definicional, é um tipo de DT²⁶ em que encontramos, via de regra, a reescritura de um outro texto, o texto de referência a partir do qual é, geralmente, moldada, extraída ou adaptada...” (FINATTO, 1998). Os contextos definitórios e explicativos, anteriormente citados, foram buscados já durante a elaboração desses *corpora* (o que não significa, necessariamente, que todos os textos os possuam). Os textos reunidos representam a seleção de dois tipos de linguagem: de especialistas para especialistas (textos acadêmicos e manuais técnicos de alguns produtos) e de especialistas para leigos (textos jornalísticos explicativos, com características de enciclopédia, e outros manuais técnicos).

A definição, partindo de uma abordagem lingüística, cognitiva e comunicativa²⁷ pode, então, ser composta pela reescrita desses dois tipos de textos para um público-alvo que terá de lidar com as mesmas questões nas traduções do seu dia-a-dia: textos técnicos de especialista para especialista, textos jornalísticos de especialista para leigos, etc.

Alves (1996, p. 125), ao discutir as propostas de Sager (1990²⁸), um entre os vários autores que tratam do assunto (ALMEIDA; SOUZA; PINO, 2007, fazem um ótimo apanhado daqueles que trabalham no tema), nos apresenta os diversos tipos de construção da definição em si:

- a. Analítica (ou por compreensão): termos genéricos e características específicas;
- b. sinonímicas;
- c. paráfrase;

²⁶ DT, no texto da autora, remete à Definição Terminológica.

²⁷ Segundo Finatto (1998), há dois tipos de abordagem para os fenômenos terminológicos: uma linha afinada com as idéias de Sager, com enfoque lingüístico, e uma linha tradicional, com enfoque em estudos ontológicos e na Teoria do Conceito.

²⁸ SAGER, J.C. **A practical course in terminological processing**. Amsterdam: John Benjamin Publishing, 1990.

- d. síntese: descrevem e identificam relações;
- e. implicação: uso do termo em um contexto explicativo;
- f. denotação (extensão): relacionam exemplos;
- g. demonstração: definem por meio de fotografias, ilustrações, etc.

Já a norma ISO 1087 (1990, p. 4) propõe dois tipos básicos de definição (“Enunciado que descreve um conceito e permite sua diferenciação de outros conceitos dentro de um sistema de conceitos”²⁹; o conceito é “uma unidade de pensamento constituída através de abstração sobre a base de propriedades comuns a um conjunto de objetos”³⁰, p.1; minhas traduções):

4.1 definição por compreensão: definição baseada na compreensão³¹ de um conceito (Nota: para este propósito é necessário especificar o conceito genérico mais próximo que já tenha sido definido ou que possa ser assumido como de conhecimento geral e adicionar as características de restrição que delimitam o conceito a ser definido).

4.2. definição por extensão: definição baseada na enumeração exhaustiva de objetos que se referem ao conceito, ou de conceitos específicos no próximo nível de abstração.³²

Baseada nessa norma, Alves (1996) sustenta que

[a] definição por compreensão (intensional) é, certamente, a mais adequada aos trabalhos terminológicos, uma vez que ela situa o conceito no âmbito de uma classe e especifica o que o distingue de outros conceitos situados na mesma classe (p. 126).

²⁹ *Statement which describes a concept and permits its differentiation from other concepts within a system of concepts.*

³⁰ *A unit of thought constituted through abstraction on the basis of properties common to a set of objects.*

³¹ Conjunto de características que constitui um conceito. *Set of characteristics which constitutes a concept*, p. 2.

³² 4.1 intensional definition: Definition based on the intension of a concept (Note: for this purpose it is necessary to state the closest generic concept that has already been defined or can be assumed to be generally known, and to add the restricting characteristics that delimit the concept to be defined); 4.2 extensional definition: Definition based on the exhaustive numeration of the objects referred to by the concept, or of the specific concepts at the next level of abstraction.

Os termos genéricos que nela aparecem não pertencem a um domínio em particular e introduzem o conceito do termo que está sendo descrito. Como exemplo desses termos, temos: elemento, estrutura, fase, fórmula, método, parte, procedimento, processo, sistema, símbolo, subárea, técnica, etc.

Gianni³³ (1993, *apud* FINATTO, 1994) acrescenta à norma 1087 a definição contextual, que reproduz o contexto em que o termo ocorre, e a definição enciclopédica, orientada à descrição exaustiva da coisa nomeada. A construção da definição proposta para os verbetes desta tese não inclui esses elementos, mas eles reaparecem na composição da microestrutura (através dos Paradigmas Pragmático e Enciclopédico), alargando os horizontes semânticos do termo em questão.

Na figura 28 abaixo, para o termo *processor*, por exemplo, podemos verificar como se dá a construção da definição no banco. Essa página representa a segunda página de inserção de dados. Nela são visualizados, num primeiro momento, na parte superior, os exemplos extraídos pelos pesquisadores na primeira tabela e seus respectivos conceitos iniciais; a partir dos mesmos, os traços distintivos serão agrupados em linhas (parte inferior). O terminógrafo deverá agrupar traços que abarquem as mesmas idéias, nas diferentes linhas, nas mesmas colunas (como *performs* e *execute*, no exemplo abaixo).

³³ GIANNI, E. O Paradigma definicional lexicográfico e terminográfico. In: **Cadernos do IL**, 10. Porto Alegre: UFRGS, 1993. p. 45-56.

http://jr.icmc.sc.usp.br - Vocabulário Técnico Online - Mozilla Firefox

English

Termo: processor

Passo Anterior Salvar Sair sem salvar

Contextos Conceito Final / Definição

Contextos

Exemplo	Conceito	Fonte
1 The "brain" of any computer is the central processing unit (CPU). This component does all the calculations and performs 90 percent of all the functions of a computer. There are many different types of processors for computers—so many, in fact, that you will learn about them later in this chapter in the section "Identifying Purposes and Characteristic of Processors."	does all the calculations and performs 90% of all the functions of a computer	Wiley 08/06/2007

Dados Traços Distintivos Semântica Termo Equivalente Termos Remissivos Informações Enciclopédicas

Traços Distintivos

Nova Coluna — Não use aspas duplas

	A	B	C	D	E	F	G	H	I
1	calculates	performs	functions						
2				flat squar...	circuits e...				
3						speed mea			
4		execute					CPU	control	programs

Figura 28. Contextos, conceitos e seleção de traços distintivos.

Desses traços distintivos, oriundos dos contextos do *corpus*, serão tirados o conceito final e a definição do termo. O conceito final é montado tendo em vista os vários traços agregados nas diferentes linhas e colunas (note-se que esses traços, ao comporem a definição, podem não estar na mesma ordem em que aparecem nas linhas e colunas). O trabalho do terminógrafo, aqui, pode representar variações na construção do conceito final (diferentes pesquisadores percebem diferentes traços ou os agrupam de formas variadas).

O ponto de partida adotado para a construção do conceito final é o uso do paradigma GPDE (gênero próximo, diferença específica) adotado pelo GETerm (ALMEIDA; SOUZA; PINO, 2007), ou seja, a definição começa pelo termo superordenado (hiperônimo) em relação àquele que está sendo descrito e, em seguida,

são apresentados os semas específicos do termo em questão. O termo superordenado (normalmente fornecido pelos exemplos do *corpus*) é imediatamente superior ao termo em questão e não representa, necessariamente, uma subárea da árvore de campo. Quando os excertos do *corpus* não forneceram o termo superordenado, foram usados dois termos genéricos dentro da área de hardware: *máquina* ou *equipamento* ou ainda descritores mais genéricos, como *tipo de* (produzindo enunciados como *verbetes x: é um tipo de máquina que...*). Como a maioria dos termos levantados a partir do *corpus* são substantivos, estipulou-se que o primeiro descritor também deveria ser um substantivo. Continuando com o exemplo da figura 28, em português, tem-se para o termo *processador* a seguinte definição:

tipo de chip, parte principal do computador, também conhecido como CPU (Central Processing Unit); executa instruções (processos) e dados (através da UC - unidade de controle) e operações lógicas e aritméticas (através da UAL - unidade aritmética e lógica).

Outra decisão tomada pelo GETerm, e aqui também adotada, foi a quantidade de semas que compõem a diferença específica. Como a elaboração do conceito final representa uma pluralidade de concepções sobre o termo, retiradas dos vários excertos do *corpus*, decidiu-se pela agregação do maior número possível de semas na sua definição, mesmo que essa tenha um caráter parcialmente enciclopédico. Em virtude disso, ao contrário do que propõem alguns pesquisadores, como Dubuc³⁴ (1999, *apud* ALMEIDA; SOUZA; PINO, 2007), sobre a construção da definição, a mesma não obedeceu, necessariamente, ao preceito da concisão. Campos (1994, p. 46) comenta que quanto mais curta e concisa é uma definição, mais difícil ela se torna. Como no exemplo acima apresentado, quando da construção de um conceito com mais de uma sentença,

³⁴ DUBUC, R. **Manual práctico de terminología**. Chile: RiL Editores, 1999.

preferiu-se, no intuito de tornar o entendimento mais claro, o uso de ponto-e-vírgula em vez de longas sentenças separadas apenas por vírgulas.

Tem razão Fraústo (1990, p.1) ao afirmar que “... a construção de uma terminologia, informatizada ou não, implica uma forte componente criativa”. “Engessar” a construção do conceito final, elaborado a partir de um *corpus* que sempre pode ser atualizado, limita as possibilidades de reconstrução desse conceito. Além disso, qual a garantia de que normas internacionais referentes à terminografia hoje vigentes serão adequadas para essas novas composições de *corpora*? A preocupação primeira do terminógrafo, então, deve ser com o seu público-alvo. Por essa razão, tendo em vista os tradutores e a dinamicidade da língua, preferi trabalhar com esse componente criativo: não há razão para “engessar” a construção da definição.

É importante notar, também, a diferenciação feita entre conceito final e definição no banco de dados, como mostrado na figura 29 abaixo. Cabré propõe que a definição deva obedecer aos critérios previamente estabelecidos quando do planejamento da microestrutura da obra (1993, p. 207-213). A definição, no presente caso, pode refletir esses critérios e modificar (ou não) os dados fornecidos pelo conceito final, adequando o texto final da definição às características da microestrutura da obra em desenvolvimento. No caso do campo de computação, neste trabalho, decidiu-se pela duplicação, em todos os casos, do conceito final na definição. Levando em conta o componente criativo, não se constatou a necessidade de haver uma construção da definição totalmente igual, ou seja, padronização, para todos os verbetes.

http://jr.icmc.sc.usp.br - Vocabulário Técnico Online - Mozilla Firefox

Termo: processor

Passo Anterior Salvar Sair sem salvar

Contextos **Conceito Final / Definição**

Conceito Final / Definição

Conceito Final: flat square sliver of silicon device composed of circuits etched on its surface; it controls the computer and execute programs; also known as CPU (central processing unit)

Definição: flat square sliver of silicon device composed of circuits etched on its surface; it controls the computer and execute programs; also known as CPU (central processing unit)

Dados **Traços Distintivos** Semântica Termo Equivalente Termos Remissivos Informações Enciclopédicas

Traços Distintivos

Nova Coluna — Não use aspas duplas

	A	B	C	D	E	F	G	H	I
1	calculates	performs	functions						
2				flat squar...	circuits e...				
3						speed mea			
4		execute					CPU	control	programs

Concluído

Figura 29. Os traços distintivos (parte inferior) e a criação do conceito final e definição (parte superior).

Barbosa (1994) já alerta que um conjunto noêmico (conceito) de uma língua e seu respectivo conjunto léxico-semântico nem sempre encontram os equivalentes perfeitos em outra língua, demandando, por parte do lexicógrafo/terminógrafo, uma nota explicativa de caráter enciclopédico. Uma característica importante resultante dessa observação pode ser encontrada nas definições levantadas: elas não são, necessariamente, uma a tradução da outra; representam, na verdade, o processo de descrição, em cada língua e em cada momento, de um determinado conceito. Têm razão González-Jover e Sierra (2007) quando mencionam que:

[...] o dicionário não só há de lhe oferecer aquela informação que se considera necessária e suficiente sobre os conceitos, mas também deverá incluir informação “enciclopédica”, relativa ao conhecimento extralingüístico. Neste sentido, o dicionário especializado bilíngüe destinado ao tradutor tem que conter informações que podem superar os limites do estritamente necessário e suficiente: há de fazer as vezes

de um dicionário monolíngüe no qual apareçam uma definição e contextos de usos real do termo validado que assegurem a confiança da informação e, assim mesmo, há de conter outras informações como sinônimos, variações significativas determinadas pelo contexto, categoria gramática, relações conceituais, etc., que sirvam de ajuda ao tradutor no processo de tomada de decisões que é a tradução.

(tradução minha)³⁵

Em outras palavras, os levantamentos das Fichas Terminológicas, no banco de dados, foram pensados como monolíngües: o que os une, como vocabulário bilíngüe, são os Paradigmas de Forma Equivalente e a visualização na página de consulta (VoTec). Com essa abordagem, o banco de dados pode funcionar, também, para a construção de vocabulários monolíngües. A diferenciação no modo de consulta se dará a partir da criação de diferentes páginas (*sites*), com diferentes características de busca, que acessam o banco.

Vejamos exemplos de definições bilíngües de alguns termos que estão disponíveis para consulta na página elaborada para a tese:

1. computador (*computer*): máquina capaz de realizar tarefas de tratamento de informações em alta velocidade, com precisão, formada por circuitos e dispositivos./ *device with a microprocessor that receives inputs, processes them and displays results.*
2. impressora (*printer*): periférico composto por placa lógica, motor, engrenagens, correias, eixo, carro de impressão e cabeça de impressão; imprime fotos, desenhos, textos, tabelas, dados, relatórios, etc.; podem ser coloridas ou branco e

³⁵ [...] el diccionario no sólo ha de ofrecerle aquella información que se considera necesaria y suficiente sobre los conceptos, sino que también deberá incluir información de la denominada «enciclopédica», relativa al conocimiento extralingüístico. En este sentido, el diccionario especializado bilingüe destinado al traductor ha de contener informaciones que pueden superar los límites de lo estrictamente necesario y suficiente: ha de hacer las veces de un diccionario monolingüe en el que aparezcan una definición y contextos de uso real del término validados que aseguren la fiabilidad de la información y, asimismo, ha de contener otras informaciones como sinónimos, variaciones significativas determinadas por el contexto, categoría gramatical, relaciones conceptuales, etc., que sirvan de ayuda al traductor en el proceso de toma de decisiones que es la traducción.

preto; os três tipos principais são: matricial, jato de tinta e laser./ *peripheral device that prints two-dimension graphics, texts, color charts and photographs in hard-copy paper; the two main types are laser and inkjet.*

3. memória (*memory*): chip (formado por células idênticas, na forma de linhas e colunas) ou área para armazenamento de dados; unidade exibida em Mbytes, cujo desempenho é medido pela taxa de transferência; pode ser real ou virtual, associada ao Windows/ *place where the computer stores current programs and data (information) in use; it's connected to the processor by its address and data buses.*

Os exemplos mostram que as definições, em ambas as línguas, não são idênticas e oferecem semas complementares para o melhor entendimento do conceito. Isso se dá em virtude de uma das particularidades do campo da computação, ou seja, a língua em que uma nova tecnologia ou processo foram criados em primeiro lugar. Os textos criados para a confecção de manuais ou os artigos jornalísticos representam, naquela língua e naquele momento, o estágio de desenvolvimento da técnica. Textos em outras línguas, sobre o mesmo assunto, podem ter sido produzidos na mesma época ou não. Esse é um dos fatores (o outro seriam as diferenças, várias, entre as fontes) que gera as diferenças nas definições: enquanto os textos de um termo na língua X, na qual o produto/tecnologia/processo foi criado há alguns anos, apresentam poucos semas para a construção da definição, outros textos, na língua Y, que descrevem esse produto/tecnologia/processo hoje, apresentam uma quantidade maior de semas. Há claramente uma tendência a vulgarizar (passagem da terminologia para a língua comum; BARBOSA, 1998) os termos com o passar do tempo e eles, quando da sua apresentação nos textos, passam de um tipo de contexto para outros: de contextos definitórios,

tornam-se explicativos e depois associativos, como o termo *computador*. Não há a necessidade, por exemplo, de se explicar o que é um computador em países nos quais a tecnologia já existe há décadas; o inverso se dará em países onde a tecnologia ainda é incipiente. No caso do *corpus* de computação aqui apresentado, significa que termos mais antigos, em inglês, podem apresentar, nos contextos, menos semas para a construção da definição do que os equivalentes, em português.

6.7 Paradigma Informacional (PI)

O próximo passo, ainda na mesma janela (parte inferior), com a visualização dos contextos em destaque, é a extração de várias informações morfossintáticas (figura 30): classe gramatical, número, gênero e possíveis siglas ou acrônimos³⁶, entrada por extenso (caso o termo seja uma sigla ou acrônimo), variações morfossintáticas, o número da acepção³⁷. São inseridos também os campos que localizam o termo em relação ao *corpus* de onde foi tirado: a frequência do termo na listagem de todas as palavras e o número total de vezes em que ele aparece.

³⁶ “Muitos dos termos sintagmáticos, por razões de economia discursiva, são reduzidos sob a forma de siglas, termos formados pelas letras iniciais de cada elemento do sintagma, ou de acrônimos, termos que se constituem com a formação de sílabas extraídas do sintagma, geralmente as iniciais” (Alves, 2006).

³⁷ Embora obras terminológicas tendam a apresentar definições monossêmicas, preferimos inserir esse campo. Algumas áreas, que já apresentam uma grande diversidade de terminologia, como a informática, podem vir a apresentar alguns termos polissêmicos em suas diferentes subáreas. Embora a definição do termo seja única em cada uma dessas subáreas (ou seja, são monossêmicas), numa consulta genérica à área geral, mais de uma definição poderia ser apresentada.

Dados

Ontologia: Computação > Hardware

Categoria Gramatical: Substantivo Número: Singular

Gênero: Masculino Sigla/Acrônimo: CPU, UCP

Entrada por Extenso:

Var. Morfossintáticas:

Acepção Nº: 1

Cópus

Posição na Ordem de Freqüência: 67 Nº de Ocorrências do termo: 1212

18/07/2007 15:59 © 2007 FFLCH - ICMC Jr.

Concluído

Figura 30. Informações morfossintáticas do termo “processador”

6.8 Paradigma Semântico (PS)

O Paradigma Semântico expõe as relações de hiperonímia, hiponímia, co-hiponímia, antonímia e sinonímia do termo em questão com outros termos disponíveis nos contextos extraídos, caracterizando possíveis remissivas. Além disso, o pesquisador poderá cruzar referências com obras já publicadas, verificando se o termo é dicionarizado, se há definições coincidentes, a fonte da definição e a definição dicionarizada em si. A tela abaixo (figura 31) apresenta as possibilidades acima descritas para o termo “processador”:

Fig. 31. Informações semânticas e referência do mesmo termo/lexema em outras obras, termo *processador*.

6.9 Paradigma de Forma Equivalente (PFE)

Cada termo deve ter o seu equivalente na outra língua. É o preenchimento desse campo (figura 32), conforme descrito anteriormente, que possibilitará a visualização dos vocabulários como bilíngües (o assunto é melhor descrito na apresentação da página de consulta, capítulo 7.1). Os equivalentes, já detectados numa fase anterior (capítulo 4), precisam ser, agora, apenas associados.

Figura 32. Termo equivalente a “processador”, em inglês (em destaque).

6.10 Remissivas

Termos remissivos (remissivas são relações, traçadas dentro da obra lexicográfica/terminológica, entre os termos, mantendo a coerência semântica dessa [FROMM, 2002]), levantados a partir dos campos semânticos acima citados (sinônimos, antônimos, hipônimos, co-hipônimos e hiperônimos), só poderão ser incluídos dentre de um verbete se os mesmos já fizerem parte do banco como outros verbetes. Embora sejam discriminados, também, na composição da microestrutura, eles só funcionarão como *hyperlink* na mesma se forem inseridos (figura 33) como elementos de pesquisa remissiva.

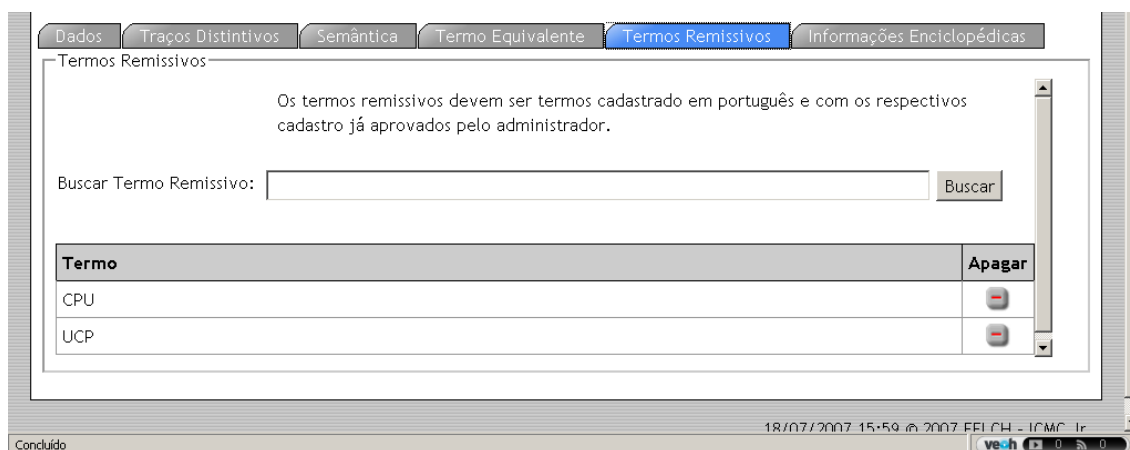


Figura 33. Termos remissivos para *processador*.

6.11 Paradigma Enciclopédico (PE)

Há a possibilidade, também, de inserção de informações enciclopédicas no banco (figura 34). Embora a tela ofereça a possibilidade de inclusão de áudio e vídeo, essas facilidades ainda não foram implementadas na versão atual do programa.

Definição: O processador é a parte mais fundamental para o funcionamento de um computador

Artigo: processador

Fonte: Wikipedia

Link: <http://pt.wikipedia.org/wiki/Processador> Tipo: Imagem

Adicionar

Definição	Artigo	Fonte	Link	Apagar
O processador é a parte mais fundamental para o funcionamento de um computador. Processadores são circuitos digitais que realizam operações como: cópia de dados, acesso a memórias e operações lógicas e matemáticas.	Processador	Wikipedia		

Figura 34. Tela de inserção e visualização de informações enciclopédicas.

6.12 O Enunciado Terminográfico resultante

Somando todas as possibilidades de paradigmas acima apresentadas, a construção do banco de dados oferece a seguinte fórmula para a construção do Enunciado Terminográfico do termo:

$$\mathbf{Termo = \{+ entrada + enunciado terminográfico (\pm PI + PD \pm PS + PFE + PP \pm Remissivas \pm PE)\}}$$

O sinal \pm representa a opcionalidade (nem todos os campos podem ser preenchidos em virtude da carência de informações apresentadas pelos exemplos) e o sinal $+$ representa a obrigatoriedade. A ordem dos componentes do enunciado terminográfico acima apresentado não representa, necessariamente, a ordem de visualização na página de consulta, já que a mesma apresenta diferentes possibilidades de construção da microestrutura. Pode-se tomar como exemplo o termo “processador”:

processador. (CPU, UCP). *Hardware.* **s.m.s.** tipo de chip, parte principal do computador, também conhecido como CPU (Central Processing Unit); executa instruções (processos) e dados (através da UC - unidade de controle) e operações lógicas e aritméticas (através da UAL - unidade aritmética e lógica). Ex.: Processador (UCP). A UCP é a parte principal do computador responsável pelo processamento e execução de programas armazenados na memória principal. *Sinônimos:* CPU, UCP. *Hipônimo de:* computador. *Hiperônimo de:* UAL (unidade aritmética e lógica); UC (unidade de controle). *Co-hipônimos:* chip. *Veja Também:* chipset, CPU, UCP. **Cópus:** *Posição na Ordem de Frequência:* (67); *Nº de Ocorrências do termo:* (1212). **Informações Enciclopédicas:** O processador é a parte mais fundamental para o funcionamento de um computador. Processadores são circuitos digitais que realizam operações como: cópia de dados, acesso a memórias e operações lógicas e matemáticas. Em: *Processador* – [Wikipedia](#)

7. O aprendiz e o VoTec

Apresento, a seguir, uma descrição pormenorizada da página de consulta (VoTec), disponibilizada, provisoriamente, em <<http://jr.icmc.sc.usp.br/~comet/dic>>. Defendo, quanto ao aprendizado dos tradutores iniciantes, dois níveis: o inicial, baseado num trabalho de aprendizagem dos paradigmas que compõem uma microestrutura por meio da página de consulta e o avançado, baseado no trabalho direto com o banco de dados.

7.1 O aprendiz iniciante: usando a página de consulta

Para todos os consulentes de obras terminográficas eletrônicas, o levantamento de dados de um banco qualquer se dá mediante ferramentas desenvolvidas visando a extração de dados específicos ou gerais. Proponho, nesta tese, uma página como fonte de treinamento de alunos de tradução na área de vocabulários técnicos. O aprendiz iniciante tem acesso aos dados do banco, porém somente para consulta. A inovação proposta será a forma de consultar o banco através da página. A construção da página (ambiente web) foi feita em conjunto com a construção do banco de dados.

7.1.1 Apresentação da página e modo de consultar

A página inicial de consulta, assim como todas as subseqüentes, foi desenhada como página monolíngüe (português ou inglês) com acesso direto à outra língua de exibição: basta clicar no canto superior direito para trocar a língua da tela. Além do

modo de exibição normal, o site pode ocupar, também, toda a tela do computador (basta clicar em “tela cheia/full screen”). São apresentadas aqui essas duas telas (figuras 35 e 36); serão trabalhadas na seqüência, porém, somente as telas em português:

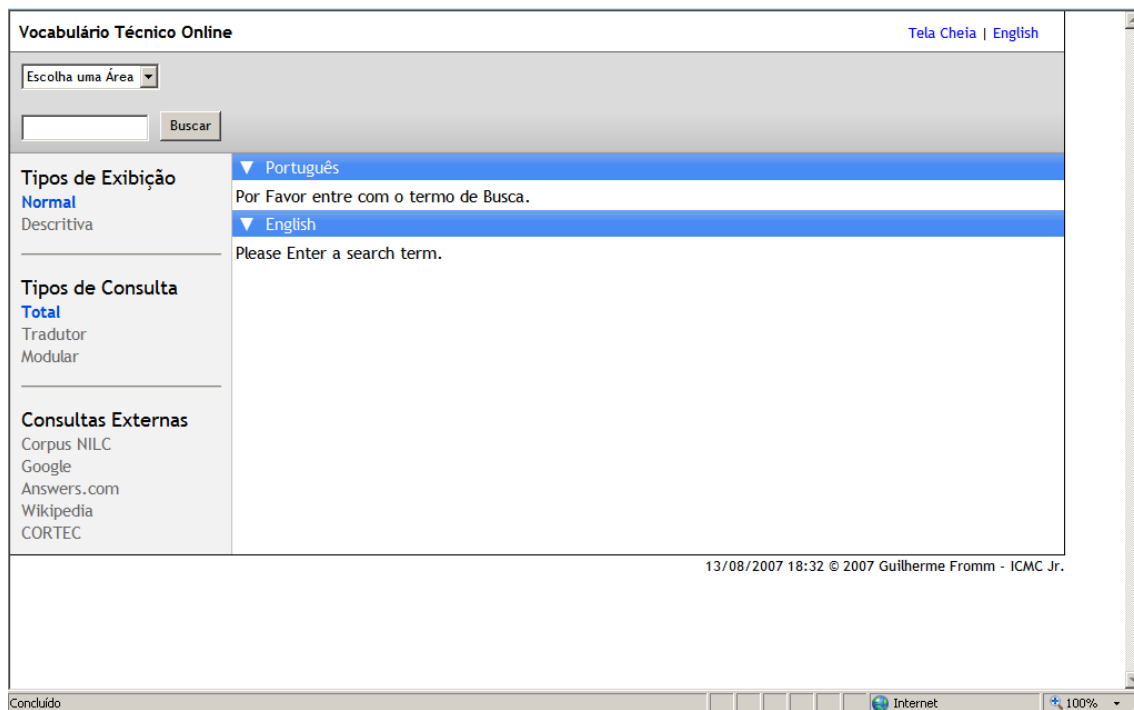


Figura 35. Tela inicial de consulta do site, em português.

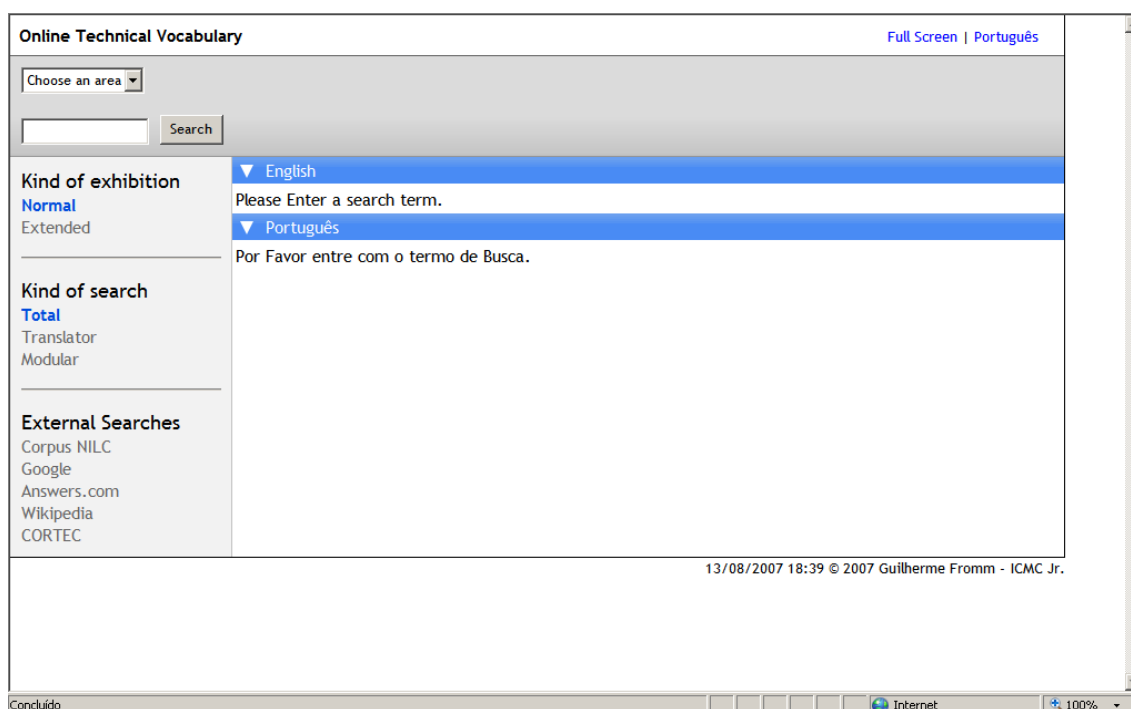


Figura 36. Tela inicial de consulta do site, em inglês.

Aubert (2001, p. 43) coloca que o livre-comércio, num mundo cada vez mais globalizado, pressupõe uma bidirecionalidade na comunicação e, portanto, os materiais terminográficos também devem ser bidirecionais; modelos como aqueles que colocam a língua estrangeira como fonte e a língua vernácula como alvo estão superados. Campos (1994, p. 49) comenta que os dicionários bilíngües têm uma vantagem sobre os monolíngües: uma simples tradução pode evitar uma explicação prolixa.

Tendo os pensamentos dos dois autores como mote, pode-se notar, nas figuras acima, que a base da visualização do site, seja qual for o modo de exibição ou tipo de consulta, é o Paradigma da Forma Equivalente (tradução). No lado direito da tela, aquela com a maior área de visualização, as barras horizontais (em azul, no site) delimitam o espaço dedicado a cada língua para o mesmo termo. Convém comentar, agora, os modos de exibição e os tipos de busca.

7.1.2 Modos de exibição: normal e descritivo

Novos dicionários eletrônicos costumam utilizar melhor os recursos que os computadores oferecem na hora de consultar um termo. Leffa (2006, p. 336), ao comparar o dicionário eletrônico ao tradicional em papel, comenta que “o conteúdo pode ser o mesmo, mas o suporte fará a diferença, quer seja na rapidez do acesso, quer seja na forma como os dados são apresentados”.

Uma forma de apresentação de dados, já presente em dicionários eletrônicos gerais de língua, como o Dicionário Eletrônico Houaiss da Língua Portuguesa (2002), é a interativa (descritiva, nesta tese). Não havendo a necessidade de se “comprimir” as informações numa microestrutura espacialmente pequena, no formato tradicional (figura 37), pode-se ampliá-la até os limites da tela ou além. O Houaiss, ao utilizar o modo interativo (figura 38), dá especial destaque às várias acepções de um mesmo verbete, disponibilizando-as de forma sucessiva na vertical e utilizando, inclusive, várias cores. O dicionário denomina esses dois tipos de visualizações como *modo tradicional* e *modo interativo*.

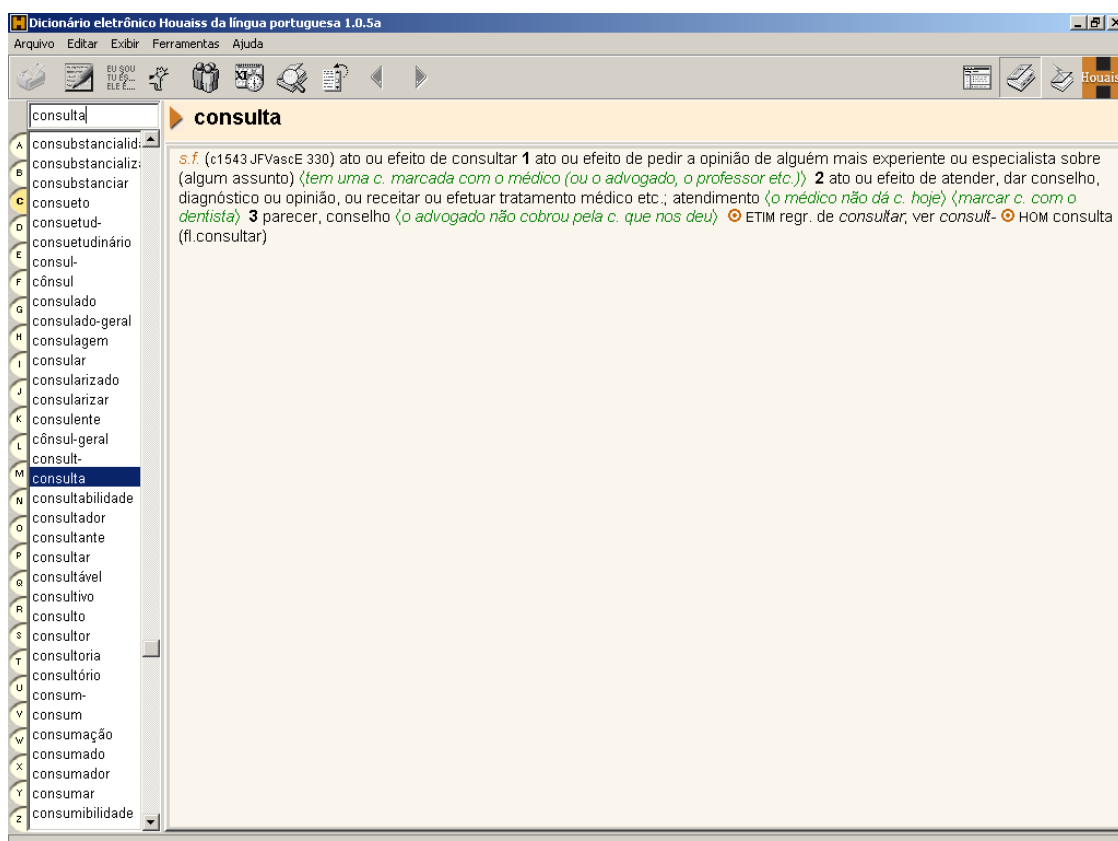


Figura 37. Dicionário Houaiss. Verbetes *consulta*, modo de visualização tradicional.

Além dos recursos de visualização apresentados, o dicionário permite que todas as palavras apresentadas, em qualquer definição, funcionem como *links* para a sua própria definição (por meio do recurso de hipertexto). Essa funcionalidade, embora bastante desejável, não se aplica ao presente estudo. Isso se dá por não se tratar de um dicionário geral de língua, mas um vocabulário terminológico, ou seja, não há como ligar qualquer palavra que compõe a definição de um termo à sua respectiva definição, já que o vocabulário é composto apenas de termos da área em questão.

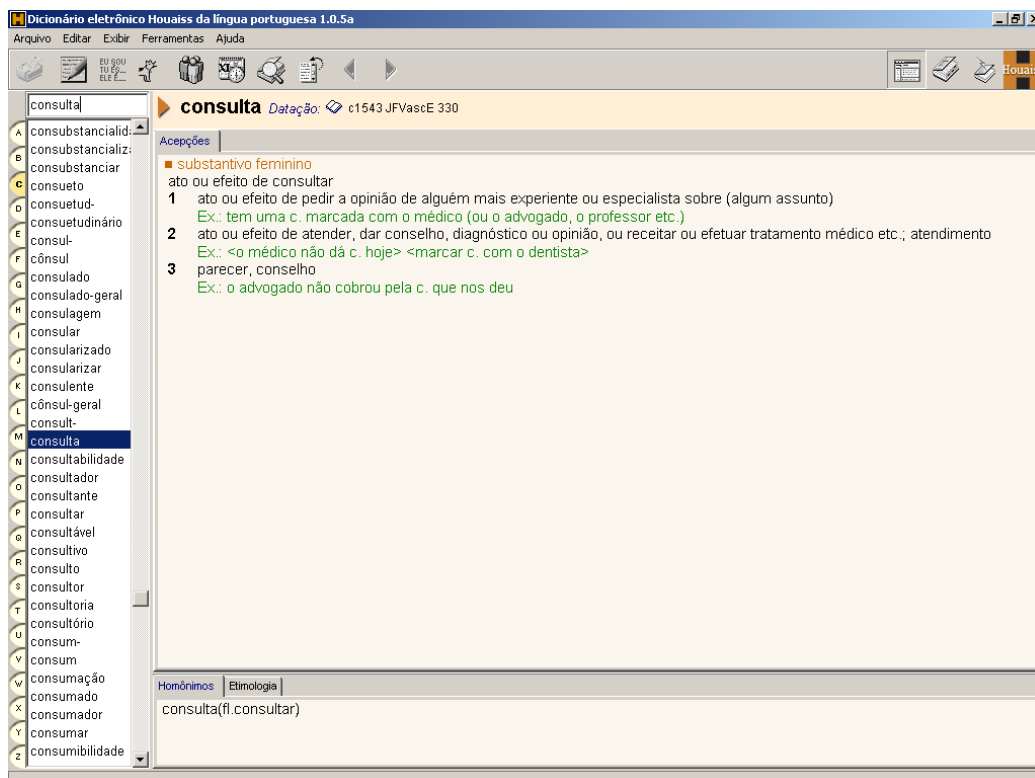


Figura 38. Dicionário Houaiss. Verbetes *consulta*, modo de visualização interativo.

Tomando parte desse exemplo de diferentes visualizações como um paradigma, adotamos o mesmo modelo, porém com denominações diferentes: tipos de exibição normal, que reproduz o leiaute de uma obra impressa, e descritivo, onde cada item do enunciado terminográfico é disposto em linhas diferentes. Para mudar o formato, basta clicar numa das duas opções do lado esquerdo da tela (figuras 39 e 40).

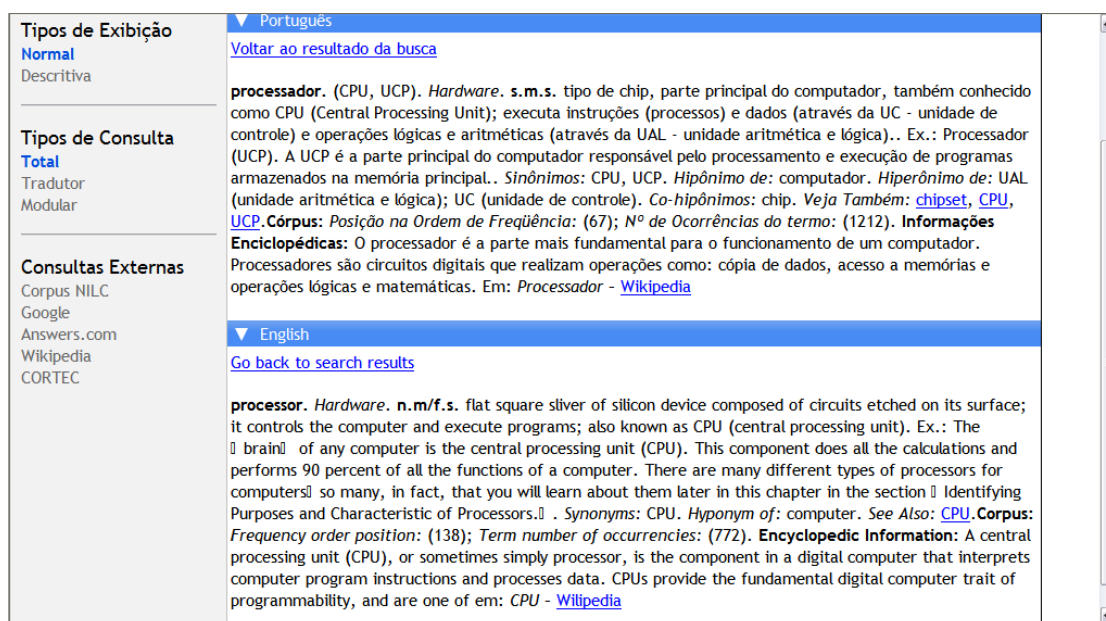


Figura 39. Termo *processador*. Tipo de exibição normal.

Vocabulário Técnico Online Tela Cheia | English | Ajuda

Computação Escolha uma área

processador Buscar

Tipos de Exibição

Normal
Descritiva

Tipos de Consulta

Total
Tradutor
Modular

Consultas Externas

Corpus NILC
Google
Answers.com
Wikipedia
CORTEC

▼ Português

[Voltar ao resultado da busca](#)

processador

tipo de chip, parte principal do computador, também conhecido como CPU (Central Processing Unit); executa instruções (processos) e dados (através da UC - unidade de controle) e operações lógicas e aritméticas (através da UAL - unidade aritmética e lógica).

Abreviação/Acrônimo: CPU, UCP

Categoria Gramatical: substantivo

Gênero: masculino

Número: singular

Cópus

- Posição na Ordem de Frequência: 67
- Nº de Ocorrências do termo: 1212

Ontologia: [Computação](#) > [Hardware](#)

Variações Morfosintáticas: *Nada encontrado*

Sinônimos: CPU, UCP

Antônimos: *Nada encontrado*

Hipônimo de: computador

Co-hipônimos: chip

Hiperônimo de: UAL (unidade aritmética e lógica); UC (unidade de controle)

Exemplos

1. Processador (UCP). A UCP é a parte principal do computador responsável pelo processamento e execução de programas armazenados na memória principal.
2. UCP


Sua função consiste em coordenar, controlar e realizar todas as operações (execução de instruções) do sistema.
3. UCP

Suas partes principais são as seguintes:

 - a unidade aritmética e lógica (UAL): realiza o cálculo real ou o processamento de dados (realiza as operações aritméticas e lógicas).
 - a unidade de controle (UC): controla o movimento de dados e instruções dentro e fora da UCP e controla a operação da UAL, de forma adequada e sincronizada.

[3 mais...](#)

Informações Enciclopédicas

 O processador é a parte mais fundamental para o funcionamento de um computador. Processadores são circuitos digitais que realizam operações como: cópia de dados, acesso a memórias e operações lógicas e matemáticas. Em: *Processador* - [Wikipedia](#)

Veja Também

[chipset](#) - [CPU](#) - [UCP](#)

▼ English

[Go back to search results](#)

PROCESSOR

Concluído Internet 100%

Figura 40. Termo *processador*. Tipo de exibição descritivo.

7.1.3 Consultas internas: total, tradutor e modular

Os tipos de consulta foram divididos em três (figuras 41, 42 e 43). A consulta total reflete a disposição de todos os campos disponíveis no banco de dados e tem um caráter mais lexicográfico. A consulta tradutor reflete os campos mais usados pelos tradutores, conforme pesquisa prévia. O campo modular é a ferramenta desenvolvida para que o aluno aprenda a trabalhar com a microestrutura do termo em questão (através de sua manipulação) e, a partir desse treinamento, adquira uma melhor compreensão de como utilizar uma obra terminográfica.

7.1.3.1 Consulta Total

A consulta total apresenta todos os campos que existem no banco de dados, com exceção das informações acerca de termos equivalentes em dicionários já publicados (por uma questão de direito autoral). Veja um exemplo, no tipo de exibição normal, para o termo *chipset* (figura 41):

The screenshot shows the 'Vocabulário Técnico Online' interface. At the top, there is a search bar with 'Computação' selected in the dropdown and 'Escolha uma área' in another. Below the search bar is a 'Buscar' button. The interface is divided into several sections:

- Tipos de Exibição:** Includes 'Normal' (selected) and 'Descritiva'.
- Tipos de Consulta:** Includes 'Total' (selected), 'Tradutor', and 'Modular'.
- Consultas Externas:** Lists 'Corpus NILC', 'Google', 'Answers.com', 'Wikipedia', and 'CORTEC'.

The main content area displays the definition for 'chipset' in Portuguese:

chipset. *Placa-mãe. m.s.* conjunto de circuitos existentes na placa-mãe que apoiam o processador, gerenciando o mesmo e também responsáveis pelas informações de reconhecimento do hardware. Ex.: Chipset são circuitos de apoio ao processador (este sozinho não consegue controlar todo funcionamento da placa-mãe) que gerência praticamente todo funcionamento da placa mãe (controle de memória cache, DRAM, controle do buffer de dados, interface com a CPU, etc). É responsável pelas informações necessárias ao reconhecimento de hardware (armazenadas na sua memória ROM).. *Hipônimo de:* chip. *Veja Também:* [memória](#), [placa-mãe](#), [processador](#). **Córpus:** *Posição na Ordem de Frequência:* (389); *Nº de Ocorrências do termo:* (313). **Informações Enciclopédicas:** De uma forma geral, um chipset (anglicismo que significa grupo de chips) é o cérebro de uma placa de circuitos. Em: *chipset* - [Wikipedia](#)

Below this, there is a section for 'English' with a link 'Go back to search results' and another definition for 'chipset' in English:

chipset. *Motherboard. n.m/f.s.* chip or circuitry that connects the microprocessor to the rest of the motherboard and memory, expansion cards and onboard peripherals. Ex.: A chipset is a collection of chips or circuits that perform interface and peripheral functions for the processor. This collection of chips is usually the circuitry that provides interfaces for memory, expansion cards, and onboard peripherals and generally dictates how a motherboard will talk to the installed peripherals.. *See Also:* [bus](#), [motherboard](#). **Corpus:** *Frequency order position:* (724); *Term number of occurrences:* (188). **Encyclopedic Information:** In computing, the term chipset is commonly used to refer to a set of specialized chips on a computer's

Figura 41. Termo *chipset*, tipo de exibição normal, tipo de consulta total.

7.1.3.2 Consulta Tradutor

A consulta no formato tradutor reproduz os campos que a pesquisa com os profissionais da área revelou serem os mais procurados, ou seja, área de especialidade, tradução, sinônimos, definição, etc. O Paradigma Definicional é a base; o mesmo é precedido pelos elementos mais trabalhados pelos tradutores no Paradigma Informacional (área, abreviatura, categoria gramatical, gênero e número, diferenças ortográficas) e sucedido pelos paradigmas Pragmático (com a apresentação de um único exemplo), Semântico (sinônimos) e a Remissiva. Continua-se com o mesmo termo (figura 42), embora ele não forneça todos os campos acima citados:

The screenshot shows the 'Vocabulário Técnico Online' interface. At the top, there's a search bar with 'Computação' selected and a 'Buscar' button. Below the search bar, there are two columns of filters. The left column, 'Tipos de Exibição', has 'Normal' selected. The 'Tipos de Consulta' section has 'Tradutor' selected. The 'Consultas Externas' section lists various sources. The main content area shows the definition for 'chipset' in Portuguese and English. The Portuguese definition is: 'chipset. Placa-mãe. m.s. conjunto de circuitos existentes na placa-mãe que apoiam o processador, gerenciando o mesmo e também responsáveis pelas informações de reconhecimento do hardware. Ex.: Chipset são circuitos de apoio ao processador (este sozinho não consegue controlar todo funcionamento da placa-mãe) que gerência praticamente todo funcionamento da placa mãe (controle de memória cache, DRAM, controle do buffer de dados, interface com a CPU, etc). É responsável pelas informações necessárias ao reconhecimento de hardware (armazenadas na sua memória ROM).. Veja Também: [memória](#), [placa-mãe](#), [processador](#).' The English definition is: 'chipset. Motherboard. n.m/f.s. chip or circuitry that connects the microprocessor to the rest of the motherboard and memory, expansion cards and onboard peripherals. Ex.: A chipset is a collection of chips or circuits that perform interface and peripheral functions for the processor. This collection of chips is usually the circuitry that provides interfaces for memory, expansion cards, and onboard peripherals and generally dictates how a motherboard will talk to the installed peripherals.. See Also: [bus](#), [motherboard](#).' At the bottom of the page, there is a copyright notice: '24/11/2007 15:34 © 2007 Guilherme Fromm - ICMC Jr.'

Figura 42. Termo *chipset*, tipo de exibição normal, tipo de consulta tradutor.

7.1.3.3 Consulta Modular

A consulta modular exhibe a microestrutura conforme a preferência/necessidade do consulente. No lado esquerdo (figura 43) são disponibilizados os paradigmas citados no capítulo 7, porém com nomes não técnicos: Pré-Definição (Paradigma Informacional), Definição (Paradigma Definicional), Exemplificação (Paradigmas Pragmático e Enciclopédico) e Relações (Paradigma Semântico).

The screenshot shows the 'Vocabulário Técnico Online' interface. At the top, there are dropdown menus for 'Computação' and 'Escolha uma área', a search input field, and a 'Buscar' button. The main content area is divided into two language sections: Portuguese and English. The Portuguese section shows the definition of 'chipset' as a set of circuits on the motherboard that manage the processor and memory. The English section shows the definition of 'chipset' as a collection of chips that provide interfaces for memory, expansion cards, and onboard peripherals. The sidebar on the left contains various navigation options for different types of queries and definitions.

Figura 43. Termo *chipset*, tipo de exibição normal, tipo de consulta modular.

Pelo exemplo na figura 43, nota-se que os campos categoria gramatical, ontologia, *corpus*, informações enciclopédicas e hiperônimo foram desativados, ou seja, não constam mais da microestrutura. Campos que ainda estão selecionados, porém não aparecem na microestrutura, não têm informações a serem disponibilizadas para esse termo. O aprendiz descobre, através da manipulação da microestrutura, todos os componentes do Enunciado Terminográfico (conforme item 7.9) aqui propostos. Isso o torna apto a perceber as diversas possibilidades de estruturas em outras obras do gênero. Ao configurar esse tipo de consulta, o aprendiz (e o tradutor profissional) pode, também, aumentar a velocidade de sua tradução ao acessar somente os campos que lhe interesse.

7.1.4 Consultas externas

Além da informação enciclopédica disponível em cada microestrutura, o aluno pode consultar outras fontes, na Internet, para apreender melhor o sentido do termo buscado. No intuito de treinar o estudante a trabalhar com diferentes tipos de fontes, foram selecionadas cinco, entre milhares disponíveis, para consulta a partir do site. Cada uma tem a sua peculiaridade:

1. *Corpus* do NILC: acesso direto ao site do Projeto NILC (Núcleo Interinstitucional de Linguística Computacional – ICMC/USP/São Carlos). O *corpus* aqui disponível é um *corpus* geral de língua portuguesa e fornece ferramentas de análise lexical para acessar o conjunto de textos. As buscas devem ser feitas em português (embora o *corpus* possa apresentar algumas palavras em inglês), quer a interface do VoTec esteja em português ou inglês;

2. site de busca Google: o instrumento, por excelência, de busca na Internet; como no item anterior, tanto a interface em português quanto a do inglês do VoTec levam a uma busca na página em português do Google (a partir dessa página, novas buscas podem ser realizadas);
3. site Answers.com: um dos vários dicionários online disponíveis na Internet; as pesquisas aqui realizadas, nas interfaces em português ou inglês, conduzem a uma página em inglês. O site fornece descrições dicionarizadas, enciclopédicas e termos equivalentes em várias línguas para os lexemas ou termos consultados.
4. Wikipedia (também Wikipédia): a maior enciclopédia online gratuita do mundo; diferentemente dos itens anteriores, a interface do VoTec escolhida remete a páginas diferentes na Wikipedia (interface em português/ busca em português/ página da Wikipedia em português; interface em inglês/ busca em inglês/ página da Wikipedia em inglês)
5. CORTEC: o site de consulta aos *corpora* técnicos bilíngües disponíveis no Projeto COMET; permite consultar, através do uso de ferramentas de análise lexical, os *corpora* já disponíveis no site do projeto.

Os quatro primeiros abrem um campo de consulta logo abaixo do seu nome. Basta o aluno inserir o termo desejado e uma nova página se abrirá, já com a pesquisa do termo realizada. O quinto abre direto uma página para consulta. Não é necessária a consulta a nenhum termo do VoTec para acessar essas páginas.

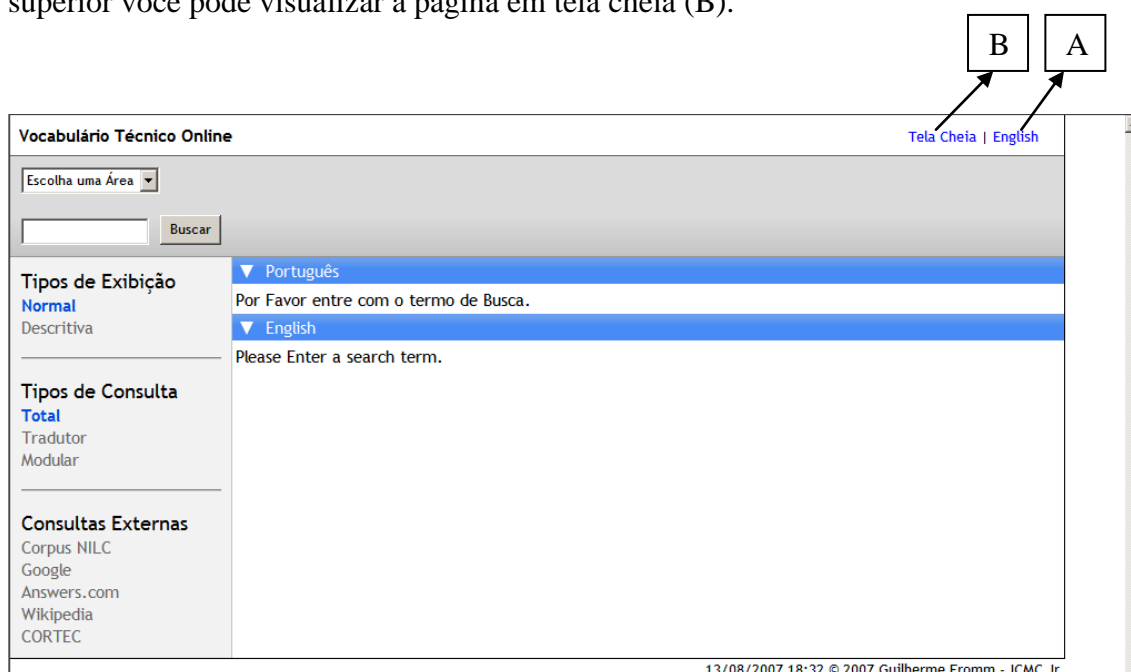
7.1.5 A ajuda online

É apresentada, a seguir, uma reprodução da página de ajuda disponível no site³⁸.

A versão em inglês da mesma página consta do apêndice F. O objetivo da mesma é tornar o consulente auto-suficiente na sua consulta, sem a necessidade de treinamento.

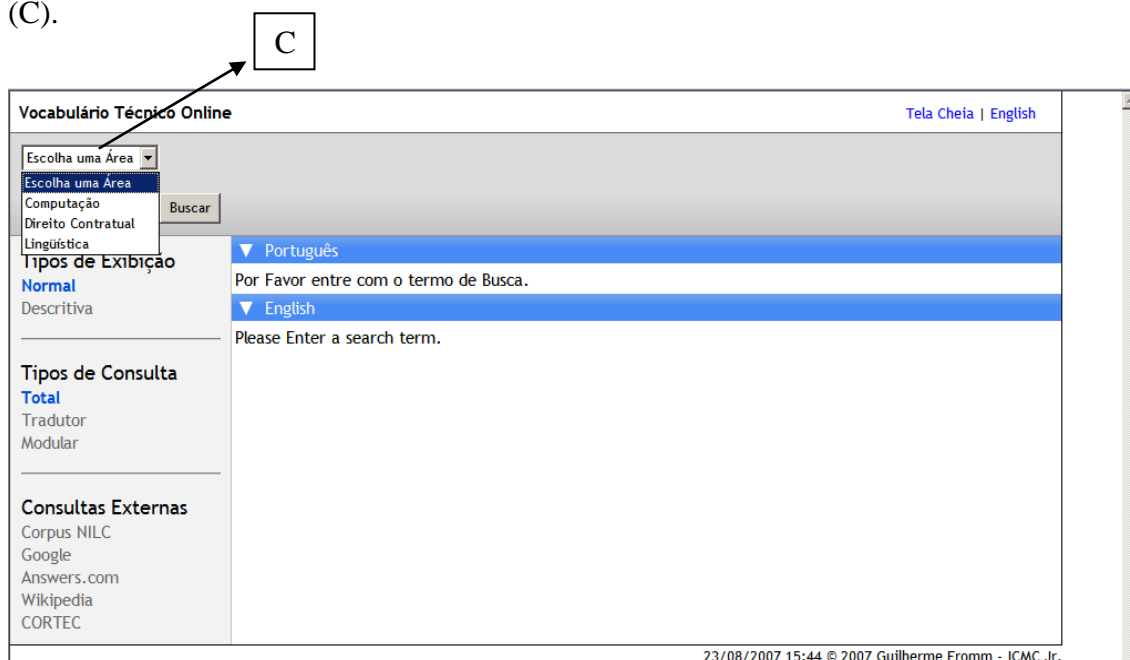
Apresentação da página e modo de consultar

Você pode escolher a língua em que quer visualizar a página: português ou inglês. Basta clicar no canto direito superior para mudar a língua (A). Também nesse canto direito superior você pode visualizar a página em tela cheia (B).

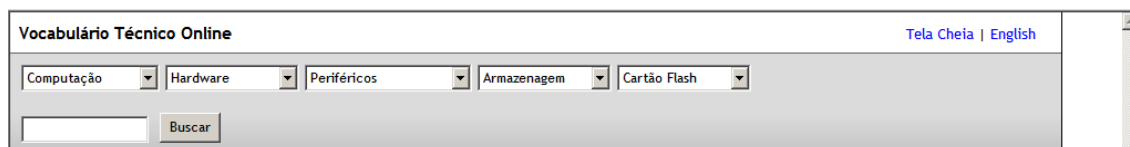


³⁸ Note-se que numa página da Internet, não há quebras. O texto apresenta-se, portanto, numa única página e as figuras não estão divididas, como podem ser apresentadas aqui.

Para começar a consulta, escolha uma área técnica no menu do canto esquerdo superior (C).

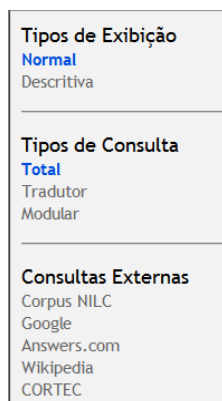


Você pode, ao selecionar uma área, descer até mais quatro níveis dentro dela (sua subdivisão interna):

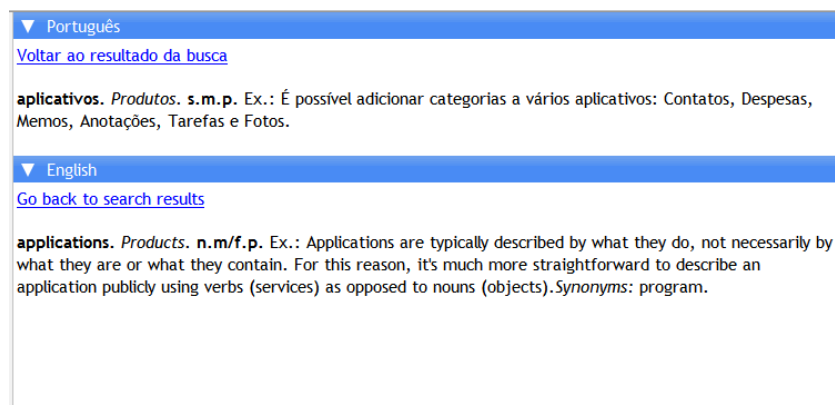


A tela de consulta

No lado esquerdo da tela você encontra os menus de visualização, tipos de consulta e consulta externa.



No lado direito é mostrado o verbete em português e inglês.



Tipos de exibição: normal e descritiva

Os modos de exibição representam dois momentos na história dos dicionários:

1. Exibição normal, que mostra o verbete da mesma forma que os dicionários em papel.

Vocabulário Técnico Online Tela Cheia | English

Computação Escolha uma área

Buscar

<p>Tipos de Exibição</p> <p>Normal</p> <p>Descritiva</p> <hr/> <p>Tipos de Consulta</p> <p>Total</p> <p>Tradutor</p> <p>Modular</p> <hr/> <p>Consultas Externas</p> <p>Corpus NILC</p> <p>Google</p> <p>Answers.com</p> <p>Wikipedia</p> <p>CORTEC</p>	<p style="background-color: #0070C0; color: white; padding: 2px;">▼ Português</p> <p>Voltar ao resultado da busca</p> <p>processador. (CPU, UCP). <i>Hardware. s.m.s.</i> tipo de chip, parte principal do computador, também conhecido como CPU (Central Processing Unit); executa instruções (processos) e dados (através da UC - unidade de controle) e operações lógicas e aritméticas (através da UAL - unidade aritmética e lógica).. Ex.: Processador (UCP). A UCP é a parte principal do computador responsável pelo processamento e execução de programas armazenados na memória principal.<i>Sinônimos:</i> CPU, UCP. <i>Veja Também:</i> CPU, UCP.</p> <p style="background-color: #0070C0; color: white; padding: 2px;">▼ English</p> <p>Go back to search results</p> <p>processor. <i>Hardware. n.m/f.s.</i> flat square sliver of silicon device composed of circuits etched on its surface; it controls the computer and execute programs; also known as CPU (central processing unit). Ex.: The brain of any computer is the central processing unit (CPU). This component does all the calculations and performs 90 percent of all the functions of a computer. There are many different types of processors for computers so many, in fact, that you will learn about them later in this chapter in the section Identifying Purposes and Characteristic of Processors. <i>Synonyms:</i> CPU. <i>See Also:</i> CPU.</p>
---	---

23/08/2007 15:44 © 2007 Guilherme Fromm - ICMC Jr.

2. Exibição descritiva: os novos dicionários, em formato eletrônico, disponibilizam essa forma de consulta onde cada item da descrição do verbete aparece em linhas diferentes. Note, neste exemplo, que nem todos os campos estão preenchidos. Isso se deve às informações que conseguimos levantar no *corpus* de estudo para preencher os dados.

Em Português

▼ Português

[Voltar ao resultado da busca](#)

processador

tipo de chip, parte principal do computador, também conhecido como CPU (Central Processing Unit); executa instruções (processos) e dados (através da UC - unidade de controle) e operações lógicas e aritméticas (através da UAL - unidade aritmética e lógica).

Abreviação/Acrônimo: CPU, UCP

Categoria Gramatical: substantivo

Gênero: masculino

Número: singular

Córpus

- Posição na Ordem de Frequência: 67
- Nº de Ocorrências do termo: 1212

Ontologia: [Computação](#) > [Hardware](#)

Variações Morfossintáticas: Nada encontrado

Sinônimos: CPU, UCP

Hipônimos de: computador

Co-hipônimos: chip

Hiperônimos de: UAL (unidade aritmética e lógica); UC (unidade de controle)

Exemplos

1. Processador (UCP). A UCP é a parte principal do computador responsável pelo processamento e execução de programas armazenados na memória principal.
2. UCP

Sua função consiste em coordenar, controlar e realizar todas as operações (execução de instruções) do sistema.

3. UCP

Suas partes principais são as seguintes:

- a unidade aritmética e lógica (UAL): realiza o cálculo real ou o processamento de dados (realiza as operações aritméticas e lógicas).

- a unidade de controle (UC): controla o movimento de dados e instruções dentro e fora da UCP e controla a operação da UAL, de forma adequada e sincronizada.

[3 mais...](#)

Informações Enciclopédicas



O processador é a parte mais fundamental para o funcionamento de um computador. Processadores são circuitos digitais que realizam operações como: cópia de dados, acesso a memórias e operações lógicas e matemáticas. Em: *Processador* - [Wikipedia](#)

Veja Também

[CPU](#) - [UCP](#)

Em Inglês

▼ English

[Go back to search results](#)

processor

flat square sliver of silicon device composed of circuits etched on its surface; it controls the computer and execute programs; also known as CPU (central processing unit)

Acronym: *Nothing found*

Part of Speech: noun

Gender: neutral

Number: singular

Corpus

- **Frequency order position:** 138
- **Term number of occurrences:** 772

Ontology: [Computing](#) > [Hardware](#)

Morphosyntatic Variants: *Nothing found*

Synonyms: CPU

Antonyms: *Nothing found*

Synonyms: CPU

Antonyms: *Nothing found*

Hyponyms of: computer

Co-hyponyms: *Nothing found*


Hypernyms of: *Nothing found*

Examples

1. The "brain" of any computer is the central processing unit (CPU). This component does all the calculations and performs 90 percent of all the functions of a computer. There are many different types of processors for computers so many, in fact, that you will learn about them later in this chapter in the section "Identifying Purposes and Characteristic of Processors."
2. In essence, a CPU is a flat square sliver of silicon with circuits etched on its surface. This chip is linked to connector pins and the whole contraption encased some form of packaging - either ceramic or plastic - with pins running along the flat underside or along one edge. The CPU package is connected to a motherboard via some form of CPU interface, either a slot or a socket
3. The speed of the processor is generally described in clock frequency (MHz or GHz).

[1 more...](#)

Encyclopedic Information

 A central processing unit (CPU), or sometimes simply processor, is the component in a digital computer that interprets computer program instructions and processes data. CPUs provide the fundamental digital computer trait of programmability, and are one of em: CPU - [Wikipedia](#)

See Also
[CPU](#)

23/08/2007 15:44 © 2007 Guilherme Fromm - ICMC Jr.

OBS.: as informações enciclopédicas não são exibidas no modo normal.

Consultas internas: total, tradutor e modular

Os tipos de consulta foram divididos em três:

Total: apresenta todos os campos disponíveis (veja exemplo acima).

Tradutor: apresenta os campos considerados mais importantes pelos tradutores profissionais (note que nem todos os campos do exemplo anterior aparecem aqui). Veja o exemplo:

Tipos de Exibição	▼ Português
Normal	Voltar ao resultado da busca
Descritiva	processador
Tipos de Consulta	tipo de chip, parte principal do computador, também conhecido como CPU (Central Processing Unit); executa instruções (processos) e dados (através da UC - unidade de controle) e operações lógicas e aritméticas (através da UAL - unidade aritmética e lógica).
Total	Abreviação/Acrônimo: CPU, UCP
Tradutor	Categoria Gramatical: substantivo
Modular	Gênero: masculino
Consultas Externas	Número: singular
Corpus NILC	Ontologia: Computação > Hardware
Google	Variações Morfossintáticas: <i>Nada encontrado</i>
Answers.com	Sinônimos: CPU, UCP
Wikipedia	Antônimos: <i>Nada encontrado</i>
CORTEC	Exemplos

	<ol style="list-style-type: none"> 1. Processador (UCP). A UCP é a parte principal do computador responsável pelo processamento e execução de programas armazenados na memória principal. 2. UCP <p>Sua função consiste em coordenar, controlar e realizar todas as operações (execução de instruções) do sistema.</p> 3. UCP <p>Suas partes principais são as seguintes:</p> <ul style="list-style-type: none"> - a unidade aritmética e lógica (UAL): realiza o cálculo real ou o processamento de dados (realiza as operações aritméticas e lógicas). - a unidade de controle (UC): controla o movimento de dados e instruções dentro e fora da UCP e controla a operação da UAL, de forma adequada e sincronizada. <p>3 mais...</p> <p>Veja Também CPU - UCP</p>
--	---

Modular: nessa consulta, você escolhe os campos que deseja visualizar. Veja o exemplo abaixo. Nele selecionamos o tipo de exibição descritiva e excluimos da seleção (ao clicar nos campos em azul do lado esquerdo, eles ficam cinzas) os campos gênero, variações morfossintáticas, forma por extenso, *corpus*, exemplos, informações enciclopédicas e remissiva.

<p>Tipos de Exibição Normal Descritiva</p> <hr/> <p>Tipos de Consulta Total Tradutor Modular</p> <hr/> <p>Pré-definição Abreviatura Categoria Gramatical Gênero Número Ontologia Variações Morfossintáticas</p> <hr/> <p>Definição Definição Forma por Extenso Corpus</p> <hr/> <p>Exemplificação Exemplos Informações Enciclopédicas</p> <hr/> <p>Relações Remissiva Sinônimos Antônimos Hipônimos Co-hipônimos Hiperônimos</p>	<p>▼ Português</p> <p>Voltar ao resultado da busca</p> <p>processador</p> <p>tipo de chip, parte principal do computador, também conhecido como CPU (Central Processing Unit); executa instruções (processos) e dados (através da UC - unidade de controle) e operações lógicas e aritméticas (através da UAL - unidade aritmética e lógica).</p> <p>Abreviação/Acrônimo: CPU, UCP</p> <p>Categoria Gramatical: substantivo</p> <p>Número: singular</p> <p>Ontologia: Computação > Hardware</p> <p>Sinônimos: CPU, UCP</p> <p>Antônimos: <i>Nada encontrado</i></p> <p>Hipônimos de: computador</p> <p>Co-hipônimos: chip</p> <p>Hiperônimos de: UAL (unidade aritmética e lógica); UC (unidade de controle)</p> <p>▼ English</p> <p>Go back to search results</p> <p>processor</p> <p>flat square sliver of silicon device composed of circuits etched on its surface; it controls the computer and execute programs; also known as CPU (central processing unit)</p> <p>Acronym: <i>Nothing found</i></p> <p>Part of Speech: noun</p> <p>Number: singular</p>
---	--

Clique nas possibilidades do lado esquerdo da tela e monte a sua estrutura de acordo com as suas necessidades!

Consultas externas

Além dos diversos tipos de consultas internas disponíveis no vocabulário, você também pode fazer consultas externas a sites diversos. Segue o resumo de cada um deles; basta clicar em cada um, esperar abrir a caixa de inserção, inserir o termo desejado e clicar <enter> no seu teclado:

The screenshot shows a web interface for the word "processor". On the left, there are sections for "Relações" (Remissiva, Sinônimos, Antônimos, Hipônimos, Co-hipônimos, Hiperônimos) and "Consultas Externas" (Corpus NILC, Google, Answers.com, Wikipedia). Each external search option has a text input field and a "Buscar" button. At the bottom left, there is a "CORTEC" link. On the right, the word "processor" is defined as a "flat square sliver of silicon device composed of circuits etched on its surface; it controls the computer and execute programs; also known as CPU (central processing unit)". Below the definition, there are fields for "Acronym", "Part of Speech", "Number", "Ontology", "Synonyms", "Antonyms", "Hyponyms of", "Co-hyponyms", and "Hypernyms of".

Corpus NILC: o termo que for digitado aqui será buscado no *corpus* técnico do projeto NILC (Núcleo Interinstitucional de Linguística Computacional) e será apresentado numa estrutura KWIC (o termo aparecerá no meio da página).

Google: o termo será buscado no mais famoso site de busca da Internet, em português.

Answers.com: o termo será buscado no site, misto de dicionário e enciclopédia, em inglês.

Wikipedia: o termo será buscado na mais famosa enciclopédia online, em português.

CORTEC: site que apresenta vários *corpora* técnicos. Lá você pode fazer buscas por listagens de palavras, concordanciador, etc.. Ao clicar neste link, uma nova página será aberta.

BOM DIVERTIMENTO!

Guilherme Fromm

guifromm@uol.com.br

OBS.: se você tem alguma dúvida sobre termos usados para descrever as palavras (como *Hiperônimo*, por exemplo), consulte a área de Linguística.

7.2 O aprendiz avançado: usando o banco de dados

O uso da página de consulta por aprendizes pode ser um primeiro passo para o aprendizado de Terminologia, mas não é o único.

Cabré (1999, *apud* ARAÚJO, 2001, p.207)³⁹ propõe uma classificação para o grau de contato do tradutor com a terminologia. O primeiro nível seria o *passivo*, em que o profissional apenas consulta fontes terminológicas, mas precisa ter um mínimo de formação para saber como usá-las (o que propõe esta tese através da página do VoTec); no segundo nível, o tradutor passa a ser *minimamente ativo* e começa a compilar termos e criar bases de dados (o que pode ser alcançado pela manipulação do banco de dados). Krieger (2006, p. 203), seguindo a mesma linha de raciocínio, ao comentar sobre o ensino de Terminologia e sua aplicação por tradutores, afirma que um importante passo para esses profissionais é a produção de glossários, dicionários técnicos e bancos de dados terminológicos.

Há uma grande oportunidade, aqui, para que alunos mais avançados, que já dominem o uso de ferramentas para tradutores (minimamente ativos), como a desenvolvida para esta tese, partam para um segundo nível de aprendizado que vai ao encontro da idéia das autoras: a elaboração de terminologia através da construção de vocabulários técnicos. Num curso de Terminologia para aprendizes de tradução, por exemplo, todas as fases apresentadas anteriormente nesta tese podem ser trabalhadas por professores (que dominem todas as etapas) em sala de aula, passo a passo:

³⁹ CABRÉ, M. T. **La Terminologia** – representación e comunicación. Barcelona: IULA/UFP, 1999.

1. escolha de uma área que o aluno domine ou com que tenha afinidade, pois é essencial que o mesmo se sinta familiarizado com os textos e vocabulário que vai trabalhar nesse primeiro momento;
2. a construção da árvore de campo, que envolve uma pesquisa de toda a área e com profissionais dispostos a ajudar. Essa construção pode delimitar domínios ou subáreas que apresentam maiores problemas e que necessitem a elaboração de uma terminologia própria;
3. a seleção do(s) *corpus* (*corpora*), pois, partindo de exemplos reais, o aluno consegue encontrar, facilmente, exemplos que sirvam de base para a construção de uma terminologia. Há a necessidade, neste caso, do aluno dispor de um computador com acesso à Internet e espaço de armazenamento interno para dar conta da tarefa (conforme mencionado anteriormente, o banco armazena somente os contextos selecionados a partir do *corpus*, e não o *corpus* em si);
4. a extração dos termos a partir desses *corpora* implica o aprendizado de uma ou várias ferramentas de análise lexical que podem despertar no aluno a vontade de fazer pesquisas mais aprofundadas; como no caso anterior, o banco de dados não possui essa facilidade; há, portanto, a necessidade dessa(s) ferramenta(s) estar(em) instalada(s) no computador do aluno;
5. o uso do banco de dados em si: a inscrição do aluno (por parte do administrador) no banco de dados como pesquisador, a visita a todas as telas com a descrição de seu funcionamento, a inserção, para cada termo, dos excertos de textos retirados dos *corpora*, a construção dos paradigmas e da definição, a visualização final no VoTec.

Todos esses passos podem ser orientados por um professor que tenha passado por um treinamento adequado, pois, como já citado por Araújo (2002), muitos não têm noções básicas de Terminologia. Esse treinamento pode ser elaborado partindo dos dados apresentados nesta tese. A grande vantagem, no caso, seria a constante expansão do banco através do trabalho colaborativo entre pesquisadores, professores e alunos.

8. Avaliação da ferramenta

Após o desenvolvimento do banco de dados e da página de consulta, foi feita uma pesquisa com o público-alvo da tese: os aprendizes de tradução. Assim como apresentado no capítulo 5, vou descrever o que me motivou a fazer a pergunta e tabular os resultados para verificar a eficácia da proposta. Essa pesquisa foi feita por meio de um questionário respondido em papel (a versão completa da folha de pesquisa encontra-se no apêndice G), distribuído para 62 alunos de três cursos de Tradução (bacharelado) na região metropolitana de São Paulo. Para o exercício da tradução, um texto real, retirado da Internet (disponível em: <http://www.geocities.com/ferds3/bus_structures.htm>, acessado em: 21/07/2007) foi disponibilizado aos alunos:

There are several requirements and standards associated with PnP. The BIOS, operating system, motherboard, chipset, and device must be PnP compatible. Windows 95 and 98 are fully PnP compatible, where Windows NT is only partially compatible. All PCI and AGP devices are PnP, but ISA devices aren't necessarily and VESA Local Bus devices never were. Remember that all PCI cards are PnP, but not all PnP cards are PCI.

O objetivo era fazer com que achessem os termos no VoTec, traduzissem o texto e depois refletissem, a partir do questionário, sobre o processo. Ao contrário da pesquisa anterior, totalmente informatizada, os dados aqui descritos foram tabulados manualmente. Junto com a pesquisa dos alunos, foi distribuída uma mini-pesquisa aos professores. É extremamente importante notar que essa pesquisa com os alunos foi conduzida sem prévia explicação, por parte do professor que a aplicou, sobre o funcionamento do site. Também é importante notar, como aconteceu na pesquisa

anterior, que os gráficos apresentam divergências quanto ao total de respondentes. Isso se dá porque alguns alunos simplesmente não responderam a uma ou outra pergunta.

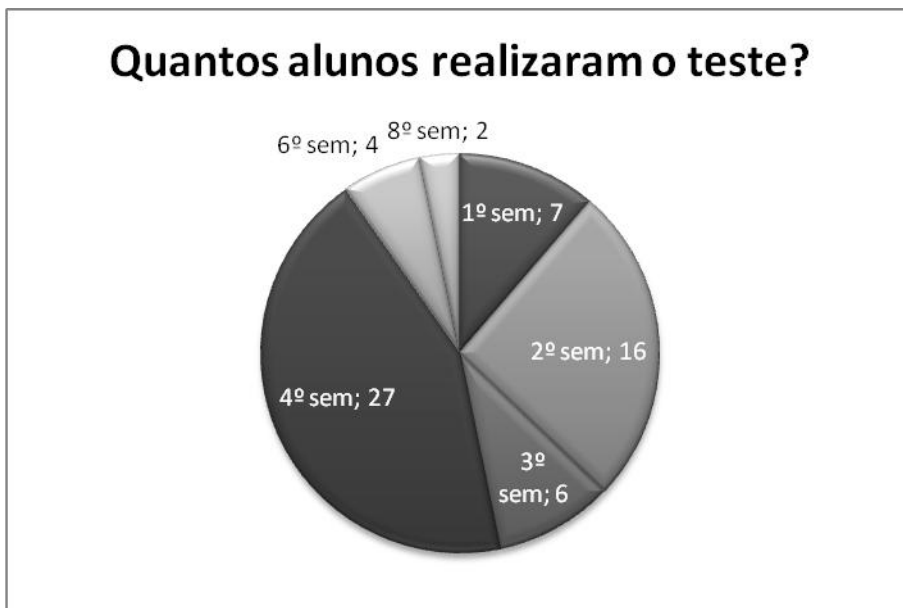
8.1 As perguntas, suas motivações e a relevância para a tese

São analisadas, a seguir, as perguntas apresentadas aos alunos, a motivação para a sua criação, os gráficos com os resultados obtidos e um comentário sobre a relevância desses resultados para a tese. Toda a tabulação dos resultados foi criada numa planilha do Excel.

1. Há quantos semestres você cursa Tradução?

Motivação: o objetivo, aqui, é separar as respostas dos alunos por semestre cursado. Tenta-se descobrir possíveis diferenças entre o grau de aprendizado e a compreensão/uso do site. Essa pergunta gerou a criação de oito pastas na planilha (já que os cursos pesquisados, todos de graduação, apresentam, no máximo, oito semestres), uma para cada semestre, mais uma planilha totalizadora, com os resultados gerais. São apresentados, a partir da pergunta dois, somente os gráficos com os dados totalizados; os demais podem ser conferidos no apêndice H ou na planilha inserida no CDROM que acompanha a tese. No comentário sobre a relevância para a tese, porém, sempre que necessários, todos os dados das seis planilhas preenchidas serão analisados.

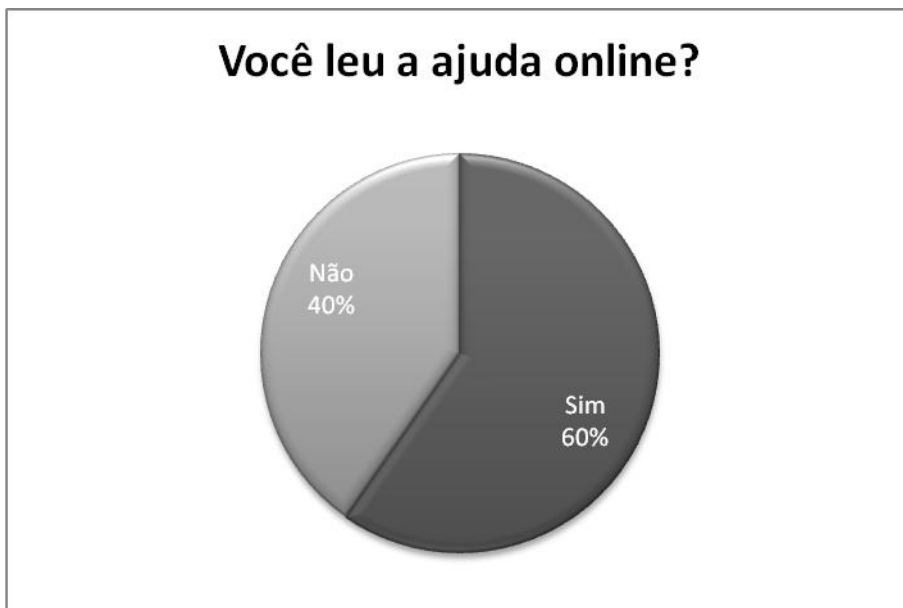
Resultado:



Relevância para a tese: não participaram da pesquisa alunos dos quinto e sétimo semestres; embora o número ideal de alunos devesse se concentrar nos dois primeiros semestres, já que representam os aprendizes com menos experiência, nota-se que a grande concentração de respostas situou-se entre alunos do segundo e quarto semestres.

2. Você leu a ajuda online antes de trabalhar no site?

Motivação: descobrir o quanto a ajuda online (descrita no capítulo anterior) é necessária para os aprendizes. Embora muitos estejam começando seus estudo de Tradução, a experiência individual no uso da Internet pode tornar essa ajuda desnecessária.

Resultado:

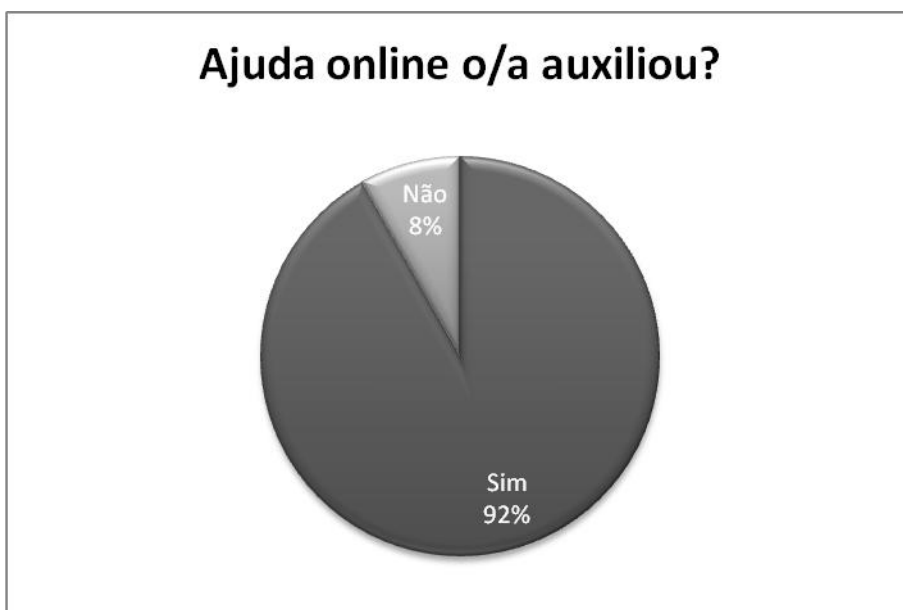
Relevância para a tese: a maioria leu a ajuda online. A maior concentração de leitores da mesma se dá no terceiro e quarto semestres. Como o objetivo da tese é oferecer uma ferramenta que, dentro do possível, seja auto-explicativa, o número considerável de consulentes que visitaram a página de ajuda indica que os mesmos ainda não têm muito contato com este tipo de tecnologia⁴⁰, porém se mostraram dispostos a ler a página, o que indica uma certa curiosidade sobre o funcionamento da mesma e a utilidade que ela possa vir a ter. É interessante notar também que, em termos de porcentagem, há mais alunos entre o primeiro e segundo semestres (respectivamente 40% e 50%) que não leram a ajuda do que nos outros semestres. Respostas da pergunta quatro indicam familiaridade desses alunos com o uso do computador.

⁴⁰ A pergunta 4 retoma a questão e explica como o contato prévio com páginas semelhantes é essencial para aqueles que responderam de forma negativa esta questão.

3. A ajuda online auxiliou você a trabalhar com o site?

Motivação: verificar se aqueles que usaram a ajuda, apontados pela pergunta anterior, obtiveram uma orientação correta de como funciona o site

Resultado:



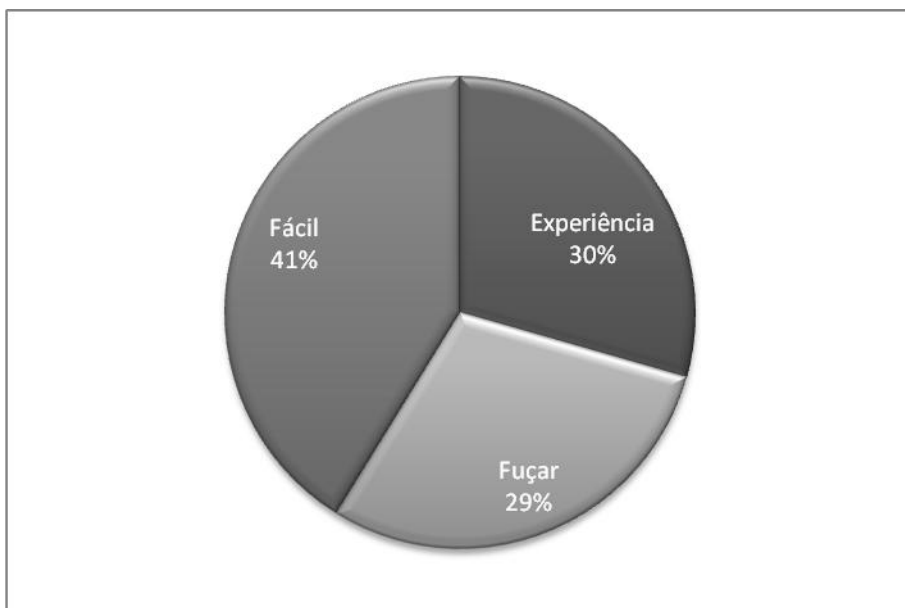
Relevância para a tese: indica que a informação levantada para a ajuda online, na opinião da maioria dos consulentes (inclusive alunos do primeiro e segundo semestres), é adequada para travarem um primeiro contato com uma ferramenta informatizada de Terminologia.

4. Se você não usou a ajuda online para trabalhar com o site, é por que:

Motivação: descobrir quais os pontos que mais se destacam para o não uso da ajuda online. Essa questão sugeria três possibilidades, por mim aventadas (e baseadas na minha experiência como professor de prática de tradução), das quais o aluno deveria escolher somente uma. As motivações para essas respostas são:

- a. *você tem experiência com sites que oferecem dicionários online*; descobrir se o aluno já havia trabalhado com sites similares na Internet;
- b. *é mais rápido aprender “fuçando” do que lendo instruções*; alguns alunos já possuem uma familiaridade tão grande com o computador, seja pela experiência na área, seja pela experiência de vida, que dispensam o uso do auxílio escrito, assim como algumas pessoas dispensam o auxílio de um manual de instrução de um aparelho qualquer;
- c. *é fácil de aprender a usar o site*; mesmo sem experiência prévia, o site mostra-se de fácil manuseio e não há a necessidade de ler instruções.

Resultado:



Relevância para a tese: embora, na teoria, somente 25 alunos (aqueles que responderam *não* na pergunta 2) devessem responder a essa pergunta, 34 acabaram por fazê-lo. O gráfico indica que não há dificuldade em descobrir, por conta própria e sem ajuda, as facilidades que o site apresenta. Indica também que, por mais que os alunos sejam novatos no mundo da Tradução, eles não o são quanto ao uso da Internet, cada vez mais

popular no Brasil. As páginas da Internet repetem muitos padrões visuais e de consulta e essa questão foi levada em conta na elaboração do VoTec. É uma página “limpa”, sem muitas figuras, gráficos ou acessórios que desviem a atenção do foco em questão e segue a página do programa de busca Google, que é uma referência nesse campo.

5. Você já conhecia termos como *Hiperônimo*, *Hipônimo*, *Co-hipônimo*?

Motivação: os termos acima são usados no Paradigma Semântico da microestrutura de cada verbete. A idéia, aqui, é verificar se os alunos já trabalharam com os conceitos em algum momento no curso de Letras e/ou Tradução e se os mesmos são claros para os consulentes.

Resultado:



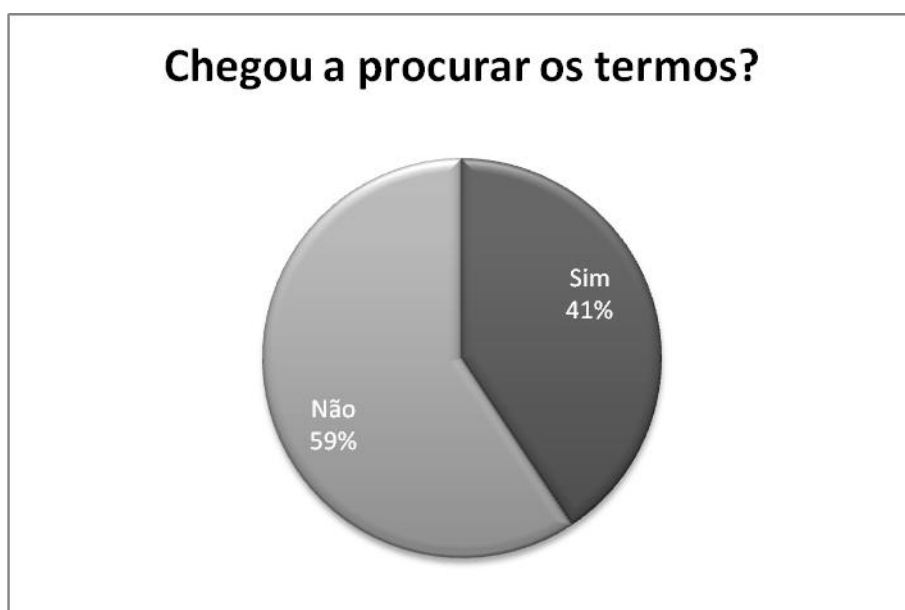
Relevância para a tese: poucos tiveram ou acreditam que tiveram contato com esses termos pertencentes à área de Linguística. Mesmo alunos do sexto (75%) e oitavo (50%)

semestres, que já passaram da metade do curso, não sabiam sobre os termos. Embora não tenham sido analisados os currículos dos cursos, a disciplina de Lingüística, normalmente, é uma das primeiras a ser ministrada. A ausência (ou o esquecimento, por parte dos alunos) do conhecimento de termos ligados ao Léxico sugere a pouca importância dedicada ao tema nesses cursos.

6. Se você não conhecia os termos citados na pergunta 5, você chegou a procurá-los no próprio site, na seção de Lingüística?

Motivação: o objetivo é verificar se os alunos fizeram uma relação, já proposta na ajuda online, entre as diferentes áreas disponíveis para análise no site (computação e lingüística) e perceberam que a solução para sua dúvida se encontrava no próprio site (configurando uma metalinguagem).

Resultado:

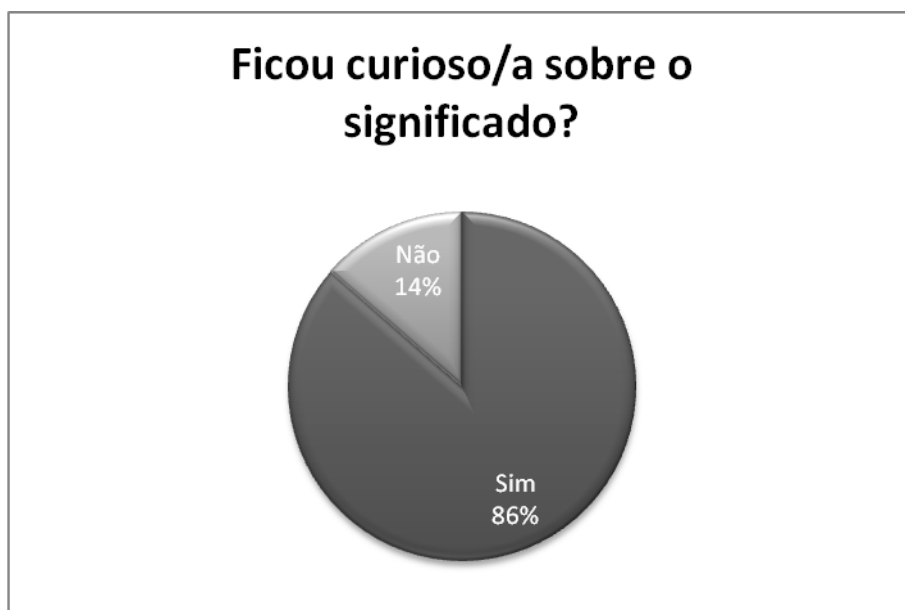


Relevância para a tese: percebe-se aqui que a não-necessidade de saber o significado dos termos que aparecem no Paradigma Semântico, para uma primeira compreensão do verbete, faz com que a maioria dos consulentes não se interessem pela sua busca. Muitos, por não terem lido a página de ajuda, também não perceberam que os termos poderiam ser buscados dentro da própria página.

7. Se não conhecia os termos da questão 5, você chegou a ficar curioso sobre o significado deles?

Motivação: descobrir se o uso de termos técnicos da área de Linguística na microestrutura suscita algum tipo de dúvida quanto ao seu significado ou se esse uso passa despercebido pelos usuários.

Resultado:

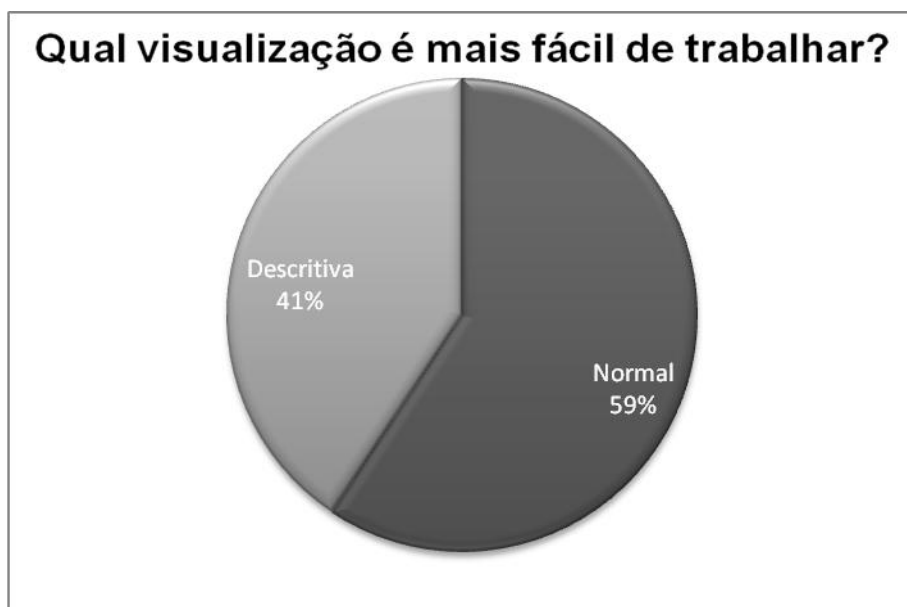


Relevância para a tese: não só a necessidade, mas também a curiosidade leva à consulta nas obras lexicográficas/terminográficas. Parece que o significado até despertou a curiosidade dos alunos sobre os termos que aparecem na microestrutura, conforme a pergunta anterior, mas eles não se propuseram a buscá-los, seja na própria página, seja em outra fonte. Há uma evidente contradição entre os resultados desta pergunta e a anterior.

8. Qual tipo de visualização é mais fácil de trabalhar, descritiva ou normal?

Motivação: verificar se os alunos já estão acostumados a trabalhar com diferentes processos de visualização que não aquele encontrado nos vocabulários ou dicionários tradicionais em papel.

Resultado:



Relevância para a tese: embora o modo de visualização descritivo já faça parte de várias obras eletrônicas, muitas obras disponíveis na Internet ainda usam o padrão de visualização tradicional das obras impressas. Esse padrão, para a maioria dos alunos (59%) , ainda é o mais fácil de consultar. O número daqueles que preferem uma visualização mais estendida, por outro lado, não pode ser desconsiderado (41%). Notou-se uma diferença significativa entre os dados verificados no primeiro e oitavo semestres: 100% dos alunos do primeiro preferiram a visualização normal, enquanto 100% dos alunos do oitavo preferiram a visualização descritiva. Isso sugere diferentes graus de contato com materiais do tipo. Interessante destacar, também, a quantidade de consulentes que preferiram a visualização descritiva no segundo semestre (56%), diferenciado-se bastante dos semestres subsequentes (terceiro, quarto e sexto).

9. Dentre os três tipos de pesquisa (total, tradutor, modular), qual você achou mais fácil para trabalhar?

Motivação: verificar se os alunos se sentem mais à vontade para trabalhar com uma construção da microestrutura que é tradicional nos dicionários correntes, uma construção aventada pelos profissionais da área ou uma construção em que eles possam manusear a criação dessa microestrutura.

Resultado:

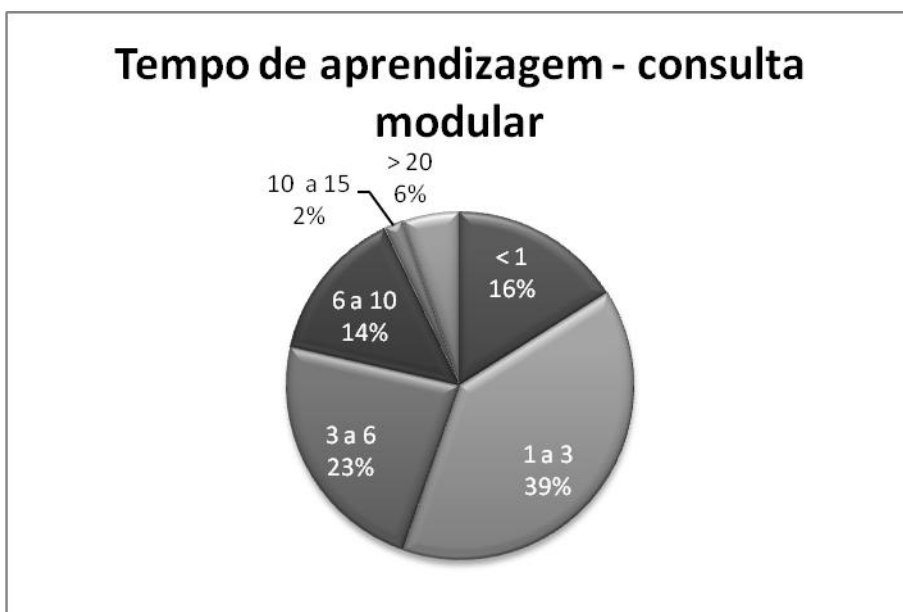


Relevância para a tese: a maioria dos alunos prefere, seguindo a tendência apresentada na questão anterior, seguir com o tipo de pesquisa que apresente uma microestrutura idêntica às obras que estão acostumados a pesquisar. Poucos alunos parecem ter entendido uma das propostas do site, que é facilitar a sua busca. Apenas alguns se identificaram com a microestrutura proposta pelos tradutores profissionais e somente três entenderam (quarto e sexto semestres) que a consulta modular pode moldar a pesquisa às suas necessidades.

10. Quanto tempo você demorou para entender como funciona o tipo de consulta “modular”?

Motivação: verificar o tempo de aprendizagem para lidar com um novo tipo de microestrutura, criada pelos próprios usuários.

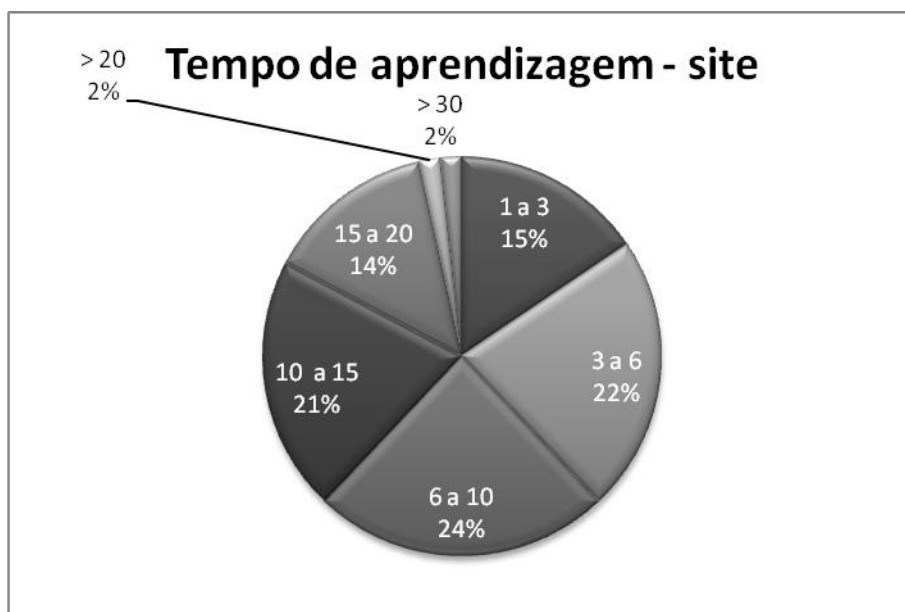
Resultado:



Relevância para a tese: embora não tenham escolhido a consulta modular como a sua favorita, a maioria dos aprendizes não levou mais do que 10 minutos para completar a curva de aprendizagem (tempo necessário para o aprendizado de um programa) desta parte da página. Isso prova que o manuseio da nova proposta apresentada não se mostrou de difícil compreensão. Apenas um aluno demorou entre 10 e 15 minutos (sexto semestre) para aprender como a página funciona e três alunos (um do terceiro e dois do quarto semestres) demoraram mais de 20 minutos. Por outro lado, surpreende a quantidade de alunos que entenderam o funcionamento da proposta em menos de um minuto (três alunos do quarto semestre e um do sexto).

11. Quanto tempo você demorou para entender como funciona o site como um todo?

Motivação: verificar o quanto o usuário, sem a ajuda do professor ou monitor, demorou para aprender a entender e manusear o site.

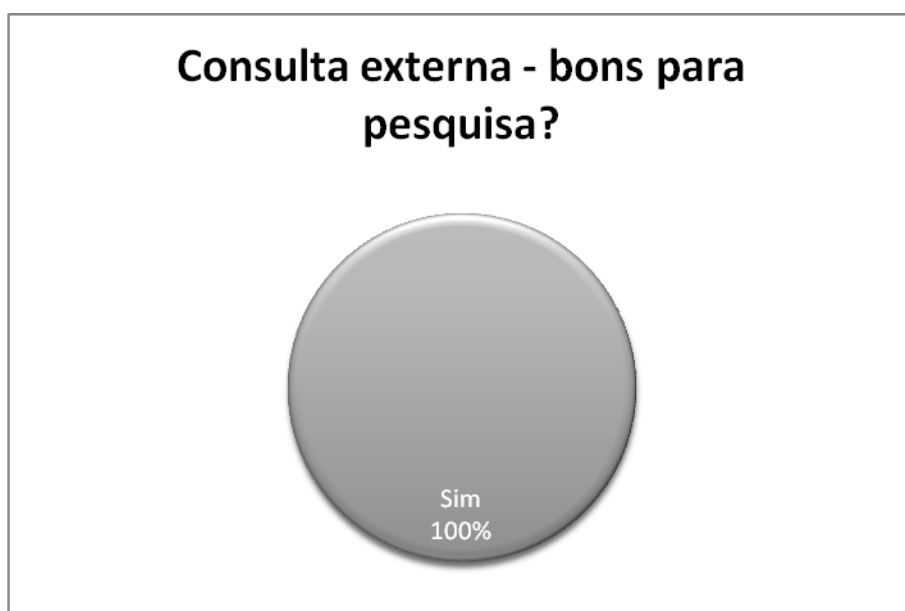
Resultado:

Relevância para a tese: a curva de aprendizado, para a maioria dos alunos, se deu em até 15 minutos. Como uma possível nova ferramenta e considerando a não interferência por parte do docente, o VoTec apresenta um tempo de aprendizado curto. Dos que levaram mais de 15 minutos, um está no segundo semestre (15 a 20 minutos) e nove no quarto semestre (sete levaram entre 15 e 20 minutos, um levou mais de 20 minutos e outro mais de 30 minutos).

12. Você considera os sites de consulta externa (lado esquerdo inferior) bons para pesquisas extras?

Motivação: aferir se minhas escolhas, na composição de quais dos sites que deveriam constar como *links*, foram acertadas ou não para o profissional ligado à Tradução. A pergunta considera todos os sites como um único bloco.

Resultado:

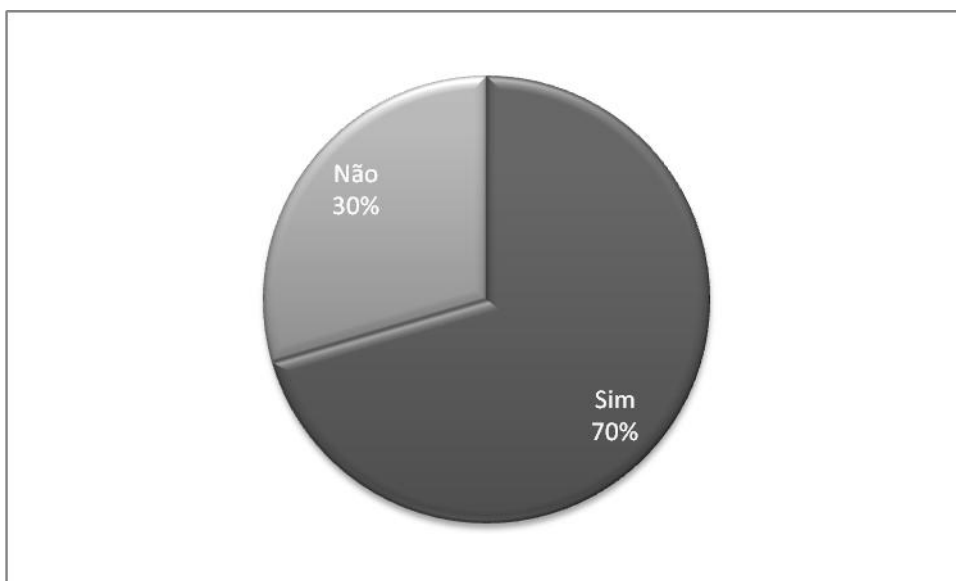


Relevância para a tese: unanimidade entre os respondentes, os *links* para consultas externas se mostraram uma decisão acertada, especialmente porque indicam páginas que podem aumentar o conhecimento do alunos sobre o uso de ferramentas informatizadas.

13. Os termos técnicos na área de computação disponíveis no site foram suficientes para a sua tradução?

Motivação: verificar se o aluno conseguiu consultar todos os termos técnicos, ligados à área de computação, que constavam do texto a ser traduzido.

Resultado:



Relevância para a tese: embora todos⁴¹ os termos técnicos do campo de computação presentes no texto constassem no banco de dados, muitos não conseguiram encontrá-los no VoTec, sugerindo dificuldade na digitação dos mesmos na página de consulta ou pouca familiaridade com páginas deste tipo. Muitos alunos relatam também, conforme os comentários expostos na questão 17, problemas de visualização por causa da fonte adotada, o que pode ter contribuído para que 30% deles não achassem todos os termos. Destacam-se, também, os alunos do terceiro semestre, onde 50% deles não encontraram as palavras.

14. Quais palavras constavam do texto a ser traduzido e não foram encontradas no próprio site?

⁴¹ Por erro de digitação, o termo *barramento* não contava com a definição. No exemplo apresentado, contudo, havia um contexto definatório.

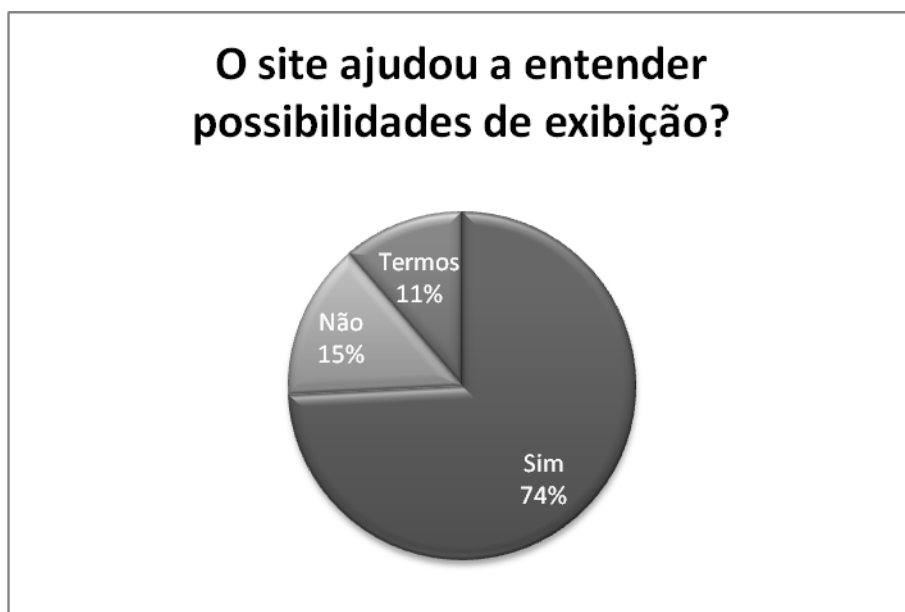
Motivação: verificar se os alunos indicariam apenas palavras não-técnicas da área de computação, não disponíveis para consulta, ou se indicariam todas as palavras com as quais tiveram dificuldade no ato da tradução.

Resultado: entre as várias respostas obtidas, destacam-se *chipset, PnP, AGP, BIOS, VESA, PCI*.

Relevância para a tese: embora todos os termos acima citados estivessem disponíveis para consulta, muitos alunos não os encontraram; vários outros comentários também indicam que os consulentes procuraram termos não-técnicos (como *standard, device, requirement*), o que pode indicar o não entendimento da proposta da página.

15. O site o ajudou a entender as diversas possibilidades de exibição de um termo técnico?

Motivação: fazer com que o aluno reflita sobre o aprendizado que o site proporcionou quanto aos possíveis modos de exibição da microestrutura numa obra terminológica.

Resultado:

Relevância para a tese: a maioria percebeu que há mais de uma maneira de exibir a microestrutura de um termo técnico. Os alunos podem ter preferido uma forma em detrimento de outras, mas conseguiram notar as diferenças. Entre as respostas citadas na opção *em termos*, temos:

- a. não dá o significado dos termos usados na definição, como barramentos;
- b. muitos termos queriam dizer a mesma coisa, o que confunde qualquer pessoa leiga no assunto;
- c. pouco prático, um pouco confuso;
- d. muita demora para resposta da pesquisa;
- e. é difícil achar a palavra;
- f. é necessário em pouco de conhecimento na área;

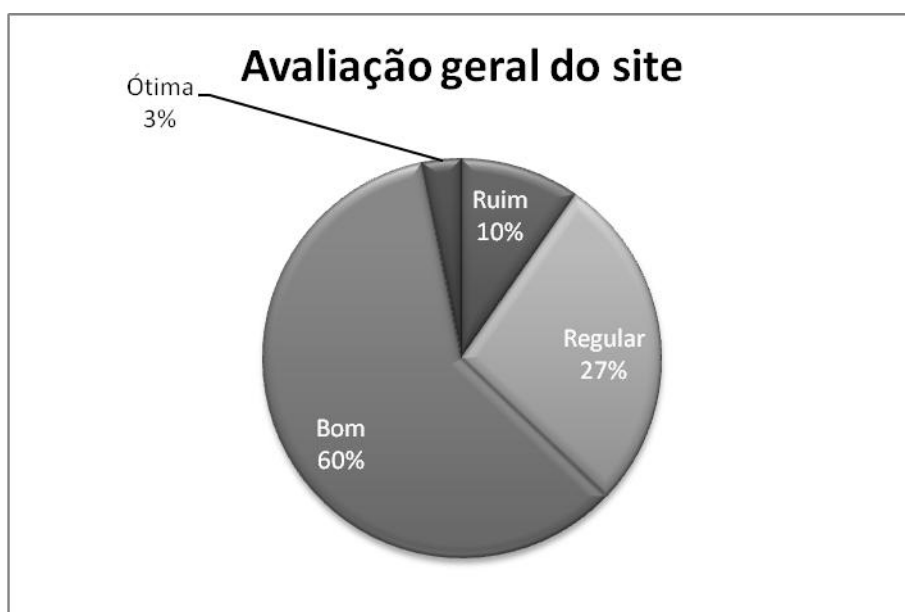
Ao analisar esses comentários, percebe-se que alguns alunos tiveram dificuldades com o manuseio da página (como em *a* e *e*), outros não notaram que muitos verbetes funcionam como remissiva para outro (*b*). O computador que estava sendo usado para o teste ou a conexão com a Internet também pode ter influenciado no resultado (como em *d*). Alguns alunos podiam se sentir despreparados para trabalhar em

áreas que não dominam (*f*) e outros, simplesmente, podem não ter gostado da página como um todo (*c*).

16. De modo geral, como você avalia o site tendo em vista a ajuda que oferece ao aprendiz de tradução?

Motivação: compreender se a proposta foi bem aceita pelos consulentes ou não. Havia uma escala de quatro valores (ruim, regular, bom, ótimo) disponível para escolha.

Resultado:



Relevância para a tese: com 63% de aprovação (bom e ótimo), é possível considerar que a proposta foi bem entendida e recebida pelos consulentes. As respostas negativas, obviamente, fornecem pistas sobre o que pode ser melhorado na página e também indicam algumas deficiências, por parte dos alunos. Essas deficiências podem,

futuramente, render um trabalho que vise melhorar os currículos dos cursos pesquisados.

17. Gostaria de tecer algum comentário sobre o site?

Motivação: obter respostas diretas dos alunos quanto às suas impressões sobre a página e/ou sugestões para o seu aprimoramento.

Resultado: apresento alguns dos vários comentários recebidos.

- a. Acredito que o site seria melhor se não houvessem [sic] tantos filtros de pesquisa , pois para quem não entende muito de determinados termos técnicos leva certo tempo para efetuar a pesquisa;
- b. as palavras estão muito ‘escondidas’;
- c. os termos técnicos são bons, mas a parte de tradução deixa a desejar;
- d. confuso, no início, estamos ‘viciados’ a pesquisar no Google;
- e. o layout é um pouco complicado. Poderia ter um pouco mais de cor. O conteúdo é bom;
- f. o site é muito bom, o campo para pesquisas extras é ótimo, pois todo tradutor precisa de mais de uma fonte;
- g. seria interessante se as palavras fossem colocadas juntamente com os contextos da área para melhorar a explicação do termo;
- h. às vezes as informações do site não são suficientes. Tive que fazer pesquisas extras para ajudar no ato tradutório;
- i. o site é muito confuso e não me ajudou a esclarecer todos os termos;
- j. para quem não é conhecedor de computação é impossível realizar uma tradução correta/legível com base nesse site;
- k. a fonte utilizada é de difícil visualização, dificultando a utilização do mesmo;
- l. ferramenta que pode auxiliar bem, já que tem o que é necessário para pesquisa num só lugar;
- m. a fonte é muito ruim, de difícil visualização. Faltou ‘exemplos’ para PnP;
- n. muito interessante como proposta para ajudar o tradutor;
- o. eu ainda não entendi como funciona, pode ser que este computador não ajude... mas como entendi que o texto trata da compatibilidade de hardware e software em um computador, tentei limitar a busca primeiro, colocando computação – hardware – componentes internos – placa mãe, ou seja a tradução já estava lá no caminho

percorrido de motherboard. Mas tentei obter algum resultado na pesquisa dessa palavra e não cheguei a lugar algum. Esse site de busca deveria ser todo reformulado, começando pelo tipo de letra, cansativa, difícil de ler, chega a dar dor de cabeça. Outra: depois que eu faço a escolha pelo campo de pesquisa e limito a busca utilizando os outros campos, eu não consigo voltar atrás e buscar apenas por computação por exemplo, sem ter que fechar e abrir o site novamente. Ah, acabei de ver algumas explicações para os termos que existem na busca e a explicação é bem legal, mas não encontro nenhum outro termo como as siglas que estão no texto;

p. a visualização não é de fácil compreensão;

q. não gostei da estrutura do site o formato das letras torna difícil a visualização;

r. ainda estou muito no início da tradução, então foi difícil, iria precisar de muito mais tempo para entender os ícones [sic] do site como um todo e o seu funcionamento;

s. a fonte escolhida é cansativa para leitura. Um leigo na área de info teria problemas na hora de procurar termos específicos, uma vez que escolhida a opção 'computação' surge outra dentre 'hardware' ..., e outra e mais outra... se eu não sei do que se tratam [sic] fica difícil especificar;

t. sim, a fonte do site de apoio pode causar confusão (confundir c com e);

u. o site é bom sim, mas tive pouco tempo para explorar e conhecer melhor;

v. talvez outra fonte. As letras atrapalham um pouco. Poderia ser Times Roman ou Verdana;

x. A parte de português da busca parece estar traduzida pois foram usados muitos termos que nunca havia ouvido falar, como 'barramento' que seria simplesmente 'slot' mas ao pé da letra;

y. achei interessante a construção do site relacionada à distribuição de significados por área de pesquisa (computação, etc.).

Relevância para a tese: esses comentários indicam diversas melhorias que podem ser incluídas numa próxima versão da página. Podemos dividi-los, basicamente, em cinco categorias;

a. Elogios (*f, l, n, u, y*): embora demonstrem uma boa aceitação por parte dos consulentes, não fornecem detalhes para a melhora da página;

b. visualização (*b, k, m, p, q, s, t, v*): aparece como um problema recorrente em vários comentários, sugerindo que a fonte usada (Trebuchet) pode não estar instalada em todos os computadores. Embora essa fonte esteja associada aos navegadores mais novos e com eles seja instalada, navegadores mais antigos apresentam problemas na visualização, alterando o leiaute da página e dificultando seu entendimento;

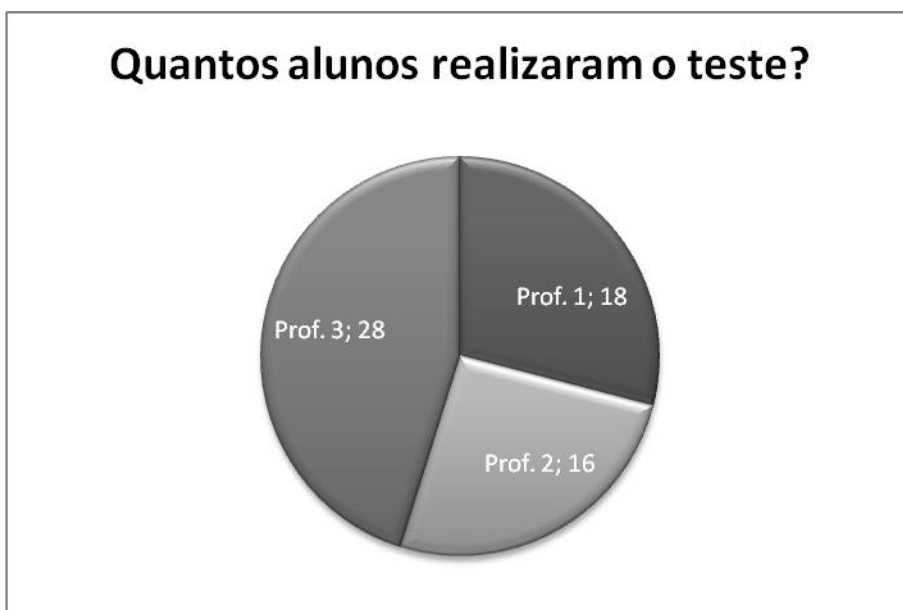
- c. insatisfação com a qualidade do material apresentado (c, h, j, x): representa uma opinião dos alunos;
- d. questões de leiaute (*d, e, i, r*): indicam sugestões para uma futura reformulação no visual da página;
- e. questões de construção da microestrutura (*g*): embora a consulta modular resolva parcialmente esse problema, é uma opção a ser analisada (alguns dicionários impressos apresentam esse tipo de microestrutura);
- f. menus *drop-down* (*a, o, s*): a dificuldade encontrada por esses alunos (dois, por sinal, não leram a ajuda disponível no site), indica que a questão dos menus que vão surgindo pode atrapalhar a consulta para alguns. Acredito que uma melhor descrição do uso desse recurso na ajuda online, numa versão futura do site, possa eliminar o problema;

8.2 A mini-pesquisa com os professores

1. Quantos alunos realizaram o teste?

Motivação: apenas para confirmar se o número de folhas recebidas estava em conformidade com o número de alunos presentes na sala.

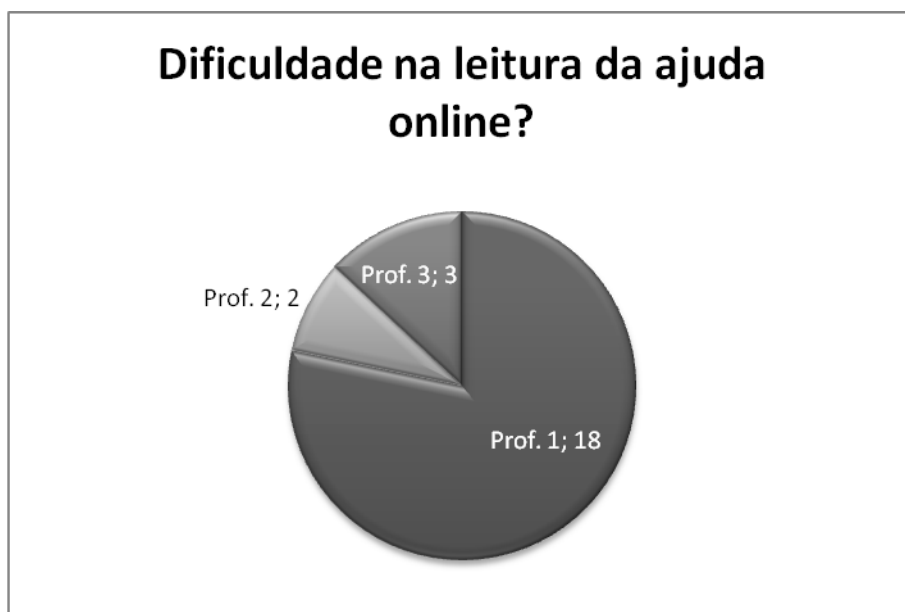
Resultado:



2. Algum aluno encontrou dificuldade na leitura da ajuda online, ou seja, ele/ela lhe fez alguma pergunta? Se sim, quantos alunos?

Motivação: verificar se, apesar da indicação expressa (lida pelo professor antes do início da pesquisa) de não solicitar ajuda ao professor, algum aluno fez alguma pergunta e/ou expôs algum tipo de dificuldade relacionada ao site.

Resultado:



Relevância para a tese: a grande discrepância entre os dados do professor 1 e dos professores 2 e 3 sugere a existência de algum problema quando da aplicação da pesquisa: dos 18 alunos desse primeiro professor, 12 usaram a ajuda online e desses, todos responderam que a pesquisa os ajudou (pergunta 3 da pesquisa acima). Os outros professores tiveram poucas requisições de ajuda, o que indica uma facilidade no manuseio do site pela maioria.

3. Foi fácil para você entender como funciona a estrutura do site?

Motivação: medir o grau de entendimento, por parte do professor, das funcionalidades do site.

Resultado:



Relevância para a tese: o entendimento do funcionamento da página pelos professores foi total. Isso indica que, para aqueles que têm mais experiência com busca de Terminologia, o site não apresenta nenhum tipo de dificuldade para consulta. Caso algum professor decida trabalhar o VoTec com os alunos durante uma aula, parece não haver a necessidade de um treinamento prévio desse professor.

4. De modo geral, como você avalia o site tendo em vista a ajuda que oferece ao aprendiz de tradução?

Motivação: descobrir, perante os olhos do profissional que já trabalha com tradução e educação, se o site é funcional para o propósito para o qual foi desenhado. Há a possibilidade, já disponibilizada para os alunos, de escolher entre ruim, regular, bom e ótimo.

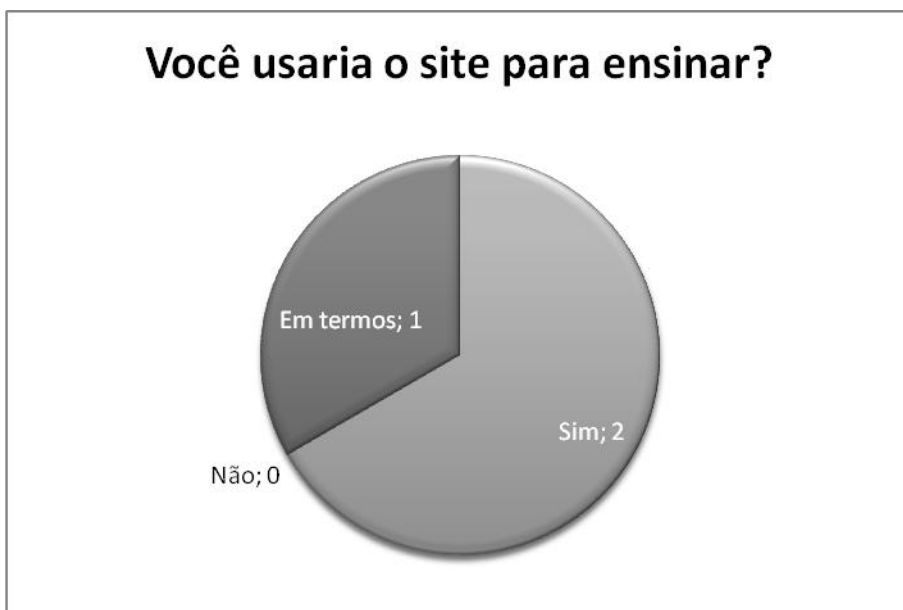
Resultado:

Relevância para a tese: a página de consulta foi julgada como boa para 2 dos 3 professores. Um dos professores, na verdade, inseriu mais um campo na resposta (muito bom), logo houve três tipos de classificação. Tendo concordado em partes ou no todo com a estruturação da página, os professores parecem levar em consideração o uso do mesmo para ensinar Terminologia.

5. Você usaria esse site para ensinar seus alunos a trabalhar com terminologia bilíngüe e manuseio de vocabulários eletrônicos?

Motivação: descobrir se o profissional responsável pelas aulas de Prática de Tradução em geral usaria o VoTec para ensinar seus alunos a trabalharem com obras terminográficas. A resposta pode ser afirmativa, negativa ou ‘em termos’, o que sugere um comentário por parte do professor.

Resultado:



Relevância para a tese: apesar das restrições, os professores aprovaram a página para ser usada como ferramenta no ensino de Tradução/Terminologia. Um comentário foi inserido para a resposta *sim*, outro para a resposta *em termos*, o terceiro professor não expressou comentários. Interessante notar, no comentário do professor 1, que parece não ter havido uma compreensão do objetivo da página: ela nunca se propôs a ser um “Santo Graal” da tradução, apenas se apresenta como uma ferramenta, entre tantas disponíveis, para consulta a vocabulários. O objetivo geral é conscientizar o alunos das possibilidades de consulta disponíveis nas várias ferramentas ou páginas da Internet com a qual vai trabalhar e acelerar a sua produção.

Comentários dos professores

Professor 1

O “dicionário” online não é suficiente para solução dos problemas tradutórios. O tradutor terá de se valer de outros conhecimentos e estratégias complementares para

chegar a um resultado positivo. Na realidade, nenhum glossário, dicionário ou programa dá conta da complexidade do fenômeno tradutório. A competência em tradução não se resume a listas de termos técnicos (mesmo se acompanhadas de ótimas explicações). Tais ferramentas ajudam sem dúvida o tradutor, porém outros tipos de conhecimento, habilidades e estratégias serão necessárias para a produção de um texto na língua de chegada. A título de exemplo: observei que um de meus alunos teve maior facilidade na realização da tarefa proposta; o motivo: ele lida com computação e informática e sabe que a maioria dos termos são usados na forma do idioma inglês e que as explicações oferecidas são apenas para [ilegível] o pesquisador e não uma solução tradutória. [grifos do autor].

Nota: nenhuma outra informação foi passada ao professor, além da carta convite (enviada por e-mail e reproduzida no APÊNDICE G) e o parágrafo introdutório que constava na folha de sua pesquisa. Os demais e-mails foram apenas para acertar o horário de aplicação da pesquisa.

Professor 2

Agiliza a pesquisa; como toda a ferramenta, precisa ser usada para que se ganhe familiaridade.

O objetivo de toda essa pesquisa era levantar as dificuldades que os estudantes de tradução têm com o Léxico, verificar seu grau de contato com ferramentas informatizadas e o grau de satisfação, tanto dos alunos quanto dos professores, com o VoTec. As respostas evidenciaram problemas na competência lingüística desses alunos, especialmente porque estudam num bacharelado em Tradução e têm a disciplina Lingüística implantada em seus cursos. O contato com ferramentas informatizadas depende do semestre cursado (alunos novatos tiveram menos experiência com essas ferramentas) e a experiência de vida de cada respondente. A proposta foi aceita pela maioria, embora vários comentários tenham sido tecidos; os mesmos apontam futuros ajustes e reformulações da página e do banco de dados.

9. Considerações finais

Face aos problemas dos alunos de Tradução em utilizar obras de referência, em especial, terminológicas, pretendi construir uma ferramenta – o VoTec - que proporcionasse um primeiro contato aos aprendizes do curso de Tradução com uma obra de referência e, em especial, com a área de Terminologia. Acredito que esse objetivo tenha sido alcançado. Exponho, a seguir, os passos que levaram à criação da ferramenta e o porque da minha crença nos resultados obtidos.

A partir do perfil do aprendiz de Tradução, um aluno que, a princípio, pretende melhorar seu domínio na língua inglesa, foram também examinadas as problemáticas inerentes ao ensino de Tradução, no Brasil e no mundo, com destaque para alguns autores que tratam do uso de dicionários. Essas análises revelaram que, embora o ensino de Terminologia já faça parte do currículo de vários cursos, a questão das ferramentas de trabalho do tradutor ainda é pouco explorada.

Dentre as competências (como a lingüística, por exemplo) esperadas de um tradutor, a Competência Tradutória, em especial o item Apoio Externo (uso de obras lexicográficas/terminográficas, Internet, aconselhamento com profissionais da área, etc.), provou ser uma das competências mais problemáticas. Foram então, analisados alguns modelos de vocabulários técnicos impressos e eletrônicos, apresentando suas características e verificando como essas poderiam contribuir para a construção de uma nova ferramenta. Além desses, novas tecnologias e programas voltados ao tradutor também foram verificados com o objetivo de verificar os avanços tecnológicos disponíveis na área. Conclui-se pela necessidade da interface dos estudos de Terminologia e Tradução para os aprendizes, pois num mundo com cada vez mais

traduções em áreas técnicas, comprovado por uma pesquisa com tradutores profissionais, não há como dissociá-las.

A proposta para a construção da ferramenta previa que a elaboração da microestrutura de cada termo fosse totalmente baseada em *corpora* técnicos. Tomando o campo de computação como base para “alimentar” o banco de dados criado, foram elaborados dois *corpora* monolíngües na área de hardware. Para delimitar bem as áreas e subáreas, levantou-se uma taxonomia do campo. Cada subárea do *corpus* continha, no mínimo, vinte mil palavras, número julgado suficiente para a construção das definições exigidas pelo banco de dados. Em seguida os *corpora* foram examinados com um programa de análise lexical, com o objetivo de extrair candidatos a termos e todas as informações a eles pertinentes. No decorrer do processo de construção do banco de dados, percebeu-se a necessidade da construção de mais *corpora* bilíngües, dessa vez na área de Lingüística, com o objetivo de apresentar uma metalinguagem para que o aprendiz pudesse entender o que significa cada campo da página de consulta. Essa busca pode ser feita acessando-se a área de Lingüística no primeiro menu.

A construção da página baseou-se nas necessidades aventadas pelos tradutores profissionais conforme pesquisa levantada junto a esses profissionais sobre as estruturas, dentro de uma obra terminográfica, que consideravam as mais importantes para o exercício da tradução.

Entre uma abordagem lexicográfica ou terminográfica para a construção de uma obra, optou-se por uma visão lexicográfica para a construção do banco de dados, que disponibiliza uma grande quantidade de campos para o preenchimento de dados, podendo outros campos virem a ser incluídos. Para a página de consulta - ou o VoTec - foi escolhida uma abordagem híbrida: o consulente pode acessar a base de dados com uma visão mais voltada para os dicionários gerais de língua ou com uma visão mais

comum entre as obras terminográficas disponíveis. Como toda obra, lexicográfica ou terminográfica, pressupõe a construção de uma Ficha Terminológica, essa foi construída já visando sua transposição para o mundo digital: o banco de dados. Foram apresentados todos os passos da construção desse banco com a exibição de suas telas: opções de cadastramento de pesquisadores e obras, a inserção da taxonomia (ou ontologia) de um campo, a inserção de contextos retirados dos corpora, etc. Importante são os paradigmas (partes que compõem a microestrutura de cada verbete) que serviram como base para a construção da microestrutura do verbete: Pragmático (exemplos), Definicional (vários modelos para a construção do mesmo foram discutidos), Informacional (relações morfossintáticas), Semântico (relações semânticas entre um determinado verbete e outros disponíveis no banco), Forma Equivalente (tradução, que serviu de base para a construção da página), Enciclopédico (informações de caráter mais abrangente) e o sistema de remissivas.

Uma das singularidades do VoTec são seus modos de exibição, diferenciando o normal (tradicional, disponível na maioria das obras de consulta) e o descritivo (ou interativo, mais moderno e com uma microestrutura desmembrada em diferentes linhas). Além disso, o consulente dispõe de três maneiras de consultar o VoTec: total (reproduz a microestrutura como um enunciado lexicográfico, com quase todos os campos disponíveis no banco de dados), tradutor (apresenta um enunciado mais terminográfico, com a disponibilização dos resultados levantados na pesquisa com os tradutores) e modular (forma híbrida em relação às duas anteriores, em que o consulente escolhe como montar a microestrutura), que representa uma inovação proposta pela tese. O VoTec também conta com uma ajuda online, cujo objetivo é evitar que o aluno necessitasse de ajuda externa, ou seja, do professor.

Para avaliar a ferramenta, foi proposta uma tarefa de tradução a alguns grupos de aprendizes de graduação, ao final da qual deveriam responder a um questionário. Esse mostrou que a maioria dos entrevistados aprovou a página como um todo. Aos professores perguntou-se se usariam o VoTec para o ensino de Terminologia nas aulas de Tradução e a resposta obtida foi afirmativa.

Além dos elogios colhidos, o mais importante foram as diversas críticas recebidas, pois as mesmas indicam os novos passos que devem ser tomados em futuras versões do projeto.

Entre esses novos passos, posso citar a reformulação das fontes disponíveis na página, já que muitos casos relataram problemas na visualização do site; alguns alunos questionaram, também, o leiaute e a dificuldade em trabalhar com os menus *drop down* do VoTec. O banco de dados, também, pode ser melhorado, e bastante; erros de concordância (como aqueles da figura 33) e ortografia necessitam de urgente correção. Embora uma discussão se o banco deveria ter um caráter terminográfico ou lexicográfico tenha sido levada a cabo, e o segundo tenha sido escolhido como base, ainda há espaço para a inclusão de diversos outros tipos de campos, não necessariamente voltados para tradutores: diferenças diatópicas, diafásicas, diacrônicas, pronúncia, etc. De um projeto voltado para aprendizes de tradução, pode-se chegar a um projeto de banco de dados híbrido lexicográfico/terminográfico: a diferenciação se daria na construção das páginas de consulta, e não no banco. Diferentes páginas poderiam ser construídas de acordo com o público-alvo em questão; o banco de dados seria sempre o mesmo.

A opção da construção de um banco de dados em um padrão que pode ser facilmente alterado (de acordo com a documentação constituída para o mesmo) não foi aleatória: novas ferramentas podem ser acrescentadas para automatizar, sempre que

possível, partes do processo. Muitos pesquisadores já se dedicam a criar essas ferramentas e os exemplos mais bem acabados são aquelas desenhadas para a extração automática de *corpus* e, a partir desse, da extração automática de candidatos a termos. Devido aos constantes avanços da técnica, deve-se considerar esta proposta como um ponto de partida para novas empreitadas na área que combinem Inteligência Artificial, Mineração de Dados e criação de Ontologias e Taxonomias.

Uma dessas novas empreitadas possíveis seria a confecção de um curso, voltado aos professores da área de Tradução, que abarcasse todo o processo de construção de Terminologia que se verificou nesta tese: elaboração de árvores de campo, levantamento de corpora, treinamento com o banco de dados e um melhor detalhamento do funcionamento da ferramenta. Isso criaria professores “proficientes” para trabalhar em todas as etapas na criação de Terminologia. Esse professor poderia, posteriormente, trabalhar todas as fases aqui propostas com os seus alunos e tornar todo este projeto uma grande ferramenta colaborativa, com a constante atualização e inclusão de novas terminologias no banco de dados.

É de vital importância destacar, também, que o VoTec, como ferramenta, não está voltado somente para os aprendizes. Sua estrutura foi configurada como um vocabulário eletrônico totalmente funcional, isto é, ele pode servir de ponto de partida para tradutores profissionais como uma primeira opção de vocabulário eletrônico. Sua grande vantagem, ao contrário de muitas ferramentas (especialmente aquelas disponíveis, gratuitamente, na Internet), é a base científica com a qual os verbetes foram elaborados. Eles não foram constituídos com base na intuição ou na experiência profissional de um único pesquisador; representam o estado atual da terminologia adotada por vários profissionais de um determinado campo do saber. Essa atualidade da terminologia já está imbuída na constituição da ferramenta: conforme mencionado,

prevêm-se constantes atualizações com o acréscimo de textos aos *corpora* levantados com conseqüente alteração das microestruturas disponibilizadas. E por estar acessível a todos na Internet, gratuitamente, cumpre o papel de funcionar como um portal inicial para o exercício diário do Tradutor, aprendiz ou profissional.

Referências bibliográficas

ALMEIDA, G. M. B.; ALUISIO, S. M.; OLIVEIRA, L. H. M. A terminologia na era da informática. **Ciência e Cultura**, v. 58, n. 2. 2006. Disponível em: <http://cienciaecultura.bvs.br/scielo.php?script=sci_arttext&pid=S0009-67252006000200016&lng=en&nrm=iso>. Acesso em : 15/08/2007.

ALMEIDA, G. M. B., PINO, D. H. P., SOUZA, D. S. L. A definição nos dicionários especializados: proposta metodológica. **RI Term – Debate Terminológico**, n. 3, janeiro 2007. Disponível em: <http://www.riterm.net/revista/n_3/Art_Barcellos_Almeida.pdf>. Acesso em: 15 abril 2007.

ALMEIDA, M. B.; BAX, M. P. Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. **Ciência da Informação**, Brasília: IBCT, v. 32, n. 3., 2003.

ALVES, F. A formação de tradutores a partir de uma abordagem cognitiva. **TradTerm**. São Paulo: Humanitas, n. 4.2, 1997.

ALVES, I. M. Definição terminológica: da teoria à prática. **TradTerm**. São Paulo: Humanitas, n. 3, 1996.

_____. A renovação lexical nos domínios de especialidade. **Ciência e Cultura**, v. 58, n. 2. Disponível em: <http://cienciaecultura.bvs.br/scielo.php?pid=S0009-67252006000200013&script=sci_arttext&tlng=pt>. Acesso em: 23 julho 2007.

ARAÚJO, L. A. **De big-bangs a buracos negros no universo da tradução no Brasil**: um estudo sobre o papel da terminologia na prática tradutória e na formação de tradutores. Campinas, 2001. 279 f. Tese (Doutorado em Lingüística Aplicada) - Instituto de Estudos da Linguagem (IEL)/UNICAMP.

ARROJO, R. **Oficina de tradução**. São Paulo: Ática, 2003.

ATKINS, S.B.T., FILLMORE, C.J., JOHNSON, C.R. Lexicographic relevance: selecting information from corpus. **International Journal of Lexicography**. Londres, n. 16.3, p. 251-280, 2003.

AUBERT, F. H. **Introdução à metodologia da pesquisa terminológica bilíngüe**. São Paulo: Humanitas, 1996.

_____. Tradução técnico-científica e terminologia: um ensaio exploratório de uma via de mão dupla. **TradTerm**, São Paulo: Humanitas, n. 4, 2001.

BACELLAR, F. **Elementos para a elaboração de um dicionário terminológico bilíngüe em Ciências Agrárias**. São Paulo: 2002. 200 f. Tese (Doutorado em Lingüística) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo.

BARBOSA, M. A. Da microestrutura dos vocabulários técnico-científicos. In: **Anais do IV Encontro Nacional da ANPOLL**. Recife: Anpoll, 1989.

_____. Da microestrutura de vocabulários técnico-científicos bilíngües: para um microssistema terminológico de ecologia e meio ambiente. In: **IV Simpósio RITerm**, 1994. Disponível em: < <http://www.riterm.net/actes/4simposio/barbosa.htm>>. Acesso em: 11 maio 2007.

_____. Terminologização, vocabularização, cientificidade, banalização: relações. **Acta Semiotica et Linguistica**, São Paulo: Plêiade, v. 7, 1998.

_____. Dicionário, vocabulário, glossário: concepções. In: ALVES, I. M. (org.). **A constituição da normalização terminológica no Brasil**. São Paulo: FFLCH/CITRAT, 2001.

BERBER SARDINHA, A. **Lingüística de corpus**. Barueri: Manole, 2004.

BERGERON, M., KEMPA, C. **Vocabulaire d’Internet: terminologie des technologies de l’information: vocabulaire anglais-français**. Montreal: Office de la langue française, 1995.

BIDERMANN, M.T.C. **Teoria Lingüística**. 2. ed. São Paulo: Martins Fontes, 2001.

BOWKER, L., BARLOW, M. Bilingual concordancers and translation memories: a comparative evaluation. In: **Second International Workshop on Language Resources for Translation Work, Research and Training**. 2004. Disponível em: <http://www.ifi.unizh.ch/cl/yuste/lr4trans-2/WKS_PAPERS/8.pdf>. Acesso em: 10 outubro 2007.

BRAGA, R. C. G. **Para a elaboração de um vocabulário sistemático monolíngüe da telefonia celular**. São Paulo, 2000. Dissertação – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo.

BRITISH NATIONAL CORPUS (BNC). Listagem de palavras, disponível em: <<http://www.lexically.net/downloads/version4/downloading%20BNC.htm>>. Acesso em: 12 maio 2007.

CABRÉ, M. T. **La terminología**. Teoría, metodología, aplicaciones. Barcelona: Editorial Antártida/Empúries, 1993. p. 281-282.

CAMPOS, M. G. **Sobre la elaboración de diccionarios monolingües de producción**. Granada: Universidad, 1994.

CORREIA, M. (coord.). **Dicionário de termos do comércio electrónico**. ILTEC (Instituto de Lingüística Teórica e Computacional). Disponível em: http://www.iltec.pt/projectos/em_curso/dtce.html. Acessado em 15 maio 2007.

CORTEC (*Corpus Técnico Científico*). Disponível em **COMET** (*Corpus Multilíngüe para Ensino e Tradução*): <<http://www.fflch.usp.br/dlm/comet/>>. Acesso em 17 julho 2007.

DARIN, L. C. M. Pesquisa em sala de aula: a prática de tradução na prática. **TradTerm**. São Paulo: Humanitas, n. 4.2, 1997.

DTS. **Mtech – Dicionário eletrônico Michaelis técnico**. São Paulo: DTS Software, 1997.

FINATTO, M. J. B. Caracterização de paradigmas definicionais terminológicos. In: **IV Simpósio RITerm**, 1994. Disponível em: <<http://www.riterm.net/actes/4simposio/bocorny.htm>>. Acesso em: 11 maio 2007.

_____. Microestrutura no dicionário terminológico: bases teóricas e orientações metodológicas em direção ao seu tratamento formal. In: **VI Simpósio RITerm**, 1998. Disponível em: <<http://www.riterm.net/actes/6simposio/finato.htm>>. Acesso em: 11 maio 2007.

FRAÚSTO, I. P. Terminologia: uniformização vs. Criatividade. In: **II Simpósio RITerm**, 1990. Disponível em: <<http://www.riterm.net/actes/2simposio/frausto.htm>>. Acesso em: 11 maio 2007.

FROMM, G. **Proposta para um modelo de glossário de informática para tradutores**. São Paulo, 2002. Dissertação (Mestrado em Lingüística). Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo.

_____. Ferramentas de análise lexical computadorizadas: uma aplicação prática. **Revista Factus**. Taboão da Serra: FTS, n. 3, p. 153-164, 2004.

GAVENSKI, M. M. Microsis: uma experiência no gerenciamento de dados terminológicos. In: KRIEGER, M. G.; MACIEL, A. M. B (org.). **Temas de terminologia**. Porto Alegre/São Paulo: Ed. Universidade/UFRGS/Humanitas/USP, 2001.

GIL, J. R. B. & PYM, A. **Technology and translation** (a pedagogical overview). Disponível em: <http://www.tinet.org/~apym/online/translation/BiauPym_TechnologyAndTranslation.pdf>. Acesso em: 19 julho 2007.

GONÇALVES, J. L. V. R. & MACHADO, I. T. N. Um panorama do ensino de tradução e a busca da competência do tradutor. **Cadernos de Tradução**, Santa Catarina: UFSC, n. 17, 2006.

GONZÁLES-JOVER, A. G. & SIERRA, C. V. **Aspectos metodológicos para la elaboración de diccionarios especializados bilingües destinados al traductor**. Disponível em: <http://www.ua.es/personal/chelo.vargas/Documentos/GomezYVargas_Toledo.pdf>. Acesso em: 25 abril 2005.

GOYOS JR., D. N. **Dicionário jurídico Noronha**. São Paulo: Observador Legal Editora, 2000. CD-ROM

HAENSCH, G. Tipología de las obras lexicográficas e Aspectos prácticos de la elaboración de diccionarios. In: ETTINGER, S. et alii. **La lexicografía**. De la lingüística teórica a la lexicografía práctica. Madrid: Gredos, 1982.

HOUAISS, A. **Dicionário eletrônico Houaiss da língua portuguesa**. São Paulo: Objetiva, 2002.

ILTEC. **Dicionário de termos do comércio exterior**. Disponível em: <http://www.iltec.pt/projectos/em_curso/dtce.html>. Acessado em: 11 maio 2007.

INTERNATIONAL STANDARD ORGANIZATION. ISO 1087. Genebra, 1990.

JAATINEN, H. & JÄÄSKELÄINEN, R. Introducing IT in translator training: experiences from the COLC project. In: **Translation Technology and its teaching**. 2006. Disponível em: <http://isg.urv.es/library/papers/JaatinenJaaskelainen_IntroducingIT.pdf>. Acesso em: 19 julho 2007.

KRIEGER, M. G. Do ensino da terminologia para tradutores: diretrizes básicas. **Cadernos de Tradução**, Santa Catarina: UFSC, n. 17, 2006.

KRIEGER, M. G.; FINATTO, M. J. B. **Introdução à terminologia**: teoria e prática. São Paulo: Contexto, 2004.

LARA, M. L. G. Diferenças conceituais sobre termos e definições e implicações na organização da linguagem documentária. **Ciência da Informação**, Brasília, v. 33, n. 2, p. 91-96, 2004. Disponível em: < <http://www.scielo.br/pdf/ci/v33n2/a09v33n2.pdf>>. Acesso em: 27 julho 2007.

LEFFA, V. J. O dicionário eletrônico na construção do sentido em língua estrangeira. **Cadernos de Tradução**, Florianópolis: UFSC, n. 18, 2006.

MACIEL, A. M. B. Termisul e terminótica. In: *KRIEGER, M. G.; MACIEL, A. M. B (org.). Temas de terminologia.* Porto Alegre/São Paulo: Ed.Universidade/UFRGS/Humanitas/USP, 2001.

MARINOTTO, O. **Para a elaboração de um vocabulário especializado bilíngüe (inglês/português) da linguagem da aviação**: manutenção de aeronaves, controle de tráfego aéreo e operações aéreas. São Paulo, 1995. Tese (Doutorado em Linguística) - Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo.

MILTON, J. Por que estudantes brasileiros escolhem cursos de tradução? **TradTerm**, São Paulo: Humanitas, n. 6, 2000.

MIRANDA, F. B., FARIAS, V. S. Informações discretas e discriminantes no artigo léxico. **Cadernos de Tradução**, Florianópolis: UFSC, n. 18, 2006.

PAGANO, A. Estratégias de busca de subsídios externos: fontes textuais e recursos computacionais. In: ALVES, A., MAGALHÃES, C., PAGANO, A. **Traduzir com autonomia**: estratégias para o tradutor em formação. São Paulo: Contexto, 2003.

PERROTI-GARCIA, A. J. **Vocabulário para ortodontia e ortopedia funcional dos maxilares**: português/inglês/inglês/português.. São Paulo: SBS, 2003. (Série Mil e Um Termos).

PYM, A. **On the market as a factor in the training of translators**. 1993. Disponível em: <<http://www.tinet.org/~apym/on-line/translation/market.html>>. Acesso em: 19 julho 2007.

_____. **Redefining translation competence in an electronic age.** In defence of a minimalist approach. 2002. Disponível em: <<http://www.tinet.org/~apym/online/training/competence.pdf>>. Acesso em: 17 julho 2007.

_____. Asymmetries in the teaching of translation technology. In: **Translation Technology and its teaching.** 2006. Disponível em: <http://www.tinet.org/~apym/online/training/Pym_TeachingAsymmetries.pdf>. Acesso em: 19 julho 2007.

REPPEN, R.; IDE, N.; SUDERMAN, K. **American National Corpus (ANC) - Second release.** Philadelphia: Linguistic Data Consortium, 2005. CD-ROM.

ROST, H. G. **Termos técnicos de seguros:** dicionário trilingüe. 2003. [s.n.].

SCHMITZ, J. R. Algumas considerações sobre a dificuldade da utilização de dicionários de aprendizes dentro e fora da sala de aula. In: Vilson Leffa. (Org.). **As Palavras e a sua Companhia.** Pelotas: Editora Educat, 2000.

SCOTT, M. **WordSmith Tools.** Versão 4. Disponível em: <<http://www.lexically.net/wordsmith/>>. Acesso em 17 junho 2007.

SOWA, J. F. **Building, sharing and merging ontologies.** Tutorial. 1999. Disponível em: <<http://www.jfsowa.com/ontology/ontosar.htm#s6>>. Acesso em: 22 abril 2007.

STOLZE, R. Indicadores de qualidade para a avaliação de traduções no âmbito da didática. **TradTerm**, São Paulo: Humanitas, n. 4.1, 1997.

TAGNIN, S. E. O. **Corpora:** o que são e para quê servem. Minicurso. São Paulo, 2004.

_____. Um *corpus* multilíngüe para ensino e tradução – o Comet: da construção à exploração. **TradTerm**, São Paulo: Humanitas, n. 10, 2004.

ZERFASS, A. Evaluating translation memory systems. LREC-2002: **Third International Conference on Language Resources and Evaluation.** Workshop: Language resources for translation work and research. Las Palmas, 2002; p.49-52. Disponível em: <<http://www.mt-archive.info/LREC-2002-Zerfass.pdf>>. Acessado em: 15 junho 2007.

APÊNDICES

APÊNDICE A. Questionário para os tradutores profissionais, em inglês.

Translators' Research

Hello. This research aims to supply data for my Ph.D. Thesis. These data are about the way translators work with their main tool: the dictionary. It also aims to analyze other help sources for translators. You will need no more than 5 minutes to answer it. Thank you very much in advance and, if you want to comment or criticize anything, feel free to do so.

Guilherme Fromm

1. How many dictionaries do you usually need for a search?

- 1
- 2
- 3
- 4
- more than 4

2. What kind of dictionaries do you use most? (more than one answer possible)

- general monolingual
- general bilingual
- technical monolingual
- technical bilingual
- technical multilingual

3. Among the following search ways to find a word, what's the efficiency of each one? Classify them from the least important (1) to the most important (5).

- | | | | | | |
|-----------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|
| Printed Dictionaries | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| CD-ROM Dictionaries | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| Internet Dictionaries | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| Glossaries Download | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| Internet Glossaries | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| Internet Forum | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| Experts advice | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |

- On-line Corpus 1 2 3 4 5
- Internet search engines (Google, etc.) 1 2 3 4 5
- Similar printed texts 1 2 3 4 5

4. When you are searching for translations in a bilingual dictionary, what's the most common option for you?

- search just the translation of the word
- search the definition of the word
- search the examples given in the entry
- search the translation and the definition of the word
- search the definition and examples given in the entry
- search the translation, definition and examples given in the entry

5. What kind of bilingual or multilingual dictionary do you use most?

- one that gives just the word's translation
- one that gives the word's translation and examples it in the target language
- one that gives the word's translation and definition in the target language
- one that gives the word's translation and definition in the target language and exemplifies in the source language
- one that gives the word's translation, definition, examples and shows cross references (see also:)
- one that gives the word's definition in the source language and translation and examples in the target language
- one that combines all possibilities above

6. What are the names of the 3 dictionaries you use most of the time?

7. How would you classify the information that a dictionary entry can present before the definition of a word, them according to your task as a translator? Classify it from the least important (1) to the most important (5).

- Acronyms and Abbreviations 1 2 3 4 5
- Word Class (part of speech) 1 2 3 4 5
- Gender (male, female, neutral) 1 2 3 4 5
- Number (singular, plural) 1 2 3 4 5
- Pronunciation, stress 1 2 3 4 5

Translation (s)	<input type="radio"/>	1	<input type="radio"/>	2	<input type="radio"/>	3	<input type="radio"/>	4	<input type="radio"/>	5
Verbal conjugation (irregular verb forms)	<input type="radio"/>	1	<input type="radio"/>	2	<input type="radio"/>	3	<input type="radio"/>	4	<input type="radio"/>	5
Verbal classification (grammar codes: transitive, intransitive, etc.)	<input type="radio"/>	1	<input type="radio"/>	2	<input type="radio"/>	3	<input type="radio"/>	4	<input type="radio"/>	5
Homonyms	<input type="radio"/>	1	<input type="radio"/>	2	<input type="radio"/>	3	<input type="radio"/>	4	<input type="radio"/>	5
Synonyms	<input type="radio"/>	1	<input type="radio"/>	2	<input type="radio"/>	3	<input type="radio"/>	4	<input type="radio"/>	5
Etymology	<input type="radio"/>	1	<input type="radio"/>	2	<input type="radio"/>	3	<input type="radio"/>	4	<input type="radio"/>	5
Language Levels (popular, slang, regionalisms, etc.)	<input type="radio"/>	1	<input type="radio"/>	2	<input type="radio"/>	3	<input type="radio"/>	4	<input type="radio"/>	5
Spelling differences	<input type="radio"/>	1	<input type="radio"/>	2	<input type="radio"/>	3	<input type="radio"/>	4	<input type="radio"/>	5
Spelling division	<input type="radio"/>	1	<input type="radio"/>	2	<input type="radio"/>	3	<input type="radio"/>	4	<input type="radio"/>	5
Pragmatics (usage notes)	<input type="radio"/>	1	<input type="radio"/>	2	<input type="radio"/>	3	<input type="radio"/>	4	<input type="radio"/>	5
Technical area	<input type="radio"/>	1	<input type="radio"/>	2	<input type="radio"/>	3	<input type="radio"/>	4	<input type="radio"/>	5
Word's lifetime (neologism, archaisms, etc.)	<input type="radio"/>	1	<input type="radio"/>	2	<input type="radio"/>	3	<input type="radio"/>	4	<input type="radio"/>	5

8. How would you classify the information that a dictionary entry can present after the definition of a word, them according to your task as a translator? Classify it from the least important (1) to the most important (5).

Examples	<input type="radio"/>	1	<input type="radio"/>	2	<input type="radio"/>	3	<input type="radio"/>	4	<input type="radio"/>	5
Synonyms	<input type="radio"/>	1	<input type="radio"/>	2	<input type="radio"/>	3	<input type="radio"/>	4	<input type="radio"/>	5
Antonyms	<input type="radio"/>	1	<input type="radio"/>	2	<input type="radio"/>	3	<input type="radio"/>	4	<input type="radio"/>	5
Sayings, quotations, idioms with this word	<input type="radio"/>	1	<input type="radio"/>	2	<input type="radio"/>	3	<input type="radio"/>	4	<input type="radio"/>	5
Cross references (see also:)	<input type="radio"/>	1	<input type="radio"/>	2	<input type="radio"/>	3	<input type="radio"/>	4	<input type="radio"/>	5
Other word classes (ex.: adjective, adverb, etc.) of the same word	<input type="radio"/>	1	<input type="radio"/>	2	<input type="radio"/>	3	<input type="radio"/>	4	<input type="radio"/>	5
Prepositions that occur with the verbs	<input type="radio"/>	1	<input type="radio"/>	2	<input type="radio"/>	3	<input type="radio"/>	4	<input type="radio"/>	5
Phrasal Verbs	<input type="radio"/>	1	<input type="radio"/>	2	<input type="radio"/>	3	<input type="radio"/>	4	<input type="radio"/>	5
Cultural Notes	<input type="radio"/>	1	<input type="radio"/>	2	<input type="radio"/>	3	<input type="radio"/>	4	<input type="radio"/>	5
Drawings	<input type="radio"/>	1	<input type="radio"/>	2	<input type="radio"/>	3	<input type="radio"/>	4	<input type="radio"/>	5
Derived words (without definitions)	<input type="radio"/>	1	<input type="radio"/>	2	<input type="radio"/>	3	<input type="radio"/>	4	<input type="radio"/>	5
Pragmatics (usage notes)	<input type="radio"/>	1	<input type="radio"/>	2	<input type="radio"/>	3	<input type="radio"/>	4	<input type="radio"/>	5

9. How many meanings, inside the same entry, do you consider ideal?

- 1
- 2
- 3
- 4
- as many as necessary

10. In your opinion, what's the easiest way to consult homonyms (as bank)?

- clustered in the same entry
- put into different entries

11. In your opinion, what's the easiest way to consult words with more than one meaning, e.g. horse?

- clustered in the same entry
- put into different entries

12. In your job as translator, what are the main areas you deal with?

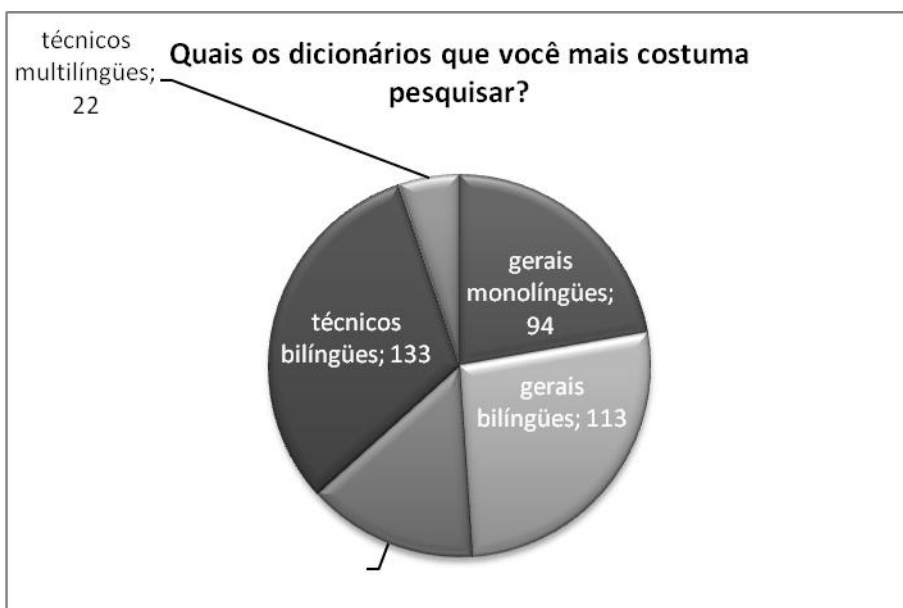
- Automotive
- Business Administration
- Education
- Humor
- Information Technology (computers)
- Information Technology (telecommunications)
- Insurance
- Legal (in general)
- Linguistics
- Literary
- Medical (in general)
- News (in general)
- Sports
- Technical (in general)
- Technical-industrial
- Tourism
- Other (s)

APÊNDICE B. Pesquisa com os tradutores. Perguntas e gráficos de respostas.

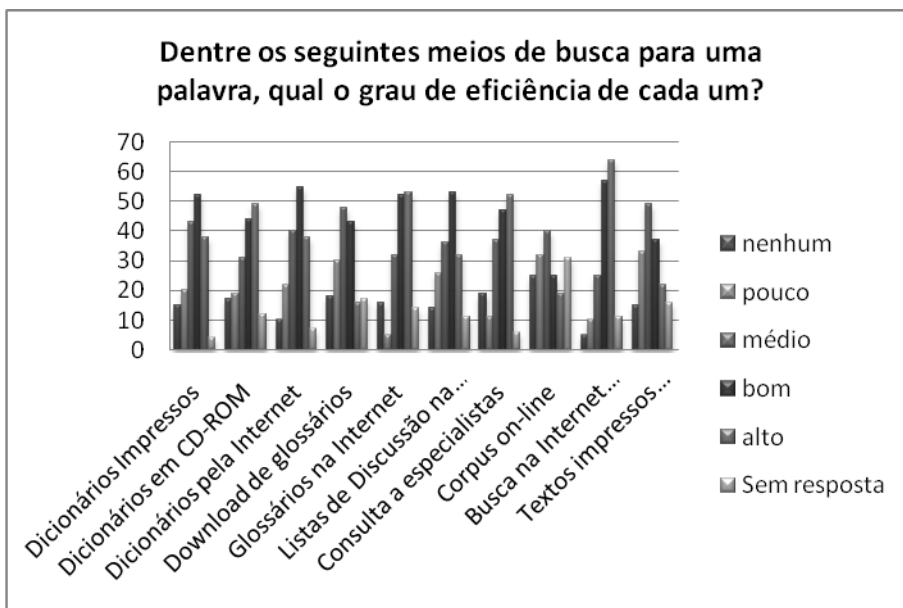
1. Quantos dicionários você costuma usar durante uma consulta?



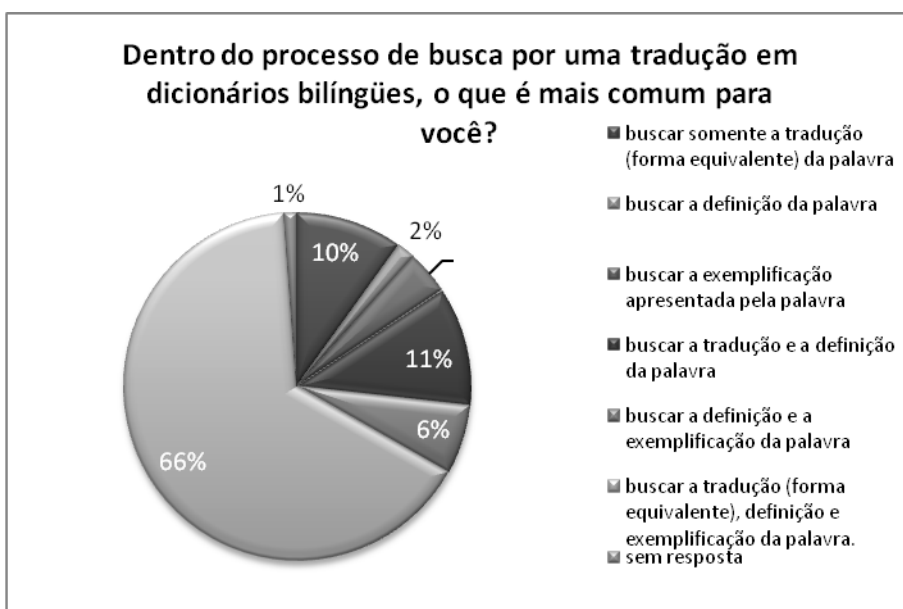
2. Quais os dicionários que você mais costuma pesquisar? (mais de uma resposta possível).



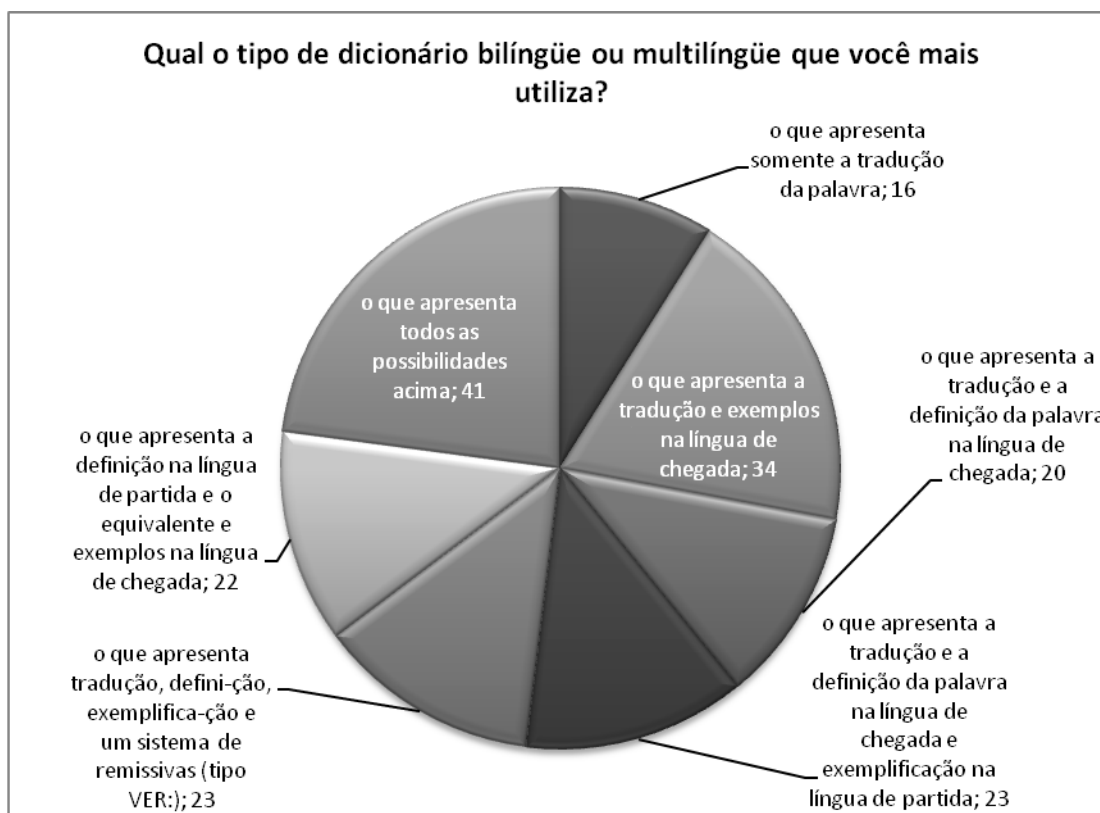
3. Dentre os seguintes meios de busca para uma palavra, qual o grau de eficiência de cada um? Classifique: 1 = nenhum, 2 = pouco, 3 = médio, 4 = bom e 5 = alto.



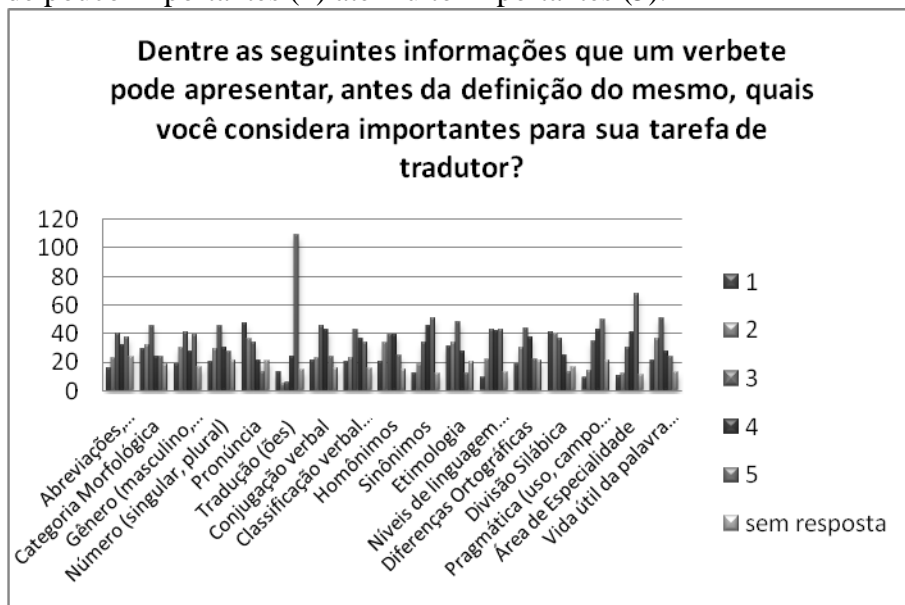
4. Dentro do processo de busca por uma tradução em dicionários bilíngües, o que é mais comum para você?



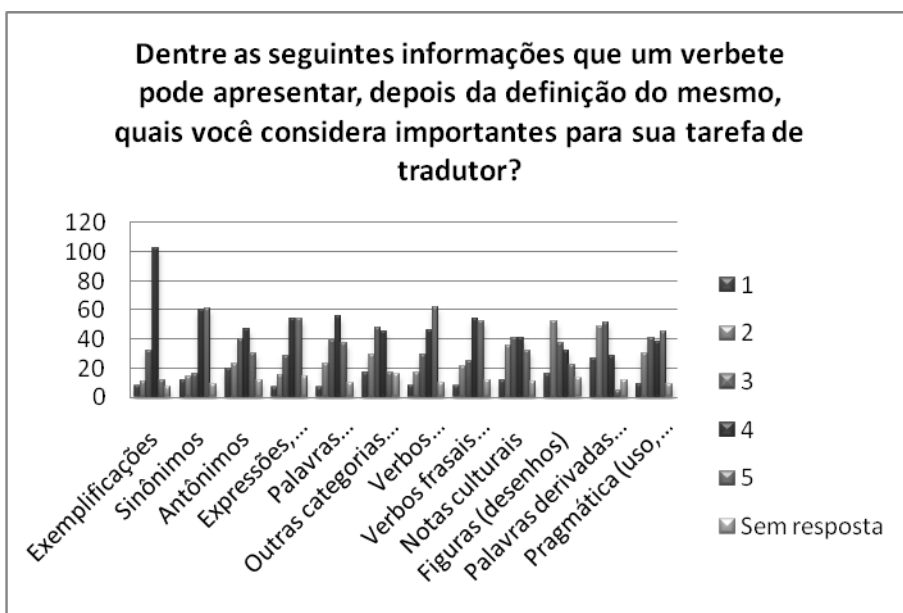
5. Qual o tipo de dicionário bilíngüe ou multilíngüe que você mais utiliza?



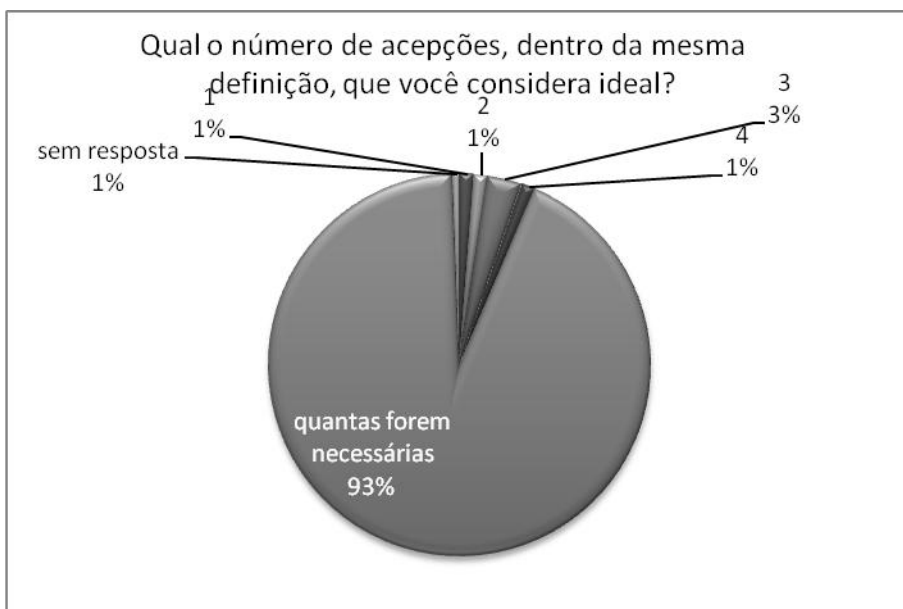
7. Dentre as seguintes informações que um verbete pode apresentar, antes da definição do mesmo, quais você considera importantes para sua tarefa de tradutor? Classifique-os de pouco importantes (1) até muito importantes (5).



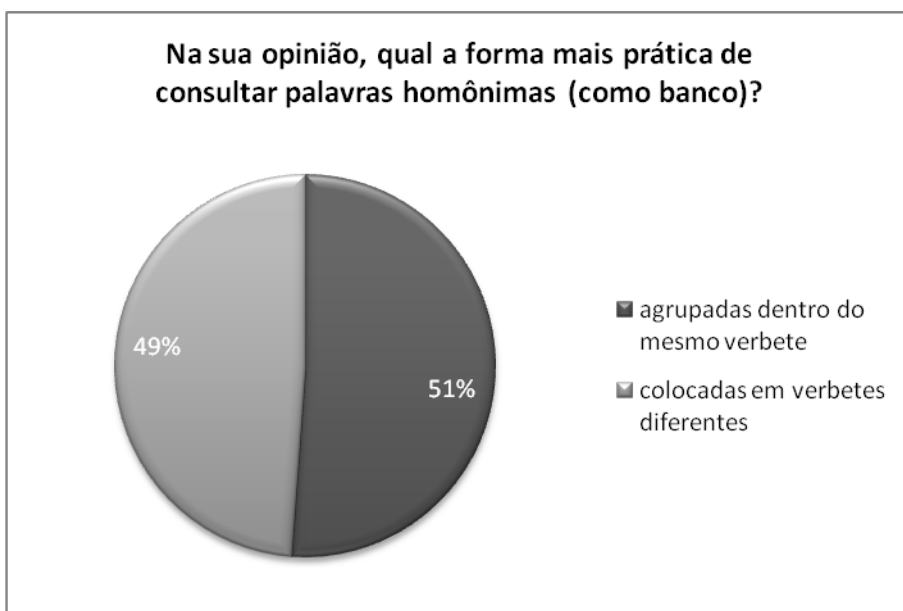
8. Dentre as seguintes informações que um verbete pode apresentar, depois da definição do mesmo, quais você considera importantes para sua tarefa de tradutor? Classifique-os de pouco importantes (1) até muito importantes (5).



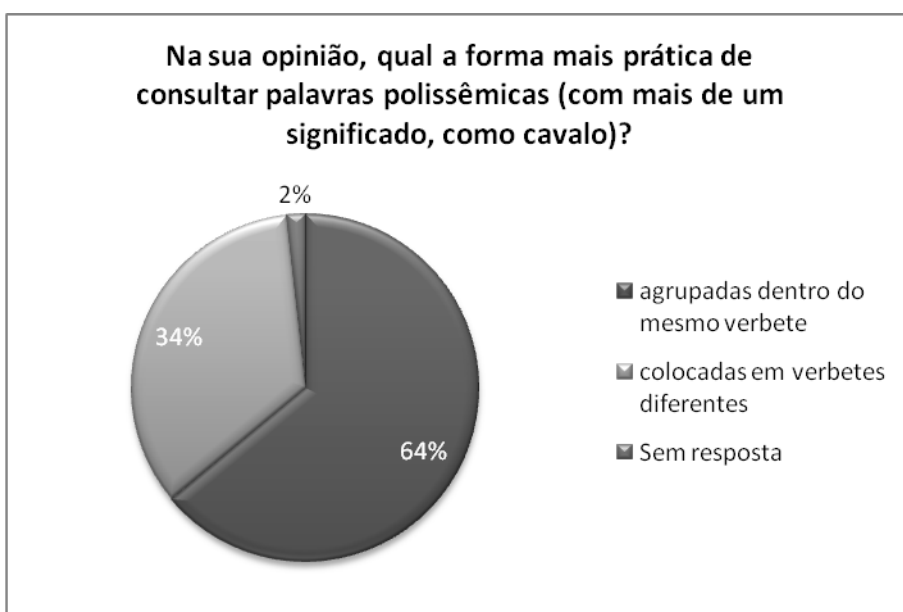
9. Qual o número de acepções, dentro da mesma definição, que você considera ideal?



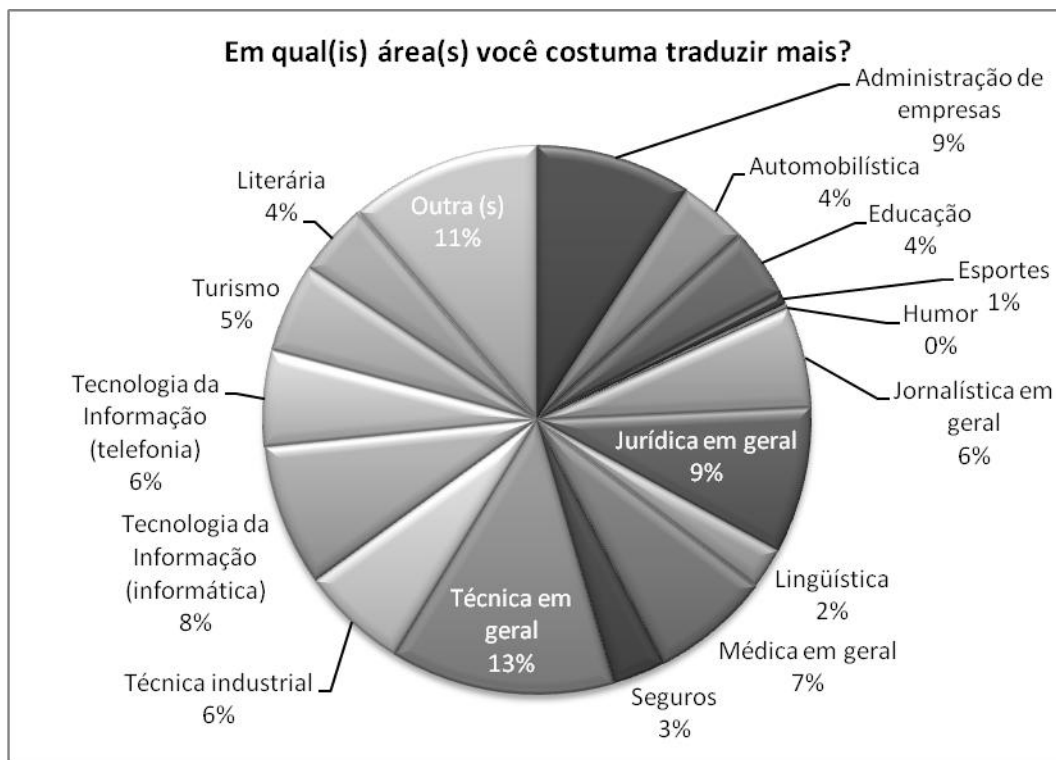
10. Na sua opinião, qual a forma mais prática de consultar palavras homônimas (como banco)?



11. Na sua opinião, qual a forma mais prática de consultar palavras polissêmicas (com mais de um significado, como cavalo)?



12. Em qual(is) área(s) você costuma traduzir mais?



APÊNDICE C. Relação dos dicionários aventados pelos tradutores

(pergunta 6 da pesquisa com os tradutores)

ABBY Lingvo	Collins-Pons DE>EN dico	Dietl, Dictionary of Legal, Commercial, Political Terms
ACADEMIA(MEDICAL DICTIONARY)	Collins-Robert French-Englisch	Dorland
AHD	CollinsRobertFrench	Dorland, Medical
AL-MAWRID	Concise Oxford English dic.	DRAE (Spanish Royal Academy Dictionary)
Alkalay HEB-ENG-HEB	der kleine Muret Sanders	Duden
American Heritage	Dic Eletronico Houaiss	Duden Pictorial Dictionaries
American Heritage Dictionary	Dic Francês-Português da Porto Editora	Duden Universalwoerterbuch
angol-magyar, Orszagh/Magay	Dic Maxi Michaelis	Duden, Die deutsche Rechtschreibung.mannheim e.a, 2000
Antonio Houaiss Ingl-Port	Dic Michaelis - UOL	Economía y Empresa
Aurélio	Dic. Jurídico	Eijiro (online bilingual Japanese to English)
Aurélio Eletrônico	Dic. Medico Ilustrado Dorland ingles>portugues	Ernst
Aurélio XXI	Dic. Tec. de Negocios	Ernst (tech. Spanisch - German)
Babylon	Dic. Termos Tecnicos - Luiz Mendes Antas	Ernst De-FR
Babylon on-line (multilanguage)	Dicc. Técnico Collazo	Ernst Dictinnaire général de la technique industrielle
Bantas	Dicc.para Ingenieros Robb	Ernst En-FR
Barrons law, finance, business	Diccionario crítico de dudas en inglés-español de medicina	Ernst GER-ENG
Becher, Wörterbuch der Rechts- und Wirtschaftssprache, Teil I/II. München, 1999	Diccionario de la técnica industrial, Dr. Ing. Richard Ernst, Herder	Ernst technological Dict.
Beigbeder Atienza EN>ES	Diccionario de María Moliner	Ernst Woerterbuch der industriellen Technik
Beuth: Wörterbuch technischer Begriffe mit Definitionen nach DIN	Diccionario de términos económicos, financieros y comerciales, Alcaraz Varó y Hughes, Ariel	Ernst Wörterbuch der industriellen Technik
Blacks Law Dic	diccionario de terminos juridicos	Ernst Wörterbuch der industriellen Technik ES-DE DE-ES
Bookshelf	Diccionario de Terminos Juridicos - Alcaraz & Hughes	Ernst Wörterbuch der industriellen technik German-English
Brandstetter	Diccionario de uso de la lengua española	Ernst, Brandstetter-Verlag
Britannica CD	Diccionario de uso del español - Moliner	Ernst, Wörterbuch der industriellen Technik
Brokhaus	Diccionario Jurídico - Cabanellas	Ernst, Wörterbuch der industriellen Technik, Band V /VI. Wiesbaden, 2000
Cabanellas, Legal	Dicionário Aurélio	EURODIC
Cambridge	Diccionario de la Real Academia de la lengua española	Eurodicam
Cobuild	Dicionário de Medicina e Saúde - Luis Rey	Eurodicatum
Cobuild (cd)	DICIONARIO DE ODONTOLOGIA	eurodicautom
Cobuild Series	Dicionário de Termos de Negócios	Eurodicotaum
Collins	Dicionário Houaiss	Eurodictaucom (internet)
Collins Cobuild	Dicionário Técnico P-I	Europa - Eurdicautom
Collins DE>EN	Dicionários e Enciclopédias jurídicas	Everest
Collins English Dictionary	Dictionary of Drives	Fachwörterbuch Logistik und Supply Chain Management
Collins English-English	Dictionary of Engineering - Ernst	Felonline
Collins Millenium Edition (English)	Dictionary of Engineering - Schmidt	FIG
Collins Robert	Dictionnaire Juridique Moors	Garzanti
Collins Robert English French		GDT
Collins Robert English-French		GDT (Grand Dic Termonogique)
Collins Robert French		
Collins Robert French-English		
Collins Spanish-English		
Collins SpanishEnglish		
Collins The English Dictionary		
Collins Unabridged English-German		

- general (spanisch -german) (PONS)
 Gladstone
 Gladstone Dictionary of Medicine
 Gladstone Medical Dictionary
 Glosario del traductor publico
 Glosario Internacional del Traductor
 - Marina Orellana
 Glosario internacional para el
 traductor - Orellana
 Google (Internet)
 Grand dictionnaire (internet,
 french/english)
 Grand Dictionnaire Langenscheidt
 Grand Dictionnaire Terminologique
 Grand Dictionnaire Terminologique
 (on line)
 Granddictionnaire.com
 Grande dizionario tecnico Hoepli
 Brandstetter tedesco-italiano e
 italiano -tedesco
 Gross: Chemie und chemische
 Technik (German-English)
 Hachette E. I. 2003
 Hakusuisha nihongo-supeingo
 Hamblock & Wessels GER-ENG
 Hamblock-Wessels
 Wirtschaftswörterbuch
 Hamblock/Wessels Großwörterbuch
 Wirtschaftsenglisch
 Harper Collins Unabridged German-
 English
 Harraps Standard
 Hazon Garzanti IT<>ENG
 Herbst (economic)
 Hoepli (general)
 Hoepli (technical)
 Hoepli - marolli IT<>ENG
 houaiss
 Houaiss (bilíngüe I-P)
 Houaiss eletrônico
 Houaiss Ingles/Port
 Houaiss Português
 Houaiss Sinonimos/Antonimos
<http://www.felonline.nl/>
<http://www.wkkl.nl/lexicon/>
 Hueber Taschenwörterbuch Recht
 GER-ENG
 iccionari de terminos economico-
 finaieros
 Internet
 Ivens-Ferraz (Petróleo)
 Jurídico - Noronha
 Kenkyu-sha New College Japanese-
 English dictionary
 Kich-Ehlers: Economia e Direito -
 port/ale/port
 Kluwer Groot Polytechnisch
 Kluwer´s
 Kojien (Japanese monolingual)
 Kucera Wörterb. d. exakten
 Naturwissenschaften und der
 Technik
 Langeenscheidt - Großwörterbuch
 Englisch-Deutsch/Deutsch-Englisch
 Langendorf Juridisch woordenboek
 Langenscheidt
 Langenscheidt (several)
 Langenscheidt (various)
 Langenscheidt e-Fachwoerterbuch
 der Technik und angewandte
 Wissenschaften DE>EN
 Langenscheidt Handwörterbuch
 Englisch-Deutsch
 langenscheidt technical
 Langenscheidt Technik und
 angewandte Wissenschaften
 Langenscheidts Fachwörterbuch der
 industriellen Technik ES-DE DE-ES
 Langenscheidts Großwörterbuch
 LaRousse
 Larousse francais-spagnol español-
 frances
 Larousse Spanish French
 Le Docte
 Le Grand Dictionnaire
 terminologique
 Le grand Robert
 le petit Larousse illustré
 Le Petit Robert
 Le Petit Robert monolingual French
 Le Pons
 Le Robert
 Le Robert & Collins senior
 Le Robert (French)
 Le Robert (paperback)
 LE Robert micropoche
 Leo
 LEO online
 LEO online dictionary
 Leo online-dictionary
 Lepetit Robert
 Lerobert et Collins
 lexikon
 Lexis (Bundessprachenamt CD-
 Rom)
 Lingvo
 Lingvo Russian-English
 Longman
 Longman monolingual
 Luiz Mendes Antas
 M. Webster ENG-ENG
 MacMillan (monolíngüe)
 MacMillan English Dictionary
 magyar-angol.
 Orszagh/Futasz/Kovecses
 maria moliner
 Marketing Dict.
 McGraw Hill (técnico)
 Merriam Webster
 Merriam Webster monolingual
 Merriam-Webster
 Merriam-Webster Unabridged
 Merriam-Webster: Medical
 Dictionary
 mertens - guide de la traduction
 meu dicionário pessoal
 Michaelis
 Michaelis - port/ale/port, versão do
 início do século passado , para
 tradução de textos mais antigos
 Michaelis Ing>Port
 Michaellis
 michaellis port/ing, ing/port
 Michal HEB-ENG-HEB
 Mink
 Multilex
 Multitran
 muret sanders langenscheidt
 english-german-english
 Müller: Gegenwörterbuch
 Navarro: Diccionario inglés-español
 de dudas de medicina
 OED
 Onelook
 Other technical depending on the
 subject
 own terminology database
 Oxford
 Oxford
 Oxford (various)
 Oxford Advanced Learners
 Oxford Bilingual dictionary
 English-Spanish
 Oxford Dictionary
 Oxford Duden German-English
 Oxford Encyclopedic
 Oxford English
 Oxford english-spanish-english
 Oxford informática
 Oxford Monolingual English
 Oxford Spanish Dictionary
 Oxford Spanish-English
 Oxford Superlex
 Petit Rober 1
 Petit Robert
 Picchi
 Pons
 pons english-german-english

PONS FR/DT and E/DT	Schlegelmich tech. dictionary german-italian and viceversa	van Dale NL-EN, EN-NL
Pons Globalwoerterbuch	Schlegelmilch	Van Dale NL>EN dico
Pons Grosswoerterbuch English- German	Schmitt - Technical Dic.	Van den End NL>EN legal dico
Pons-Collins	Schäfer Wirtschaftslexikon	Verlag VEB Technik (several)
PORTO	Seidel	Wahrig
Porto Editora	Sell (Técnico)	Webster
Potonnier (German-French)	Servotte: Internationaal handelswoordenboek	Webster eletrônico
Radde	siemens Electrical	webster port/ing, ing/port
RAE	spinelli cassasanta	websters
rae.es	STEDMAN	Websters ingles>portugues
Ragazzini	Stedman (Médico)	Webster´s Unabridged
Random House	Stedman: Bilingüe de ciencias médicas	Wenske Chemie German -English
random house english	Synonym Finder	Wildhagen Héraucourt German-to- English
Random house unabridged	Synonymes	Woerterbuch der industriellen Technik - Ernst, Brandstetter Verlag german/english and viceversa
Real Academia (Spanisch)	Taylor (bilíngüe P-I)	Woerterbuch Rechnungslegung und Steuern (PricewaterhouseCoopers)
Real Academia Espanhola	Taylor-Houaiss Bilingual Portuguese English	Word Web
Richard Ernst Technik	Technical Dictionary	Wordnet
Robb, Ingeneering	Termium	wordreference.com
Robert	Termium Electronic CD Version	WordWeb
Robert & Collins	The Heritage	www.archaeologicalresource.com
Robert & Collins Anglais-Français	THE UNIFIED MEDICAL DICTIONARY	www.granddictionnaire.com
Robert Collins	Thibaut Dictionnaire Fr-All	www.leo.org
robert et collins	TLFI (Trésor de la langue française) online	Wörterbuch der industriellen Technik - inglês/alemão/inglês
Romain Byrd Thilecke [legal] DE>EN	Trésor de la langue française	Wörterbuch der Rechts und Wirtschaftssprache, Becher
Romain/Bader/Byrd - Dictionary of legal and commercial terms	Ultralingua	Zahn Banking and Stock Trading
Rothe Rechtswörterbuch ES-DE DE-ES	Van Dale	Zanichelli
Routledge	Van Dale Dutch to English	Zanichelli (economics and business)
Routledge Technical Dic.	Van Dale Frans-Nederlands	Zanichelli IT<>ENG
Sachs-Vialatte	Van Dale Groot woordenboek Engels-Nederlands	Zingarelli
Schaefer Hyperbook		
Schaefer Wirtschaftswoerterbuch		

APÊNDICE D. Exemplo de Ficha Terminológica Não-computadorizada.

Entrada:	Forma Equivalente:	Cat. Gram.	Nº	Sing./Plural	Sigla/Acrônimo	Entrada por extenso	Var. Morfossimáticas	Área rede	Aceção nº única	Cópus 1219
Peer-to-peer	P2P, ponto a ponto	sm	s	s/pl	P2P					
<p>Contexto: Depois de assustar músicos e gravadoras, os programas para troca de arquivos entre internautas (conhecidos como "peer-to-peer") começam a provocar a indústria cinematográfica. A maior parte dos programas "peer-to-peer" que surgiram depois do "Napster", diferentemente da criação de Shawn Fanning, permitem que os usuários compartilhem não apenas vídeos mas também imagens e software, além de músicas em MP3.</p> <p>Contexto: Pretendemos desenvolver mais produtos interessantes baseados na tecnologia "peer-to-peer" [ponto a ponto], que permite a comunicação de usuários em locais diferentes" e chamar mais companhias para expandir o negócio.</p> <p>Contexto: Mas, para fazer tudo isso ao mesmo tempo, só com uma aplicação peer-to-peer (colega a colega ou "entre pares"). Em vez de servidores centralizados, o peer-to-peer se apóia na conexão direta entre pares, que se revezam nos papéis de cliente e servidor.</p> <p>Contexto: Peer-to-peer - Modelo de conexão no qual cada um dos equipamentos conectados tem os mesmos recursos e cada parte pode dar início a uma sessão. Na internet, refere-se a uma rede transitória que garante a um grupo de usuários com o mesmo programa acessar arquivos instalados no disco rígido de outros. É o princípio utilizado pelo Napster e programas e serviços similares. ... Discussões sobre direito autoral à parte, a verdade é que o P2P nunca mais será o mesmo</p>										
Traços Distintivos										
Conceito	A	B	C	D	E	F	G	H	I	J
1	programa		troca	arquivos	internautas					
2	tecnologia	ponto a ponto	comunicação		usuários	locais diferentes				
3			Conexão direta		Pares, colega a colega		Revezamento	Cliente	Servidor	
4	Rede transitória			Acesso de arquivos	Usuários					Disco rígido
<p>Conceito final: programas que criam uma rede transitória ponto a ponto para a troca de arquivos entre usuários de locais diferentes, revezando as máquinas desses usuários como clientes e servidores e possibilitando o acesso a arquivos nos respectivos diretórios.</p> <p>Termo Dicionarizado? () sim (x) não Definições coincidentes? () sim () não () parcial Fonte (s): Definição Dicionarizada:</p>										

APÊNDICE E. Documento de requisitos para a construção do banco e da página.

Documento de Requisitos – Dicionário eletrônico de termos terminológicos.

Nome do projeto: Dicionário de termos terminológicos.

Nome do cliente: Guilherme

Assunto: Documento de Requisitos.

Numeração: 1

Versão: 0.1

Criado por: Ednei Alves dos Santos.

Data da criação: 19/08/2006

Histórico de Revisões

Data	Versão	Descrição	Autor
19/08/2006	0.1	Criação do documento inicial	Ednei Alves dos Santos

Descrição do problema:

O cliente é Mestre em lingüística e está elaborando sua tese de doutoramento que propõe uma nova forma de consultar termos em dicionários terminológicos com a finalidade de auxiliar os estudantes de tradução em seu aprendizado. Como parte de sua tese o cliente necessita elaborar um sistema computacional que permita aos estudantes de tradução Inglês/Português utilizar o método proposto na tese para realizar suas consultas.

Descrição da Solução:

Sistema web que permita aos usuários pesquisar por um termo, obtendo como resultado o verbete do termo desejado em português, Inglês ou ambas. Permita ao usuário escolher para exibição as informações que lhe interessar sobre o verbete, ou duas outras formas de exibição chamadas tradutor e total. O sistema deve permitir também que os usuários façam consultas à duas fontes externas ao site a primeira é a sistema corpus e a segunda é a um conjunto de sites existentes na web que oferecem ferramentas que a auxiliam o trabalho de tradução. O sistema deve ainda contar com uma área administrativa que auxilie na construção e inserção de novas fichas terminológicas pelo administrador do sistema, bem como a atualização e remoção de fichas existentes.

Glossário:

Administrador – pessoa responsável por gerir o conteúdo do dicionário.

Cliente – A pessoa que encomendou o sistema.

Ficha terminológica – É uma ficha estruturada que auxilia na construção de um verbete para um dicionário terminológico.

Pesquisador – é um usuário ao qual é dada a permissão de inserir novas fichas terminológicas.

Termo equivalente – Dado um termo numa língua, o termo equivalente é o correspondente desse termo na outra língua do dicionário.

Usuários - Aprendizes de tradução inglês – Português.

Documentos anexos:

1. Ficha Terminológica Original.doc
2. Projeto Modular.doc
3. Árvore de Domínio da Computação.doc

Obs: Estão anexos a esse documento todos os e-mails trocados com o cliente.

E-mails escritos pelo redator do documento estão sobre a forma emailnumero_data_de_envio, e-mails escritos pelo cliente em resposta estão nos arquivos resposta-emailnumero_data_chegada.

Requisitos funcionais:**Controle de acesso:**

1. Haverá apenas um único administrador no sistema.
2. O sistema deve permitir que o administrador possa cadastrar pesquisadores.
 - 2.1. O cadastro é sempre feito pelo administrador e os seguintes campos devem ser preenchidos para o cadastro de um pesquisador: nome, sexo, nacionalidade, endereço, telefone, instituição.
3. O Acesso à área administrativa será privado ao Administrador e aos pesquisadores e será feito mediante autenticação de nome de usuário e senha.
4. O administrador poderá inserir novas fichas terminológicas, atualizar e remover fichas existentes no sistema.
5. Os pesquisadores poderão inserir novas fichas terminológicas
 - 5.1. As fichas inseridas pelos pesquisadores devem ser aprovadas pelo administrador antes de serem integradas ao sistema.
 - 5.2. Caso uma ficha seja reprovada, o sistema deve enviar uma notificação via e-mail para o pesquisador que inseriu a ficha solicitando a revisão da mesma.

Área administrativa.

1. Inclusão de novas áreas e sub áreas.
 - 1.1. Cada área pode ter n sub áreas, e cada sub área pode ter suas próprias sub área, a árvore pode crescer até o 5 nível contando a área como raiz. Essa árvore é

chamada Árvore de domínio (para melhor entendimento ver anexo 3 –Árvore de Domínio da Computação.doc)

2. Inserção de novas fichas terminológicas:

2.1. Inserção das características gerais do termo.

2.1.1. A inserção de uma nova ficha terminológica deve ser iniciada pela entrada do nome do termo, e pela seleção das áreas e subáreas (cinco níveis no total) ao qual o termo pertence.

Obs: O cliente sugeriu o uso de Combo Box (menu Drop Down) para a seleção de áreas e subáreas.

2.1.2. Em seguida os Contextos devem ser inseridos, o administrador pode adicionar quantos contextos desejar, deve ser possível ao administrador visualizar todos os contextos relacionados aos termos de forma prática para que o administrador possa preencher as características restantes do termo (a exceção da definição, ver item 2.2). Cada contexto é formado pelos seguintes campos: Exemplo, Conceito, Fonte, Data de coleta, Data de inserção.

2.1.3. O administrador a partir de visualizando dos contextos preencherá o formulário contendo os campos: Entrada, Língua, Entrada equivalente, Número da aceção, Posição de frequência no corpus, Formas Equivalentes, Categoria Gramatical, Gênero, Número, Possibilidades de número, Sigla, Acrônimo, Entrada por extenso, Variações morfossintáticas, Relações de hiperonímia, Relações de hiponímia, Relações de co-hiponímia, Relações de antonímia, Relações de sinonímia, Remissiva Remissiva.

2.1.4. O sistema deve ainda permitir as entradas: Termo dicionarizado, Definições coincidentes, Fonte da definição, Definição dicionarizada. Para essas entradas a visualização dos Contextos não é mais necessária

Obs: O cliente gostaria que todas as características do termo, com exceção da definição, fossem preenchidos num mesmo formulário. Quanto à visualização dos contextos, ela é necessária para a extração dos valores dos campos da ficha terminológica citados no item 1.1.3, por isso o cliente necessita visualizar ao mesmo tempo todos os contextos e os campos do formulário da ficha. Sugeri ao cliente a divisão da página em duas áreas horizontalmente, na área de cima se encontrariam os Contextos como pode haver muitos contextos essa área poderia conter uma barra de rolagem, que permitisse o administrador a rolar o texto para baixo visualizando todos os conceitos sem perder a visão do formulário que seria colocado na área de baixo.

2.2. Inserção da Definição

2.2.1. A definição é composta pelos traços distintivos, pelo conceito final e pela definição.

2.2.2. O sistema deve fornecer novamente uma forma prática de visualização que permita ao administrador visualizar a lista dos n conceitos inseridos anteriormente, sem perder a visualização da tabela de traços distintivos. A tabela de traços distintivos deverá ter n linhas uma para cada conceito, e m colunas. A primeira coluna da tabela deve conter a numeração dos conceitos de 1 a n, ao clicar em um dos números o conceito relacionado a esse número deve ser mostrado ao administrador para que esse possa

extrair os traços distintivos do conceito. A tabela pode ter no máximo 30 colunas diferentes, o sistema deve inicialmente mostrar apenas algumas colunas, mas permitir que o cliente aumente o número de colunas até o máximo de trinta. (para melhor compreensão da tabela de traços distintivos ver ficha terminológica original no anexo 1).

- 2.2.3.** A partir da visualização da tabela de traços distintivos, o administrador preencherá os dois campos restantes conceito final e definição.

Obs: O cliente novamente gostaria que o conceito pudesse ser visualizado durante a construção da tabela e que a tabela pudesse ser visualizada durante o preenchimento dos campos conceito final e definição, pois isso ajudaria os profissionais da área a desempenhar o seu trabalho. Sugeri que usássemos uma estrutura parecida com a página anterior, ou seja, divisão em duas áreas uma com os Conceitos e outra com a tabela, incluindo a funcionalidade que o texto da área de conceito seria rolado automaticamente para Contexto desejado quando o administrador clicasse em um dos números da tabela.

- 2.2.4.** Permitir a inserção de um conjunto de informações chamada enciclopédico formado pelas seguintes informações: definição, artigo, fonte, foto, vídeo e imagens.

Obs: serão armazenados apenas referenciais (links) para os recursos multimídia (sons, vídeos, imagens) apresentados pelo site.

3. Atualização das fichas terminológicas já inseridas.
4. Remoção de fichas existentes.

Área de consultas:

1. O sistema deve permitir que o usuário faça consultas a fontes internas e externas.
2. Consultas internas.

Obs: Consultas a fontes internas são consultas cujo resultado é um verbete construído a partir das fichas terminológicas existentes no sistema.

- 2.1.** O sistema deve permitir que o usuário entre com o nome do texto e escolha o idioma desejado no caso de uma consulta monolíngüe (a escolha do idioma é necessária porque pode haver termos que tem o mesmo nome nos dois idiomas, um exemplo é o termo USB). O idioma pode ser português, inglês ou ambos no caso de uma consulta bilíngüe.

- 2.2.** O resultado da consulta será um verbete que pode ser visualizado de 3 formas: TOTAL, TRADUTOR e MODULAR. (Ver anexo 2 – Projeto Modular).

- 2.2.1.** Consulta Modular:

- 2.2.1.1.** O sistema deve permitir que o usuário possa alterar o CONTEÚDO do resultado da consulta de forma a exibir apenas os campos do seu interesse, possibilitando assim que o usuário construa o seu verbete de acordo com a sua necessidade (Para melhor compreensão ver slide – 4 do anexo 2 e/ou o arquivo resposta-email3_16-08-2006).

Obs: A consulta modular é o coração do sistema, pois é esse método de consulta que é proposto na tese do cliente. O cliente imagina o sistema de forma bastante dinâmica, com campos “aparecendo” e “desaparecendo” do corpo do verbete, ou seja, ele não imagina o usuário tendo que apertar o botão de submissão de consulta toda vez que um campo for somado ou subtraído ao verbete.

2.2.1.2. O sistema deve exibir o número total de exemplos ao usuário e permitir que o usuário escolha o número de exemplos que deseja exibir por página, e ter mecanismos de navegação entre os exemplos.

2.2.2. Consulta Tradutor

2.2.2.1. O sistema devera exibir um verbete formado por um conjunto de campos especificado pelo cliente, esse conjunto de campos foi determinado através de uma pesquisa que avaliou quais os campos mais utilizados por tradutores profissionais durante a execução de uma tradução.

2.2.2.2. O sistema deve exibir apenas 3 exemplos neste modo de consulta.

2.2.3. Consulta Total

2.2.3.1. O sistema deverá exibir o verbete formado por todos os campos disponíveis.

2.2.3.2. O sistema deve exibir apenas 3 exemplos neste modo de consulta.

2.3. O sistema deve exibir o nome do campo quando o usuário deslizar o mouse sobre o valor do campo (parecido com aqueles sistemas de ajuda que ao deslizar o ponteiro do mouse por cima de um ícone aparece uma pequena mensagem explicando qual a função do ícone).

2.4. O sistema deve exibir o valor de cada campo com uma cor específica.

2.5. Na consulta bilíngüe o sistema deve apresentar dois verbetes, o verbete para o termo pesquisado e o verbete do termo equivalente (ver glossário) lado-a-lado com a estrutura de exibição paralela e simétrica. (para melhor compreensão ver Slide-3 do anexo 2) .

2.6. Os campos termo equivalente e remissivos são preenchidos por termos para os quais existiram fichas terminológicas no sistema. O sistema deve permitir que o cliente faça buscas recursivas, ou seja, deve permite que ao clicar em um item do campo termo equivalente ou remissivos o usuário obtenha o verbete desses termos (como se originalmente tivesse feito uma busca para os mesmos).

3. Consultas externas.

Obs: Consultas a fontes externas são consultas cujo resultado é pesquisados em sistemas externos (outros web sites).

3.1. O sistema deve possibilitar a submissão de consultas às ferramentas do corpus. (um site desenvolvido no NILC que implementa algumas ferramentas lingüísticas)

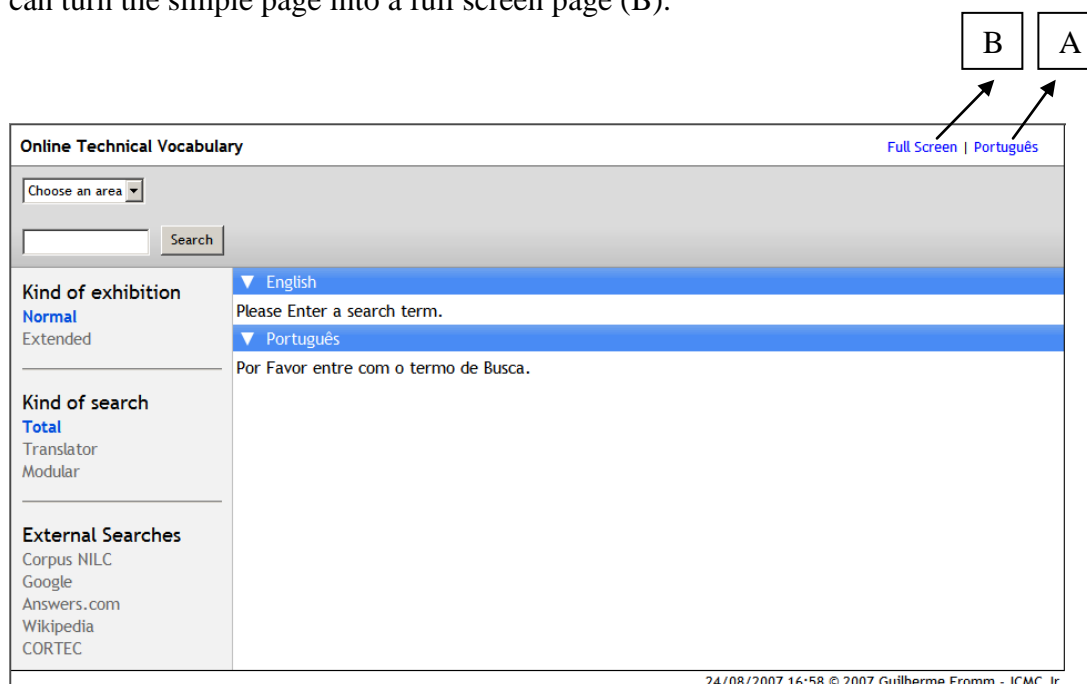
3.2. O sistema deve possibilitar a submissão de consultas a outros sites que disponibilizam ferramentas de tradução. Os sites utilizados serão especificados pelo cliente.

APÊNDICE F. Ajuda online em inglês.

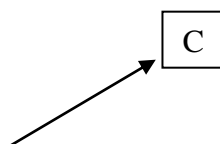
Online Help

Site presentation and search way

You can choose the language you wish to view the site in: English or Portuguese. Just click on the upper right hand corner to change the language (A). In the same corner, you can turn the simple page into a full screen page (B).



To start the search, choose a technical area from the menu in the upper left hand corner (C).



When you select an area, you can go further down into four more sub-fields (its internal division) for a more specific search.

Search screen

On the left side of the screen you find the view menus: kind of exhibition, kind of search and external search.

On the right side, the entry is shown in English and Portuguese.

Kind of exhibition: normal and extended

The kinds of exhibition represent two different moments in dictionary history:

1. Normal exhibition, which shows the entry in the same way as paper dictionaries.

Online Technical Vocabulary Full Screen | Português

Computing Choose an area

Search

Kind of exhibition
Normal
Extended

Kind of search
Total
Translator
Modular

External Searches
Corpus NILC
Google
Answers.com
Wikipedia
CORTEC

▼ English
[Go back to search results](#)

processor. *Hardware. n.m/f.s.* flat square sliver of silicon device composed of circuits etched on its surface; it controls the computer and execute programs; also known as CPU (central processing unit). Ex.: The "brain" of any computer is the central processing unit (CPU). This component does all the calculations and performs 90 percent of all the functions of a computer. There are many different types of processors for computers so many, in fact, that you will learn about them later in this chapter in the section "Identifying Purposes and Characteristic of Processors." . *Synonyms:* CPU. *Hyponyms of:* computer. *See Also:* **CPU.** *Corpus: Frequency order position:* (138); *Term number of occurrences:* (772). **Encyclopedic Information:** A central processing unit (CPU), or sometimes simply processor, is the component in a digital computer that interprets computer program instructions and processes data. CPUs provide the fundamental digital computer trait of programmability, and are one of em: CPU - [Witipedia](#)

▼ Português
[Voltar ao resultado da busca](#)

processador. (CPU, UCP). *Hardware. s.m.s.* tipo de chip, parte principal do computador, também conhecido como CPU (Central Processing Unit); executa instruções (processos) e dados (através da UC - unidade de controle) e operações lógicas e aritméticas (através da UAL - unidade aritmética e lógica).. Ex.: Processador (UCP). A UCP é a parte principal do computador responsável pelo processamento e execução de programas armazenados na memória principal.. *Sinônimos:* CPU, UCP. *Hipônimos de:* computador. *Hiperônimos de:*

2. Extended exhibition: the new dictionaries, in electronic format, present each item of the entry description in different lines. Note, in the example below, that not all the fields are filled in. This happens because the information extracted from the *corpus* is not enough to complete all the available fields.

In English

▼ English
[Go back to search results](#)

processor

flat square sliver of silicon device composed of circuits etched on its surface; it controls the computer and execute programs; also known as CPU (central processing unit)

Acronym: *Nothing found*

Part of Speech: noun

Gender: neutral

Number: singular

Corpus

- **Frequency order position:** 138
- **Term number of occurrences:** 772

Ontology: [Computing](#) > [Hardware](#)

Morphosyntactic Variants: *Nothing found*

Synonyms: CPU

Antonyms: *Nothing found*

Synonyms: CPU

Antonyms: *Nothing found*

Hyponyms of: computer

Co-hyponyms: *Nothing found*

Hypernyms of: *Nothing found*

Examples

1. The "brain" of any computer is the central processing unit (CPU). This component does all the calculations and performs 90 percent of all the functions of a computer. There are many different types of processors for computers so many, in fact, that you will learn about them later in this chapter in the section "Identifying Purposes and Characteristics of Processors."
2. In essence, a CPU is a flat square sliver of silicon with circuits etched on its surface. This chip is linked to connector pins and the whole contraption encased some form of packaging - either ceramic or plastic - with pins running along the flat underside or along one edge. The CPU package is connected to a motherboard via some form of CPU interface, either a slot or a socket
3. The speed of the processor is generally described in clock frequency (MHz or GHz).

[1 more...](#)

Encyclopedic Information



A central processing unit (CPU), or sometimes simply processor, is the component in a digital computer that interprets computer program instructions and processes data. CPUs provide the fundamental digital computer trait of programmability, and are one of em: *CPU* - [Wikipedia](#)

See Also

[CPU](#)

23/08/2007 15:44 © 2007 Guilherme Fromm - ICMC Jr.

In Portuguese

▼ Português

[Voltar ao resultado da busca](#)

processador

tipo de chip, parte principal do computador, também conhecido como CPU (Central Processing Unit); executa instruções (processos) e dados (através da UC - unidade de controle) e operações lógicas e aritméticas (através da UAL - unidade aritmética e lógica).

Abreviação/Acrônimo: CPU, UCP

Categoria Gramatical: substantivo

Gênero: masculino

Número: singular

Córpus

- **Posição na Ordem de Frequência:** 67
- **Nº de Ocorrências do termo:** 1212

Ontologia: [Computação](#) > [Hardware](#)

Variações Morfosintáticas: *Nada encontrado*

Sinônimos: CPU, UCP

Hipônimos de: computador

Co-hipônimos: chip

Hiperônimos de: UAL (unidade aritmética e lógica); UC (unidade de controle)

Exemplos

1. Processador (UCP). A UCP é a parte principal do computador responsável pelo processamento e execução de programas armazenados na memória principal.
2. UCP

Sua função consiste em coordenar, controlar e realizar todas as operações (execução de instruções) do sistema.

3. UCP

Suas partes principais são as seguintes:

- a unidade aritmética e lógica (UAL): realiza o cálculo real ou o processamento de dados (realiza as operações aritméticas e lógicas).

- a unidade de controle (UC): controla o movimento de dados e instruções dentro e fora da UCP e controla a operação da UAL, de forma adequada e sincronizada.

[3 mais...](#)

Informações Enciclopédicas



O processador é a parte mais fundamental para o funcionamento de um computador. Processadores são circuitos digitais que realizam operações como: cópia de dados, acesso a memórias e operações lógicas e matemáticas. Em: *Processador* - [Wikipedia](#)

Veja Também

[CPU](#) - [UCP](#)

Internal searches: total, translator and modular

The possible searches were divided into three kinds:

Total: this shows all the available fields (see example above).

Translator: this shows the fields which were considered more important by professional translators (note that not all the fields from the previous example are shown here). See the following example:

<p>Kind of exhibition Normal Extended</p> <hr/> <p>Kind of search Total Translator Modular</p> <hr/> <p>External Searches Corpus NILC Google Answers.com Wikipedia CORTEC</p>	<p>▼ English</p> <p>Go back to search results</p> <p>processor</p> <p>flat square sliver of silicon device composed of circuits etched on its surface; it controls the computer and execute programs; also known as CPU (central processing unit)</p> <p>Acronym: <i>Nothing found</i></p> <p>Part of Speech: noun</p> <p>Gender: neutral</p> <p>Number: singular</p> <p>Ontology: Computing > Hardware</p> <p>Morphosyntatic Variants: <i>Nothing found</i></p> <p>Synonyms: CPU</p> <p>Antonyms: <i>Nothing found</i></p> <p>Examples</p> <p>1. The brain of any computer is the central processing unit (CPU). This component does all the calculations and performs 90 percent of all the functions of a computer. There are many different types of processors for computers so many, in fact, that you will learn about them later in this chapter in</p>
--	--

the section **Identifying Purposes and Characteristic of Processors.**

- In essence, a CPU is a flat square sliver of silicon with circuits etched on its surface. This chip is linked to connector pins and the whole contraption encased some form of packaging - either ceramic or plastic - with pins running along the flat underside or along one edge. The CPU package is connected to a motherboard via some form of CPU interface, either a slot or a socket
- The speed of the processor is generally described in clock frequency (MHz or GHz).

[1 more...](#)

See Also
[CPU](#)

Modular: in this kind of search, you choose the fields you want to see. See the following example. We selected, in it, the extended kind of exhibition and deselected (when you click the mouse on the blue fields on the left side, they turn gray) the fields gender, morphosyntactic variants, complete word, *corpus*, examples, encyclopedic information and cross-reference.

Full Screen | Português

Computing ▾ Choose an area ▾

Search

Kind of exhibition
Normal
Extended

Kind of search
Total
Translator
Modular

Pre-definition
Acronym
Part of Speech
Gender
Number
Ontology
Morphosyntactic Variants

Definition
Definition
Complete Word
Corpus

Example
Examples
Encyclopedic Information

Relações
Cross-references
Synonym
Antonym
Hyponyms
Co-hyponyms
Hypernyms

English

[Go back to search results](#)

processor

flat square sliver of silicon device composed of circuits etched on its surface; it controls the computer and execute programs; also known as CPU (central processing unit)

Acronym: *Nothing found*

Part of Speech: noun

Gender: neutral

Number: singular

Ontology: [Computing](#) > [Hardware](#)

Synonyms: CPU

Antonyms: *Nothing found*

Hyponyms of: computer

Co-hyponyms: *Nothing found*

Hypernyms of: *Nothing found*

Português

[Voltar ao resultado da busca](#)

processador

tipo de chip, parte principal do computador, também conhecido como CPU (Central Processing Unit); executa instruções (processos) e dados (através da UC - unidade de controle) e operações lógicas e aritméticas (através da UAL - unidade aritmética e lógica).

Abreviação/Acrônimo: CPU, UCP

Categoria Gramatical: substantivo

External Searches Corpus NILC Google Answers.com Wikipedia CORTEC	Gênero: masculino
	Número: singular
	Ontologia: Computação > Hardware
	Sinônimos: CPU, UCP
	Antônimos: <i>Nada encontrado</i>
	Hipônimos de: computador
	Co-hipônimos: chip
	Hiperônimos de: UAL (unidade aritmética e lógica); UC (unidade de controle)

24/08/2007 17:53 © 2007 Guilherme Fromm - ICMC Jr.

Click on the left side of the screen and build your structure according to your needs!

External searches

Apart from the internal searches available, you can also do external searches on many sites. Here you find a summary of each of them; to access them, just click on the desired one, wait for the insertion box to appear, type the desired term and click <enter> on your keyboard:

Kind of exhibition Normal Extended	Go back to search results
Kind of search Total Translator Modular	<p>processor. <i>Hardware.</i> n.m/f.s. flat square sliver of silicon device composed of circuits etched on its surface; it controls the computer and execute programs; also known as CPU (central processing unit). Ex.: The brain of any computer is the central processing unit (CPU). This component does all the calculations and performs 90 percent of all the functions of a computer. There are many different types of processors for computers so many, in fact, that you will learn about them later in this chapter in the section Identifying Purposes and Characteristic of Processors. Synonyms: CPU. Hyponyms of: computer. See Also: CPU. Corpus: <i>Frequency order position:</i> (138); <i>Term number of occurrences:</i> (772). Encyclopedic Information: A central processing unit (CPU), or sometimes simply processor, is the component in a digital computer that interprets computer program instructions and processes data. CPUs provide the fundamental digital computer trait of programmability, and are one of em: <i>CPU</i> - Wikipedia</p>
External Searches Corpus NILC <input type="text"/> <input type="button" value="Search"/> Google <input type="text"/> <input type="button" value="Search"/> Answers.com <input type="text"/> <input type="button" value="Search"/> Wikipedia <input type="text"/> <input type="button" value="Search"/> CORTEC	<p style="background-color: #e0e0e0; padding: 2px;">▼ Português</p> <p>Voltar ao resultado da busca</p> <p>processador. (CPU, UCP). <i>Hardware.</i> s.m.s. tipo de chip, parte principal do computador, também conhecido como CPU (Central Processing Unit); executa instruções (processos) e dados (através da UC - unidade de controle) e operações lógicas e aritméticas (através da UAL - unidade aritmética e lógica).. Ex.: Processador (UCP). A UCP é a parte principal do computador responsável pelo processamento e execução de programas armazenados na memória principal.. Sinônimos: CPU, UCP. Hipônimos de: computador. Hiperônimos de: UAL (unidade aritmética e lógica); UC (unidade de controle). Co-hipônimos: chip. Veja Também: CPU, UCP. Corpus: <i>Posição na Ordem de Freqüência:</i> (67); <i>Nº de Ocorrências do termo:</i> (1212). Informações Enciclopédicas: O processador é a parte mais fundamental para o funcionamento de um computador. Processadores são circuitos digitais que realizam operações como: cópia de dados, acesso a memórias e operações lógicas e matemáticas. Em: <i>Processador</i> - Wikipedia</p>

25/08/2007 13:33 © 2007 Guilherme Fromm - ICMC Jr.

Corpus NILC: the term you type here will be searched in the general *corpus* of the NILC project (Núcleo Interinstitucional de Linguística Computacional) and it will be shown through a KWIC structure (the term appears in the middle of the page). *Corpus* available in Portuguese only.

Google: the term will be searched in the most used search engine on the Internet.

Answers.com: the term will be searched in this site, a mix of dictionary and encyclopedia, in English.

Wikipedia: the term will be searched in the most famous online encyclopedia. In Portuguese.

CORTEC: site that hosts many technical *corpora*. You can search for wordlists, concordance lines, and n-grams. When you click on this link, a new page will open. *Corpora* available in Portuguese and English.

HAVE FUN!

Guilherme Fromm

guifromm@uol.com.br

NOTE: if you have any doubts about terms used to describe words in this site (like *Hyperonym*), search the Linguistics area.

APÊNDICE G. Questionário com o público-alvo/professores. Versão completa.

Pesquisa – Tese Guilherme Fromm

Páginas 1/2: Alunos

1. Há quantos semestres você cursa Tradução?
2. Você leu a ajuda online antes de trabalhar no site?
() Sim (vá para a pergunta 3) () Não (vá para a pergunta 4)
3. A ajuda online auxiliou você a trabalhar com o site?
() Sim () Não
4. Se você não usou a ajuda online para trabalhar com o site, é por que:
você tem experiência com sites que oferecem dicionários online;
é mais rápido aprender “fuçando” do que lendo instruções;
é fácil de aprender a usar o site.
5. Você já conhecia termos como *Hiperônimo, Hipônimo, Co-hipônimo*?
() Sim (vá para a pergunta 8) () Não (vá para a pergunta 6)
6. Se você não conhecia os termos citados na pergunta 5, você chegou a procurá-los no próprio site, na seção de Linguística?
() Sim () Não
7. Se não conhecia os termos da questão 5, você chegou a ficar curioso sobre o significado deles?
() Sim () Não
8. Qual tipo de visualização é mais fácil de trabalhar?
() Normal () Descritiva
9. Dentre os três tipos de pesquisa (total, tradutor, modular), qual você achou mais fácil para trabalhar?
() Total () Tradutor () Modular
10. Quanto tempo você demorou para entender como funciona o tipo de consulta “modular”?
 - a. Menos de 1 minuto
 - b. De 1 a 3 minutos

De 3 a 6 minutos

- d. De 6 a 10 minutos
- e. De 10 a 15 minutos
- f. De 15 a 20 minutos
- g. Mais de 20 minutos

11. Quanto tempo você demorou para entender como funciona o site como um todo?

- a. De 1 a 3 minutos
- b. De 3 a 6 minutos
- c. De 6 a 10 minutos
- d. De 10 a 15 minutos
- e. De 15 a 20 minutos
- f. Mais de 20 minutos
- g. Mais de meia hora
- h. Mais de uma hora

12. Você considera os sites de consulta externa (lado esquerdo inferior) bons para pesquisas extras?

- Sim Não

13. Os termos técnicos na área de computação disponíveis no site foram suficientes para a sua tradução?

- Sim (vá para a pergunta 15) Não (vá para a pergunta 14)

14. Quais palavras constavam do texto a ser traduzido e não foram encontradas no próprio site?

15. O site o ajudou a entender as diversas possibilidades de exibição de um termo técnico?

- Sim Não Em termos (explique porque)

16. De modo geral, como você avalia o site tendo em vista a ajuda que oferece ao aprendiz de tradução?

- a. ruim
- b. regular
- c. bom
- d. ótimo

17. Gostaria de tecer algum comentário sobre o site?

Pesquisa – Tese Guilherme Fromm

Carta convite para os professores

Olá a todos,

para quem não me conhece, sou orientando da profa. Dra. Stella Tagnin (FFLCH/USP) e estou finalizando a minha tese de doutorado. Foi desenvolvido, para a mesma, um site e, agora, necessito fazer uma pesquisa com o público-alvo: os aprendizes de tradução. Preciso da ajuda de todos que trabalhem com prática de tradução inglês/português em laboratório (com acesso à Internet) para poder realizar essa pesquisa. Esse primeiro e-mail é só para sondagem: envio mais informações na sequência para aqueles que puderem ajudar. Se vocês não trabalham com prática de tradução, mas conhecem quem o faça, gostaria que vocês me indicassem nomes.

Obrigado, desde já, pela ajuda,

Guilherme Fromm

Página de instruções e pesquisa para os professores

Caros professores, antes de aplicar a pesquisa, por favor, leiam o seguinte parágrafo para os alunos. Muito obrigado.

Prezados alunos. Meu nome é Guilherme Fromm e sou doutorando em Estudos Lingüísticos e Literários em Inglês pela USP. Estou desenvolvendo um site que visa ajudar alunos dos cursos de Tradução a aprender a manusear sua principal ferramenta de trabalho: o dicionário. Gostaria que vocês me ajudassem a avaliar o sistema por meio de um teste prático: a tradução de um parágrafo de uma área de especialidade. Depois da tradução do parágrafo, por favor, respondam ao questionário que seu professor (sua professora) distribuirá e devolvam ao mesmo. Muito obrigado. Se quiserem enviar sugestões, sintam-se à vontade: guifromm@uol.com.br (por na lousa).

Instruções para a aplicação

- Por favor, indique aos alunos que existe uma ajuda online que eles podem consultar, ou não, antes de iniciarem a tradução.
- Explique que o site não é comercial e, portanto, somente alguns termos (que constam no texto a ser traduzido) foram inseridos até o momento.
- **NÃO** ajudem o aluno com a tradução nem expliquem como o site funciona.

Estas perguntas são para você, professor.

1. Quantos alunos realizaram o teste?

2. Algum aluno encontrou dificuldade na leitura da ajuda online, ou seja, ele/ela lhe fez alguma pergunta? Se sim, quantos alunos?

sim não

Quantidade de alunos: _____

3. Foi fácil para você entender como funciona a estrutura do site?

sim não

4. De modo geral, como você avalia o site tendo em vista a ajuda que oferece ao aprendiz de tradução?

a. ruim

b. regular

c. bom

d. ótimo

5. Você usaria esse site para ensinar seus alunos a trabalhar com terminologia bilíngüe e manuseio de vocabulários eletrônicos?

Sim

Não

Em termos (explique porque)

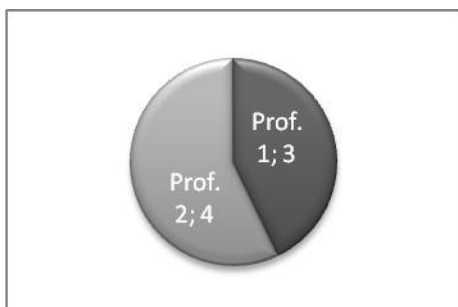
Por favor, sintam-se à vontade para mandar comentários e sugestões:
guifromm@uol.com.br

APÊNDICE H. Resultados completos da pesquisa com o público-alvo.

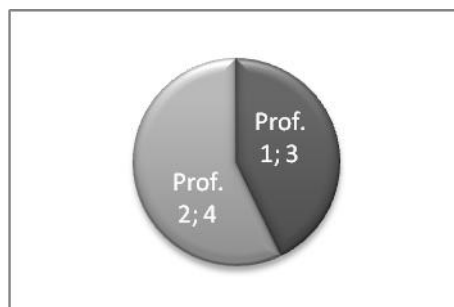
Pergunta 1

Quantos alunos realizaram o teste?

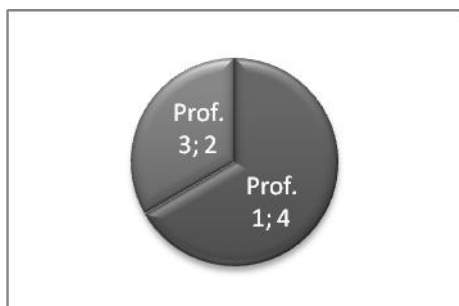
Semestre 1



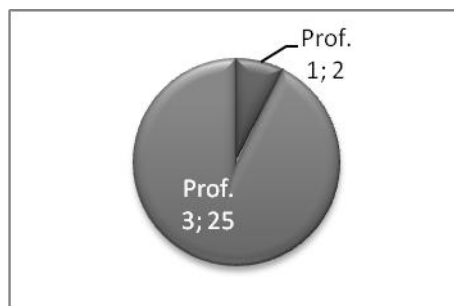
Semestre 2



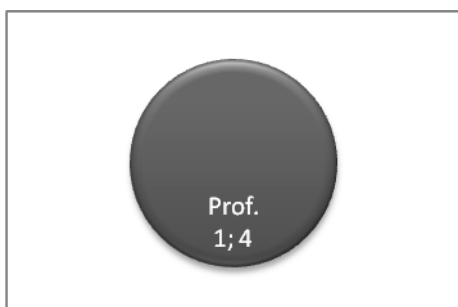
Semestre 3



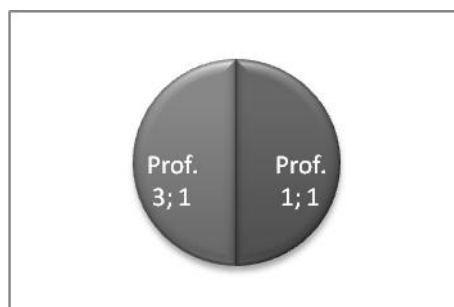
Semestre 4



Semestre 6



Semestre 8

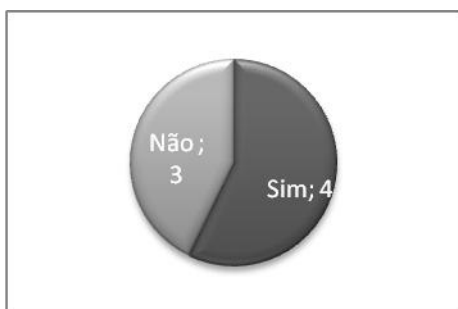


Pergunta 2

Você leu a ajuda online?

Semestre 1

Semestre 2



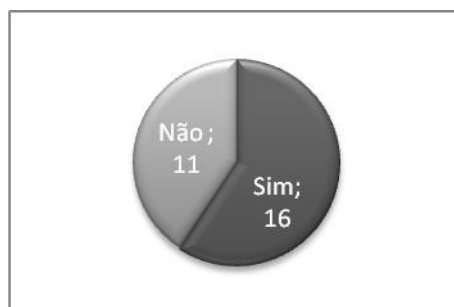
Semestre 3



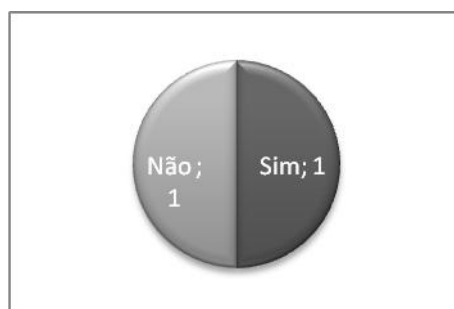
Semestre 4



Semestre 6



Semestre 8



Pergunta 3

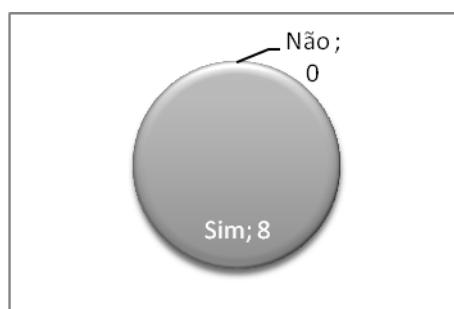
A ajuda online o/a auxiliou?

Semestre 1

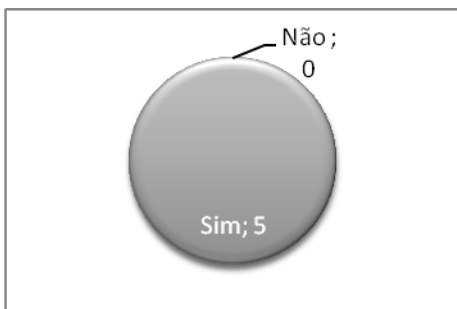


Semestre 3

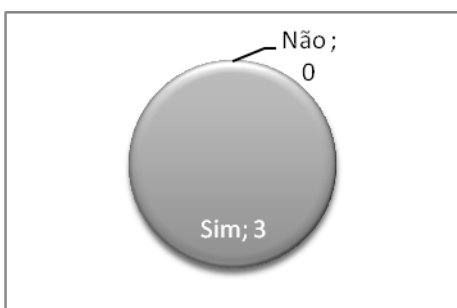
Semestre 2



Semestre 4



Semestre 6



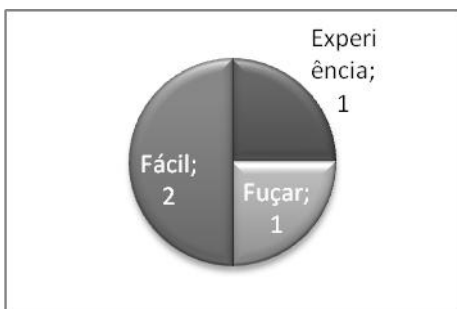
Semestre 8



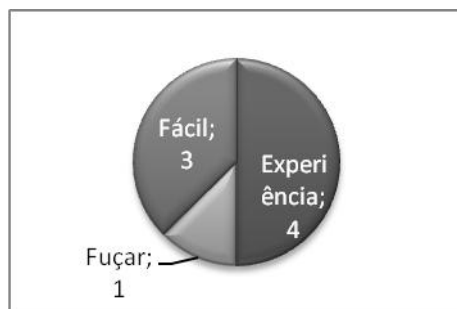
Pergunta 4

Não usou a ajuda online porque:

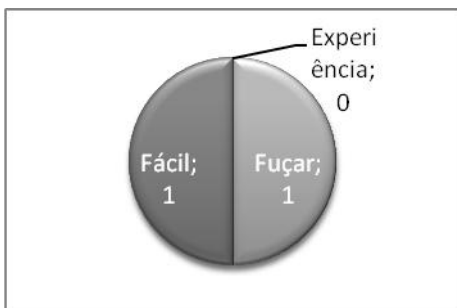
Semestre 1



Semestre 2



Semestre 3

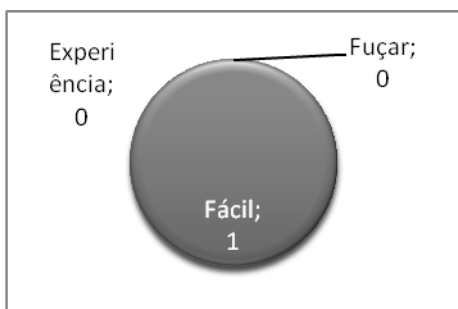


Semestre 4



Semestre 6

Semestre 8



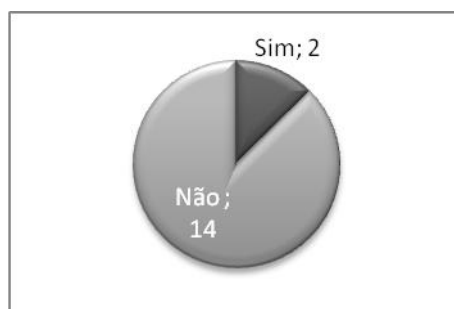
Pergunta 5

Você já conhecia termos como Hiperônimo, Hipônimo, Co-hipônimo?

Semestre 1



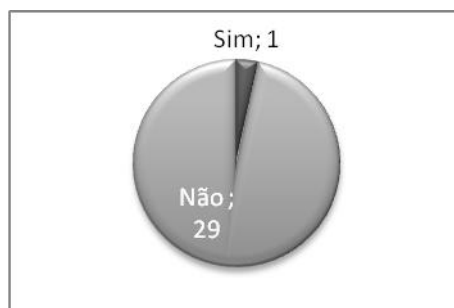
Semestre 2



Semestre 3



Semestre 4



Semestre 6



Semestre 8



Pergunta 6

Se você não conhecia os termos citados na pergunta 5, você chegou a procurá-los no próprio site, na seção de Lingüística?

Semestre 1



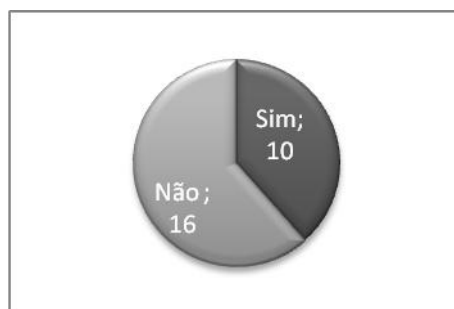
Semestre 2



Semestre 3



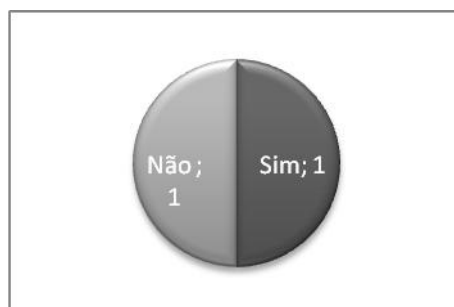
Semestre 4



Semestre 6



Semestre 8

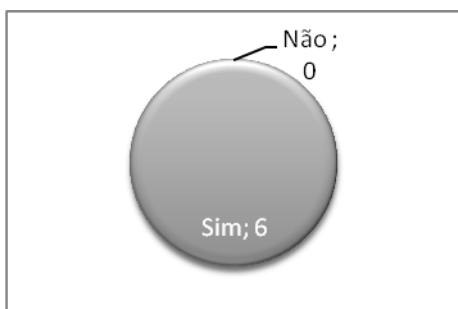


Pergunta 7

Se não conhecia os termos da questão 5, você chegou a ficar curioso sobre o significado deles?

Semestre 1

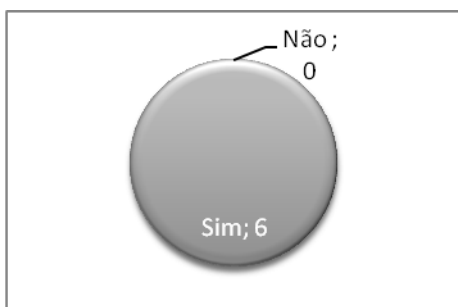
Semestre 2



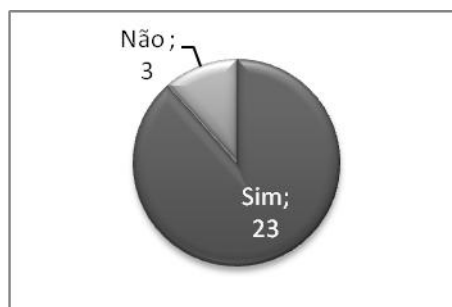
Semestre 3



Semestre 4



Semestre 6



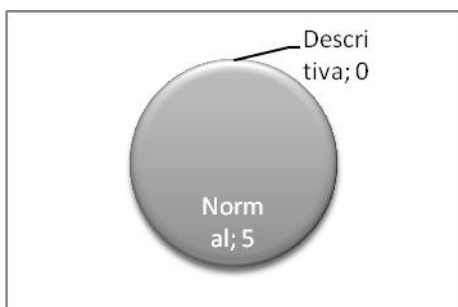
Semestre 8



Pergunta 8

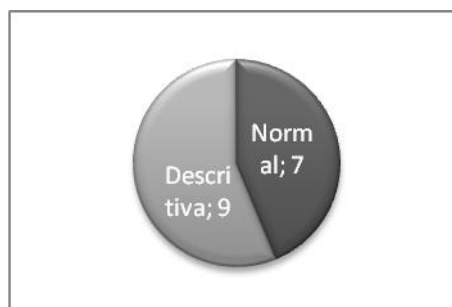
Qual tipo de visualização é mais fácil de trabalhar?

Semestre 1

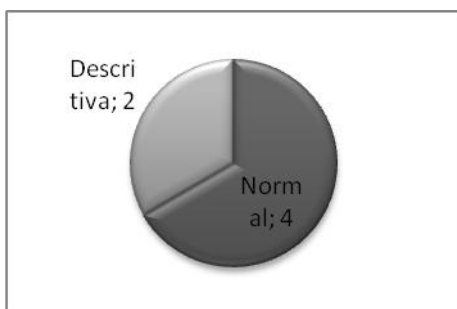


Semestre 3

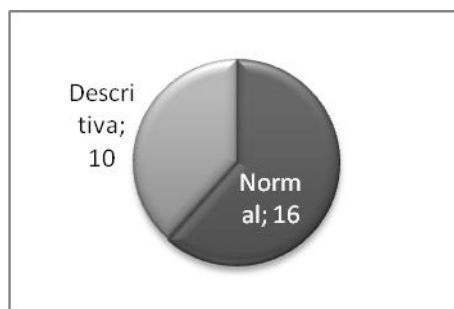
Semestre 2



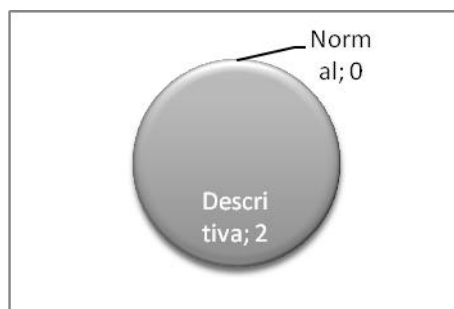
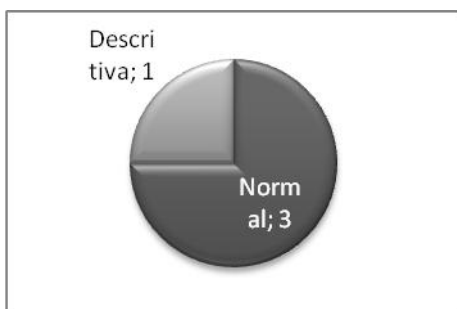
Semestre 4



Semestre 6



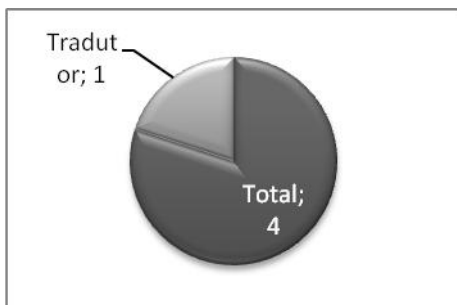
Semestre 8



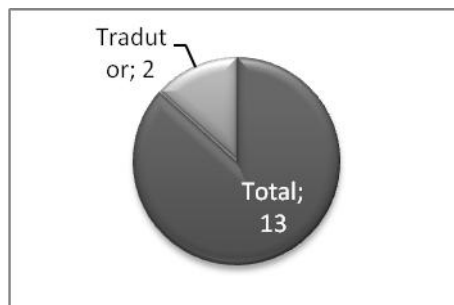
Pergunta 9

Dentre os três tipos de pesquisa (total, tradutor, modular), qual você achou mais fácil para trabalhar?

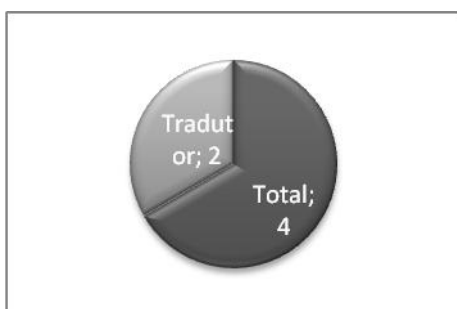
Semestre 1



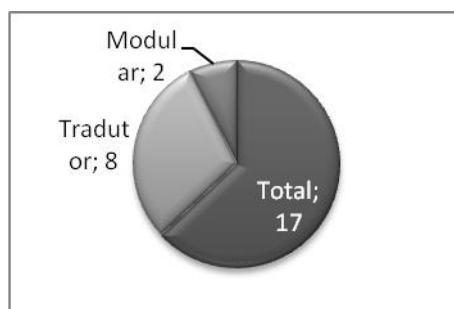
Semestre 2



Semestre 3

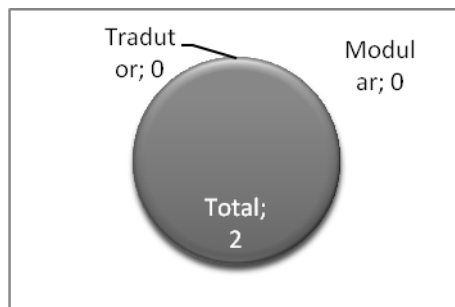
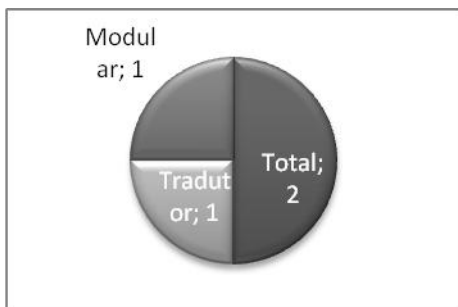


Semestre 4



Semestre 6

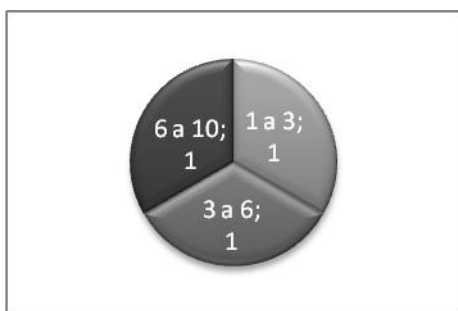
Semestre 8



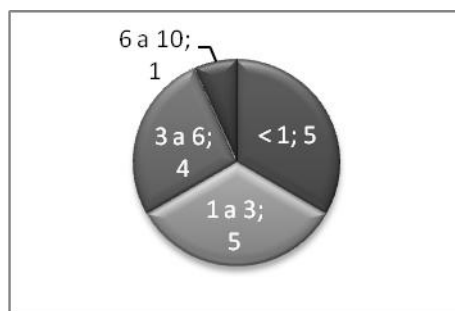
Pergunta 10

Quanto tempo (em minutos) você demorou a entender como funciona o tipo de consulta “modular”?

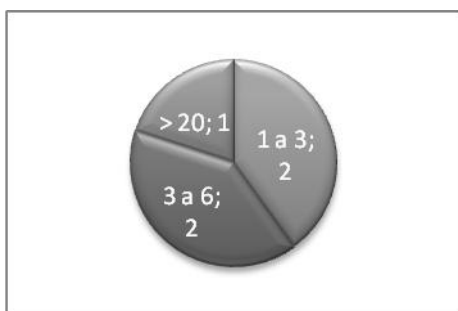
Semestre 1



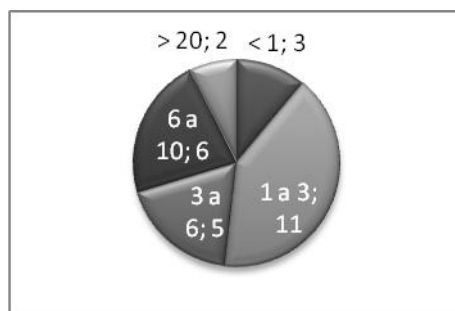
Semestre 2



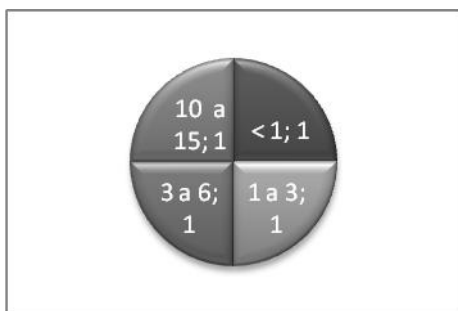
Semestre 3



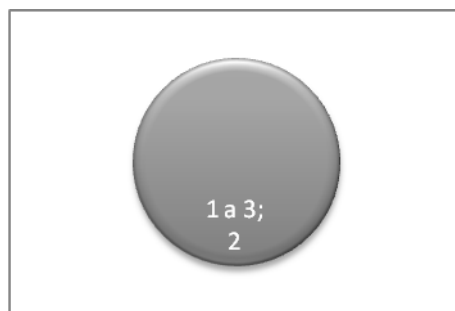
Semestre 4



Semestre 6



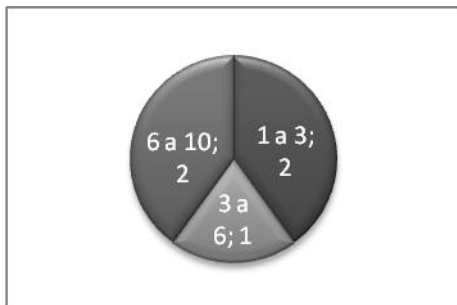
Semestre 8



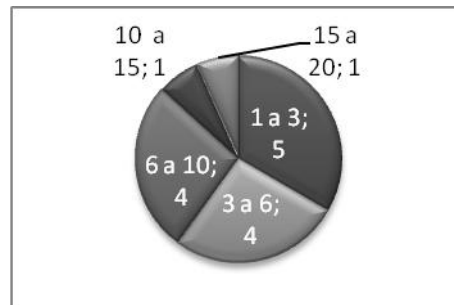
Pergunta 11

Quanto tempo (em minutos) você demorou a entender como funciona o site como um todo?

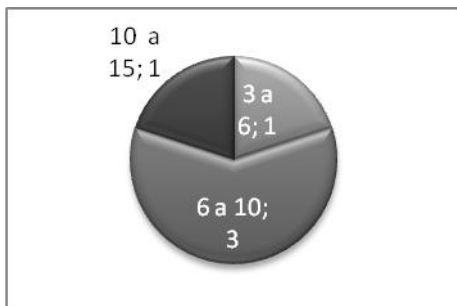
Semestre 1



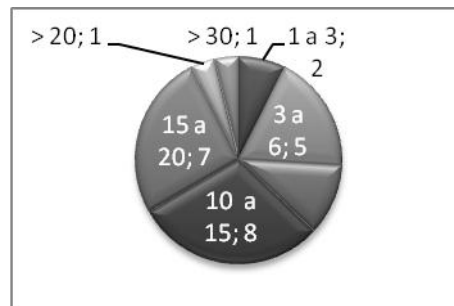
Semestre 2



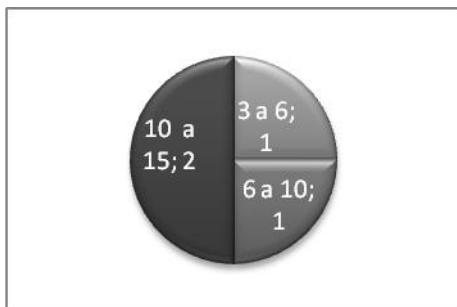
Semestre 3



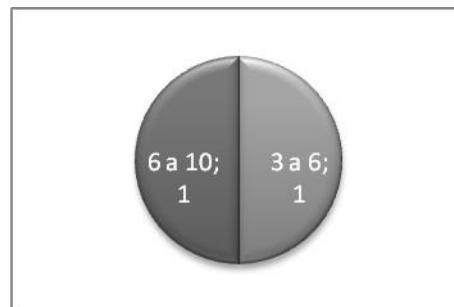
Semestre 4



Semestre 6



Semestre 8



Pergunta 12

Você considera os sites de consulta externa (lado esquerdo inferior) bons para pesquisas extras?

Semestre 1

Semestre 2



Semestre 3



Semestre 4



Semestre 6



Semestre 8



Pergunta 13

Os termos técnicos na área de computação disponíveis no site foram suficientes para a sua tradução?

Semestre 1



Semestre 3

Semestre 2



Semestre 4



Semestre 6



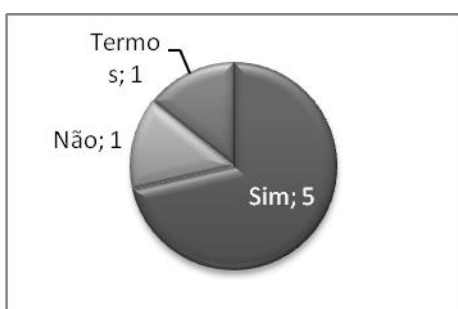
Semestre 8



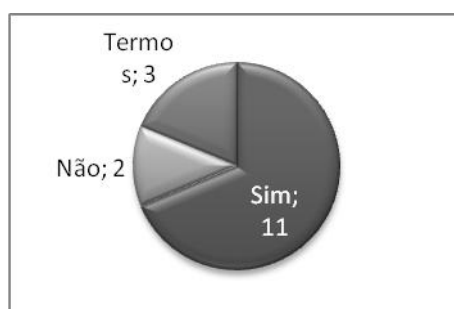
Pergunta 15

O site o ajudou a entender as diversas possibilidades de exibição de um termo técnico?

Semestre 1



Semestre 2

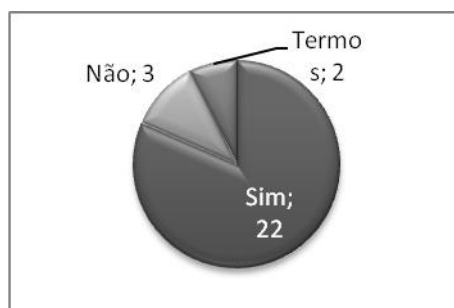


Semestre 3



Semestre 6

Semestre 4



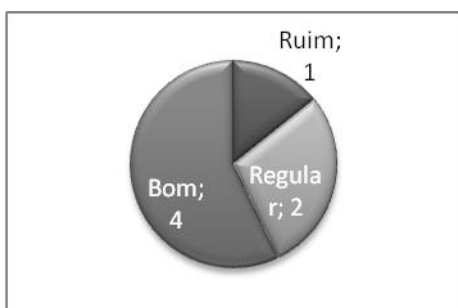
Semestre 8



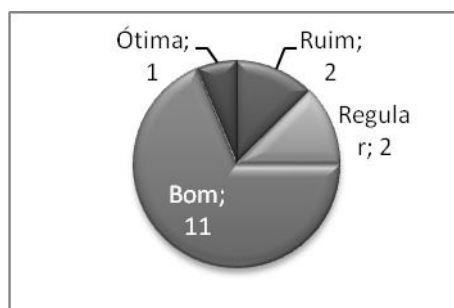
Pergunta 16

De modo geral, como você avalia o site tendo em vista a ajuda que oferece ao aprendiz de tradução?

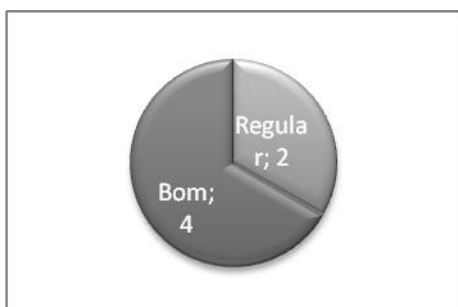
Semestre 1



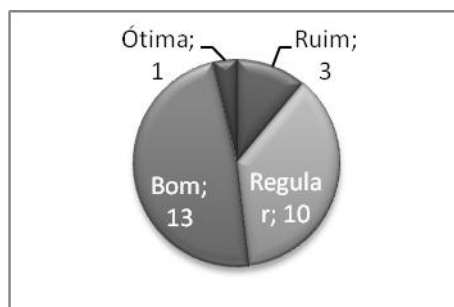
Semestre 2



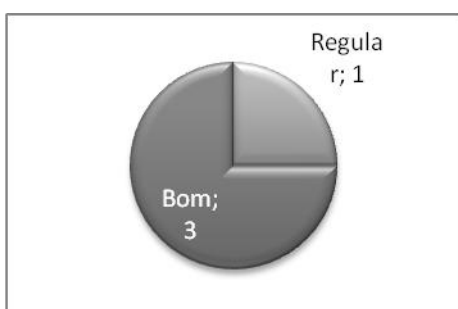
Semestre 3



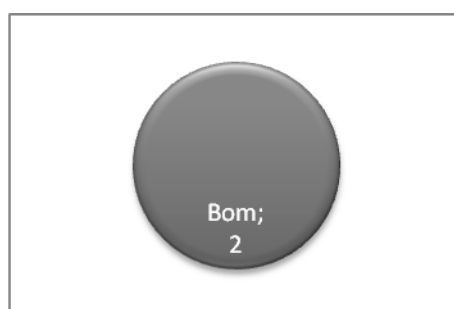
Semestre 4



Semestre 6



Semestre 8



ANEXO

ANEXO A. Organização das possíveis entradas de um dicionário bilíngüe.

	Data type	Mode	Information Content	Function	User
1	lemma forms	SL	lexical form(s) of the HW ¹ /subheadword	helps user find the information being sought	enc ³ SL dec ³ TL
2	phonetic transcription	code IPA	how the HW is pronounced	helps the non-native speaker pronounce the word correctly	enc TL
3	grammar form	code	part of speech, gender, etc. of HW	helps user find the information being sought	enc SL dec TL
4	sense or /subsense + counter	alph /num code	this is a distinct sense or subsense of the HW	helps user find the information being sought	enc SL dec TL
5	grammar usage item	SL + TL	grammatical complementation of HW in this sense & its translation	helps TL user use SL item correctly helps SL user identify the sense of the HW	enc TL ⁴ enc SL
6	TL equivalent	TL	this is TL equivalent of HW in this sense	helps TL user understand helps both users translate	dec TL enc SL
7	gloss	TL	an explanation of HW in this sense	helps TL user understand helps both users translate	dec TL enc SL
8	Typical example ⁵ + translation	SL + TL	this is how the HW in this sense is typically used & translated	helps SL user identify the sense of the HW reassures SL user trying to translate SL item helps TL user use SL item correctly	enc SL enc TL
9	problematic example ⁶ + translation	SL + TL	the HW in this context has a specific TL equivalent	helps SL user identify the sense of the HW helps SL user avoid translating error	enc SL
10	idiomatic example ⁷ + translation	SL + TL	the HW and context have this specific TL equivalent	helps TL user understand helps both users translate	dec TL enc SL
11	diatechnical label	code	HW in this sense belongs to this semantic domain of (Music, Science etc.)	helps both users select correct TL equivalent helps SL user identify the sense of the HW	dec TL enc SL
12	stylistic label	code	using the SL or TL item in this sense is in (literary etc.) style	helps both users translate helps TL user understand helps SL user identify the sense of the HW	enc SL dec TL
13	register label	code	using the SL or TL item in this sense is in (informal etc.) register	helps both users translate helps TL user understand helps SL user identify the sense of the HW	enc SL dec TL
14	diatopic label	code	the SL or TL item in this sense belongs to X regional variety of the language	helps both users translate helps TL user understand helps SL user identify the sense of the HW	enc SL dec TL
15	diachronic label	code	the SL or TL item in this sense is (obsolete / old-fashioned etc.)	helps both users translate helps TL user understand helps SL user identify the sense of the HW	enc SL dec TL
16	evaluative label	code	using the SL or TL item in this sense is (pejorative etc.)	helps both users translate helps TL user understand helps SL user identify the sense of the HW	enc SL dec TL
17	sense indicator	SL	synonym or paraphrase of HW in this sense / other brief sense clue	helps SL user identify the sense of the HW	enc SL
18	collocators	SL	typical subjects / objects of HW verbs, nouns modified by HW adjectives etc.	helps both users translate helps SL user identify the sense of the HW	enc SL dec TL
19	collocators	TL	typical subjects / objects of TL equivalent verbs, nouns modified by TL equivalent adjectives etc.	helps both users translate	enc SL dec TL
20	cross-reference	SL	this other definiendum is relevant to the HW in this sense	helps users find the information being sought	enc SL dec TL

Figura 1. Organização das possíveis entradas de um dicionário bilíngüe (Atkins, 2003).