

UNIVERSIDADE FEDERAL DE SANTA CATARINA

LARISSA LEPRE

**A ELABORAÇÃO DE GLOSSÁRIOS BILÍNGÜES PARA A INTERPRETAÇÃO DE
TEXTOS EM INGLÊS COM BASE EM UM *CORPUS* PARALELO.**

**Florianópolis
2007**

LARISSA LEPRE

**A ELABORAÇÃO DE GLOSSÁRIOS BILÍNGÜES PARA A INTERPRETAÇÃO DE
TEXTOS EM INGLÊS COM BASE EM UM *CORPUS* PARALELO.**

Dissertação de mestrado apresentada ao
Programa de Pós-Graduação em Estudos da
Tradução, da Universidade Federal de Santa
Catarina, como parte dos requisitos para a
obtenção do título de Mestre em
Estudos da Tradução.

Área de concentração: Lexicografia,
Tradução e Ensino de Línguas.

ORIENTADOR: Prof. Dr. Marco Antonio E. da Rocha.

Florianópolis
2007

Ficha Catalográfica

Lepre, Larissa

A Elaboração de glossários bilíngües para a interpretação de textos em inglês com base em um *corpus* paralelo / Larissa Lepre. - Florianópolis: UFSC / Nome ou sigla da Unidade, 2007.
67 f.; 31 cm.

Orientador: Marco Antonio E. da Rocha
Dissertação (mestrado) – Universidade Federal de Santa Catarina, UFSC, Programa de Pós-Graduação em Estudos da Tradução, 2007.

1. Tradução. 2. Lexicografia – Dissertação. I. Rocha, Marco. II. Universidade Federal de Santa Catarina. III. Título.

Larissa Lepre

**A ELABORAÇÃO DE GLOSSÁRIOS BILÍNGÜES PARA A INTERPRETAÇÃO DE
TEXTOS EM INGLÊS COM BASE EM UM *CORPUS* PARALELO.**

Esta Dissertação foi julgada adequada à obtenção do título de Mestre em XXXXXX e aprovado em sua forma final pelo Curso de CCCCCCCCCCCCC, da Universidade Federal de Santa Catarina.

Florianópolis, 10 de dezembro de 2007.

Professor e orientador: Marco Antonio E. da Rocha, Dr.

DEDICATÓRIA

Dedico este trabalho a meus pais, Antônio Salvador Lepre e Rosa Nicolini Lepre, por terem me dado suporte durante todos esses anos de estudos e, principalmente, pela inspiração de vida e demonstração de força interior, as quais me incentivaram na busca de novos desafios.

AGRADECIMENTOS

Ao Programa de Pós-Graduação em Estudos da Tradução – PGET, em especial ao meu orientador, Prof. Dr. Marco Rocha, pela oportunidade de realização de trabalhos em minha área de pesquisa.

Ao corpo docente, pela dedicação e excelente contribuição por meio dos cursos ministrados.

Aos colaboradores da PGET, pela dedicação em nos prestar o melhor serviço, nos momentos de necessidade.

Substantivos

Faca é faca

Pão é pão

Fome é fome

Amor é amor

Estranho desígnio das coisas

De serem exatamente elas

Quando as olhamos sem paixão

(Tanussi Cardoso)

RESUMO

Esta dissertação foi produzida com o objetivo de apresentar uma metodologia para a elaboração de um glossário bilíngüe com base em um *corpus* de domínio técnico. O objetivo do glossário, produto desta metodologia, é dar suporte a acadêmicos e pesquisadores que se interessem pela tradução especializada, pelo ensino de línguas e a todos os que se interessam pela interpretação de termos técnicos em geral. Dessa maneira, apresenta-se aqui a metodologia, junto com os exemplos extraídos de *corpora* fundamentados na Fisiologia do Exercício, uma das disciplinas inseridas no ensino da Educação Física.

PALAVRAS-CHAVE: Tradução. Terminologia. Lingüística de *Corpus*. Lexicografia. Fisiologia do Exercício.

ABSTRACT

This paper aims to describe a methodology for the construction of a bilingual glossary based on a domain-specific *corpus*. The bilingual glossary, result of said methodology, would have the purpose of giving support to academic students, specialized translation, language teaching, and to all researches who is interested in the interpretation of technical terms in general. The theoretical basics, the methodology itself and examples of extraction of data from the physical education domain *corpus* are presented.

KEYWORDS: Translation. Terminology. Corpus Linguistics. Lexicography. Physical Education.

Lista de Figuras

Figura 1 - Tela inicial do *WordSmith*

Figura 2 - As trinta primeiras palavras-chave encontradas no *corpus* em inglês

Figura 3 – Ferramenta *Concord* do *Wordsmith*:

Figura 4 - Tela principal do Concordanciador do *WordSmith*

Figura 5 - Janela de seleção de *corpus*

Figura 6 - Janela de *corpora* selecionados

Figura 7 - Exemplo de busca da concordância de um item lexical

SUMÁRIO

| | |
|--|-----------|
| 1 INTRODUÇÃO | 12 |
| 1.1 Unificação de termos | 12 |
| 1.2 Lingüística de <i>corpus</i> | 13 |
| 1.3 A falta de glossários técnicos na Educação Física | 15 |
| 2 REVISÃO DA LITERATURA | 20 |
| 2.1 Sobre os conceitos que permeiam a Lingüística de <i>Corpus</i> | 20 |
| 3 JUSTIFICATIVA | 25 |
| 3.1 Problema a ser estudado | 25 |
| 3.2 Relevância da proposta para os estudos de termos da Educação Física | 26 |
| 4 OBJETIVOS | 27 |
| 4.1 Glossário ou dicionário | 27 |
| 4.2 O conceito de tradução | 29 |
| 4.3 Sobre a tradução | 30 |
| 5 METODOLOGIA | 33 |
| 5.1O uso do <i>WordSmith</i> | 33 |
| 5.2 Tipologia e variação dos textos a incluir | 35 |
| 5.3 Coleta do material a ser estudado | 36 |
| 5.4 Dimensão dos <i>corpora</i> selecionados | 38 |
| 5.5 Procedimentos utilizados na elaboração do glossário | 40 |
| 5.6 Características sobre traduções | 50 |
| 6 RESULTADOS | 53 |
| 6.1 Organização das entradas | 53 |
| 7 CONSIDERAÇÕES FINAIS | 67 |
| REFERÊNCIAS | 69 |

1 INTRODUÇÃO

1.1 Unificação de termos

No início deste trabalho, serão abordadas algumas questões relevantes sobre o processo que diz respeito à unificação de alguns termos das diversas áreas científicas. Também é discutido o espaço que a Lingüística de Corpus ocupa nesse universo, além de temas como a carência de glossários técnicos em áreas como a Educação Física.

A expansão científica e tecnológica, juntamente com o fenômeno da globalização, tem gerado mudanças constantes nos termos utilizados em diferentes áreas do conhecimento humano. Uma das mudanças notáveis em algumas áreas é que muitos termos técnicos não são mais cognatos, ou seja, não possuem raízes provenientes do grego ou do latim, como o item lexical *exercise*, por exemplo, um termo muito corrente nos textos aqui analisados. Esse termo se assemelha muito ao seu correspondente *exercício* em língua portuguesa, auxiliando, dessa maneira, o leitor na tarefa interpretativa.

A maior parte dos termos técnicos e científicos que utilizamos foram convencionados no século XVIII. Em geral, sua formação tem semelhança com a formação de unidades da língua portuguesa, pois são oriundos do grego ou do latim. Por outro lado, no caso da Informática, conhecimento bastante mais moderno, os termos cunhados se diferenciam bastante da configuração lexical do português e exibem uma influência anglófona marcante.

Assim, se um leitor brasileiro com pouca proficiência em inglês tentar ler um texto técnico de Informática em inglês, enfrentará a dificuldade de não conseguir depreender o significado de grande parte da terminologia, visto que sua morfologia pouco exhibe elementos aatinos ou gregos. Esse mesmo mesmo leitor, que não tenha conhecimentos especializados de Informática nem de Medicina, teria um pouco mais de facilidade para interpretar a terminologia de textos médicos, dada a herança latina e grega muito presente na cunhagem das expressões.

A proximidade ou semelhança morfológica entre uma terminologia técnica qualquer e os padrões de formação de palavras do português será um fator bastante positivo para a troca e a circulação de conhecimentos. A situação oposta, entretanto, é um fator complicador a mais, considerando-se o cenário de uma comunicação científica e técnica que se faz entre usuários de línguas diferentes, mesmo que o inglês tenha um estatuto de *língua franca*. Assim, reconhecer esse quadro poderá ser importante para que se facilite, por exemplo, a compreensão de leitura de textos técnicos em língua estrangeira por parte de brasileiros, sem contar a necessidade de se reconhecer as terminologias e suas diferentes realizações e usos.

Após uma atuação de três anos como instrutora de inglês no universo acadêmico, ministrando para os mais variados cursos e níveis de conhecimento dos alunos, a autora desta pesquisa pôde perceber que muitos alunos se interessam e buscam auxílio na interpretação de seus textos acadêmicos em inglês, pois acreditam que as melhores obras de suas áreas não são traduzidas para o português.

A partir da percepção de dificuldade de compreensão de terminologias em língua estrangeira e também em língua portuguesa, escolheu-se fazer um levantamento dos termos mais importantes encontrados em alguns textos referentes à Fisiologia do Exercício. Desse levantamento, é proposta a criação de um glossário bilíngüe português/inglês para usos de aprendizes de leitura de textos técnicos em inglês e também para instrutores de idiomas, pois as trocas de informações incorretas, nessa área, podem acontecer com bastante frequência. A má compreensão das terminologias pode ter sérias conseqüências para os profissionais dessa área, já que se trata de uma das subáreas da saúde, sobretudo no caso de aplicação de procedimentos para pacientes ou para o desenho de treinamentos de atletas.

1.2 Lingüística de *Corpus*

É nesse cenário que se impõe essa importante disciplina do universo lingüístico, denominada Lingüística de *Corpus*. É uma das áreas mais exploradas nos últimos anos, transformando-se no mais recente paradigma dentro da pesquisa lingüística, ao propor um estudo da língua baseado na observação empírica de grandes coletâneas de textos, os *corpora*, criteriosamente compilados por meio do uso de concordanceadores, como o programa *WordSmith Tools*, por exemplo, de autoria de Mike Scott (1999), obtido através da Internet.

Por serem capazes de analisar criteriosamente uma grande quantidade de textos extensos por meio de máquinas, como os computadores, esses concordanceadores, utilizados na Lingüística de *Corpus*, servem para auxiliar pesquisadores interessados em estudos lingüísticos em geral, como professores de idiomas, tradutores, alunos e todos os profissionais que se interessam pela leitura de textos em inglês. Além disso, podem servir para gerar instrumentos de referência universais: dicionários e glossários.

É nesse contexto que as próprias metodologias para análise estatística, lexical e gramatical em textos tornam-se itens de aprendizagem importantes, pois ainda é recente a pesquisa terminológica baseada em *corpora* técnico-científicos com o apoio de *softwares*.

Berber Sardinha (2000, p. 3) aponta, que havia *corpora* antes do computador e ilustra o verdadeiro significado da palavra *Corpus*, a qual é *Corpo*, um conjunto de documentos (conforme o dicionário Aurélio).

Uma das grandes contribuições para a Lingüística de *Corpus*, nos estudos de textos em inglês, foi a construção do *Corpus Brown*. Antes dele, na Antiguidade e na Idade Média, os *corpora* eram coletados, mantidos e analisados manualmente. Um exemplo desses *corpora* manuais são algumas citações da Bíblia.

A busca de *corpora* para fins de pesquisas lingüísticas iniciou-se com o *Brown corpus*. Antes disso, a perda de tempo e recursos financeiros com a busca de registros lingüísticos era algo comum, fato notado por Berber Sardinha (2000, p. 1) quando afirma que “Há 35 anos, as dificuldades de se informatizar um conjunto de textos eram tremendas. Vale lembrar, por exemplo, que os textos tiveram de ser transferidos para o computador por meio de cartões, perfurados um a um, tal era a tecnologia da época”.

Foi, porém, nos anos de 1980, com o aparecimento dos microcomputadores e o advento da expansão de ferramentas de processamento e de *corpora*, que ocorreu uma nova onda de mudanças, o que fez com que as pesquisas baseadas em *corpus* adquirissem maior influência em pesquisas lingüísticas.

De todos os centros que utilizam a Lingüística de *Corpus* para pesquisas lingüísticas, o mais desenvolvido encontra-se na Grã Bretanha, além dos países escandinavos (Noruega, Suécia e Dinamarca), onde ela se faz presente há anos. Tanto que, mesmo tendo sido pioneiro, dentre os *corpora* de grande extensão na Lingüística de *Corpus*, o BNC (*British National Corpus*) é uma coleção de 100 milhões de exemplares escritos e falados, retirados de uma ampla fonte de dados, elaborada para representar um extenso vocabulário do inglês britânico,

do final do século XX, falado e escrito. Sua última versão é a da edição BNC XML, criada em 2007.

É uma coleção de exemplares de 100 milhões de textos da língua escrita e falada retirados de um amplo leque de fontes, criados para fazer uma ampla representação do inglês britânico, falado e escrito, do final da década do século 20. Sua última versão foi a edição BNC XML, gravada em 2007.¹

Neste trabalho, mesmo estando propensas a iminentes alterações, tão comuns no universo tradutório, as citações em inglês serão aqui traduzidas e suas versões originais aparecerão no final das páginas em que constam.

Também os grandes dicionários de inglês baseados na Lingüística de *Corpus* encontram-se na Grã-Bretanha tais como: *Oxford*, *Cambridge*, *Collins*, *Longman* e outros de renome.

Uma grande contradição que se faz marcante é que, embora os Estados Unidos sejam o país onde há um maior número de centros de pesquisas e de recursos de informática, a Lingüística de *Corpus* não é tão comum quanto à Lingüística gerativa-transformacional, a qual ocupa papel de grande destaque nos departamentos de Lingüística, conflitando-se com a Lingüística de *Corpus*.

1.3 A falta de glossários técnicos na Educação Física

Alguns lingüistas, como Gaudin (1993), acreditam na insuficiência dos instrumentos de referência, tais como glossários, visto que não expressam a realidade dos usos dos termos, já que o seu ideal seria uma análise contextual deles. Entretanto, acredita-se glossários são muito importantes. Em que pesem suas limitações, são úteis para uma melhor compreensão dos textos, pois inúmeros termos técnicos muitas vezes são encontrados apenas em dicionários especializados. Os dicionários comuns de língua muitas exibem apenas recortes muito genéricos do sentido da terminologia técnica mais específica de uma área de conhecimento. E, quando o aprendiz de leitura de textos técnicos em língua estrangeira procurar alguns desses termos isoladamente em um dicionário não específico, ficará bem difícil compreendê-los.

¹ It is a million word collection of samples of written and spoken language from a wide range of sources, designed to represent a wide cross-section of British English from the later part of the 20th century, both spoken and written. The last edition is the BNC XML, released in 2007.

Some-se a isso o fato de os dicionários estrangeiros para aprendizes não-nativos em geral privilegiarem palavras de maior frequência de uso na linguagem cotidiana, não técnica.

Krieger e Finatto (2004, p. 179) afirmam que “Infelizmente, é inegável a carência de instrumentos de referência em português brasileiro, o que obriga tradutores a atuarem como verdadeiros ‘pesquisadores-exploradores’ das linguagens especializadas. Desse modo, produzem seus próprios materiais de apoio e ‘fazem’ glossários”.

Essa solução, porém, pode ter um resultado insatisfatório pelo fato de esses pesquisadores não possuírem as adequadas metodologias e ferramentas para essa tarefa. Isso, muitas vezes, gera trabalhos de má qualidade na tentativa de proporem soluções para dar conta da lacuna desses instrumentos de referência. Além disso, seria preciso considerar também que instrumentos dirigidos para profissionais de tradução teriam um perfil diferenciado dos instrumentos voltados para a compreensão de leitura em situação de ensino/aprendizagem em língua estrangeira.

Uma das áreas atingidas pela falta de instrumentos bilíngües para consulta é a Fisiologia do Exercício, uma área da saúde que está relacionada com a Biologia. Estuda a função de todos os órgãos e sistemas fisiológicos durante as atividades físicas e está inserida nas grades curriculares dos cursos de Educação Física, Fisioterapia, Medicina e Nutrição. Os estudos podem variar desde funções biológicas, no âmbito molecular, até o comportamento de um organismo completo ou de populações de indivíduos.

As atividades físicas também podem envolver desde praticantes de esportes amadores até profissionais ou qualquer pessoa que faz uso do corpo físico em movimento. Isso envolve até mesmo as atividades cotidianas, o lazer, o trabalho, aquelas realizadas em ambientes como o espaço exterior e a grandes altitudes ou profundidades.

Antes de discorrer sobre o porquê da Fisiologia do Exercício ter sido a disciplina selecionada para o estudo aqui desenvolvido, faz-se necessário salientar que ela é uma disciplina importante em muitos cursos de Educação Física, por ser uma das áreas amplamente estudadas em cursos de pós-graduação em muitas universidades brasileiras, tais como: Universidade Gama Filho e Universidade Veiga de Almeida, ambas no Rio de Janeiro, e no Centro de Estudos de Fisiologia do Exercício (CEFE), órgão suplementar da Universidade Federal de São Paulo, Escola Paulista de Medicina, na UFSCar, dentre outras, e no mundo.

Além do mais, a Fisiologia do Exercício está sendo estudada há vários séculos de forma disseminada em outras áreas do conhecimento. Na década de 1920, com a criação do

Laboratório de Harvard, começaram a surgir laboratórios de Fisiologia especializados em temas relacionados com atividades físicas, inclusive militares. Esses laboratórios receberam progressivamente mais atenção a partir da Segunda Guerra, disseminando-se nas universidades com cursos de formação para Educação Física, Medicina e Fisioterapia.

Contudo, Tojal (2004) faz referência a um grave erro presente em muitos desses cursos:

Sem querer historiar a existência da Educação Física no Brasil, o que pretendo neste momento é declarar que devido ao já longo período de atuação nessa área, considero que tem faltado à Educação Física aquela cientificidade necessária a toda profissão de nível superior, ou seja, falta que seus profissionais consigam se expressar através de termos científicos e técnicos adequados identificados com os diferentes conhecimentos que nela e através dela sejam desenvolvidos. O que se percebe é que essa área vive e mesmo viveu sempre do que toma emprestado de outras áreas, como biologia, psicologia, sociologia, pedagogia, deixando de demonstrar qualquer rigor científico que lhe pertença enquanto uma nova área ou campo de conhecimento.

Um dos problemas atuantes na falta de cientificidade, aí observada, diz respeito à lacuna gerada pela falta de termos científicos e técnicos adequados aos diversos fenômenos existentes na área, atrapalhando, dessa forma, a efetiva comunicação de seus profissionais.

Há também o fato de os representantes das subdisciplinas da Ciência do Esporte possuírem diferentes interesses de pesquisas e diferentes formações profissionais, fazendo aparecer diferenças substanciais nos objetivos e métodos de trabalho, o que torna difícil alcançar um consenso. Devido a isso, surgem as subáreas da Ciência do Esporte, as quais são: Pedagogia do Esporte, Psicologia do Esporte, Sociologia do Esporte, Medicina do Esporte, História do Esporte, Filosofia do Esporte, Biomecânica do Esporte, Fisiologia do Exercício, Teoria do Treinamento, Jornalismo Esportivo e outras que começam a se integrar.

Se tais dificuldades surgem entre profissionais dentro do mesmo país, o que dizer de profissionais que se encontram em diferentes países, com idiomas e cultura diferentes?

Devido a todas essas questões abordadas, tentou-se auxiliar alunos de áreas que trabalham com a Fisiologia do Exercício na construção de seus próprios glossários terminológicos, utilizando artigos retirados da internet, na tentativa de facilitar a leitura de seus textos em inglês, utilizando a Terminologia, ciência pouco explorada nos estudos lingüísticos. Vale lembrar que, no decorrer deste trabalho, pode-se perceber que, para se obter um maior êxito na criação de seu glossário, é necessário que o profissional interessado no processo de criação deste glossário, tenha o mínimo de conhecimento dos termos mais importantes da sua área, na língua de partida e de chegada, além de um pouco de conhecimento dos diversos programas e corpora disponíveis.

Na tentativa de sanar alguns dos problemas cruciais decorrentes da falta de interpretação de termos especializados das diferentes áreas do conhecimento humano, foi proposta a criação de um glossário individual, realizado com o auxílio de programas concordanceadores utilizados em computadores, que são peças fundamentais da Lingüística de *Corpus*.

Em vista da limitação momentânea de um enfoque total da Lingüística de *Corpus*, devido à sua expansão, o grande desafio será a demonstração da metodologia da elaboração de um glossário que consiga oferecer subsídios para mostrar como criá-lo, e não fazer uma abordagem específica da Lingüística de *Corpus*.

Como se trata da confecção de glossários realizada por pesquisadores por meio de ferramentas eletrônicas oferecidas pela Lingüística de *Corpus*, vê-se a importância de algum conhecimento computacional prévio por parte do pesquisador interessado, o que nem sempre é realidade, pois muitos alunos, principalmente de áreas como a Educação Física, possuem conhecimentos, às vezes, muito precários, devido à falta de uso ou, até mesmo, de acesso a programas de microcomputadores.

É importante ressaltar que o presente trabalho não é um manual prático de como utilizar o *Wordsmith*, programa de análise de *corpus*, mas uma tentativa de esclarecimento da importância da definição de um procedimento metodológico de base estatística, tipicamente embutido em um programa de análise de *corpus*, para a caracterização de um gênero textual, buscando seu entendimento pela elaboração de glossários pertinentes à área estudada.

É, portanto, em função de todas as implicações acima explicadas, que se pode afirmar a contribuição da proposta da elaboração de um glossário bilíngüe, no par lingüístico inglês/português voltado para a interpretação de textos da Fisiologia do Exercício, em que o uso de ferramentas eletrônicas tem papel preponderante nesse tipo de elaboração.

O fato de o público-alvo da pesquisa ter sido profissionais interessados na interpretação de textos em inglês da área da Fisiologia do Exercício, não se nega sua utilização por outros profissionais, tais como: lingüistas, tradutores, professores de línguas, pesquisadores e estudantes de línguas e áreas correlatas, com conhecimentos básicos de inglês e computação (nível usuário).

No propósito de divulgar um método para a geração de glossários bilíngüe por meio da Lingüística de *Corpus* e tendo o *Wordsmith* como ferramenta principal, este trabalho está dividido em seis capítulos, desde seu projeto até o início das primeiras experimentações de seu

verbetes, organizados como: 1) Introdução, onde há uma breve apresentação dos assuntos abordados neste trabalho; 2) Revisão da Literatura, onde há uma apresentação da colaboração de pesquisadores do assunto e suas teorias; 3) Justificativa, onde é explanado porque o tema apresentado foi escolhido em detrimento de outros, relacionados à Linguística de *Corpus*; 4) Objetivos, onde estão apresentados alguns dos passos a serem percorridos para a montagem do glossário; 5) Metodologia, capítulo onde será vista uma breve apresentação da construção de um verbete de glossário desde sua determinação como palavra-chave à análise dos resultados de concordância, com a identificação dos contextos de coocorrência, a identificação das traduções em português e o resultado final sob a forma de um verbete de glossário. 6) Resultados, onde se demonstram os itens lexicais provenientes de textos eletrônicos, utilizados como entrada neste glossário. 7) Considerações Finais, onde será feita uma breve descrição do que foi feito durante esta pesquisa.

2 REVISÃO DA LITERATURA

2.1 Sobre os conceitos que permeiam a Lingüística de *Corpus* e os *corpora*

Embora pareçam estar fragmentadas, devido à variedade de informações presentes nesta revisão da literatura, essas informações se fazem necessárias, para um maior entendimento do assunto abordado.

Já que as definições e conceitos são itens lexicais de grande representatividade no meio técnico, deve-se tomar muito cuidado para não presumir que o leitor já saiba uma série de conceitos aqui tratados. Faz-se então necessário um detalhamento maior em relação a esses conceitos.

O glossário funciona como uma amostra sucinta, e não como uma obra de consulta para uma pesquisa mais aprofundada, da área de fisiologia do exercício.

Vale ressaltar que, em caso de um prazo mais longo para a efetivação deste trabalho, seria acrescentado também outros itens no glossário confeccionado, além da tradução e exemplos dos respectivos itens. Seria realizada também, a ilustração desses itens, além da classificação da classe de palavras e a inserção de uma quantidade maior de exemplos, com as expressões mais comuns desses exemplos. Mas, isto levaria muito tempo, e seria um trabalho mais focado, para a confecção de dicionário, e não no de glossários, já que o dicionário, produto final da lexicografia, deve providenciar a solução de dúvidas de ortografia, de regência verbal e significados, além de obter importantes funções, tais como: a fonética, a classificação de palavras, a etimologia e o uso das diferentes listas de consulta de medidas, códigos e abreviaturas, tornando-o assim, uma obra complexa para se elaborar.

Isto poderia gerar a dúvida se questões como: se a ilustração e classificação de classe de palavras não seriam mais apropriadas para a confecção de dicionários, e não na de glossários. A efetiva resposta para este questionamento aparecerá bem detalhada, logo abaixo na seção de objetivos a serem atingidos, no decorrer deste trabalho.

No presente capítulo, é feita uma breve retrospectiva da visão de alguns autores, com relação a temas como *corpora* e Lingüística de *Corpus*, esclarecendo alguns conceitos referentes a esses temas.

O termo Lingüística de *Corpus* não causa espanto a estudantes da Lingüística, nem a pesquisadores do meio acadêmico, mas ainda é uma área pouco explorada, que está se desenvolvendo gradativamente, ao longo dos anos. Um dos motivos pelo qual essa área ainda é pouco explorada na realidade brasileira é que sua ferramenta principal – o computador – nem sempre é o instrumento de trabalho de todos os brasileiros.

Para se entender exatamente o conceito atual de *corpus* e de Lingüística de *Corpus*, é necessário um resgate dos conceitos que os permeiam.

Berber Sardinha (2000) relaciona o sentido original da palavra *corpus* à *corpo*, conjunto de documentos (segundo o dicionário Aurélio). Um dos efeitos para se definir *corpus* é visto por Berber Sardinha (2000), quando ele nos remete à idéia de que há um problema na definição do *corpus*, quanto à originalidade dos textos, pois, ao mesmo tempo em que afirma a necessidade dos textos ou *corpora* de estudo serem autênticos e naturais, eles são analisados de uma forma superficial, ao passo que são utilizados para fins de pesquisa lingüística. Isso faz com que pareçam superficiais para alguns pesquisadores do assunto.

Quanto às características dos *corpora*, Berber Sardinha (2004, p. 19) relata que, para viabilizá-los, devem-se respeitar alguns critérios, como:

1- Os textos devem ser autênticos, produzidos por falantes nativos e não elaborados com o propósito de serem analisados lingüisticamente.

2- O conteúdo deve ser autêntico e natural, obedecendo a regras estabelecidas por seus criadores, para, assim, atingir seu objetivo.

3- O objeto de análise deve representar um idioma ou variedade lingüística.

Berber Sardinha (2004, p. 6) também menciona a idéia de a Linguística de *Corpus* estar sendo explorada com fins comerciais, além de acadêmicos:

Também no âmbito empresarial há um interesse crescente nas aplicações comerciais de estudos baseados em *corpora*. Devem-se destacar as parcerias entre empresas e dicionários: a norma é a associação de um centro de pesquisa em lingüística de *corpus* com uma editora.

Desses projetos comerciais, o primeiro, e já desativado, foi o Cobuild, uma parceria entre a Universidade de Birmingham (Grã Betanha) e a Editora Collins, quando foram elaborados dicionários, gramáticas e livros didáticos para o ensino de inglês.

Outros projetos comerciais mantidos por empresas estão se tornando cada vez mais comuns, onde há centros de pesquisas baseadas em corpus para o processamento automático de textos, informatização de grandes bases de dados e montagem de sistemas inteligentes de reconhecimento de voz e gerenciamento de informação.

Das empresas que mais investem no tema estão as de telecomunicações e algumas de renome internacional, tais como: Xérox, Microsoft e Cannon. Apesar de não ser uma novidade no mercado, a criação de glossários também deverá agilizar a vida de muitos profissionais do meio empresarial e acadêmico que desejam ler textos da sua área em inglês.

Quanto à seleção dos *corpora* de estudo, Johansson (1998) menciona que podem ser: (a) *corpora* formados por textos originais produzidos em duas ou mais línguas; e (b) *corpora* formados por textos originais e respectivas traduções, os quais também podem ser denominados *corpus* de referência e *corpus* de estudo.

Para a criação dos glossários propostos neste trabalho, vale ainda lembrar da Terminologia, uma das áreas abordadas neste trabalho.

Lara (2004) menciona que o *termo*, unidade mínima da Terminologia, se distingue da palavra por ser utilizado no interior de uma linguagem de especialidade. Ao contrário da palavra “comum” que é mais polissêmica, o termo tende a exibir menor polissemia. Assim, um termo que é uma unidade inserida em um determinado discurso, o que, muitas vezes colabora, para não haver ambigüidades lingüísticas, fator indesejável no universo científico especializado. Desta maneira, os termos técnicos possuem uma maior tendência em ser monossêmicos, ou seja, possuírem somente um sentido somente para cada termo em uma situação de uso bastante determinada. Um exemplo dessas determinações são os diferentes sentidos específicos de *erosão* em Geologia e em Odontologia.

Como mencionado anteriormente, o termo aparece relacionado a um significante do universo do léxico científico, ou seja, retratado em uma linguagem de especialidade, ao contrário da palavra que faz parte de todo um universo do léxico da língua. Esses termos, no entanto, enquanto designações verbais, são associados a determinados conceitos. O reconhecimento de conceitos e de suas designações é o ponto de partida do trabalho terminológico, o que deriva no reconhecimento de definições terminológicas.

Krieger e Finatto (2004) relatam a importância da Terminologia, que pode ser vista como uma disciplina cujo objetivo é descrever o uso de termos técnicos na esfera internacional. Há, entretanto, várias perspectivas teóricas em Terminologia, hoje uma área de estudos marcada pela pluralidade de enfoques que vão dos mais prescritivos aos mais descritivos. A Teoria Geral da Terminologia (TGT), por exemplo, é um tipo de teoria que visa reconhecer e promover a padronização de termos, nacional e internacional, nas diversas áreas e subáreas do conhecimento humano. Há também a perspectiva denominada Sócio-Terminologia, cujo maior ponto de interesse é justamente a variação terminológica. Essa perspectiva ocupa-se de descrever as situações de uso em que há uma diferença entre os termos padronizados e os “vulgares” ou mais disseminados em uma dada comunidade de especialistas ou técnicos em situações de comunicação menos formais. Um exemplo disso seriam os usos verificados de *cordas vocais* e de *pregas vocais* em textos de Medicina. Essa variação é entendida, nessa perspectiva, como algo inerente à linguagem humana em uso.

Segundo as autoras, a aplicação dos estudos de Terminologia encontra-se, com maior ênfase, na produção de glossários e dicionários, pois, mesmo que a definição dos termos traga uma “situação textual dicionarística”, é seu conhecimento algo merecedor de destaque. Krieger e Finatto (2004) afirmam, ainda, a esse respeito que a falta de homogeneidade e sistematicidade na organização e na maneira de registrar a informação sobre uma dada terminologia pode impedir a comunicação entre sistemas de diferentes organizações, tanto nacional como internacionalmente.

Foi então, na tentativa de buscar certo entendimento e organização dos diferentes termos usados nas inúmeras ciências e técnicas, que a Terminologia se integrou com a Lingüística de *Corpus*, sendo utilizada por esta última com grande repercussão em sua estrutura organizacional.

Outra razão apontada por Krieger e Finatto (p. 82) sobre a origem dos termos científicos é que ela também “está vinculada ao fato das ciências captarem a essência dos fenômenos observados, buscando uma imagem que reflita essa essência”. É a partir desse pressuposto que se cria uma linguagem diferenciada da língua comum para os termos técnicos.

Dessa maneira, a presença de componentes de origem grega e latina caracteriza a terminologia de algumas áreas. Conforme Martí (1991, p. 81), tal presença ocorre mais nas ciências experimentais e exatas e menos nos domínios técnicos, como o vocabulário da Mecânica. Por outro lado, a análise dos constituintes dos termos, observando componentes

como radical, afixos e sufixos, permite um reconhecimento da formação lexical terminológica, típica de diferentes áreas do conhecimento.

Apesar de o século XVIII ter sido a época de destaque referente às nomenclaturas, foi no século XX que um estudo mais aprofundado das terminologias passou a ser realizado. No entanto, ainda não há uma organização e um levantamento efetivo e homogêneo de todos os termos referentes às diversas áreas do conhecimento humano.

A Terminologia é também bastante presente nos estudos sobre *Language for Specific Purposes* (LSP), estudos que se dedicam a reconhecer os fenômenos da comunicação científica e técnica. Uma terminologia engloba todo o universo referente aos componentes lexicais especializados de determinadas áreas, um meio inevitável de expressão e comunicação profissional.

É indispensável que o profissional ou futuro profissional de uma especialidade, o qual deseje para elaborar seu próprio glossário de termos, saiba reconhecer e utilizar os recursos hoje disponíveis para recolher e organizar um conjunto de dados úteis e confiáveis. Não lhe bastará somente o conhecimento “técnico” para a criação de glossários da língua inglesa, também será necessário o conhecimento do sentido bastante específico de alguns desses termos em inglês e de seus equivalentes em português, além do conhecimento das ferramentas principais de Informática para obter um resultado satisfatório.

Sobre esse assunto, Baker (1995) ressalta a importância de um conhecimento prévio de recursos informatizados na manipulação da terminologia para a elaboração de glossários atualizados por meio de *corpora* digitais, sem haver uma mera republicação de listas prévias. Por ser mais amigável, o trabalho com recursos informatizados com uso de *corpora* digitais está se expandindo cada vez mais. Segundo a autora, isso ocorre com maior destaque na Terminologia onde os textos são retirados de *corpora* eletrônicos, e não retirados de listas prévias. Além da Terminologia, os *corpora* computadorizados também estão se tornando populares na área da Tradução Informatizada.

Sager (*apud* BAKER 1995, p. 239) ainda detalha várias aplicações da análise de *corpus* relacionadas à Terminologia. Isso inclui:

complementar os léxicos computadorizados, processar textos a serem traduzidos e comparar às entradas terminológicas informatizadas para identificar novos itens. Os itens identificados podem ser utilizados, tanto para a complementação das entradas, quanto para ajudar os terminologistas a esvaziar suas entradas. Certificar-se de que todas as possíveis variáveis de um termo estão esgotadas. Demonstrar o

comportamento lingüístico dos termos. Identificar mudanças na frequência ou uso de termos através da análise de extensos corpora estatísticos.²

² supplementing machine-translation lexicons. Processing texts which are to be translated and comparing them to machine-readable terminology holdings to identify new items. Items thus identified can be used both to supplement the holdings and to help terminologists pre-empt translators queries. Ensuring that all possible variants of a term are covered. Demonstrating the linguistic behaviour of terms. Identifying changes in the frequency or usage of terms by means of statistical analyses of large corpora.

3 JUSTIFICATIVA

3.1 Problema a ser estudado

Algo que se procura deixar claro nesta seção é a importância que o presente estudo tem para o universo acadêmico, como um todo.

Uma questão que poderá surgir na mente de muitos alunos é se não seria mais prático buscar textos da área da Fisiologia do Exercício que já apresentassem as respectivas traduções para o português, assim não haveria a necessidade de se coletarem textos na internet, realizando todo o esforço necessário à compilação de um glossário, sempre que se quisesse compreendê-los. Ou até mesmo em ferramentas de tradução *on-line*, como dicionários ou tradutores. Isso, com certeza, seria bem mais prático se todas as vezes que fosse necessário ler um texto de uma determinada área científica, o aprendiz pudesse encontrar sua tradução para o português, na internet ou em livros impressos. A resposta à primeira questão é que, ao se procurarem as palavras que se têm dúvidas em dicionários genéricos, *on-line* ou impressos, muitas vezes não serão encontradas, pois, como esses dicionários tratam dos termos gerais falados na língua, não tratarão de termos mais específicos de uma determinada área do conhecimento humano.

Dessa forma, um questionamento que poderá surgir na mente de muitas pessoas é por que alguém se objetivaria a realizar essa tarefa, ao invés de adquirir um dicionário específico da área, já elaborado?

Sobre esse questionamento, um fator que merece destaque é que nem sempre há dicionários especializados das diversas áreas do conhecimento disponíveis no mercado, e quando os há, são de alto custo por serem provenientes de outros países, o que impede que muitos estudantes possam adquiri-los.

3.2 Relevância da proposta para os estudos de termos da Educação Física

A relevância do tema desta dissertação pode ser afirmada em relação à importância da proposta de elaboração de glossários bilíngües, no par lingüístico inglês-português, em uma área pouco explorada no contexto brasileiro, a Lingüística de Corpus. O projeto configura-se como merecedor da atenção do universo acadêmico, pois virá a contribuir para a tarefa de interpretação/leitura de textos da área de Fisiologia do Exercício, assim como de todas as áreas de interesse, como a tradução, a Lexicografia e o ensino de línguas estrangeiras.

Quando se consulta um dicionário especializado de uma área específica, muitos itens de interesse aparecem em um contexto desprovido de situações reais de uso, podendo tornar sua compreensão impossível. No entanto, quando se elabora o próprio glossário, ou dicionário, pode-se inseri-lo nas situações de uso em que essas palavras realmente ocorrem, por meio de exemplos encontrados nos próprios *corpora* de estudo, como será demonstrado na seção de Análise e resultados obtidos, onde se demonstram os itens lexicais provenientes de textos eletrônicos, utilizados como entrada no glossário.

Os textos eletrônicos são textos passíveis de leitura por máquina. Sua análise é feita por programas computacionais específicos, como o *WordSmith* (SCOTT, 1999). Porém, não é qualquer coletânea de frases soltas, de trechos desconexos ou conjunto de dados legíveis por computador que podem ser considerados como tal. Para serem vistos assim, devem ser uma coletânea de dados lingüísticos provenientes da linguagem natural.

Dessa forma, na elaboração do glossário, decidiu-se buscar textos da internet, local ideal em que o conhecimento é acessível sem limitação de tempo e espaço, de onde se retirou o *corpus* de análise, pois nela encontra-se uma gama de textos digitalizados em inglês. O fato de os textos já estarem em formato digital é de suma importância, já que se trata de Lingüística de *Corpus*, em que os textos são analisados com o auxílio de programas concordanceadores. Além disso, se fossem utilizados livros impressos como fonte de dados, haveria a necessidade de escaneá-los e corrigi-los por completo, tarefa muito árdua e demoradamente demorada, pois os textos, quando escaneados, geram inúmeras imperfeições, trazendo a necessidade de revisá-los.

4 OBJETIVOS

4.1 Glossário ou dicionário

No presente capítulos, são abordados algumas diferenças de uso entre um glossário e um dicionário e a relação da tradução com a Lingüística de *Corpus* e a Terminologia.

Esta pesquisa visa apresentar uma metodologia para a elaboração de um glossário bilíngüe com base em um *corpus* do domínio científico da área da Fisiologia do Exercício.

Glossário, segundo Krieger e Finatto (2004, p. 51), “Costuma ser definido como repertório de unidades lexicais de uma especialidade com suas respectivas definições ou outras especificações sobre seus sentidos”.

O glossário ainda pode ser visto como o resultado das ocorrências de um texto específico, no qual se armazenam as palavras-ocorrências dele extraídas. É devido ao fato de manifestar a especificidade de um determinado texto ou discurso que geralmente se encontra no fim de alguns livros, buscando esclarecer o uso do léxico ou expressões mais rebuscadas utilizadas pelo autor, auxiliando o pesquisador na interpretação das palavras-chave da obra.

O fator-chave de um glossário é o vocabulário, ou seja, o conjunto de palavras existentes num ambiente mais restrito, proveniente de um lugar e tempo determinados, e não o léxico, o conjunto das palavras fundamentais de uma língua.

Já o dicionário terminológico ou técnico-científico “É uma obra que registra o conjunto de termos de um domínio oferecendo primordialmente informações conceituais e, por vezes, lingüísticas. Caracteriza-se por uma cobertura exaustiva de itens lexicais” (KRIEGER e FINATTO, 2004, p. 51).

Por ser uma área que está intimamente ligada ao tratamento das palavras e termos, a Lexicografia não poderia deixar de ser aqui mencionada, pois se trata da ciência responsável pela confecção de dicionários.

Segundo Welker (2004, p. 11), a Lexicografia tem dois sentidos:

- Lexicografia prática – designa a “ciência”, “prática”, ou mesmo “arte” de elaborar dicionários.
- Lexicografia teórica – emprega-se, freqüentemente, em línguas como o inglês, francês e alemão, o termo *metalexigrafia*. [...] O estudo de problemas ligados à

elaboração de dicionários, a crítica de dicionários, a pesquisa da história da lexicografia, a pesquisa do uso dos dicionários e ainda a tipologia.

Geralmente a utilização de glossários se faz presente no meio técnico. Os termos técnicos são palavras oriundas de uma determinada área científica que visa unificar sua utilização no universo daquela área, fazendo com que sejam os mesmos em qualquer lugar do mundo. Esse tem sido o propósito da Terminologia, na tentativa da unificação de termos universais e é de suma importância no que concerne aos glossários.

Um grande elemento facilitador, na busca da unificação de termos técnicos, é que eles, em geral, são monossêmicos, isto é, possuem somente um significado, o que não ocorre com a maioria dos lexemas de um idioma.

Embora a Lexicografia e a Terminologia sejam áreas correlatas, há uma distinção de grande importância entre elas, pois a Lexicografia se ocupa da criação dos dicionários, e a Terminologia se ocupa da descrição dos termos técnicos que serão inseridos nesses dicionários ou glossários terminológicos. (KRIEGER e FINATTO, 2004, p. 47).

O Michaelis, segundo Welker (2004, p.24), “lista léxico, vocabulário, glossário simplesmente como sinônimos de dicionário”.

Da forma definida acima, não há nenhuma distinção entre dicionário e glossário, o que faz o leitor acreditar em tal possibilidade. Visto sob a ótica de um lexicógrafo, essa afirmativa não corresponde à verdade dos fatos, pois, como já foi mencionado acima por Krieger e Finatto, o dicionário é uma obra bem mais complexa e, por vezes, mais completa do que um glossário, já que engloba todo o universo lexical da língua, e não um universo lexical menor, com menos informações a respeito dos itens abordados.

Em vista das questões abordadas, acredita-se que, por ser mais prático e menos custoso, o glossário seja mais apropriado para os fins desta pesquisa, já que não exige muitos itens e outros profissionais para fazer pesquisa dos verbetes que serão inseridos nele. Para sua elaboração, basta um levantamento das ocorrências do item em um texto e das acepções que aparecem naquele texto.

Devido ao fato de os glossários não exigirem tantas informações detalhadas referentes a um termo quanto um dicionário, muitos livros possuem um glossário no final explicando o significado de algumas palavras utilizadas pelo autor em sua obra, as quais poderiam gerar maiores problemas de compreensão.

Como já afirmado por Krieger e Finatto (2004), o processo de geração de um dicionário terminológico é longo; na maioria das vezes, lento, e demanda um grande volume de pesquisa textual prévia (p. 127).

A elaboração de dicionários torna-se lenta devido ao fato de que um dicionário de língua engloba o universo dos lexemas, os quais são as unidades-padrão dos dicionários. E os lexemas são vocábulos polissêmicos, ou seja, correspondem a vários significados.

Devido a sua maior complexidade, ao se procurarem as diferentes tipologias e classificações dos inúmeros dicionários existentes, encontrar-se-ão alguns, dentre os quais, mencionados por Welker (2004, p. 42), classificados como obras gerais (G) e obras especiais (E), subdivididos conforme informações lingüísticas (l) ou informações factuais (f):

G/l: dicionário geral, para aprendizes, histórico, ortográfico; thesaurus;
 E/l: dicionário de estrangeirismos, de idiomatismos, de arcaísmos, de provérbios;
 G/f: enciclopédia, almanaque, catálogo, bibliografia, guia de viagem, atlas;
 E/f: dicionário de música, de história, de arte, de física; banco terminológico.

Há ainda os dicionários monolíngües e bilíngües/multilíngües.

Welker (2004, p. 43) ainda confirma um fato já mencionado anteriormente a respeito dos dicionários monolíngües ao apontar que “O fato de a maioria ser monolíngüe deve-se a razões econômicas ou à falta de procura por bilíngües especializados”.

Uma solução muito útil e proveitosa utilizada por alguns tradutores tem sido a proposta neste trabalho. Krieger e Finatto (p. 177) observam que “Desse modo, a prática terminológica, principalmente pela via da produção de glossários e dicionários mono e multilíngües, tem sido bastante aproveitada pelo tradutor de textos técnicos e científicos e também pelos gestores de informação”.

4.2 O conceito de tradução

Assim como a Lingüística de *Corpus*, na elaboração de glossários, a tradução também é uma das disciplinas determinantes para sua realização, tais como a Lexicografia e a Terminologia.

O fato de a terminologia existente em muitas áreas técnicas do inglês conter termos cognatos provenientes da mesma raiz do grego e do latim, muitas vezes faz acreditar que a

análise de *corpora* que possuem termos técnicos em sua essência não seja uma tarefa que exija maiores esforços.

Ao contrário do que muitos acreditam, a tradução não é uma tarefa de fácil realização, não bastando que o tradutor domine ou tenha fluência nos dois idiomas em questão, apenas para traduzir as palavras de um idioma ao outro.

O raciocínio equivocado em torno da atividade tradutória, como uma atividade modesta, faz com que muitos não vejam a grande importância que essa tarefa constitui nas inovações técnicas, literárias e científicas.

4.3 Sobre a tradução

A tradução também nunca recebeu a devida importância no meio social e acadêmico.

Também se pode notar isso na didática dos cursos de línguas estrangeiras, onde a tradução é vista apenas como um recurso utilizado para se atingirem objetivos mais específicos, como a interpretação de textos e o ensino de palavras isoladas.

Até o presente momento, o que se tem feito em relação a esses problemas provenientes de textos com interpretação mais complexa, os quais demandam um maior grau de exigência na sua prática, é buscar mecanismos que disfarçam essas dificuldades, recorrendo à omissão de algumas partes do texto ou desmistificando a verdadeira essência do texto.

Ao se analisarem os paradigmas postulados como solução adotada para as deficiências subjacentes aos textos estranhos à cultura, chega-se à conclusão que, muitas vezes, a verdadeira essência do texto de origem se perde em meio a tentativas ineficientes adotadas por alguns tradutores. Isso ocorre, principalmente, no meio literário, onde muitos escritores utilizam os mais variados estilos e figuras de linguagem como seu estilo pessoal.

Há ainda várias outras questões envolvendo o conceito de boas e más traduções, do que é traduzível e intraduzível, provenientes de diversos autores, mas não serão abordadas por não fazerem parte do foco central deste trabalho.

Embora todos os fatos explanados acima se refiram a traduções de textos literários, acredita-se que, na esfera que concerne aos textos técnicos, essas questões sejam de cunho mais simples, já que se concentram em termos isolados, dentro de um universo científico pré-estabelecido. Isso, portanto, não implica a afirmação de que os termos fazem parte de um

universo paralelo formando uma língua à parte, diferente daquela considerada como léxico geral da língua.

Mesmo sendo retirados de muitos *sites* conhecidos, também se deve tomar cuidado em relação à qualidade de algumas traduções de textos eletrônicos de *sites* bilíngües. Deve-se tomar cuidado com a qualidade das traduções de textos provenientes de alguns *sites* desconhecidos, pois alguns tradutores, ou indivíduos que se julgam tradutores, realizam todos os tipos de barbaridades em suas traduções, muitas vezes se esquecendo da responsabilidade que envolve a arte de traduzir.

Isso vale para a navegação em qualquer *site* disponível aos internautas, já que não há nenhum órgão responsável pela avaliação constante e garantia de qualidade das traduções dos textos. A avaliação dessa qualidade só depende da crítica dos usuários.

Apesar de muito útil no cotidiano de um tradutor, as ferramentas de uso, como o computador e outras tecnologias, tais como os tradutores *on-line*, também devem ser utilizadas com cuidado, pois, mesmo com todos os seus atributos, elas não substituem o bom senso e a criatividade do ser humano.

Assim como as ferramentas citadas acima, a criação de tradutores e glossários *on-line*, como meios para auxiliar profissionais que necessitam compreender textos bilíngües, também está sendo algo bem proveitoso no meio empresarial. Segundo dados de uma pesquisa realizada pelo grupo *Delphi* (2005) a maioria dos usuários de computador gasta mais de duas horas por dia buscando as informações necessárias ao seu trabalho, o que corresponde a 25% de sua jornada diária. Dessa forma, foi criado o *Babylon Enterprise*, que possui como atual cliente a Petrobras. Antes da invenção desse tradutor, era necessário que o usuário deixasse a tela de trabalho para acionar o tradutor, que ainda exigia a digitação correta da palavra.

A partir de 1997, o *Babylon* começou a oferecer esse serviço de forma mais ágil. Com apenas o deslizar do *mouse* na palavra, no termo ou na expressão, aparecia o resultado na tela com um *pop-up*. O *Babylon Enterprise* funciona da mesma maneira, a fim de facilitar a gestão de conhecimentos dentro de uma empresa, otimizando o tempo de seus colaboradores.

Para obter um resultado satisfatório, a empresa criou glossários com as informações consideradas mais importantes, por meio das ferramentas da plataforma. Não houve nenhum tipo de restrição de conteúdo na criação desses glossários, os quais dependeram do que cada departamento considerava mais relevante. Essa invenção, denominada *point-and-click* (aponte, clique e está pronto), foi vista como uma revolução no sistema de tradução em computadores.

É importante notarmos que as tecnologias mencionadas acima servem como solução se houver glossários *on-line* como perspectiva, muito úteis no meio empresarial, no qual o tempo e o dinamismo devam ser fatores imperativos, e não para glossários impressos que apenas se utilizam de programas tecnológicos para sua criação, como o que foi criado neste trabalho.

Após a descoberta das ferramentas computacionais, um problema notado na observação de muitos desses glossários, tanto os *on-line* quanto os impressos, foi sua efetivação por muitos pesquisadores, os quais não possuíam a adequada metodologia e ferramenta para essa tarefa, gerando assim inúmeras quantidades de informações errôneas a respeito de importantes dados inerentes à confecção desses glossários, gerados por meio de programas. Um dado que muitas vezes acaba sendo negligenciado por muitos desses pesquisadores que objetivam elaborar glossários, é o levantamento real dos itens que podem ser considerados chave, o que ocorre devido à falta da habilidade necessária para o procedimento correto de uso dos programas utilizados pela *Linguística de Corpus*.

5 METODOLOGIA

5.1 O uso do *WordSmith*

Seguindo os passos propostos nesta dissertação, tentou-se demonstrar a metodologia da criação de glossários como elemento facilitador na leitura de textos em inglês.

Krieger e Finatto (2004) relatam que, embora a falta de materiais bilíngües para consulta de linguagens especializadas e a criação de glossários para a exploração de novos termos não sejam novidade, as metodologias com o apoio de *softwares* utilizadas para análise estatística, lexical e gramatical em textos baseados em *corpora* ainda são vistas como algo recente.

Para isso, faz-se necessária a verificação do uso de concordâncias como algo fundamental no trabalho de pesquisa e criação de glossários para tradutores e alunos. A pesquisa realizada foi constituída por meio de exemplos trabalhados junto a um programa de computador denominado *Wordsmith Tools*.

O *WordSmith* é um programa concordanceador de análise de *corpus*, de autoria de Mike Scott (1999), usado nas Ciências Sociais para comparação, elaboração das palavras-chave e concordância dessas palavras em um *corpus*. Esse programa consegue executar operações de listagem de palavras e concordância, dentre outras.

As ferramentas básicas utilizadas no *WordSmith* são: os geradores de listas de palavras (*WordList*), de palavras-chave (*KeyWords*) e de linhas de concordância (*Concord*). O domínio de todos esses recursos, no entanto, exige do usuário conhecimentos específicos e intensa prática.

Esses recursos foram os mais utilizados neste trabalho, já que o objetivo se ateu à confecção de glossários bilíngües e suas respectivas traduções, como elemento facilitador na interpretação de textos em inglês da área da Fisiologia do Exercício.

Foram, então, coletados textos da *World Wide Web*, diretamente de *sites* bilíngües da área de Fisiologia do Exercício.

Por se tratar de uma ferramenta informatizada e utilizada para a seleção dos termos com relevância ao assunto procurado, permite a análise de dados estatísticos facilitando assim o processo, que define o termo que deve constar no glossário, pelo levantamento das palavras de

maior frequência, gerando assim uma menor margem de erros comparada ao mesmo processo, caso fosse feito manualmente. Tal processo levaria mais tempo para ser concluído e acarretaria maior gasto, pois necessitaria de um maior número de profissionais envolvidos.

Embora muitos não vejam as reais necessidades de se fazer a demonstração das diversas janelas existentes no programa utilizado, acredita-se que tal processo seja necessário por se tratar de um trabalho prático visando à realização de um glossário, por meio de um programa informatizado, o *WordSmith*, Mike Scott, (1999), não familiar a muitas pessoas.

Por vezes, o pesquisador interessado não possui nenhum conhecimento prático de uso das diferentes ferramentas e das funções existentes no programa, tornando-se útil a demonstração das janelas impressas com suas explicações em português, como foi realizado neste trabalho.

Caso o aluno de nível básico procure as informações de funcionamento do programa na ferramenta *Help* (Ajuda), existente no programa, encontrará dificuldades, pois estão em inglês.

Na tentativa de fazer uma sucinta apresentação do programa trabalhado, fez-se necessária a divisão das etapas a serem seguidas em passos, como demonstrados na Figura 1:

Figura 1 - Tela inicial do *WordSmith*



Neste trabalho, foram utilizados quatro *corpora* para a elaboração do glossário, sendo dois de análise: um em inglês, com suas respectivas traduções em português, e outros dois de referência: um em português e outro em inglês.

As opções relativas ao desenho dos *corpora* articulam-se em torno dos seguintes critérios de seleção:

5.2 Tipologia e variação dos textos a incluir

Os de referência serão *Brown University Standard Corpus of Present-day American English*, desenvolvido por pesquisadores da *Brown University* com 1.000.000 de palavras e o *corpus* em português denominado Núcleo Interinstitucional de Lingüística Computacional (NILC), criado em 1993, pela Universidade de São Paulo (USP) de São Carlos, com 40 milhões de palavras.

Ambos são *corpora* cujos textos são, portanto, amostras da linguagem natural, contendo desde artigos jornalísticos até textos literários e científicos, das mais diversas dimensões. Esses *corpora* foram escolhidos, devido à sua dimensão, pois no momento da seleção de amostras da linguagem, deve-se priorizar os mais representativos do idioma, sem deixar de lado, o tipo de usuário a ser atingido.

O NILC é um *corpus* representativo da língua portuguesa, com uma variedade de textos jornalísticos.

Já o Brown foi um *corpus* considerado bem representativo da língua inglesa, da universidade americana Brown, pois foi o primeiro *corpus* eletrônico em inglês a ser implantado em pesquisas baseadas em *corpus* eletrônico. Amplamente utilizado na Lingüística Computacional e elaborado por meio de uma criteriosa seleção de textos, possui em torno de um milhão de palavras do inglês atual, retiradas de uma imensa variedade de fontes.

Um recurso de referência merecedor de destaque é o *corpus* denominado *corpus* de referência, o qual serve como parâmetro para o levantamento e analogias dos dados existentes em seu conteúdo. Alguns desses exemplos de *corpus* de referência da língua portuguesa são:

- Banco de português com 233 milhões de palavras. Privilegia o português brasileiro, escrito e falado. Construído pelos pesquisadores da PUC/SP.
- COMET (*Corpus* multilíngüe para ensino e tradução) com 5 milhões de palavras. Refere-se ao português escrito comparável com o inglês. Construído por pesquisadores da USP.
- *Corpus* UNESP/Araraquara/Usos do Português* com 200 milhões de palavras, privilegiando o português brasileiro escrito. Elaborado por pesquisadores da UNESP, Araraquara.

- CR-LW (*Corpus* de referência Lácio-Web)* com 5 milhões de palavras. Privilegia o português escrito. Composto por pesquisadores da USP, NILC.

Outros exemplos de *corpora* de estudos de língua inglesa são:

- *Brown Corpus (Brown University Standard Corpus of Present-day American English)*, de 1964. Contém 1 milhão de palavras em inglês americano escrito.
- *SEU Corpus (Survey of English Usage)*, de 1989. É composto por 1 milhão de palavras em inglês britânico, escrito e falado.
- *LLC (Lancaster-Oslo-Bergen)*, de 1978, contendo 1 milhão de palavras em inglês britânico escrito.

Para mais exemplos de *corpora* das línguas portuguesa e inglesa, consultar a obra *Linguística de corpus*, de Tony Berber Sardinha (vide capítulo das Referências).

Muitos podem ser os *corpora* escolhidos como referência desses idiomas, mas, nesta pesquisa, foram escolhidos o NILC, da língua portuguesa, e o *Brown*, da língua inglesa, que foram utilizados no levantamento dos verbetes, junto ao *corpus*, escolhidos como *corpus* de estudo, denominado *Copacabana runners*.

Segundo Berber Sardinha (2004, p. 8), além do *Brown*, outros *corpora* de referência da língua inglesa são: o BNC e o *Bank of English*.

O que faz notar se a ocorrência de determinada palavra possui papel relevante em um idioma é quando ela também ocorre, algumas vezes, no *corpus* de referência e não simplesmente no *corpus* da área estudada, pois o de referência, como dito anteriormente, é mais genérico e, conseqüentemente, menos influenciado por palavras específicas, ao contrário do de análise que é totalmente influenciado pelos termos da área específica.

Corroborando a idéia de gênero textual, conta-se com a abordagem de Biber (1998), referente a registros da língua (termos específicos), os quais podem ser bem restritos, dependendo do contexto dos falantes ou autores.

Segundo o artigo: *El registro, el género y la tipología textual* do site www.descargas.cervantesvirtual.com, o conceito de registro surge rapidamente nos estudos de tradução. Desde 1965, Catford defendia a necessidade de dividir a linguagem de alguma forma, já que o tamanho e a heterogeneidade da língua impediam os estudos de caráter descritivo, operativos e pedagógicos de serem realizados de uma forma operativa.

O autor menciona que em situações rotineiras, como uma conversa ou na escrita de dissertações acadêmicas, esses registros tendem a ser mais genéricos, ao contrário do que ocorre em textos técnicos.

De acordo com Baker (1993), as traduções devem manter uma relação de equivalência a mais próxima do original, tanto no aspecto semântico quanto no aspecto formal. Baker (*ibid*, p. 248) afirma que tais mudanças geram a necessidade de utilização de *corpora* computadorizados nos estudos da tradução.

Como ainda serão vistos, alguns registros ou termos não terão equivalentes em português, o que, segundo Catford (1965), é um dos problemas centrais da prática de tradução.

Como colaboração na busca desses registros e de suas respectivas equivalências, Berber Sardinha (2004) remete à noção de concordância, uma relação de exemplos de determinada palavra ou expressão, com seus respectivos exemplos nos contextos em que aparecem, realizada por uma das ferramentas do programa *WordSmith*, denominada *Concord*, para a confecção do glossário.

O *Concord* serve para fazer a busca das ocorrências de um item específico dentro dos diferentes contextos em que aparece no *corpus*, acompanhado do texto ao seu redor, com o uso da ferramenta *Search Word*, onde se digita a palavra que se deseja procurar. Há também a probabilidade de se fazer a concordância com outras palavras, as quais aparecerão nas linhas de Concordância, com seus devidos contextos.

Essa ferramenta é considerada importante porque exhibe as expressões e os contextos em que essa palavra aparece. Isso também pode ser confirmado com Krieger e Finatto (2004) em que uma análise abrangente de termos, textos e da linguagem especializada deve ir além de uma simples observação de palavras isoladas.

5.3 Coleta do material a ser estudado

Quanto à coleta do material utilizado para estudo, foi do *site* sobre Fisiologia do Exercício, denominado *Copacabana runners*, que foram retirados pequenos artigos científicos e suas respectivas traduções, totalizando 15.783 palavras/*tokens*, sendo sete textos em inglês e sete em português. As traduções dos textos, utilizadas como *corpus* paralelo, ou de referência, já estão digitalizadas e são do mesmo *site*.

É relativamente fácil encontrar materiais sobre a Fisiologia do Exercício ou temas afins em um dos maiores mecanismos de busca disponíveis na internet o: www.google.com com um tamanho considerável, contendo os mais variados tipos de textos de quase todas as áreas, o que faz com que se torne praticamente impossível não encontrar qualquer espécie de texto ou *site* em seu acervo.

Embora haja uma emergente preocupação com a saúde e com a estética, fatores tão relevantes na sociedade atual, e uma forte demanda de muitos internautas por *sites* relacionados a temas como emagrecimento com saúde, atividades físicas, o *Copacabana runners* foi o único *site* que disponibiliza seus textos em diversos idiomas. O site dispõe de pequenos artigos sobre saúde e condicionamento físico, de fácil compreensão para todos os leitores interessados no assunto e foi o único site encontrado na área de fisiologia do exercício, com a opção de tradução para vários idiomas, o que possibilitou seu uso como corpus de estudo, neste trabalho. O site é de responsabilidade de Hélio Fontes (1999-2008) e foi originalmente escrito em português e traduzido para diversos idiomas tais como: sueco, italiano, espanhol, inglês, alemão e francês.

Nota-se, portanto, uma lacuna na presença de outros *sites* como opção ao leitor que deseja as informações em outras línguas.

A maneira de se escolherem os termos será selecionando um determinado termo e designando suas diferentes variantes socioprofissionais encontradas no *corpus* de estudo. Vale lembrar que, ao invés de buscar os textos impressos, torna-se mais interessante buscá-los já digitalizados, assim como se fez neste trabalho, pois essa seria uma tarefa muito custosa, que levaria muito tempo para corrigir os prováveis erros, oriundos de textos escaneados.

Abaixo estão os textos escolhidos como corpora de estudo na dissertação e que são específicos da área de interesse:

- a) *Concepts of Exercise Physiology for Runners* (Conceitos de Fisiologia do Exercício para Corredores)
- b) *Herbs for Runners?* (Ervas para Corredores?)
- c) *So why is Broccoli so healthy?* (Por que brócolis é tão saudável?)
- d) *The food guide pyramid* (Pirâmide alimentar e dieta equilibrada)
- e) *Common Training Mistakes* (Erros comuns na corrida)
- f) *Exercise and Bone Health* (Exercícios físicos e a saúde dos ossos)
- g) *Eating for Peak Performance* (Alimentação para maximizar a performance)

5.4 Dimensão dos *corpora* selecionados:

Segundo Berber Sardinha (2004, p.25), há três abordagens nas quais se inserem as classificações de extensão de um *corpus*:

1) Impressionística: Derivada das afirmações de peritos da área, os quais criaram e exploraram *corpora*. Dentre eles, encontra-se Leech, que diz ser 1 milhão o mínimo desejado, e Aston, que considera de 20 a 200 mil palavras, como um *corpus* pequeno e de 100 milhões ou mais como um de tamanho grande. Há, ainda, Sinclair que diz que o tamanho do *corpus* depende do quanto a tecnologia da época lhe permitir.

2) Abordagem histórica: Berber Sardinha (2004, p.26) sugere a seguinte classificação com base na observação dos *corpora* utilizados, durante quatro anos, em conferências de Lingüística de *Corpus*:

| Tamanho em palavras | Classificação |
|-----------------------|---------------|
| Menos de 80 mil | Pequeno |
| 80 a 250 mil | Pequeno-médio |
| 250 mil a 1 milhão | Médio |
| 1 milhão a 10 milhões | Médio-grande |
| 10 milhões | Grande |

3) Abordagem Estatística: Fundamentada em teorias com base estatística. Biber utiliza fórmulas matemáticas a fim de identificar mínimas quantidades de itens lexicais, gêneros e textos que formariam uma amostra representativa.

O *corpus* analisado neste trabalho pode ser considerado de pequeno porte, de acordo com a classificação mencionada por Berber Sardinha (2004), em sua abordagem histórica, que discute a importância de um *corpus* possuir, no mínimo, 10.000 palavras para ser analisado.

Biber (1998) também ressalta a importância da inclusão de um *corpus* de referência com grande número de gêneros textuais na comparação de *corpora* realizada pelo *WordSmith*, pois quanto maior for o *corpus*, mais variado e imparcial ele será, contendo menos termos específicos de uma determinada área, os quais possuem muitos jargões, e não a realidade de uso do idioma em situações rotineiras.

Um dos motivos pelo qual se faz necessário a comparação do *corpus* de estudo com um de referência da língua, é que, se o levantamento dos itens fosse realizado somente com os *corpora* de estudo, não se teria como julgar se o resultado obtido nesse levantamento decorreria simplesmente do fato de a análise ter sido realizada com textos específicos da área, gerando, assim, um resultado parcial. Ao se analisarem *corpora* das mais variadas origens e gêneros, como os de referência, por exemplo, obtém-se um resultado mais satisfatório, pois neles há uma quantidade mais extensa de palavras existentes na língua, as quais funcionam como um modelo padrão do idioma estudado.

Os textos contidos nos *corpus* de referência servirão para comparar a diferença entre as freqüências de cada uma das palavras pesquisadas no *corpus* de análise e no *corpus* de referência, facilitando assim a busca de características específicas dessa palavra na análise contrastiva. Dessa forma, será possível saber se as ocorrências de maior freqüência dos termos técnicos analisados são constantes na língua como um todo.

As palavras de baixa freqüência em uma língua têm probabilidade de baixa ocorrência no *corpus*. Para que um *corpus* seja considerado representativo, é necessário que seja extenso para que as palavras de baixa ocorrência tenham a chance de aparecer com mais freqüência, as quais, muitas vezes, podem ser as palavras com um significado de maior representatividade.

Dessa forma, o leitor deverá estar se perguntando se, no caso da elaboração de um glossário, o que mais importa é dar prioridade a todas as palavras mais freqüentes no texto. A resposta a essa pergunta é não, já que muitas vezes uma palavra, tal como um artigo, pode aparecer mais de 100 vezes em um *corpus*, e um determinado verbo da língua surgir apenas umas 10 vezes, com mais de um sentido. Mesmo com uma freqüência maior do que os verbos, alguns itens lexicais, tais como os artigos e as preposições, são considerados menos relevantes, pois trazem menos carga semântica consigo, demonstrando que nem sempre a freqüência de um item lexical será o fator decisivo na escolha dos verbetes a serem inseridos no glossário.

5.5 Procedimentos utilizados na elaboração do glossário

Após tecer algumas considerações a respeito da tipologia e dimensão dos *corpora* estudados e dando continuidade aos procedimentos metodológicos, seguir-se-á adiante, dividindo essa etapa em passos a serem percorridos pelo pesquisador interessado na criação do próprio glossário:

1º PASSO - um passo a ser percorrido será o de incluir neste glossário exemplos retirados dos textos eletrônicos mencionados acima e suas respectivas traduções para o português, de forma a enriquecer o material de consulta do acadêmico na tradução e interpretação de textos técnicos.

2º PASSO – Gerando-se o resultado do processamento da ferramenta *Wordlist* do programa *WordSmith*, aparece a lista de todas as palavras encontradas nos sete artigos sobre Fisiologia do Exercício, com suas frequências respectivas em inglês.

3º PASSO - Em seguida, utiliza-se o comando *Find the Keywords in a text*, que exibe uma lista de *Keywords* (palavras-chave), com maior precisão dos itens que realmente serão úteis ao glossário, pois a lista de palavras (*Wordlist*) analisa todos os itens lexicais que estão inseridos no *corpus* de pesquisa, e a de palavras-chave (*Keywords*) mostra somente as mais importantes, segundo critérios dos testes de significância.

A ferramenta intitulada *Keywords* faz a comparação da lista de palavras consideradas mais importantes com um *corpus* de referência, que no caso do inglês é o *Brown*, o que não acontece quando se aciona a ferramenta *Wordlist*, pois essa não utiliza um *corpus* de referência para contrastar com o *corpus* de análise.

Na Figura 2, há os trinta primeiros exemplos retirados da lista elaborada por meio do comando *Keywords* dos *corpora* em inglês:

Figura 2 - As trinta primeiras palavras-chave encontradas no *corpus* em inglês

| N | WORD | FREQ. | LST % | BROWN1.TXT | | KEYNESS | P |
|----|----------|-------|-------|------------|------|---------|---|
| | | | | FREQ. | % | | |
| 01 | YOUR | 137 | 1,67 | 914 | 0,08 | 568,1 | 0 |
| 02 | BONE | 50 | 0,61 | 37 | | 381,1 | 0 |
| 03 | EXERCISE | 51 | 0,62 | 57 | | 360,6 | 0 |

| | | | | | | | |
|----|---------------|-----|------|-------|------|-------|---|
| 04 | YOU | 141 | 1,72 | 3.234 | 0,27 | 281,5 | 0 |
| 05 | HEART | 52 | 0,63 | 174 | 0,01 | 277,8 | 0 |
| 06 | RUNNING | 44 | 0,54 | 124 | 0,01 | 247,8 | 0 |
| 07 | GLYCOGEN | 23 | 0,28 | 0 | | 229,6 | 0 |
| 08 | RATE | 47 | 0,57 | 222 | 0,02 | 223,1 | 0 |
| 09 | TRAINING | 39 | 0,47 | 157 | 0,01 | 195,9 | 0 |
| 10 | MUSCLES | 26 | 0,32 | 30 | | 182,6 | 0 |
| 11 | CALORIES | 20 | 0,24 | 7 | | 168,8 | 0 |
| 12 | PROTEIN | 23 | 0,28 | 25 | | 163,5 | 0 |
| 13 | CARBOHYDRATES | 15 | 0,18 | 0 | | 149,7 | 0 |
| 14 | DIET | 20 | 0,24 | 20 | | 144,5 | 0 |
| 15 | CARBOHYDRATE | 15 | 0,18 | 1 | | 142,2 | 0 |
| 16 | RUN | 35 | 0,43 | 245 | 0,02 | 141,8 | 0 |
| 17 | FOODS | 23 | 0,28 | 51 | | 138,6 | 0 |
| 18 | ENERGY | 26 | 0,32 | 104 | | 130,9 | 0 |
| 19 | GRAMS | 18 | 0,22 | 18 | | 130 | 0 |
| 20 | PYRAMID | 14 | 0,17 | 2 | | 127,7 | 0 |
| 21 | MUSCLE | 19 | 0,23 | 44 | | 113,1 | 0 |
| 22 | FOOD | 26 | 0,32 | 154 | 0,01 | 113 | 0 |
| 23 | BEATS | 13 | 0,16 | 4 | | 111,3 | 0 |
| 24 | PACE | 18 | 0,22 | 43 | | 106,2 | 0 |
| 25 | MAX | 14 | 0,17 | 12 | | 104 | 0 |
| 26 | FIBERS | 15 | 0,18 | 23 | | 99 | 0 |
| 27 | DENSITY | 16 | 0,19 | 35 | | 96,7 | 0 |
| 28 | CAN | 60 | 0,73 | 1.751 | 0,15 | 96,1 | 0 |
| 29 | TWITCH | 11 | 0,13 | 3 | | 95,3 | 0 |
| 30 | HERBS | 11 | 0,13 | 3 | | 95,3 | 0 |

Quando se observa a tabela dos 30 primeiros itens do *keywords* exibida pelo programa, observa-se que os exemplos que podem gerar uma maior dificuldade em sua interpretação são os itens de número:

- 1- *Bone* - osso
- 5- *Heart* - coração
- 6- *Running* - corrida
- 8- *Rate* - frequência
- 16- *Run* - correr
- 17- *Foods* - refeições
- 21- *Muscles* - músculos
- 23- *Beats* - batidas
- 24- *Pace* - ritmo
- 25- *Max* - máximo
- 28- *Can* - pode
- 29- *Twitch* - Contração

Esses seriam os itens que causariam maiores conflitos em seu entendimento, devido ao fato de não possuírem nenhuma semelhança com termos existentes na língua portuguesa.

Dessa maneira, na tentativa de ilustrar melhor o significado real de alguns itens mais complexos, há os exemplos abaixo de cada item, os quais foram retirados dos *corpora* de estudo em inglês com a finalidade de inserir o item de interesse em um determinado contexto.

As outras palavras podem ser mais fáceis de se interpretar pela semelhança com o português e pelo contexto em que estão inseridos, como no item de nº 21 - *Muscles*, o qual é interpretado devidamente como músculos, facilitando a compreensão do leitor, de nível mais básico.

5º PASSO - Com base na tabela apresentada, torna-se necessária a observação dos termos que aparecem com maior frequência no *corpus* de análise da segunda coluna, o qual será comparado com o de referência da quarta, pela proporção da frequência com que cada palavra ocorre dentro dos *corpora* de análise e de referência.

No final da lista, na última coluna mostrada pelo programa, a probabilidade da diferença entre as frequências demonstradas pela comparação entre os dois *corpora* ser devida ao acaso, ou seja, ser mera coincidência, sem maior importância, será medida por meio dos testes de significância: o qui-quadrado e razão de probabilidades logarítmicas, operações estatísticas realizadas automaticamente pelo programa.

Na coluna denominada *Keyness* da tabela apresentada acima, verificou-se o grau de importância dos efeitos observados, os quais expressam o resultado do teste de significância, denominada razão de probabilidades logarítmicas, feito pelo *Wordsmith*. Tal teste é realizado a partir da comparação entre as palavras-chave no *corpus* de estudo e da frequência em que as mesmas palavras aparecem no *corpus* utilizado como referência.

Tomando o item lexical *Your* como exemplo, não se deve deixar de considerar as diferentes dimensões entre os dois *corpora*, utilizados na análise contrastiva, pois o item aparece 137 vezes no *corpus* de análise, referente a 1,69% do total de palavras do *corpus* e 914 vezes no *corpus* utilizado como referência, o que corresponde a 0,08% do total de palavras do *corpus*. Isso deve ser considerado como um resultado de considerável importância, pois os *corpora* de referência são bem maiores do que os de análise, gerando um grau de centralidade (*Keyness*) de 571,4, onde *p* (a probabilidade) da importância registrada nesse item ser devida ao acaso é menor que uma em um milhão, expressa na última coluna.

Outro fato importante demonstrado na tabela acima é que, embora a caracterização de do conteúdo demonstrado na tabela e na língua observada geralmente seja feita por meio de itens lexicais de carga semântica vinculada a aspectos temáticos desse gênero, as preposições, artigos e conjunções dificilmente caracterizam um gênero textual, pois seu conteúdo semântico pouco acrescenta ao léxico.

É interessante também não direcionar este estudo somente para termos isolados, mas para todo o texto relacionando, dessa maneira, os três objetos da Terminologia: termo, fraseologia e definição, pois eles podem auxiliar na busca do sentido do termo, dentro de um contexto.

Como destacam Krieger e Finatto (2004), nas concepções da definição de um termo, é necessário distingui-lo da palavra, a qual é composta de forma e conteúdo, ao contrário do termo, o qual se deve buscar o conceito lingüístico correspondente ao signo, implicando significado dentro de um sistema de significantes. Segundo as autoras, a definição pode ser vista como um elemento-chave no enunciado do conhecimento especializado, tecnológico ou científico, por estabelecer seus elementos mais significativos. É por isso que é um dos principais objetos de estudo da Terminologia.

Já a fraseologia, segundo Krieger e Finatto (2004, p. 88), “é um fenômeno de colocação, bem como de constituição de sintagmas e locuções”, um enunciado-texto, que deve conter significados de termos ou expressões técnicas especializados nos conceitos de uma área específica. Assim sendo, sua aplicação torna-se um requisito na elaboração de glossários e dicionários.

É nesse cenário que se percebe que, mesmo sendo de grande auxílio para alunos universitários, e até mesmo para profissionais, algumas técnicas provenientes de didáticas para a leitura de textos em inglês, como as do *English for Specific Purposes (ESP)*, utilizadas no ensino do inglês em muitas universidades brasileiras, ajudam, mas não são suficientes para atingirem um pleno desempenho na sua interpretação.

Figura 3 – Ferramenta Concord do *Wordsmith*:



Outro comando de grande relevância no programa utilizado é o *Concord* (Figura 3), pois ele faz um levantamento das diferentes ocorrências em que aparecem determinados termos, o qual auxiliará na escolha dos que devam aparecer juntos na lista.

6º PASSO - Para fazer a concordância de uma palavra na geração de um glossário, deve-se primeiramente observar as de maior frequência na lista denominada *Keywords* e gerar sua concordância na ferramenta denominada *Concord* pela associação das palavras mais frequentes.

Tal comando vai demonstrar todas as ocorrências dessas associações nos mais distintos contextos, o que também ajudará a localizar expressões recorrentes da área.

Se o interesse se restringe a buscar uma determinada expressão, pode-se procurar essa expressão sem a necessidade de buscar em todos os exemplos do item pesquisado encontrados na lista realizada pelo *Concord*.

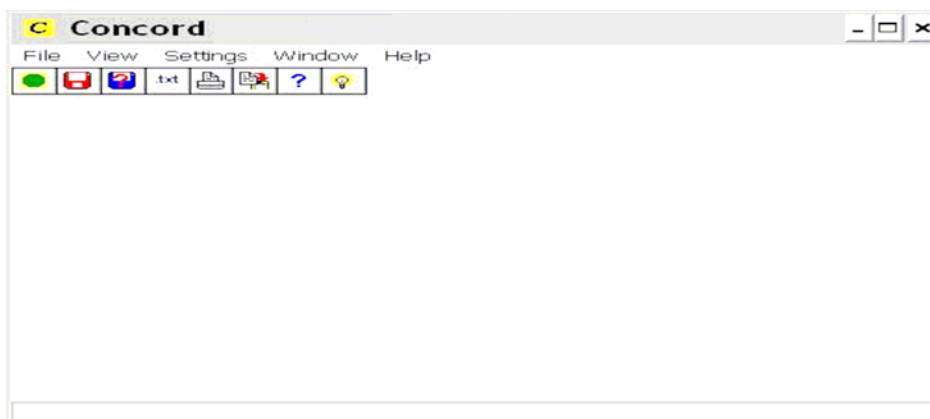
Um exemplo que será visto na lista de concordância demonstrada mais adiante no 8º PASSO é a associação das palavras *exercise* e *physiology*.

Isso auxilia o pesquisador a verificar os diferentes contextos em que ocorre uma determinada palavra, mostrando as palavras que a acompanham, já que o concordanceador a isola do todo, o que é de grande benefício para os estudos da tradução, exibindo diversas situações em que o item lexical aparece no *corpus*, ajudando assim a verificar a tradução mais apropriada para determinada palavra. Tal fato já não ocorre em dicionários, pois os exemplos

fornecidos em muitos deles muitas vezes não se aplicam ao dia-a-dia das pessoas e podem não ser da área de interesse do pesquisador, contendo muitas vezes termos não muito usuais.

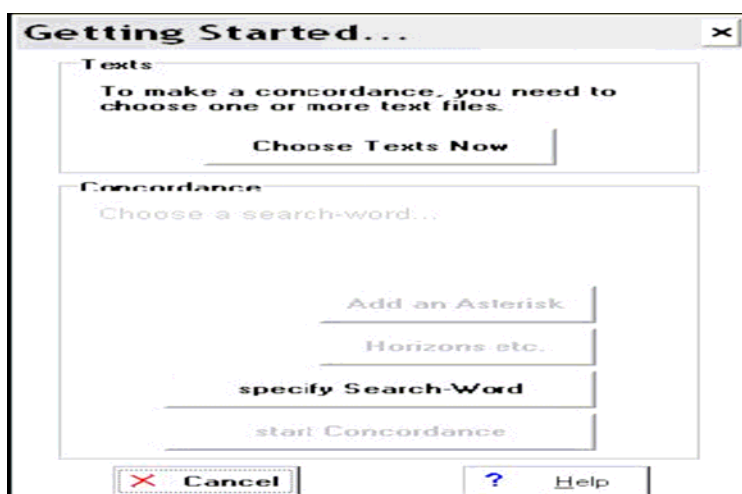
Na Figura 4, pode ser visto um leiaute do concordanceador.

Figura 4 - Tela principal do Concordanciador do *WordSmith*:



7º PASSO - A seguir, deve-se clicar em *File* e em seguida em *Start*, onde será aberto o item *Getting Started*, e em *Choose Texts Now*, o qual fornecerá meios de escolher os textos em que há interesse para pesquisar um item lexical. Em seguida, deve-se acionar o comando *Specify Search-Word* (Tabeka 5) e digitar quais as palavras que se quer saber se coocorrem.

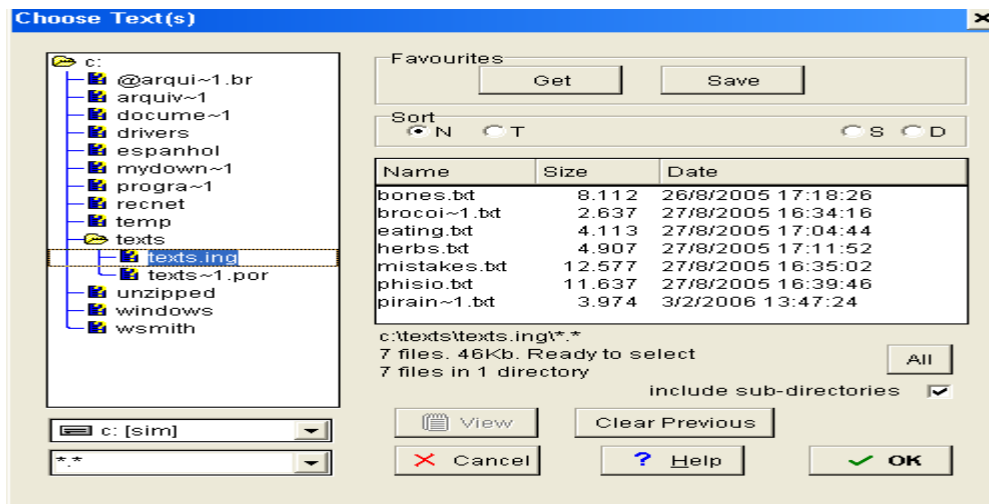
Figura 5 - Janela de seleção de *corpus*



Há também a opção de mudar os textos e as palavras selecionadas para análise em *Change Selection* ou *Change Search -Word*. Basta acioná-los e escolher os textos e palavras de interesse.

8º PASSO - O próximo passo será o de clicar em *Choose texts now* (Figura 6) para fazer a escolha dos *corpora* de interesse:

Figura 6 - Janela de *corpora* selecionados



Há a opção de se escolher um texto ou mais da pasta selecionada, ou de se escolher todos, apenas clicando no item *All*, como demonstrado na figura acima. Após esse procedimento, basta clicar em *OK* para voltar a janela *Getting Started*.

A seguir, apresenta-se um exemplo de concordância em que aparece a palavra *exercise* junto de *physiology*, extraída do *corpus* de análise, com suas ocorrências:

| N | Concordance % | Set Tag | Word No. | File |
|---|--|----------|----------|-----------------------------|
| 1 | Iso the stuff of <i>exercise physiology</i> . | Exercise | 45 | c:\texts\txts.ing\phsio.txt |
| 2 | concepts from <i>exercise physiology</i> for runners: | | 106 | c:\texts\txts.ing\phsio.txt |
| 3 | ace next time. <i>Exercise physiology</i> says your ti | | 757 | c:\texts\txts.ing\phsio.txt |
| 4 | of <i>Exercise Physiology</i> for Runners | | 3 | c:\texts\txts.ing\phsio.txt |
| 5 | physiology. <i>Exercise physiology</i> is the scienc | | 47 | c:\texts\txts.ing\phsio.txt |

Nos exemplos anteriores, têm-se as concordâncias das palavras *exercise* e *physiology*, especificadas juntas no *Concord*, onde esses itens ocorrem juntos. Se houver somente a observação da lista de palavras ou a de palavras-chave não há como saber que tais palavras coocorrem, ao menos que haja uma leitura de todos os textos detalhadamente. É graças a essa ferramenta que os alunos podem retirar exemplos de todos os contextos onde aparecem os itens lexicais de interesse.

A seguir, os exemplos em que se digitou o item lexical *exercise* separado da palavra *physiology*:

| FILE | CONCORDANCE % | WORD N° |
|---------------------------------------|--|---------|
| 01 c:\texts\texts.ing\bones.txt | sitive bone benefit from <i>exercise</i> , with greater lo 77 | 995 |
| 02 c:\texts\texts.ing\bones.txt | sitive bone benefit from <i>exercise</i> , with greater lo 28 | 359 |
| 03 c:\texts\texts.ing\mistakes.txt | competition. Mike is an <i>exercise physiologist</i> wh 60 | 1.319 |
| 04 c:\texts\texts.ing\eating.txt | lved in heavy resistance <i>exercise</i> or prolonged en 84 | 558 |
| 05 c:\texts\texts.ing\bones.txt | ss is complete, regular <i>exercise</i> can have a prot 35 | 456 |
| 06 c:\texts\texts.ing\bones.txt | ss is complete, regular <i>exercise</i> can have a prot 85 | 1.092 |
| 07 c:\texts\texts.ing\bones.txt | e loss in older adults. <i>Exercise</i> and Peak Bon 55 | 698 |

Caso se deseje utilizar qualquer uma das frases acima como exemplo para o glossário, deve-se clicar com o botão direito do *mouse* em cima da frase que interessa e em seguida no item *View text*, em que uma janela se abrirá com o texto em que consta o item de interesse. O item pesquisado aparecerá destacado nas linhas em que constar. Os exemplos utilizados no glossário foram retirados dos textos por meio desse procedimento.

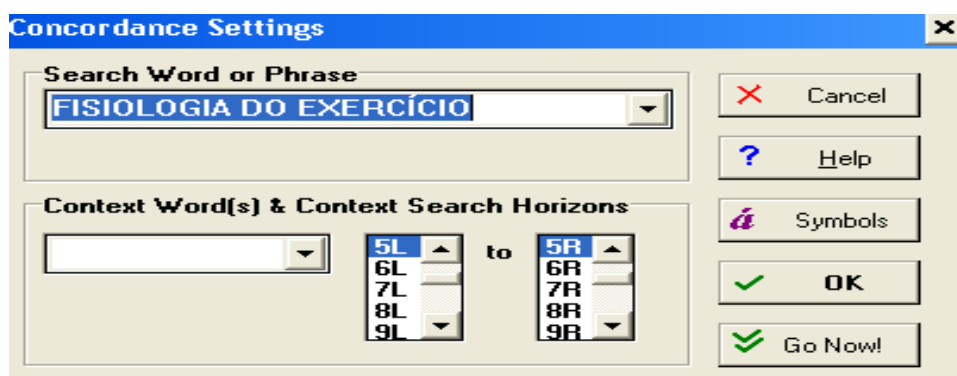
A partir da ferramenta *Concord* do *Wordsmith* exibida acima, tem-se uma idéia das coocorrências de um dos itens mais freqüentes da tabela, *exercise*, o qual figura tanto como substantivo, na maioria dos casos abaixo, quanto verbo, o que se pode observar na linha 33.

Digitando apenas a palavra *exercise*, não se obteria o mesmo resultado da concordância com a associação dos itens *exercise* e *physiology*, conforme demonstrado acima.

A diferença de maior destaque é que, digitando um item lexical somente no comando *Specify Search-word* do *Concord*, se obteria todos os exemplos de associações presentes no item especificado, e não somente o item de interesse.

Ainda no exemplo visto, o item *exercise* coocorre algumas vezes junto à palavra *physiology/physiologist*, nos itens 2, 23 e 24, o que remete ao objeto principal de estudo, a Fisiologia do Exercício, demonstrando a importância de tal vocábulo constar no glossário.

Figura 7 - Exemplo de busca da concordância de um item lexical



Como se pode observar na Figura 7, quando se faz a busca de uma determinada palavra, ela se situa no meio, rodeada de todos os contextos em que aparece. Na tabela acima, o nódulo, a palavra que está sendo analisada, foi selecionado para aparecer no meio, com as cinco palavras que se situam à sua direita e com as cinco que se situam à esquerda, seus colocados. Tanto o termo nódulo quanto o termo colocado foram mencionados por Welker (2004, p. 140).

Resultado das ocorrências em que aparece a palavra “fisiologia” ocorre junto com “exercício”, no *corpus* em português:

| N | Concordance Set | Tag | Word No. | File | % |
|----|------------------------------------|---------------------------------|----------|------|-----|
| 01 | stão seis conceitos da | <i>fisiologia do exercício</i> | pa | | 117 |
| | c:\texts\texts~1.por\ fisiop~1.txt | | 6 | | |
| 02 | isiologia do exercício. | <i>Fisiologia do exercício</i> | é | | 54 |
| | c:\texts\texts~1.por\ fisiop~1.txt | | 3 | | |
| 03 | tante da próxima vez. A | <i>fisiologia do exercício</i> | di | | 796 |
| | c:\texts\texts~1.por\ fisiop~1.txt | | 40 | | |
| 04 | ão os componentes da | <i>fisiologia do exercício.</i> | | | 51 |
| | c:\texts\texts~1.por\ fisiop~1.txt | | 3 | | |
| 05 | Conceitos de | <i>Fisiologia do Exercício</i> | p | | 2 |
| | c:\texts\texts~1.por\ fisiop~1.txt | | 0 | | |

9º PASSO: Esse será o último passo a ser percorrido dentro do universo do *WordSmith* para a conclusão do glossário. Trata-se da utilização do programa como meio facilitador na busca de traduções para os termos estudados. Essa ferramenta é de imenso valor, já que o foco principal reside em facilitar a leitura e a interpretação de textos técnicos em inglês por alunos com um nível básico de inglês.

Para encontrar as traduções de cada sentença alinhadas uma a uma, sem ter que ficar comparando os textos e suas traduções à parte, há o recurso *Viewer and Aligner*, encontrado na janela do *Tools*, como uma das opções do comando *Utilities*. Após as etapas mencionadas acima, o próximo passo será clicar em *File Start*, depois em *Source Text* e selecionar a opção *All Files* e procurar os textos de interesse no diretório C, disponível na janela à direita.

Se primeiramente se deseja encontrar um texto em inglês, deve-se selecioná-lo na pasta do diretório C em que está salvo. Em seguida, deve-se acionar o comando *Translation*, que aparecerá numa janela à direita, onde será selecionado o mesmo texto em português, e, finalmente, clicar no item *View Dual Text* que se abrirá em seguida.

O resultado final desse procedimento será uma janela em que aparecerão os textos alinhados com suas respectivas traduções abaixo. Esse procedimento torna-se bem mais prático e eficiente na busca por traduções das palavras-chave analisadas pelo programa e inseridas no glossário do que se tivesse que buscar tais termos em dicionários genéricos, em *sites* específicos sobre o assunto, ou até mesmo comparando os textos de análise impressos.

5.6 Características sobre traduções

Para se obter sucesso na busca pela tradução correta dos termos de interesse, é necessário um conhecimento mínimo do inglês por parte do pesquisador, já que terá que induzir os significados corretos dos itens estudados.

No quadro de exemplos do texto com sua tradução, que será visto abaixo, há um caso específico de inversão de estrutura de vocábulos, muito comum no inglês, o que exige desembaraço do leitor na busca de sua tradução, pois, no exemplo, o item *even* (até) inicia a sentença do texto em inglês, sendo colocado após o sujeito em sua tradução.

| | |
|----|--|
| 19 | <!--L1, S 10-->Even muscle activity in areas distant from the spine and hip can positively impact bone density at these sites. |
|----|--|

| | |
|----|---|
| 20 | <!--L2, S 10-->Atividade muscular, até em áreas distantes da espinha e quadril, podem ter impacto positivo na densidade óssea dessas áreas. |
|----|---|

Nesse mesmo exemplo, há ainda dois outros termos que exigem do pesquisador um entrosamento com a língua inglesa, pois os vocábulos *spine* (espinha) e *hip* (quadril) e *sites* (locais) podem aparecer em textos de Fisiologia do Exercício.

Tais termos como *spine* e *hip* muitas vezes fogem ao controle da interpretação de alunos com nível básico, pois sua tendência é sempre buscar termos que possuam alguma semelhança com o português, ou seja, os cognatos.

Outro exemplo de inversão na estrutura da sentença pode ser observado com a palavra *physically active youth*, a qual se torna “jovem fisicamente ativa”, em sua tradução:

| | |
|----|--|
| 21 | <!--L1, S 11-->Not all physically active youth reap the skeletal benefits of exercise. |
|----|--|

| | |
|----|---|
| 22 | <!--L2, S 11-->Nem toda jovem fisicamente ativa colhe os efeitos benéficos do exercício ao esqueleto. |
|----|---|

Além das situações citadas acima, dependendo dos *corpora* estudados, há outras situações em que também é importante um conhecimento prévio do idioma por parte do pesquisador, como no exemplo em seguida:

| | |
|----|--|
| 15 | <!--L1, S 8-->Activities that are high impact and weight-bearing, such as running, jumping, soccer and volleyball, appear to be the most beneficial to the skeleton. |
| 16 | <!--L2, S 8-->Atividades de alto impacto e de levantamento de peso como correr, pular, jogar futebol e vôlei parecem ser as mais benéficas para o esqueleto. |

Na sentença demonstrada acima, há alguns itens que podem confundir um leitor com nível básico, pois, desde cedo, aprende-se, nas escolas, que todos os verbos com o final *ing* representam o gerúndio em inglês, o que nem sempre acontece. No caso do exemplo demonstrado acima, os verbos *running* e *jumping* representam sua forma no infinitivo, e não no gerúndio. Para saber disso, não é necessário que o leitor tenha sólidos conhecimentos da gramática da língua, mas que seja um leitor capaz de fazer as inferências necessárias na busca das traduções equivalentes.

A seguir, exemplos de como constarão os textos de interesse no inglês com suas traduções alinhadas:

| | |
|---------------------------|--|
| 94 Sentences and headings | |
| 01 | <!--L1, S 1-->Exercise and Bone Health |
| 02 | <!--L2, S 1-->Exercícios físicos e a saúde dos ossos |
| 03 | <!--L1, S 2-->Physical activity is an essential factor in bone health. |
| 04 | <!--L2, S 2-->Atividade física é uma fator essencial para a saúde dos ossos. |
| 05 | <!--L1, S 3-->The skeletal benefits of exercise have been demonstrated throughout the life cycle. |
| 06 | <!--L2, S 3-->Os benefícios dos exercícios para o esqueleto têm se manifestado por todo o ciclo de vida. |
| 07 | <!--L1, S 4-->Exercise can positively affect peak bone mass in children and adolescents; has been shown to help maintain or even modestly increase bone density in adulthood and; can assist in minimizing age related bone loss in older adults. |
| 08 | <!--L2, S 4-->Exercícios podem: afetar positivamente o pico de massa óssea em crianças e adolescentes; auxiliar a manter ou mesmo promover um modesto aumento na densidade óssea em adultos; auxiliar na diminuição da perda de massa óssea devida à idade em adultos mais velhos. |
| 09 | <!--L1, S 5-->Exercise and Peak Bone Mass By age 18, skeletal growth is nearly complete. |

| | |
|----|--|
| 10 | <!--L2, S 5-->Exercícios e o Pico de Massa Óssea. Em torno dos 18 anos o crescimento do esqueleto está quase completo. |
| 11 | <!--L1, S 6-->The amount of peak bone mass achieved and the subsequent rate of bone loss are the key determinants of bone density later in life. |
| 12 | <!--L2, S 6-->A quantidade do pico de massa óssea alcançado, e a subsequente taxa de perda são os fatores-chave para a massa óssea depois na vida. |
| 13 | <!--L1, S 7-->There is substantial evidence that exercise plays an important role in the acquisition of bone mass in youth. |
| 14 | <!--L2, S 7-->Há evidência substancial de que os exercícios físicos desempenham um papel importante na aquisição de massa corporal na juventude. |

Como se pode notar também, o *WordSmith* não é o programa mais indicado para quem se interesse por buscar traduções, pois requer a versão do mesmo texto em português, para poder alinhá-los. Além disso, o programa não alinha as palavras isoladamente, mas só exibe as traduções das sentenças alinhadas, como demonstradas acima. Seu papel principal, portanto, reside na busca de palavras-chave encontradas por meio de análises estatísticas de frequência, e não no fornecimento das traduções de palavras.

Dois desses exemplos, no *corpus* de estudo, são os itens de nº 135 e 136 do *Keywords*:

135- *Dehydration*

136- *Dehydrated*

Na tradução do item 135, encontra-se o equivalente: Desidratação, e na tradução do item 136, o equivalente encontrado aparece como Desidratado. Embora ambos os itens sejam derivados do mesmo verbo, essas palavras possuem sentidos diferentes.

6 RESULTADOS

6.1 Organização das entradas

NOTAS REFERENTES À ORGANIZAÇÃO DOS VERBETES:

Uma questão que preocupa os lexicógrafos, na confecção de seus dicionários, é qual a melhor maneira de se explorarem as diversas possibilidades de apresentação de um verbete. Por se tratar de um glossário e devido a uma infinidade de possibilidades, optar-se-á por uma sucinta apresentação de cada um dos verbetes, que irão conter:

- 1- Título
- 2- Tradução para o português
- 3- Exemplos elucidativos dos itens.
- 4- Indicação das características incomuns de caráter informativo, em suas reais situações de uso, o que, segundo Biber *et all* (1998), é considerado o estudo de uso da língua.

ORDEM DOS VERBETES: O glossário obedece à ordem alfabética das palavras consideradas “chave”, com a relação de suas palavras derivadas abaixo. As palavras derivadas foram listadas abaixo do item em todas as suas formas exploradas nos *corpora* de estudo.

PALAVRAS-CHAVE/KEY WORDS: a versão das palavras-chave em inglês aparece em itálico e negrito e suas traduções na forma usual, começando pela margem esquerda.

Nos exemplos, os itens em inglês aparecem em negrito e itálico, e na tradução o mesmo item aparece em itálico e em negrito.

OBSERVAÇÃO: Não foram fornecidas informações como: pronúncia, proveniente da Fonética, as definições das palavras, questões gramaticais como a sintaxe, sinônimos e antônimos e formas incomuns de conjugações verbais como as dos *Phrasal verbs*

ESTILO: As diferentes situações em que se usa uma palavra ou expressão foram listadas abaixo dos exemplos, como forma de ilustrar as características dos itens estudados.

Outro fato importante é que, embora, na maioria das vezes, os itens lexicais encontrados no corpus em inglês tenham um correspondente em português, isso nem sempre ocorre, pois haverá casos em que não existirá um equivalente para determinado termo, fazendo com que o tradutor opte por traduzi-lo ou não.

Um exemplo de ilustração desse fato está no título de um dos artigos Alimentação para maximizar a performance, utilizado para análise em português, onde o tradutor decide manter o termo performance, ao invés de procurar um equivalente ou sinônimo, como desempenho. Acredita-se que o tradutor decidiu manter essa palavra na língua de partida, pois seu uso já se consolidou na língua de chegada, testado no *corpus* de referência.

Uma outra maneira de auxiliar na tradução desses itens seria buscar correspondentes no *corpus* de análise, tentando substituí-los por outros sinônimos ou derivados, o que no caso do item lexical *performance* poderia ser traduzido como desempenho.

Há também casos em que o tradutor opta por domesticar os termos para o idioma de chegada, alterando-os em sua forma.

| | |
|----|---|
| N | 254 Sentences and headings |
| 91 | <!--L1, S 46-->Your muscles then become more acidic. |
| N | 254 Sentences and headings |
| 92 | <!--L2, S 46-->Então sua musculatura ficará mais ácida. |

EX: *muscles* = musculatura

No exemplo da tradução do item acima, o recurso utilizado pelo tradutor é conhecido como modulação, pois o tradutor opta por traduzir o item *muscles* como “musculatura” ao invés de utilizar a palavra “músculos”.

Segundo Vinay e Darbelnet (*apud* Shuttleworth e Cowie, 1997) é um termo utilizado para se referir a um dos sete métodos de tradução. É um tipo de tradução oblíqua, a qual não envolve o uso do paralelo língua-mãe – língua-fonte. Vinay e Darbelnet definem modulação como sendo uma variação da forma da mensagem, obtida por uma mudança do ponto de vista.

Segundo os mesmos autores, algumas modulações são fixas e inseridas em dicionários, outras são ensinadas. E ainda há as livres, não fixadas pelo uso.

Mesmo sabendo que há essas e inúmeras outras técnicas e recursos que podem vir a ser utilizados na tradução, procurou-se fornecer o máximo de informações a potenciais usuários no glossário de uma forma bem prática e objetiva para dar maior clareza a leitores de textos acadêmicos em inglês, que necessitem de um auxílio na compreensão de termos pertinentes à sua área de conhecimento.

Os exemplos demonstrados são como os itens aparecerão no glossário gerado neste trabalho, com as respectivas traduções e exemplos para esclarecer seu uso.

É importante que o pesquisador interessado na criação de seu próprio glossário tenha consciência da necessidade de se levantar uma quantidade bem maior de termos, pois os demonstrados aqui seriam de pouca utilidade como instrumento de pesquisa numa variedade de textos e servem somente como caráter ilustrativo.

Como mencionado na seção de Procedimentos Metodológicos, no 4º PASSO da elaboração do glossário não se levou em consideração os artigos e as preposições, levantados pelas ferramentas do *Keywords*, por não possuírem uma carga semântica de grande importância no léxico da língua.

Para que se possa criar um glossário mais extenso, com maior quantidade de jargões, é necessário que se colete uma quantidade bem extensa do *corpus* de estudo para que se possa aumentar o tamanho do objeto de estudo.

Já que o intuito deste trabalho é apenas demonstrar a metodologia de elaboração de glossários por meio do KeyWords, não se tornam necessárias as listas extensas de verbetes do glossário em sua totalidade. Dessa forma, seguem abaixo somente os primeiros itens do glossário elaborado:

| | |
|---------------------|---|
| Bone | Exercise and Peak <u>Bone</u> Mass. |
| Osso | Exercícios e o <i>Pico de Massa Óssea</i> |
| Bone density | Hormonal status, body composition, nutritional intake and medication use are also important predictors of <u>bone density</u>. |
| Densidade óssea | Situação hormonal, composição corporal, ingestão de nutrientes e uso de medicamentos também são influências importantes para a <i>densidade óssea</i> . |
| Bone healthy | Physical activity is an essential factor in <u>bone health</u>. |
| Saúde dos ossos | Atividade física é um fator essencial para a <i>saúde dos ossos</i> . |

| | |
|------------------|---|
| Bone loss | Once the phase of accelerated <u>bone loss</u> is complete, regular exercise can have a protective effect on bone. |
| Perda óssea | Assim que a fase de <i>perda óssea</i> acelerada está completa, exercício físico regular pode ter um efeito de proteção aos ossos. |
| Bone mass | There is substantial evidence that exercise plays an important role in the acquisition of <u>bone mass</u> in youth. |
| Massa óssea | Há evidência substancial de que os exercícios físicos desempenham um papel importante na aquisição de <i>massa corporal</i> na juventude. |
| Beats | Heart rate during running varies by a few beats from day-to-day. |
| Batidas | A frequência cardíaca durante a corrida varia em alguns batimentos de dia para dia. |

| | |
|------------------------------|--|
| Beats from day-to-day | Heart rate during running varies by a few <i>beats from day-to-day</i> . |
| Batidas no dia-a-dia | A frequência cardíaca durante a corrida varia em alguns batimentos de dia para dia. |
| Calories | If you run the marathon in 3 hours, therefore, you will take in about 600 <i>calories</i> during the race. |
| Calorias | Então, caso corra a maratona em 3 horas, tomará em torno de 600 calorias durante a corrida. |
| Calories per hour | Each gram of carbohydrate contains 4.1 calories, so you will be taking in 200 <i>calories per</i> |

| | |
|--------------------------------|---|
| | <i>hour.</i> |
| Calorias por hora | Cada grama de carboidrato contém 4,1 calorias, então você estará tomando 200 calorias por hora. |
| Carbohydrates | If you drink 7 oz. every 15 minutes (that's about the maximum amount that will empty from your stomach) of a drink containing 6% glucose, you will take in about 48 grams of carbohydrate per hour. |
| Alimentos ricos em carboidrato | Diferentes alimentos ricos em carboidratos podem afetar o seu nível de energia de forma variada. |
| Exercise Exercício | Not all physically active youth reap the skeletal |

| | |
|----------------------------|---|
| | <p>benefits of <i>exercise</i>.</p> <p>Nem toda jovem fisicamente ativa colhe os efeitos benéficos do exercício ao esqueleto.</p> |
| Exercise physiology | <p>Concepts of <i>Exercise Physiology</i> for Runners</p> |
| Fisiologia do exercício | <p>Conceitos de Fisiologia do Exercício para Corredores</p> |
| Diet | <p>Choose a <i>diet</i> low in fat, saturated fat, and cholesterol to reduce your risk of heart attack</p> |

| | |
|-------------------------|---|
| | and certain types of cancer and to help you maintain a healthy weight. |
| Dieta | Escolha uma dieta baixa em calorias, com gordura saturada e colesterol para reduzir o risco de ataque cardíaco, alguns tipos de câncer e à manter um peso saudável. |
| Downhill | When you run <i>downhill</i> your muscles contract eccentrically. |
| Ladeira abaixo | Quando você corre na descida, seus músculos contraem excentricamente. |
| Downhill running | He no doubt developed some |

| | |
|--------------------|---|
| | muscle soreness each time, but by marathon day, his legs were used to both uphill and downhill running. |
| Corrida na descida | Ele sem dúvida ficou com dores musculares depois de cada treino, porém no dia da maratona suas pernas estavam acostumadas a correr, tanto na subida, quanto em corridas na descida. |
| Energy | A high-carbohydrate diet increases stores of glycogen, the <i>energy</i> for muscles, and improves overall athletic performance. |
| Energia | Uma dieta rica em carboidratos eleva as reservas de glicogênio, a <i>energia</i> para os |

| | |
|-------------------------------|---|
| | músculos, e melhora a performance atlética em geral. |
| The Energy Diet | To have enough energy you need to consume enough energy. |
| A Dieta Energética | Para ter energia suficiente, você precisa consumir energia suficiente. |
| Energy production | Protein, therefore, supplies less than 10% of your <i>energy needs</i> , and then only if you are already low on glycogen. |
| Produção de energia | As proteínas desempenham um papel pequeno na produção de energia, contribuindo somente de 5% a 10% da energia usada durante os exercícios |

| | |
|----------------------|---|
| | prolongados. |
| Energy source | Fat is definitely an important energy source, particularly for athletes involved in prolonged, low-intensity activity. |
| Fonte de energia | Definitivamente gordura é uma importante fonte de energia, particularmente para atletas envolvidos em atividades longas e de baixa intensidade. (para atividades curtas de alta intensidade, carboidrato é a fonte primária de energia. |
| Foods | A healthy diet includes foods from all five major food groups in the Food Guide Pyramid. Balance. |

| | |
|-------------------------|--|
| Comida | Uma dieta variada inclui alimentos diferentes nos 5 grandes grupos da Pirâmide que juntos atendem às recomendações nutricionais. |
| Variety of foods | It takes in consideration seven guidelines: The Seven Guidelines: Eat a <i>variety of foods</i> to get the energy, protein, vitamins, minerals, and fiber you need for good health. |
| Variedade de comidas | Isto leva sete princípios em consideração: Os sete princípios: Varie os alimentos para ganhar energia, proteínas, vitaminas, minerais e as fibras que você necessita para a sua saúde. |

| | |
|-------------------------------------|---|
| <p>Protein-rich foods</p> | <p><i>The ergogenic diet</i> is based on the US Department of Agriculture's widely published food guide pyramid, which includes five basic groups: grains, fruits, vegetables, dairy foods, and <i>protein-rich foods</i>.</p> |
| <p>Alimentos ricos em proteínas</p> | <p>A <i>dieta ergogênica</i> está baseada na amplamente divulgada Guia da Pirâmide Alimentar publicada pelo Departamento da Agricultura dos Estados Unidos, a qual inclui cinco grupos básicos: frutas, vegetais, laticínios, e alimentos ricos em proteínas.</p> |

| | |
|------------------------|---|
| Glycogen | Your body can only store a limited amount of <i>glycogen</i> . |
| Glicogênio | Seu corpo pode armazenar somente uma quantidade limitada de glicogênio. |
| Glycogen stores | Similarly, long runs train your muscles to rely more on fat and less on carbohydrates at a given speed so your glycogen stores last longer. |
| Reservas de glicogênio | De forma similar, corridas longas treinam seus músculos a contarem mais com gorduras e menos com carboidratos numa dada velocidade, de modo que suas reservas de glicogênio durem mais. |

| | |
|-------------------------------|---|
| Grams | If you drink 7 oz. every 15 minutes (that's about the maximum amount that will empty from your stomach) of a drink containing 6% glucose, you will take in about 48 grams of carbohydrate per hour. |
| Gramas | Se você bebe 200ml cada 15 minutos (essa é a quantidade máxima que será esvaziada do seu estômago) de uma bebida contendo 6% de glicose, terá em torno de 48 gramas de carboidrato por hora. |
| Grams of carbohydrates | Moderate-glycemic-index foods may be added for the next 18 to 20 hours, with a goal |

| | |
|------------------------|--|
| | of consuming at least 600 <i>grams of carbohydrate</i> during the 24 hours after an intense workout or competition Fat. |
| Gramas de carboidratos | Alimentos com moderado índice glicêmico podem ser adicionados nas próximas 18 a 20 horas com o objetivo de consumir pelo menos 600 gramas de carboidratos durante às 24 horas seguintes ao treino intenso ou competição. |
| Grams of fiber | It is loaded with Vitamin A, high in folacin, potassium, calcium, and it has 5 <i>grams of fiber</i> (floret and stalk offers the highest fiber) per cup. |

| | |
|-------------------------|--|
| Gramas de fibra | Brócolis também contém Vitamina A, folacin, potássio, cálcio e 5 gramas de fibras alimentares por xícara. |
| Grams of protein | The average person needs approximately 0.8 grams of protein per kilogram bodyweight per day. |
| Gramas de proteína | Uma pessoa comum precisa aproximadamente de 0,8 gramas de proteína por kg de peso corporal por dia. |
| Heart | Training at 3K to 5K race pace will work your cardiovascular system to its limit, which will |

| | |
|----------------------------|--|
| | <p>help increase the stroke volume of your <i>heart</i> and improve your muscles' ability to use oxygen to produce energy aerobically.</p> |
| <p>Coração</p> | <p>Treinar no ritmo de corrida de 3km a 5km trabalhará seu sistema cardiovascular ao seu limite, o que ajudará a elevar o volume de bombeamento do seu <i>coração</i> e melhorar a habilidade dos seus músculos utilizarem oxigênio para produzir energia aerobicamente.</p> |
| <p>Heart attack</p> | <p>Choose a diet low in fat, saturated fat, and cholesterol to reduce your risk of <i>heart attack</i> and certain types</p> |

| | |
|---------------------------|---|
| | of cancer and to help you maintain a healthy weight. Choose a diet moderate in sugars. |
| Ataque cardíaco | Escolha uma dieta baixa em calorias, com gordura saturada e colesterol para reduzir o risco de ataque cardíaco, alguns tipos de câncer e à manter um peso saudável. |
| Heart rate | <i>Heart rate</i> increases at high temperatures. |
| Frequência cardíaca | A <i>frequência cardíaca</i> eleva-se com altas temperaturas. |
| Heart rate monitor | Using a heart monitor is a good way to prevent yourself from training too hard |

| | |
|---------------------------------|--|
| | on your easy days. |
| Monitor de frequência cardíaca | Utilizar um monitor de frequência cardíaca é uma boa forma de prevenir que treine muito forte nos dias leves. |
| Heart rate reserve | Keep your heart rate below 75% of your maximal heart rate (or 70% of your <i>heart rate reserve</i>) and let your body recover to allow high quality workouts on your hard training days. |
| Reservas de frequência cardíaca | Mantenha sua frequência cardíaca abaixo de 75%, da frequência máxima atingida (ou 70% da sua reserva cardíaca) e deixe o seu |

| | |
|------------------------------------|---|
| | <p>corpo recuperar-se, para permitir treinamentos de qualidade, em seus dias de treinamento mais pesados.</p> |
| <p>Heart rate zone</p> | <p>This means that if you set your <i>heart rate zones</i> based on your morning heart rates, and train in the afternoon, then you will train a bit less intensely than planned.</p> |
| <p>Zona de frequência cardíaca</p> | <p>Isto significa que se você arrumar suas zonas de frequência cardíaca, baseadas em seus batimentos cardíacos matutinos, e no treino, no período da tarde, então, você irá treinar um pouco menos intensamente do que o planejado.</p> |

| | |
|------------------------|---|
| Muscles | A high-carbohydrate diet increases stores of glycogen, the energy for <i>muscles</i> , and improves overall athletic performance. |
| Músculos | Uma dieta rica em carboidratos aumenta o armazenamento de glicogênio, a energia de músculos e melhora a performance do atleta. |
| Muscle activity | Even <i>muscle activity</i> in areas distant from the spine and hip can positively impact bone density at these sites. |
| Atividade muscular | Até mesmo as atividades musculares em locais distantes da coluna vertebral e da cintura, podem |

| | |
|----------------------------|--|
| | causar impacto na densidade óssea desses locais, positivamente. |
| Muscles of glycogen | If you exercise for longer than an hour, you can begin to deplete your <i>muscles of glycogen</i> . |
| Músculos de glicogênio | Se você se exercitar por mais de uma hora, você poderá encher os seus músculo de glicogênio. |
| Muscle weakness | Increased <i>muscle weakness</i> can compound the problem of low bone density by increasing the risk of falls and fracture. |
| Pyramid | What is the <i>Food Pyramid</i> ? Flexible and full of options, the Pyramid gives simple and easy orientation that |

| | |
|----------|--|
| | will help you choose what and how much to eat from each food group. |
| Pirâmide | O que é Pirâmide Alimentar? Flexível e com muitas variações, a Pirâmide oferece uma simples e fácil orientação que irá lhe ajudar em quê e no quanto comer de cada grupo alimentar. |

Quando se encontra um modelo de glossário como o demonstrado acima, com a tradução de seus termos e algumas frases retiradas dos textos estudados, tem-se a impressão de que tal processo não é nada complexo e que sua elaboração pode ser realizada em um curto período de tempo, pelo método simples.

Porém, observando o respectivo glossário acima, pode-se notar que nem todas as palavras são cognatas, (com raiz grega/latina), o que faz com que alguns termos se tornem mais complexos na sua interpretação. É justamente nesse aspecto que reside a real dificuldade de alguns alunos quando tentam interpretar essas palavras, as quais diferem totalmente da morfologia de palavras oriundas do grego e do latim.

No glossário confeccionado, foi a partir da tradução e da demonstração dos termos equivalentes aos cognatos e principalmente aos não-cognatos dos *corpora* analisados que se tornou mais clara e abrangente a leitura e interpretação de termos próprios da Fisiologia do Exercício e áreas relacionadas. Os exemplos demonstrados no contexto de situação de uso

também são de grande valia, pois permitem aos alunos tanger uma maior noção do real significado dos termos.

Isso ocorrerá com uma simples consulta, feita pelos alunos, aos itens constantes no glossário, na leitura de um texto de sua área e de dúvidas em relação a esses termos. Ou seja, o glossário que se elaborou funcionará como um manual de instruções, para consultas em caso de dúvidas, quando for lido algum texto específico da Fisiologia do Exercício em inglês.

Para a interpretação desses termos, na ausência de dicionários específicos impressos, acredita-se que a solução vigente para esse caso seria a pesquisa em algum dicionário genérico impresso, dicionários *on-line*, ou, até mesmo, em ferramentas de busca como a do *Google*.

Como foi visto na Justificativa, tudo isso seria muito mais prático se já houvesse um dicionário impresso com uma quantia razoável de termos técnicos da área estudada. Entretanto, nem sempre isso é possível, devido a algumas questões já levantadas no início desta dissertação, como a falta de recursos necessários para a aquisição desses dicionários por parte de muitos alunos e, em muitos casos, devido à própria inexistência de um dicionário específico daquela área.

Às vezes, encontra-se em *sites* de busca, como o *Google*, a tradução da palavra, proveniente de algum tradutor *on-line*, ou até mesmo em *sites* que fazem parte de seu universo científico, dentre os quais se consegue inferir seu significado, mas nem sempre isso ocorre.

Os leitores mais ávidos por respostas rápidas provavelmente deverão estar se perguntando se não seria mais rápido e eficaz procurar os significados desses termos em dicionários *on-line*, já que são muitos os recursos desse tipo disponíveis gratuitamente na internet.

Um exemplo de dicionário *on-line* que é facultado aos internautas de forma gratuita e completa é o www.answers.com, que pode ser visto, também, como uma enciclopédia *on-line*, pois, além das traduções do termo buscado, há explicações desse termo em seus diversos aspectos, com figuras e seus equivalentes. Há ainda outro recurso admirável nessa enciclopédia, que é a possibilidade de se ouvir a correta pronúncia do item buscado.

Devido a todos esses recursos encontrados no dicionário mencionado acima, esse seria o *site* perfeito para se buscar as traduções dos itens de interesse, sem a necessidade de se deparar com o trabalho de seguir os passos exigidos para a criação do glossário proposto nesta dissertação. Mas o fato de essa enciclopédia ser monolíngüe, com o inglês como única língua dominante, dificulta, ou até mesmo impede, que muitos usuários possam desfrutá-la por

completo. Além do mais, a origem do dicionário é genérica, e não específica de determinada área.

Um outro exemplo de tradutor *on-line* é o *Babel Fish* do site Alta Vista, outro recurso rápido e eficiente em alguns casos. O motivo de esse tradutor ser eficiente, somente em alguns casos, é que ao se procurar um termo de uma área específica, muitas vezes ficar-se-á sem uma resposta precisa, pois a dimensão dos *corpora* nele existente não é o suficiente para buscar palavras muito específicas.

Um exemplo disso é a palavra *Heart rate* que foi digitada depois de acionado o comando Inglês para Português disponível no programa, que forneceu a palavra “Taxa de coração” como resposta. Respostas como essa geram certa confusão ao usuário, porque não corresponde ao sentido real da palavra, que é “Batimento cardíaco”.

A parte boa desse programa é que ele possui diversas opções de traduções de um idioma para o outro, proporcionando maior agilidade ao pesquisador, quando consegue encontrar o real sentido que corresponde à tradução do item pesquisado.

Se for digitado “dicionários de inglês *on-line*” na barra de pesquisa do site www.google.com, encontram-se os seguintes dicionários como resposta:

- 1 - <http://www.wordreference.com/>
- 2- <http://www.yourdictionary.com/>
- 3- <http://www.freedict.com/onldict/por.html> (dicionário inglês-português)
- 4- <http://www.freesearch.co.uk/dictionary> (com expressões idiomáticas)

Ao se digitar a expressão *heart rate* no dicionário do item número 1, será encontrada a definição da expressão com sua categoria semântica e algumas expressões relacionadas ao tema. Em alguns casos também há a pronúncia do vocábulo pesquisado.

O dicionário que consta no item número 3 é bilíngüe, mas também não demonstrou ter a capacidade necessária de armazenamento de palavras para uma pesquisa mais específica, como a de que se trata aqui. Pode-se notar isso quando se digita a expressão *Heart rate* em inglês e solicita-se sua tradução para o português, não encontrando nenhum registro para essa expressão.

Já o item número 4 mencionado acima é monolíngüe e mostra as diferentes expressões que contém o item procurado, sem conter as explicações do seu significado. O que se encontra,

ao se clicar em qualquer dessas expressões relacionadas ao tema de interesse, são os *sites* que contêm o termo buscado e estão relacionados a ele.

Além de todos os dicionários mencionados acima, há ainda o *Babel fish*, considerado popular no mundo virtual, do *site* www.altavista.com, que funciona como um tradutor bilíngüe disponível em diversos idiomas. Esse tradutor é bem útil quando se está *on-line* e necessita-se a tradução de uma palavra, mas não é tão útil quando se trata de frases inteiras.

O motivo de o *Babel fish* não ser ideal para sua utilização na tradução de frases, ou até mesmo em textos inteiros, é que, assim como todos os outros dicionários *on-line* bilíngües, ele não possui a mesma capacidade de reflexão do ser humano, por ser realizada por uma máquina. Dessa forma, não se pode esperar que a máquina consiga relacionar todas as palavras e sentenças associando-as como se faz na tradução.

Embora, na perspectiva da atividade tradutória, a máquina dificilmente se equipare ao ser humano, os glossários elaborados por meio de *corpora* eletrônicos, com levantamento de frequências de palavras, atingem resultados bem mais satisfatórios do que os glossários realizados sem um levantamento de suas principais ocorrências, apesar de mais trabalhosos e demorados.

O motivo pelo qual os glossários elaborados por meio de *corpora* eletrônicos podem ser considerados mais úteis é que, embora mais complexos, uma vez realizados com uma quantia considerável de termos técnicos, servirão como instrumentos únicos de consulta em casos de dúvidas. Isso faz desaparecer a necessidade de ter que ficar perdendo tempo buscando a tradução das palavras que causarem dúvidas.

Há também os glossários mais concisos, que contêm apenas itens lexicais em inglês, com suas traduções para o português. É o caso do modelo de glossário com termos relacionados à internet, o qual se baseia somente na colocação desses termos com suas respectivas traduções, sem as suas situações de uso. Isso se dá por se tratar de uma área prática, com imensa rapidez de informações, pois, se obtivesse todos os exemplos dos contextos em que cada item aparecesse, seria gasto muito tempo, tanto dos elaboradores quanto dos leitores desses glossários.

Eis um modelo de glossário referente a termos da internet retirado do *site*: <http://visibilidade.net/tutorial/glossario-informatica.html>:

| |
|---|
| <p><i>Anchor text</i>, texto da âncora. O texto da âncora apresenta uma curta descrição do conteúdo apontado. Por exemplo, no apontador seguinte o texto da âncora é "Visibilidade em motores de busca".</p> |
| <p><i>Atuning</i>, adaptação. É o oposto de <u>tuning</u>, ou seja, o utilizador adapta-se ao desempenho do sistema. Por exemplo, se uma determinada página demora muito tempo para carregar o utilizador habitua-se a executar outra tarefa enquanto espera.</p> |
| <p><i>Backlinks; inlinks; incoming links</i>, ligações de entrada, ligações recebidas. Conjunto dos <i>links</i> que apontam para uma página da <i>Web</i>.</p> |
| <p><i>Backtrack</i>, retroceder. Voltar a um estado anterior.</p> |
| <p><i>Banner</i>, manchete. Anuncio embutido numa página web.</p> |
| <p><i>Bookmark</i>, favorito. Funcionalidade de um <i>browser</i> que permite guardar URLs de páginas para que sejam rapidamente acedidas no futuro.</p> |
| <p><i>Banned page</i>, página banida ou excluída. Significa que a página (ou <i>site</i>) foi banida dos resultados de um motor de busca por ser prejudicial ao bom funcionamento do motor de busca, porque usou <i>Spamming</i>.</p> |

O glossário deste trabalho, em sua estrutura, mantém uma grande semelhança com o glossário demonstrado acima por aquele ser tão prático e conciso quanto esse, constando as informações de maior interesse que são os termos técnicos e suas equivalências em português.

A diferença crucial que se pode observar em relação ao nosso glossário e ao aqui apresentado é que, embora o glossário, em sua forma final, seja visto como simples e conciso, os verbetes apresentados nessa área apresentam maior complexidade, na busca de suas traduções, pois são termos bem mais distintos da língua portuguesa, ao contrário da fisiologia do exercício. Acredita-se que essa maior complexidade na compreensão dos termos da área

ocorra devido ao fato de que esta é bem mais recente no universo acadêmico e está em constante mudança.

Glossários informatizados como esses da área da informática são bem mais recorrentes e disponíveis em *sites* de buscas. Acredita-se que essa maior facilidade em encontrá-los deve-se ao fato da urgência de alguns internautas em se adaptarem aos novos termos do mundo virtual em constante mudança e pela maioria dos seus termos tomarem emprestadas palavras do inglês, idioma universal e dominante no meio eletrônico.

Os glossários referentes à informática e áreas tecnológicas devem sofrer constantes atualizações, pois seus termos estão sempre mudando, conforme as exigências do mercado de trabalho.

7 CONSIDERAÇÕES FINAIS

Como foi demonstrado no decorrer deste trabalho, o uso de ferramentas eletrônicas tem papel preponderante nesse tipo de elaboração. Assim como este trabalho visa contribuir com futuros pesquisadores interessados no assunto, espera-se que esses também venham a contribuir com a divulgação e amplitude de novos horizontes referentes à Lingüística de *Corpus* aplicada à Terminologia.

Para que o pesquisador possa atingir seu objetivo de criar glossários, há a necessidade de certa familiarização com a Lingüística de *Corpus*, na busca das ferramentas necessárias para sua construção. A primeira ferramenta e, muitas vezes, uma das mais complexas é a seleção dos *corpora* em inglês, com suas traduções. Esse tem sido um dos maiores desafios dos pesquisadores interessados, pois encontrar as versões dos textos traduzidas para o português é uma das mais complexas questões envolvidas no processo vigente.

Outra ferramenta que se faz útil ao pesquisador é o conhecimento dos comandos existentes no programa utilizado para a confecção do glossário, no caso o *WordSmith*. O conhecimento dos comandos pode ser adquirido por meio do item “Ajuda”, no menu do programa. A questão de maior dificuldade encontrada na “Ajuda” é que as explanações são feitas todas em inglês, o que exige certo conhecimento do idioma, por parte do usuário.

Conforme demonstrado no decorrer do trabalho, quando se seguem os passos aqui propostos para a criação de um glossário, pode-se notar que o programa *WordSmith* não é um programa recomendado para se fazerem traduções de textos, mas sim para se fazer o levantamento dos itens mais importantes dos textos analisados. A tradução dos itens já deve ser inserida nos textos do programa; o que este faz é somente alinhar os idiomas selecionados, conforme demonstrado na tabela da página 55, o que, mais uma vez, exige do pesquisador certo conhecimento do idioma-alvo, já que ele terá que fazer as respectivas traduções dos itens levantados pelo programa. Nesse processo, quem decide quais as traduções que julga conveniente colocar em seu glossário é o próprio pesquisador, que vai retirar os itens levantados pelo programa, partindo sempre do mais freqüente para o menos freqüente, selecionando os itens por ele considerados importantes. Assim, no caso de os artigos e preposições serem apontados como os mais freqüentes no levantamento de itens mais importantes, elaborado pelo programa, o pesquisador deverá optar se vai colocá-los no

glossário ou não. Neste glossário, optou-se por inserir os substantivos, pois foram considerados como portadores de uma grande carga semântica. Já, os artigos, conjunções e preposições possuem um valor significativo de menor peso, se comparado aos substantivos e verbos de um idioma.

Neste trabalho, acredita-se ter atingido o objetivo de realçar o papel de destaque que a Lingüística de *Corpus* possui junto à Terminologia das diversas áreas científicas, auxiliando, dessa forma, no tratamento e propagação de seus termos.

Referências:

- BAKER, Mona. **Corpora in Translation Studies: an overview and some suggestions for future research**. Target. 7:2 223-243, 1995.
- BARBETTA, P. A. **Estatística aplicada às ciências sociais**. Florianópolis: Ed. da UFSC, 1994.
- BERBER SARDINHA, T. **Linguística de Corpus**. São Paulo: Manole, 2004.
- CATFORD, J. C. **A linguistic theory of translation: an essay in applied linguistics**. London: Oxford University Press, 1965.
- CRUZ, T. D.; SILVA V. A.; ROSAS M. **Inglês.com.textos para informática**. São Paulo: Ed. Disal, 2003.
- DURÃO, B. A. de B. A. **Linguística contrastiva: teoria de prática**. Londrina: Ed. Moriá, 2004.
- GAUDIN, F. **Pour une socioterminologie. Des problèmes semantiques aux pratiques institutionnelles**. Rouen: Publication de l'Université de Rouen, 1993.
- JAKOBSON, R. **Linguística e comunicação**. São Paulo: Ed. Cultrix, 1999.
- KRIEGER, G. da M.; FINATTO B. J. M. **Introdução à Terminologia: teoria e prática**. São Paulo: Ed. Contexto, 2004.
- ROCHA, Marco. **Métodos com base em corpus no processamento de linguagens naturais**. Não publicado.
- ROCHA, Marco. **Métodos estatísticos comuns em linguística de corpus: uma visão geral**. UFSC
- WALSH, A. **Statistics for the social sciences**. Nova York: Harper and Row, 1990.
- PARTINGTON, A. **Patterns and Meanings: using corpora for English language research and teaching**. Amsterdam / Philadelphia: John Benjamins Publishing Company, 1998.
- SINGER, P. ET AL. **Linguística**. (Leitura)- Cadernos Puc nº 16. São Paulo: Ed. Cortez, 1983.
- SHUTTLEWORTH, M; COWIE M. **Dictionary of translation studies**. Manchester: Ed. St Jerome, 1997.
- TORRES, C. M. **Paradoxos da Tradução**. Diário Catarinense. Florianópolis, 21 maio de 2005.
- WALSH, A. **Statistics for the social sciences**. Nova York: Harper and Row, 1990.
- WELKER, A. H. **Uma pequena introdução à Lexicografia**. Brasília: Thesaurus, 2004.
- DELTA: Documentação de Estudos em Linguística Teórica e Aplicada. Disponível em:

<www.scielo.br/scielo.php?pid=S010244502000000200005&script=sci_arttext&tlng=pt>. Acesso em: março de 2007.

Conselho Federal de Educação Física. Disponível em:

<www.confef.org.br/extra/conteudo/default.asp?id=%7B40260DA6-2A17-4122-9B9A-395937FA5CEC%7D>. Acesso em: novembro de 2005.

Babylon. Disponível em:

<www.babylon.com/display.php?id=115&tree=7&level=3&show_id=1754 (03/06/2005)>. Acesso em: março de 2005.

Copacabana runners. Disponível em: <www.copacabanarunners.net>. Acesso em: abril de 2005.

Instituto Aleixo. Disponível em: <www.fisiologiadoexercicio.com.br>. Acesso em: abril de 2005.

British National Corpus. Disponível em: <<http://www.natcorp.ox.ac.uk/corpus/> Maintained by: **BNC Webmaster** (bnc-queries@rt.oucs.ox.ac.uk) . © 2005, University of Oxford. >. Acesso em: julho de 2006.

ANEXO

Itens que fazem parte do glossário

| | |
|------------------------------|---|
| Bone | Exercise and Peak <u>Bone</u> Mass. |
| Osso | Exercícios e o <i>Pico de Massa Óssea</i> |
| Bone density | Hormonal status, body composition, nutritional intake and medication use are also important predictors of <u>bone density</u>. |
| Densidade óssea | Situação hormonal, composição corporal, ingestão de nutrientes e uso de medicamentos também são influências importantes para a <i>densidade óssea</i> . |
| Bone healthy | Physical activity is an essential factor in <u>bone health</u>. |
| Saúde dos ossos | Atividade física é um fator essencial para a <i>saúde dos ossos</i> . |
| Bone loss | Once the phase of accelerated <u>bone loss</u> is complete, regular exercise can have a protective effect on bone. |
| Perda óssea | Assim que a fase de <i>perda óssea</i> acelerada está completa, exercício físico regular pode ter um efeito de proteção aos ossos. |
| Bone mass | There is substantial evidence that exercise plays an important role in the acquisition of <u>bone mass</u> in youth. |
| Massa óssea | Há evidência substancial de que os exercícios físicos desempenham um papel importante na aquisição de <i>massa corporal</i> na juventude. |
| Beats | Heart rate during running varies by a few <i>beats</i> from day-to-day. |
| Batidas | A frequência cardíaca durante a corrida varia em alguns batimentos de dia para dia. |
| Beats from day-to-day | Heart rate during running varies by a few <i>beats from day-to-day</i> . |

| | |
|--------------------------------|---|
| Batidas no dia-a-dia | A frequência cardíaca durante a corrida varia em alguns batimentos de dia para dia. |
| Calories | If you run the marathon in 3 hours, therefore, you will take in about 600 <i>calories</i> during the race. |
| Calorias | Então, caso corra a maratona em 3 horas, tomará em torno de 600 calorias durante a corrida. |
| Calories per hour | Each gram of carbohydrate contains 4.1 calories, so you will be taking in 200 <i>calories per hour</i> . |
| Calorias por hora | Cada grama de carboidrato contém 4,1 calorias, então você estará tomando 200 calorias por hora. |
| Carbohydrates | If you drink 7 oz. every 15 minutes (that's about the maximum amount that will empty from your stomach) of a drink containing 6% glucose, you will take in about 48 grams of carbohydrate per hour. |
| Alimentos ricos em carboidrato | Diferentes alimentos ricos em carboidratos podem afetar o seu nível de energia de forma variada. |
| Exercise Exercício | Not all physically active youth reap the skeletal benefits of <i>exercise</i> . Nem toda jovem fisicamente ativa colhe os efeitos benéficos do exercício ao esqueleto. |
| Exercise physiology | Concepts of <i>Exercise Physiology</i> for Runners |
| Fisiologia do exercício | Conceitos de Fisiologia do Exercício para Corredores |
| Diet | Choose a <i>diet</i> low in fat, saturated fat, and cholesterol to reduce your risk of heart attack and certain types of cancer and to help you maintain a healthy weight. |
| Dieta | Escolha uma dieta baixa em calorias, com gordura saturada e colesterol para reduzir o risco de ataque cardíaco, alguns tipos de câncer e à manter um peso saudável. |
| Downhill | When you run <i>downhill</i> your muscles contract eccentrically. |
| Ladeira abaixo | Quando você corre na descida, seus músculos contraem excentricamente. |
| Downhill running | He no doubt developed some muscle soreness each time, but by marathon day, his legs were used to both uphill and downhill running. |

| | |
|---------------------------|---|
| Corrida na descida | Ele sem dúvida ficou com dores musculares depois de cada treino, porém no dia da maratona suas pernas estavam acostumadas a correr, tanto na subida, quanto em corridas na descida. |
| Energy | A high-carbohydrate diet increases stores of glycogen, the energy for muscles, and improves overall athletic performance. |
| Energia | Uma dieta rica em carboidratos eleva as reservas de glicogênio, a energia para os músculos, e melhora a performance atlética em geral. |
| The Energy Diet | To have enough energy you need to consume enough energy. |
| A Dieta Energética | Para ter energia suficiente, você precisa consumir energia suficiente. |
| Energy production | Protein, therefore, supplies less than 10% of your energy needs , and then only if you are already low on glycogen. |
| Produção de energia | As proteínas desempenham um papel pequeno na produção de energia, contribuindo somente de 5% a 10% da energia usada durante os exercícios prolongados. |
| Energy source | Fat is definitely an important energy source, particularly for athletes involved in prolonged, low-intensity activity. |
| Fonte de energia | Definitivamente gordura é uma importante fonte de energia, particularmente para atletas envolvidos em atividades longas e de baixa intensidade. (para atividades curtas de alta intensidade, carboidrato é a fonte primária de energia. |
| Foods | A healthy diet includes foods from all five major food groups in the Food Guide Pyramid. Balance. |
| Comida | Uma dieta variada inclui alimentos diferentes nos 5 grandes grupos da Pirâmide que juntos atendem às recomendações nutricionais. |
| Variety of foods | It takes in consideration seven guidelines: The Seven Guidelines: Eat a variety of foods to get the energy, protein, vitamins, minerals, and fiber you need for good health. |
| Variedade de comidas | Isto leva sete princípios em consideração: Os sete princípios: Varie os alimentos para ganhar energia, proteínas, vitaminas, minerais e as fibras que você necessita para a sua saúde. |
| Protein-rich | The ergogenic diet is based on the US Department of Agriculture's widely |

| | |
|-------------------------------|--|
| foods | published food guide pyramid, which includes five basic groups: grains, fruits, vegetables, dairy foods, and <i>protein-rich foods</i> . |
| Alimentos ricos em proteínas | <i>A dieta ergogênica</i> está baseada na amplamente divulgada Guia da Pirâmide Alimentar publicada pelo Departamento da Agricultura dos Estados Unidos, a qual inclui cinco grupos básicos: frutas, vegetais, laticínios, e alimentos ricos em proteínas. |
| Glycogen | Your body can only store a limited amount of <i>glycogen</i> . |
| Glicogênio | Seu corpo pode armazenar somente uma quantidade limitada de glicogênio. |
| Glycogen stores | Similarly, long runs train your muscles to rely more on fat and less on carbohydrates at a given speed so your glycogen stores last longer. |
| Reservas de glicogênio | De forma similar, corridas longas treinam seus músculos a contarem mais com gorduras e menos com carboidratos numa dada velocidade, de modo que suas reservas de glicogênio durem mais. |
| Grams | If you drink 7 oz. every 15 minutes (that's about the maximum amount that will empty from your stomach) of a drink containing 6% glucose, you will take in about 48 grams of carbohydrate per hour. |
| Gramas | Se você bebe 200ml cada 15 minutos (essa é a quantidade máxima que será esvaziada do seu estômago) de uma bebida contendo 6% de glicose, terá em torno de 48 gramas de carboidrato por hora. |
| Grams of carbohydrates | Moderate-glycemic-index foods may be added for the next 18 to 20 hours, with a goal of consuming at least 600 <i>grams of carbohydrate</i> during the 24 hours after an intense workout or competition Fat. |
| Gramas de carboidratos | Alimentos com moderado índice glicêmico podem ser adicionados nas próximas 18 a 20 horas com o objetivo de consumir pelo menos 600 gramas de carboidratos durante às 24 horas seguintes ao treino intenso ou competição. |
| Grams of fiber | It is loaded with Vitamin A, high in folacin, potassium, calcium, and it has 5 <i>grams of fiber</i> (florete and stalk offers the highest fiber) per cup. |
| Gramas de fibra | Brócolis também contém Vitamina A, folacin, potássio, cálcio e 5 <i>gramas de fibras alimentares</i> por xícara. |
| Grams of protein | The average person needs approximately 0.8 <i>grams of protein</i> per kilogram bodyweight per day. |
| Gramas de | Uma pessoa comum precisa aproximadamente de 0,8 <i>gramas de proteína</i> por |

| | |
|---------------------------------|---|
| proteína | kg de peso corporal por dia. |
| Heart | Training at 3K to 5K race pace will work your cardiovascular system to its limit, which will help increase the stroke volume of your <i>heart</i> and improve your muscles' ability to use oxygen to produce energy aerobically. |
| Coração | Treinar no ritmo de corrida de 3km a 5km trabalhará seu sistema cardiovascular ao seu limite, o que ajudará a elevar o volume de bombeamento do seu <i>coração</i> e melhorar a habilidade dos seus músculos utilizarem oxigênio para produzir energia aerobicamente. |
| Heart attack | Choose a diet low in fat, saturated fat, and cholesterol to reduce your risk of <i>heart attack</i> and certain types of cancer and to help you maintain a healthy weight. Choose a diet moderate in sugars. |
| Ataque cardíaco | Escolha uma dieta baixa em calorias, com gordura saturada e colesterol para reduzir o risco de ataque cardíaco, alguns tipos de câncer e à manter um peso saudável. |
| Heart rate | <i>Heart rate</i> increases at high temperatures. |
| Frequência cardíaca | <i>A frequência cardíaca</i> eleva-se com altas temperaturas. |
| Heart rate monitor | Using a heart monitor is a good way to prevent yourself from training too hard on your easy days. |
| Monitor de frequência cardíaca | Utilizar um monitor de frequência cardíaca é uma boa forma de prevenir que treine muito forte nos dias leves. |
| Heart rate reserve | Keep your heart rate below 75% of your maximal heart rate (or 70% of your <i>heart rate reserve</i>) and let your body recover to allow high quality workouts on your hard training days. |
| Reservas de frequência cardíaca | Mantenha sua frequência cardíaca abaixo de 75%, da frequência máxima atingida (ou 70% da sua reserva cardíaca) e deixe o seu corpo recuperar-se, para permitir treinamentos de qualidade, em seus dias de treinamento mais pesados. |
| Heart rate zone | This means that if you set your <i>heart rate zones</i> based on your morning heart rates, and train in the afternoon, then you will train a bit less intensely than planned. |

| | |
|-----------------------------|--|
| Zona de frequência cardíaca | Isto significa que se você arrumar suas zonas de frequência cardíaca, baseadas em seus batimentos cardíacos matutinos, e no treino, no período da tarde, então, você irá treinar um pouco menos intensamente do que o planejado. |
| Muscles | A high-carbohydrate diet increases stores of glycogen, the energy for <i>muscles</i> , and improves overall athletic performance. |
| Músculos | Uma dieta rica em carboidratos aumenta o armazenamento de glicogênio, a energia de músculos e melhora a performance do atleta. |
| Muscle activity | Even <i>muscle activity</i> in areas distant from the spine and hip can positively impact bone density at these sites. |
| Atividade muscular | Até mesmo as atividades musculares em locais distantes da coluna vertebral e da cintura, podem causar impacto na densidade óssea desses locais, positivamente. |
| Muscles of glycogen | If you exercise for longer than an hour, you can begin to deplete your <i>muscles of glycogen</i> . |
| Músculos de glicogênio | Se você se exercitar por mais de uma hora, você poderá encher os seus músculo de glicogênio. |
| Muscle weakness | Increased <i>muscle weakness</i> can compound the problem of low bone density by increasing the risk of falls and fracture. |
| Pyramid | What is the <i>Food Pyramid</i> ? Flexible and full of options, the Pyramid gives simple and easy orientation that will help you choose what and how much to eat from each food group. |
| Pirâmide | O que é Pirâmide Alimentar? Flexível e com muitas variações, a Pirâmide oferece uma simples e fácil orientação que irá lhe ajudar em quê e no quanto comer de cada grupo alimentar. |

Um recurso de referência merecedor de destaque é o *corpus* denominado *corpus* de referência, o qual serve como parâmetro para o levantamento e analogias dos dados existentes em seu conteúdo. Alguns desses exemplos de *corpus* de referência da língua portuguesa são:

- Banco de português com 233 milhões de palavras. Privilegia o português brasileiro, escrito e falado. Construído pelos pesquisadores da PUC/SP.
- COMET (*Corpus* multilíngüe para ensino e tradução) com 5 milhões de palavras. Refere-se ao português escrito comparável com o inglês. Construído por pesquisadores da USP.
- *Corpus* UNESP/Araraquara/Usos do Português* com 200 milhões de palavras, privilegiando o português brasileiro escrito. Elaborado por pesquisadores da UNESP, Araraquara.
- CR-LW (*Corpus* de referência Lácio-Web)* com 5 milhões de palavras. Privilegia o português escrito. Composto por pesquisadores da USP, NILC.

Outros exemplos de *corpora* de estudos de língua inglesa são:

- *Brown Corpus (Brown University Standard Corpus of Present-day American English)*, de 1964. Contém 1 milhão de palavras em inglês americano escrito.
- *SEU Corpus (Survey of English Usage)*, de 1989. É composto por 1 milhão de palavras em inglês britânico, escrito e falado.
- *LLC (Lancaster-Oslo-Bergen)*, de 1978, contendo 1 milhão de palavras em inglês britânico escrito.

Para mais exemplos de *corpora* das línguas portuguesa e inglesa, consultar a obra *Linguística de corpus*, de Tony Berber Sardinha (vide capítulo das Referências).

Muitos podem ser os *corpora* escolhidos como referência desses idiomas, mas, nesta pesquisa, foram escolhidos o NILC, da língua portuguesa, e o *Brown*, da língua inglesa, que foram utilizados no levantamento dos verbetes, junto ao *corpus*, escolhidos como *corpus* de estudo, denominado *Copacabana runners*.

Uma das principais diferenças entre o *corpus* de referência e o de estudo é que o primeiro serve como padrão do idioma escolhido e parâmetro da língua de partida, quando comparada ao *corpus* de estudo, na língua de chegada.

Os *corpora* de estudo podem ser qualquer coleção de textos, orais ou escritos, selecionados pelo pesquisador, como meio de análise de um determinado assunto.