

**O novo universo de trabalho do tradutor:
Ferramentas e Fontes de Pesquisa Eletrônicas**

Felipe Miranda Portela

Felipe Miranda Portela

**O novo universo de trabalho do tradutor:
Ferramentas de Trabalho e Fontes de Pesquisa Eletrônicas**

Monografia submetida ao
Departamento de Letras
Estrangeiras Modernas da
Universidade Federal de Juiz de
Fora, como parte dos requisitos
para a obtenção do grau de
Bacharel em Letras: Ênfase em
Tradução – Inglês, elaborada sob
a orientação da Profa. Dra. Maria
Clara Castellões de Oliveira.

Juiz de Fora
Faculdade de Letras
Universidade Federal de Juiz de Fora
Janeiro de 2007

BANCA EXAMINADORA

Profa. Dra. Maria Clara Castellões de Oliveira (orientadora)

Prof. Ms. Adauto Villela

Profa. Dra. Maria Cristina Lobo Name

AGRADECIMENTOS

À Profa. Dra. Maria Clara Castellões de Oliveira,

Por dedicar-se ao curso de Bacharelado em Letras: Ênfase em Tradução – Inglês, ajudando a mantê-lo ativo e pela orientação paciente e precisa.

Ao Prof. Ms. Adauto Villela,

Pelo espírito empreendedor e renovador trazido ao Bacharelado e pelas inúmeras oportunidades de trabalho oferecidas a todos da turma.

À Profa. Dra. Miriam Lúcia Volpe,

Pela confiança, carinho e paciência que sempre teve comigo.

SUMÁRIO

INTRODUÇÃO.....	6
 CAPÍTULO 1 HISTÓRICO DA TRADUÇÃO AUTOMÁTICA.....	11
 CAPÍTULO 2 FONTES DE PESQUISA ELETRÔNICAS.....	18
2.1 Portais de busca.....	19
2.2 Páginas plurilíngües.....	23
2.3 Dicionários <i>on-line</i>	24
2.4 Corpora eletrônicos.....	25
2.5 Glossários.....	30
2.6 Correio eletrônico.....	31
 CAPÍTULO 3 FERRAMENTAS ELETRÔNICAS DE AUXÍLIO AO TRADUTOR.....	32
3.1 Tradução automática.....	33
3.2 Tradução Assistida por Computador.....	36
 CONSIDERAÇÕES FINAIS.....	47
 REFERÊNCIAS.....	50

A atividade tradutória vem sofrendo transformações profundas desde o início do século XX, a partir do surgimento do computador. Muitas vezes pensou-se que a profissão do tradutor estaria com os dias contados, uma vez que esse seria totalmente substituído pela máquina. Os anos se passaram e a máquina de traduzir não conseguiu desempenhar com perfeição o papel a ela destinado. Em pleno século XXI, a pessoa do tradutor continua imprescindível e quase não se questiona mais a substituição da tradução humana pela automática. O que vimos surgir nas últimas décadas foi uma gama de sistemas computacionais destinados a funcionar interativamente com os tradutores humanos.

A partir da década de 1990, com a criação da Internet, os computadores se fixaram como uma ferramenta absolutamente indispensável a qualquer profissional da tradução, oferecendo um universo de pesquisa gigantesco e sempre crescente ao qual o acesso é incrivelmente rápido e prático. Neste trabalho, pretendo fazer um apanhado geral das ferramentas e fontes de pesquisa informatizadas, hoje indispensáveis a realização de traduções com qualidade e num tempo hábil, condizentes com as exigências do mercado.

A opinião dos tradutores sobre todas essas mudanças varia enormemente. Nem sempre é fácil adaptar-se a estações de trabalho e modos de pesquisa em constante transformação. Em *Conversas com Tradutores* (2003), uma série de entrevistas com tradutores brasileiros renomados, compilada por Ivone Benedetti e Adail Sobral, vários entrevistados com experiência em diversas áreas da tradução foram convidados a dissertar sobre temas relevantes ao universo da tradução como atividade profissional no mundo contemporâneo, entre eles, a formação do tradutor, o mercado, a remuneração e o papel da crítica. Indagados sobre a questão da globalização, quase a totalidade dos entrevistados atestou sua grande influência não só na tradução, mas em todas as áreas do conhecimento. A tradutora de textos médicos Regina Alfarano (p. 33-43) ressalta a facilidade com que equipes médicas multinacionais trabalham juntas, a partir de partes distantes do planeta, e os benefícios

trazidos pelo intercâmbio de informações e pesquisas científicas. No campo da tradução em si, ela cita o emsimesmamento do tradutor, que tem em seu computador a troca de idéias necessárias ao ofício. Alfarano fala da descartabilidade da informação e sinaliza o risco da perda de identidade trazido pelos tradutores automáticos. O tradutor técnico Erik Borten, com experiência em várias multinacionais de grande porte, ressalta as vantagens mercadológicas trazidas pela globalização, que possibilita que um profissional atue em diversos mercados a custos reduzidos. Paulo Henriques Britto, tradutor de poemas, cita os novos campos de trabalho, como a localização de *software*, legendagem e dublagem, áreas que se apóiam completamente nas ferramentas de trabalho eletrônicas. Por sua vez, a tradutora juramentada Lúcia Helena França destaca a facilidade para atuar em diferentes mercados ao mesmo tempo devido à Internet e, como muitos de seus colegas, reclama do fator tempo, que muitas vezes acaba por comprometer a qualidade dos trabalhos. Ela atenta para o fato de que a maior rapidez de trabalho possibilitada pelos *softwares* de auxílio ao tradutor e pelas pesquisas online acaba por causar uma maior urgência do mercado, que exige dos profissionais prazos muitas vezes desumanos. De forma parecida, a tradutora nas áreas literária e médica Elaine Alves Trindade, que realiza também dublagem e legendagem, ressalta as desvantagens acarretadas pela globalização, destacando entre elas a queda na qualidade dos trabalhos e na remuneração dos profissionais. O problema, segundo Trindade, advém do fato de pessoas sem qualquer qualificação estarem atuando como tradutores, o que acaba por tirar o espaço de verdadeiros profissionais. Ela deixa claro que percebe essa situação como passageira, pois aqueles que vêem a tradução somente como uma forma de aumentar a renda não terão afinho o bastante para permanecer muito tempo nessa profissão que exige esforço e comprometimento constantes, mesmo com a utilização exaustiva de ferramentas e fontes de pesquisa eletrônicas.

A amplitude da visão exposta pelos diversos profissionais entrevistados traz uma nova luz a uma antiga questão: a invisibilidade do tradutor. Ivone Benedetti (p. 26-28) dá nova interpretação ao conceito, delimitando duas faces diferentes da invisibilidade: a textual e a social. A primeira diz respeito à fluência e à domesticação do texto, a segunda trata da condição anônima do tradutor. Benedetti ressalta o paradoxo presente no conceito da invisibilidade textual. Quando o texto traduzido é considerado tão fluente, tão bem escrito, que não pareça aos leitores uma tradução, então a intervenção do tradutor torna-se mais evidente, tornando-o mais visível. Desse modo, embora socialmente invisível, o tradutor torna-se textualmente mais reconhecível. As ferramentas eletrônicas de auxílio à tradução são muitas vezes acusadas de causar uma desvalorização a figura do tradutor, contribuindo para o fenômeno da invisibilidade. Porém, sabemos que o mercado e as condições de trabalho atuais exigem que os profissionais da tradução se familiarizem com esses recursos e saibam valer-se das vantagens trazidas por eles.

Este trabalho tratará dos principais recursos eletrônicos disponíveis hoje em dia para o auxílio à tradução. Pretendemos abordar essas ferramentas de forma bem prática, de modo a auxiliar tradutores, profissionais ou em formação, que desejam familiarizar-se com esse amplo universo de ferramentas e fontes de pesquisa, cada vez mais indispensáveis aos profissionais dessa área.

No primeiro capítulo, traçaremos um histórico geral desses recursos, desde seu surgimento, na década de 1930 até os *softwares* mais recentes. No segundo capítulo, abordaremos de forma mais detalhada as inúmeras fontes de pesquisa disponíveis na Internet e algumas formas práticas de utilizá-las. O terceiro e último capítulo será dividido em duas partes. Na primeira, trataremos dos *softwares* de tradução automática, comentando suas limitações e utilidades aos tradutores; nessa parte, incluímos uma análise comentada de uma tradução feita por um dos programas de tradução automática mais conhecidos atualmente.

Na segunda parte, abordaremos os *softwares* de auxílio ao tradutor, também conhecidos pela sigla CAT (*computer-aided translation*), dos quais o recurso mais proeminente é a memória de tradução, e analisaremos o modo de funcionamento de um dos representantes mais simples e funcionais da categoria.

Capítulo 1

Histórico da Tradução Automática

De acordo com informações extraídas do artigo “Machine Translation”, de Harold Somers (1998) e da monografia de conclusão do Bacharelado em Letras – Ênfase em Tradução da FALE-UFJF, de Luiz Carlos Moreira Rocha (1990), traçarei um breve histórico das ferramentas eletrônicas de auxílio à tradução, desde os primeiros tradutores automáticos até os mais recentes *softwares* que utilizam memórias de tradução.

As primeiras tentativas do homem de automatizar o processo tradutório datam do início do século XX. Os primeiros protótipos da máquina de traduzir surgiram na União Soviética em 1933 e, por volta de 1946, repercutiram na Inglaterra e nos Estados Unidos. Porém, só na década de 1950 houve um movimento maior por parte de várias nações para criar um conjunto de programas que possibilitasse a tradução automática (em inglês, *machine translation* – MT). Desde o início, um dos maiores interessados foi o governo estadunidense que, embalado pelo sonho de obter essas traduções com finalidades comerciais e estratégicas, concedeu bolsas a um grupo de cientistas para levar adiante esse projeto. Em 1951, o logicista e filósofo da linguagem Yehoshua Bar-Hillel tornou-se o primeiro pesquisador em tempo integral da tradução automática (MT), no MIT (*Massachusetts Institute of Technology*), um centro universitário de educação e pesquisa localizado em Cambridge, Massachusetts, nos EUA. A primeira demonstração de um sistema de tradução automática aconteceu em Nova Iorque, em janeiro de 1954, com um sistema de tradução russo-inglês desenvolvido conjuntamente por pesquisadores da IBM e da Universidade de Georgetown. Grupos de pesquisas foram aparecendo em outros países e, em 1956, aconteceu a primeira conferência internacional sobre tradução automática, na qual estavam presentes pesquisadores americanos, ingleses e canadenses, com contribuições da União Soviética.

Nos dez anos seguintes, as atividades nesse campo se espalharam pelo mundo, embora particularmente nos EUA, onde 20 milhões de dólares foram investidos em pesquisas relativas à MT. Esse alto nível de financiamento – em grande parte militar – deveu-se à

necessidade de produzir versões de um mesmo texto em diferentes línguas, principalmente no contexto da Organização das Nações Unidas no período da Guerra Fria. Investimentos em menor escala também foram destinados a programas de pesquisa na Inglaterra, França e Japão. Na União Soviética, assim como nos Estados Unidos, a pesquisa sobre MT foi largamente financiada, com vários grupos envolvidos.

Inúmeras técnicas foram propostas durante este período, mas a predominante foi a chamada “substituição direta baseada em dicionário” (SOMERS, 1998), na qual algumas poucas análises do texto fonte eram seguidas por buscas em dicionários, substituição pelos termos alvo e rearranjo da ordem das palavras. Alguns grupos utilizavam técnicas de maior sofisticação lingüística, porém, a abordagem formal da lingüística proposta por Noam Chomsky no fim da década de 1950 não afetou a pesquisa sobre a tradução automática até bem mais tarde. Outras abordagens incluíam análise distribucional estatística, enquanto alguns grupos se concentravam nas questões teóricas envolvidas na criação do que viria a ser chamado de “sistemas de segunda geração” (SOMERS, 1998).

O ano de 1964 foi um ponto crucial na história da MT, com a formação do Comitê Consultivo sobre Processamento Lingüístico Automático (ALPAC, do inglês *Automatic Language Processing Advisory Committee*) para relatar os progressos da pesquisa em MT. O documento resultante, o mal afamado relatório ALPAC (ALPAC, 1966), concluiu que a tradução automática era mais lenta, menos precisa e duas vezes mais cara que a tradução humana, e que não havia nenhum prospecto imediato ou previsto da real utilidade das MT.

Embora amplamente criticado e acusado de usar informações desatualizadas, fazer cálculos errados nas estimativas de custos e ignorar resultados favoráveis, o relatório teve um efeito devastador sobre os financiamentos nos EUA, e má repercussão em todo o mundo. Na realidade, o relatório ALPAC também contém informações positivas nem sempre citadas,

como a recomendação de recursos eletrônicos para a tradução e financiamentos para mais pesquisas em Lingüística Computacional.

Após a publicação do relatório ALPAC, grupos de pesquisadores da tradução automática passaram a revisar as técnicas básicas que vinham utilizando, e terminaram por desenvolver o chamado ‘método indireto’: a abordagem transferencial, que envolve análise estrutural do texto de entrada, um mapeamento bilíngüe em nível abstrato e síntese do texto em língua-alvo; ou, ainda, a abordagem interlingual, na qual o estágio de transferência bilíngüe é evitado pelo uso de uma representação universal mais abstrata (SOMERS, 1998, p. 141). Aproximadamente na mesma época, Bar-Hillel escrevia que uma tradução totalmente automática de alta qualidade só seria possível se computadores pudessem ter acesso ao que hoje chamamos de ‘conhecimento real de mundo’ (BAR-HILLEL citado por SOMERS, 1998, p. 141). Por outro lado, podemos afirmar, ainda com base no artigo “Machine Translation”, de Somers (1998), que alguns dos sistemas de tradução automática que sobreviveram ao ALPAC foram designados com uma arquitetura essencialmente de primeira geração.

As pesquisas em MT mais significativas dos dez anos posteriores ao ALPAC concentraram-se na abordagem indireta e em outras características da concepção de segunda geração, como o uso de técnicas lingüísticas e computacionais mais sofisticadas. Esse foi um período de reflexão durante o qual o único trabalho de pesquisa em MT nos EUA foi financiado por iniciativa privada, e o investimento geral nessa área foi drasticamente reduzido. De acordo com Somers (1998), a figura mais importante desse período foi Bernard Vauquois e seu grupo GETA em Grenoble, cuja influência se fez notar em grupos de estudo em Montreal e Kyoto, trazendo alguns resultados promissores.

Os resultados mais notáveis desse período foram o sistema MÉTÉO, desenvolvido pelo grupo TAUM em Montreal, que substituíu os tradutores humanos na tarefa excepcionalmente tediosa de traduzir boletins meteorológicos do inglês para o francês, e o SYSTRAN, um

sistema híbrido de primeira e segunda geração que foi desenvolvido por Peter Toma, na Califórnia, e posteriormente usado pela Força Aérea Norte-Americana e pela NASA, que precisava de traduções ‘brutas’ do russo para o inglês. Havia também uma versão inglês-francês do SYSTRAN que a Comissão de Comunidades Européias (CEC) desenvolveu em Luxemburgo. Essas experiências bem-sucedidas encorajaram a CEC em 1978 a iniciar as discussões que levariam ao financiamento do Eurotra, que permanece ainda hoje como o maior projeto de tradução automática já empreendido.

O renascimento da tradução automática estava quase completo no fim da década de 1970, com vários projetos em desenvolvimento e trabalhos acadêmicos sendo publicados em toda a Europa. No Japão, pesquisadores obrigados a desenvolver sistemas de computador capazes de lidar com a escrita japonesa, começaram a voltar sua atenção à tradução automática.

Nessa época, os primeiros programas de tradução automática estavam começando a aparecer no mercado, a maioria dos quais era baseada na arquitetura de primeira geração. Contudo, nenhum desses sistemas era direcionado ao público monolíngüe; eles tinham que ser usados por tradutores, que se sentiram ameaçados pela nova tecnologia que, em qualquer caso, não facilitou muito o trabalho deles.

Nos dez anos seguintes, o campo da MT começou a amadurecer, principalmente quando as limitações da tecnologia de segunda geração foram reconhecidas e pesquisadores começaram a explorar modos de superar essas limitações. Avanços na informática também tiveram grande efeito. Em particular, o surgimento de programas interativos e computadores pessoais pequenos possibilitou que homem e máquina trabalhassem em conjunto para criar traduções, resultando no desenvolvimento de novos tipos de programas para auxiliar o trabalho do tradutor.

As propostas para as primeiras ferramentas eletrônicas que constituiriam o que hoje são os *softwares* de tradução assistida por computador remontam ao fim da década de 1970. A chamada tradução automática interativa (em inglês, *interactive machine translation – IMT*) teve sua primeira aparição em 1973, no sistema *Mind* de Kay (KAY citado por LANGLAIS, FOSTER & LAPALME, 2000, p. 276), no qual o papel do usuário era auxiliar o computador na análise do texto fonte, respondendo perguntas sobre o sentido das palavras, elipses, concordância, etc. A maior parte dos trabalhos posteriores sobre IMT seguiu essa linha, concentrando-se em melhorar o processo de pergunta-resposta com a diminuição e adaptação das perguntas. Apesar do progresso nesse sentido, os sistemas desse tipo são geralmente inapropriados como ferramentas para tradutores habilitados, já que o usuário serve apenas como conselheiro, e é o componente de tradução automática que tem o controle sobre o ato tradutório em si.

Segundo Hutchins (1998), na década de 1980, vários protótipos de programas de computador para uso de tradutores continuaram a se desenvolver. Esse desenvolvimento culminou com o surgimento, na década de 1990, de uma gama de programas conhecidos como "*translators' workstations*" (numa tradução literal, estações de trabalho dos tradutores). O componente mais importante desses *softwares* era a memória de tradução (em inglês, *translation memory – TM*), que muitos pensaram que substituiria a tradicional tradução automática (HUTCHINS, 1998, p. 287). Ao contrário do previsto, a última década viu surgir de forma aparentemente repentina um novo paradigma para a tradução automática. Duas novas técnicas tendo em comum uma abordagem 'empírica', bem diferentes da metodologia predominante, começaram a atrair pesquisadores. A abordagem empírica envolve o uso dos corpora (coleção de textos reunidos, de áreas variadas ou não, com um propósito específico de análise) e da estatística, em vez de regras lingüísticas e algoritmos. Uma abordagem inclui a extração de equivalentes de tradução lexicais e sintáticos de extensos corpora paralelos

(bilhões de palavras) com base na probabilidade estatística. A outra trabalha com corpora menores constituídos de exemplos-chave usados como modelos de tradução. A essa altura, já existem também sistemas híbridos que incorporam ambas as abordagens (SOMERS, 1998, p. 142).

O fato é que a tradução automática continua se desenvolvendo rapidamente, sendo aplicada hoje em dia a inúmeras atividades, como grupos de discussão na Internet, tradução de sítios, acesso a bancos de dados em línguas estrangeiras. Na Internet, encontram-se inúmeras ofertas de *softwares*, e muitos dos programas disponíveis funcionam de forma bastante interativa com o usuário, permitindo a modificação dos bancos de dados e dicionários utilizados pelos programas. A União Européia vem pesquisando e utilizando amplamente as traduções automáticas como forma de agilizar as trocas de informações entre os países membros. Porém, os textos são previamente preparados para serem "traduzíveis" pelo computador, evitando inversões sintáticas, ambigüidades, etc., uma vez que, ainda hoje, o computador se mostra inapto a traduzir automaticamente textos "normais" ou não preparados especialmente para esse fim.

Paralelamente, a tradução assistida por computador continua a se desenvolver, e todos os anos o mercado vê surgirem novas ofertas de *softwares* que prometem auxiliar cada vez mais os profissionais da tradução, inclusive programas que combinam ferramentas de auxílio ao tradutor com tradução automática.

Capítulo 2

Fontes de Pesquisa Eletrônicas

Neste segundo capítulo, faremos um apanhado das principais fontes de pesquisa que podem ser úteis aos profissionais da tradução, oferecendo uma breve descrição sobre como utilizar cada uma delas da forma mais proveitosa:

Trataremos dos seguintes recursos:

- portais de busca, dos quais o representante mais popular e abrangente é o *Google*;
- páginas que disponibilizam seu conteúdo em vários idiomas, constituindo "lexinomas espontâneos";
- versões *on-line* de dicionários renomados (tanto monolíngües quanto bilíngües);
- corpora monolíngües e bilíngües desenvolvidos principalmente por universidades e outras instituições de ensino, abrangendo cada vez mais temas e tipos de texto;
- glossários de todos os tipos e dimensões, que normalmente consistem em listas comparativas de palavras isoladas sobre um tema específico;
- correio eletrônico, que pode facilitar e tornar mais rápidas consultas diretas com profissionais e expertos em áreas específicas do conhecimento humano.

Passemos então para uma análise mais detalhada de cada uma destas amplas fontes de informação disponibilizadas na Internet e sobre as formas mais viáveis de utilização delas pelos profissionais da tradução.

2. 1 Portais de busca

Os inúmeros portais de busca disponíveis na Internet são ferramentas úteis àqueles que trabalham com tradução de textos técnicos e científicos. O grande volume de informação indexada por esses portais de pesquisa pode auxiliar enormemente a fase inicial do trabalho do tradutor: a busca de termos e conceitos específicos de cada área do saber. O tradutor pode, através desses portais, buscar textos já produzidos na língua em questão sobre o tema desejado. Por exemplo, se um profissional precisa verter para o inglês um texto escrito no

Brasil sobre espécies de jacaré existentes no Pantanal, pode ser de grande ajuda obter textos em inglês sobre animais da mesma família existentes em outras partes do mundo. Para isso, basta que ele procure por palavras chave referentes ao tema, como *alligator*, *zoology*, etc. Depois, escolhe-se dentre os resultados aqueles que podem conter informações lingüísticas úteis, como tipo de vocabulário usado, nomenclatura usada na descrição da anatomia, fisiologia e comportamento do animal em questão.

No caso específico de textos acadêmicos e científicos, há uma ferramenta preciosa para pesquisa e busca de material de referência: o Portal de Periódicos da CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), que reúne artigos acadêmicos publicados por centenas de universidades em todo o mundo, trazendo, portanto, artigos em dezenas de línguas, todos contendo um pequeno resumo (*abstract*, *resumé*) em uma das línguas principais (inglês, francês, espanhol). Uma limitação do uso dessa fonte de pesquisa é que o portal é acessível somente a partir de instituições de ensino, sendo vetado o acesso a computadores domésticos.

Mas a utilidade dos portais de busca vai muito além. Portais como o *Google* podem ser uma excelente fonte de pesquisa lingüística (ortográfica, sintática, semântica). Um tradutor pode usá-los das mais diferentes formas, como sugerido abaixo:

- Pesquisar a ortografia e definição de palavras:

Se um profissional tem dúvidas sobre a grafia de um termo qualquer na língua inglesa, por exemplo, basta digitar a palavra entre aspas no campo de pesquisa e observar a quantidade de resultados obtidos. Para uma precisão maior, ele deve utilizar as "ferramentas de idioma" disponíveis no *Google*. Poderá então escolher visualizar somente os resultados de língua inglesa, ou ainda, somente os resultados de língua inglesa de páginas australianas, e assim por diante. Se o tradutor tem dúvida sobre a grafia da palavra equivalente a "contração", por exemplo, ele deve digitar no campo de pesquisa a grafia que ele julga correta e verificar os

resultados. Se digitar: *contration* surgirão 15.500 sítios em inglês (de qualquer país) que mencionam a palavra *contration* com esta grafia errada. Parece muito, mas não se comparado ao número de páginas em inglês que trazem a grafia correta: 17.200.000. Neste caso, o próprio *Google* apresentou a seguinte sugestão: "Você quis dizer: **contraction**" com um link para a busca da palavra grafada corretamente, cujos resultados incluíam informações semânticas sobre o termo. Assim, comparando-se o número de ocorrências, e verificando os resultados, pode-se chegar a uma conclusão confiável sobre a grafia correta. Esse tipo de busca pode ser muito útil no caso de jargões e termos técnicos.

Vale mencionar que o *Google* oferece também um serviço de busca por definições de palavras, disponibilizando informações de vários dicionários. Para buscar definições de palavras, basta digitar no campo de pesquisa "define: " seguido da palavra a ser pesquisada. São oferecidas definições gerais e específicas advindas de dicionários especializados nos mais variados assuntos. Esse serviço funciona perfeitamente para a língua inglesa, mas é ainda insuficiente quando se trata de línguas menos difundidas, como o português. A título de exemplo, a busca por "define: contraction" trouxe cerca de 20 definições para o termo da língua inglesa, incluindo exemplos de expressões relacionadas ao termo: eccentric contraction concentric contraction isometric contraction length contraction muscle contraction contraction coefficient uterine contraction isotonic contraction thermal contraction contraction joint. A busca da palavra equivalente em português, "define: contração" não encontrou resultados, embora o serviço seja oferecido também para o português.

- Buscar jargões e expressões cristalizadas:

Um outro recurso útil disponibilizado pelos portais de busca é a pesquisa de concordância e expressões cristalizadas. Representam uma forma prática de confirmar o uso de expressões e estruturas frasais típicas de cada língua. Podem ser usados sempre que há

dúvida sobre o uso de uma expressão, a ordem das palavras em frases. Por exemplo, ao traduzir para o inglês um texto científico, um tradutor depara-se com a seguinte expressão: "condições normais de temperatura e pressão". Supondo-se que o ele não conheça a expressão correspondente na língua inglesa, ele tem duas alternativas: traduzir 'ao pé da letra', ou buscar na Internet a expressão correta utilizada em inglês. Para isso, basta digitar no campo de pesquisa do Google as palavras: *temperature pressure* (sem aspas ou qualquer caractere entre elas). Na página de resultados, o quarto item traz no título: "*Standard conditions for temperature and pressure*", o equivalente exato em língua inglesa para a expressão em questão. Se mesmo assim o tradutor tiver dúvida sobre a pertinência desta expressão, basta abrir a página e conferir a explicação para ver que ela é a expressão correta a ser usada.

Há ainda alguns caracteres especiais que podem auxiliar a busca nesses portais. As aspas são usadas sempre que se deseja buscar por uma expressão específica, por exemplo, "*translation memory*". Se as palavras são digitadas no campo de pesquisa, sem aspas, o portal traz como resultado todas as páginas que contenham uma ou as duas palavras em qualquer ordem. Se a pesquisa é feita com os termos entre aspas, os resultados trazem somente as páginas que contenham essas duas palavras nessa ordem. O asterisco serve como caractere coringa, podendo ser usado na busca de expressões (sempre entre aspas) nas quais um termo é incerto ou desconhecido. Por exemplo, a busca por "*computer * translation*" traz, entre os primeiros resultados, páginas contendo expressões como: *computer aided translation*, *computer-assisted translation*, *computer-based translation*, *computer-controlled translation*.

Os portais de busca são também a melhor maneira de encontrar referências a personagens históricos, lugares, obras de arte, marcas registradas, nomes populares de animais e quaisquer outras expressões que requerem traduções precisas. Se bem utilizados, esses portais podem conduzir o profissional aos locais onde ele poderá encontrar as informações lingüísticas necessárias para o seu trabalho.

2. 2 Páginas plurilíngües

Existem na Internet incontáveis páginas, sobre os mais variados assuntos, que apresentam seu conteúdo em mais de um idioma. Mesmo aquelas que trazem o conteúdo em somente um idioma costumam trazer *links* para páginas estrangeiras que tratam do mesmo assunto, o que pode constituir uma fonte útil de pesquisa lingüística.

Bergenholtz/Kaufman (1997) cunharam o termo *lexinoma*, que designa uma obra lexicográfica multifuncional e plurilíngüe que combina um avançado dicionário enciclopédico e eletrônico com uma série de outras possibilidades que a Internet oferece. Segundo Tarp (2002), um *lexinoma* inclui as funções lexicográficas de recepção, produção e tradução de textos sobre uma determinada disciplina, tendo uma apresentação nivelada por grau de especialização e contendo textos em vários idiomas, para servir tanto a leigos quanto a expertos e semi-expertos de diferentes línguas maternas. Um *lexinoma* contém, além disso, *links* a outras páginas relevantes ao tema e utiliza, como meio de acesso rápido à informação, o serviço de busca tanto dentro do próprio sítio quanto na Internet como um todo. Trata-se, portanto, de uma completa obra lexicográfica que pode servir, entre muitos outros usuários, aos tradutores de textos especializados.

Nos últimos tempos, as páginas na Internet de certas empresas, organizações e instituições vêm se aproximando de forma espontânea do conceito de um *lexinoma*. Essas páginas não se baseiam em critérios lexicográficos e, por conseguinte, não procuram de forma consciente cumprir com as funções de recepção, produção e tradução de textos. Têm um enfoque próprio e bem diferente, pois buscam, principalmente, divulgar sua organização, empresa, produto, campo de investigação, etc. Porém, com os dados que incluem, aos poucos se aproximam do conceito de um *lexinoma*, daí surge, então, o conceito de "*lexinoma espontâneo*".

Um dos melhores exemplos desse tipo de fonte de informação é o sítio elaborado para a Associação Dinamarquesa da Indústria Eólica: <http://www.windpower.org>. Ao entrar no sítio, o visitante tem imediatamente a opção de continuar em um dos seguintes idiomas: dinamarquês, inglês, alemão, francês e espanhol. A partir daí, todas as demais páginas são idênticas em todos os idiomas, a menos que se tenha introduzido alguma novidade em dinamarquês e inglês que ainda não tenha sido traduzida para as demais línguas. Todas as traduções foram realizadas por engenheiros bilíngües. O sítio apresenta ainda um glossário com cerca de 150 termos e seus equivalentes nas outras línguas e um serviço de busca interna. Uma das desvantagens desse tipo de "dicionário" é a dificuldade em conseguir, a partir dele, informações explícitas sobre as propriedades sintáticas de um termo, mas utilizando-se o serviço de busca com paciência, às vezes se consegue algum ou outro dado sobre essas propriedades. Mesmo com algumas deficiências, esse tipo de obra lexicográfica constitui-se num instrumento adequado à tradução profissional. (TARP, 2002, p. 257-274)

2. 3 Dicionários *on-line*

Todos os principais dicionários possuem edições eletrônicas, sejam elas disponibilizadas *on-line* ou vendidas em forma de CD-ROM. A maioria deles disponibiliza gratuitamente versões compactas de seus conteúdos a todos os usuários da rede, a exemplo do *Compact Oxford English Dictionary* (http://www.askoxford.com/dictionaries/compact_oed/) e do *Cambridge Advanced Learner's Dictionary* (<http://dictionary.cambridge.org/>), em língua inglesa. Merece destaque o portal Bússola Escolar (<http://www.bussolaescolar.com.br/dicionarios.htm>), que oferece uma ampla série de dicionários *on-line* de uso irrestrito e ainda glossários específicos das mais variadas áreas do conhecimento, constituindo-se numa fonte útil de pesquisa.

A versão eletrônica do dicionário de língua inglesa *Collins COBUILD* (disponível para compra em www.collins.co.uk) também merece destaque por fazer uso de um corpus de textos em inglês de 4 milhões de palavras (*The Bank of English*) como banco de referências. Desse modo, além de pesquisar definições de termos isolados, pode-se também ter acesso aos trechos de textos contendo a(s) palavra(s) buscada(s), o que permite ao tradutor inferir sobre o uso das palavras, sua forma de organização nas sentenças, etc.

Entre os dicionários eletrônicos comercializados, o *Babylon* (www.babylon.com) – um indexador que engloba vários dicionários – pode ser uma ferramenta prática, pois funciona acoplado a uma série de dicionários e enciclopédias disponíveis para *download* (pago) como, por exemplo, *Michaelis* (em várias línguas), *Oxford*, *Britannica*, e é ativado por uma tecla de atalho escolhida pelo usuário. Pode-se, por exemplo, configurar “Ctrl+botão direito do mouse” como tecla de atalho, então posiciona-se a seta do mouse sobre a palavra a ser consultada, utiliza-se a tecla configurada e o *Babylon* abre uma pequena janela trazendo todas as acepções da palavra contidas em seus dicionários. Se o computador estiver conectado à Internet, ele busca também definições extras disponíveis em seus bancos de dados *on-line*. O *Babylon* inclui ainda um conversor de medidas e moedas. Encontra-se disponível para compra em diversos pacotes, incluindo diferentes dicionários e idiomas. Pode-se também baixar uma versão de teste, que funciona perfeitamente, mas por tempo limitado.

2. 4 Corpora

A Linguística de Corpus propõe um estudo da língua baseado na observação empírica de grandes coletâneas de textos, os corpora, criteriosamente compiladas por meio do uso de ferramentas computacionais. Apresentamos aqui os conceitos compilados por Sardinha (2002) acerca desse importante universo de pesquisa linguística. O posicionamento corrente entre os pesquisadores da tradução e os linguistas de corpus é o de que tanto os estudos tradutológicos

como área acadêmica de pesquisa, quanto à prática tradutória, têm muito a ganhar com um contato maior com a Lingüística de Corpus.

Sardinha cita a posição de Hunston (2002) para ilustrar o pensamento dos lingüistas de corpus:

Os Corpora [...] têm mais a oferecer aos tradutores do que pode parecer à primeira vista. Eles não apenas fornecem evidências de como as palavras são usadas e quais as traduções possíveis para uma determinada palavra ou locução, mas também permitem uma percepção aprofundada do processo e natureza da tradução em si (HUNSTON citado por SARDINHA, 2002, p. 15).

Entre os tradutores, o reconhecimento do valor de corpora eletrônicos não é muito diferente. Para Tymoczko (1998), a influência do uso de corpus na tradução tem vários efeitos positivos, entre eles o de permitir um maior intercâmbio de dados entre pesquisadores e praticantes, além de modernizar a área, trazendo-a mais perto do que se espera da pesquisa contemporânea: "Os estudos da tradução com base em corpus mudam tanto qualitativa quanto quantitativamente o conteúdo e os métodos da disciplina de Estudos da Tradução, de forma adequada à era da informação" (TYMOCZKO citada por SARDINHA, 2002, p. 16).

Boa parte do interesse em corpora eletrônicos na área de tradução provém de necessidades reais de tradução automática de grande volume de dados. É o caso da União Européia, que, como parte de sua existência, tem de lidar com documentos que precisam ser traduzidos para as várias línguas da comunidade. Diversos projetos importantes de tradução assistida por corpus têm sido financiados pela União Européia (HUNSTON citado por SARDINHA, 2002, p. 17).

Porém, mesmo em centros onde a Lingüística de Corpus está altamente desenvolvida, como na Grã-Bretanha, a interface com a tradução ainda é restrita. Um exemplo bem sucedido dessa interface é o TEC (*Translational English Corpus*), um corpus de inglês contemporâneo constituído de textos traduzidos para o inglês a partir de uma variedade de línguas fonte. O

TEC é mantido pelo Centro de Estudos Interculturais e da Tradução da Universidade de Manchester, na Inglaterra e pode ser acessado gratuitamente na rede (www.umist.ac.uk/ctis/research/TEC/tec_home_page.htm). As pesquisas que deram origem ao TEC partem do princípio de que a tradução é um tipo distinto de evento comunicativo, caracterizada por seus próprios objetivos, pressões e contextos de produção. Sendo assim, as traduções não são normalmente recebidas da mesma forma que textos originais na maioria das sociedades contemporâneas, embora essa recepção seja diferente para cada contexto social e histórico. Fenômenos desse tipo podem então deixar traços na linguagem produzida por tradutores, cuja análise pode lançar luz sobre os processos de formação do texto de saída. Essa questão não interessa somente aos estudiosos da tradução, mas a todos os lingüistas, uma vez que grande parte da linguagem que todos consumimos regularmente hoje em dia é traduzida. O TEC serve de base a uma ampla gama de estudos em duas áreas principais: o modo como o padrão do texto traduzido difere do texto não traduzido no mesmo idioma e a variação estilística entre tradutores individuais. Exemplos de ambos os tipos de pesquisa podem ser encontrados na página da Universidade de Manchester.

No Brasil, ainda é recente a tentativa de união dessas duas áreas. Entre os corpora disponíveis para consulta *on-line*, destacamos o Projeto COMET (Corpus Multilíngüe para Ensino e Tradução) da Faculdade de Filosofia, Letras e Ciências Humanas da USP. O COMET se subdivide em três corpora (<http://www.fflch.usp.br/dlm/comet/>, acessado em 25/10/2006):

- CorTec (Corpus Técnico-Científico): é um corpus comparável de textos técnicos e/ou científicos originalmente escritos em português brasileiro e em inglês. Esse corpus privilegiará quatro áreas: Direito Comercial, Informática, Ortodontia e Meio Ambiente, que serão ampliadas sistematicamente até cada uma atingir a meta de um milhão de palavras. Abrigará, eventualmente, corpora produzidos pelos alunos do Curso de Especialização em

Tradução e por outros pós-graduandos do programa de Estudos Lingüísticos e Literários em Inglês.

- CorTrad (Corpus de Tradução): subdivide-se em Literário e Juramentado. O Corpus Literário é composto de contos traduzidos do inglês e seus respectivos originais. O Corpus Juramentado será constituído de textos cedidos pela Junta Comercial de São Paulo por meio de contrato de comodato com a USP.

- CoMAprend (Corpus Multilíngüe de Aprendizes): será constituído de redações dos alunos da graduação e dos cursos de extensão das áreas do Departamento de Letras Modernas: alemão, espanhol, francês, inglês e italiano.

Segue uma seleção de corpora do português e do inglês divulgada pelo sítio do Projeto COMET (<http://www.fflch.usp.br/dlm/comet/>, acessado em 25/10/2006):

- Lácio Web: sítio brasileiro com um corpus contemporâneo de língua geral, subdividido em subcorpora que representam vários gêneros e tipos textuais - textos disponíveis para *download* e para serem usados com as próprias ferramentas do sítio.

- Banco do Português: sítio brasileiro com um corpus de português contemporâneo, atualizado constantemente (em 2004, possuía 223 milhões de palavras). Apenas uma amostra está disponível para consulta e uso com as próprias ferramentas do sítio.

- COMPARA: site português com um corpus paralelo – originais e respectivas traduções – de/para as diversas variantes do português e do inglês.

- CetenFOLHA (Corpus de Extratos de Textos Eletrônicos NILC/Folha de São Paulo): corpus de cerca de 24 milhões de palavras em português brasileiro retirados do jornal "A Folha de São Paulo" - os textos podem ser baixados via FTP / HTTP ou consultados no link do projeto AC/DC.

- CetemPUBLICO (Corpus de Extratos de Textos Eletrônicos MCT/Público): corpus de aproximadamente 180 milhões de palavras em português europeu retirados do jornal

português "Público" - disponibiliza as versões anotada e sem anotação para consulta *on-line* e envia CD com o corpus mediante solicitação.

- TychoBrahe: sítio brasileiro com um corpus de português histórico.
- PHPB (Para Uma História do Português do Brasil - RJ): sítio brasileiro com corpus diacrônico: transcrições de impressos do século XIX e edições diplomático-interpretativas de manuscritos dos séculos XVIII e XIX. Não oferece nenhuma ferramenta de análise e os textos devem ser consultados *on-line*, um a um.
- BNC (British National Corpus): um extenso corpus de inglês moderno.
- ICAME (International Computer Archive of Modern and Medieval English): Organização Internacional de Linguistas e Cientistas da Informação trabalhando com textos legíveis por máquinas em língua inglesa. O objetivo dessa organização é coletar e distribuir informações sobre materiais em inglês disponíveis para processamento eletrônico e sobre pesquisas lingüísticas com esse material, para compilar corpora de textos em inglês em formas legíveis por máquinas, e tornar esse material disponível para instituições de pesquisa.
- Corpus de Inglês Americano Profissional, Falado: construído a partir de uma seleção de transcrições de interações em ambientes profissionais, contém dois subcorpora, cada um com um milhão de palavras: um subcorpus consiste sobretudo de discussões acadêmicas, o segundo contém transcrições de coletivas à imprensa da Casa Branca. O corpus está disponível comercialmente e há uma amostra com 50.000 palavras disponível *on-line*.

Alguns dos corpora listados acima permitem acesso restrito gratuito a seu conteúdo, porém, a maioria exige cadastramento e pagamento. Contudo, os corpora não se destinam a pesquisa de termos ou expressões para realização de traduções isoladas, se não a análises mais profundas sobre a natureza da tradução, a relação entre as línguas etc. Portanto, servem mais como fontes de pesquisa aos Estudos da Tradução que à prática tradutória isolada. Assim sendo, a partir de Magalhães (2001), podemos afirmar que a contribuição da lingüística de

corpus à tradução acontece a médio e longo prazo, servindo de base mais aos estudiosos da tradução como disciplina teórica do que aos que trabalham com a tradução prática.

2. 5 Glossários

Glossário, como trataremos aqui, é uma obra terminológica bilíngüe que agrupa os termos pertinentes e representativos de uma determinada área, normalmente apresentando verbetes na língua fonte em ordem alfabética, seguidos dos termos correspondentes na língua alvo. Apesar de trazerem somente palavras isoladas, os glossários podem ser úteis na tradução de textos técnicos e científicos, por sua especificidade e precisão. São particularmente práticos quando se utiliza um dos *softwares* de auxílio ao tradutor dos quais trataremos adiante, pois a maior parte deles permite o carregamento de glossários durante a tradução, e a inserção, mais ou menos automática, dos termos contidos no glossário. Nesse caso, precisa-se que os glossários estejam formatados de maneira tal a permitir seu funcionamento num desses programas. Alguns deles permitem até a criação de um glossário (através de teclas de atalho) durante a tradução de determinado texto, podendo ser esse glossário utilizado posteriormente em outros projetos.

Os glossários disponíveis para consulta *on-line* podem ser de ajuda quando trazem termos técnicos e jargões de difícil entendimento, mas sua interação direta com programas de auxílio ao tradutor só é possível se o usuário formatá-los para tal. A UNIBERO (Centro Universitário Ibero-Americano) disponibiliza em seu sítio (http://www.unibero.edu.br/glossarios_def.asp) uma série de glossários para *download* gratuito. O já citado portal Bússola Escolar oferece, na página <http://www.bussolaescolar.com.br/dicionarios.htm>, glossários específicos de várias áreas.

Apesar de não trazerem informações sintáticas e apresentarem somente palavras isoladas (não dando uma noção exata de seu uso e contexto), os glossários têm seu lugar como fonte

de pesquisa e auxílio aos tradutores, uma vez que freqüentemente são a única fonte de informação sobre determinados jargões e termos técnicos e científicos.

2. 6 Correio eletrônico

Os *e-mails* representam um meio rápido e eficiente de contatar expertos e conhecedores de diferentes áreas de conhecimento. Não raro, um tradutor sente a necessidade de confirmar com um experto as informações que obteve, seja para averiguar a confiabilidade de suas fontes de pesquisa ou para eliminar a dúvida entre duas ou mais possibilidades de tradução. Hoje em dia, a forma mais prática de fazê-lo é certamente o correio eletrônico.

Capítulo 3

Ferramentas Eletrônicas de Auxílio ao Tradutor

A diferença entre os mais recentes *softwares* utilizados por tradutores, na chamada tradução assistida por computador, e os conhecidos tradutores automáticos é que esses últimos funcionam de forma pré-programada, não-controlada pelo usuário, traduzindo textos inteiros, sem levar em conta contexto, situação, público-alvo do texto, etc., muitas vezes produzindo resultados equivocados. Os programas de auxílio ao tradutor, por outro lado, funcionam em completa interação com o profissional, e na verdade é ele quem traduz, utilizando os recursos do programa. A utilização de ambos requer que o texto fonte esteja em formato digital. Se o cliente só puder fornecer o material impresso (o que muitas vezes acontece), pode-se utilizar um *scanner* para fazer uma cópia digital do material e depois submetê-lo a um *software* de reconhecimento de caracteres (normalmente fornecido com o *scanner*) para transformá-lo num arquivo de texto.

É necessário ressaltar que os programas em questão são de grande utilidade na realização de traduções técnicas, ou seja, textos jornalísticos, manuais, textos científicos, menus, etc., e dificilmente seriam de alguma ajuda na tradução de textos literários, como poemas, contos ou romances. Ao traduzir textos desse último tipo, o profissional deve ter muita atenção ao estilo e à linguagem utilizada, tomando um cuidado enorme com cada detalhe do texto, sutilezas que os programas, por mais funcionais que sejam, não têm como oferecer.

Passaremos então a uma análise mais detalhada de alguns dos principais programas de tradução automática e tradução assistida por computador.

3. 1 Tradução automática

Tradução automática (*machine translation - MT*) é aquela desempenhada completamente pela máquina, sem nenhuma intervenção do tradutor. Existem inúmeros *softwares* de tradução automática e, apesar de possuírem modos de funcionamento diferentes, todos apresentam resultados bem parecidos. Possuem glossários de termos e bancos de regras sintáticas de

ambas as línguas em sua memória e, a partir desses dados, vertem os textos de uma língua a outra. Porém, não dão conta de polissemia, ambigüidade, estruturas sintáticas truncadas, entre outros fatores presentes em todos os tipos de texto. É desnecessário dizer, portanto, que esses *softwares* produzem textos estranhos e muitas vezes ininteligíveis, sendo sempre indispensável uma revisão e correção por parte do tradutor humano. A utilidade desses programas ao profissional da tradução é a economia de tempo e digitação que eles permitem, pois o profissional pode começar o trabalho com o corpo do texto já digitado (mesmo que cheio de erros), fazendo somente as modificações necessárias. Portanto, até os tradutores automáticos, tão criticados a princípio, podem representar uma economia de trabalho ao profissional. Uma vez realizada a pré-tradução pelo programa, basta que o tradutor repasse o texto corrigindo os erros, que certamente não serão poucos.

Há quem diga que o uso desse tipo de programa não representa uma economia de fato, pois a maioria traduz tão mal que é preciso modificar tudo e, além disso, o programa não "aprende" a partir das correções realizadas, e continua sempre cometendo os mesmos erros. Fica, então, a cargo de cada profissional saber quais recursos podem ser realmente úteis em cada projeto e quais representariam mais uma perda de tempo do que uma facilitação do trabalho.

Apresentamos a seguir um teste rápido com um dos mais conhecidos programas de tradução automática, o *Globalink Power Translator Pro*, para avaliar a qualidade da tradução inglês-português executada por ele. Submetemos ao programa o seguinte texto escrito originalmente em inglês sobre diabetes, disponibilizado no sítio do jornal *The New York Times* como informação complementar de uma notícia sobre diabetes na Índia:

Diabetes

Diabetes mellitus is a group of diseases characterized by high blood sugar levels that result from the body's inability to make or use insulin, a hormone produced by the pancreas that plays a vital role in metabolism.

Symptoms include increased thirst and urination, hunger, weight loss, fatigue, and blurred vision. Diabetes can lead to debilitating and life-threatening complications including blindness, memory problems, kidney disease, heart disease, nerve damage, and amputations.

The most common type is type 2 diabetes, previously called adult onset diabetes. It usually develops because the body fails to use insulin properly. It occurs in people, including children, who are overweight; other risk factors include high cholesterol, high blood pressure, ethnicity, and a family history of diabetes. Treatment includes a healthy diet, weight loss, and regular exercise. Many patients require daily insulin injections.

Type 1 diabetes, formerly called juvenile diabetes, usually develops in childhood. It is caused by the inability of the pancreas to produce insulin. Genetic predisposition combined with exposure to viruses are the main risk factors. Treatment consists of carefully monitored insulin replacement, typically via needles or a special pump.

A small percentage of pregnant women develop gestational diabetes, sometimes labeled type 3, as a result of changing hormonal levels. Blood sugars often return to normal after delivery, but almost half of the women who experience gestational diabetes develop type 2 diabetes later in life.

- A tradução automática executada pelo *Power Translator* foi a seguinte:

Diabete

Mellitus de diabete é um grupo de doenças caracterizado por níveis de açúcar de sangue altos que são o resultado da inabilidade do corpo fazer ou usar insulina, um hormônio produzido pelo pâncreas que faz um papel vital em metabolismo. Sintomas incluem sede aumentada e urinação, tenha fome, perda de peso, fadiga, e visão borrada. Diabetes podem conduzir a debilitar e complicações vida-ameaçadoras inclusive cegueira, problemas de memória, doença de rim, doença de coração, dano de nervo, e amputações.

O tipo mais comum é tipo 2 diabetes, diabete de adulto onset previamente chamada. Normalmente desenvolve porque o corpo não usa insulina corretamente. Acontece em pessoas, inclusive crianças que são preponderantes; outros fatores de risco incluem colesterol alto, pressão alta, ethnicity, e uma história familiar de diabete. Tratamento inclui uma dieta saudável, perda de peso, e exercício regular. Muitos pacientes requerem injeções de insulina diárias.

Diabete 1 diabete, diabete juvenil antigamente chamada, normalmente desenvolve em infância. É causado pela inabilidade do pâncreas para produzir insulina. Predisposição genética combinada com exposição a vírus é os fatores de risco principais. Tratamento consiste em substituição de insulina cuidadosamente monitorada, tipicamente por agulhas ou uma bomba especial.

Uma porcentagem pequena de mulheres grávidas desenvolve diabete de gestacional, tipo 3 às vezes etiquetado, como resultado de mudar níveis hormonais. Açúcares de sangue voltam freqüentemente a normal depois de entrega, mas quase a metade das mulheres que experimentam diabete de gestacional desenvolve tipo 2 diabetes em vida.

No geral, as informações contidas no texto original podem ser apreendidas nessa tradução, porém são freqüentes os erros de concordância e ordenação de palavras. O texto em português traz algumas passagens sem coesão e traduções equivocadas de termos. O programa foi incapaz de lidar com a polissemia da palavra *delivery*, que no contexto queria dizer "parto", mas foi traduzida como "entrega". Já a palavra *type*, recorrente no texto, sempre com o mesmo significado, foi traduzido ora por "tipo" ora por "digite", esta última ocorrendo em início de frase, interpretada erroneamente pelo programa como um imperativo. Isso mostra uma tentativa mal-sucedida do programa em lidar com a polissemia. A tradução traz as informações principais, mas tão mal apresentadas que seria difícil considerá-la um informativo confiável sobre diabetes. O programa traduz corretamente expressões como “*high blood sugar levels*”, fazendo as inversões necessárias e reconhecendo a função da cada palavra da expressão. No entanto, houve expressões mais simples das quais a tradução automática não deu conta, por exemplo “*later in life*”.

Há tradutores que, para economizar tempo e digitação, submetem os textos a uma tradução automática antes de iniciar o trabalho. Muitos profissionais questionam se esse tipo de expediente representa uma economia real e se ele afeta a qualidade do texto final. Cabe a cada tradutor averiguar quais projetos poderiam ser ajudados por esse tipo de ferramenta e quais se adaptam mais a outras formas de trabalho. Essa averiguação deve ser feita em função de critérios tais como a função do texto, o público leitor alvo, o grau de especificidade e o tipo de texto, por exemplo.

3. 2 Tradução Assistida por Computador

A partir de informações extraídas de artigos disponibilizados por Vera e Danilo Nogueira no sítio <http://www.nogueiratranslations.com.br/>, abordaremos pontos relevantes relativos aos programas de tradução assistida pelo computador, também chamados de

programas de auxílio ao tradutor. Esse tipo de *software* é amplamente utilizado pelos profissionais, sendo que grande parte das agências de tradução exigem o uso dessa ferramenta, para economizar tempo e tornar as traduções mais uniformes. Diferente dos programas de tradução automática, aqui é o profissional quem traduz, cabendo à máquina auxiliá-lo, tornando seu trabalho mais rápido e organizado. A vantagem mais conhecida desses *softwares* é o uso das chamadas memórias de tradução, que geram um corpus bilíngüe com as sentenças traduzidas no programa. Durante a tradução, o programa consulta esse corpus automaticamente para oferecer sugestões, mas pode-se também consultá-lo manualmente. A maioria dos *softwares* de auxílio ao tradutor permite também inverter o sentido do corpus, de modo a usar, por exemplo, o corpus inglês > português como português > inglês. Os bons programas administram também glossários e até advertem quando o termo sugerido é negligenciado. Permitem a tradução de arquivos em diversos formatos além de .doc, a saber, arquivos html, ppt, xls, xml, entre outros.

Muitos defendem que esse tipo de programa só é útil quando o texto tem várias repetições, o que não é verdade, pois além dos recursos de memórias de tradução, esses *softwares* desempenham uma série de outras tarefas:

- segmentam o texto, organizando melhor o material a ser traduzido e impedindo que o tradutor salte sem querer uma frase ou trecho. No entanto, se quiser saltar, não há problema;
- transcrevem (semi-)automaticamente nomes próprios e números;
- preservam a formatação original;
- gerenciam glossários, automaticamente exibindo os termos e as traduções à medida que vão aparecendo no texto, permitindo que o tradutor insira a tradução que achar melhor.

Alguns dos programas de tradução assistida por computador disponíveis no mercado são: DejáVu (www.atril.com), SDL Trados (www.sdl.com), StarTransit (www.star-group.net) e Wordfast (www.wordfast.net).

Em função de minha própria experiência, passo a detalhar o funcionamento do *Wordfast*, que, dos programas mencionados, é o mais simples, funcional e barato. É o programa ideal para iniciantes, pois além de desempenhar bem todas as funções a que se propõe, funciona em conjunto com vários outros programas e é disponibilizado gratuitamente (na versão gratuita o tamanho dos arquivos de memória com que se pode trabalhar é limitado, mas suficiente para a maioria dos projetos).

O *Wordfast* funciona "dentro" do *MS Word*, pois na verdade não chega a ser um programa, e sim um conjunto de macros, ou seja, comandos pré-programados para desempenhar funções específicas no *MS Word*. O usuário deve baixar o programa no sítio oficial (<http://www.wordfast.net>), juntamente com o manual disponibilizado em vários idiomas, e seguir as instruções para instalá-lo.

Uma vez instalado, o programa se apresenta como uma barra de ferramentas adicional do *MS Word*. Nela estão todos os comandos necessários para utilizar o *Wordfast*, embora seja bem mais prático comandá-lo através das teclas de atalho, que o usuário proficiente deve dominar. Quando se passa a seta do mouse sobre os "botões" de comando do *Wordfast*, o programa exibe o nome do comando e entre parênteses a tecla de atalho correspondente a ele, o que faz com que com o usuário atento se familiarize rapidamente com todas elas.

Abre-se então o texto a ser traduzido e inicia-se o *Setup* do programa, na tecla chamada *Wordfast* (símbolo *f*). O menu de configurações se abre, para que o tradutor faça os ajustes iniciais antes de começar seu trabalho. A primeira coisa a fazer é sempre configurar glossários e memórias de tradução. No caso de um usuário iniciante, deverá criar uma nova memória, clicando em *Translation Memory > TM > New TM*, escolhendo as línguas fonte e alvo da memória e salvando a nova memória (como um arquivo .txt) com o nome que quiser, na pasta que quiser (Figura 1). Lembramos que ao lidar com um grande volume de textos, memórias e glossários, é imprescindível manter os arquivos organizados, com nomes reconhecíveis e

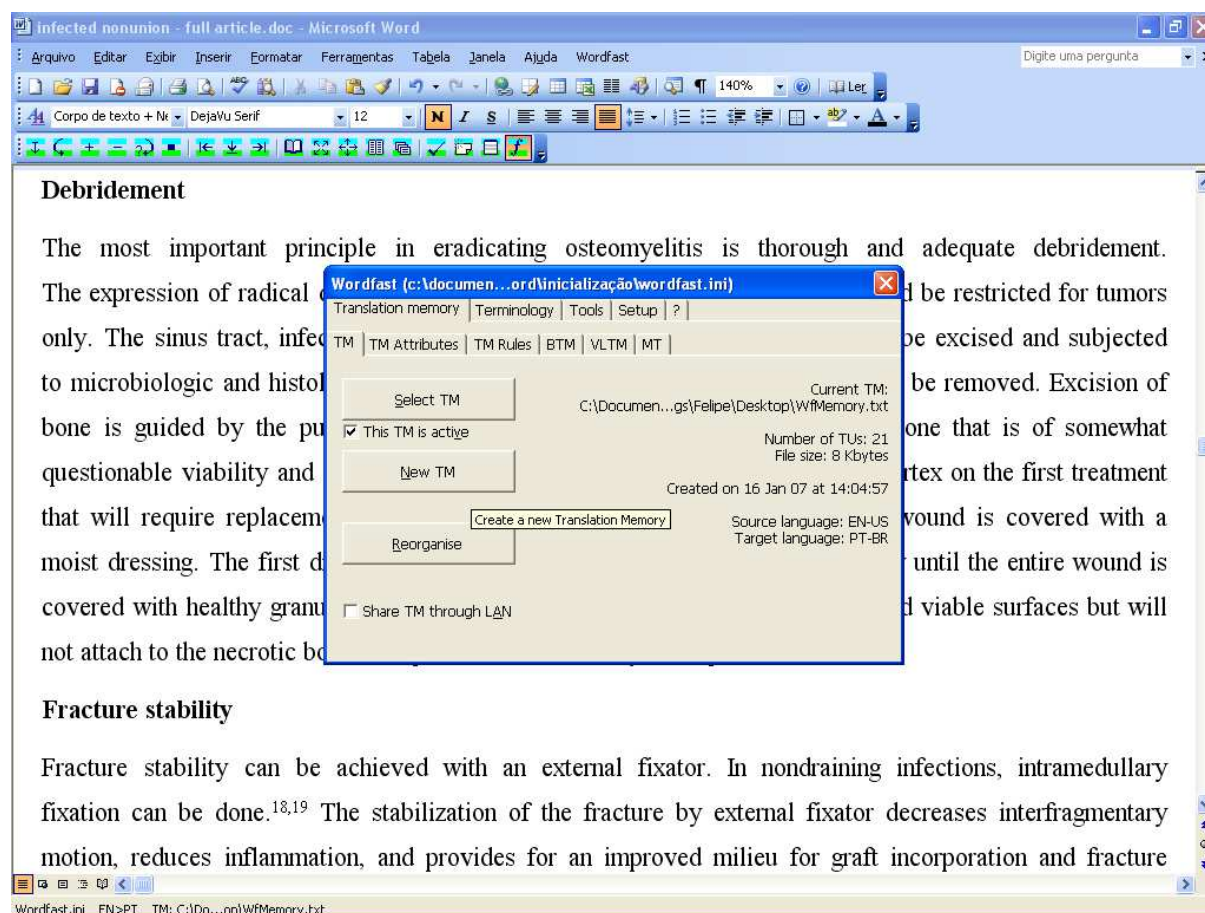


Figura 1: Janela de configurações do Wordfast, na seção *Translation Memory*, mostrando a criação de uma nova memória de tradução. Acima do espaço destinado ao texto, nota-se a barra de ferramentas do Wordfast.

pastas ordenadas, para que se possa recuperar facilmente memórias e glossários antigos que podem ser úteis na tradução de textos sobre assuntos semelhantes.

No contexto das agências de tradução, que normalmente dividem um texto para muitos tradutores, usa-se fornecer ao tradutor, junto com o projeto, uma memória de tradução e glossários correspondentes, para garantir a uniformidade do trabalho. Nesse caso, em vez de criá-los, basta que o tradutor os abra respectivamente em *Translation Memory > TM > Select TM* e *Terminology > Glossary 1 > Select Glossary* antes de iniciar a tradução.

O Wordfast suporta glossários em formato .txt com até 200.000 verbetes. Palavras na língua fonte e alvo são separadas por marcas de tabulação. O exemplo a seguir é um trecho (verbetes com "e") de um glossário feito para auxiliar a versão de um artigo científico sobre o molusco *Leptinaria unilamellata*:

efeito estimulante	stimulant effect
efeito residual	residual effect
em jejum	fasting
embriões	embryos
enriquecida	enriched
enterramento	inhumation
epifragma	epiphragm
espécie	species
espécies	species
estado de vigília	state of vigil
esterilizada	sterilized
estivação	aestivation
estratégia de controle químico	chemical control strategy
estrutura do meio ambiente	environmental structure
estímulos externos desfavoráveis	adverse external stimuli
experimento*	experiment*
extrato butílico	butyl extract
extrato etéreo	ethereal extract
extrato etílico	ethyl extract
extrato*	extract*

Não é necessário que os verbetes estejam em ordem alfabética, pois o *Wordfast* tem uma ferramenta chamada *Reorganize Glossary* que cuida disso. Ademais, os glossários permitem a utilização do asterisco como caractere curinga, o que facilita a tradução das variações de gênero e número.

Se ao começar uma tradução o profissional não dispõe de um glossário pronto sobre o assunto, ele pode criar um a partir do *Wordfast*. Basta que abra o texto no *MS Word* e na janela de configurações do *Wordfast* (tecla *f*) escolha *Terminology > Glossary 1 > New Glossary*. O programa vai então oferecer as opções de criar um glossário de amostra (opção "sim") ou criar um glossário vazio (opção "não"). Seleciona-se a opção "não" e ele pede que salvemos o novo glossário (Figura 2). Feito isso, pode-se vasculhar o texto a procura não só do vocabulário mais difícil para incluí-lo no novo glossário, como também das palavras e expressões mais repetidas no texto, pois sua inclusão no glossário pode significar uma boa economia de tempo e digitação.

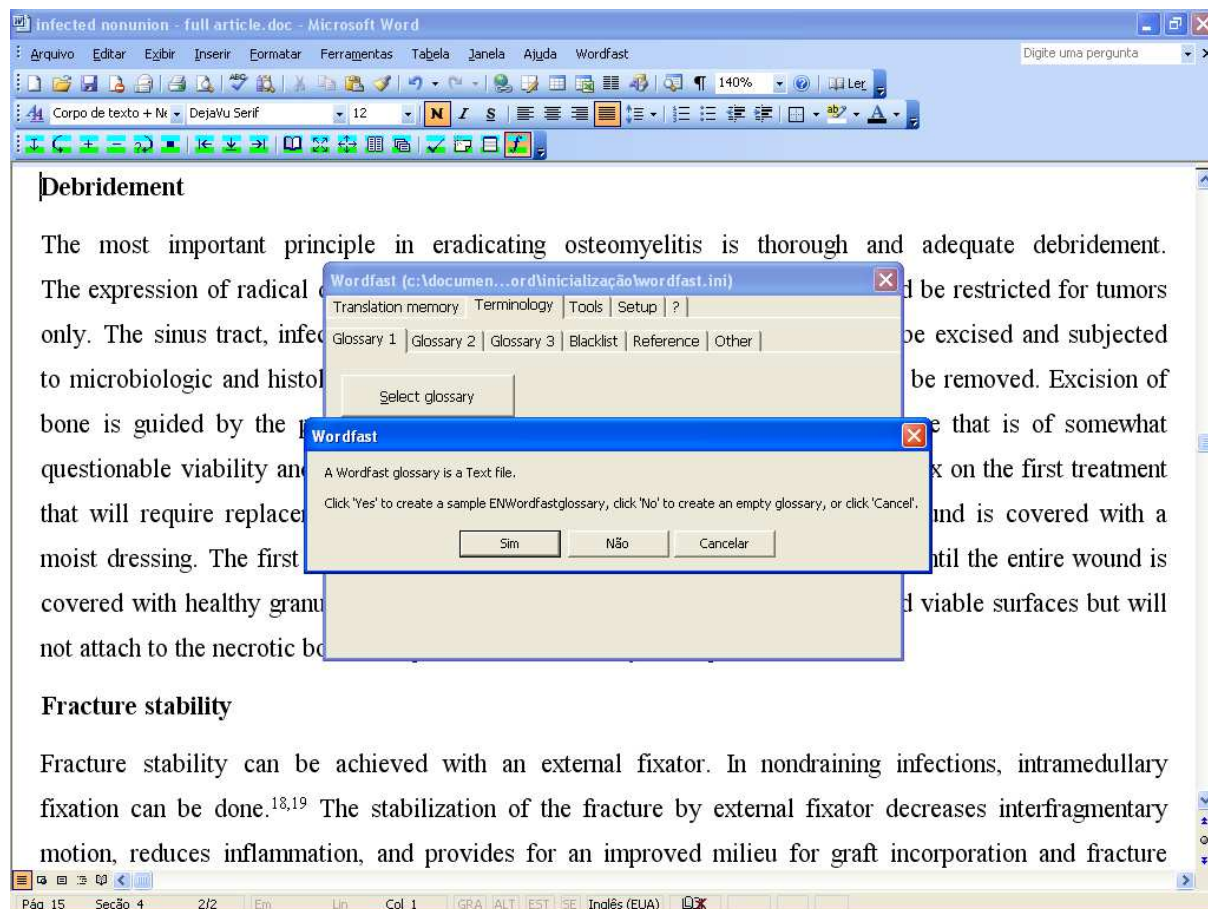


Figura 2: Janela de configurações do Wordfast, na seção *Terminology*, mostrando a criação de um novo glossário.

Para adicionar uma palavra ao glossário mantém-se o cursor sobre ela e usa-se a tecla de atalho "Ctrl, Alt, T". Uma pequena janela é aberta com uma linha para a palavra fonte e outra para o correspondente na língua alvo. Basta preenchê-las e clicar em "OK" (Figura 3). Recomendo dar especial atenção a essa fase inicial da tradução, criando um glossário o mais completo possível, pois assim, na hora da digitação do corpo do texto, não é necessário parar todo o tempo para buscar o sentido de palavras.

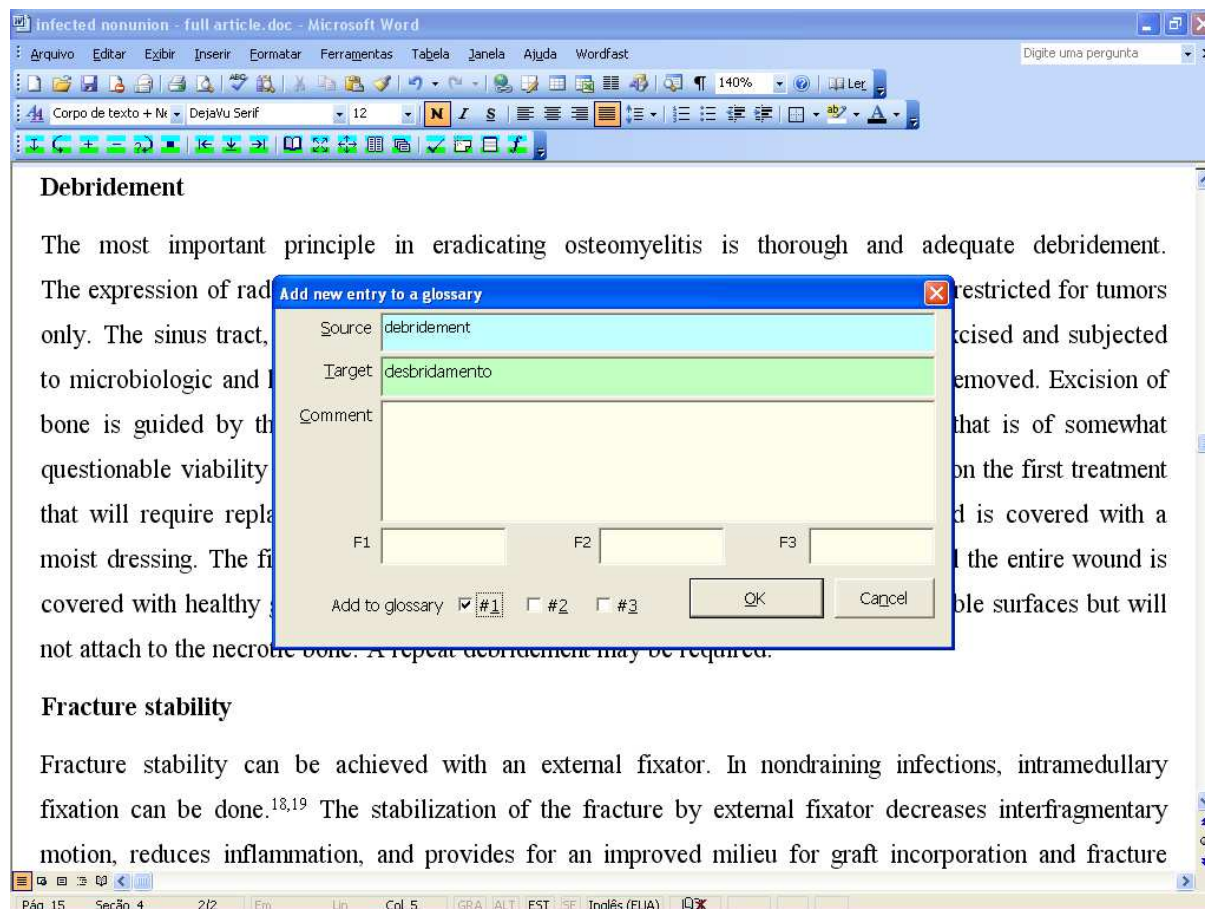


Figura 3: Ao utilizar a tecla de atalho “Ctrl, Alt, T” abre-se a janela *Add new entry to a glossary*, com espaços destinados ao termo na língua de origem (azul) e alvo (verde).

Feitas as configurações iniciais, começa-se a tradução propriamente dita. O *Wordfast* segmenta o texto baseando-se na pontuação, criando unidades de tradução que normalmente equivalem a um período. Para abrir o primeiro segmento usa-se o atalho "Alt, seta para baixo" ou a tecla "Start translation/Next segment" da barra de tarefas. O primeiro segmento do texto aparece então dentro de um retângulo azul e abaixo um retângulo cinza destina-se a digitação do texto na língua alvo (Figura 4). Após digitar a tradução, usa-se "Alt, seta para baixo" para passar ao próximo segmento. Sempre que no segmento aberto houver palavras presentes no glossário, elas aparecerão marcadas de azul mais escuro e serão transformadas em "placeables", isto é, elementos que podem ser inseridos no texto por meio de teclas de atalho, não precisando ser digitados todas as vezes que aparecem. Basta usar as teclas "Ctrl, Alt, seta esquerda" para selecionar diferentes "placeables" de um segmento e "Ctrl, Alt, seta para

baixo" para inserir a tradução dos termos (Figura 5). O usuário pode configurar o programa para tratar como "*placeable*" qualquer palavra em letra maiúscula, números, caracteres especiais etc. Para isto, o *Wordfast* oferece um recurso chamado *Pandora's Box*, uma série de configurações adicionais que podem facilitar ainda mais o trabalho do profissional. Além disso, o programa armazena na memória cada segmento traduzido, e uma vez aberto um segmento que contenha algum trecho semelhante a algo já digitado anteriormente, a mesma tradução é inserida automaticamente, e a cor do retângulo destinado à tradução muda, de acordo com a porcentagem de semelhança entre os trechos.

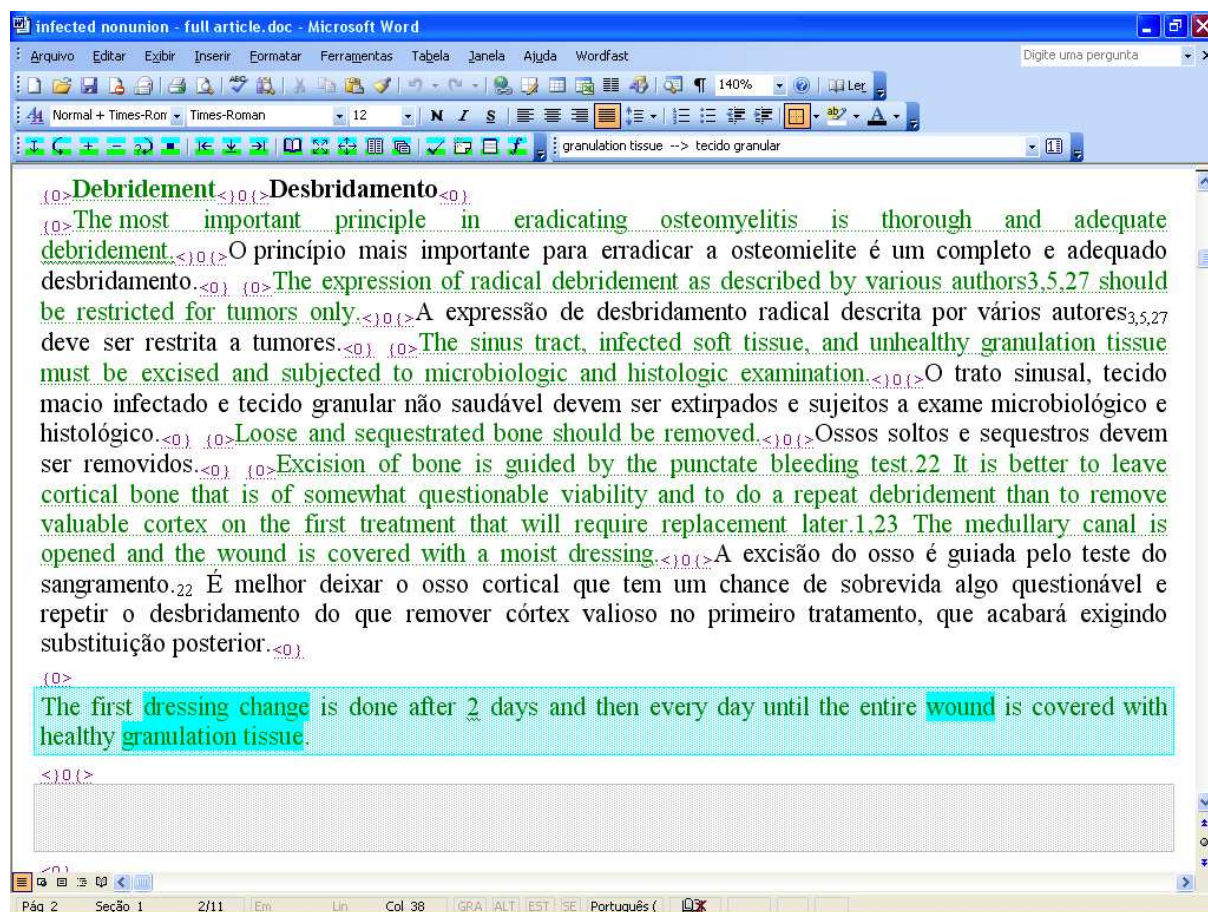


Figura 4: O retângulo azul traz o segmento de texto na língua fonte aberto, com os *placeables* destacados em azul mais escuro e, abaixo, o retângulo cinza destina-se à digitação da tradução.

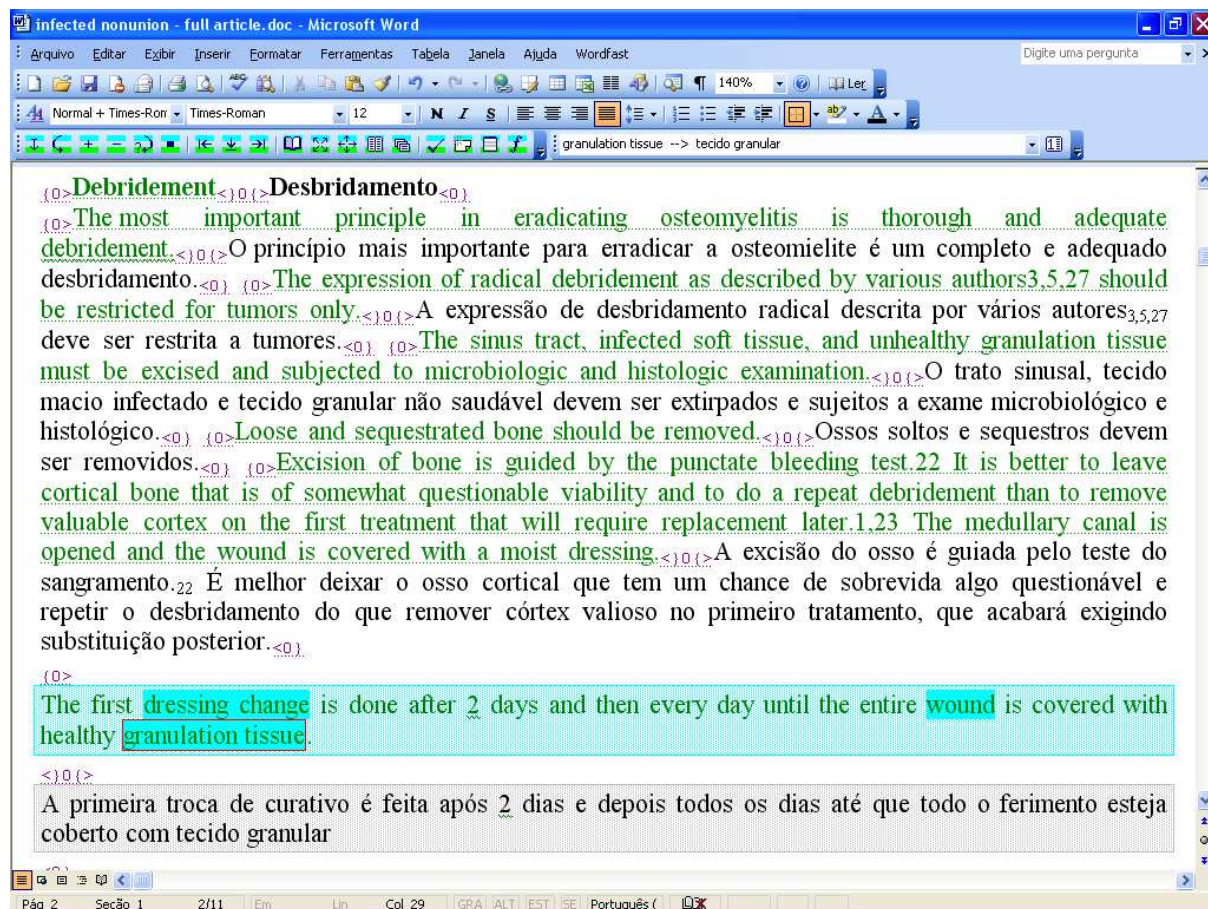


Figura 5: Aqui, nota-se um *placeable* selecionado (com um retângulo vermelho em volta) cuja tradução acaba de ser inserida pela tecla de atalho “Ctrl, Alt, seta para baixo”.

Ao fim da tradução, tem-se um arquivo com o texto original ocultado e o texto traduzido ocupando seu lugar, com a mesma formatação. Utiliza-se a tecla de atalho "Ctrl, vírgula" para alternar entre a visualização de texto original seguido por texto traduzido (entre eles, caracteres de marcação em roxo, como mostra a Figura 6) e somente texto traduzido (Figura 7). Para salvar um arquivo final somente com a tradução (eliminando o texto original do arquivo) executa-se a função *Clean up* (Ctrl, Alt, q).

Entre seus inúmeros recursos adicionais, o *Wordfast* disponibiliza também um pacote de ferramentas chamado *PlusTools*, que, entre inúmeras funções, serve para extrair textos de arquivos .pdf mantendo a formatação original e organizando-os num arquivo .doc. Este recurso é útil na tradução de artigos científicos e outros trabalhos disponibilizados somente em .pdf.

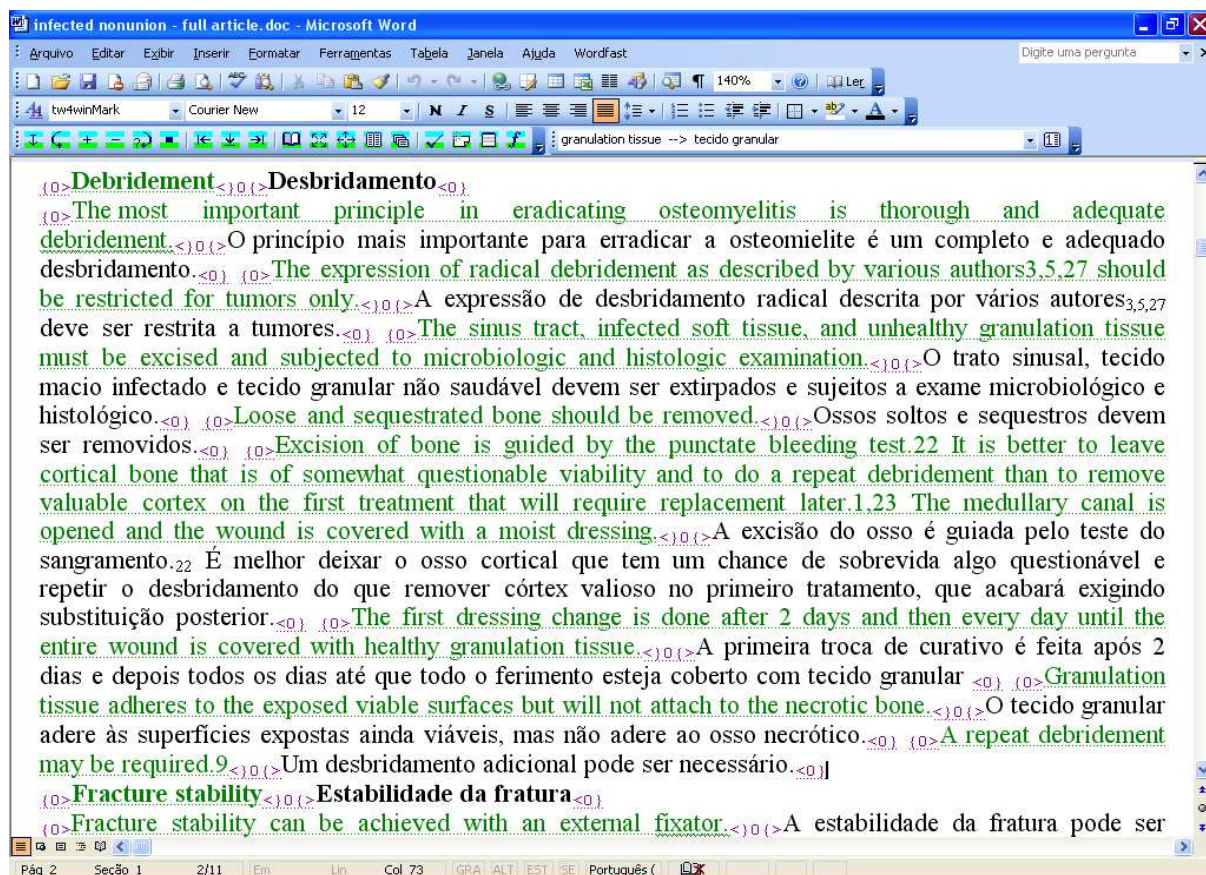


Figura 6: Visualização do texto original (em verde) seguido pelo texto traduzido. Entre eles, caracteres de marcação em roxo.

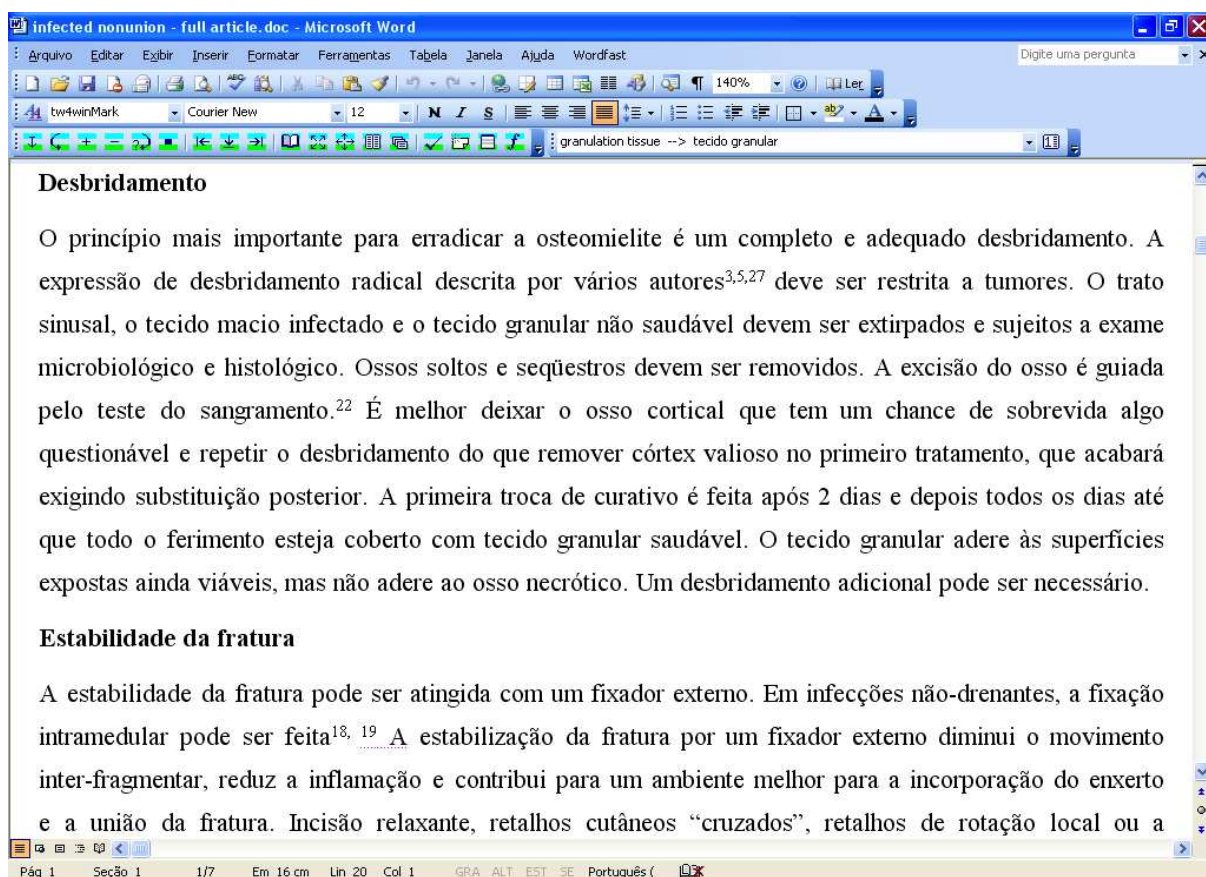


Figura 7: Visualização do texto traduzido somente. Versão final do documento após *Clean-up*. (Ctrl, Alt, Q).

Pode-se também usar o *Wordfast* em conjunto com um tradutor automático, por exemplo, o *Power Translator*. No menu de configurações em *Translation Memory > MT* marque a opção "*Menu, Sub-menu for MT*" e no espaço correspondente, escreva o nome dos comandos do programa de tradução utilizado, no caso do *Power Translator*, os comandos são "Traduzir, Oração". Assim configurado, cada segmento aberto será traduzido automaticamente, cabendo ao tradutor somente averiguar a tradução feita e modificar o necessário. Com o auxílio de um glossário marcando as palavras mais específicas do texto (que o tradutor automático tem menos chances de traduzir corretamente) pode-se facilmente corrigir os erros de vocabulário da tradução automática, criando-se um modo de trabalho prático para a tradução de textos informativos.

Esse modo de trabalho pode representar uma economia de tempo sem perda na qualidade, uma vez que a maior parte do trabalho manual será poupada; porém é preciso considerar antes a que tipos de texto essa tradução semi-automática se adapta bem. Em alguns casos, ela seria mais uma perda de tempo que uma economia, por exemplo, em textos com excesso de vocabulário especializado.

Neste trabalho sobre o ofício do tradutor e os inúmeros auxílios prestados a ele pela informática, lidamos com um universo de pesquisa amplo e em constante transformação. A infinidade de programas e fontes de pesquisa disponíveis hoje em dia torna impossível tratar desse assunto de forma exaustiva. O que fiz aqui foi meramente sugerir alguns dos recursos com os quais estou mais familiarizado e descrever o modo de funcionamento de cada um, de modo a produzir um material de consulta que sirva aos tradutores, profissionais ou em formação, interessados em dinamizar seu modo de trabalho e adequar-se aos prazos e exigências do mercado. Apresento uma abordagem prática e simplificada das ferramentas, de forma a contemplar aqueles que ainda estão “por fora” desse universo de *softwares* e fontes de pesquisa, recursos cada vez mais necessários aos que fazem (ou tentam fazer) da tradução a sua fonte de sobrevivência.

O universo dos programas e sítios de pesquisa aqui descritos muda tão constantemente que se torna difícil fazer uma descrição completa e atualizada dos recursos por eles oferecidos. Sendo assim, busquei ressaltar as características que considero mais úteis e práticas em minha rotina de trabalho, principalmente na tradução de textos acadêmicos. No caso das fontes de pesquisa, como sítios de busca e dicionários *on-line*, citei aquelas que costumo usar em meu trabalho, incluindo também alguns sítios utilizados por colegas tradutores. Para os *softwares* de auxílio aos tradutores, segui a mesma linha. Tratando-se de uma abordagem pragmática dessas ferramentas, limitei-me a descrever-lhes o funcionamento e principais vantagens, sem aprofundar-me em críticas e conclusões teóricas.

A forma com que cada um conduz as pesquisas em sua rotina de trabalho é algo extremamente pessoal, e cada profissional sabe o ritmo que lhe permite um rendimento maior. É certo que as ferramentas aqui apresentadas podem ser adaptadas a inúmeros modos de trabalho diferentes dos descritos por mim. Busco ressaltar a versatilidade e adaptabilidade

desses *softwares*, e deixo claro que os itens descritos aqui correspondem a uma ínfima parte dos recursos disponíveis ao usuário proficiente da Internet.

Ao analisarmos o histórico das ferramentas eletrônicas criadas para auxiliar a tradução, vimos as inúmeras tentativas já empreendidas pelo homem de criar máquinas aptas a substituir o ser humano no processo tradutório. Vimos também a impossibilidade dessa empreitada e a criação de *softwares* mais "realistas", que buscam auxiliar o profissional da tradução, tornando seu trabalho menos repetitivo e sujeito a erros. Porém, foi há pouco mais de dez anos que a profissão do tradutor foi totalmente revolucionada, com o surgimento da Internet. O universo de pesquisa em constante expansão fez surgir um novo perfil de tradutor, não mais preocupado em acumular glossários, livros e cadernos de anotações, mas um pesquisador profissional, totalmente familiarizado com o ambiente virtual, apto a vasculhar exaustivamente a Internet em busca de textos de referência e definições de termos. Com os computadores e a Internet, não há mais necessidade de acumular informações, uma vez que a memória virtual o faz de forma muito mais segura. Isso obviamente não significa que os tradutores atuais não devam ter um grande conhecimento nas línguas com que trabalham, mas que não precisam ter memorizados termos técnicos de inúmeras áreas, desde que saibam onde e como pesquisar, sempre que necessário.

Na última década, muitos artigos têm sido produzidos em todos os idiomas sobre os temas discutidos aqui. Todos os profissionais da tradução que desejam aprofundar seu conhecimento e técnica nessa área devem estar atualizados com as publicações mais recentes sobre o assunto, bem como participar de grupos de discussões, seminários e congressos na área. A troca de idéias e informações é fundamental numa profissão que promove, por excelência, a comunicação entre as pessoas e a difusão do conhecimento.

1. Livros, artigos e monografias:

HUTCHINS, John. The Origins of the Translator's Workstation. In: *Machine Translation 13*. Netherlands: Kluwer Academic Publishers, 1998. P. 287-307

LANGLAIS P., FOSTER G. & LAPALME G. Unit Completion for a Computer-aided Translation Typing System. In: *Machine Translation 15*. Netherlands: Kluwer Academic Publishers, 2000, p. 267-294

MAGALHÃES, Célia. Pesquisas textuais/discursivas em tradução: o uso de *corpora*. In: PAGANO, Adriana (Org.). *Metodologias de pesquisa em tradução*. Belo Horizonte: Faculdade de Letras, UFMG, 2001. p. 90-116.

ROCHA, L.C. *A tradução na era da informática*. Monografia de conclusão do Bacharelado em Letras – Ênfase em Tradução, DLEM, FALE-UFJF, 1990 (Mimeo.)

SARDINHA, T. B. Corpora eletrônicos na pesquisa em tradução. In: TAGNIN, S. E. O. (Org.). *Cadernos de Tradução: Corpora e Tradução*. Florianópolis: NUT, 2002, v. 1, n. 9, p. 15-59.

SOMERS, H. Machine Translation. In: BAKER, M (ed.). *Routledge Encyclopedia of Translation Studies*. New York, 1998, p. 134-149

STRANZY, Philip. *Encyclopedia of Linguistics*. New York: Fitzroy Dearborn, 2005, vol.1, pp. 124-126

TARP, Sven: Propuestas para la traducción especializada mediante un sitio web. In: *TradTerm: Revista do Centro Interdepartamental de Tradução e Terminologia*, vol 8, pp 257-274, 2002.

2. Páginas da Internet:

<http://www.fflch.usp.br/dlm/comet/> (acessado em 25/10/2006)

http://www.unibero.edu.br/glossarios_def.asp (acessado em 30/10/2006)

<http://www.bussolaescolar.com.br/dicionarios.htm> (acessado em 30/10/2006)

<http://www.nogueiratranslations.com.br/> (acessado em 15/11/2006)

<http://www.windpower.org> (acessado em 10/12/2006)

http://www.askoxford.com/dictionaries/compact_oed/ (acessado em 13/01/2007)

<http://dictionary.cambridge.org/> (acessado em 13/01/2007)

<http://www.collins.co.uk/Corpus/CorpusSearch.aspx> (acessado em 15/01/2007)

http://www.umist.ac.uk/ctis/research/TEC/tec_home_page.htm (acessado em 16/01/2007)