

1 Introdução

A semântica dos compostos nominais em língua inglesa tem sido objeto de investigação de muitas pesquisas no âmbito da Linguística e fora dela também. A interpretação das relações semânticas entre os constituintes de um composto é difícil, porque não é facilmente depreendida a partir de uma análise composicional dos substantivos que fazem parte da estrutura; muitas vezes a compreensão das relações semânticas depende do conhecimento de mundo e é idiossincrática (o que faz sentido para um falante pode não fazer para outro); há poucas pistas sintáticas; os compostos oscilam entre a transparência semântica (composicionalidade) e a opacidade (lexicalização), e qualquer substantivo pode ser combinado com um outro e formar um composto que faça sentido em determinado contexto.

Ademais, mesmo que os termos de um composto não sejam ambíguos, como *woman e doctor* e *Chicago e flights*, por exemplo, a combinação entre essas palavras pode ser ambígua. Segundo Finin (1980:310), “a woman doctor may be a doctor who is a woman or a doctor whose patients are women. Similarly, Chicago flights may be those bound for Chicago, coming from Chicago or even those making a stop in Chicago”¹.

O presente trabalho tem como objetivo principal investigar a semântica dos compostos nominais retirados de um texto técnico, bem como analisar, do ponto de vista sintático e semântico, os resultados tradutórios em português.

1.1 Relevância do tema

Os compostos nominais do tipo N + N são estruturas bastante recorrentes em língua inglesa, principalmente em textos especializados. A precisão terminológica exigida em textos técnicos, além da abundância da utilização dos compostos nominais nesse tipo de texto, torna a investigação das relações semânticas entre os termos dos compostos relevante e necessária.

A vasta bibliografia sobre a sintaxe e a semântica dos compostos nominais em língua inglesa contrasta com a reduzida literatura sobre o tema em língua

¹ “woman doctor pode ser um médico do sexo feminino ou um(a) médico(a) que atenda mulheres. Igualmente, Chicago flights pode se referir a voos rumo a Chicago, vindos de Chicago ou que fazem escala em Chicago”.

portuguesa. O número de publicações em português é inversamente proporcional à complexidade do tema e aos problemas que surgem na tradução.

Em relação aos textos técnicos, embora haja, segundo Leipnitz (2005), uma grande demanda desse tipo de texto pelo mercado de tradução, a produção de material de pesquisa sobre tradução técnica é relativamente pequena no Brasil, se comparada à de textos ditos literários. Dessa forma, este trabalho pretende contribuir para a pesquisa em terminologia, através do estudo sobre a característica dos compostos nominais em língua inglesa (oriundos da terminologia técnica), as dificuldades de tradução desse tipo de estrutura para o português e, especialmente, sobre a caracterização sintática e semântica das estruturas correspondentes em língua portuguesa.

Apesar da pouca pesquisa sobre o assunto, a tradução técnica é “a que mais remunera o profissional tradutor, exigindo um bom conhecimento de linguagens especializadas e familiaridade com suas terminologias” (*idem*:13).

O principal problema reside no fato de que muitas traduções técnicas são feitas por profissionais de determinada área; por exemplo, traduções médicas são feitas por médicos, traduções na área de engenharia, por engenheiros etc. Parte desses profissionais não são tradutores, por isso, em muitas situações, não têm um conhecimento linguístico especializado nem na língua de partida, nem na de chegada. Dessa forma, é muito comum que a preocupação com a terminologia não seja igual ao cuidado com a língua-alvo. Ademais, detalhes linguísticos — mas não menos importantes do que detalhes terminológicos — passam despercebidos (como uma possível diferença semântica entre traduzir um composto nominal por N + Adj. ou por N + SP, como em ‘diferença potencial’ e ‘diferença de potencial’ para *potential difference*).

Por outro lado, um profissional com formação em tradução deve prezar pelo bom uso da língua. Quando o profissional traduz algum material de determinada área técnica, ele deve tentar suprir a falta de conhecimento especializado através de consultas a fontes confiáveis e/ou a profissionais que atuam no ramo, ou, a longo prazo, através de estudos a fim de se especializar minimamente em alguma área técnica. Esse tipo de profissional deve enfrentar a falta de conhecimento técnico e minimizá-la, ou até mesmo supri-la, através de

estudos que possibilitem uma melhor compreensão da formação dos compostos e das relações semânticas entre os constituintes.

A visão de tradução que permeia este trabalho está diretamente relacionada à experiência da tradutora. Embora os Estudos da Tradução tragam discussões bastante profícuas e muito bem embasadas sobre a atividade tradutória e de como o seu resultado está diretamente relacionado ao sujeito tradutor, com suas contingências, visão de mundo, interpretação do texto fonte etc., a prática tradutória revela que o tradutor precisa tomar decisões bem objetivas, visando atender o público-alvo, respeitar ou até mesmo acatar ordens de editoras e clientes, nunca deixando de lado o texto fonte. Ainda que muitos teóricos critiquem uma visão dicotômica da linguagem, ou seja, de que haveria uma tradução certa e outra errada, e abominem a visão de que os sentidos são estáveis, o tradutor sabe que os seus possíveis deslizes e erros serão sempre apontados. Em se tratando de um texto técnico então os erros ficam ainda mais visíveis e comprometem bastante o resultado tradutório.

É importante frisar que a complexidade dos compostos nominais não aguça a curiosidade apenas dos tradutores ou profissionais de Letras, mas sim de engenheiros, matemáticos, físicos e médicos, por exemplo, que lidam diariamente com a terminologia técnica (seja com o termo original ou sua tradução), e se questionam sobre as relações semânticas entre os constituintes dos compostos e os resultados tradutórios. Por exemplo, para o termo *central limit theorem*, um aluno de doutorado em engenharia eletrônica da UFRJ optou por utilizar a tradução “teorema central do limite”. Seu orientador pediu-lhe que trocasse para “teorema do limite central”, por esta ser a tradução mais difundida. Não satisfeito, o aluno recorreu a um livro de probabilidade, escrito em português por um americano, que esclarece exatamente a natureza e a relevância do problema. Em *Probabilidade: um curso em nível intermediário* (2004), o autor Barry James afirma:

na expressão ‘leite de vaca puro’, é obvio, pela concordância, que não estamos opinando sobre a vaca. Mas em espanhol, onde ‘leche’ é feminina, a expressão seria altamente ambígua. Temos o mesmo problema em português com a expressão ‘**Teorema do Limite Central**’, atualmente muito usado no Brasil. Assim como está, esta frase dá a nítida impressão de que é o limite que seja central, o que na realidade não faz sentido. Por isso, optei por ‘**Teorema Central do Limite**’, para afastar qualquer dúvida sobre o que seja central. É interessante

que a frase em inglês, ‘**Central Limit Theorem**’ é também altamente ambígua, e acho que isto explica a tradução de uso corrente no Brasil.²

Embora o exemplo fornecido pelo autor apresente, respectivamente, um modificador adjetivo não predicativo e um substantivo, além do núcleo, pode-se dizer que a natureza do problema é a mesma da dos compostos contendo apenas modificadores nominais, uma vez que os adjetivos não predicativos apresentam características dos substantivos, muitas vezes funcionando como tal, conforme será mostrado no Capítulo 3.

Esta dissertação visa divulgar as dificuldades de tradução originadas a partir dos compostos nominais, além de contribuir para uma maior compreensão da estrutura, e despertar uma maior consciência das relações semânticas subjacentes. Dessa forma, este trabalho é relevante tanto para tradutores quanto para professores e alunos de inglês, alunos de Letras ou até mesmo das áreas exatas.

Acrescente-se ainda que um estudo semântico dos compostos nominais é de grande interesse para a Linguística Computacional. Os estudos em PLN (Processamento de Linguagem Natural) já apresentam grandes avanços quanto à anotação morfossintática automática. Segundo Teixeira (2009), há ferramentas que incluem informações morfossintáticas de maneira automática. Por exemplo, para a frase “O livro é de Pedro”, é possível atribuir a etiqueta N (do inglês *noun*) para “livro”, de modo a identificar a palavra como substantivo e Subj (do inglês *subject*) para classificar a expressão “o livro” como sujeito da frase. No entanto, de acordo com o que consta no *site* sobre o PLN no Brasil, o PLN BR, “[u]m dos aspectos de anotação linguística de *corpus* menos desenvolvido é relacionado com a anotação semântica”.

Em língua inglesa, Girju *et alii* (2007) relatam que, além das relações semânticas, os pesquisadores têm demonstrado grande interesse na interpretação automática dos compostos nominais. Há uma série de publicações sobre interpretação semântica dos compostos nominais do tipo N + N que utilizam diversas metodologias, como a análise sintática através de técnicas estatísticas (Lauer e Dras, 1994; Pustejovsky *et alii*, 1993); o aprendizado de máquinas (Ó Segahdha, 2007); métodos probabilísticos (Lauer, 1995; Lapata e Keller, 2004;

² Grifos meus.

Girju, 2006; Turney 2006) e a indexação semântica latente (Buckridge & Sutcliffe, 2002).

1.2 Objetivos específicos

Através da investigação da semântica dos compostos nominais, pretende-se propor uma classificação dos compostos nominais técnicos, partindo da categorização semântica dos compostos nominais não técnicos adotada em diferentes trabalhos e fazendo as devidas adaptações. Pretende-se também analisar sintática e semanticamente as estruturas correspondentes em português, destacando a distribuição dos termos e a função das preposições.

Cabe também verificar os padrões semânticos observados entre os constituintes dos compostos, assim como o grau de transparência semântica das estruturas em língua inglesa.

No caso dos compostos com múltiplos modificadores nominais, observar-se-á também se há uma hierarquização das relações semânticas, ou seja, se uma relação é mais geral do que outra, se uma impossibilita a ocorrência da outra em determinado composto etc.

Ademais, visa-se agrupar os compostos de acordo com o mesmo núcleo ou o mesmo modificador, de modo a observar se núcleos ou modificadores iguais implicam relações semânticas iguais, e se, no caso de núcleos iguais, os modificadores especializam o sentido do núcleo, podendo agrupar esses compostos em ontologias.

1.3 Organização do trabalho

Além deste capítulo introdutório e o de conclusão, a presente dissertação apresenta ainda mais sete capítulos.

O Capítulo 2 aborda, em um primeiro momento, o conceito de composto nominal, apresenta uma breve caracterização sintática da estrutura, mostra a classificação dos compostos em endocêntricos e exocêntricos, e discute a questão da produtividade e recursividade dos compostos.

O Capítulo 3 apresenta a revisão de dois estudos clássicos sobre a semântica dos compostos nominais. Uma pesquisa foi realizada por Judith Levi e a outra por Beatrice Warren. Ambos os trabalhos foram publicados em 1978.

Após a abordagem dos estudos clássicos, o Capítulo 4 apresenta a Teoria do Léxico Gerativo, de Pustejovsky (1995), que deu um novo fôlego aos estudos sobre a semântica dos compostos nominais e serviu de inspiração para que pesquisadores da área da Linguística Computacional e campos afins desenvolvessem métodos que auxiliassem a interpretação semântica automática ou semiautomática dos compostos nominais. Também será mostrada a proposta mais recente de caracterização dos compostos nominais de Warren (2003), tendo como parâmetro a estrutura *qualia* da Teoria do Léxico Gerativo.

Apresentadas as principais características dos compostos nominais em inglês, assim como algumas propostas de classificação semântica, é necessário analisar as possibilidades tradutórias em português do ponto de vista estrutural, assim como a especificidade de cada estrutura. Para tanto, no Capítulo 5, recorreu-se a gramáticas normativas (Bechara, 2003; Cunha & Cintra, 2001; Rocha Lima, 2001), descritivas (Perini, 1995) e funcionais (Neves, 2000) e a um trabalho dedicado apenas aos estudos dos SNs (Perini *et alii*, 1996) em língua portuguesa. Ademais, verificam-se a distribuição dos termos de um SN em português e as relações semânticas expressas pelas preposições em língua portuguesa, além do grau de transparência semântica das mesmas.

O Capítulo 6 mostra as características e especificidades de um texto técnico, assim como a escolha do *corpus* e a metodologia utilizada para selecionar os dados. Além disso, são apresentados os critérios utilizados para manter ou descartar determinados compostos pré-selecionados para compor o *corpus*.

O Capítulo 7 apresenta uma proposta de classificação dos compostos selecionados, mostrando as características das categorias semânticas pertinentes para analisá-los, assim como exemplos de compostos nominais técnicos e não técnicos inseridos em cada categoria.

O Capítulo 8 apresenta a análise e discussão dos dados. Além disso, o capítulo traz a discussão sobre a característica dos substantivos integrantes dos compostos, visando a identificação de padrões semânticos nas combinações encontradas; a transparência semântica e opacidade dos compostos do *corpus*, e a

análise comparativa entre compostos cujos núcleos ou modificadores são iguais, de modo a verificar se núcleos ou modificadores iguais implicam relações semânticas iguais em compostos representantes de um domínio específico. A noção de ontologia, muito importante para a Linguística Computacional, já que auxilia a busca semântica, também é discutida no capítulo. As ontologias organizam hierarquicamente o conhecimento de um domínio específico.

Após a Conclusão (Capítulo 9), que lista possíveis futuros desdobramentos da pesquisa, apresentam-se as Referências Bibliográficas (Capítulo 10) e os Anexos, que contêm o *corpus* e as tabelas que agrupam, primeiramente, compostos com núcleos iguais e, em um segundo momento, com modificadores iguais.