

9 Conclusões

O objetivo principal deste trabalho foi analisar sintática e semanticamente os compostos nominais de língua inglesa oriundos de um domínio específico — a área de engenharia elétrica/eletrônica — e os resultados tradutórios em português.

Baseado em pesquisas sobre a semântica dos compostos nominais em língua inglesa, nesta dissertação, propôs-se uma tipologia que desse conta dos compostos nominais pertencentes à terminologia técnica. A proposta de classificação aqui sugerida foi inspirada nos estudos clássicos sobre a semântica dos compostos nominais (Levi, 1978; Warren, 1978) e em estudos mais recentes (a maioria inserida dentro da perspectiva da Linguística Computacional), e adaptada de acordo com a natureza dos dados.

Além da classificação dos 127 compostos nominais técnicos em nove categorias (PERTENÇA/CONSTITUIÇÃO; GÊNERO-ESPÉCIE; COMPARAÇÃO; MEDIDA; LOCAL; TÓPICO; PROPÓSITO; CAUSA/ORIGEM e USO), foram analisadas as estruturas utilizadas nas traduções para o português, assim como o valor semântico das preposições nos sintagmas preposicionados.

A categoria PERTENÇA/CONSTITUIÇÃO, que envolve as relações de posse, meronímia (todo-parte/parte-todo) e constituição (conteúdo-contidente e material), foi a mais produtiva no *corpus* analisado, o que vai ao encontro dos resultados obtidos nas pesquisas de Levi (1978) e Warren (1978). Segundo Levi, *have* ‘ter’ e *be* ‘ser’ são dois dos verbos mais vagos e, ao mesmo tempo, mais usados na língua inglesa. Muitas vezes esses dois verbos aparecem nas paráfrases dos compostos nominais inseridos na categoria PERTENÇA/CONSTITUIÇÃO (*company assets – the company HAS assets* ‘ações da empresa’/*copper coin – a coin which IS cooper* ou *a coin made of cooper* “moeda de cobre”). Ainda segundo a autora, diferentemente do predicado IN, que requer um substantivo que se refira a um local para funcionar como modificador do núcleo (*office friends* ‘amigos do escritório’), o predicado HAVE, que engloba em suas paráfrases os verbos *have* e *be*, seleciona mais livremente os argumentos nominais.

Em seguida, aparece a categoria de GÊNERO-ESPÉCIE, que estabelece uma relação hipônimo-hiperônimo. Esta relação é mais abrangente do que as outras. De acordo com a classificação aqui proposta, o composto *bronze statue* ‘estátua de bronze’, seria inserido na categoria PERTENÇA/CONSTITUIÇÃO. No entanto, se levarmos em consideração que uma estátua pode ser feita de diversos materiais, chega-se à conclusão de que é possível dizer *silver statue* ‘estátua de prata’, *ebony statue* ‘estátua de mármore’ etc. Sendo assim, *bronze statue* seria um tipo de estátua e poderia ser incluída também na categoria GÊNERO-ESPÉCIE, quando comparada a outras estátuas. A relação de gênero-espécie é importante para a construção de ontologias.

As categorias COMPARAÇÃO e MEDIDA também foram produtivas no *corpus*, respectivamente, porque a abstração dos conceitos torna necessário fazer comparações com algo palpável ou visível, ou ainda, usar convenções simbólicas e gráficas como sistema de representação desses conceitos, e porque o texto remete a muitos valores e relações, tornando importante a noção de medida.

As categorias LOCAL; TÓPICO; PROPÓSITO CAUSA/ORIGEM e USO foram as menos produtivas no *corpus* analisado.

A partir da categorização dos dados, foi possível observar uma possível limitação de combinação entre as categorias em compostos com múltiplos modificadores, e até de hierarquização das mesmas. No entanto, seria necessário um estudo mais aprofundado e um *corpus* maior para se confirmar tais hipóteses.

Como muitos compostos apresentavam o mesmo núcleo ou modificador, buscou-se verificar se núcleos ou modificadores iguais implicavam relações semânticas iguais entre os constituintes dos termos. Observou-se que muitas relações eram realmente iguais, principalmente em compostos com núcleos iguais.

A partir da organização dos compostos quanto à similaridade dos núcleos e dos modificadores, e inspirado no trabalho de Johnston *et alii* (1996), pensou-se na noção de ontologias locais, ou seja, na ideia de que haveria uma hierarquização entre os termos. *Transformers*, por exemplo, seria um termo mais geral, ou melhor, o hiperônimo de *air-core*, *iron-core* e *ferrite-core transformers*. Estes seriam o hipônimo de *transformers*. A combinação dos diferentes modificadores com o mesmo núcleo especifica o núcleo.

Para os compostos N + N em inglês, as estruturas utilizadas na tradução foram, em sua maioria, N + de + N. Nesse tipo de estrutura, a relação semântica

depende mais dos termos que antecedem e sucedem a preposição, uma vez que esta é mais vazia semanticamente do que “com”, “em e “para”, por exemplo. As estruturas N + de + art. + N; N + Adj.; N + N e N também foram utilizadas.

A principal dificuldade encontrada pela tradutora ao traduzir a estrutura N + N do inglês para o português era a falta de elementos (como, por exemplo, preposições) que tornassem as relações entre os termos mais clara. Atrelada à concisão das estruturas N + N, havia a falta de conhecimento técnico da tradutora, que tinha apenas o conhecimento linguístico para tentar interpretar semanticamente os compostos, além de, obviamente, a ajuda de profissionais especializados. No caso dos compostos com múltiplos modificadores nominais, a dificuldade era ainda maior, pois era necessário realizar uma análise da estrutura de constituintes a fim de identificar os agrupamentos nominais e as relações de modificação.

É possível dizer, portanto, que este trabalho é multidisciplinar, na medida em que propõe um diálogo entre diferentes vertentes linguísticas, e é interlingual, já que aborda questões e problemas relacionados à tradução de compostos nominais do inglês para o português. Sua maior contribuição talvez seja a análise semântica dos compostos nominais técnicos, sob uma perspectiva contrastiva.

9.1 Futuros desdobramentos

A análise semântica apresentada neste trabalho não pretende ser exaustiva. Há ainda muito a ser feito nessa área, principalmente sob uma perspectiva interlingual. É preciso verificar se as categorias semânticas aqui propostas são realmente adequadas para o tratamento dos compostos nominais oriundos da terminologia técnica; ampliar a discussão do agrupamento dos compostos em ontologias; e observar com mais critério uma possível hierarquização das combinações de categorias em compostos com múltiplos modificadores nominais.

Dado o número expressivo de trabalhos sobre a semântica dos compostos nominais inseridos no âmbito da Linguística Computacional (Macdonald, 1982; Isabelle, 1984; Rosario & Hearst, 2001; Buckeridge & Sutcliffe, 2002; Girju, 2007), parece um caminho natural investigar a semântica desse tipo de estrutura em uma perspectiva contrastiva, à luz dos estudos computacionais.

Deve-se destacar que, em língua portuguesa, ainda há pouca pesquisa sobre anotação semântica. Há também pouca pesquisa sobre a interpretação semântica automática de compostos nominais numa perspectiva contrastiva entre o inglês e o português, por exemplo¹⁰¹. No entanto, é importante citar a grande quantidade de trabalhos sobre a discussão de compostos em língua portuguesa (Smarsaro, 2004; Dias, 1994; Oliveira & Freitas, 2006).

Dessa forma, um dos possíveis desdobramentos é verificar em que medida uma caracterização das relações semânticas entre elementos que integram compostos nominais do tipo N + N pode contribuir para o aperfeiçoamento da anotação semântica, de ferramentas de tradução automática e de mecanismos de busca.

Para tanto, será necessário ampliar a investigação da estrutura e semântica dos compostos nominais retirados de um texto técnico, tendo em vista o arcabouço teórico fornecido pela Linguística Computacional.

Mais precisamente, poder-se-á

- (i) verificar se a tipologia aqui proposta apresenta-se como alternativa adequada para a caracterização das relações semânticas dos elementos que integram compostos nominais “técnicos”, tendo em vista um tratamento computacional dessas estruturas;
- (ii) investigar como os recursos de tradução automática processam as expressões multivoculares, com vistas a identificar os principais problemas decorrentes das diferenças estruturais e da natureza de modificação nas duas línguas;
- (iii) verificar até que ponto as ferramentas disponíveis são eficientes em realizar uma análise semântica automática dos multivocábulos (compostos nominais) e identificar principais falhas resultantes da tradução automática de compostos nominais;
- (iv) prover um tratamento dessas estruturas que permita minimizar as falhas na tradução automática;
- (v) contribuir para o aprimoramento da anotação semântica automática no *corpus* em português;

¹⁰¹ Vale mencionar que Girju (2007) publicou um trabalho sobre a interpretação semântica de compostos nominais utilizando *corpora* em língua inglesa e nas línguas românicas, propondo assim, um tratamento “entre-línguas”.

- (vi) investigar as relações de meronímia, sinonímia, hiperonímia e hiponímia e a importância das mesmas na construção de ontologias de compostos nominais “técnicos”;
- (vii) contribuir para a construção de ontologias, que, por sua vez, auxiliam na criação de glossários, tesouros etc.;
- (viii) contribuir para uma melhor interpretação semântica automática de compostos com múltiplos modificadores em língua inglesa, que, segundo Girju *et alii* (2005) ainda não é satisfatória, de modo a aperfeiçoar os resultados tradutórios em português utilizando ferramentas automáticas ou semiautomáticas.

Um outro possível desdobramento seria verificar os resultados tradutórios de outros tradutores para os compostos compilados, de modo a comparar com as estruturas aqui mostradas e verificar se os outros sujeitos tiveram as mesmas dúvidas da tradutora, se identificaram as traduções já canonizadas ou os compostos que apresentavam mais de uma leitura etc. Seria pertinente também mostrar a especialistas as diferentes traduções, para que eles avaliassem e justificassem a opção mais adequada para o contexto ou o domínio em questão. Dessa forma, poder-se-ia fazer um paralelo entre as informações semânticas contidas nos compostos e o uso das mesmas, ou seja, se as traduções preferidas pelos especialistas refletem a relação semântica entre os termos.