

6 O *corpus*

Neste capítulo serão discutidos a escolha do *corpus*, a metodologia usada para selecionar os dados e os critérios utilizados para manter ou excluir os compostos pré-selecionados para compor o *corpus*. Antes, porém, serão abordadas as características dos textos técnicos, uma vez que o *corpus* faz parte de um domínio especializado.

6.1 Textos técnicos e terminologia

Normalmente, os textos técnicos apresentam uma linguagem bastante repetitiva. O mesmo termo aparece inúmeras vezes no texto. Sendo assim, é necessário que o tradutor fique atento a essas repetições e tente reproduzi-las também na tradução. O mesmo termo técnico traduzido por várias expressões dá a impressão de que se está falando de coisas diferentes, e isso é facilmente depreendido por um leitor mais atento ou um especialista no assunto.

Um erro de tradução pode ser muito mais prejudicial em textos técnicos do que em textos literários. Paulo Rónai (1952) *apud* Ottoni afirma que “[u]m erro na versão de uma peça de Shakespeare, quando muito, indignará um crítico; mas na de uma bula de remédio ou de um formulário de materiais de construção pode ter consequências imprevisíveis” (1998:92). Sendo assim, preza-se por um cuidado quase que “cirúrgico” com os termos técnicos. Erros de terminologia ficam muito mais visíveis do que uma interpretação mal feita (e consequente tradução equivocada) de um trecho de um livro, por exemplo.

Deve-se ressaltar, no entanto, que há uma falsa sensação de estabilidade dos sentidos dos termos técnicos. A tradução de um texto técnico, juntamente com a pesquisa terminológica, revela que a terminologia é dinâmica e longe de ser consensual. Além disso, percebe-se nitidamente que algumas variações de terminologia estão condicionadas à adoção ou não do termo por parte de algumas universidades. Uma instituição prefere uma construção, a outra já utiliza uma expressão diferente etc. Por exemplo, para o termo *four-wire system*, foram encontradas as seguintes traduções: “sistema com quatro fios” (fonte: *site* da

UNICAMP), “sistema a quatro fios” (fonte: *site* da UFRJ) e “sistema de quatro fios” (fonte: *site* da UnB). A única diferença é em relação à preposição utilizada.

Diferentemente do que se imagina, a tradução técnica, apesar de não ser um exercício literário, apresenta os mesmos problemas linguísticos de uma tradução dita literária, uma vez que em ambas as modalidades de texto o tradutor está lidando com a linguagem. Segundo Rónai, (1952) *apud* Ottoni (1998:94), “[a]inda que o tradutor supere todas as dificuldades da terminologia técnica, só poderá fazer trabalho satisfatório se manusear com igual eficiência os termos não técnicos: verbos, pronomes, conjunções e preposições”.

Dessa forma, o cuidado com a língua em traduções técnicas deve ser o mesmo do que em traduções literárias. Ainda que a linguagem técnica seja mais enxuta e objetiva do que a tradução literária, é importante manter a fluência da língua de chegada, ou seja, os textos técnicos devem se adequar às normas da língua-alvo.

Deve-se destacar que os textos técnicos também estão condicionados a fatores culturais, à pessoa do tradutor, aos envolvidos na revisão (suas crenças, instituições a que pertencem), à situação de recepção etc. Logo, ainda que o conhecimento específico seja muito importante para a realização de uma tradução técnica, ele “não pode garantir, sozinho, a qualidade da tradução de um texto. Afinal, a produção e a tradução de textos não técnicos também envolvem conhecimentos específicos sobre os mecanismos de produção de sentido na linguagem” (Leipnitz, 2005:72).

6.2 *Corpus* e metodologia

Nesta seção serão apresentadas as características do *corpus* analisado na presente dissertação, assim como a metodologia utilizada para criar o *corpus* e definir o escopo do trabalho.

Como o objetivo do presente trabalho é apresentar uma análise semântica dos compostos nominais e os problemas enfrentados na tradução para o português, seria necessário compilar um *corpus* em que houvesse a versão original em inglês e a respectiva tradução para o português. A princípio, pensou-se em procurar um texto na área das ciências humanas que tivesse sido traduzido para o português e, a partir dele, extrair os compostos nominais e verificar as estruturas utilizadas na

tradução. No entanto, como a autora estava traduzindo um livro técnico durante o primeiro ano de sua pesquisa e levava muitos exemplos e questionamentos às reuniões com seus orientadores, achou-se pertinente extrair os dados do livro em questão. Dessa forma, a autora poderia discutir com mais propriedade as dúvidas surgidas, as opções de tradução, e justificar as escolhas de tradução (tendo em vista a opinião dos revisores, as traduções canonizadas e as coerções da editora). Além disso, verificou-se que abordar a semântica dos compostos nominais técnicos sob uma perspectiva contrastiva seria um ineditismo⁶⁹.

6.2.1 Caracterização do *corpus*

Corpus “é um conjunto de dados linguísticos [...] sistematizados segundo determinados critérios, suficientemente extensos em amplitude e profundidade, de maneira que sejam representativos da totalidade do uso linguístico ou de algum de seus âmbitos” (Sanchez, 1995 *apud* Berber Sardinha, 2000). Os dados devem ser autênticos, ou seja, não devem ter sido produzidos com o propósito de serem analisados, e a pesquisa deve ter por finalidade o estudo linguístico.

A definição de *corpus* atualmente está associada ao uso e à análise do material em computador. No entanto, isso nem sempre foi possível. Embora hoje a maioria dos trabalhos que utilizam *corpora* seja auxiliada por ferramentas e programas de computador, no passado, os *corpora* não eram eletrônicos, mas sim coletados, expandidos e analisados manualmente.

Pode-se dizer que o *corpus* organizado para a análise dos dados relevantes no presente estudo é apenas um recorte de uma obra maior — um livro técnico —; é especializado, na medida em que compreende compostos nominais retirados de uma área específica — engenharia elétrica/eletrônica —, e tem por finalidade auxiliar a classificação das relações semânticas entre os constituintes dos compostos (através da aplicação de uma tipologia utilizada para termos não técnicos e adaptada para o tipo de compostos selecionados). Ademais, através do *corpus*, pretende-se verificar as estruturas utilizadas na tradução para o português, o exame das relações semânticas presentes na construção em português (se foram

⁶⁹ A pesquisa de Lílian Teixeira (2009) aborda a semântica dos compostos nominais identificados em edições da revista National Geographic em inglês e das respectivas edições traduzidas para o português sob uma perspectiva interlingual.

mantidas ou não, se são implícitas ou explícitas, no caso das preposições), e investigar a relação semântica presente em compostos com o mesmo núcleo ou com o mesmo modificador (se há a prevalência da mesma relação na maioria dos casos).

Em *The Acquisition and Interpretation of Complex Nominals* (1995), Johnston, Boguraev e Pustejovsky fazem uma distinção entre *corpus* aberto e *corpus* fechado. A diferenciação entre os dois tipos se refere à estruturação, aquisição e manutenção do *corpus*. O *corpus* aberto pode ser expandido sem que haja a necessidade de verificar a consistência da forma do texto etiquetado. Já no *corpus* fechado, qualquer informação nova deve ser compatível com as informações já fornecidas ao leitor no *corpus* anterior. Ademais, o *corpus* aberto apresenta pouca restrição à inclusão de material. Para o *corpus* de um texto jornalístico, por exemplo, qualquer artigo publicado pelo jornal poderia ser incluído. Já o *corpus* fechado apresenta estrutura e objetivos mais bem definidos e função delimitada (um manual de instrução auxilia os usuários a manusearem corretamente um utensílio, por exemplo). Segundo os autores, os *corpora* de jornais, como o *Wall Street Journal* e o *Financial Times* são abertos; já o *corpus* de um manual é fechado.

É possível afirmar que o *corpus* analisado nesta dissertação é um *corpus* fechado, na medida em que o texto técnico em questão apresenta estrutura e funções bem definidas e em que a expansão do texto técnico exige que o assunto e as informações sejam consistentes com o conteúdo já presente no *corpus* antes de este ser realimentado.

Para este trabalho, foi necessário delimitar ao máximo o *corpus*, uma vez que o *corpus* em potencial, conforme será mostrado mais à frente, era bastante extenso. Sendo assim, optou-se por um *corpus* fechado com possibilidades de futura expansão. É possível expandir o *corpus*, aumentando o seu escopo (utilizando todos os compostos que aparecem no livro, por exemplo), mudando os critérios estabelecidos e que serão discutidos mais adiante etc.

Deve-se ressaltar que os resultados obtidos através da análise do *corpus* elencado aqui não pretendem reproduzir o resultado da coleta de dados em um *corpus* maior (o livro inteiro, por exemplo), nem representam a quantidade e variedade de compostos encontrados no livro; muito menos retratam a linguagem técnica como um todo. O material selecionado para fins de pesquisa teve de ser

bastante delimitado e, pode-se dizer, é apenas uma pequena amostra do que foi encontrado pela tradutora.

6.2.2 Metodologia empregada na delimitação e análise do *corpus*

A escolha do *corpus* deve obedecer a critérios claros, definidos pelos seus criadores, de modo que “corresponda às características que se deseja dele” (Sardinha, 2000:339). Não há um modelo a ser seguido igualmente por todos; o importante é, conforme já mencionado, selecionar um *corpus* autêntico e se ater a critérios pré-estabelecidos. Nesta subseção serão apresentados os critérios estabelecidos para a seleção do *corpus*.

Sardinha (2003:33) observa que:

a tradução busca desenvolver metodologias de análise de corpus próprias, que tendem a distanciá-la da Linguística de Corpus. Além disso, à medida que o uso de corpora se torna mais comum na pesquisa em tradução, começam a surgir trabalhos que propõem metodologias específicas de exploração de corpora para fins tradutológicos.

A presente pesquisa é, ao mesmo tempo, baseada em e conduzida por *corpus*. Os dados servem tanto para confirmar ou refutar uma hipótese ou teoria (neste caso, uma proposta de tipologia), quanto para servir de ponto de partida para a análise das relações semânticas previamente propostas. A escolha das categorias foi baseada em trabalhos anteriores, mas com vistas a dar conta dos compostos nominais técnicos. A aplicação da tipologia foi sendo ajustada de acordo com os dados.

Como será mostrado a seguir, a metodologia utilizada na coleta e análise do *corpus* foi guiada pelas características do material impresso e pelas condições que esse tipo de material permitia. Além disso, a metodologia foi adequada ao objetivo principal do trabalho: analisar as relações semânticas dos compostos nominais técnicos em inglês e o resultado tradutório em português tanto do ponto de vista estrutural quanto do semântico.

O *corpus* analisado nesta dissertação foi retirado de um texto técnico traduzido por mim. O livro *Circuit Analysis: Theory and Practice*, publicado pela antiga editora Thomson Delmar Learning e atual Cengage Learning, encontra-se na quarta edição.

O original apresenta 960 páginas em um único volume. Por questões editoriais, a tradução foi dividida em dois volumes. A tradutora recebeu apenas o original impresso, não tendo acesso a arquivo algum. Como o original é volumoso, seria bastante trabalhosa a transferência do material para mídias eletrônicas.

A princípio, pensou-se em procurar na internet arquivos em pdf, referentes à quarta edição do livro em língua inglesa, que pudessem ser utilizados com o auxílio de aplicativos, de modo a identificar semiautomática ou automaticamente os compostos e facilitar a visualização e análise dos mesmos. Através desses programas, é mais fácil quantificar o número total de compostos; o tipo de composto; o número de compostos que apresentam o mesmo núcleo, ou o mesmo modificador; identificar o contexto, comparar o original com a tradução e, a partir daí fazer generalizações, analisar os tipos de estruturas privilegiadas para determinado composto etc. No entanto, encontrou-se apenas material referente à terceira edição, que era bastante diferente da quarta edição.

Dessa forma, não seria possível utilizar um *corpus* paralelo, em que se trabalha com o texto original e a tradução, muito menos trabalhar com um *corpus* alinhado, de forma que as linhas do primeiro e do segundo sejam correspondentes. De acordo com Teixeira (2009:83), este tipo de *corpus* é bastante utilizado pelos estudos da tradução, “quando é necessário comparar semelhanças e diferenças entre original e traduções ou até mesmo comparar as diferentes traduções de uma mesma obra.”

A partir do banco de dados de um *corpus* paralelo, é possível criar métodos probabilísticos que auxiliem a tradução automática. Nas palavras de Teixeira (2009:83), “[p]or meio de estudos de frequência, é possível construir uma base de dados terminológica de textos especializados”.

Na impossibilidade de utilizar qualquer programa de computador, elencar todos os compostos de uma obra extensa e de construir um *corpus* paralelo, cogitou-se a hipótese de selecionar aleatoriamente alguns capítulos e, através da leitura do livro, extrair manualmente os compostos nominais. Concordou-se que um universo de 100 a 150 compostos seria o ideal para o presente trabalho. Chegou-se à conclusão, entretanto, que a escolha aleatória de alguns capítulos não refletiria a riqueza das construções composicionais presentes no livro, nem cobriria todos os assuntos abordados na obra. Como o livro, dividido em dois

volumes na edição brasileira, foi traduzido pela mesma pessoa, mas teve a revisão técnica de dois profissionais (um revisor diferente para cada volume), seria interessante que houvesse compostos referentes aos dois volumes, uma vez que os resultados tradutórios em português são de suma importância para esta dissertação.

O revisor técnico exerce um papel importante neste tipo de publicação, uma vez que fica responsável pela precisão técnica (tanto dos termos quanto do assunto em si) — algo muito valorizado na área científica e tecnológica. Dessa forma, em se tratando de um tradutor não especialista no assunto, a função do revisor técnico fica ainda mais evidente; na verdade, se torna crucial.

Como a tradução é resultado não apenas da troca de códigos linguísticos de uma língua para outra, mas fundamentalmente de realidades sócio-culturais — que envolvem a cultura receptora, as contingências da língua e cultura de chegada (recomendações, padronizações e práticas diferentes da cultura-fonte), as crenças, normas e coerções dos envolvidos (tradutores, revisores e editora) —, é fundamental que se apresentem exemplos traduzidos pelo mesmo indivíduo, ainda que em momentos diferentes, mas cancelados por duas pessoas, obviamente com crenças e histórias distintas, o que influencia no produto final.

Tendo em vista a necessidade de englobar termos técnicos do livro todo, pensou-se em consultar o índice remissivo. Como o índice remissivo apresenta uma lista dos principais assuntos relatados em uma obra, juntamente com a indicação do número da página, o resultado seria uma seleção de compostos que refletissem os principais assuntos abordados e talvez os compostos de maior ocorrência.

Ademais, a utilização do índice remissivo tornaria mais fácil ater-se ao escopo do trabalho, ou seja, identificar a relação semântica de termos técnicos. A utilização do texto inteiro talvez induzisse a inclusão de compostos nominais não técnicos, fugindo, assim, do foco da pesquisa. Ficaria mais fácil, portanto, identificar o composto no índice remissivo e, através da indicação da página, consultar a expressão no corpo do texto e identificar o contexto em que está inserida.

Como o índice remissivo apresentava muitos termos repetidos, optou-se também por consultar a seção *Key Terms* ‘Termos-chave’ presente no início de cada capítulo, de modo a chegar a um universo de 100 a 150 compostos. Tal seção

continha os principais termos dos capítulos, o que forneceria um panorama dos termos-chave do livro todo.

Após a seleção manual dos compostos, verificou-se a ocorrência de cada um deles no corpo do texto. Muitos compostos apareciam diversas vezes no índice. O procedimento adotado nesses casos foi o de considerar aleatoriamente apenas uma ocorrência dos compostos repetidos e consultar seu contexto.

O *corpus* foi organizado da seguinte maneira: na primeira coluna aparece o termo em inglês, no singular (mesmo que no texto apareça no plural); na segunda, o contexto em que o composto destacado em itálico está inserido, juntamente com a indicação da página — de modo a facilitar a interpretação semântica em alguns casos —; na terceira, a relação semântica subjacente entre os termos; por último, a tradução para o português. Os compostos foram distribuídos em duas tabelas: uma contendo apenas a estrutura N + N e a outra com os múltiplos modificadores nominais. As duas tabelas estão disponíveis nos anexos I e II.

Foram descartadas as construções que apareciam de uma forma no índice ou na seção *Key Terms* e de outra no corpo do texto. Por exemplo, o termo *series capacitor* ‘capacitor série’ ou ‘capacitor em série’ aparecia dessa forma no índice remissivo. No título de uma subseção constava *capacitors in series*, assim como no corpo do texto. É interessante observar que, apesar de os títulos de capítulos, seções e subseções serem extremamente breves e, portanto, um local bastante propício para o uso de compostos, neste caso, os autores do livro optaram por uma construção menos concisa e mais transparente do ponto de vista das relações semânticas.

Há uma grande quantidade de compostos em títulos e legendas de figuras, já que essas seções são curtas e servem para destacar o assunto e aglomerar o maior número de informações no menor espaço possível. Segundo Downing (1975) *apud* Levi (1978:58), “[c]ompounds are ideally suited for use in situation where there is a premium on brevity, yet no appropriate unitary item exists”⁷⁰. Ainda, é possível fazer uma analogia entre a função das manchetes de jornais e os títulos de livros, capítulos, seções etc. De acordo com Quirk *et alii* (1985:1335):

⁷⁰ “os compostos são ideais em situações em que é necessário que se seja breve e em que não se encontra um item lexical unitário”.

[t]he flouting of the given/new principle may be explained not only as a striving for brevity in a headline with limited space but also as ‘premeditated inversion’ of the given/new sequence in order to arouse the reader’s interest in going on to read the fully explicit text of the news item⁷¹.

Sendo assim, como há uma tendência maior em se usar compostos em títulos, descartaram-se as ocorrências em títulos de capítulo, seção, subseção e legendas de figura. Muitas vezes os compostos apareciam apenas no título, e no corpo do texto havia um sintagma com uma pós-modificação (por exemplo, *battery capacity* ‘capacidade da bateria’ constava em um subtítulo, mas no texto aparecia *the capacity of batteries*). Alguns compostos, como *voltage subscripts* ‘subscritos de tensão’, não apareciam no corpo do texto, pois eram parafraseados: *If we wish to subscribe the voltage between two points (say points a and b in a circuit), then we express such a voltage in a subscripted form*⁷²[...] ⁷³ Robbins & Miller (2007:129).

Foram descartados também os compostos cujos modificadores eram nomes próprios. Em termos técnicos, é muito comum homenagear cientistas designando um efeito, um teorema, uma lei etc. com o nome de seu inventor ou principal pesquisador. No entanto, normalmente, para se recuperar a relação semântica entre os constituintes desse tipo de composto, seria necessário explicar todo um contexto. Por exemplo, o termo *Fourier series*, cuja tradução é ‘série de Fourier’, é uma homenagem ao francês John Baptiste Joseph Fourier. Para compreender a relação entre série e Fourier, é necessário saber que Fourier desenvolveu um ramo da matemática usado para expressar qualquer forma periódica como uma série infinita de formas de onda senoidais. Percebe-se, portanto, que tais casos envolvem muito conhecimento “enciclopédico” em oposição ao estritamente “dicionarizado”. Dessa forma, qualquer tentativa de sistematização ou identificação de determinada relação semântica é em vão.

Como não podem ser originados a partir da supressão de um predicado (preposição ou verbo) que expresse a relação semântica entre os constituintes,

⁷¹ “a desobediência ao princípio dado/novo pode ser explicada não apenas como uma necessidade de brevidade em uma manchete com um espaço limitado, mas também como uma ‘inversão premeditada’ da sequência dado/novo, de modo a fazer com que o leitor se interesse em ler integralmente o texto do artigo.”

⁷² Grifo meu.

⁷³ “Quando queremos representar a tensão entre dois pontos, (por exemplo, os pontos *a* e *b* em um circuito), expressamos tal tensão na forma subscrita.”

esses compostos estão fora do escopo do estudo de Levi (1978). Warren (1978), apesar de dedicar um capítulo sobre construções com nomes próprios, reconhece que os compostos formados com substantivos próprios têm a função de identificar ou nomear exclusivamente determinada entidade. Segundo a autora, “[b]ecause of this we think of names as devoid of dictionary meaning. They cannot be used productively. Their reference has to be learnt in each individual case and is not predictable, even though they may be repeated” (1978:218).⁷⁴

O processo de formação das estruturas contendo nomes próprios foge à característica fundamental dos compostos nominais: a recursividade, que possibilita que compostos possam ser expandidos, formando estruturas cada vez maiores. Dessa forma, as estruturas que continham nome próprio como modificador foram excluídas do escopo deste trabalho.

Além disso, foram descartadas estruturas que, apesar de não apresentarem nomes próprios como modificadores, também exigem a recuperação de uma série de informações para se conseguir interpretar a relação entre um constituinte e outro. Para compreender o significado de *valence shell* ‘camada de valência’, é necessário saber que a camada de valência é a mais externa do átomo ou o último nível de distribuição dos elétrons que geralmente participam das ligações químicas por serem os mais externos (Wikipédia). A relação entre os dois constituintes do termo não pode ser expressa por um predicado, por isso exige toda uma contextualização do tipo “enciclopédico”.

Pelo mesmo motivo, foram descartados compostos em que, apesar de o conjunto dos termos representar uma razão ou medida, a relação entre os constituintes só era recuperada através de uma referência externa e que exigia um conhecimento mínimo do assunto. O resultado de *power factor* ‘fator de potência’ é expresso por um valor angular, e o de *quality factor* ‘fator de qualidade’ por uma razão; no entanto, deve-se saber que “fator de potência” se refere à razão entre a potência real e a aparente e que o “fator de qualidade” se refere à razão entre as potências reativa e média. A relação entre os termos é “enciclopédica” e não “predicativa” e nada diz sobre a real referência dos compostos; sendo assim, esse tipo de composto está fora do escopo do presente trabalho.

⁷⁴ “Por isso se pensa que os nomes não apresentam um significado dicionarizado. Não podem ser usados produtivamente. Suas referências devem ser depreendidas individualmente e não são previsíveis, embora possam se repetir.”

As construções que tinham como afixo um genitivo também foram excluídas do *corpus* desta dissertação. A razão da exclusão deve-se à natureza dos compostos. Segundo Quirk *et alii* (1985:1330), os modificadores nominais apresentam uma relação tão forte com o núcleo que a impressão que se tem é que modificador e núcleo formam uma palavra só.

Além disso, Levi (1978) defende que os compostos nominais, ou melhor, os grupos nominais complexos (segundo a nomenclatura adotada pela autora) se comportam sintaticamente como substantivos e não como sintagmas nominais. Leech & Svartvik (1975) *apud* Gregorim (1982:89) relatam que, “embora o caso genitivo seja geralmente discutido como parte da classe dos substantivos, seria melhor, sob vários aspectos, estudá-lo como um tipo de sintagma nominal”. Sendo assim, como os compostos nominais se comportam como substantivos, não podem, diferentemente dos SNs, ser interrompidos pelo morfema genitivo. Logo, a estrutura do tipo (the) *group's recent attacks* ‘recentes ataques do grupo’ seria um exemplo de SN, mas não de composto nominal.

Warren (1978:8) afirma que não há um consenso quanto a classificar estruturas com genitivo como compostos ou não. De acordo com a autora, há linguistas como Bergsten (1911) que consideram as construções com genitivo como compostos fracos, uma vez que são morfologicamente iguais aos grupos sintáticos, os SNs. Ainda segundo a autora, linguistas como Marchand (1969) acreditam que ambas as construções são morfologicamente distintas, uma vez que o -s em compostos seria um sufixo derivacional e não mais um morfema.

Apesar do reconhecimento da falta de critérios claros para estabelecer se construções com genitivo poderiam figurar como compostos, Warren afirma que exemplos como *driver's seat* ‘assento do motorista’ e *sculptor's chisel* ‘cinzel do escultor’ (1978:9), presentes no *corpus* selecionado pela autora, se encaixam, respectivamente, nas classes semânticas de propósito e posse.

Para o propósito deste trabalho, será adotada a noção de extrema coesão entre os constituintes dos compostos e, portanto, a impossibilidade de inserção do sufixo genitivo e de adjetivos predicativos (*universal bias circuit* ‘circuito da polarização universal’, mas não **bias universal circuit*, por exemplo) entre um termo e outro da estrutura, conforme apresentado mais detalhadamente na subseção 2.1.2.

Os compostos resultantes da nominalização do núcleo são bastante profícuos em língua inglesa. Em exemplos do tipo *birth control* ‘controle de natalidade’ e *cell decomposition* ‘decomposição da célula’, o núcleo é um verbo nominalizado e o pré-modificador é o sujeito ou objeto direto do verbo, ou, no caso dos múltiplos modificadores, tanto o sujeito quanto o objeto direto. Embora sejam estruturas extremamente produtivas e interessantes do ponto de vista sintático, elas reproduzem uma relação entre sujeito e verbo e/ou verbo e objeto direto; logo, uma relação oracional. Dessa forma, transformando o núcleo deverbal em um verbo, é fácil depreender se o modificador é o sujeito ou o objeto do verbo. Em *birth control* ter-se-ia, segundo a perspectiva de Levi (1978), uma estrutura profunda do tipo “the act of controlling birth”, ou, seguindo um raciocínio menos elaborado, porém igualmente eficaz, “to control birth”; logo, *birth* seria o objeto direto do verbo. Já em *cell decomposition* ter-se-ia algo como “the act of cells decomposing, ou simplesmente *cells decompose*, em que *cell* seria o sujeito do verbo *decompose*.

De acordo com Grimshaw (1993), a relação entre nomes e verbos vem sendo um importante objeto de investigação desde Chomsky (1970). Segundo a autora, os substantivos “flutuam” entre uma interpretação em que têm uma função argumental e outra em que não têm. Isso ocorre porque os substantivos apresentam uma natureza ambígua, ou seja, não se comportam de maneira uniforme. Para Meyer (1994:64-65),

[d]izer que uma forma nominalizada tem interpretação verbal equivale a dizer que ela está, através do significado da base verbal, estabelecendo uma atribuição para um único elemento, da mesma maneira que o verbo. Dizer que uma forma nominalizada tem interpretação nominal significa dizer que ela está nomeando, referindo elementos do universo de discurso do enunciado em que está inserida.

Isto posto, para o propósito deste trabalho, os compostos cujos núcleos são substantivos deverbais e em que núcleo e modificador preservam a relação de sujeito e verbo e/ou verbo e objeto direto estão, portanto, fora do escopo do presente estudo. Em um primeiro momento, o termo *branch-current analysis* ‘análise de corrente nos ramos’ havia sido selecionado para o *corpus*; porém, como análise e corrente apresentam uma relação de verbo e complemento (análise da corrente – “analisar a corrente”), optou-se por descartar esse tipo de exemplo.

No entanto, os compostos nominais cujos núcleos representem objetos físicos, como *transformers* em *iron-core transformers* ‘transformadores com núcleo de ferro’ e *circuit breaker* ‘disjuntor’⁷⁵ serão mantidos. Em casos como esses, percebe-se que o núcleo deverbal perdeu sua referência verbal e passou a designar um objeto. No exemplo mostrado acima, o substantivo “transformador” é intransitivo, pode ser substituído por outro substantivo e apresenta características sintático-semânticas de um núcleo; logo, os modificadores *iron-core* funcionam como especificadores do núcleo e não como complemento ou sujeito do mesmo.

Muitos trabalhos, como os de Levi (1978), Lauer (1996), Buckeridge & Sucliffe (2002) e Girju *et alii* (2005), destacam que, devido à natureza recursiva dos compostos nominais, esse tipo de estrutura pode ser expandida. Embora a estrutura N + N seja a mais comum, há compostos com múltiplos modificadores. Os compostos com mais de um modificador são facilmente encontrados em textos técnicos, porque há uma necessidade de se aglomerarem conceitos fundamentais em estruturas concisas (Johnston, Boguraev & Pustejovsky, 1995).

Além disso, Quirk *et alii* (1985:1338) afirmam que um manual técnico, escrito para um público bem restrito, digamos para cientistas de pesquisas espaciais, pode apresentar compostos longos e complexos, como *Apollo Block II fuel cell voltage current VI characteristics*. Os mesmos cientistas que consultam o manual em questão jamais utilizariam esse composto enorme, com oito modificadores, em uma conversa informal, por exemplo.

Neste trabalho, optou-se por incluir no *corpus* compostos com múltiplos modificadores, porque, além de serem comuns em textos técnicos, geram ambiguidades estruturais, conforme já mostrado na seção 5.6 e que será retomado no capítulo 8. Sendo assim, a tarefa de analisar as relações semânticas entre os constituintes do composto torna-se ainda mais complicada nos compostos com múltiplos modificadores, porque, antes de estabelecer as relações semânticas, é necessário identificar se há uma modificação à direita, como em [*car* [*radio equipment*]] ‘equipamento de rádio para o carro’, ou à esquerda, como em [[*car radio*] *equipment*] ‘equipamento para o rádio do carro’.

⁷⁵ Disjuntor é um interruptor “que se desliga automaticamente, interrompendo um circuito elétrico, quando a corrente excede o valor máximo ou não alcança o valor mínimo aceitável” (*Dicionário Houaiss*, 2001).

Muitas vezes, a desambiguação se dá apenas no contexto. Segundo Warren, “[n]o interpretation depends solely on linguistic knowledge”⁷⁶ (2003:15). O contexto fornece pistas, principalmente, sobre o assunto em questão, além de limitar as múltiplas interpretações de determinado composto. Por exemplo, para interpretar corretamente os compostos *apple juice seat* e *GM car*, é necessário haver no contexto algum elemento indicando que “a cadeira está situada em frente ao suco de maçã, que está na mesa” [Downing (1977) *apud* Girju *et alii*, 2005] e saber que a GM é “uma montadora de carros”, respectivamente. Quanto mais opaca é a relação entre os constituintes de um composto, mais dependente do contexto ele é.

Dessa forma, em se tratando de um assunto especializado, o contexto é uma boa pista para a compreensão de determinado composto, ainda mais quando não se domina a terminologia.

Segundo a perspectiva da Teoria do Léxico Gerativo, apresentada na seção 4.1, o contexto desempenha um papel fundamental para se interpretar uma palavra ou expressão. As estruturas gerativas utilizadas na construção do significado limitam-no de acordo com um contexto (coerção de significado); limitam o tipo de operação de um item lexical (conexão seletiva) e geram novos significados a partir da colaboração entre as estruturas (cocomposição) (Pinto, 2002:68-69).

Foram excluídos do *corpus* construções em inglês do tipo Adj. + N (Ex.: *standard symbols*), salvo quando se tratava de uma estrutura com múltiplos modificadores, em que também havia pelo menos um substantivo modificador [Ex.: (*universal*) *time constant curves*]. Nesses casos, optou-se por estabelecer a relação semântica apenas entre os constituintes nominais, já que o adjetivo serviria somente como especificador, classificador ou quantificador do composto inteiro, como em *universal bias circuit* ou de um modificador, como em *plastic film capacitor* ‘capacitor de filme plástico’.

No *corpus*, decidiu-se manter a ocorrência do composto coordenado (*rise and fall times*), em que dois substantivos modificam ao mesmo tempo e de forma independente um único núcleo. A tradução também apresenta uma estrutura coordenada (tempos de subida e descida).

⁷⁶ “nenhuma interpretação depende exclusivamente de conhecimento linguístico”.

Os compostos em inglês compilados no *corpus* apresentavam a estrutura N + N ou, no caso dos múltiplos modificadores, N + N + N.... Cumpre observar que também foram analisados como N + N substantivos como *bandwidth* ‘largura de banda’ e *horsepower* ‘cavalo-vapor’, em que se pode depreender claramente a presença de dois substantivos em sua composição. Essa decisão foi tomada dada a variação na grafia dos termos em inglês, como em *honey beehoneybee* ‘abelha produtora de mel’. Em Levi (1978), o composto aparece grafado separado. Já no LONGMAN Online, o mesmo termo é grafado junto. No livro técnico consultado, pode-se comprovar a falta de um critério claro. O composto *bandwidth* é grafado junto e *pulse width* ‘largura de pulso’ separado.

Optou-se por manter os compostos que apresentavam como um dos modificadores um numeral. Os numerais têm uma função quantificadora quando são periféricos (desempenhando o papel de adjunto e não de núcleo); logo, na classificação da relação semântica, eles não modificavam diretamente o núcleo, mas sim quantificavam o modificador, como em *four-wire system* ‘sistema com quatro fios’, *zero-impedance conductor* ‘condutor com impedância nula’ (neste caso, o numeral foi traduzido por um adjetivo em português) e *two-wattmeter method* ‘método dos dois wattímetros’, por exemplo. Além disso, em algumas traduções para o português, o numeral em inglês é traduzido como um prefixo, como em *three-phase circuit*. Na tradução, junta-se o prefixo “tri” ao adjetivo “fásico”, formando a estrutura “circuito trifásico”, ou seja, que tem três fases. Como os numerais quantificam apenas o modificador e não estabelecem uma relação de modificação com o núcleo, os compostos que apresentavam um numeral foram tratados como expressões N + N e não como estruturas com múltiplos modificadores.