

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

**O USO DOS VERBOS MODAIS EM MANUAIS DE AVIAÇÃO EM INGLÊS:
UM ESTUDO BASEADO EM CORPUS**

SIMONE SARMENTO

**PORTO ALEGRE
2008**

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE LETRAS
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS
ÁREA: ESTUDOS DA LINGUAGEM
ESPECIALIDADE: TEORIAS DO TEXTO E DO DISCURSO
LINHA DE PESQUISA: LEXICOGRAFIA E TERMINOLOGIA:
RELAÇÕES TEXTUAIS**

**O USO DOS VERBOS MODAIS EM MANUAIS DE AVIAÇÃO EM INGLÊS: UM
ESTUDO BASEADO EM CORPUS**

SIMONE SARMENTO

ORIENTADORA: PROFa. DRa. MARIA DA GRAÇA KRIEGER

Tese de Doutorado em Teorias do Texto e do Discurso,
apresentada como requisito para a obtenção do título de
Doutor pelo programa de Pós-Graduação em Letras da
Universidade Federal do Rio Grande do Sul.

PORTO ALEGRE

2008

AGRADECIMENTOS

Durante esses anos em que realizei essa tese de doutoramento, muitas pessoas foram importantes para que eu pudesse concluir esta obra.

Antes de agradecer essas pessoas, gostaria de agradecer à PUCRS e ao Conselho Britânico por terem viabilizado a importante etapa de estudo durante um ano na Lancaster University, onde me familiarizei e me apaixonei pela Lingüística de Corpus. Agradeço especialmente à Profa. Dra. Cristina Perna por ter “comprado essa briga” e ter conseguido meu afastamento.

Agradeço especialmente:

Aos professores do Programa de Pós-Graduação em Letras da UFRGS, em especial à Prof. Dra. Maria José Bocorny Finatto pelos valiosos ensinamentos e à Profa. Dra. Anna Maria Becker Maciel pelas importantes sugestões no exame de qualificação.

Aos colegas do Departamento de Letras Estrangeiras da PUCRS pelas palavras amigas e momentos de descontração nas horas mais difíceis desta caminhada. Agradeço especialmente à Ana Bocorny, Ana Ibaños, Aline Pacheco, Cristina Perna, Heloisa Delgado, Regina Buchweitz, Vera Müller e “aos Rossas”. Todos vocês contribuíram muito para o resultado final desta tese.

Às amigas Ana Luiza Freitas e Cristiane Kilian pela leitura cuidadosa e carinhosa.

À Gêssica Gomes, a quem muito admiro, por disponibilizar o guia *Simplified English* e esclarecer dúvidas a respeito da redação de manuais.

Aos meus amigos Ana Flávia e Marcelo, Bira e Maria Tereza, Nenê e Regina, Titi e Pretto, Júlia e Juarez e Rosele por fazerem com que meus filhos não sentissem tanto a minha ausência.

À minha avó Gertrudes, pelo carinho infinito e por me manter sempre alimentada.

Aos meus pais, que sempre com muito amor, me deram a oportunidade de chegar até aqui.

Ao meu querido namorado Caio, por toda a ajuda, por estar do meu lado, por se interessar e reconhecer meu trabalho. Eu não teria conseguido chegar ao fim sem ti.

E em especial aos meus filhos Luiza e Vitor pela paciência em esperar o “fim da tese” que parecia nunca chegar, por entenderem que teriam “menos mãe” durante esse período e pela alegria e felicidade que me proporcionam diariamente.

Enfim, agradeço a todos aqueles que estiveram perto de mim durante esses anos e que, de alguma forma, contribuíram para que eu pudesse realizar este trabalho. Muito obrigada.

“Tudo vale a pena quando a alma
não é pequena”

Fernando Pessoa

RESUMO

Este trabalho trata do uso dos verbos modais em manuais de aviação em inglês sob uma perspectiva da lingüística de corpus. O objetivo é gerar subsídios que permitam elaborar materiais didáticos que reflitam as estruturas lingüísticas como aparecem em seu contexto de uso. Parte-se da hipótese de que o uso dos verbos modais no corpus de estudo apresenta peculiaridades com relação à freqüência de ocorrência, estrutura sintática e sentidos. O corpus de estudo é composto de três manuais técnicos da aeronave BOEING 737, sendo dois manuais de operações destinados a pilotos e um manual de manutenção destinado a mecânicos. A pesquisa é realizada em três etapas. Na primeira etapa é realizado um contraste entre o corpus de estudo, um corpus de referência de inglês geral e um corpus de contraste composto de manuais de sistemas operacionais para computadores pessoais. Esse contraste visa a verificar as diferenças de ocorrências entre os verbos modais e as suas estruturas sintáticas. Depois, são realizadas comparações entre os três manuais que compõem o corpus de estudo também com relação às freqüências dos modais e suas estruturas sintáticas. Na terceira etapa são verificadas as colocações principais de cada VM no manual de manutenção e em um dos manuais de operações. Outro objetivo desta investigação é contrastar o manual de manutenção e um livro didático que tem por objetivo ensinar inglês técnico para mecânicos de aeronaves. Nesse sentido, é feita uma análise do livro no que tange os verbos modais e sugeridas algumas modificações. O trabalho reúne referenciais teóricos da Lingüística de Corpus, Lingüística das Linguagens Especializadas e Verbos Modais. Os resultados do trabalho mostram que há diferenças entre a presença dos modais nos três corpora pesquisados e principalmente entre os três manuais que compõe o corpus de estudo. O estudo mostra a importância de realizar pesquisas baseadas em corpus para revelar as peculiaridades de uma linguagem especializada e assim fornecer subsídios para a elaboração de material didático com propósitos específicos mais de acordo com as necessidades do público alvo.

Palavras-chave: Lingüística de Corpus, Inglês para Aviação, Verbos Modais, Linguagem Especializada

ABSTRACT

This thesis is about the use of modal verbs in aviation manuals in the light of corpus linguistics. The goal is to gather subsides to enable the creation of classroom material which presents the linguistic structures as they appear in their natural context. The hypothesis is that the use of modal verbs in the corpus has peculiarities in terms of frequency, syntactic patterns, and meanings. The aviation corpus is composed of three manuals from the BOEING 737 aircraft: two operations manuals directed to pilots and one maintenance manual directed to mechanics. The research has been carried out in three stages. In the first stage, the aviation corpus is contrasted to a reference corpus and another specialized corpus composed of personal computer operational systems. Such contrast aims at verifying the differences in the number of modal tokens and their syntactic patterns. In the second stage, the three manuals that make up the aviation corpus are compared in terms of modal frequencies and their syntactic patterns. In the third stage, the main collocations of each modal are checked. This thesis also aims at contrasting the maintenance manual and a course book designed to teach aviation English to mechanics. Thus, an analysis of the modal verbs in the course book is carried out and changes to the book are suggested. This thesis brings together theoretical assumptions from Corpus Linguistics, the Linguistics of Specialized languages and Modal Verbs. The results point out that there are differences in the use of modal verbs in the three investigated corpora and mainly among the three manuals that compose the aviation corpus. This study shows the importance of corpus based research to reveal the peculiarities of a specialized language and, therefore, provide input to produce specific pedagogical material which better meets the needs of the target public.

Key words: Corpus Linguistics, Aviation English, Modal Verbs, Specialized Language

SUMÁRIO

1 INTRODUÇÃO.....	11
2 LINGÜÍSTICA DE CORPUS: HISTÓRIA, METODOLOGIA, CAMPOS DE APLICAÇÃO.....	22
2.1 PANORAMA HISTÓRICO.....	22
2.2 CARACTERÍSTICAS DA LINGÜÍSTICA DE CORPUS.....	24
2.3 TIPOS DE CORPORA.....	25
2.4 TERMOS DA LINGÜÍSTICA DE CORPUS.....	29
2.5 ANÁLISE DE CORPORA.....	32
2.6 COLOCAÇÃO, PADRONIZAÇÃO E FRASEOLOGIA.....	34
2.7 VARIAÇÃO E DISTRIBUIÇÃO.....	42
2.8 O DESENVOLVIMENTO DE MATERIAL DIDÁTICO.....	43
2.9 PONTOS DE DIVERGÊNCIA.....	49
2.10 LIMITAÇÕES.....	52
2.11 A OPÇÃO PELA LINGÜÍSTICA DE CORPUS.....	53
3 LINGUAGENS ESPECIALIZADAS: MÚLTIPLAS PERSPECTIVAS.....	55
3.1 TERMINOLOGIA.....	55
3.2 LINGUAGENS ESPECIALIZADAS.....	58
3.2.1 O texto especializado e os gêneros textuais.....	62
3.2.2 A linguagem técnica.....	66
3.2.2.1 <i>Os manuais técnicos: abordagem</i> <i>descritiva.....</i>	67
3.2.2.2 <i>Os manuais técnicos: abordagem</i> <i>prescritiva.....</i>	72
4 OS VERBOS MODAIS.....	79
4.1 MODALIZAÇÃO: DEFINIÇÕES E CLASSIFICAÇÕES.....	79
4.2 OS VERBOS MODAIS EM INGLÊS.....	86
4.2.1 A forma.....	87
4.2.2 O uso.....	88
4.2.3 A problemática da negação.....	104

4.2.4 Os atos de fala	107
4.2.5 Distribuição.....	110
5 OS CORPORA DE ANÁLISE E DE SUA SELEÇÃO.....	118
5.1 CORPUS DE ESTUDO.....	118
5.2 CORPUS DE REFERÊNCIA.....	122
5.3 CORPUS DE CONTRASTE.....	123
6 METODOLOGIA.....	126
7 DESCRIÇÃO E ANÁLISE DOS DADOS.....	145
7.1 PRIMEIRA ETAPA: CORPUS DE ESTUDOS X CORPUS DE REFERÊNCIA X CORPUS DE CONTRASTE.....	145
7.1.1 Contrastes.....	145
7.1.2 Considerações Parciais	154
7.2 SEGUNDA ETAPA: DESCRIÇÃO DO CORPUS DE ESTUDO.....	156
7.2.1 Contrastes.....	156
7.2.2 Considerações Parciais.....	162
7.3 TERCEIRA ETAPA: PADRÕES E REGULARIDADES.....	165
7.3.1 Maintenance Manual (MM).....	167
7.3.1.1 Can.....	167
7.3.1.2 Could.....	172
7.3.1.3 May.....	175
7.3.1.4 Might.....	178
7.3.1.5 Must.....	178
7.3.1.6 Shall.....	182
7.3.1.7 Should.....	183
7.3.1.8 Will.....	186
7.3.1.9 Would.....	190
7.3.2 Operations Manual (OM).....	190
7.3.2.1 Can.....	190
7.3.2.2 Could.....	194
7.3.2.3 May.....	196
7.3.2.4 Must.....	199

7.3.2.5 <i>Shall</i>	202
7.3.2.6 <i>Should</i>	202
7.3.2.7 <i>Will</i>	205
7.3.2.8 <i>Would</i>	209
8 SÍNTESE DESCRITIVA E CONSIDERAÇÕES	210
8.1 SÍNTESE DO MM.....	210
8.1.1 Maintenance Manual e o Simplified English	220
8.1.2 English for Aircraft	226
8.2 SÍNTESE DO OM.....	232
9 CONCLUSÕES	241
10 REFERÊNCIAS	253
ANEXO A- LOG LIKELIHOOD	260
ANEXO B- TAXONOMIA DOS VERBOS	261

1 INTRODUÇÃO

Este trabalho tem origem na minha experiência como professora de inglês para profissionais e alunos da aviação. Minha experiência teve início em 1991 em uma empresa aérea, a VARIG, onde atuei por seis anos lecionando inglês para pilotos, mecânicos e técnicos de aeronaves e pessoal de terra, ou seja, agentes de reserva e *check-in*. Ainda na VARIG, participei da elaboração e criação do Curso de Ciências Aeronáuticas da Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS) em 1993, o primeiro da América Latina. Esta tese é, portanto, mais uma etapa de uma longa trajetória de estudo e reflexão sobre a linguagem usada na área da aviação, a respeito da qual sempre observei uma grande carência de estudos acadêmicos sistematizados sobre as diversas sublinguagens¹ que perfazem o total da “linguagem da aviação”.

No intuito de contribuir para suprir essa lacuna, o presente trabalho se dedica a investigar a incidência e a funcionalidade dos verbos modais no segmento da linguagem da aviação representado pelo texto do tipo “manuais de aeronaves em inglês”. Com base nos resultados da investigação, busco colher subsídios para elaborar materiais didáticos que reflitam a linguagem efetivamente usada nos manuais.

Pesquisas na área das linguagens especializadas (TRIMBLE, 1985) apontam para o fato de que²

os textos [...] e técnicos são escritos baseados em certas suposições (ou pré-suposições) culturais relacionadas ao tipo e quantidade de informações gramático-retóricas que os redatores acreditam que os leitores compartilhem. [...] Nossa pesquisa mostra que grande parte dos falantes (leitores) não-nativos de inglês não possui esse *background* cultural que os habilite a ter essas informações pressupostas de forma a entender os textos técnicos. (ibid, p. 114)

Trimble (1985, p. 115) sugere que os elementos que parecem causar maior dificuldade para esses alunos são (nesta ordem): (1) distinções entre voz ativa x passiva na retórica das descrições e instruções; (2) uso dos verbos modais na retórica das instruções; (3) uso não-

¹ Uma sublinguagem é uma variedade da linguagem usada em uma certa ciência ou tecnologia que não somente é menor do que o total de uma língua, mas é também mais sistemática na sua estrutura e significados. É um subsistema de uma língua que se comporta essencialmente como o total desta língua, ao mesmo tempo em que é limitado a um domínio específico. Cada sublinguagem possui uma gramática e um léxico distinto (KITREDGE, 1982).

² A tradução das citações são de minha autoria.

padrão (e o não uso) do artigo definido na retórica das descrições e instruções; e (4) a escolha do tempo verbal na retórica das descrições. De acordo com os elementos elencados por Trimble, o número dois, verbos modais (doravante VM), é precisamente o foco deste estudo. Investigações anteriores (SARMENTO, 2005a; SARMENTO, 2005b; GABRIELATOS e SARMENTO, 2006) mostram que os modais são freqüentemente usados em orações passivas. Assim, tanto o primeiro, quanto o segundo fator que mais apresentam dificuldades nos textos técnicos, segundo Trimble (1985) são estudados aqui. O autor menciona ainda que os elementos gramaticais que são geralmente atrelados a características retóricas específicas, são suficientemente padronizados de forma a permitir que generalizações sejam feitas. Portanto, como já dito, é através dos VMs, entre outras unidades lexicais, que os redatores de manuais técnicos poderão expressar as diretrizes a respeito do que os leitores³ devem fazer. Dessa forma, é importante conhecer o real funcionamento dessas unidades “*in vivo*” para poder desenvolver materiais didáticos que melhor atinjam as necessidades dos alunos, pois, os materiais didáticos existentes não parecem representar a linguagem presente nesses manuais.

Como esta tese trata, ainda que de modo indireto, do ensino de inglês instrumental, é preciso ressaltar uma das características desse tipo de curso que também me motivou a estudar os textos mais profundamente; diferentemente dos cursos de inglês geral, em que os alunos passam longos anos estudando a língua, os alunos dos cursos específicos buscam resultados rápidos. Esses alunos, muitas vezes já estão no mercado de trabalho e sua necessidade de usar o inglês é iminente. Nesse contexto, o trabalho do professor é constantemente testado. Assim, conhecer com mais precisão as características da linguagem da aviação torna-se fundamental.

A opção pelos estudos dos VMs nos manuais de aviação foi principalmente (mas não unicamente) motivada pelo importante papel que essas unidades apresentam em manuais técnicos. Os textos dos manuais técnicos, acima de tudo, determinam um tipo de ação, expressando obrigações, possibilidades e conseqüências. Os VMs costumam ser usados para expressar essas funções comunicativas. Ou seja, são elementos lexicais fundamentais nesse tipo de texto.

Além disso, os VMs têm recebido minha atenção desde o início da minha experiência na área da aviação. Meus colegas mais experientes, ainda na VARIG, costumavam dizer que o VM *should*, nos manuais, era usado como sinônimo de *must*, uma vez que manuais não são

³ Os leitores desses manuais são mecânicos de aeronave e pilotos.

lugar de sugestões e conselhos (uso mais prototípico de *should*), mas de ordens e instruções claras. Aos poucos notei outro uso de *should* nos mesmos manuais: no início de frases, expressando uma condição. Assim como em grande parte dos estudos e reflexões sobre os VMs, tanto eu, como meus colegas, não conseguíamos perceber a complexidade que envolvia tais unidades. Observei ainda uma incidência maior de um ou outro modal em cada um dos manuais que trabalhávamos. Mas, na época, eram apenas suposições. Daí por que neste trabalho procuro ver o real papel dos VMs nos manuais, seus sentidos, usos e funções.

Apresento agora os objetivos da presente tese.

No intuito de alcançar um melhor conhecimento do segmento da linguagem da aviação caracterizada pelo tipo manual, busco investigar os VMs em três manuais da aeronave BOEING 737, a aeronave mais voada no Brasil: um manual de operações (*Operations Manual*, doravante OM) um guia de referência rápida (*Quick Reference Handbook*, doravante QRH), ambos destinados a pilotos, e um manual de manutenção (*Maintenance Manual*, doravante MM) destinado a mecânicos de aeronaves.

Utilizo a metodologia da lingüística de corpus (LdC) por várias razões: para poder dar conta de todas as ocorrências dos VMs nesse corpus por mim denominado Corpus de Aviação (doravante AC); por acreditar que o fator frequência não é aleatório; e para verificar se diferentes padrões sintáticos podem trazer consigo diferentes sentidos de uma mesma palavra.

Buscando uma compreensão dos manuais de aviação, descrevo a incidência e a funcionalidade dos VMs em manuais de aviação da Boeing. Tento mostrar em que medida a presença maior ou menor de VMs pode contribuir para caracterizar o gênero manual de aviação frente a um corpus de referência de inglês geral (British National Corpus, ou BNC) e a um corpus de contraste de inglês específico composto de dois manuais para operação e instalação de sistemas operacionais de computadores pessoais (Operation Systems Manual, ou OSC). Em um segundo momento, é realizado um contraste entre os manuais que compõem o corpus de estudo para caracterizar cada um deles. Termina a análise com a investigação detalhada de cada VM em cada um dos três subcorpora com o objetivo de oferecer subsídios para futura elaboração de material didático voltado ao ensino de inglês para aviação.

Diante disso, constitui o objetivo geral desta pesquisa:

- Analisar a incidência e a funcionalidade dos verbos modais em um corpus composto de manuais de aviação sendo dois de operação e um de manutenção para gerar subsídios para a elaboração de material didático.

Para alcançar esse objetivo, sigo o seguinte roteiro metodológico:

- i) Contrastar as freqüências de cada VM entre o corpus de estudo (AC), o corpus de referência (BNC) e o corpus de contraste (OSC), observando se há diferenças significativas no uso dos VMs.
- ii) Contrastar as freqüências das estruturas sintáticas em que os VMs se encontram entre o corpus de estudo (AC), o corpus de referência (BNC) e o corpus de contraste (OSC), observando se há diferenças significativas na freqüência de uso das estruturas.
- iii) Verificar as diferenças quanto à freqüência de cada VM e de cada estrutura em cada um dos três manuais que compõe o corpus de estudo.
- iv) Identificar as colocações de cada VM em cada um dos corpora.

Dessa forma, este trabalho pretende oferecer condições para melhor entender o *modus dicendi* próprio dos manuais de aviação, considerando que os textos desses manuais são considerados portadores do conhecimento sobre as aeronaves, e não são passíveis de questionamentos por parte dos destinatários. Possuem um caráter altamente deôntico. A intenção não é fazer um estudo de todos os componentes que integram a linguagem da aviação, mas obter elementos que possam auxiliar na sua descrição e melhor compreensão para assim poder elaborar materiais didáticos para o ensino de ESP (*English for Specific Purposes*)⁴ mais condizente com a realidade dos textos estudados.

Pretendo ainda contribuir para os estudos da linguagem especializada que extrapolam o estudo restrito a termos e ainda para o estudo dos VMs nas linguagens especializadas e para a elaboração de materiais didáticos lingüísticos com finalidades específicas.

Com base nos objetivos acima, a questão geral de pesquisa do presente trabalho é:

⁴ Inglês com Propósitos Específicos ou, ainda, Inglês Instrumental.

Qual é a incidência e a funcionalidade dos verbos modais no segmento da linguagem da aviação representado pelo texto do tipo manual técnico operacional e de manutenção?

A observação prévia dos textos levou-me a formular a hipótese que norteia este trabalho em relação ao uso dos VMs:

- O uso dos VMs é peculiar no corpus de estudo, tanto em relação à frequência de uso e distribuição quanto em relação aos sentidos e funções.

Para dar conta dos objetivos, reúno as seguintes perspectivas teóricas: a) os estudos lingüísticos sobre os VMs em inglês; b) a Lingüística de Corpus (LdC); e c) a Lingüística das Linguagens Especializadas. Esses referenciais serão brevemente introduzidos nesta seção e serão aprofundados nos capítulos, 2, 3, e 4.

A modalização é um assunto bastante complexo. Entre as várias perspectivas que tratam sobre a modalização, priorizo os estudos sobre o uso, principalmente os relacionados à LdC. Com relação à modalização, Oliveira (2003) e Coates (2005) sugerem que numa língua existem formas diversas de expressar um mesmo tipo de modalidade. Em contrapartida, uma mesma forma pode expressar diferentes modalidades, ou seja, os conceitos modais podem ser expressos nas línguas naturais através de uma grande variedade de formas.

Faz-se necessário distinguir dois tipos de modalidade: a epistêmica e a raiz⁵. Para Coates (1995), a modalidade epistêmica diz respeito às suposições do falante ou à avaliação de possibilidades, e, na maioria dos casos, indica a segurança, ou falta de segurança, que um falante possui na verdade da proposição expressa. A modalidade raiz, por sua vez, engloba significados como permissão e obrigação, assim como possibilidade e necessidade. Ou seja, a modalidade raiz é muitas vezes utilizada nos atos de fala (SEARLE, 1969): convidar, pedir, dar uma ordem, etc.

Na língua inglesa, a modalização, tanto a epistêmica, quanto a raiz, pode ser expressa através de verbos como *know*, *believe*, *allow*, *have to*, *etc*; advérbios tais como *possibly*, *necessarily*, *probably*; adjetivos como *possible*, *probable*, *capable*, a forma imperativa, entre outras. Há também alguns tempos gramaticais que podem criar situações alternativas, como é o caso do futuro e do condicional. Entretanto, a modalização, especialmente na linguagem

⁵ São várias as definições de modalidades oferecidas por diversos autores, entretanto *epistêmica* e *raiz* parecem ser os tipos mais aceitos pelos estudiosos do assunto.

escrita, acontece “principalmente através dos auxiliares modais, como *may, can, shall, must*” (CRYSTAL, 1985, p. 174).

Os VMs em inglês, foco prioritário desta tese, formam uma subclasse dos verbos, com características sintáticas, semânticas e pragmáticas próprias. Para melhor definir a classe dos VMs, ou entender os ambientes nos quais um VM é usado, deve-se lançar mão de vários níveis lingüísticos; (i) o puramente sintático; (ii) o semântico; (iii) o pragmático, ou seja, a função do VM, as suposições compartilhadas entre destinador e destinatário, a situação social adotada pelos participantes no discurso, a impressão que o falante quer causar no seu interlocutor; (v) o fator distribucional, entre outros. Para R. Lakoff (1973), a estrutura dos modais é bem mais complexa do que tem sido considerada. Por serem complexos e por serem considerados ambíguos⁶, os VMs estão freqüentemente presentes nas discussões sobre polissemia, havendo, conforme Oliveira (2003) movimentações sobre dois eixos. Sobre um eixo, um único item lexical, como por exemplo, *could*, pode expressar habilidade, permissão, assim como possibilidade. Por outro lado, esses mesmos significados poderão ser expressos por outros VMs, ou seja, a idéia expressa por *could* poderá por vezes ser transmitida por *might, can may*, entre outros.

A diferença de uso entre um modal ou outro pode ser bastante tênue, ou, por vezes, invisível ou até mesmo (quase) inexistente, como é possível notar nos exemplos abaixo retirados de diferentes *websites* (utilizando o mecanismo de busca Google). O sintagma escolhido para busca foi “*test result _____ occur*”:

- (1) *One needs to be aware, however, that a false-positive **test result can occur**.*
- (2) *...a negative **test result could occur** in an exposed, infected kitten that ...*
- (3) *A negative **test result might occur** if the level of extracted antigen in a sample is below the sensitivity of the test.*
- (4) *A false positive **test result may occur** if the urine is contaminated with...*

Os contextos dos exemplos acima parecem muito semelhantes, sendo difícil visualizar limites claros entre os usos dos diferentes VMs. Contudo, nessa breve busca, nota-se algumas diferenças entre os quatro sintagmas: a busca com o modal *can*, resultou em 1450 ocorrências, o *may*, 2290, o *could*, 160 e, por fim, o *might*, apenas 85. Ou seja, há uma grande diferença na

⁶ Ambigüidade é a propriedade de certas frases realizadas que apresentam vários sentidos. A ambigüidade pode ser do léxico, quando certos morfemas léxicos têm vários sentidos (DUBOIS et.al., 1978, p. 45)

freqüência de uso de cada VM nesse sintagma. O próximo passo seria observar se um dos VMs é mais ou menos freqüente em algum determinado gênero textual, ou conforme o tipo de sujeito, ou ainda, se existe uma força ilocucionária⁷ diferente no uso de cada um dos quatro modais. Portanto, é necessário que se leve em consideração diversos aspectos inerentes aos VMs para que se possa estabelecer uma tipologia de uso e assim oferecer subsídios para a melhor elaboração de materiais didáticos com propósitos específicos. Devido a essa complexidade inerente aos VMs, acredito que eles podem contribuir para uma dificuldade de compreensão por parte dos leitores dos manuais, especialmente para os leitores que não possuem uma boa proficiência na língua inglesa.

O fator distribucional⁸ recebeu mais atenção nos últimos anos com o avanço das pesquisas realizadas com o auxílio de computadores, isto é, com a popularização da lingüística de corpus (doravante LdC). Vários pesquisadores têm investigado (COATES, 1983; BIBER et al., 1999; KENNEDY, 2002; GABRIELATOS e SARMENTO, 2006; DERNARDIN, 2008), por exemplo, como os VMs se distribuem em diferentes corpora e quais as estruturas sintáticas em que se encontram. Biber et al. (1999), conduziram a mais abrangente pesquisa baseada em corpus sobre a língua inglesa até o momento, que deu origem ao Longman Grammar of Spoken and Written English (LSWE⁹). Os autores encontraram uma grande variação na freqüência de uso dos VMs dependendo do ambiente: conversação, ficção, texto jornalístico, ou texto acadêmico; e da variante nacional: inglês americano (IA) ou inglês britânico (IB). Similarmente, Kennedy (2002), ao analisar a distribuição dos VMs no BNC¹⁰ encontrou uma grande variação no uso dos VMs. Ambos os estudos concluíram que a distribuição dos VMs varia largamente dependendo do tipo de texto ou registro. Essas pesquisas estão em consonância com os postulados de Hoffmann (2004, p. 80) que afirma que a especificidade de uma linguagem especializada é expressa, entre outros fatores, pela “freqüência de determinadas manifestações lingüísticas”.

⁷ Força ilocucionária é um termo proposto por Austin (1962) e indica como uma locução (sentença) deve ser entendida, ou seja, como uma pergunta, uma asserção, uma promessa, um pedido de desculpas, etc. É uma questão relevante para aspectos relacionados à educação e cultura, pois é necessário observar que ao ler um texto, recai sobre o leitor o ônus de compreender as intenções do redator por meios exclusivamente lexicais.

⁸ Por fator distribucional refiro-me a quais modais são utilizados em quais tipo de texto e com qual freqüência.

⁹ O LWSE foi baseado em um corpus de 40 milhões de palavras, escrito ou oral, de IA ou IB.

¹⁰ O BNC (British National Corpus) contém aproximadamente 100 milhões de palavras, de textos oral e escritos, de IB.

Kennedy (2002) analisou ainda as estruturas sintáticas em que os VMs se encontram. Há, a exemplo da variação de distribuição dos VMs por gênero textual¹¹, uma diferente preferência por estruturas específicas, dependendo dos VMs e dos gêneros textuais em que eles se encontram. Ou seja, esses ambientes sintáticos dos VMs, são mais ou menos freqüentes nos diferentes gêneros textuais. Os manuais de aviação constituem um gênero textual específico. Dessa forma, acredito ser importante investigar a freqüência dos VMs nesses manuais, assim como seus padrões sintáticos, uma vez que, em consonância com os pressupostos da LdC, a freqüência de itens lexicais não é considerada aleatória, e sim motivada por necessidades comunicativas entre outros fatores.

Além do fator freqüência, a LdC é também responsável pela investigação de aspectos lingüísticos que costumavam ser desconsiderados ou receber pouca atenção, como por exemplo o estudo das colocações, entendidas como a tendência com que certas palavras co-ocorrem com outras e da padronização, definida como a regularidade expressa na recorrência sistemática de unidades co-ocorrentes de várias ordens, lexical, gramatical, sintática e pragmática. As listas de colocações podem ressaltar os diferentes significados de uma palavra, ou seja, os diferentes significados de uma palavra tenderão a ter diferentes colocados. Para que sejam definidos os padrões de uma palavra, faz-se necessário averiguar as palavras e as estruturas freqüentemente associadas a ela que de alguma forma refletem no seu significado. Desconheço estudos sobre a linguagem da aviação que tenham investigado esses fatores.

É importante ressaltar que pilotos e mecânicos ao redor do mundo, independente das suas línguas maternas, necessitam ler e entender em detalhe as informações contidas nos manuais de aeronaves, que são obrigatoriamente escritos na língua inglesa. Traduções não são aceitas pelos fabricantes, como a Boeing e a Airbus. Além disso, diferentemente de outros manuais, como por exemplo, um manual de computador, um erro na interpretação dos manuais pode ter conseqüências graves e muitas vezes fatais, como já relatado no boletim da Organização da Aviação Civil Internacional¹² (2004). Desse modo, há uma intensa e legítima preocupação por parte dos fabricantes de aeronaves em redigir esses manuais da forma mais clara e unívoca possível. Essa preocupação é traduzida, dentre outras formas, através de comitês organizados para elaborar normas que guiem a redação desses manuais, como o

¹¹ Os gêneros são classes de eventos comunicativos que possuem características de estabilidade, reconhecimento de termos, etc. Para Swales (1990, p.46) eventos comunicativos compreendem “não somente o próprio discurso e seus participantes, mas também o papel do discurso e seu ambiente de produção e recepção” (Swales, 1990, p. 46).

¹² OACI, ou International Civil Aviation Organization (ICAO)

comitê do *Simplified English* (SE)¹³, que elabora regras sobre como redigir a documentação de manutenção de aeronaves.

Para esse comitê, existe uma dificuldade por parte dos leitores dos manuais em entendê-los corretamente. Há também uma opinião geral de que os manuais não são escritos de forma clara e concisa, pois, como já mencionado anteriormente, há uma notável carência de pesquisas lingüísticas que tratam da descrição das diversas sublinguagens da aviação. Assim, apesar de haver essa preocupação em melhorar a redação dos manuais, as decisões a respeito da padronização dessa linguagem são tomadas com base nas opiniões e impressões dos membros dos comitês especializados, geralmente sem respaldo acadêmico ou científico, e de uma forma lingüisticamente ingênua.

Outra conseqüência dessa carência de pesquisas é a proporcional carência de materiais didáticos voltados ao ensino de ESP para pilotos e mecânicos. Até agosto de 2008, havia apenas dois livros de qualidade reconhecida pelos profissionais que atuam nessa área: *English for Aircraft Maintenance* (SHAWCROSS, 1992) e *Airpeak* (ROBERTSON, 1988). Mais recentemente, outros dois livros foram lançados¹⁴, o *English for Aviation* (ELLIS e GERIGHTY, 2008) e *Aviation English* (EMERY e ROBERTS, 2008). O *English for Aircraft Maintenance* é relacionado com esta pesquisa, pois trata do ensino de vocabulário e estruturas presentes em manuais de manutenção de aeronaves, voltado a mecânicos. Os outros três livros têm o foco no ensino da comunicação oral, mais especificamente entre pilotos e controladores de tráfego aéreo.

Esses fatores e percepções subsidiaram alguns estudos preliminares sobre o assunto (SARMENTO, 2005a; SARMENTO, 2005b; GABRIELATOS e SARMENTO, 2006). Esses estudos, dentro da abordagem da LdC, seguiram um enfoque quantitativo, mas já apontam uma grande variação na freqüência de uso (por VM e estrutura sintática) entre os diferentes

¹³ O *Simplified English* é um conjunto de regras sintáticas e semânticas que tem por objetivo oferecer diretrizes para que os redatores de manuais técnicos de aviação usem uma linguagem simplificada e padronizada. Entre as regras estabelecidas pelo *Simplified English*, encontram-se normas explícitas sobre quais VM usar e em quais estruturas sintáticas. Esse assunto será debatido no capítulo 3 desta tese. No manual da ICAO os VM estão incluídos entre as estruturas lingüísticas importantes de serem ensinadas aos pilotos.

¹⁴ O lançamento desses livros foi motivado por uma determinação da ICAO. Desde março de 2008 é requerida uma certificação em proficiência na língua inglesa para pilotos que realizam vôos internacionais e para controladores de vôos que trabalham em aeroportos internacionais. No Brasil cabe à Agência Nacional de Aviação Civil (ANAC), como autoridade de aviação civil, a responsabilidade de avaliar e certificar pilotos de aviões e helicópteros. Para que tais pilotos comprovem proficiência lingüística, deverão se submeter a um teste de inglês, o *Santos Dumont English Assessment*.

manuais de aviação, quanto entre o Corpus de Aviação¹⁵ (AC) e outros corpora de inglês geral.

Com relação aos modais, verifiquei (SARMENTO, 2007) o tratamento concedido ao VM *can* em dois dicionários monolíngües para aprendizes avançados de inglês. Apesar de os dois dicionários serem baseados em grandes corpora, há uma notável discrepância em relação à ordem das subentradas, ao número e ao tipo de definições. Essa discrepância pode ter origem na própria palavra *can* e sua complexidade, o que, mais uma vez, mostra a necessidade de um melhor entendimento dessas unidades lingüísticas.

A lacuna de estudos descritivos sobre a linguagem da aviação, o caráter complexo dos VMs, a falta de materiais didáticos na área, assim como os resultados preliminares dos outros estudos que realizei me conduziram a esta tese. Por isso, pretendo que esta pesquisa oportunize o preenchimento de uma lacuna no que diz respeito ao melhor conhecimento dos VMs, à descrição dos manuais de aviação, e a conseqüente qualificação da minha prática docente.

Para alcançar os objetivos acima, foi adotada uma abordagem que permite a descrição e a explicação do uso dos VMs nos três manuais que compõem o corpus de estudo. Nesse sentido, além desse capítulo introdutório, a tese é constituída de mais oito capítulos que descrevo a seguir.

No segundo capítulo, *Lingüística de Corpus*, delimito a área, abordo termos importantes, aponto diferentes perspectivas, apresento seus campos de aplicações e examino a contribuição da LdC para as pesquisas lingüísticas e para a produção de material didático com propósitos específicos.

No terceiro capítulo, intitulado *Linguagens especializadas: múltiplas perspectivas*, discuto vários aspectos e teorias relacionadas ao estudos das Linguagens Especializadas, entre eles, Texto Especializado e o conceito de Gêneros Textuais. Nas duas últimas seções trago algumas considerações sobre a linguagem dos manuais técnicos e sobre a linguagem da aviação.

¹⁵ Nesses trabalhos anteriores usei o mesmo nome para o corpus de estudo que estou utilizando na presente tese. Entretanto, nesses estudos preliminares o corpus utilizado era bem menor do que o utilizado atualmente, em torno de 120 mil palavras, comparado com mais de 400 mil palavras agora, devido a dificuldade de acesso a esse tipo de texto.

O quarto capítulo, *Os verbos modais*, dedica-se ao estudo das definições, classificações e características da área. São abordados diferentes pontos de vista, incluindo a modalização na linguagem especializada e sob uma perspectiva da LdC.

No quinto capítulo, *Os Corpora de Análise e sua Seleção*, justifico o motivo que me levou a escolher essa área (aviação) e esse gênero textual (manuais). Em um segundo momento, descrevo o corpus de estudo (AC), o corpus de referência (BNC) e o corpus de contraste (OSC).

No sexto capítulo, *Metodologia*, apresento os procedimentos de análise adotados e as etapas percorridas. Mostro assim os procedimentos para coleta e observação dos dados, os métodos utilizados para fazer as observações contrastivas e os procedimentos realizados para investigar a padronização de cada um dos Verbos Modais.

No sétimo capítulo, *Descrição e Análise dos Dados*, mostro os dados encontrados no contraste dos três corpora, assim como no contraste entre os três manuais (dois manuais para pilotos e um para mecânicos) que compõem o corpus de estudo. Apresento também os dados encontrados no levantamento da lista de colocados de cada um dos VMs e as funções comunicativas predominantes de cada um dos itens lexicais investigados. São elaboradas tabelas detalhadas com os resultados encontrados para cada uma das etapas. Ainda destaco os aspectos mais significativos dessa investigação.

No oitavo capítulo, *Síntese Descritiva e Considerações*, discuto os resultados revelados no capítulo sete e comparo o uso dos VMs no MM com o guia *Simplified English* e *O English for Aircraft*. Por fim, no último capítulo, *Conclusões*, faço uma retomada dos aspectos mais relevantes do trabalho, apresento algumas perspectivas de investigações futuras e esboço algumas sugestões para a elaboração de material didático e redação de manuais de aviação, baseadas nos resultados obtidos.

2 LINGÜÍSTICA DE CORPUS: HISTÓRIA, METODOLOGIA, CAMPOS DE APLICAÇÃO

Neste capítulo pretendo delimitar e caracterizar a Lingüística de Corpus (doravante LdC); definir termos importantes da LdC; apontar as diferentes perspectivas da LdC; apresentar seus campos de aplicações; mostrar suas limitações; e, acima de tudo, examinar a contribuição da LdC para as pesquisas lingüísticas, em especial para o presente trabalho e para a produção de material didático.

2.1 PANORAMA HISTÓRICO

Desde os anos 60, os corpora eletrônicos têm sido considerados um valioso recurso para o estudo lingüístico. Apesar de o seu uso ainda ser motivo de controvérsia (HUNSTON, 2002; McENERY e GABRIELATOS, 2006), sua contribuição ao ensino de línguas assim como à lingüística é amplamente reconhecida. Essa contribuição foi primeiramente sentida na lingüística inglesa devido ao trabalho pioneiro de corpora de língua inglesa, como o Brown Corpus (FRANCIS e KUCERA, 1964) que deu origem à grande parte dos trabalhos produzidos utilizando corpora desde então.

O Brown Corpus foi o primeiro corpus computadorizado compilado para fins de pesquisa lingüística. Entretanto, sua importância torna-se ainda maior se levarmos em consideração o fato de essa compilação ter acontecido em um momento em que o paradigma cada vez mais dominante, liderado por Noam Chomsky, era totalmente contrário ao registro e à pesquisa do desempenho lingüístico. Para os lingüistas gerativo-transformacionais, o estudo da língua deve descrever não o que os falantes fazem com a língua (desempenho), mas o que os falantes sabem sobre uma língua (competência). Ou seja, fundamenta-se no estudo da linguagem através da introspecção para a verificação dos modelos de funcionamento da linguagem. Os dados estão na mente do lingüista, e acessíveis através da introspecção. As teorias são verificadas através de frases inventadas, muitas vezes pelo próprio pesquisador. Esse paradigma racionalista perdurou como predominante por um longo período, distanciando grande parte das pesquisas lingüísticas dos estudos descritivos de desempenho. Chomsky, entre outros, discordava do uso de corpora e dos modelos de competência probabilísticos baseados em estatística, derivados do estudo do desempenho lingüístico. Kennedy (1998) relata que em 1958 Chomsky teria supostamente argumentado a respeito da inadequação do uso de corpora para a descrição de regras gramaticais ao dizer que “algumas frases não acontecem por serem óbvias demais, outras por serem falsas, e outras por serem indelicadas”

(ibid, p. 18). Cinco anos depois, Chomsky teria trivializado a relevância de análises estatísticas de frequência de ocorrência de itens lingüísticos ao dizer que, apesar de *I live in New York*¹⁶ ser mais freqüente do que *I live in Dayton, Ohio*, esse fato não teria qualquer relevância para a descrição ou teorias lingüísticas (KENNEDY, 1998; BERBER SARDINHA, 2004). Foi então, nesse ambiente acadêmico bastante hostil ao uso de corpora, que Nelson Francis e Henry Kucera iniciaram o que deve ter sido a enorme tarefa de compilar um corpus sincrônico de aproximadamente um milhão de palavras representativas do inglês escrito publicado nos Estados Unidos em 1961. O trabalho foi finalizado em 1964 com velocidade surpreendente tendo em vista os recursos da época. O Brown Corpus foi então disponibilizado em fita de computador acompanhado do respectivo manual para o usuário.

Nascia mais uma dicotomia lingüística¹⁷: de um lado os lingüistas neo-firthianos, britânicos na sua maioria, que tratam os corpora como repositórios de instâncias do uso real da língua, nos quais os exemplos que se encaixam, ou não, em uma teoria, ou que sustentam um ponto de vista em uma discussão devem ser selecionados querendo o pesquisador ou não (McENERY e GABRIELATOS, 2006). Do outro lado, temos os chomskianos americanos, que Fillmore (1992) denominou de “lingüista de poltrona”¹⁸, ou seja, lingüistas que buscam na intuição e nos exemplos introspectivos (em grande parte inventados pelos próprios pesquisadores) suas únicas fontes de dados. Eles sentam em uma poltrona confortável esperando por algum *insight* lingüístico interessante. Ocasionalmente, o lingüista de poltrona tomaria nota desses *insights* que o levaram mais próximo de entender como a linguagem funciona (BERBER SARDINHA, 2000). Para Fillmore, ao se encontrarem, o lingüista de poltrona pergunta ao de corpus “por que eu deveria acreditar que o que você me diz é interessante?”, enquanto o de corpus pergunta “por que eu deveria acreditar que o que você me diz é verdadeiro?”. Entretanto, Fillmore acredita que os lingüistas deveriam fazer uso dos dois paradigmas, apesar das dificuldades, uma vez que ambas as teorias têm muito a contribuir para a área.

¹⁶ Berber Sardinha (2004, p. 132) afirma que *I live in New York* é apenas supostamente mais freqüente, uma vez que não se sabe se os falantes dessa cidade usam efetivamente esta frase, em quais ocasiões e em que contextos.

¹⁷ São várias as dicotomias conhecidas na área da lingüística, grande parte delas originadas por Ferdinand de Saussure, para citar algumas: Langue X Parole, Sincronia X Diacronia, Sintagma X Paradigma, Significante X Significado e Léxico X Gramática.

¹⁸ *Armchair linguist*.

2.2 CARACTERÍSTICAS DA LINGÜÍSTICA DE CORPUS

O termo “lingüística de corpus” (LdC) é entendido (MCENERY e WILSON, 1996) como o estudo da linguagem baseado em exemplos da vida real. A LdC não é um ramo da lingüística como a sintaxe, a semântica ou a pragmática, que concentram-se na descrição ou explicação de algum aspecto da língua em uso (RAYSON, 2002). A LdC é uma metodologia¹⁹ que pode ser aplicada a uma grande variedade de estudos lingüísticos, ou ainda ao ensino de línguas, ou seja, é uma das várias maneiras de fazer lingüística.

Biber, Conrad e Reppen (1998. p. 4) listam as características essenciais da lingüística baseada em corpus:

- É empírica, ou seja, analisa os padrões reais de uso em textos naturais;
- Utiliza uma grande coletânea de textos²⁰ (um corpus, com princípios de coleta pré-estabelecidos) como base para análise;
- Faz um extenso uso de computadores para análise, podendo também utilizar técnicas automáticas e interativas;
- Depende de técnicas analíticas quantitativas e qualitativas.

Os corpora são usados para gerar conhecimento empírico sobre uma língua, que pode complementar, ou muitas vezes suplantam, informações provenientes de fontes de referência e introspecção (LEECH, 1991, 1992). Stubbs (2001) acrescenta ainda que a LdC vê a linguagem como sendo um sistema probabilístico, ou seja, embora muitas combinações e características lingüísticas sejam possíveis, nem todas são prováveis de ocorrer. Dessa forma, por ser uma técnica adequada à análise estatística, os corpora podem fornecer informações sobre a frequência relativa de muitos aspectos da língua. Berber Sardinha afirma que “o mais importante da diferença de frequências entre traços é não serem aleatórios” (2004, p. 31). Se essas diferenças fossem aleatórias, o fator frequência não seria significativo e não adicionaria informações a respeito da estrutura da língua. Contudo, grupos de características lingüísticas apresentam uma variação sistemática em textos específicos, variações oriundas de situações comunicativas específicas. A variação sistemática, ou seja, a recorrência de traços lingüísticos (colocação, coligação, padrão sintático, entre outros) indica que a linguagem é padronizada (*patterned*) e motivada por diversos fatores além das necessidades comunicativas. Por

¹⁹ Existe certo debate a cerca do real status da LdC que será discutido na seção 2.10.

²⁰ Berber Sardinha (2004, p. 17) salienta que no lugar de “textos” a expressão “porções de linguagem” parece mais adequada devido aos problemas relacionados à delimitação do conceito de texto. Para o autor, na LdC, pode-se considerar um artigo científico, seu resumo inicial ou ainda um trecho de um diálogo como textos.

exemplo, ao escolher o determinante *the* as escolhas das palavras subseqüentes são automaticamente limitadas, isto é, adjetivo, advérbio ou substantivo. Outros fatores influenciam a seleção de palavras, tais como, a proficiência lingüística do autor, colocações, tópico, tipo de texto, e, no caso deste trabalho, normas de redação para os manuais de aviação. Os padrões apresentam regularidades e variações sistemáticas em variedades textuais, dialetais, etc. A verificação dessas regularidades não pode ser alcançada através da intuição de um falante nativo. Somente a observação empírica de dados reais, em diferentes contextos de uso pode fornecer essa informação. Cabe assim dizer, que a freqüência de ocorrência de traços lingüísticos, não constitui uma constatação trivial, como havia afirmado Chomsky.

2.3 TIPOS DE CORPORA

Um corpus pode ser definido como uma “coletânea de exemplos naturais de linguagem, que consistem desde algumas frases até conjuntos de textos escritos ou gravações orais que foram coletados para serem usados como base para pesquisa lingüística” (HUNSTON, 2002, p. 2). Mais recentemente, a palavra “corpus” (cujo termo mais comumente aceito para o plural é corpora) tem sido usada para referir-se a coletâneas de textos (ou partes de textos) que são armazenadas e que podem ser acessadas por meio de computadores. Textos escritos retirados de jornais ou revistas podem ser escaneados, retirados de um CD ou da internet. Textos orais, como conversas, são gravados e posteriormente transcritos; ou seja, são copiados palavra por palavra de forma que os textos dessas conversas possam ser alimentados em um computador. Torna-se assim possível analisar a língua contida no corpus através de softwares específicos para o estudo lingüístico, como por exemplo, o *Wordsmith Tools*, utilizado neste trabalho.

Entende-se por “exemplos naturais de linguagem” (conforme citação anterior) aqueles exemplos que não tenham sido produzidos, ou criados, para serem utilizados em um corpus. Berber Sardinha (2004) acrescenta que a idéia de natural inclui também o fato de a linguagem ser produzida por humanos, excluindo, assim, programas de geração de textos. Entretanto, apesar de os textos serem naturais, um corpus é um objeto artificial, pois foi criado com a finalidade específica da pesquisa.

Não há uma especificação do tipo de conteúdo que um corpus deveria conter. Um corpus pode conter desde a obra completa de Shakespeare, até instruções expressas nas caixas de sabão em pó, ou textos jornalísticos sobre o Grêmio Football Porto-Alegrense no ano em que (quase) foi campeão brasileiro. Com relação à dimensão, não há um consenso quanto ao

tamanho mínimo ou máximo aceito para um corpus. Segundo Hoffmann (1998[2007]), as interpretações sobre o tamanho mínimo necessário para os corpora nas pesquisas linguísticas divergem amplamente. Na pesquisa de linguagem especializada, já foram obtidos resultados úteis com amostras de 35.000 palavras, mas sugere a dimensão de 200.000 palavras como mínima. O autor defende que o tamanho dependerá dos objetivos da pesquisa e do tipo de corpus. Desta forma, “não há nenhuma fórmula matemática amplamente aceita que informe a quantidade ou distribuição de palavras ou textos que um corpus deva ter para ser representativo” (BERBER SARDINHA, 2000, p. 104). Entretanto, a maior parte das palavras tem frequência de ocorrência muito baixa e para que elas apareçam em um corpus é necessário que ele possua um grande número de palavras. O mesmo pode ser dito com relação aos diferentes sentidos ou significados de uma mesma palavra: há os mais e os menos frequentes. Os sentidos mais raros terão uma maior probabilidade de aparecer em um corpus maior.

A questão da representatividade envolve ainda conhecer o “todo” que, no caso da linguagem, não é conhecido. Deve-se tentar dividir esse todo estimado em partes. Por exemplo, um corpus de “linguagem jornalística” deve incluir diferentes tipos de jornais, os populares e os mais tradicionais, por exemplo. Deve também incluir textos das diferentes seções, como variedades, esportes, editoriais, negócios, entre outras. Para ser considerado representativo e equilibrado, um corpus desse tipo deveria incluir um número aproximado de palavras em cada categoria: negócios nos populares, negócios nos tradicionais, esporte nos populares, esporte nos tradicionais, etc.

Conforme Hunston (2002), outro detalhe relacionado ao tamanho de um corpus é a velocidade e a eficiência do software de acesso a esse corpus, assim como a capacidade do computador de acessá-lo. Se, por exemplo, obter a listagem das formas do presente e passado do verbo *to be* levar mais de alguns minutos, o pesquisador pode preferir utilizar um corpus menor, cujo resultado pode ser considerado tão confiável quanto o de um corpus maior, mas para o qual o software trabalhará muito mais rapidamente.

O objetivo da pesquisa também influencia o tamanho que um corpus necessita ter. Carter e McCarthy (1995, p. 143) afirmam que para estudar gramática na linguagem falada, um corpus relativamente pequeno pode ser suficiente, pois as palavras gramaticais tendem a ser muito frequentes. Por outro lado, itens de baixa frequência necessitam de um corpus bem maior.

Os corpora são geralmente coletados com base em um projeto de pesquisa lingüística específico em mente, tal como fornecer informações sobre frequências para verbetes de dicionários, ou produzir material didático para o ensino de língua estrangeira, um dos propósitos deste trabalho. Algumas vezes, contudo, os corpora são coletados sem um propósito específico e são disponibilizados como um recurso da língua geral para lingüistas, professores de línguas, lexicógrafos, entre outros.

Há vários tipos de corpora dependendo do tamanho, propósito e forma como foram compilados. Sinclair (1995) sugere a seguinte lista de tipos de corpora:

-Corpus Geral- Um corpus contendo muitos tipos de texto. Pode incluir linguagem escrita, falada ou ambas; textos produzidos em um país ou vários. Por ser de cunho geral, muito provavelmente esse tipo de corpus não será representativo de nenhum “todo” (como por exemplo, um corpus de aviação), mas incluirá o maior tipo de textos possível. Um corpus de língua geral precisa ser muito maior do que um corpus específico. É muitas vezes utilizado como contraste em relação aos corpora mais especializados. Por essa razão, são por vezes denominados de Corpora de Referência. Um dos corpora mais famosos da língua inglesa é o British National Corpus (BNC), que será utilizado como referência neste trabalho.

-Corpus Monitor- Corpus projetado para verificar mudanças atuais em uma língua. Esse tipo de corpus é alimentado anualmente, mensalmente, ou até mesmo diariamente, aumentando de tamanho rapidamente. Entretanto, a proporção de tipos de texto mantém-se constante, de forma que cada período de tempo possa ser comparado com o anterior. Um exemplo é o Bank of English, que atualmente conta com cerca de 400 milhões de palavras.

-Corpus Comparável- Dois (ou mais) corpora em línguas diferentes (inglês e português, por exemplo) ou em diferentes variedades de uma língua (português do Brasil e de Portugal, por exemplo). São compilados seguindo as mesmas diretrizes, ou seja, conterão a mesma proporção de gêneros, por exemplo, textos jornalísticos, romances, conversas informais, etc. Podem ser usados por tradutores ou por aprendizes para identificar diferenças e equivalências em cada língua. O exemplo mais citado desse tipo de corpus é o ICE (International Corpus of English), que contém mais de um milhão de palavras de diversas variedades da língua inglesa.

-Corpus Paralelo- Dois (ou mais) corpora em línguas diferentes contendo textos que foram traduzidos de uma língua para outra (por exemplo, um romance traduzido do inglês

para o português), ou textos que foram produzidos simultaneamente em duas ou mais línguas (por exemplo, normas da União Européia).

Além dos tipos de corpora citados acima, Hunston (2002, p.14) adiciona ainda os seguintes:

-Corpus de Aprendiz- Uma coletânea de textos-redações produzidos por aprendizes de uma língua. O propósito desse tipo de corpus é identificar em que aspectos os aprendizes diferem entre si e em relação a falantes nativos (em comparação a um corpus de falantes nativos²¹). Provavelmente, o mais conhecido seja o International Corpus of Learner English (ICLE), que consiste em redações escritas em inglês por falantes de várias línguas nativas (Português, Francês, Alemão, etc.).

-Corpus Pedagógico- Corpus que consiste na linguagem a qual um aprendiz é exposto. Pode consistir de livros didáticos e gravações. Esse tipo de corpus pode, por exemplo, ser comparado a um corpus de linguagem autêntica (produzida sem propósitos pedagógicos) para verificar se o aprendiz está sendo exposto à linguagem útil e natural.

-Corpus Histórico ou Diacrônico- Um corpus de textos de diferentes períodos de tempo. É utilizado para averiguar o desenvolvimento de certos aspectos de uma língua através dos tempos.

-Corpus Especializado²²- Um corpus contendo um tipo específico de texto (ou gênero), tal como resumos (*abstracts*), artigos acadêmicos sobre um assunto específico, conversas telefônicas, etc. Esse tipo de corpus tem por objetivo ser representativo de certo tipo de texto, ou linguagem. É comumente compilado pelo próprio pesquisador para refletir o tipo de

²¹ Neste estudo me refiro várias vezes à dicotomia “falante nativo” x “falante não nativo”, uma vez que vários autores da área de LdC se referem a esse aspecto. Apesar de estar além do escopo do trabalho, considero importante marcar a minha posição e entendimento a cerca dessa dicotomia. Tradicionalmente, essa visão dicotômica serviu ao propósito de privilegiar e contemplar somente as variantes lingüísticas dos falantes nativos. Da mesma forma, essa dicotomia também favoreceu e privilegiou o professor de língua estrangeira falante nativo da língua, considerando, assim, o professor não-nativo como inferior. Ou seja, nessas visões, o falante nativo é o padrão, o “correto”, e o não nativo deveria se esforçar ao máximo para atingir esse padrão nativo, que é geralmente inatingível (e indesejável). Há, entretanto, um novo entendimento sobre essa questão (a dicotomia “falante nativo” e “falante não nativo”), focalizando não as falhas e fraquezas dos falantes não-nativos, mas sim o sucesso das interações entre falantes que se comunicam em uma determinada língua independentemente de suas línguas maternas (SCHLATTER et al., 2004). Esse novo entendimento (discutido na Sociolingüística Interacional) derruba, da mesma forma, a superioridade relativa do professor de língua falante nativo, justificando que o professor não-nativo pode, muitas vezes, por conhecer a cultura e a língua materna desse aluno, melhor dar conta das particularidades e necessidades desse aluno, facilitando, assim, o aprendizado. Nesse sentido, as pesquisas realizadas contrastando corpora de aprendizes e de falantes nativos, antes de apontar erros, ou desvios da forma padrão, devem verificar, até que ponto esses “desvios”, ou diferenças (palavra preferida por mim), realmente interferem no sucesso das comunicações.

²² Os corpora especializados serão melhor explicados na seção 2.8.

linguagem que quer investigar. Não há limite para o grau de especialização envolvido, mas há parâmetros para limitar o tipo de texto incluído. Esse é o tipo de corpus utilizado para esta tese, o Corpus de Aviação (*Aviation Corpus*), compilado pela autora, contendo textos de três diferentes manuais da aeronave 737 produzida pela Boeing.

2.4 TERMOS DA LINGÜÍSTICA DE CORPUS

Definirei a seguir termos utilizados nas pesquisas baseadas em corpus e que serão importantes para o entendimento do presente estudo.

- *Type, token e hapax*

Para explicar esses termos, tomarei o seguinte parágrafo sobre *rotorcraft*²³ como exemplo:

If there is provision for a second pilot, the rotorcraft must be controllable with equal safety from either pilot position. Flight and powerplant controls must be designed to prevent confusion or inadvertent operation when the rotorcraft is piloted from either position. Inflight leakage of rain or snow that could distract the crew or harm the structure must be prevented.

A tabela abaixo (baseada no parágrafo acima) serve para explicar o significado de *tokens*, *types* e *hapax*:

Tabela 2.4.A: Tokens

Palavra	No. de ocorrências
<i>The</i>	4
<i>Must</i>	3
<i>Be</i>	3
<i>Or</i>	3
<i>It</i>	2
<i>Pilot</i>	2
Todas as outras 42 palavras restantes	1

Há 59 palavras, ou seja, 59 seqüências de letras separadas por espaço ou pontuação. Isto é, há 59 *tokens*. Entretanto, muitas dessas palavras ocorrem mais de uma vez. Contando cada um dos itens acima repetidos somente uma vez de forma que somente palavras diferentes

²³ Tipo de aeronave com hélices: girocôptero.

sejam contadas, tem-se um total de 48 palavras, ou *types*. As palavras que ocorrem apenas uma vez são chamadas *hapax*²⁴.

O parágrafo sobre *rotorcraft*, portanto, compreende um corpus de 59 *tokens*, 48 *types*, incluindo 42 *hapax*. Em um corpus muito pequeno, como o parágrafo supracitado, a razão *type/token* (forma/item) será alta. Em um corpus maior, haverá relativamente mais *tokens* (itens) para cada *type* (forma), uma vez que há mais repetição de palavras individuais em textos longos.

- *Lema e formas da palavra*

Para definir o conceito de “lema”, refiro-me novamente ao parágrafo sobre *rotorcraft*. Alguns pesquisadores, por exemplo, consideram *is* e *be* como sendo a “mesma” palavra, uma vez que *is* é apenas uma conjugação de *be*, sendo assim possível dizer que *be* e *is* pertencem ao mesmo lema: *be*. Entretanto, alguns pesquisadores dizem que duas palavras não pertencem ao mesmo lema quando pertencem a diferentes classes, como é o caso do substantivo *pilot* e do verbo *piloted*. Para a LdC, a noção de lema torna-se conveniente quando se deseja observar apenas os substantivos, excluindo-se adjetivos ou verbos, por exemplo. Para que se possa utilizar a lematização²⁵, é necessário que um corpus seja anotado. Essa noção será explicada a seguir.

- *Anotação*

Anotação refere-se aos procedimentos realizados para adicionar informações às palavras em um corpus. Essas informações podem ser adicionadas automaticamente (somente por um programa de computador) ou manualmente (por alguém trabalhando no programa de computador). Há diferentes formas de anotação, tais como a etiquetagem morfossintática, a anotação sintática, entre outras.

A etiquetagem morfossintática (*Part-of-speech tagging* ou *POS tagging*) é a adição de um código a cada palavra em um corpus, indicando a classe gramatical (substantivo, verbo, adjetivo, etc.)

²⁴ *Hapax* origina-se do grego e significa “algo dito apenas uma vez”.

²⁵ A lematização é amplamente usada nos dicionários de língua geral, em que os verbos aparecem sempre na forma infinitiva, e os substantivos no singular. Os dicionários técnicos, por sua vez, preferem incluir as unidades lingüísticas como elas realmente aparecem nos textos. Da mesma forma, seguidores da abordagem *corpus-driven* acreditam que cada forma de uma mesma palavra possua diferentes colocações e usos, preferindo assim, a palavra ortográfica.

Anotação sintática (*parser*) constitui a análise de um texto em constituintes, ou seja, a identificação da estrutura sintática (sintagma nominal, verbal, etc) de cada frase. Esse tipo de anotação, a exemplo da etiquetagem morfossintática, pode ser realizada automaticamente ou manualmente. Contudo, um bom nível de precisão é difícil de ser obtido automaticamente.

Outros tipos de elementos podem ainda ser anotados, tais como: entonação, em corpora de linguagem oral; anáforas, que identifica os itens de coesão e seus referentes; semântica, que é a categorização de palavras e sintagmas em campos semânticos (avião = meio de transporte, piloto = profissão).

No presente estudo, foi utilizada a etiquetagem morfossintática no corpus de referência com relação aos VMs, pois vários deles possuem palavras homônimas, por exemplo: *may*=VM e o mês de “maio”, *can*=VM e “lata”. No corpus de estudo, tal anotação não foi necessária, uma vez que não havia homônimos com a mesma entrada ortográfica. É importante ressaltar que as anotações não são neutras e que há sempre uma teoria embasando os diferentes tipos de anotação²⁶.

- *Colocação*

Colocação refere-se à forma na qual duas ou mais palavras são tipicamente usadas juntas. Por exemplo, usa-se *heavy rain*, mas não *heavy sun*. Nesse contexto, *heavy* é colocado de *rain*, ou *heavy* e *rain* são colocados²⁷.

- *Coligação*

Coligação refere-se à associação entre itens lexicais e gramaticais. Por exemplo, *start* é mais comum com sintagmas nominais e orações-*ing*, enquanto *begin* é mais usado com um complemento *to*” (BERBER SARDINHA, 2004, p. 40).

- *Prosódia semântica*

Prosódia semântica é o termo usado para referir a palavra ou expressão usadas em um contexto específico de tal forma que a palavra/expressão adquira certa conotação daquele contexto. Um exemplo seria a expressão *sit through* (HUNSTON, 2002, p. 141), que é geralmente usada com itens que indicam algo longo e cansativo, ou seja, assume uma conotação de algo tedioso.

²⁶ Esse assunto será amplamente discutido na seção 2.9.

²⁷ As colocações serão mais detalhadamente descritas na seção 2.6.

2.5 ANÁLISE DE CORPORA

Um corpus, como já dito anteriormente, é um repositório de textos digitais. Para que seu conteúdo seja acessado, é necessário que haja recursos, ou ferramentas para tal. Os corpora maiores, como o BNC, geralmente possuem seus próprios recursos ou ferramentas de acesso. Outros corpora, por exemplo, o Corpus de Aviação (AC), necessitam ser armazenados e acessados através de programas específicos para a descrição lingüística, como é o caso do *WordSmith Tools* (SCOTT, 1996)²⁸ ou do Corpógrafo²⁹. Em qualquer uma das formas acesso, os recursos mais utilizados nas investigações lingüísticas são:

- Concordâncias;
- Lista de frequência de palavras;
- Lista de colocados³⁰

- *Concordâncias*

O concordanciador é provavelmente a ferramenta computacional mais utilizada para processar informações em um corpus. Um concordanciador é um programa que busca, em um corpus, uma palavra selecionada ou um sintagma, apresentando todas as ocorrências daquela palavra ou sintagma no centro da tela do computador com as palavras que as antecedem ou seguem à esquerda e à direita, isto é, o co-texto. A palavra selecionada que aparece no centro da tela é chamada de nóculo ou palavra nóculo (*node* ou *node-word*). O material é disposto de forma a facilitar a visualização dos padrões da palavra-nóculo. Assim, as observações de padrões como colocação, coligação e prosódia semântica são otimizados. No exemplo abaixo foi utilizado o corpus Brown, a palavra nóculo escolhida foi o VM *must*, e a busca³¹ foi aleatória:

```

1 ce, one of two alternative courses must be taken: _1._ Five
2 cent of the voters in each county must sign petitions requestin
3 llot, or _2._ The Republicans must hold a primary under the
4 sersion, and ADC dependency". # "MUST SOLVE PROBLEM" # The mont
5 Co&, committee chairman. "We must solve the problems which
6 negative side of the balance sheet must be set some disappointme
7 entirely different societies, and must be treated as such withou
```

²⁸ O *WordSmith Tools* foi a ferramenta usada neste trabalho e será referida novamente no capítulo sobre metodologia.

²⁹ O Corpógrafo é um software de acesso livre disponível em <http://poloclup.linguateca.pt/ferramentas/gc/>

³⁰ Apesar de estarem sendo citadas nesta seção, a listas de colocados serão discutidas na seção seguinte, 2.6.

³¹ Esse tipo de busca é também conhecido como KWIC – *Key Word in Context*

A mesma busca pode ser realizada com as palavras ordenadas alfabeticamente à direita da palavra nódulo (*sort right*):

```

1   inue in his chosen profession, he must abandon his own code and
2   and to which even law enforcement must accommodate itself. One
3   union members under contract, it must accomplish its payroll b
4   of time, a method of preservation must accomplish the destructi
5   ating the antennae and receivers, must account for much of the d
6   ades across Berlin the free world must acquiesce in dismemberme
7   turning point in history, and we must act decisively and wisely

```

Ou, ainda, com as palavras ordenadas à esquerda (*sort left*):

```

1   es cleaning and drying equipment a must for modern gin operat
2   ry style in fine fashion and is a must for those who want to col
3   education (read "reading") was "a must". He moved in a "highly
4   ques. The platform accelerometers must be slightly modified for
5   erials and library accommodations must be planned. In the secon
6   was modern, large, on five acres. Must have cost plenty. The St
7   firm that Christian love-in-action must first justify and then d

```

- *Lista de freqüência de palavras*

Uma lista de freqüência é simplesmente uma lista de todas as formas, ou vocábulos (*types*) em um corpus juntamente com o número de ocorrências de cada forma/vocábulo (*tokens*). A lista pode ser classificada por ordem de freqüência, com as formas mais freqüentes em primeiro lugar, ou, ainda, alfabeticamente. Essa listagem pode também ser lematizada ou não. A comparação de listas de freqüência pode fornecer informações interessantes sobre os diferentes tipos de textos, uma vez que para a LdC, textos são formatados por textos anteriores, através de repetições ou através de rotinas e convenções. “Os textos são historicamente herdados”. (STUBBS, 1996, p. 34).

Essa comparação é especialmente importante entre corpora especializados. Kennedy (1998, p. 102) salienta que “quanto mais especializado” for um corpus (inglês acadêmico comparado com inglês geral, ou inglês para economia, comparado com inglês acadêmico geral) maior será o número de palavras lexicais (ou com mais conteúdo) entre as mais freqüentes. Nesse sentido, o autor menciona que em um corpus de economia, 18, entre as 50 palavras mais freqüentes são lexicais; enquanto em um corpus de inglês acadêmico geral, somente três, entre as 50 mais freqüentes são lexicais; no corpus Birmingham, por exemplo, somente *said* (considerado lexical) está entre as 50 mais freqüentes, as outras 49 são consideradas palavras gramaticais.

Na comparação entre dois corpora, somente o resultado normalizado não é prova suficiente de que o resultado é significativo, ou seja, que não é aleatório. A aplicação de testes estatísticos pode fornecer resultados até 99% confiáveis de que as diferenças são motivadas, ou seja, não são aleatórias.

Conforme Rayson (2002), o teste estatístico com melhores resultados para a comparação da frequência de palavras ou expressões entre dois corpora é o *Log-Likelihood* (doravante LL). Se o resultado obtido após a aplicação do LL foi de 6,63 ou mais, a probabilidade de a diferença entre os dois corpora ter acontecido aleatoriamente é de menos de 1%. Dessa forma, o pesquisador pode estar 99% confiante de que o resultado é significativo. Esse resultado é geralmente expresso como $p < 0.01$ ³².

Palavras que são significativamente (com base estatística) mais freqüentes em um corpus que em outro são também conhecidas como “palavras chave” (*keywords*). O *WordSmith Tools* (SCOTT, 1996) inclui um recurso que compara dois corpora (geralmente um maior, mais geral, e outro menor, mais especializado) automaticamente³³.

Segundo Hoffmann (1988[2004]), o aspecto freqüência é de especial interesse para o estudo das linguagens especializadas. O autor defende a idéia de que a especificidade de uma sublinguagem “se expressa geralmente por parâmetros quantitativos, isto é, pela freqüência de determinadas manifestações lingüísticas” (HOFFMANN, 1988[2004], p. 80). Como este trabalho trata sobre a incidência e a funcionalidade dos VMs em um corpus especializado, acredito que a quantificação dessas unidades seja de especial importância para o melhor conhecimento da linguagem dos manuais de aviação.

2.6 COLOCAÇÃO, PADRONIZAÇÃO E FRASEOLOGIA

Firth, com sua célebre frase “*You shall know a word by the company it keeps*” mostrou ao mundo da lingüística a importância dos estudos descritivos da linguagem, especialmente, a importância do co-texto de uma palavra de forma a conhecê-la. Nesse sentido, as colocações habituais das palavras são simplesmente os acompanhantes desta palavra. Ligado ao conceito de colocação, há os conceitos de fraseologia e padronização. Apesar de esses conceitos não serem necessariamente dependentes da LdC, os apresento em detalhes nesta seção, pois, ao

³² A fórmula do LL está na íntegra no anexo A.

³³ Para tanto, os dois corpora devem poder ser armazenados e acessados pelo *WordSmith Tools*. Esse não é o caso do BNCWeb, que somente pode ser acessado através de seus próprios recursos. Por essa razão, esse recurso não será utilizado neste estudo.

referir-me à padronização e à fraseologia, estarei referindo aos conceitos relacionados à LdC. Da mesma forma, por serem termos polissêmicos, considero importante estabelecer de forma clara e precisa o que se entende por padronização e fraseologia.

Como já mencionado anteriormente, colocação refere-se à tendência com que as palavras co-ocorrem com outras. A palavra *toy* (brinquedo), por exemplo, co-ocorre com *children* (crianças) mais frequentemente do que com *men* (homens) e *women* (mulheres) (HUNSTON, 2002, p. 68). Hunston considera essa colocação como motivada, pois há uma explicação lógica para tal. Outras colocações, no entanto, como *strong tea* (chá forte) e *powerful car* (carro potente), não possuem uma motivação aparente.

As colocações podem ser observadas com números absolutos, constituindo uma “associação entre itens lexicais” (Berber Sardinha, 2004, p. 200), mas tornam-se mais confiáveis se medidas estatisticamente. Dessa forma, é possível verificar até que ponto a relação palavra-nóculo e colocado não é aleatória, ou seja, uma “associação não aleatória entre itens lexicais” (ibid). Para Berber Sardinha (2004, p. 201): “Uma associação não-aleatória é aquela que é mais comum do que o esperado. Para saber se uma associação entre palavras não é aleatória, precisamos de apoio estatístico, na forma de medidas estatísticas de associação”.

Programas que calculam a colocação em números absolutos, contam as ocorrências de todas as palavras que ocorrem em um certo horizonte, por exemplo, quatro palavras para à esquerda da palavra-nóculo e quatro para à direita. É importante ter acesso a uma grande quantidade de dados para calcular a colocação, principalmente da língua geral, pois assim, haverá mais confiabilidade de que as colocações mais significativas são realmente mostradas.

São vários os testes de significância utilizados. Os mais comuns são o qui-quadrado (*chi-square*), o informação-mútua (*mutual information*), o *z-score* e a razão do LL. Rayson (2002) e Gómez (2002) reportam que o qui-quadrado torna-se não confiável quando a frequência esperada é muito pequena (menos do que cinco), possivelmente superestimando a significância de palavras muito frequentes e/ou ao comparar um corpus relativamente pequeno com outro muito maior. Para Rayson (2002), o LL é preferido sobre o qui-quadrado. Além disso, Gómez (2002) menciona que a Informação Mútua tende a superestimar o grau de associação quando os eventos são raros. Com base nas sugestões desses autores, e tendo em vista que o objetivo deste trabalho é colher subsídios para a elaboração de material didático,

não sendo assim relevante desvendar colocações raras, uso o LL para medir o grau de associação entre o nódulo e seus colocados.

Um dos importantes usos das informações provindas das listas de colocações é ressaltar os diferentes significados de uma palavra. Hunston (2002, p. 76) lista os colocados do verbo *leak*. Algumas das associações são relacionadas ao significado físico de *leak*: *oil*, *water*, *gas*, *roof*; enquanto outras são associadas ao sentido metafórico: *information*, *report*, *memo*, *confidential*. Em outras palavras, a lista de colocados fornece uma espécie de perfil semântico das palavras envolvidas.

Sinclair (1991) sugere que quando duas palavras de freqüências diferentes são consideradas como colocados, a colocação tem um valor diferente na descrição de cada uma dessas palavras. Se uma palavra “a” é duas vezes mais freqüente do que uma palavra “b”, cada vez que elas ocorrerem juntas é duas vezes mais importante para “b” do que para “a”, pois o evento é responsável por duas vezes a proporção da ocorrência de “b” com relação à “a”. Ou seja, quando todas as ocorrências de “a” com “b” são calculadas, um resultado é registrado no perfil de “a” e outro resultado é registrado no perfil de “b”. Dessa forma, Sinclair considera separadamente esses dois tipos de colocação, e utiliza o termo “nódulo” para a palavra que está sendo estudada e “colocado” para qualquer palavra que ocorre no ambiente específico de um nódulo. Cada palavra consecutiva em um texto é assim, nódulo e colocado, embora nunca ao mesmo tempo. Quando uma palavra mais freqüente “a” se coloca com uma palavra menos freqüente “b”, Sinclair denomina de “colocação descendente”, e o contrário de “colocação ascendente”. Segundo o autor, a colocação ascendente não possui um valor estatístico significativo, e a maioria das palavras tendem a ser elementos de estruturas gramaticais ou hiperônimos. A colocação descendente, por sua vez, demonstra uma análise semântica da palavra.

Conforme Berber Sardinha (2004), “De um modo geral, a padronização é a regularidade expressa na recorrência sistemática de unidades co-ocorrentes de várias ordens (lexical, gramatical, sintática, etc.)” (ibid, p. 47). Isto é, para que sejam definidos os padrões de uma palavra, faz-se necessário averiguar as palavras e as estruturas freqüentemente associadas a ela que de alguma forma refletem no seu significado. Para Berber Sardinha (2004), “padrão” e “fraseologia” são muitas vezes utilizados como sinônimos. Dessa forma, há por vezes o emprego de expressões como “a fraseologia da palavra X” referindo-se aos padrões observáveis da palavra em questão. Seja qual for o termo utilizado para essa

descrição, ela é considerada de extrema relevância para o ensino de língua estrangeira, pois aspectos como naturalidade e fluência são demonstrados por meio de padrões.

Antes de aprofundar o assunto como exposto por Berber Sardinha (2004) e Sinclair (1991) (a seguir), acredito que seja importante fazer uma ressalva quanto ao termo “fraseologia”, uma vez que se trata de palavra polissêmica, assumindo diferentes sentidos. Altenberg (1998) afirma que a fraseologia é uma parte obscura de um idioma. Ela engloba o lado convencional, ao invés do lado produtivo ou regido por regras da língua, e envolve variados tipos de unidades compostas e expressões “predefinidas”, tais como expressões idiomáticas, frases feitas e locuções.

De acordo com Gläser (1998), uma unidade fraseológica é um grupo lexicalizado de palavras, reproduzível, composto por pelo menos duas palavras, que possui uma relativa estabilidade sintática e semântica, que pode ter características idiomáticas, que pode conter conotações, e que pode ter uma função enfática ou intensa em um texto. Unidades fraseológicas são ainda entendidas como um grupo de palavras estáveis com significado parcial ou totalmente não-composicional, por exemplo, “bater as botas”, que no português brasileiro vigente significa “morrer”. Esse significado pode ser considerado um sinônimo das “expressões idiomáticas”. Há ainda quem inclua os provérbios nessa categoria.

Na área da Terminologia, há as Unidades Fraseológicas Especializadas (UFE). Para Bevilacqua (2001, p. 106), as UFE são entendidas como “unidades sintagmáticas que incluem um termo entre os seus elementos, que possuem um determinado grau de fixação e uma frequência relevante em um conjunto de textos ou em um âmbito especializado”. A pesquisadora cita como exemplos “instaurar um inquérito” e “cometer um delito”. Bevilacqua as difere das unidades livres afirmando que elas possuem um grau de fixação nulo e uma frequência mínima.

“Fraseologia aeronáutica” é entendida como um sistema lingüístico padrão e livre de ambigüidades artificialmente criado para facilitar a comunicação entre pilotos e controladores de tráfego aéreo (ROBERTSON, 1988). A fraseologia aeronáutica³⁴ é um procedimento estabelecido com o objetivo de assegurar a uniformidade das comunicações radiotelefônicas,

³⁴ Apesar de tratar da linguagem da aviação, considero importante ressaltar que o conceito de fraseologia a ser utilizado neste trabalho não é de forma alguma relacionado à fraseologia aeronáutica, pois o corpus de estudo contém unicamente manuais escritos da Boeing, nos quais tal fraseologia não ocorre.

reduzir ao mínimo o tempo de transmissão das mensagens e proporcionar autorizações claras e concisas. Esse tipo de comunicação tem como principal meta a segurança das vidas de passageiros e pilotos. Em outras palavras, fraseologia aeronáutica é a comunicação realizada entre pilotos e controladores de tráfego aéreo em vários momentos do voo, como por exemplo, quando um piloto necessita pedir autorização para decolar, pousar, ou informações sobre a pista em uso, entre outros vários casos. Mostro a seguir um exemplo de uma situação comunicativa:

PRJTS> *Good Evening Curitiba Ground, Papa Romeo Juliet Tango Sierra.*

SBCT_GND> *Papa Romeo Juliet Tango Sierra, Curitiba Ground, good evening, go ahead.*

PRJTS> *Juliet Tango Sierra, request Flight Plan Clearance to São Paulo, Flight level two nine zero, roger DELTA information.*

SBCT_GND> *Juliet Tango Sierra, Flight Plan received, standby for copy.*

Sinclair (1991), por sua vez, considera a fraseologia como a base da descrição lingüística desafiando outras visões sobre a linguagem. Para o autor:

- Não há distinção entre padrão sintático e significado;
- A língua possui dois princípios de organização, o princípio idiomático (*idiom principle*) e o princípio da livre escolha (*open-choice principle*);
- Não há distinção entre léxico e gramática.

Se uma palavra possui vários sentidos, cada sentido tende a ser associado com um conjunto diferente de padrões. Por exemplo, quando o adjetivo *mobile* é usado para qualificar coisas, significando “que pode ser carregado” (*can be moved*), geralmente precede o substantivo, por exemplo *mobile unit*, *mobile library*. Quando o mesmo adjetivo é usado para qualificar pessoas, significando “não impedido de mover-se por doença ou falta de recursos” (*not prevented from moving by disability or lack of resources*), geralmente sucede o verbo de ligação, por exemplo *I’m still very mobile* (HUNSTON, 2002, p. 139). Sinclair (1991) acredita que essa noção de fraseologia possa substituir a palavra isolada como unidade para o ensino de vocabulário, simplificando, dessa forma, a tarefa do aprendiz, uma vez que cada item lexical conteria mais informações sobre o seu uso.

Ainda nesse sentido, palavras com o mesmo padrão tendem a compartilhar aspectos comuns de significado. Hunston (2002, p. 140) cita o caso da seqüência: “verbo seguido por

substantivo seguido por *as* seguido por substantivo”. Nesses casos, a associação entre padrão e significado é tão forte que o significado parece pertencer à frase inteira, e não a cada palavra individual. Verbos com esse padrão parecem significar “fazer com que alguém ou alguma coisa seja ou pareça ser algo”. Por exemplo:

-he described it as a legalised theft;

-he revealed himself as a man of deep culture;

-I would like to appoint you as managing director.

Da mesma forma, a prosódia semântica somente pode ser conhecida através da observação de um grande número de ocorrências de uma palavra/expressão, pois baseia-se no uso típico da palavra/expressão. Como a prosódia semântica nem sempre faz parte do conhecimento consciente de um falante (seja ele nativo ou não), pode não ser ensinada, mas muitas vezes pode consistir em um aspecto importante da linguagem. O ensino de vocabulário deveria levá-la em consideração. Entretanto, isso só poderá acontecer se a abordagem for fraseológica e não baseada na palavra.

Relacionado ao conceito de fraseologia, Sinclair (1991, p. 115) estabelece dois princípios organizadores da língua, simultaneamente alternativos e complementares, a partir dos quais é possível interpretar o significado das palavras: (i) o princípio da livre escolha (*open-choice principle*), em que o falante tem como única restrição a gramaticalidade do enunciado; (ii) o princípio idiomático (*idiom principle*), em que o falante tem à sua disposição um grande número de grupos de palavras pré-construídos (ainda que possam apresentar alguma variação, nomeadamente no plano lexical, flexional ou de ordem das palavras). Quando uma elocução não pode ser interpretada sob o princípio idiomático, o usuário da língua recorrerá ao princípio da livre escolha. Hunston (2002) menciona a dificuldade em provar, ou não, a existência do princípio idiomático, mas argumenta que algumas seqüências de palavras notoriamente constituem fraseologias, ou combinatórias, como atestado em corpora maiores. Nesse sentido, para a autora, não é irracional supor que tais combinatórias serão codificadas e decodificadas como entidades únicas, e não como uma composição de significados de palavras individuais.

Sinclair (1991) sugere que quaisquer grupos ou seqüências de palavras são construídos e entendidos a luz de um dos dois princípios, mas nunca de ambos simultaneamente. Ou seja, o significado pode ser construído pela frase como um todo, operando de acordo com a fraseologia convencional, ou pelas palavras individualmente, operando de acordo com as

regras gramaticais. A escolha entre os princípios idiomáticos ou da livre escolha tornam a ambigüidade teoricamente possível; o fato de que apenas um ou outro princípio seja empregado por um usuário da língua em um momento, explica por que a ambigüidade é raramente um problema para falantes ou ouvintes. Hunston (2002) oferece *grasp the point* como exemplo. Para a autora, a frase é ambígua. De acordo com o princípio idiomático, ela significaria “entender a idéia principal de algo” (*understand the main idea of something*); interpretado de acordo com o princípio da livre escolha, *grasp* combinado com qualquer objeto sólido, significaria “segurar a ponta de algo” (*take hold of the sharp end of something*). Entretanto, falante e ouvinte utilizariam apenas um dos princípios, (possivelmente) eliminando a ambigüidade.

Sinclair (1991) salienta que entender linguagem como fraseologia (uma visão obtida observando-se uma grande porção de linguagem, e não palavra por palavra), necessita a rejeição de léxico e gramática como entidades separadas. Para o autor não há uma diferença essencial entre “palavras lexicais” e “palavras gramaticais”. Além disso, os padrões observáveis de itens lexicais são observações sobre o léxico e a gramática.

A visão mais tradicional sobre a linguagem é de que palavras lexicais são facilmente distinguíveis das palavras gramaticais, e que fatos lexicais, tais como colocação, são separados de fatos gramaticais tais como transitividade. A distinção entre palavras gramaticais e lexicais é baseada em algumas noções: de que palavras gramaticais são mais freqüentes que palavras lexicais; que palavras gramaticais são mais facilmente relacionadas paradigmaticamente, ao passo que palavras lexicais sintagmaticamente; e que palavras gramaticais não possuem significado próprio, mas que palavras lexicais possuem. Algumas palavras gramaticais possuem meios formais de identificação. Os verbos modais³⁵ em inglês, por exemplo, não flexionam, mas os verbos auxiliares (de uma forma geral), sim. Ainda sobre as palavras gramaticais na língua inglesa, essas são as únicas com menos de três letras no inglês escrito³⁶.

Entretanto, as diferenças listadas no parágrafo acima são desafiadas por evidências provenientes das pesquisas com corpora. Embora as palavras gramaticais sejam geralmente as mais freqüentes em listas de palavras, e as palavras lexicais sejam menos freqüentes, nem todas as palavras gramaticais são mais freqüentes do que as palavras lexicais. No BNC, por

³⁵ Os verbos modais parecem ocupar uma posição intermediária no continuum palavra gramatical e palavra lexical.

³⁶ Com exceção da palavra “ax” do inglês americano que quebra essa regra.

exemplo, *and, it, is, was, I, that, you, be, he e are* são as mais freqüentes. Entretanto, o verbo lexical *said*, é mais freqüente do que *up, in, did*, entre outras palavras gramaticais. Sinclair (1999) também ressalta que algumas palavras gramaticais muito freqüentes tais como *a*, participam do princípio idiomático da mesma forma que palavras lexicais. Um exemplo de *a* nessa situação seria a expressão *come to a head*, em que o artigo indefinido *a* não estaria em contraste com o artigo definido *the*. Ou seja, o comportamento paradigmático é anulado, pois se comporta da mesma forma que a palavra lexical *head* na mesma expressão. Hunston (2000³⁷, apud HUNSTON 2002) salienta que a fraseologia pode distinguir entre os significados dos VMs *may* e *must* da mesma forma que entre palavras lexicais.

A terceira distinção tradicional é que palavras gramaticais não possuem significados, ao contrário de palavras lexicais. Hunston (2002, p. 150) lembra que a palavra *lexical point*, em expressões como *from your point of view, the point is that e from that point on* tem um significado isolado bastante fraco. Ao passo que a palavra gramatical *would* deveria ser ensinada como tendo um significado próprio e não como parte de uma abstração gramatical chamada de “condicional”. Sinclair resume o princípio idiomático afirmando que “o usuário de uma língua tem disponível um grande número de sintagmas semi-pré construídos que constituem escolhas únicas, embora pareça que os sintagmas possam ser analisados em segmentos “(SINCLAIR, 1991, p. 110).

Para o autor, o fenômeno do princípio idiomático parece refletir a recorrência de situações semelhantes na vida e relações humanas; pode assim ilustrar uma tendência natural à economia de esforço; ou pode ser em parte motivado pelas exigências da conversação em tempo real. Independentemente de sua motivação, esse fenômeno foi relegado a uma posição inferior na maior parte dos estudos lingüísticos por não se encaixar no modelo do princípio da livre escolha.

Neste trabalho, fraseologia será utilizada com o significado semelhante àquele de padrão ou padronização, como proposto por Berber Sardinha (2004). Uma vez que, com base em Sinclair (1991), acredito que se uma palavra possui vários sentidos, cada sentido tende a ser associado com um conjunto diferente de padrões e que palavras com um mesmo padrão podem compartilhar aspectos comuns de significado. Ademais, Hunston (2002) sugere que é

³⁷ HUNSTON, S. Phraseology and the modal verb: a study of pattern and meaning. In: Heffer, C. e Saunston, H. *Words in Context: a tribute to John Sinclair on his retirement*. University of Birmingham CD-ROM, 2000

precisamente esse tipo de estudo que pode ajudar na distinção dos significados e usos dos modais.

2.7 VARIAÇÃO E DISTRIBUIÇÃO

Um dos aspectos do estudo lingüístico que tem sido assistido pelo desenvolvimento de corpora é o estudo da variação entre linguagens produzidas em diferentes situações. Apesar de haver uma longa tradição na investigação de diferentes registros³⁸ e de gêneros, os corpora adicionaram uma nova dimensão aos tipos de pesquisa que podem ser conduzidas. O estudo da variação é essencialmente o estudo de comparações entre discursos produzidos em momentos diferentes, com diferentes objetivos ou propósitos, por diferentes grupos de pessoas, ou sob diferentes condições.

Enquanto diferenças claras e significativas são facilmente identificadas entre diferentes registros, é também possível identificar diferenças claras e significativas intra-registros. Hunston (2002) menciona diversos trabalhos que encontraram diferenças entre artigos acadêmicos de diferentes disciplinas, (BIBER et al., 1998³⁹; GLEDHILL, 1995⁴⁰). A autora declara que a variação poderá ser encontrada onde for procurada, desafiando, dessa forma, a validade dos estudos em registros amplos, como os trabalhos de Mindt (2000)⁴¹. Todavia, a própria autora contra argumenta afirmando que essas categorias amplas servem como base para estudo desse tipo.

São vários os parâmetros a serem utilizados entre os registros/gêneros. Um desses parâmetros é a frequência de palavras. Muitas palavras não são similarmente distribuídas entre diferentes registros, mas ocorrem mais freqüentemente em um ou outro registro. Biber et al. (1999, p. 376)⁴² mostram que o verbo lexical *get* é o mais freqüente em conversações, mas bastante infreqüente nos registros escritos (ficção, notícias, prosa acadêmica). Por outro lado, *make* é o verbo lexical mais freqüente na prosa acadêmica, e apenas o décimo primeiro na

³⁸ Um registro é um subconjunto de uma linguagem utilizado para um propósito específico ou em um contexto social específico. É importante salientar que, principalmente na literatura sobre a LdC, as noções de “registro” e “gênero” se sobrepõem, não havendo uma distinção clara entre elas.

³⁹ BIBER, D.; CONRAD, S.; REPPEN, R. *Corpus linguistics: investigating language structure and use*. Cambridge, Cambridge University Press, 1998.

⁴⁰ GLEDHILL, C. Collocation and Genre Analysis. The Phraseology of Grammatical Items in Cancer Research Abstracts and Articles. 1995. In S. Botley, J. Glass, T. McEnery, & A. Wilson (Eds.), *Proceedings of the Teaching and Language Corpora*, UCREL Technical Papers 9: 108-126. 1996.

⁴¹ MINDT, D. *An Empirical Grammar of The English Verb System*. Berlin: Cornelsen, 2000.

⁴² Biber et al. (1999) consideram quatro abrangentes registros em sua gramática baseada no *Longman Spoken and Written English Corpus (LSWEC)*: notícias, prosa acadêmica e ficção (todos registros escritos) e conversação (naturalmente, registro oral).

conversação (ibid, p. 375). Similarmente, vários VMs, nomeadamente *can*, *will*, *would*, são bem mais freqüentes na conversação do que nos registros escritos (ibid, p. 8).

Características gramaticais são, da mesma forma, distribuídas diferentemente entre registros. Interrogativas, por exemplo, “são 47 vezes mais freqüentes em conversação do que em prosa acadêmica ou notícias, mas apenas quatro vezes mais freqüentes do que em ficção” (BIBER et al., 1999, p. 211). De acordo com os autores, a alta freqüência de frases interrogativas na conversação é natural, considerando que a situação é interativa, com um constante dar-e-receber entre os participantes. Similarmente, a presença de diálogos pode explicar a relativa alta freqüência das interrogativas em ficção. Por outro lado, notícias e prosa acadêmica não são interativas fazendo, naturalmente, menos uso da forma interrogativa.

2.8 O DESENVOLVIMENTO DE MATERIAL DIDÁTICO

Desde o início da década de 80, o ensino de uma língua estrangeira voltou-se para o processo de comunicação, quando foram introduzidos os conceitos de “abordagem comunicativa” e “competência comunicativa”⁴³. Esse reposicionamento afastou o ensino de línguas estrangeiras de um foco mais tradicional na forma. Ou seja, o aprendizado passou a ser baseado em “como se fazer coisas com a língua”. De certa forma, os livros-texto concentravam-se mais no tipo de atividade apresentada, que deveria ser “comunicativa”. Assim, apesar de ter havido um tipo de ruptura com o tradicional foco na forma, os livros-texto são, de uma maneira geral, bastante pautados na “intuição nativa” de seus autores, que, como já mencionado neste capítulo, possuem apenas uma visão parcial e pessoal de uma língua. Além disso, os livros-texto ainda são baseados nas gramáticas normativas. Mais recentemente, parece estar surgindo o reconhecimento de que a LdC pode ser um valioso recurso para autores de livros-texto e outros materiais didáticos por que fornecem uma visão detalhada sobre como uma língua é realmente usada. A LdC pode oferecer informações relacionadas a vocabulário, gramática, formalidade e informalidade, diferenças entre a linguagem escrita e falada, como as pessoas começam e terminam uma conversa, entre outros

⁴³ Hymes (1979) foi o primeiro a incorporar a dimensão social ao conceito de “competência” anteriormente atrelada quase que unicamente à Chomsky. Ao acrescentar comunicativo ao termo *competência*, demonstrou claramente estar preocupado com o uso da língua. Assim, para Hymes, não é o bastante que o indivíduo saiba e use a fonologia, a sintaxe e o léxico da língua para caracterizá-lo como competente em termos comunicativos. É preciso que, além disso, que esse indivíduo saiba e use as regras do discurso específico da comunidade na qual se insere. O indivíduo demonstra possuir competência se sabe quando falar, quando não falar, e a quem falar, com quem, onde e de que maneira. Deve-se a Hymes igualmente, a ampliação do conceito de competência para incluir a idéia de “capacidade para usar”, unindo desta forma as noções de “competência” e “desempenho” que estavam bem distintas na dicotomia proposta por Chomsky.

aspectos. Desta forma, estudos baseados em corpora podem sugerir os itens lingüísticos e processos que serão mais provavelmente encontrados por usuários de uma língua (ou sub-linguagem) e que, portanto, merecem mais investimento em termos de tempo (HOFFMANN, 1998[2007]).

Para a utilização de um corpus na elaboração de material didático (livros, polígrafos ou exercícios) é necessário, primeiramente, decidir (no caso da língua inglesa) quanto ao tipo e variedade de inglês que servirá como base para a elaboração do material, uma vez que corpora diferentes apresentarão palavras diferentes e, freqüentemente, diferentes usos e funções das palavras a serem ensinadas. A palavra *nice*, por exemplo, é uma das quinze palavras mais freqüentes no inglês falado (McCARTEN, 2007). Entretanto, ela é bastante rara no inglês acadêmico escrito, ocorrendo sempre em citações de literatura ou em entrevistas. Portanto, a escolha (ou a compilação) de um corpus pode afetar as palavras a serem incluídas nos materiais didáticos, assim como seus sentidos e usos.

Ainda sobre o que ensinar, tanto o ensino de ESP⁴⁴, quanto o ensino de Inglês Geral, englobam uma grande variedade de propósitos específicos. Provavelmente o que os distingue é que o ESP possui um número menor de variedades, tipos de texto e situações comunicativas. Conforme Gavioli (2005, p. 23), buscar os itens básicos a serem ensinados em um curso de ESP é um problema crucial para os professores e diretamente associado com cada ambiente de ensino. No caso de inglês acadêmico, por exemplo, há uma enorme variedade de situações de ensino e também de necessidades dos alunos. Ou seja, enquanto ler livros e artigos em inglês é uma grande necessidade para estudantes de medicina que moram em países de fala não-inglesa, alunos de medicina estrangeiros morando na Inglaterra, Estados Unidos, ou qualquer outro país onde o inglês seja a língua de instrução, terão a necessidade de trabalhar o inglês oral.

Com relação a formas de classificar ESPs, alguns autores, como Dudley-Evans e St. John (1998), utilizam categorias bastante amplas para classificar os tipos de ESP: Inglês para Médicos, Inglês para Negócios, etc. Entretanto, os corpora especializados tendem a representar uma porção muito mais restrita de textos especializados, como por exemplo, manuais didáticos de Química (AZEREDO, 2007; FINATTO et al., 2003), ou sobre resíduos ambientais (KILIAN, 2007).

⁴⁴ *English for Specific Purposes*, que em alguns casos é referido como “Inglês Instrumental”. Contudo, Inglês Instrumental é também usado para referir-se ao ensino unicamente da habilidade de leitura em detrimento as outras habilidades.

Um corpus especializado significa, assim, um corpus projetado de forma a representar uma amostra de uma sub-linguagem. Dependendo da pesquisa, ou dos propósitos de ensino, os corpora podem variar quanto ao grau de especialização. Biber (1993, p. 245) sugere os seguintes aspectos para caracterizar os diferentes tipos de corpora:

Quadro 2.8.A: Caracterização dos tipos de corpora

Canal	Escrito/ Falado
Formato	Publicado/ Não-publicado
Cenário	Institucional / Público/ Privado-pessoal/ Outro
Receptor	Quantos / Outro / Si próprio / Presente / Ausente / Grau de Interatividade / Grau de conhecimento compartilhado
Produtor	Sexo / Idade / Ocupação / Pessoa Reconhecida / Instituição
Factualidade	Informação Factual / Imaginativo / etc.
Propósito	Convencer / Entreter / Informar / Instruir / Explicar / Narrar / Descrever / Registrar / Revelar-se / Expressar Atitudes / Opinar / Expressar Emoções / Relacionar-se / Outros

Gavioli (2005) exemplifica os diferentes graus de especialização dos corpora mencionando um corpus de Hepatite C e um corpus sobre Inglês Acadêmico de Medicina. Para a autora, o corpus sobre Hepatite C deveria conter um número de artigos que tratam sobre o tópico, escritos por autores de importância reconhecida na área e, se possível, de ambos os sexos. O corpus sobre a linguagem acadêmica da medicina, deveria considerar a linguagem escrita e a oral, publicada, institucional, com interlocutores presentes e ausentes, em cenários interativos e não-interativos, com diferentes graus de conhecimento compartilhado. Além disso, os produtores dos textos deveriam representar ambos os sexos e os tópicos deveriam cobrir várias áreas da medicina. Apesar de os dois corpora serem especializados, o primeiro pode ser considerado muito mais especializado e restrito.

Um corpus de textos especializado pode ser usado para gerar rapidamente uma série de itens que podem caracterizar os textos especializados em questão. As listas de frequência (tanto de palavras quanto de concordâncias) podem auxiliar na tomada de decisão quanto ao que ensinar e em qual ordem (KENNEDY, 1998; GAVIOLI, 2005). Expressões mais raras em um dado contexto podem ser ensinadas em um momento posterior. Entretanto, nem sempre é possível ensinar primeiramente somente as palavras mais frequentes de uma língua (ou sub-linguagem) por várias razões. Algumas podem ser culturalmente inapropriadas, ou não adequadas para o contexto de sala de aula. Adicionalmente, algumas palavras/expressões podem apresentar uma dificuldade no aprendizado até que alunos tenham um conhecimento mais abrangente da língua. Por outro lado, os alunos podem necessitar outras

palavras/expressões além das mais frequentes nos corpora de inglês geral. A palavra *homework* não está entre as mil palavras mais frequentes da língua inglesa falada nesses corpora, mas é muito necessária e frequente no contexto de sala de aula. Por essas razões, o critério frequência deve servir apenas como um guia nesses casos.

Todavia, apesar da importância significativa de dados quantitativos no ensino de Inglês Geral, a relevância desse tipo de informação parece ser ainda mais importante no ensino de ESP. A natureza quantitativa de certos aspectos característicos de ESP tem sido tradicionalmente reconhecida em vários aspectos, tais como, a recorrência de itens lexicais específicos ou de estruturas gramaticais, tais como sintagmas nominais na linguagem da medicina (GOTTI, 1991). Halliday (1988) e Hoffman (1988[2004], 1991, 1998[2007]) definem uma linguagem especializada, ou sub-linguagem, como uma linguagem funcional, ou um registro, e descrevem os registros em termos quantitativos. De acordo com Halliday, “um registro é um aglomerado de características associadas que possuem uma tendência maior que o acaso de co-ocorrerem” (HALLIDAY, 1988, p. 162).

Hoffmann (1998[2007]) menciona que resultados quantitativos e estatísticos devem ser “amplamente considerados no ensino de línguas estrangeiras” como um todo, mas “principalmente no ensino das linguagens especializadas” (ibid, p. 74). Conforme o autor é importante escolher e delimitar o material de forma que o aluno possa aprender, em menor tempo, a quantidade máxima possível de conhecimentos, práticas e habilidades. Um “*minimum*” deve conter as ocorrências lexicais e gramaticais mais úteis. A quantidade de unidades lingüísticas necessárias para solucionar certas tarefas de comunicação constitui “o cerne do ensino de línguas estrangeiras.” Para Hoffmann (1998[2007]), esse *minimum* deve passar por uma criteriosa e rigorosa seleção.

Gavioli (2005, p. 55) cita o trabalho de Barber (1962)⁴⁵ realizado em um corpus de aproximadamente 23.000 palavras formado por três artigos acadêmicos de três domínios diferentes (engenharia, química e astronomia). Barber (devido às limitações tecnológicas da época) calculou manualmente dados como o comprimento médio de uma frase, ocorrência dos VMs mais frequentes e ocorrência de diferentes tempos e aspectos verbais. Swales (1988⁴⁶,

⁴⁵ BARBER, C. L.(1962) *Some Measurable Characteristics of Modern Scientific Prose*. In: J. Swales (ed.). 1985. *Episodes in ESP*. A Source and Reference Book on the Development of English for Science and Technology. Oxford: Pergamon Press Ltd. pp.1-1.

⁴⁶ SWALES, J. *Episodes in ESP*. Hertfordshire: Prentice Hall, 1988.

apud GAVIOLI, 2005) sugere que esse trabalho foi usado como argumento contrário ao ensino do presente e passado contínuo em cursos de inglês acadêmico, uma vez que os usos desses tempos verbais mostraram-se extremamente baixos.

Os corpora especializados geralmente apresentam mais ocorrências de palavras especializadas do que os corpora gerais. Por exemplo, a palavra *transaminase* ocorre 61 vezes em um pequeno corpus sobre Hepatite C (GAVIOLI, 2005), mas apenas 17 vezes no BNC completo (100 milhões de palavras). Entretanto, apesar de os corpora especializados fornecerem mais termos técnicos do que os corpora de inglês geral, os corpora especializados sozinhos não podem provar se aquelas unidades realmente caracterizam a linguagem representada no corpus. Isso ocorre porque um gênero específico, e a língua geral, não são categorias distintas. Características freqüentes que emergem da análise de um corpus específico podem não ser somente uma característica daquele gênero, mas sim de um macro-gênero (GAVIOLI, 2005; AZEREDO, 2007). Para verificarmos quais os aspectos que efetivamente caracterizam um corpus que representa um gênero ou registro, faz-se necessário comparar o corpus de estudo, com um corpus de um gênero semelhante e, finalmente, com um corpus de língua geral. Nesse sentido, Biber et al. (1998, p. 36) afirmam que “A análise de um registro requer uma abordagem comparativa: necessita-se de uma referência para sabermos até que ponto uma característica lingüística em um registro é raro ou comum”. Os pesquisadores sugerem que as características lingüísticas co-ocorrem em textos porque funcionam de formas similares. Em outras palavras, tomando-se um *corpus* de inglês falado como exemplo, contrações (ex. *'ll*, *'s*), falsos inícios (ex. *Ok, so we need to put*) e o uso de palavras de conteúdo geral (ex. *thing*) são todas relacionadas às pressões da produção em tempo real. Dessa forma, suas co-ocorrências fornecem indicações de discursos produzidos em tempo real (BIBER et al., 1999, pp. 10-11).

As colocações também são muito importantes na elaboração de materiais didáticos. Listas de palavras separadas, como as descritas acima, sugerem possíveis itens lexicais a serem incluídos no *syllabus*. As concordâncias dessas palavras mostram como elas são utilizadas em um corpus, fornecendo evidências adicionais sobre as suas características, específicas da linguagem especializada em questão. As concordâncias das palavras mais freqüentes em um corpus podem sugerir as formas típicas nas quais essas palavras são utilizadas e ressaltar suas colocações freqüentes. Gavioli (2005) afirma que as colocações não são pertinentes somente para os substantivos, mas também para outras categorias gramaticais. Para McCarten (2007), encontrar as colocações de verbos não lexicais, como *have*, *get*, *make*

e *do* é de especial importância. Esses verbos não parecem possuir um significado (lexical) próprio, mas adquirem um significado a partir das palavras com as quais se colocam. O verbo *make*, por exemplo, assume um significado diferente em cada uma das seguintes expressões; *make a cake*, *make a decision* e *make fun of*. Dessa forma, é importante ensinar esses verbos juntamente com seus colocados, ao invés de tentar identificar e distinguir significados básicos que, em alguns casos, pode ser difícil ou até mesmo impossível. Eu incluiria os VMs na lista de verbos que assumem um papel diferente dependendo de seus colocados, justificando a importância deste estudo.

A verificação dos padrões gramaticais de palavras e/ou expressões é (a exemplo das colocações e listas de frequência) relevante na construção de materiais didáticos para completar a fraseologia, ou padronização da palavra. A observação desses padrões é de especial importância no que tange os complementos verbais, ou seja, itens e estruturas que devem seguir, ou que geralmente seguem um verbo específico, tal como objetos ou verbos no infinitivo. Essa verificação pode responder a perguntas a cerca das formas que seguem certos verbos, especialmente quando a intuição falha. Kennedy (2002) mostrou que os VMs podem ocorrer em nove estruturas diferentes. Entretanto, o autor mostrou que há duas formas gramaticais predominantes em termos de frequência. Sendo que os outros padrões correspondem a menos de 10% das ocorrências dos VMs no BNC⁴⁷.

Através de um estudo baseado em corpus, pode-se verificar o vocabulário mais utilizado em certas estruturas gramaticais. Por exemplo, um corpus pode desvendar os verbos mais frequentemente utilizados na estrutura *past continuous* (*was-ing*). McCarten (2007) e Kennedy (1998) listam os dez verbos mais frequentes nesta estrutura: *going*, *thinking*, *talking*, *doing*, *saying*, *trying*, *telling*, *wondering*, *looking* e *working*. Os autores sugerem que o vocabulário deve ser ensinado na estrutura que ocorre de forma a maximizar o aprendizado.

Stubbs (1994) compara características presentes em vários tipos de livros com corpora de linguagem geral. Suas pesquisas buscam observar como os textos codificam suas posições ideológicas através do uso de características específicas. Um de seus estudos compara dois livros de geografia, um geral e o outro com um foco mais específico em ambientalismo. O foco do estudo são os verbos ergativos, isto é, verbos que podem ser usados de uma forma transitiva, intransitiva, ou passiva. Um exemplo é o verbo *close* (STUBBS 1994, p. 205):

⁴⁷ Este estudo de Kennedy (2002) serviu como ponto de partida para a presente tese e é referido novamente no Capítulo 4, sobre Modalização e no Capítulo 6, quando descrevo os dados deste trabalho.

several firms have closed their factories (transitivo)

factories have been closed (passivo)

factories have closed (intransitivo)

Conforme o autor, esse tipo de verbo pode ter um papel primordial na atribuição do agente, atribuindo, ou não, responsabilidade aos sujeitos. Os verbos ergativos são utilizados de uma forma transitiva muito mais frequentemente no livro com foco ambientalista (52%) do que no de geografia geral (23%). Todavia, essa proporção é oposta para as formas intransitivas, com 38% no livro ambientalista e 51% no inglês geral. Esses dados sugerem que as construções transitivas dos verbos ergativos podem não ser aleatórias no livro ambientalista, e que o uso dos verbos ergativos nessa forma pode estar servindo para atribuir responsabilidade aos sujeitos. Ao aprofundar a pesquisa, o autor compara seus dados com os do corpus LOB. O autor sugere que o uso dos ergativos no LOB é, de fato, semelhante ao do livro ambientalista, com 49% usado na forma transitiva e 31% na forma intransitiva. Stubbs reinterpreta a diferença entre os dois livros de geografia afirmando que, uma vez que o uso dos verbos ergativos no livro ambientalista é semelhante ao uso no inglês geral, não é o livro ambientalista que faz um uso demasiado das formas transitivas, mas sim o livro de geografia geral que utiliza demasiadamente as formas intransitivas. Os resultados demonstram que é através da tripla comparação que o analista poderá interpretar as ocorrências com maior precisão e definir as especificidades (BIBER et al.; 1998; GAVIOLI, 2005; STUBBS, 1994).

Hunston (2002, p. 198) sugere que para o professor de ESP, a questão de o “que ensinar” opondo-se a “como ensinar” pode ser de significância especial. Alunos que precisam ser ensinados a escrever artigos acadêmicos, devem estar conscientes quanto às características salientes de tais gêneros. A autora observa que o estudo baseado em corpus pode complementar os estudos de gênero (SWALES, 1991). Alguns estudos enfocam fraseologias típicas (GLEDHILL, 1995), enquanto outros estudos focam a maneira como o escritor relaciona-se com seus leitores (HYLAND, 1998)⁴⁸.

2.9 PONTOS DE DIVERGÊNCIA

Até o presente momento, apresentei as principais características da LdC com seus propósitos, suposições e bases teóricas. Contudo, considero importante apontar as diferentes perspectivas metodológicas que constituem as pesquisas baseadas em corpus.

⁴⁸ HYLAND, K. Hedging in Scientific Research Articles. Amsterdam: Benjamins, 1998.

Começo com a discussão sobre o status da LdC: seria uma metodologia ou um ramo independente da lingüística? Tognini-Bonelli (2001, p.1) defende que a LdC “vai muito além de um papel puramente metodológico” tendo se tornado uma “disciplina” independente. McEnery et al. (2006), por outro lado, sustentam que a LdC é uma metodologia, e não um ramo independente da lingüística como a fonética, a sintaxe e a pragmática, que descrevem ou explicam um certo aspecto do uso da linguagem. A LdC não se restringe a um aspecto específico e pode ser empregada na exploração de quase todas as áreas de pesquisa lingüística. Por ser um sistema complexo de métodos e princípios sobre a aplicação de corpora no estudo e no ensino/aprendizado de línguas, a LdC naturalmente “possui um *status* teórico” (ibid, p. 8). Entretanto, *status* teórico não é teoria *per se*. A metodologia qualitativa usada nas ciências sociais também possui uma base teórica e um conjunto de regras relacionadas, por exemplo, a como elaborar um questionário. Todavia, ela ainda assim é chamada de metodologia sob a qual teorias podem ser desenvolvidas. McEnery et al. (2006) afirmam que o mesmo acontece com a LdC. Considero que a LdC seja uma metodologia com uma vasta gama de aplicação através de muitas áreas e teorias da lingüística.

Há alguns outros “pontos de tensão” (McENERY e GABRIELATOS, 2006) entre diferentes perspectivas inseridas dentro do paradigma metodológico da LdC. Os principais pontos de divergência dizem respeito à relação entre teoria e dados, o uso de anotação (etiquetagem) e ao papel da intuição nas pesquisas lingüísticas. Para alguns pesquisadores, a formalização dessas divergências recai principalmente sobre a polarização *corpus-based* e *corpus-driven* (TOGNINI-BONELLI, 2001).

Segundo Tognini-Bonelli (2001), a abordagem denominada *corpus-based*, é utilizada principalmente para explicar, testar, ou exemplificar teorias e descrições que tenham sido formuladas antes do advento dos grandes corpora. Para a autora, essa abordagem possui alguns pontos fracos, como por exemplo, a base intuitiva. Outros pontos críticos seriam o isolamento dos dados que não se encaixam na teoria, ou ainda, fazer com que os dados se encaixem na teoria através da anotação dos dados de acordo com a teoria. Para Tognini-Bonelli (2001), essa abordagem parece favorecer o estudo da gramática, uma vez que não prescinde da utilização da etiquetagem morfossintática. McEnery e Gabrielatos (2006) sugerem que a abordagem *corpus-based* estaria associada ao trabalho de Leech (1991) ou Halliday (1991).

O paradigma *corpus-driven* é o de que qualquer afirmação teórica deve ser totalmente consistente e refletir diretamente a evidência fornecida pelo corpus. A versão forte dessa tendência está associada ao trabalho de Sinclair (1991). Esse tipo de pesquisa visa a descobrir fatos sobre a linguagem livre da influência de quadros teóricos pré-existentes, que, de acordo com os lingüistas que seguem essa tendência, se apóiam na intuição do falante sendo pouco abrangentes e nada confiáveis. Como consequência, as pesquisas são realizadas em corpora não anotados, uma vez que a anotação supostamente imporá uma taxonomia teórica restritiva sobre os dados. Outra característica é que não há uma distinção entre léxico e gramática⁴⁹, pois essa distinção também exigiria o uso de categorias lingüísticas pré-existentes que podem não ser corroboradas pelos dados provenientes do corpus. O ponto de partida na abordagem *corpus-driven* é sempre a palavra ortográfica, e não uma categoria gramatical ou um campo semântico, por exemplo.

O papel da anotação não alcança um consenso entre os pesquisadores. Uma das críticas refere-se ao fato de que a imposição de categorias baseadas em uma certa teoria lingüística pode gerar dificuldades para encontrar evidências que refutem essa mesma teoria ou ainda em descobrir características que a teoria não prevê. Leech (1997) a partir de um ponto de vista, e Sinclair (2004) do extremo oposto, questionam o valor da anotação, ou seja, se ela adicionaria valor a um corpus (Leech) ou se restringiria a quantidade de informação ao atribuir uma etiqueta inalterável nos casos em que uma palavra não pertence claramente a uma ou outra categoria lingüística (Sinclair). Hunston (2002) questiona ainda se os pesquisadores que fazem uso de corpora disponíveis estão “conscientes quanto às suposições teóricas que embasam diferentes formas de anotação” (ibid, p. 67).

Apesar dos pontos de controvérsia até aqui apresentados, há vários pontos de convergência no que tange os processos de anotação. Há um consenso quanto à necessidade de disponibilizar também uma versão original do corpus (não anotada). Da mesma forma, há um consenso sobre a importância de os pesquisadores estarem conscientes quanto aos princípios teóricos que regem uma anotação. McEnery e Gabrielatos (2006) sugerem a existência de categorias lingüísticas comuns a todas as teorias, uma vez que mesmo as pesquisas *corpus-driven* utilizam-se das chamadas categorias tradicionais tais como “verbo”, “preposição”, “objeto”, “voz passiva”, entre outras, sem necessidade de um esclarecimento sobre essas categorias. Para Hunston (2002), o uso dessas categorias indica que, mesmo

⁴⁹ Esse assunto está mais detalhado na seção 2.6 sobre Fraseologia e Colocações.

quando não há uma anotação automática *a priori*, há uma anotação manual *a posteriori*. Quanto ao papel da anotação, o próprio Sinclair, provavelmente o maior expoente contrário ao seu uso, concorda quanto a sua utilização quando a anotação tiver como objetivo alguma aplicação fora do mundo da lingüística, ou seja, traduções automáticas, dicionários eletrônicos, entre outros aplicativos. Entretanto, para Hunston (ibid), a utilização com sucesso de anotação em alguma aplicação, sugere que a teoria na qual essa foi embasada possui contribuições valiosas a respeito do uso real da língua.

Finalizando o debate quanto aos pontos críticos da LdC, alguns tipos de pesquisa parecem só serem realizáveis em um ou outro formato de corpus, ou seja, anotado ou não. Sobretudo, formas mais complexas de investigação, especialmente de investigação gramatical, só serão possíveis quando o corpus possuir algum tipo de anotação.

2.10 LIMITAÇÕES

Após apresentar as características e os benefícios do uso da metodologia baseada em corpus. Volto as atenções agora para as limitações dos estudos que seguem a metodologia da LdC.

Um corpus não consegue informar se algum fenômeno lingüístico é possível ou não, apenas se é freqüente ou não. Por um lado, as descrições lingüísticas (especialmente da língua inglesa) estão cada vez mais concentradas no que é típico, distanciando-se das noções de boa formação, ou correção (foco das pesquisas racionalistas) (SINCLAIR, 1991, p. 17). Entretanto, a pergunta “É possível dizer isso?” ainda necessita ser respondida. Para Hunston (2002), a intuição do falante nativo ainda é a melhor maneira de responder essa pergunta.

Um corpus não consegue mostrar nada mais além de seu conteúdo. Por mais representativo que um corpus proponha-se a ser, generalizações feitas a partir de resultados de um corpus são, na verdade, extrapolações. Uma declaração sobre um corpus é uma declaração sobre aquele corpus, e não sobre a linguagem ou registro o qual o corpus representa. Dessa forma, conclusões a respeito da linguagem inferidas a partir de um corpus devem ser tratadas como deduções, não como fatos.

Um corpus pode oferecer evidências, mas não pode fornecer informações. Por exemplo, o que *something of a* significa antes de um substantivo, em expressões do tipo *something of a surprise*? Presume-se ser um “mitigador”, *something of a surprise* é uma *small*

surprise (pequena surpresa). Ou seja, um corpus apenas fornece uma abundância de exemplos ao pesquisador, mas apenas o pesquisador pode interpretá-los (HUNSTON, 2002, p. 23).

Finalmente, e, conforme Hunston (2002), a falha mais grave do uso de um corpus é que ele apresenta a língua fora de seu contexto natural. Por exemplo, os manuais utilizados para o presente estudo, possuem figuras que, devido às limitações da tecnologia disponível, tiveram de ser descartadas. Em outras palavras, transcrições de dados orais não conseguem representar fielmente informações sobre entonação, linguagem corporal e outras características paralingüísticas. Esse fato aponta para a necessidade de um corpus ser apenas uma das ferramentas, entre outras, em um estudo lingüístico.

2.11 A OPÇÃO PELA LdC

Apesar das limitações da LdC, acredito que seja a metodologia disponível mais indicada para a averiguação de dados reais sobre a linguagem quando o objetivo é gerar subsídios para a elaboração de material didático. As principais razões para a escolha da LdC são as seguintes:

- A LdC constitui um método rigoroso e criterioso para a obtenção de dados atestados da língua “*in vivo*” em que é possível acessar um conjunto de dados reais e ricos no sentido de que, se o corpus for representativo de uma certa porção de linguagem, aparecerão, de forma clara, as unidades de comunicação mais utilizadas e as menos utilizadas. Além disso, pode-se acessar seus padrões semânticos, as associações que as palavras estabelecem entre si, suas colocações, as variações das unidades lexicais, entre outras características. Por ser científico, oferece a possibilidade da verificação dos dados.

- Os corpora “simplificaram” a vida dos lingüistas. Por exemplo, um lingüista que deseje verificar

o uso dos verbos modais, pode facilmente reunir todos esses verbos modais em um só lugar para a observação. O ato de reunir evidências é simplificado, liberando os esforços do pesquisador para o ato interpretativo. (HUNSTON, 2002, p. 214).

- Os estudos baseados em corpora abriram um novo leque de opções para as investigações lingüísticas, incluindo todo o tipo de trabalhos quantitativos. Essas investigações, como já vistas neste capítulo, podem ser utilizadas para a identificação de diferenças entre os registros ou gêneros.

- Os corpora mostraram que a língua é padronizada de uma forma muito mais detalhada do que sugerido anteriormente. Regras tidas como gerais, geralmente podem ser aplicadas somente em certos contextos. Como resultado, novas idéias sobre língua emergem e velhas idéias podem necessitar reavaliação.

- Afirmações mais objetivas podem ser feitas tendo em vista observações baseadas em corpora quando comparadas a observações introspectivas. Falantes nativos podem saber uma língua perfeitamente, mas nem sempre sabem o que eles dizem ou como o fazem. Da mesma forma, há uma discrepância entre o sentido intuitivamente priorizado e o mais freqüente.

-Os corpora provêm a possibilidade da “prestação de contas total” (*total accountability*) das características lingüísticas e não apenas de traços salientes individuais (*individual salience*).

-O estudo das colocações pode ajudar a organizar o contexto em padrões principais. Pode-se utilizar o conhecimento desses padrões para acessar o comportamento da língua ou os usos de palavras específicas no texto. Isso pode facilitar a diferenciação entre os significados de uma única palavra ou ainda determinar a variação de características sintáticas.

-Por fim, em consonância com Gavioli (2005, p. 56), acredito que, enquanto métodos baseados na LdC representam um desafio para a descrição do Inglês Geral, tais métodos parecem representar uma forma natural de descrever um inglês mais específico, uma vez que o ensino e a descrição de ESP baseiam-se quase que unicamente em coletâneas de textos especializados.

3 LINGUAGENS ESPECIALIZADAS: MÚLTIPLAS PERSPECTIVAS

Este trabalho se ocupa dos fenômenos relativos à comunicação em linguagens especializadas. À exemplo de outras teses e dissertações na área (FINATTO, 2001; MACIEL, 2001a; AZEREDO, 2007, etc.) farei primeiramente um breve histórico sobre as diferentes escolas da Terminologia e seus pontos de vista. Na seqüência, apresento uma revisão de estudos sobre texto e à linguagem especializada, o texto especializado e o texto técnico, e o texto de manuais técnicos. No final do capítulo, abordo aspectos relacionados à linguagem da aviação.

3.1 TERMINOLOGIA

Deve-se a Eugen Wüster da Escola de Viena⁵⁰, em meados dos anos 30, a origem da Teoria Geral da Terminologia (TGT)⁵¹. A disciplina, segundo essa escola, tem o objetivo de controlar e harmonizar os usos terminológicos em nível mundial, para garantir uma perfeita intercomunicação científica e técnica sem ambigüidades em um plano internacional (KRIEGER e FINATTO, 2004).

De acordo com Temmerman (2000), a TGT baseava-se em algumas premissas que seus seguidores acreditavam ser inquestionáveis: que conceitos são claramente definidos através de condições necessárias e suficientes (ao contrário de definições, que possuem um caráter *fuzzy*⁵²); que a univocidade dos termos é essencial para uma comunicação livre de ambigüidades, conseqüentemente eficiente e eficaz; e que a perspectiva da terminologia é unicamente onomasiológica⁵³. A teoria wüsteriana desempenhou um importante papel no intuito de estabelecer a Terminologia como “campo de conhecimento com fundamentos epistemológicos e objeto próprio de investigação” (TEMMERMAN, 2000, p. 32). Os objetivos da TGT motivaram, entre outras coisas, a criação do Comitê Técnico 37 da ISO

⁵⁰ Wüster era engenheiro, defensor do esperanto como linguagem universal.

⁵¹ Juntamente com a escola de Viena, cabe mencionar o papel pioneiro das escolas de Praga e da Rússia. As três escolas são consideradas clássicas e apresentam objetivos comuns, que visam a padronização de termos técnicos.

⁵² Caráter *fuzzy* refere-se ao fato de, por vezes, elementos não pertencerem claramente a uma ou outra categoria, situando-se, assim, em posições intermediárias e menos definidas.

⁵³ A perspectiva onomasiológica se contrapõe à perspectiva semasiológica. Com relação a dicionários, por exemplo, isso significa que, quando o usuário vai até a obra, ele pode encontrar dois tipos de organização. Por um lado, pode consultar o dicionário a partir de campos de significação, como a família, animais, esportes e partes do corpo. Nesse caso, o dicionário possui uma orientação onomasiológica, pois leva seu consulente de uma idéia que tenha em mente até uma palavra. Por outro lado, o dicionário pode ser organizado a partir dos verbetes, geralmente em ordem alfabética, uma orientação semasiológica. Nesse caso, o consulente é levado de uma palavra até sua idéia.

(*International Standard Organization*)⁵⁴. Entretanto, essa teoria não dá conta da descrição dos termos em seu habitat natural: o texto. Devido ao seu caráter altamente prescritivo e normalizador, a TGT reduziu a atividade terminológica à recolha de conceitos e termos para sua normalização na busca da univocidade. Os termos, compreendidos como unidades de conhecimento, e não como unidades lexicais, não são considerados elementos naturais das linguagens naturais. Esse reducionismo, com o conseqüente apagamento dos aspectos comunicativos e pragmáticos, é uma das razões pelas quais o “estudo dos léxicos especializados não alcançou o devido prestígio entre os estudiosos da linguagem” (KRIEGER, 2001, p. 50).

Cabré e seus colaboradores (Universidade Pompeu Fabra-Barcelona) estão entre os pioneiros a realizar uma crítica sistematizada e fundamentada à TGT. Entre as insuficiências encontradas na teoria da TGT, esses estudiosos apontam principalmente (CABRÉ et al., 1998): (i) O método de análise, que tenta ter uma validade universal, sendo aplicado por igual a diferentes contextos geográficos, realidades socioeconômicas, culturais e lingüísticas; (ii) O estatismo, com um modelo de representação de conhecimento marcadamente estático, fruto de uma perspectiva de estudo estritamente sincrônica, não integrando em sua análise nenhum elemento que dê conta do caráter dinâmico e evolutivo do conhecimento especializado e de suas denominações; (iii) O reducionismo, primeiramente aplicado somente a áreas técnicas, como a mecânica e a engenharia, não se atualizou para ter a capacidade de descrever as realidades díspares de variadas áreas do conhecimento; (iv) O idealismo, a TGT parece querer refletir um mundo idealizado em que os conceitos (que pré existem às línguas), são criados por consenso em um laboratório; objetos são etiquetados com termos de valores supra lingüístico e supra cultural.

A partir dessas críticas, surgiu a Teoria Comunicativa da Terminologia (TCT), que visa dar conta da complexidade das linguagens de especialidade com a diversidade e o dinamismo das temáticas de trabalho e de suas situações comunicativas também diversas. Assim, a Terminologia vista como disciplina lingüística deve oferecer resposta à descrição dos códigos e dos atos comunicativos especializados reais, à explicação do funcionamento da terminologia dentro da linguagem natural e à elaboração de aplicações terminológicas

⁵⁴ O Comitê Técnico da ISO de Terminologia e outros Recursos Lingüísticos (ISO/TC37) elabora normas técnicas sobre terminologia e seus produtos, serviços, processos e sistemas lingüísticos conexos. Estas normas servem às indústrias da língua, às atividades lingüísticas que daí resultam e a toda pessoa desejosa de fornecer produtos e serviços terminológicos e lingüísticos conexos. As normas da ISO/TC37 sobre os princípios, métodos e aplicações da terminologia são a base da normalização terminológica em todos os comitês ISO e servem de guia aos outros comitês que produzem a terminologia normalizada de certa área.

diversas que cubram necessidades comunicativas também diversas. Somando-se as idéias da TCT, Krieger (2001, p. 22) afirma que

o estudo dos termos, em muito favorecido pelo desenvolvimento das teorias de texto e de discurso que, ultrapassando os limites da frase, têm maior poder explicativo para dar conta da complexidade que envolve as unidades lexicais especializadas.

Nessa afirmação, Krieger ressalta a importância do texto, que extrapola a extensão do termo, ou da frase, para um melhor entendimento do funcionamento do termo técnico-científico.

Também em reação à corrente prescritiva da Terminologia Clássica, estudiosos da linguagem técnica iniciaram outro movimento na França. A perspectiva socioterminológica, iniciada por Boulanger (1991), surgiu para enfrentar os exageros prescritivos de algumas correntes normativas. Em 1993, Auger reitera a posição de Boulanger contrariamente a um ponto de vista puramente normalizador. Em dezembro de 1993, Gaudin concretiza a idéia de uma terminologia voltada para o social. A Socioterminologia preocupa-se com diferentes níveis de comunicação, dependendo dos interlocutores, dos canais de emissão, entre outros fatores. Prioriza o evento comunicativo analisando as manifestações discursivas. Os especialistas dessa vertente da Terminologia reconhecem a existência da polissemia e da variação. Os socioterminólogos levam em consideração também os discursos orais na elaboração de produtos terminográficos por entenderem que os termos variam de acordo com o seu contexto de uso e que essas variantes orais são de igual importância⁵⁵. A esses pesquisadores, uniu-se um grupo no Canadá, que já defendia idéias semelhantes. No Brasil, temos desde 1995, a socioterminologia como subárea de conhecimento na Universidade de Brasília, coordenada por Enilde Faulstich (FAULSTICH, 2006). Entretanto, Maciel (2001b) afirma que, apesar de terem contribuído enormemente para a Terminologia descritiva, a Socioterminologia ainda não conseguiu desenvolver uma proposta teórica consistente.

Temmerman (2000), incluída entre os estudiosos da Terminologia descritiva, lança a Teoria Sociocognitiva da Terminologia (TST). A premissa básica dessa escola é de que as palavras não possuem um significado objetivo, pelo contrário, as palavras (nesse caso, os termos) podem ser entendidas em um processo lingüístico comunicativo sobre uma realidade fora da língua, que também precisa ser compreendida. Nessa perspectiva, o entendimento de uma língua não pode ser separado do entendimento do mundo. A TST propõe a Teoria dos

⁵⁵ Diferentemente de outras correntes da Terminologia moderna, como, por exemplo, a TCT, que prioriza o discurso escrito e com alto grau de especialização.

Protótipos (ROSCH, 1972) para o processo de categorização, como alternativa às definições clássicas com características necessárias e suficientes. Temmerman (ibid) reconhece a importância do estudo diacrônico dos termos para melhor compreender a natureza sociocognitiva dos termos.

Seguindo a vertente descritiva, Finatto (2004) ressalta também a importância dos trabalhos de Rey (1979)⁵⁶, Hoffmann (1982)⁵⁷ e Gambier (1987)⁵⁸. Para esses autores, o uso lingüístico deve ser a porta de entrada dos estudos terminológicos, incluindo aspectos como variação, polissemia, diacronia, entre outros.

Apesar de ter os manuais técnicos de aeronaves como textos base, o presente estudo não se situa na área da Terminologia, pois não tem o Termo, ou a Fraseologia Especializada como objetos de análise. Este trabalho está em consonância com a Lingüística das Linguagens Especializadas, uma vez que o foco específico de estudo são elementos lingüísticos dentro de um texto especializado. Dessa forma, na próxima seção aprofundarei os conceitos de Linguagem Especializada e Texto Especializado.

3.2 LINGUAGENS ESPECIALIZADAS

Tendo em vista a perspectiva tomada nesta pesquisa, é importante explicitar algumas características das linguagens especializadas. Começo por mencionar a questão relativa à proliferação de termos utilizados para denominar a área.

Oppizzi (2006, p, 61) apresenta um breve panorama sobre o assunto em um capítulo intitulado “Sublinguagem, subcódigo, código restrito, linguagem restrita, microlinguagem, tecnoleto, linguagem especializada ou LSP?”. O objetivo da autora é chamar a atenção do leitor para a proliferação de termos utilizados em relação ao recorte de linguagem que constitui o objeto de estudo deste trabalho. Assim, trago de forma sucinta as considerações de Oppizzi.

O termo “sublinguagem”, foi introduzido por Harris (1968)⁵⁹ e desenvolvido por Hirschman e Sager (1982)⁶⁰ como a linguagem de uma área ou domínio específico, com um

⁵⁶ REY, A. *La Terminologie: Nomes et Notions*. Paris: PUF, 1979.

⁵⁷ HOFFMANN, L. Probleme und Methoden der Fachsprachenforschung. In: *Wiss. Zs. Der KarlMarx-Univ. Leipzig. Gesellschafts-und sprachwiss.* Reihe 31, 25-34, 1982.

⁵⁸ GAMBIER, Y. Problèmes terminologiques des pluies acides: pour une socioterminologie. *META*, Vol 32, n°3, pp 314-320, 1987

⁵⁹ HARRIS, Z. *Mathematical Structures of Language*. New York, 1968.

vocabulário específico e convenções de uso (também referida como “subcódigo”). Dodson (1974)⁶¹ refere-se a tais variedades como “códigos restritos”, enquanto Wallace (1981)⁶² as denomina de “línguas restritas”, mas ambos os termos somente podem ser aceitos com referência a códigos que usam uma quantidade reduzida da linguagem geral com a finalidade de uma comunicação especializada. Mesmo o termo “microlínguas”, não parece ser suficiente para dar conta da riqueza de expressão das línguas especializadas, enquanto os “tecnoletos” passam a errônea idéia de serem relacionados somente ao léxico especializado (OPPIZZI, 2006, p. 62). Cortelazzo (1994) considera a especialidade não somente em termos de necessidades referenciais (o léxico especializado), mas também em termos de linguagem como discurso, levando em consideração os interlocutores e suas necessidades comunicativas. Desta forma, “línguas especiais” são uma variedade da linguagem natural, dependentes da atividade especializada do conhecimento, usadas por um grupo restrito de falantes para satisfazer as necessidades comunicativas (com especial atenção ao léxico) daquela área. Todavia, segundo Gotti (1991), o termo “línguas especiais” identifica somente línguas que adotam um código não-lingüístico, com regras e símbolos diferentes daqueles da língua geral. Gotti (ibid) cita o “Código Q”⁶³ do ramo da tecnologia da informação, embora letras, números e pontuação ainda estejam presentes. Segundo Oppizzi (2006), Ambroso (1981)⁶⁴ sugere que o termo linguagem especial está antiquado, uma vez que não é a língua que é especial, mas o propósito que é específico. Conforme Oppizzi (2006), outra expressão atualmente adotada é “*Linguagem para propósitos específicos*” (LSP – *Language for Specific Purposes*), ou “Inglês para propósitos específicos”. Entretanto, a autora defende que o termo que melhor captura a linguagem dos especialistas em suas esferas profissionais é “Línguas Especializadas”.

⁶⁰ HIRSCHMAN, L; SAGER, N. *Automatic Information Formatting of a Medical Sublanguage*. Berlin/NY, 1982.

⁶¹ DODSON, C.F. *A Restricted Language. The Languages of Air Traffic Control*. PhD thesis. University of Leeds, 1974.

⁶² WALLACE, W.D., How Registers Register: Toward the Analysis of Language Use, *International Review of Applied Linguistics*, 19:4, 267-86, 1981.

⁶³ O código Q é adotado internacionalmente por Forças armadas e trata-se de uma coleção padronizada de três letras, todas começando com a letra "Q", inicialmente desenvolvida para comunicação radiotelegráfica comercial, e posteriormente adotada por outros serviços de rádios, especialmente o radioamadorismo. Embora os códigos Q tenham sido criados quando o rádio usava apenas o código Morse, eles continuaram a ser empregados depois da introdução de transmissões por voz. Para evitar confusão, sinais de chamadas têm sido frequentemente limitados de forma a restringir sinais começando com "Q" ou tendo uma seqüência de três Qs embutidos, por exemplo: “**QAP:Na escuta**”.

⁶⁴ AMBROSO, S. *L'insegnamento lingüístico per Scopi Speciali*. A. Ciliberti (ed.). Bologna, Zanichelli, 1981.

A linguagem especializada, segundo Oppizzi (2006) focaliza três elementos do ato comunicativo:

- (1) o produtor (autor, falante, redator) e o interlocutor (leitor, ouvinte, público alvo);
- (2) o campo específico;
- (3) o uso especializado da linguagem.

A autora afirma que esses três elementos são obrigatórios para decretar uma situação comunicativa especializada. Deve englobar linguagens originadas em diferentes épocas e lugares, que pertencem a diferentes áreas de especialização e a diferentes situações comunicativas, tais como: produtor (autor) e interlocutor (leitor) são especialistas; eles são especialistas e não especialistas; eles são especialistas e o público em geral. Portanto, a linguagem especializada não é unívoca, mas possui diferentes níveis de especialização. Apresento agora a visão de outros autores.

Sager et al. (1980, p. 69) se referem à “linguagem especial”, definem como “ sistemas semióticos complexos e semi-autônomos baseados e derivados da língua geral, seu uso pressupõe um treinamento específico e é restrito à comunicação entre especialistas da mesma área, ou de áreas semelhantes.” Ou seja, os usuários de uma língua especializada são especialistas (enquanto Oppizzi reconhece diferentes níveis de interlocutores) e as situações de comunicação são do tipo formal, reguladas normalmente por critérios profissionais e/ou científicos. Para os autores, as linguagens especializadas têm variadas aplicações, assim como a língua geral, mas tendem a “neutralizar usos sociais e emotivos enquanto maximizam os usos comunicativos e classificatórios” (p. 21). Além disso, essa linguagem tem como características informar com precisão, clareza e economia.

Lothar Hoffmann é um dos grandes nomes da Lingüística do Texto Especializado e, como as suas noções de linguagem especializada, texto especializado e ensino de linguagem para propósitos específicos coincidem com as deste trabalho, será uma das principais referências. Conforme Hoffmann, linguagem especializada é

O conjunto de todos os recursos lingüísticos que são utilizados em um âmbito comunicativo, delimitado por uma especialidade, para garantir a compreensão entre as pessoas que nela trabalham. Esses recursos conformam, enquanto sublinguagem, uma parte do inventário total da língua. (HOFFMANN, 1988[2004], p. 79)

O autor acrescenta ainda que a especificidade das linguagens especializadas é mais facilmente detectada pelo léxico, ou pela terminologia. Entretanto, Hoffmann considera

importante investigar diferentes categorias gramaticais, construções sintáticas e outras estruturas textuais, assim como particularidades na morfologia, grafia, pronúncia e classificação de seus signos gráficos. As linguagens especializadas constituem-se através de textos e não de simples listas de termos ou das relações gramaticais que os organizam em frases. São os textos produzidos por uma área que perfazem a dimensão comunicacional e semiótica dessa área.

A linguagem especializada é a própria língua natural na situação de uso especializado. Não se trata de uma língua diferente dos sistemas lingüísticos naturais, e sim de um uso peculiar da língua. Tal linguagem é utilizada por interlocutores de diferentes hierarquias e graus de especialização em níveis de formalidade distintos. Existem, entretanto, peculiaridades decorrentes dos propósitos de comunicação. Destacam-se fatores inerentes à natureza do campo de conhecimento e da atividade profissional que constituem o contexto de uma área. Esse contexto articula-se através do estatuto do destinador e do destinatário, da adequação dos propósitos de comunicação e dos aspectos circunstanciais do evento comunicativo.

Como já mencionado no Capítulo 2, Hoffmann (1988[2004], 1991, 1998[2007]) ressalta a importância da frequência de itens lexicais em uma linguagem, “a especificidade das linguagens especializadas se expressa principalmente pela frequência de uso de determinados recursos lingüísticos, comprováveis com o auxílio de métodos de estatística lingüística” (ibid, 1988[2004] p. 81). A “Estatística da Linguagem”, segundo Hoffmann, é uma disciplina integrante da Lingüística que se ocupa principalmente de aspectos quantitativos do uso da linguagem. Ela “[...] complementa a descrição qualitativa da linguagem através de informações sobre a frequência de ocorrências lingüísticas, o que é útil para certas aplicações, tais como [...] o ensino de línguas estrangeiras” (HOFFMANN, 1998[2007], p. 61). A estatística da linguagem busca um critério para a significância das diferenças entre as frequências dos fenômenos lingüísticos nos diferentes estilos funcionais da linguagem. A descrição quantitativa exata pode servir de base para considerações qualitativas. Contudo, o autor sugere que traços específicos de diferentes gêneros textuais ainda não recebem a devida atenção nas pesquisas lingüísticas e que a aplicação de categorias pré-estabelecidas pode prejudicar essas pesquisas. As categorias para análise deveriam surgir dos dados (HOFFMANN, 1998[2007]). Hoffmann, desta forma, defende de um modo bastante contundente os métodos da LdC para o estudo das linguagens especializadas.

As linguagens especializadas se ordenam segundo um grau de semelhança de uso de certos recursos lingüísticos, sendo importante comparar as “fronteiras de diferentes linguagens especializadas entre si e frente a outras sublinguagens”, pois, assim, “[...] temos como resultado uma tripartição” (HOFFMANN 1988[2004] p.82):

- a) há alguns recursos lingüísticos que aparecem em todas as sublinguagens ou variantes lingüísticas;
- b) há recursos lingüísticos que aparecem em todas as linguagens especializadas;
- c) há recursos lingüísticos que aparecem em somente uma língua especializada.

Para que essa tripartição seja conhecida, faz-se necessário a comparação de pelo menos três corpora diferentes. Esse resultado pode facilitar a aprendizagem de um léxico mínimo no ensino de línguas estrangeiras instrumentais, portanto, a comparação de três corpora (um geral, e dois corpora especializados) é o procedimento metodológico adotado neste trabalho.

3.2.1 O texto especializado e os gêneros textuais

Conforme Hoffmann (1988 [2004]), as linguagens especializadas constituem-se através de textos, de forma que se torna importante emitir umas breves palavras sobre o significado de “texto”. Definindo textos de uma forma geral, Crystal (1988, p. 254) afirma que: “os textos são considerados unidades da língua com uma função comunicativa definível, caracterizada por princípios como coesão, coerência e informação”. Para o autor, um texto pode se referir a material escrito ou falado. Um texto é mais do que a soma de suas partes, incluindo, entre outros aspectos, coerência e coesão.

Pela coesão, os elementos lexicais que compõem o texto se interligam mutuamente, compondo uma seqüência lógica de acordo com as convenções gramaticais. A coerência assegura que o conteúdo expresso nos conceitos e relações presentes no texto possam ser mutuamente recuperados a partir do universo do texto.

Segundo, Beaugrande e Dressler, (1981 [ed.rev.2002]), um texto é definido por sua ocorrência natural em um contexto de comunicação, ou seja, é um evento comunicativo. Como um evento comunicativo, o texto se caracteriza por um mínimo de sete características que conformam um padrão: coesão, coerência, intencionalidade, aceitabilidade, informatividade, situacionalidade e intertextualidade. Além disso, os autores sugerem alguns princípios que controlam a comunicação textual: a eficiência de um texto, que depende do seu

uso em uma situação de comunicação; a efetividade de um texto, que depende de causar uma boa impressão e atingir seu objetivo; e a apropriação, que depende da união entre o contexto e as formas pelas quais a textualidade é mantida.

Hoffmann (1988[2004]) considera que o texto especializado é como qualquer outro texto, e que possui todas as características expostas por Beaugrande e Dressler, (1981 [ed.rev.2002]). Além disso, para Hoffmann (1991):

O texto especializado é o instrumento, e ao mesmo tempo o resultado de uma atividade comunicativa realizada em conexão com uma atividade profissional específica; ele consiste em um conjunto finito e regular de orações ou elocuições pragmática, semântica e sintaticamente coerentes e coesivas, que, por serem complexos signos lingüísticos, correspondem a proposições complexas do pensamento humano e a complexos estado-de-coisas na realidade. (ibid, p. 161)

O texto especializado, devido às elevadas exigências de precisão de sua informação, se diferencia devido a particularidades de sua macroestrutura (articulação), por relações de coerência entre seus elementos e pela utilização de unidades sintáticas, lexicais, morfológicas e gráfico-fonéticas. Isso se realiza de modo variado para cada tipo de texto. Diferentes tipos de texto são, por exemplo, artigos, resumos, contratos, manuais, instruções, cartas, etc. Essa lista, segundo Hoffmann, é infinita. Seus elementos não são claramente definidos nem exatamente demarcados entre os diferentes tipos de texto. Hoffmann (1991) defende que as características entre os textos de diferentes especialidades devem ser descritas de forma qualitativa e quantitativa através de uma análise complexa, ou seja, através da integração de todas as características distintivas importantes em diferentes níveis lingüísticos desde a macroestrutura até os morfemas. O autor sugere os seguintes aspectos em forma de *check-list* (1991, p. 162, 163): (1) macroestrutura; (2) coerência, 2.1 pragmática, 2.2 semântica, 2.3 sintática; (3) sintaxe; (4) vocabulário; (5) categorias gramaticais, 5.1 verbo, 5.1.1 modo, 5.1.2 voz, 5.1.3 tempo, 5.1.4 pessoa, 5.2 substantivos, etc.

Entretanto, Hoffman (1991)⁶⁵ acredita que uma análise realmente complexa, além dos critérios lingüísticos estruturais acima, deveria examinar também as características comunicativo- funcionais, de acordo com um segundo *check-list*: (1) variáveis sociais, 1.1 existência de competência profissional (por exemplo, especialista/especialista x especialista/não especialista), 1.2 grau de competência profissional (por exemplo,

⁶⁵ Biber (1993, p.245), estabelece uma tipologia para tipos de corpora, semelhante a de Hoffmann (1991) mas não tão abrangente. Essa tipologia é apresentada na seção 2.8 deste trabalho.

professor/aluno); (2) intenção/função textual, 2.1 informação/descrição (por exemplo, artigo), 2.2 diretrizes (por exemplo, orientação), 2.3 clarificação/instrutivo (por exemplo, regra); (3) situação, 3.1 atividade primária (não lingüística, por exemplo, pesquisa), 3.2 canal (por exemplo, escrito), 3.3 convenções (por exemplo, padrão); (4) assunto/tema, 4.1 ciência (por exemplo, ciência natural), 4.2 área específica (por exemplo, física); 4.3 objeto (eletricidade).

Ciapuscio (2000) sustenta que os textos especializados são produtos predominantemente verbais de registros comunicativos específicos, que incluem temáticas próprias de um domínio de especialidade e que respondem a convenções e tradições retóricas, que dão lugar às classes textuais específicas. Eles podem ser mais ou menos dependentes de uma cultura ou época específica. Entre os textos especializados e não especializados pode haver zonas de transição pouco nítidas; portanto, é razoável considerar a existência de um continuum. No conceito de texto especializado se incluem, “além de informações impressas e escritas, também informações orais, diálogos e discussões, entre outras” (HOFFMANN (1988[2004], p. 88).

Como já visto no capítulo 2, vários estudos baseados em corpus têm a finalidade de examinar textos de diferentes especialidades, ou diferentes gêneros. Hunston (2002) sugere que as pesquisas baseadas em corpora podem, na verdade, complementar os estudos de gênero. Entretanto, esses estudos da LdC não oferecem uma definição do que seria um “gênero textual”, por essa razão descrevo brevemente a noção de gênero segundo Swales (1990).

Para construir sua própria definição de gênero, Swales (1990) pesquisou o conceito de gênero a partir do enfoque de diferentes disciplinas: estudos folclóricos, estudos literários, lingüística e retórica. Concluiu que havia pontos em comum na forma como cada área compreendia o conceito: (1) desconfiança em relação à classificação dos gêneros e a uma postura prescritiva na sua definição: os gêneros são entidades dinâmicas, passíveis de transformações de acordo com as condições sociais e históricas em que são produzidos, por isso não se prestam a classificações fáceis ou a estudos meramente prescritivos; (2) percepção de que os gêneros são importantes para integrar o passado e o presente: como construções históricas, os gêneros mantêm certa estabilidade, mas também evoluem para responder às necessidades de um mundo em mudanças; (3) reconhecimento de que os gêneros se situam no âmbito de comunidades discursivas, nas quais são importantes as práticas e crenças de seus membros: cada comunidade tem seus objetivos particulares e, assim, é a própria comunidade

que define quais gêneros são relevantes para a consecução desses objetivos; (4) ênfase no propósito comunicativo e na ação social: situado num contexto sócio-retórico mais amplo, o gênero funciona como um meio para a realização de propósitos comunicativos (e mesmo para a identificação de tais propósitos). Swales (1990) oferece, assim, a seguinte definição de gênero:

Um gênero compreende uma classe de eventos comunicativos, cujos membros compartilham os mesmos propósitos comunicativos. Tais propósitos são reconhecidos pelos membros especialistas da comunidade discursiva de origem e, portanto, constituem o conjunto de razões (*rationale*) para o gênero. Essas razões moldam a estrutura esquemática do discurso e influenciam e impõem limites à escolha de conteúdo e de estilo. (p. 58)

Esse conceito aponta que o principal traço definidor de gênero é o propósito comunicativo compartilhado pelos membros da comunidade discursiva na qual o gênero é praticado. Assim, os gêneros aparecem como veículos comunicativos usados para a consecução de determinados fins. A noção de comunidade discursiva diz respeito àqueles que trabalham usualmente ou profissionalmente com um determinado gênero e que, deste modo, têm um maior conhecimento de suas convenções (p. 54). A noção de comunidades discursivas deve ser diferenciada daquela de comunidades de fala (HYMES, 1974), nas quais necessidades comunicativas, tais como a socialização, tendem a predominar no desenvolvimento e manutenção das características discursivas. O comportamento lingüístico tende a ser socialmente determinado. Enquanto as comunidades de fala herdaram tipicamente seus membros através do nascimento, ou adoção, as comunidades discursivas recrutam seus membros através de treinamentos ou qualificações relevantes. Para Swales (1990), os gêneros textuais são propriedades das comunidades discursivas, ou seja, gêneros pertencem a elas. Uma das condições essenciais para fazer parte de uma dada comunidade discursiva é, portanto, dominar razoavelmente os gêneros que ela detém, é ser capaz de manejar as convenções comunicativas e pragmáticas dessa comunidade. Uma lista de características que definem as comunidades discursivas inclui, segundo Swales (1990, p. 25): (a) Mecanismos de intercomunicação entre seus membros- esses mecanismos variam de acordo com a comunidade: reuniões, correspondências, *newsletters*, publicações internas, etc.; (b) Léxico específico - os itens técnicos especializados são geralmente confusos a quem não pertence a uma comunidade; e (c) um nível de membros com um grau aceitável de conhecimento relevante e expertise discursiva - os indivíduos entram como aprendizes e saem quando morrem ou de outras formas menos involuntárias (d) indivíduos podem pertencer a várias comunidades discursivas.

Nas pesquisas com corpora, como mencionado anteriormente, não parece haver uma preocupação em definir claramente a noção de gênero. Os termos “gênero” e “registro”, são usados de forma intercambiável, McEnery et al. (2006) por exemplo, posicionam-se a esse respeito: “enquanto há outras definições possíveis de registro e gênero, nós adotamos uma definição imprecisa desses termos neste livro de forma que eles não sejam excludentes” (ibid, p. 87). Hunston (2002) apresenta uma opinião semelhante afirmando que registro e gênero são usados de uma forma “menos teoricamente precisa para significar simplesmente discurso que ocorre em um contexto específico” (p, 160). Kennedy (2002) comparou o uso dos VMs nas diferentes seções que compõe o BNC escrito, são eles: Prosa Imaginativa, Ciência Pura e Natural, Ciências Aplicadas, Ciências Sociais, Textos Jornalísticos, Comércio e Finanças, Artes, Crenças e Pensamentos, Lazer. Entretanto, o autor refere-se a essas seções chamando-as de gêneros. Neste trabalho ao me referir a outros estudos, usarei “gênero” de uma forma mais geral, como sugerido por McEnery et al. (2006) e Hunston (2002). Entretanto, ao me referir ao meu próprio estudo⁶⁶ utilizarei a noção de gênero proposta por Swales.

Consoante Swales (1990), Azeredo (2007) cunha o termo “macrogênero” (p. 86), que, segundo a autora, se caracterizaria por ser mais amplo do que um gênero textual, podendo ser comparado a um “guarda-chuva”. Essa imagem torna possível visualizar o macrogênero como algo maior, que abarca diferentes gêneros textuais. Opto aqui por entender o texto dos manuais técnicos como um macrogênero textual. É macro no sentido que inclui diferentes possibilidades de manuais. Os manuais de aeronaves, assim, acredito, são apenas um tipo de manual técnico.

3.2.2 A linguagem técnica

A linguagem técnica e a linguagem científica são muitas vezes tratadas como semelhantes. Contudo, existem diferenças entre elas devido às suas diferentes funções comunicativas e contextuais, como explicam Kittredge (1982) e Sager et al. (1980).

Kittredge (1982) considera tanto a linguagem técnica, quanto a linguagem científica como uma “sublinguagem”, mas o autor as distingue com base em características externas ao texto, de acordo com o caráter unidirecional das linguagens técnicas; isto é, nas linguagens técnicas, técnicos especialistas ou analistas (que são os redatores dos manuais) podem formar uma comunidade que estabelece normas, mas os usuários dos textos são geralmente menos experientes e oferecem pouco retorno aos redatores. Por outro lado, uma comunidade

⁶⁶ Capítulos 5, 6 e 7 e 8.

científica apresenta, geralmente, uma maior troca de informações, especialmente dentro de uma mesma área de especialidade.

Para Sager et al. (1980, p. 27), a ciência é essencialmente relacionada com a extensão do conhecimento, ao passo que a tecnologia relaciona-se com a aplicação do conhecimento. A tecnologia é expositória, educacional, instrucional. Seus atos de fala têm o foco no receptor, havendo sempre uma desigualdade de conhecimento. Os autores sugerem que um dos traços lingüísticos distintivos entre a linguagem da ciência e da tecnologia seriam os “atos de fala diretivos”⁶⁷, que estariam virtualmente ausentes da linguagem da ciência, mas que são muito freqüentes na linguagem técnica.

Com base em Sager et al. (1980) e Kittredge (1982), direciono agora as atenções para o tipo de texto do presente trabalho: o texto de manuais técnicos. É importante ressaltar que existe uma grande carência de estudos descritivos e sistemáticos a respeito de manuais técnicos. Há, entretanto, várias publicações que oferecem diretrizes sobre como redigir documentos técnicos, ou seja, que seguem uma abordagem prescritiva. Na próxima seção apresento dois desses guias. O primeiro (BLAKE e BLY, 1993) não se dedica a uma área ou domínio específico, mas a manuais técnicos de uma forma geral. O segundo, *Simplified English*, oferece diretrizes bem específicas para a elaboração dos manuais de manutenção de aeronaves. Considero importante incluir esses guias por acreditar que essas diretrizes tenham, em menor (BLAKE e BLY, 1993) ou maior (*Simplified English*) grau, influenciado as redações dos manuais aqui estudados. Contudo, abro a seção me dedicando a dois estudos descritivos: Trimble (1985)- que trata das principais funções comunicativas em manuais técnicos, e Lehrberger (1982)- que descreve uma seção de um manual de manutenção de aeronaves.

3.2.2.1 Os manuais técnicos: abordagem descritiva

Kittredge (1982) sugere que os manuais técnicos apresentam basicamente duas amplas funções: descrições e instruções. Dessa forma, apresento os apontamentos de Trimble (1985) com relação a essas funções.

Conforme Trimble (1985) há três tipos de informações descritivas em um manual: descrição física, descrição de função e descrição de processo.

⁶⁷ Os atos de fala diretivos são tratados em mais detalhes no Capítulo 4.

A descrição física fornece as características de um objeto e as relações espaciais das partes do objeto entre si e com o todo, e do todo com outros objetos. Características físicas incluem: dimensão, forma, peso, cor, textura, material e volume. Relações espaciais incluem: (a) para descrições gerais, posições indicadas por termos tais como *above*, *below*, *to the right*, *close to*, etc; (b) para descrições específicas, posições indicadas por termos tais como *perpendicular to*, *1 mm from the Center*, *at an angle of 38*, etc. Seguem os exemplos conforme Trimble (1985, p. 57-58):

(1) Geral:

We will describe a simple triode as a vacuum tube with three electrodes inside an evacuated glass envelope. Right in the Center will be one electrode, the cathode. It is a specially treated tungsten wire and is referred to as the filament. Surrounding the cathode and well spaced from it is the anode, or plate.

(2) Específica:

The test section was constructed of a pure copper cylinder 2 ft 6 in long, 6 in id and 6.25 in od. Both ends of the cylinder were closed with removable Pyrex-glass plates 1/8 in thick. A fluid port was located at each end of the cylinder.

A descrição funcional, de uma forma geral, fornece ao leitor informações relacionadas a um dispositivo. Ela descreve o propósito ou uso de um objeto e a maneira em que cada uma das partes e o todo funcionam, conforme os exemplos abaixo (TRIMBLE, 1985, p. 74).

Uso ou propósito:

(1) De uma parte:

The upper dial registers the ohms for each stage.

(2) Do todo:

The FBX3 calculates the incoming frequencies.

Função:

(3) De uma parte:

Pushing lever A contracts the spring.

(4) Do todo:

The purpose of the device is to collect bottom flora.

As descrições de um processo podem ser caracterizadas como um tipo de descrição funcional, entretanto há algumas diferenças. Assim, a descrição de um processo refere-se, por definição, a uma série de etapas que são interrelacionadas de forma que cada etapa (com exceção da primeira) dependa da etapa anterior e de forma que todas as etapas conduzam a um objetivo preciso. Elas podem ser apresentadas em forma de parágrafo ou em forma de *check-lists*, conforme exemplos abaixo (TRIMBLE, 1985, p. 74):

Em forma de parágrafo:

- (1) *By turning on the current teeth, or vanes, on the timer core pass by the teeth on the pole piece so that a magnetic path is established, causing a voltage pulse to be induced in the pickup coll. This voltage pulse causes TR-3 in the amplifier to conduct.*

Em forma de *check-list*:

- (2) *The procedure for checking the operation of the current limit range is as follows:*
- (a) *Set current limit range to 0.02A.*
 - (b) *Set the meter range switch to A.*
 - (c) *Connect an 8.2K, 10% 1W resistor across the output terminals.*

Quanto às instruções, os manuais fornecem aos leitores dois tipos de informação: instruções e informações instrucionais. As instruções podem ser de dois tipos: (a) instruções diretas, que são caracterizadas pelo uso da forma imperativa dos verbos; (b) instruções indiretas, que são caracterizadas pelo uso de **verbos modais**, **voz passiva** e, com muita frequência, a combinação dos dois; isto é, **modais passivos** (grifos meus). Trimble (1985) define instruções como “discurso que diz para alguém fazer algo” (p, 96). A informação instrucional é o “discurso que auxilia as instruções fornecendo informações adicionais: cuidados, avisos, especificações, descrições e considerações teóricas”. São sempre associadas com as instruções diretas e/ou indiretas.

Por exemplo, (TRIMBLE, 1985, p. 97, 98):

Instruções diretas:

- (a) *Connect the 382A and the other supply at the + and – sens and out terminals.*

Instruções indiretas:

- (b) *The rating of the fuse should be equal to, or slightly higher than, the rated full load current.*

Informações instrucionais:

- (c) *Cautions: This circuit will short out under any overload.*
- (d) *Warnings: Do not use the intake duct as a shell for tools.*
- (e) *Theory: the VRU mounting must be a rigid and flat horizontal plate.*

Antes de apresentar a próxima pesquisa, gostaria de mencionar que, de uma forma geral, parece haver uma maior atenção destinada aos manuais de manutenção de aeronaves (destinados a mecânicos) em detrimento dos manuais de operações (destinados a pilotos). Apesar de em número reduzido, ainda é possível encontrar obras que se dediquem à linguagem da manutenção, vide Shawcross (1992), com sua obra de cunho didático, intitulada *Aviation Maintenance*⁶⁸, Lehrberger (1982) (a seguir) e o *Simplified English* (no final do capítulo). As obras dedicadas às linguagens dos pilotos têm sempre o foco na comunicação oral, isto é, nas interações entre pilotos e controladores de voo, de importância incontestável. Entretanto, durante minha prática como professora de inglês técnico para pilotos, verifiquei que a linguagem escrita dos manuais de operações possui peculiaridades diferentes de outros manuais técnicos e que essa linguagem também impõe certas dificuldades de entendimento. Nesse sentido, acredito também ser relevante descrever os manuais de operações de forma a produzir obras de cunho didático voltadas para os pilotos. Esse motivo, entre outros, levou-me a desenvolver esta pesquisa.

Lehrberger (1982) conduziu um estudo com o propósito de descrever a seção que trata do sistema hidráulico nos manuais de manutenção de aeronaves. O objetivo final da descrição do manual, para Lehrberger, é construir um aplicativo que faça a tradução automática desse manual para a língua francesa. Para alcançar esse objetivo, é importante elaborar um *parser*⁶⁹. Referindo-se a esse estudo, Kittredge (1982) afirma que os manuais de aeronaves são elaborados segundo severas normas para a organização e a busca da não ambigüidade, ou seja, os redatores e usuários tendem a ser altamente treinados. Os manuais podem ser classificados de acordo com o sistema a qual se referem: hidráulica, trem de pouso, aviônica, etc.

⁶⁸ Cujos exemplos sobre os verbos modais são apresentados na íntegra no capítulo 4.

⁶⁹ Informações sobre *parsers* estão no Capítulo 2.

O corpus de Lehrberger (1982) contém 70.000 *tokens* com 3548 *types*⁷⁰. Entretanto, Lehrberger sugere que a quantidade de diferentes *types* na linguagem dos manuais de manutenção de aeronaves como um todo pode aproximar-se a 40.000. Quando comparado ao número de verbetes no dicionário *Webster's Third* (aproximadamente 450.000) pode-se afirmar que “o vocabulário dessa sublinguagem é bastante restrito” (ibid, p. 83). As funções comunicativas básicas nesse tipo de manual são descrever: as partes da aeronave, a manutenção de sistemas hidráulicos, sistemas elétricos, as ferramentas e equipamentos de teste exigidos para tal manutenção e fornecer instruções para a manutenção.

Algumas palavras são características desse assunto: *aileron, motor, compressor, jack, filter, check, axial, e quick-disconnect*. Por outro lado, outras palavras nunca ocorrem: *parsely, meson, seduce, endocrine, hope, think, believe*, nem os pronomes pessoais, *I, me, we, us, he, she*, são usados aqui (p. 83). Interrogativas diretas também não ocorrem, *Do you have your tool kit?*, ou passado simples *The engine stopped*. Por outro lado, há a abundante inclusão da voz passiva, orações relativas, nominalizações de vários tipos, etc. Apesar do “estilo telegráfico”, orações longas e complicadas são freqüentes, o que torna o parseamento extremamente difícil, por exemplo, (p. 83):

This unit contains the fuel metering section, shutoff valve, and a mechanical governor that function as either an over speed governor for the high pressure rotor or provides manual control when the electronic computer section of the fuel control system is deactivated.

Lehrberger nota que há uma redução considerável na polissemia. Em alguns casos isso resulta em uma palavra ocorrendo em apenas uma categoria na sublinguagem, enquanto a mesma palavra ocorre em várias categorias na linguagem geral, alguns exemplos dessa restrição semântica são, *case* (só substantivo), *fine* (só adjetivo) e *cake* (só verbo). Como um *parser* explora a possibilidade de designar uma estrutura a uma seqüência de palavras para cada categoria na qual as palavras ocorrem, uma redução no número de categorias as quais as palavras individuais ocorrem resulta em menos combinatórias e menos ambigüidade, por exemplo, (p. 85):

⁷⁰ Lehrberger, entretanto, contou os *types* de uma forma lematizada, ou seja, considerando a forma básica da palavra. Neste trabalho não utilizo lemas, e sim as palavras ortográficas como realmente aparece, consoante Sinclair e a escola da Terminologia Comunicativa de Cabré. A razão para minha opção é de diferentes formas de uma mesma palavra tendem a ocorrer em estruturas gramaticais diferentes e a terem diferentes colocados. Assim, acredito que se Lehrberger houvesse feito a opção de não usar a lematização, seus resultados poderiam ser mais eficazes quando da aplicação.

Check pump case drain fitting.

No inglês geral, cada uma das palavras acima, pode ser tanto verbo como substantivo, resultando em 32 caminhos a serem explorados. Naturalmente que o *parser* deve rejeitar uma das possibilidades. Mas uma vez que *case* não é usada como verbo no corpus, a palavra é listada no dicionário somente como substantivo. Esse fato reduz o número total de combinações a serem testadas de 32 para 16.

3.2.2.2 *Os manuais técnicos: abordagem prescritiva*

Apresento agora dois guias de redação para manuais técnicos: *The Elements of Technical Writing* (BLAKE e BLY, 1993) e o *Simplified English* (AECMA, 2004).

The Elements of Technical Writing

Como já mencionado, essa obra oferece um guia para a redação de vários gêneros textuais técnicos, como por exemplo, relatórios, cartas, memorandos, manuais, entre outros. Apresentarei de uma forma concisa as informações referentes à redação de manuais somente, por ser o ponto relevante para este trabalho. Blake e Bly (1993) abrem a sua obra sugerindo que a linguagem técnica é basicamente utilitária, priorizando a precisão em detrimento do estilo. O objetivo primário da linguagem técnica é “transmitir informação técnica com precisão” (p. 4).

Na seção dedicada especificamente aos manuais técnicos, os autores definem diversos manuais, entre eles os de (p. 144):

- (a) Operações: como e porquê um equipamento funciona na prática e na teoria. Tipicamente inclui diagramas esquemáticos, tabelas com dados operacionais, curvas performáticas e especificações;
- (b) Documentação de sistemas: como um sistema foi projetado, o que ele faz e quais os seus componentes.
- (c) Manual do usuário: como usar um software, um computador, ou qualquer outro sistema ou dispositivo. Geralmente ilustrado com figuras e telas.

Os autores sugerem 10 regras para melhorar a redação dos manuais técnicos, as quais resumo a seguir:

- (1) Lembre-se que redigir um manual significa redigir instruções (p. 145). Um bom redator de manuais sabe que sua missão é “dar instruções” e não vender, impressionar ou ofuscar o leitor com seu conhecimento técnico ou estilo. A grande maioria das pessoas não aprecia ler manuais técnicos, assim quanto mais for facilitada a leitura, melhor.
- (2) Seja completo (p. 145-146). Certifique-se de que as instruções estejam completas. É melhor pressupor que os leitores possuem menos conhecimento a cerca do assunto e fornecer mais detalhes do que pressupor mais conhecimento compartilhado e não fornecer informações suficientes. Visualize seu público e redija para aquele público. É uma boa idéia expor explicitamente, na introdução do manual, para quem ele se destina.
- (3) Seja claro e seja correto (p.147). O redator técnico deve esforçar-se para tornar os manuais claro, diretos e corretos. Poucas coisas são mais frustrantes para o dono de um produto do que um manual difícil de seguir ou com informações erradas; por exemplo, instruções que direcionem para que parafusos sejam colocados em certos orifícios, quando, na realidade, tais orifícios não existem.
- (4) Evite ambigüidades (p. 147). É melhor ser repetitivo e perfeitamente claro do que muito sintético e potencialmente confuso. Se por um lado é importante tornar um manual conciso, é ainda mais importante dar instruções que possam ser seguidas com segurança. A ambigüidade em materiais instrucionais faz com que o leitor sinta-se desconfortável e nervoso. Os leitores precisam estar seguros de que estejam fazendo a coisa certa.
- (5) Use advertências (*warnings*) (p. 148). Além de informar os leitores sobre “o que fazer”, redatores devem informar quanto ao que “não fazer”. Quando certas ações podem acidentalmente: apagar um disco rígido, causar um choque elétrico no operador, expor trabalhadores a gases tóxicos ou danificar equipamentos,- leitores devem ser explicitamente advertidos sobre tais riscos. As advertências devem ser destacadas colocando-as em caixa alta, ou outra técnica gráfica destacada. As advertências mais importantes devem ser emolduradas, por exemplo, (p. 149):

WARNING: The X-9 Control Unit should be unplugged from its power source before the front panel is removed and internal components are serviced. Otherwise, a severe electrical shock may result.

- (6) Use imperativo (p.150). Manuais são escritos no modo imperativo. Seja direto. É melhor escrever “*connect the communication line*” do que a forma mais fraca, mais passiva “*The communications line should be disconnected*”.
- (7) Escolha uma organização apropriada à tarefa (p.152). O esquema organizacional de um manual deve ser óbvio e transparente, através de sumários, introdução, índice remissivo, etc.
- (8) Apresente as instruções como uma série de etapas numeradas (p. 154). Se uma operação segue um claro esquema “passo-a-passo”, as instruções podem seguir uma numeração seqüencial, como por exemplo, (p.154):
1. *Set the machine operation mode to manual.*
 2. *Quickly depress and release the injection switch.*
- (9) Use uma abordagem modular (p.155). Em uma abordagem modular as seções são numeradas seguindo um sistema hierárquico, por exemplo:
- 3.4 *Unplugging the spray head.*
 - 3.4.1 *Operation of spray head self-cleaning water-jet feature.*
- (10) Faça um *test-drive* em seu manual (p. 155). O *test-drive* pode ser feito pedindo que alguns usuários testem as instruções nos produtos.

Simplified English

Apresento agora as características do *Simplified English* (SE). Estabeleço uma relação dessas diretrizes com a literatura da área da terminologia e das linguagens especializadas. O SE é uma iniciativa da AECMA⁷¹ que estabelece diretrizes lingüísticas que devem ser seguidas pelos redatores dos manuais de manutenção de aeronaves, ou seja, trata da normatização lingüística desses manuais. Considero importante abordar o SE uma vez que suas diretrizes influenciam a redação do MM (*Maintenance Manual*). Além disso, acredito ser relevante verificar até que ponto as diretrizes com relação aos VMs são seguidas no MM.

Conforme Hoffmann (1988[2004]), a normatização terminológica mostra claramente o “desejo de interferir de uma maneira reguladora sobre as relações entre sistemas conceituais e sistemas terminológicos e o desejo de uma configuração consciente das terminologias” (ibid, p. 86). A intenção dessa normatização é otimizar a comunicação especializada, eliminar mal-entendidos para, assim, garantir uma segurança maior na comunicação entre os especialistas. Para tal, são necessárias três ações: a) alteração da língua; b) unificação; c) implementação.

⁷¹ The European Association of Aerospace Industries;

Hoffmann sugere que as instituições nela envolvidas trabalham para a implementação de suas propostas através de normas, diretrizes e recomendações. Esse tipo de ação fez surgir “a base da Teoria Geral da Terminologia” (ibid, p. 87). Nesses pronunciamentos, Hoffmann estava se referindo somente a uma padronização no nível da terminologia. O SE, entretanto, extrapola esse foco somente terminológico e inclui vários outros aspectos dos manuais técnicos de manutenção, que serão descritos a seguir.

Para Farrington (1993), são várias as razões pelas quais uma forma padronizada e simplificada de inglês tornou-se necessária no contexto da aviação:

- (a) a crescente complexidade tecnológica em aviões modernos;
- (b) os manuais estão cada vez maiores, alguns chegando a 10 000 páginas;
- (c) o caráter altamente internacional de todos os grandes projetos aeronáuticos depois do Concorde;
- (d) o fato de que aproximadamente 80% dos clientes dos fabricantes de aeronaves serem falantes não-nativos de inglês.

O autor segue afirmando que a informação em documentos técnicos deve ser precisa, completa, relevante, concisa, convincente, significativa e livre de ambigüidade. O inglês convencional não seria adequado neste contexto devido ao alto número de palavras pertencentes á língua inglesa (95 mil palavras com 150.000 definições no *Oxford Concise Dictionary*) com uma quantidade imensa de significados (13 milhões de usos para 3 milhões de palavras no *Webster's Dictionary*).

O conjunto de regras e diretrizes que compõe o SE chega a mais de 500 páginas organizadas em forma de um guia de redação técnica, com o subtítulo: A GUIDE FOR THE PREPARATION OF AIRCRAFT MAINTENANCE DOCUMENTATION IN THE INTERNATIONAL AEROSPACE MAINTENANCE LANGUAGE. A primeira versão do SE foi lançada em 1986. Esta versão a qual estou me referindo foi publicada em 2004. O guia SE é composto por duas partes: a primeira estabelece regras gerais de redação; a segunda apresenta um dicionário com as palavras que devem ser usadas. Como já mencionado, o guia é bastante abrangente e não trata somente da terminologia, mas de todos os itens lingüísticos que devem ou não ser inseridos nas documentações de manutenção. O guia contém:

- Um número limitado de palavras;
- Um número limitado de significados claramente definidos para cada palavra;
- Um número limitado de estruturas para cada palavra;

- Um conjunto de regras de redação.

Devido a dimensão do guia SE, descreverei, de forma resumida, alguns aspectos que julgo importantes para este trabalho.

A primeira parte, *writing rules*, começa com uma citação de George Orwell⁷² que fala sobre a simplificação da linguagem. Logo após, o guia oferece as regras gerais para serem seguidas pelos redatores. Um aspecto a ressaltar é que o vocabulário usado nesse guia SE tem como público alvo os redatores dos manuais, que geralmente não são lingüistas. Dessa forma, terminologias lingüísticas são evitadas e, quando utilizadas, são acompanhadas por uma explicação de forma a facilitar o cumprimento da regra pelos redatores. Algumas dessas regras, apresentadas a seguir de forma bastante resumidas, são:

- (i) univocidade, ou seja, para um significado existe uma palavra, da mesma forma, para uma palavra existe um significado e ainda, para uma palavra existe somente uma função;
- (ii) uso somente da voz ativa, apenas um entre dez usos pode ser na voz passiva, *the gear box is moved by the engine* torna-se *the engine moves the gearbox*; além de vários outros princípios prescritivos que cobrem tamanho do parágrafo, formas e tempos verbais, número máximo de palavras por frase, pontuação, etc.
- (iii) Não usar o passado particípio de um verbo juntamente com o verbo modal (chamado de *helping verb* no SE), de forma a não tornar a frase mais complexa. O SE fornece os seguintes exemplos:

- (1) WRITE: ...*you can adjust*
NOT: ...*can be adjusted*
- (2) WRITE: ...*will adjust or ... adjusts*
NOT: ...*will be adjusted*
- (3) WRITE: *Adjust the...*
NOT: ... *must be adjusted*

Na seção que trata sobre os *helping verbs* e a voz passiva, há uma explicação de como detectar uma construção passiva e de como transformar uma estrutura passiva em uma ativa.

Na Parte 2 do SE, chamada DICTIONARY, há uma lista de palavras aprovadas (*approved words*) e seus significados e uma lista de palavras não aprovadas (*unapproved words*). Essas palavras são acompanhadas de exemplos que mostram como escrever uma

⁷² *If you simplify your English, you are freed from the worst follies of orthodoxy....And when you make a stupid remark, its stupidity will be obvious, even to your self.*

descrição ou um procedimento corretamente. Nesse dicionário, palavras aprovadas aparecem em letras maiúsculas, enquanto palavras não aprovadas são apresentadas em letras minúsculas. Segue a lista dos verbos modais contidos no manual do SE⁷³:

Quadro 3.2.2.2.A: Verbos modais no *Simplified English*

Keyword (part of speech)	Assigned meaning/USE	APPROVED EXAMPLE	Not Acceptable
CAN (v) CAN, COULD	Helping verb that “means” to be “able” to	YOU CAN CLEAN THE DRAIN HOLES WITH THE CLEARING TOOL.	
CANNOT	“Can not”	YOU CANNOT FILL THE CONTAINERS WHILE THEY ARE INSTALLED IN THE AIRCRAFT.	
May (v)	CAN, POSSIBLY (adv)	IF YOU USE INCORRECT EQUIPMENT, YOU CAN CAUSE DAMAGE TO THE VANES	The vanes may be damaged by using incorrect equipment.
MUST (v) NOTE: no other form of this verb	Helping verb that shows “obligation”	THE “BAT 1” SWITCH MUST BE IN THE “ON” POSITION	
shall (v)	MUST	HOLES MUST NOT HAVE SHARP EDGES.	Holes shall not have sharp edges.
Should (v)	Must, if (CON)	PERSONNEL MUST PUT ON PROTECTIVE CLOTHING. IF A FAILURE OCCURS, STOP THE TEST.	Personnel should wear protective clothing. Should a test occur, stop the test.
WILL (v) NOTE: no other forms of this verb	Helping verb that shows simple future tense	THE TEST ALTITUDES WILL BE HIGHER.	
Would (v)	CAN	SOLVENTS THAT STAY ON THE PART CAN CAUSE CORROSION.	Solvents left on the part would corrode the part.

Essa preocupação em padronizar somente a linguagem contida nos manuais de manutenção (para mecânicos) em detrimento dos manuais de operações (para pilotos), deve-se ao fato de que (entre outros fatores), de uma forma geral, os mecânicos ao redor do mundo possuem um nível de proficiência lingüística considerado bastante inferior ao nível dos

⁷³ Os VM não são apresentados em uma seção especial. Eles aparecem em ordem alfabética juntamente com as outras palavras do Dicionário SE.

pilotos. Em um seminário promovido pela ICAEA⁷⁴ realizado em Helsinki visando à discussão de todos os aspectos relacionados ao SE, participantes demonstraram interesse em ter o SE estendido também à área de operações (Shawcross, 1992). Entretanto, membros da AECMA alegaram que havia o risco de criar um vocabulário “muito amplo” para atender às necessidades de todos os usuários na área, conseqüentemente, não atendendo a nenhuma necessidade com eficácia. Contudo, considero haver a necessidade de um certo grau de padronização também na documentação de operações, pois o *background* sociolingüístico dos pilotos é também bastante diverso.

Apesar de o guia SE não trazer referências quanto à sua origem ou seus pressupostos teóricos, nota-se a forte influência dos princípios da TGT de Wüster na elaboração das normas e diretrizes. Os esforços concentram-se na normatização da linguagem na tentativa de excluir qualquer ambigüidade e variação de uso. Além disso, Shawcross (1992) menciona que não havia lingüistas⁷⁵ no grupo de trabalho que primeiramente elaborou as regras do SE. O autor menciona que o grupo foi composto por técnicos, redatores técnicos e tradutores. Ainda no seminário de Helsinki, houve questionamentos quanto ao critério utilizado para a seleção de vocabulário e ainda sobre quais as línguas e culturas haviam sido consideradas como modelos ao acessar a compreensão do SE. Os participantes alegaram que estruturas consideradas simples para um hispânico, não eram necessariamente simples para um japonês.

Apesar de não constituir o objetivo principal deste trabalho, considero relevante verificar até que ponto os três manuais estudados seguem as diretrizes acima. Acredito que o Manual de Manutenção deva seguir as regras do SE, pois é uma determinação do órgão que controla a linguagem da aviação (ICAO). Na próxima seção trato dos verbos modais que são as estruturas lingüísticas específicas estudadas nesta tese.

⁷⁴ *International Civil Aviation English Association.*

⁷⁵ Wüster era engenheiro. E a TGT pressupõe que os termos de uma área específica devem ser estudadas por profissionais daquela área.

4 OS VERBOS MODAIS

Nesta seção pretendo tratar de vários aspectos da modalidade: suas definições, classificações e características. A complexidade semântica dos VMs é estudada por vários pesquisadores, assim como são vários os aspectos abordados e analisados. Apresentarei algumas das perspectivas que considero mais relevantes para este trabalho, sem a pretensão de exaurir o tema, mas de mostrar a grande quantidade e variedade de pontos de vista e contribuir para um melhor entendimento do tema, cujos limites são, na verdade, desconhecidos.

4.1 MODALIZAÇÃO: DEFINIÇÕES E CLASSIFICAÇÕES

Conforme R. Lakoff (1973), a semântica dos modais é bastante complexa e “deve-se permitir a existência de ambigüidade na interpretação dos modais devido a diferenças de escopo envolvendo quantificadores entre os modais” (R. LAKOFF, 1973, p. 230). Segundo Palmer (2003), não há dúvidas de que a semântica⁷⁶ e a sintaxe da modalidade apresentam um dos maiores desafios para a análise gramatical e que os VMs em inglês apresentam, ainda, várias dificuldades idiossincráticas. Parret (1988, p. 97) confirma a idéia de Palmer, pois para o autor, “o sistema modal do inglês é bem mais complexo que o do português⁷⁷”. Quirk et al. (1985), Coates (1983), entre outros, também confirmam as afirmações de Palmer e Parret e classificam os VMs como uma das áreas mais problemáticas da gramática da língua inglesa. Como resultado, esses itens lexicais são particularmente problemáticos para

⁷⁶ Há uma ampla discussão na literatura envolvendo várias pontos semânticos em relação aos modais. Entre eles, destaca-se a questão do significado mais básico ou nuclear. Coates (1983) Sweetser (1990) e Talmy (1988, 2000) entre outros, acreditam ser o significado deontico o mais básico ou nuclear, pois surgiram antes nas línguas naturais. Da mesma forma, as línguas crioulas primeiramente desenvolvem expressões de modalidade raiz e apenas posteriormente estendem-se ao domínio epistêmico e que as crianças adquirem os sentidos deonticos antes dos epistêmicos. Palmer contra-argumenta, afirmando que, ao falar sobre o significado de criaturas animadas, tal como o *pé* de uma pessoa e o *pé* do morro (devido à localização), torna-se mais fácil estabelecer um “sentido literal” ou o significado básico; entretanto, tratando-se de modais, não “parece haver um sentido literal claro e qualquer significado nuclear deve ser deduzido” (Palmer, 1990, p. 15). Myhill e Smith (1995) sugerem que “parece não haver critérios objetivos para escolher uma função ou outra como básica”, ou, ainda, e o que considero mais importante nessa discussão, “qualquer motivação óbvia para organizar os dados dessa forma”. Ou seja, conforme Myhill e Smith (1995) a busca por um significado nuclear ou básico para os VM parece ser considerada infrutífera ou, até mesmo, inútil. Há também uma discussão relacionada à polissemia. Sweetser (1990, p. 50) menciona que alguns autores “tratam os verbos modais em inglês, essencialmente como casos de homonímia e não de polissemia” pois as modalidades epistêmica e de raiz não seriam de forma alguma relacionadas”. Contudo, a autora considera os verbos modais como casos de polissemia, argumentando que eles são metaforicamente relacionados. Da mesma forma, Coates (1983) e Palmer (1990) consideram os VM como polissêmicos. Silva-Covalan (1995) defende uma posição monossêmica, principalmente para o verbo *poder* em espanhol. Sarmiento (2007) aplicou os testes diagnósticos de polissemia propostos por Ravin and Leacock (2000) para verificar os significados do modal *can* e conclui que é um caso de polissemia.

⁷⁷ O autor, na verdade, compara o sistema modal do inglês com o do francês, mas na tradução para o português, o tradutor usou a comparação entre o inglês e o português.

aqueles que pretendem aprender ou ensinar a língua inglesa. Há, conforme Oliveira (2003), movimentações semânticas dos VMs sobre dois eixos. Sobre um eixo, um único item lexical, por exemplo *could*, pode expressar habilidade, permissão, assim como possibilidade. Por outro lado, esses mesmos significados poderão ser expressos por outros VMs, ou seja, a idéia expressa por *could*, poderá por vezes ser transmitida por *might*, *can* ou por *may*, entre outros.

Além disso, Trimble (1985) declara que há uma troca nos significados dos VMs especialmente nos textos técnicos relacionados a instruções. Nesses textos, os VMs parecem adquirir significados diferentes daqueles geralmente ensinados nas gramáticas. Ao analisar o desempenho de leitura em alunos, Trimble (ibid) constatou que os redatores de textos técnicos e científicos têm certas suposições relacionadas ao tipo e quantidade de informação retórica-gramatical que eles supõem que os leitores compartilhem com eles, ou seja, informações que os leitores trazem para a leitura de um texto técnico. Entretanto, grande parte dessas suposições parece ser válida para o falante nativo de inglês, mas não para o falante não nativo. Sua pesquisa mostra que a maioria dos alunos não nativos (de inglês) não possui o conhecimento cultural apropriado que os habilite a trazer mais do que uma limitada quantidade de informações pressupostas para auxiliar na leitura dos textos técnicos. Em outras palavras, os alunos não-nativos tendem a transferir suas técnicas de leitura desenvolvidas para o inglês geral (e para a sua própria língua materna) para o texto técnico, sem realizar os devidos ajustes necessários. Cabe lembrar que os manuais de aviação estudados neste trabalho são redigidos nos Estados Unidos, possivelmente por redatores americanos, para serem lidos por pilotos e mecânicos do mundo inteiro.

Entre as diferentes perspectivas de estudos dos VMs, Coates (1983) cita os estudos de Boyd e Thorne (1969)⁷⁸, Halliday (1970)⁷⁹, Johannessen (1976)⁸⁰, Marino (1973)⁸¹, entre outros sob o viés da teoria semântica. A principal semelhança entre as metodologias supracitadas é a quase total dependência de dados introspectivos, até mesmo para citar exemplos, que dessa forma tendem a ser “claros” e “limpos” para elucidar uma dada teoria. Todavia, já é possível encontrar uma variedade de estudos de cunho empírico-descritivo, principalmente utilizando a LdC. Entre os mais extensos e significativos pode-se citar Coates (1983), Palmer (1986), Quirk et al. (1985), Biber et al. (1999), entre outros. Nesses últimos

⁷⁸ BOYD, J. E THORNE, J.P. The Semantics of Modal Verbs, *Journal of Linguistics* 5, 1, p.192, 1969.

⁷⁹ HALLYDAY, M.A.K. Functional Diversity in Language as Seen From a Consideration of Modality and Mood in English, *Foundations of Language*, 4, 225-242, 1970.

⁸⁰ JOHANNESSEN, N-L. The English modal auxiliaries : a Stratifictional Account. *Stockholm : Almqvist & Wiksell International, 1976.*

⁸¹MARINO, M. A Feature Analysis on the English Modals, *Lingua*, 32, 309-323, 1973.

estudos, o papel do analista é o de interpretar os dados provenientes dos corpora, e sob o rigor da LdC (Linguística de corpus), dar conta de “todos” os casos encontrados no corpus⁸², incluindo aqueles que não pertencem claramente a uma ou outra categoria. Ou seja, classificar exemplos menos claros e limpos.

Apesar das diferenças metodológicas entre o primeiro e o segundo tipo de estudos (no qual o presente estudo se insere), as pesquisas de ordem semântica oferecem valiosa contribuição para os estudos baseados em corpus (incluindo este trabalho), pois oferecem importantes constatações a respeito dos VMs, sendo que muitos estudos baseados em corpus ainda utilizam as categorias prescritas pelos estudos de ordem semântica, mas com a diferença da inclusão de exemplos autênticos que contribuem para reconhecer a fragilidade e a insuficiência da rigidez das categorias semânticas pré-estabelecidas. O reconhecimento dessa fragilidade e insuficiência dá origem à flexibilização no estabelecimento entre os limites das categorias, ou seja, o reconhecimento de limites *fuzzy*. Entretanto, o surgimento da LdC ainda não parece ter promovido uma suficiente quantidade de estudos detalhados, de ordem quantitativa e qualitativa que estabeleça a inter-relação entre a distribuição, padrões gramaticais, colocações, usos e significados dos VMs; em outras palavras, estudos que dêem conta da fraseologia⁸³ dos verbos modais, como sugerido por Hunston (2002), ou da “padronização” dos VMs, como denomina Berber Sardinha (2004)⁸⁴, em outras palavras, estudos que busquem estabelecer uma taxonomia dos VMs a partir dos dados (*bottom-up*). Este trabalho pretende contribuir nessa direção. Retomo agora as definições e classificações de modalização.

Modalidade, em um sentido amplo, é geralmente relacionada às noções de “necessidade” e de “possibilidade” (LYONS, 1977). São várias as propostas relativas ao

⁸² Estou aqui me referindo aos estudos dos VM baseados em corpus como sendo um grande grupo homogêneo de estudo. Entretanto, entre esses estudos, é também possível perceber diferenças metodológicas e teóricas. Primeiramente, há os estudos de Palmer (1986), que apesar de serem baseados em corpus, não levam em conta todos os exemplares de uma unidade linguística, isto é, os dados provenientes do corpus serviram para propósitos heurísticos e para cunhar exemplos. Palmer também não faz uma análise distribucional. Coates (1983 e 1995), apesar de fazer amplo uso de um corpus de inglês geral, recorre, por vezes, às categorias semânticas tradicionais pré-existentes provenientes dos estudos mais introspectivos para analisar seus dados. Naturalmente que Coates oferece uma valiosíssima contribuição aos estudos dos VM por seu pioneirismo, extensão e por ter lançado mão do conceito *fuzzy* de classificação, que a autora define como “uma categoria na qual a transição entre membro e não membro é gradual, ao invés de abrupta”. Myhill e Smith (1995, ainda nesta seção), entretanto, apresentam uma detalhada classificação (apesar de ser um estudo bem menos extenso que Coates (1993)) da modalidade deontica, cujas categorias de análise emanaram do próprio corpus. Aproximam-se de um estudo mais próximo do que pode ser chamado de *corpus driven grammar*.

⁸³ Como visto no capítulo sobre Linguística de Corpus, fraseologia é aqui entendida como a inter-relação entre o significado, a forma e o uso.

⁸⁴ Berber Sardinha (2004) padronizou, por exemplo, a partícula “só” na língua portuguesa.

número, tipos e nomes de modalidades. Alguns exemplos são as modalidades: *factual X theoretical* (LEECH, 1971, p 75-106), *truth-functional X non-truth-functional* (LAKOFF, 1972, p.232), *subjective X objective* (LYONS, 1977, p. 797). Há ainda referência às modalidades “aléticas”, “bulomaicas”, “temporais”, “avaliativas”, “causais” e “probabilísticas” (NEVES, 2006). Entretanto, a maior parte dos lingüistas reconhece a existência de pelo menos dois grandes grupos: a modalidade epistêmica e a modalidade de raiz, que engloba a deôntica e a dinâmica. (COATES, 1983; PALMER, 1986, 1990, 2003; NEVES, 2006; SILVA-CORVALÁN, 1995, entre outros). Alguns exemplos na língua inglesa dessas modalidades são (PALMER, 2003, p. 7):

Epistêmica: *They may be in the office. – They must be in the office.*

Deôntica: *They may/can come in now. – They must come in now.*

Dinâmica: *They can run very fast. – I will help you.*

Conforme os exemplos acima, a modalidade epistêmica expressa a crença dos falantes e seus diferentes graus de convicção a respeito do estado das coisas, permitindo também que os falantes expressem seus julgamentos sobre as possibilidades. Dessa forma, nos exemplos acima, o falante faz julgamentos sobre a possibilidade ou necessidade (necessariamente o caso) de que os indivíduos estejam no escritório (*in the office*). A modalidade de raiz pode ser subdividida entre deôntica e dinâmica. A modalidade deôntica, expressa uma permissão ou injunção do falante para que um ato seja realizado ou para que uma obrigação seja cumprida (HOYE, 1997). O evento é controlado por circunstâncias externas ao sujeito da oração ou, em um sentido mais restrito, à pessoa ou pessoas identificadas pelo sujeito. Nos exemplos acima, permissão é conferida com *may*, e obrigação com *must*. A dinâmica, por sua vez, trata da habilidade/capacidade ou disposição do sujeito da oração (PALMER, 1990). O controle é interno ao sujeito, como nos exemplos acima, em que a habilidade do sujeito de correr rápido (*run fast*) é expressa com *can* e a disposição do sujeito de ajudar, com *will*.

Na tentativa de diferenciar as modalidades epistêmica e deôntica, Parret (1988) refere-se a eixos semânticos, sendo o eixo epistêmico o da crença e o eixo deôntico o da obrigação, como explicitado no quadro a seguir:

Quadro 4.1.A: Diferenças entre as modalidades epistêmica e deôntica

Epistêmica	Deôntica
Certa	Obrigatória
Contestável	Facultativa
Plausível	Permitida
Excluída	Proibida

(PARRET, 1988)

Dessa forma, Parret apresenta as diferentes forças de cada tipo de modalidade, com a epistêmica, indo de “certa” até “excluída” e a deôntica indo de “obrigatória” até “proibida”. Cabe destacar que, assim como em muitas línguas, em inglês, as mesmas formas lingüísticas podem expressar tanto os significados de raiz quanto os epistêmicos, com supleção de alguns itens lexicais para a negação⁸⁵ conforme os quadros 4.1.A e 4.1.B.

Quadro 4.1.B: Itens lexicais - modalidade epistêmica

Verbos modais epistêmicos	
Possível	<i>May</i>
Não possível	<i>Can't</i>
Possível que não	<i>May not</i>
Necessário	<i>Must</i>
Não necessário	<i>May not</i>
Necessário não	<i>Can't</i>

(PALMER; 1995, p. 456)

Quadro 4.1.C: Itens lexicais - modalidade deôntica

Verbos modais deônticos	
Possível	<i>Can/May</i>
Não possível	<i>Can't/May not</i>
Possível que não	<i>Needn't</i>
Necessário	<i>Must</i>
Não necessário	<i>Needn't</i>
Necessário não	<i>Mustn't</i>

(PALMER; 1995, p. 457)

A modalidade epistêmica expressa “a opinião ou atitude com relação à proposição que a sentença expressa ou a situação que a proposição⁸⁶ descreve” (LYONS, 1977, p. 452). Palmer (1986, p. 16) afirma que essa modalidade “diz respeito às características subjetivas de um enunciado”, e que poderia ser ainda sugerido que a “subjetividade é um critério essencial para a modalidade.” Segundo Neves (2006, p. 160), a modalidade epistêmica “depende de como o mundo é.” Está relacionada ao conhecimento e a força com que o falante acredita na veracidade de uma proposição.

⁸⁵ A problemática de negação será tratada em separado a seguir.

⁸⁶ Uma proposição é o significado, o sentido, a intenção ou o conteúdo informativo de uma frase declarativa.

A noção de *hedges* é freqüentemente associada à modalização epistêmica. Assim como com as definições de modalização, as de *hedges* também variam muito em escopo. Foi possivelmente Lakoff (1972) o criador do termo. Todavia, Lakoff não estava interessado em seu valor comunicativo, mas em suas propriedades lógicas de palavras e sintagmas como *rather, largely, in a manner of speaking, very*, ou seja, em como “tornar as coisas mais ou menos *fuzzy*” (Lakoff, 1972, p.195). Para Palmer (1986, p. 63), os *hedges*, como *I think, certainly, possibly, perhaps, it seems to me* etc., são exclusivamente elementos instanciadores da modalidade epistêmica. Esses elementos são freqüentemente encontrados em textos acadêmicos,⁸⁷ mas não em manuais como o texto base deste estudo.

A modalidade deôntica implica que o ouvinte aceite o valor de verdade do enunciado para executá-lo. Engloba significados como permissão e obrigação, assim como possibilidade e necessidade. Dessa forma, a modalidade deôntica é relacionada aos atos de fala (AUSTIN, 1962; SEARLE, 1969⁸⁸): convidar, pedir, dar uma ordem, etc. É a forma “como um ato é socialmente ou legalmente circunscrito” (NEVES, 2006, p. 162). A modalidade deôntica pode ser parafraseada como “obrigatória” e “permitida”, como por exemplo:

- *You **may** go home now.*

A modalidade dinâmica, por sua vez, relaciona-se a como os “referentes de sintagmas nominais de função sujeito são dispostos em direção a um ato, em termos de habilidade e intenção” (NEVES, 2006, p. 162). Os graus dessa modalidade são parafraseados como volição e habilidade, como por exemplo:

- *He **can** play the piano very well.*

- *I **will** help you.*

Vários estudiosos defendem a exclusão do sentido dinâmico da classificação dos modais. Todavia, para Palmer (1986), a modalidade dinâmica não envolve nem a atitude nem a opinião do falante, mas mesmo assim considera importante mantê-la na tipologia das modalidades lingüísticas por considerá-la importante para o significado dos VMs.

Entretanto, devido a sua complexidade, a categorização de um VM pode tornar-se difícil tendo-se em mente as categorias semânticas estanques e pré-estabelecidas. Assim, Sarmiento (2007), consoante Coates (1983) e Palmer (1990) defendem a proposta de

⁸⁷ Vários estudos analisam a presença de *hedges* na linguagem acadêmica, como por exemplo, Gabrielatos e McEnery (2005).

⁸⁸ Os atos de fala serão novamente referidos na seção a seguir sobre o uso dos VM.

Wittgenstein (1953, in Aarts et al., 2004) denominada “semelhanças de família” (*family resemblances*) para uma categorização semântica dos VMs. Para Palmer:

É mais provável que exista um conglomerado de significados vagamente relacionados, cada qual ligado de alguma forma a pelo menos um dos outros significados, mas não necessariamente compartilhando qualquer característica em comum, ou diretamente ligado a todos os significados.

(PALMER, 1990, p. 15)

Conforme Myhill e Smith (1995), os lingüistas têm tradicionalmente feito distinções entre obrigações fortes e fracas, e, dessa forma, atribuído certos itens lexicais a esses tipos de obrigações. Em inglês, por exemplo, *must* e *have to* são associados à obrigação “forte”, enquanto *should* e *ought to* são associados à obrigação “fraca”. Palmer (1986) declara que a diferença entre as duas gradações de obrigação é que na “fraca” (*should/ought to*), o falante admite a possibilidade de o ato não se concretizar, enquanto na “forte” (*must*) o falante acredita na realização do ato. Myhill e Smith (1995) mencionam ainda os parâmetros “orientação discursiva” (PALMER, 1974), “subjatividade” (COATES, 1983; LYONS, 1977) e “envolvimento do falante” (COATES, 1983). Contudo, Myhill e Smith (1995) afirmam ter dificuldades em aplicar esses critérios de forma objetiva a dados de ocorrência natural e real. Na maior parte das ocorrências, a única evidência da força da modalização seria o próprio morfema, o que, para os autores, caracterizaria a explicação como “circular”. Dessa forma, os autores propuseram critérios alternativos para a análise das expressões relacionadas à obrigação⁸⁹ e sugerem que, além de somente registrar obrigação, essas unidades lingüísticas também expressam a avaliação do falante sobre o efeito do evento. Os critérios de classificação são baseados em se alguém é negativamente afetado pela ação expressa no sintagma verbal, e, caso, positivo, quem é negativamente afetado, falante, ouvinte ou sujeito da oração⁹⁰.

Ao conduzir esse estudo interlingüístico (inglês, chinês, hopi e hebreu bíblico), cujo objetivo principal era responder à pergunta “Em quais contextos os falantes de diferentes línguas utilizam um marcador lingüístico de obrigação?” (Ibid, p. 240), Myhill e Smith obtiveram duas importantes conclusões quanto à obrigação deôntica; a primeira é a de que “a

⁸⁹ As formas em inglês analisadas por Myhill e Smith (1995) são *have to*, *had to*, *gotta*, *should*, *should have*, *is/are supposed to*, *was/were supposed to* e *ought to*.

⁹⁰ Myhill e Smith (1995) obtiveram a seguinte classificação relacionada às categorias funcionais dos modais que expressam obrigação (Ibid:249): (a) Efeito negativo ao sujeito; (b) Inconveniência mitigada ao ouvinte; (c) Efeito negativo ao ouvinte; (d) Efeito negativo ao falante; (e) Reprimenda; (f) Contra-factual; (g) Ajuda não necessária; (h) Sugestão; (i) Ação sugerida; (j) Procedimentos e rotinas; (k) Necessidades biológicas/fisiológicas; (l) Não-codificado.

única forma de analisar suas funções é em termos de uso, e não de significado” (p. 284) e a segunda, “o uso de marcadores lingüísticos de obrigação é, obviamente, altamente relacionado à organização de uma sociedade” (p. 285). Considero essa última constatação de especial interesse para o presente estudo, uma vez que os manuais estudados são escritos pela sociedade americana para serem lidos por profissionais de várias partes do mundo e conseqüentemente, pertencentes a diferentes sociedades. Esse estudo (Myhill e Smith, 1985) corrobora as constatações de Trimble (1985) relacionadas às dificuldades que falantes não nativos de inglês parecem apresentar ao ler textos técnicos. Tanto Myhill e Smith (1985) quanto Trimble (1985) relacionam os VMs deônticos a informações culturais. Dessa forma, pode haver uma dificuldade na interpretação do ato de fala por falantes não nativos de inglês.

4.2 OS VERBOS MODAIS EM INGLÊS

A modalidade pode ser expressa por diferentes unidades lingüísticas. (OLIVEIRA, 2003; COATES, 2005; NEVES, 2006). Do mesmo modo, uma mesma forma pode expressar diferentes modalidades, ou seja, os conceitos modais podem ser expressos nas línguas naturais através de uma grande variedade de formas. Na língua inglesa, a modalização, tanto a epistêmica quanto a de raiz, pode ser expressa através de verbos como *know, believe, allow, have to, etc*; advérbios tais como *possibly, necessarily, probably*; adjetivos como *possible, probable, capable*, advérbios, tais como *certainly e possibly*, formas imperativas, entre outras. Há também alguns tempos gramaticais que podem criar situações alternativas, como é o caso do futuro e do condicional. Entretanto, a modalização, especialmente na linguagem escrita, acontece principalmente através dos auxiliares modais⁹¹, como *may, can, shall, must* (CRYSTAL, 1985; LYONS, 1977; GABRIELATOS e MCENERY, 2005).

Os VMs em inglês formam uma subclasse dos verbos, com características sintáticas, semânticas e pragmáticas próprias. Para melhor definir a classe dos VMs, ou entender os ambientes nos quais um VM é usado, deve-se lançar mão de vários níveis lingüísticos; (i) o puramente sintático, (ii) o semântico, (iii) o pragmático, ou seja, a função do VM, as suposições compartilhadas entre remetente e destinatário, a situação social adotada pelos participantes no discurso, a impressão que o falante quer causar no seu interlocutor e (iv) o fator distribucional, entre outros.

⁹¹ Nesta tese, “auxiliar modal” será utilizado como sinônimo de “verbo modal”

4.2.1 A forma

Não há um consenso sobre quais unidades lexicais incluir na categoria ‘verbo modal central’. Palmer (1990:3) considera, *will, shall, may, can, must* e *ought to* como os verbos modais centrais, classificando *dare* e *need* como mais periféricos. Biber et al. (1999), Kennedy (2002) e Coates (1983) categorizam *can, could, will, would, may, might, shall, should* e *must* como modais centrais.

Os VMs em inglês possuem características bastante específicas no que diz respeito à forma, ou seja, possuem um alto grau de gramaticalização. Diferentemente de outras línguas, tal como a portuguesa, na qual os verbos modais têm marcação de tempo (ausência de formas não-finitas) e de concordância verbal, os auxiliares modais da língua inglesa contemporânea não apresentam mais essas características⁹². Além disso, os modais precedem o verbo diretamente sem a adição do infinitivo *to*⁹³, o que geralmente acontece quando um verbo é seguido do outro na língua inglesa:

Quadro 4.2.1.A: Forma dos modais

Modal+Verbo	Verbo+Verbo
<i>I can go.</i>	<i>I want to go.</i>
* <i>I can to go.</i>	* <i>I want go.</i>

Ademais, os auxiliares modais compartilham uma série de propriedades que formam o acrônimo NICE (HUDDLESTON e PULLUM, 2002, p. 93). De acordo com esse critério, os modais são categorizados como verbos auxiliares, similarmente aos auxiliares primários *be* e *have*:

- (1)- Negação com n't⁹⁴ - exemplo: *He can't come.*
- (2)- Inversão com o sujeito na forma interrogativa – exemplo: *Can he come?*
- (3)- Coda-exemplo: *She can come and so can he.*
- (4)- Ênfase na afirmação- exemplo: *He may come.*

Palmer (1974, 1990) argumenta que como (ou talvez porque) os modais não possuem formas finitas, eles podem ocorrer somente na posição inicial:

I should begin.

⁹² De acordo com Denison (1993) e Bybee (1995), os modais em inglês são derivados de verbos que possuíam tempo e faziam concordância verbal durante um estágio bem anterior da língua. Dessa forma, o autor reconhece que a derivação histórica desses verbos ainda traz algumas implicações semânticas como *can-could, will-would, may-might, shall-should*, como formas de presente e pretérito respectivamente.

⁹³ *Ought to* é idiossincrático por ser o único verbo modal seguido pelo infinitivo *to* (PALMER, p. 1974).

⁹⁴ *Mayn't* é extremamente raro, com apenas sete ocorrências em todo o BNC.

**I begin to should.*

Ou seja, eles não co-ocorrem⁹⁵:

**I can must.*

Os verbos que parecem possuir todas as características formais para serem considerados verbos modais centrais, quais sejam, as propriedades NICE, ausência de marcação temporal e de concordância verbal são *can, could, will, would, may, might, shall, should e must*. Da mesma forma, Sager et al. (1980, p. 210) atestam ser esses os VMs freqüentemente encontrados em textos técnicos. Por essas razões, e por serem recorrentes expressões de modalização na linguagem escrita, esses são os verbos modais que eu irei considerar para análise.

4.2.2 O uso

Embora haja considerável concordância com relação às propriedades sintáticas que distinguem os modais dos outros verbos, não há critérios amplamente aceitos tampouco uma terminologia em comum para definir o seu uso. Dessa forma, apresentarei os VMs em termos de seus usos e funções conforme autores que priorizaram o estudo dos usos em suas pesquisas, como por exemplo Leech (1989), Coates (1983), Quirk et al. (1985) e Palmer (2003), assim como exemplos de Shawcross (1992) e Sager et al. (1980), para ilustrar os usos na linguagem técnica.

- *Can*

Can pode expressar “possibilidade”, como nos exemplos abaixo:

Even expert drivers can make mistakes. (It's possible for even...) (LEECH, 1989, p. 85)

Leech (ibid) afirma que esse sentido é muitas vezes encontrado na forma negativa com *cannot* ou *can't*:

He can't be working at this hour! (It's not possible...)

Às vezes, *can* (possibilidade) expressa um significado relacionado a hábitos, que pode ser parafraseado pelo advérbio *sometimes*:

⁹⁵ Denison (1993) coloca que algumas variedades não-padrão da língua inglesa permitem duplos modais, por exemplo, *You might would say that* (ibid, p. 294).

Lightning can be very dangerous. = Lightning is sometimes very dangerous. (LEECH, 1989, p. 86).

Can pode também expressar habilidade, com um sentido semelhante a *be capable of*, ou, quando for uma habilidade adquirida, com um sentido equivalente a *know how to*:

Paula can't sing, but she can play the guitar. (She knows how to...) (LEECH, 1989, p. 86)

You can work harder than this. (...are capable of...)

Segundo Leech (1989) com os verbos de “cognição inerte⁹⁶” e de “percepção inerte⁹⁷” há pouca diferença entre *being able to do something* e *actually doing something*. Em outras palavras, *I can remember*, pouco difere de *I remember* como modo de se referir ao estado de lembrar. Com os verbos de percepção inerte, a principal função é expressar uma percepção momentânea:

I can hear. (estado de percepção)

I hear. (percepção momentânea)

Não há uma clara distinção entre os significados de possibilidade e de habilidade, uma vez que “habilidade” implica “possibilidade”. Uma possível marca distintiva para o significado de “possibilidade” é sua ocorrência em orações passivas:

This game can be played by young children. (It is possible for this game...)

Entretanto, a sentença na voz ativa:

(Even) young children can play this game, pode ser interpretada como habilidade.

Shawcross (1992) confirma a semelhança entre habilidade e possibilidade apresentando os usos de *can* sob o título de “*Possibility*” (ibid, p. 93):

Use Can + Infinitive to express possibility. (It's usually better than "it's possible to/that...".)

Após essa breve explicação, Shawcross oferece os seguintes exemplos:

Can expresses physical or technical possibility, capacity etc.

The A320 can carry 172 passengers.

Coaxial cables can transmit numerous messages.

⁹⁶ *Believe, forget, hope, imagine, know, suppose, understand etc.*

⁹⁷ *Feel, hear, see, smell, taste.*

*In the event of an engine failure, the remaining engine **can** power all the hydraulic systems through the PTU.*

*The APU **can** be started up to 25,000 feet.*

O autor denomina o seguinte uso de *can* de “capacity”:

*The outer tank **can** hold 3,500kg. (ibid, p. 95)*

Em linguagem geral, o *can* também é utilizado para pedir ou conceder permissão.

*You **can** stay here as long as you like.*

***Can** I borrow your pen?*

Conforme Leech (1989) lingüistas mais tradicionais⁹⁸ consideram esse uso de *can* como “incorreto”. Entretanto, Coates (1983) e Leech (1989) sugerem que *can* é muito mais freqüentemente utilizado como auxiliar de permissão do que *may*, forma tradicionalmente considerada correta. Coates (1983), Leech (1989) e Palmer (1990) sugerem que, nesses casos, *can* e *may* são praticamente intercambiáveis, sendo que a principal diferença é o fato de *may* ser considerado mais formal e, assim, mais cortês.

- **Could**

Could pode ser encontrado como passado de *can* nas funções de “permissão”, “possibilidade” e “habilidade”.

*As Brazil was neutral, the rules of engagement would be used only after being attacked by enemy, thus Brazilian aircraft **could** only respond to attack. (permissão)*

*Fortunately in those days an aircraft **could** go from design board to production in a matter of months, not so these days. (possibilidade)*

*Only the highly experienced fighter pilots **could** fly and attack in a 45-degree climb then. (habilidade)*

Leech (1989) acrescenta que nesses usos *could* tem o significado de estado, enquanto suas paráfrases próximas *was permitted/ was allowed/ was able* denotam eventos e podem acrescentar um sentido de “cumprimento” ao significado comum do auxiliar modal.

O *could* é bastante usado para expressar significado hipotético de possibilidade.

*If you **could** fly you could land on top of Blackrock mountain.(...it would be possible for...)*

⁹⁸ Assim como pais e professores, que regularmente corrigem seus filhos e alunos.

Mas *could* também pode ser utilizado no sentido com o significado hipotético passado, principalmente quando usado em conjunto com o auxiliar *have*:

*Had you come to me sooner, I **could** have cured you.*

Ou como alternativa mais cortês de *can* em solicitações:

***Could** I see your driving license?*

Conforme Leech (1989), a forma *could* é freqüentemente empregada como substitutiva de *can* na expressão da possibilidade fatural. Esse uso aproxima-se de *may* e *might*. O efeito do auxiliar hipotético, com sua implicação de “contrário à expectativa”, é tornar a expressão de possibilidade mais provisória e cautelosa:

*There **could** be trouble at the World Cup Match tomorrow.*

Em contextos técnicos, o uso de *could* expressa possibilidade/probabilidade aplicado a fenômenos mais naturais, incidentes técnicos não planejados (ou contrário à expectativa), falhas, etc.

*Hot oil **could** cause injury.* (SHAWCROSS, 1992, p. 94)

Could pode ser empregado nas sugestões de ação futura de maneira análoga a *can*, e como torna a expressão de possibilidade mais cortês, sugerindo mais polidez:

*You **could** answer these letters for me.*

Nas negações, *couldn't* nega o auxiliar modal, e não o verbo principal:

*He **couldn't** have made that mistake. (It's not even barely possible that he made that mistake).*

- **May**

May é amplamente utilizado como “possibilidade”:

*Careful, that gun **may** be loaded. (= it is possible that it is loaded)*

Esse uso de *may* é comum em declarativas, mas não ocorre em interrogativas. Quando combinado com “verbo de evento”, *may* geralmente refere-se a uma possibilidade futura, como em: *may go, may become, may lose, may fly*, entre outros.

Leech (1989)⁹⁹ explica a diferença entre os pares *may* e *can*, em que *may* representa “possibilidade fatural” ou epistêmica (COATES 1983, 1995) e *can* representa “possibilidade teórica”, ou deôntica. Através dos exemplos oferecidos por Leech (1989, p. 93), a distinção fica clara nestas declarativas equivalentes:

*The road **may** be blocked. (Perhaps the road is blocked) (It's possible that the road is blocked)*

*The road **can** be blocked. (It's possible to block the road) (It's possible for the road to be blocked)*

A segunda sentença descreve um acontecimento teoricamente concebível, enquanto a primeira parece mais imediata, já que a probabilidade da ocorrência de um evento é considerada. As situações que elas invocam são bastante diferentes (LEECH, 1989, p. 94):

*The road **can** be blocked by the police. (and if we do this, we might intercept the criminals-* dito por um policial a outro)

*The road **may** be blocked by flood water (that explains why the guests haven't arrived-* diálogo entre marido e mulher que aguardam visitas)

Para Leech (1989), a “possibilidade fatural” é mais forte do que a “possibilidade teórica”:

*This illness **can** be fatal.*

*This illness **may** be fatal.*

A segunda situação parece ser bem mais preocupante que a primeira: *Can be fatal* é apenas uma possibilidade teórica; *May be fatal* refere-se a um evento que está ocorrendo na realidade. Ou seja, um paciente poderia sentir-se pessimista se um médico lhe comunicasse a segunda sentença.

⁹⁹ Coates (1995) sugere que há certa confusão entre as modalidades deônticas e epistêmicas quando aplicadas à “possibilidade”, pois, o contraste semântico parece ser bem mais fraco do que em outros pares epistêmicos/deônticos e que, quando a mesma unidade lingüística, por exemplo, *may*, expressa tanto possibilidade epistêmica, quanto possibilidade deôntica, há casos de fusão de significados. *Can* é o expoente mais freqüente de possibilidade deôntica em inglês, enquanto *may* é o expoente freqüente em contextos mais formais. Coates (1995) afirma que a distinção crucial entre formas que expressam possibilidade deôntica e epistêmica é que a última envolve subjetividade. Formas que envolvem subjetividade podem ser definidas como dispositivos através dos quais o falante, ao produzir a elocução, simultaneamente comenta a elocução e expressa sua atitude a respeito do que está falando. Por exemplo, *that may be yellow fever, I'm not sure*. Nesse caso, a incerteza do falante é codificada por um sintagma que reforça o modal, ou seja, *I'm not sure* (60). Coates (1995, p. 60) sugere que *well, I think there is a place where I can get a cheap kettle* (p. 57) são declarações de fato e não há subjetividade envolvida. Entretanto, se a elocução fosse *I may get a cheap kettle* a incerteza do falante estaria codificada também. O significado epistêmico é associado a estilos formais, na voz passiva e com sujeitos inanimados.

Entretanto, Leech (1989) afirma que no inglês formal *may* pode expressar tanto possibilidade teórica quanto fatural, apesar de ser incomum, confirmando a idéia de Coates (1995) de que a distinção deontica-epistêmica desse par é bem mais difícil de ser estabelecida do que nos outros pares.

Shawcross (1992) sugere apenas que *may* é usado para indicar que um evento ou uma ação podem ocorrer, ou acontecer. Semelhantemente a *could*, o autor sugere a idéia de possibilidade expressando incidentes técnicos não planejados e fenômenos naturais para o uso de *may* (ibid, p. 94):

*In dense weather, ice **may** form on the wings.*

*If there is a lot of traffic on arrival, the aircraft **may** have to be hold before landing.*

*The assembly **may** be repaired before overhaul.*

*Water **may** accumulate in the lower fuselage. (SHAWCROSS, 1992, p.95)*

Em nota, Shawcross declara que em inglês geral, *can* e *may* são intercambiáveis, mas, no Inglês Simplificado deve-se usar *can*.

Trimble (1985) sugere outro uso de *may* que, apesar de não muito freqüente, está presente nos textos técnicos, que se assemelha ao uso de *must* (deontico):

*One of the best possible nonfusible weld backings is copper. Copper should be of a sufficient mass or liquid cooled so as to dissipate the heat. For steel thickness other than gage material, a relief groove **may** be necessary. The depth of this relief groove **may** be as little as 0.02" or as much as 1/8" or more. (TRIMBLE, 1985, p. 120)*

Segundo Trimble (ibid), a leitura de *may* como *must*, não é automática. No entanto, o pesquisador explica que ao examinar a situação detalhadamente, nota-se que o soldador (*welder*) não tem a opção de fazer ou não a ranhura de alívio (*relief groove*) caso o aço tenha uma espessura diferente do material do medidor (*other than gage material*). Portanto, o a ranhura deve ser feita (*must be made*), é a única escolha. Conforme Trimble, a dificuldade para o leitor não nativo é que esse uso de *may* convive simultaneamente com os outros usos mais freqüentes. Entretanto a força de ambos é bastante diferente: um sugere algumas possibilidades, enquanto o outro estabelece apenas uma opção.

- **Might**

Might é usado como passado de *may*, apesar de raro, principalmente no IB. Assim como *could*, *might* tem significado de estado, ao contrário de suas paráfrases próximas *was permitted/ was allowed to/ was able to*, que denotam eventos (LEECH, 1989, p. 109):

*The prisoners **might** leave the camp when they wished.*

Entretanto, esse uso é mais comum no discurso indireto:

*Visitors **may** ascend the tower for 50p.-The brochure declared that visitors **might** ascend the tower for 50p.*

Might é usado para expressar “permissão hipotética” principalmente em solicitações de primeira pessoa, sendo considerado mais cortês do que *may*:

Might *I ask you for your opinion?* (LEECH: 1989, p. 139)

Might é usado para expressar um “passado hipotético”:

*If Roger was playing, Grêmio **might** win (... its's possible that Grêmiol would win...)*

O significado “hipotético passado” pode ser expresso pelo Infinitivo Perfeito:

*If Roger had been playing, Grêmio **might** have won.*

O VM pode também ser usado para expressar “possibilidade hipotética”. O auxiliar hipotético implica o sentido de “contrário à expectativa”, tornando a expressão mais tentativa e cautelosa:

*Our team **might** still win the championship.*

Might é comumente empregado nas sugestões de ação futura, de maneira análoga a *can*. Assim como em outras situações de uso, a forma hipotética *might* é mais polida tornando a expressão de possibilidade mais cautelosa:

*We **might** meet again after Christmas.* (LEECH, 1989, p. 141)

Há três significados de *might* + perfeito envolvendo possibilidade:

*You **might** have told me. (It would have been possible for you to tell me)*

*You **might** have dropped it somewhere. (It's possible that you dropped it somewhere)*

*You **might** have met him if you'd been there. (It's possible that you would have met him)*

Em todos os três exemplos acima, *might* poderia ter sido substituído por *could*, sem alteração considerável de sentido. Entretanto, *could* no terceiro exemplo expressaria um significado um tanto diferente de possibilidade teórica, em vez de fatural, e seria parafraseado por: *It would have been possible for you to meet him if you'd been there.*

Nos contextos negativos, *might* nega o verbo principal:

*He **mightn't** have made the mistake. (It's possible that he did not make the mistake)*
(LEECH, 1989, p. 141)

- **Must**

Os significados de *must* são, obrigação ou exigência (deôntico) e necessidade lógica (epistêmico).

Must como “obrigação”, geralmente implica a autoridade do falante sobre o ouvinte.

*You **must** be back by ten o'clock. (You're obliged by me to...)*

Muitas vezes, esse significado de *must* não possui essa característica de autoridade e pode significar uma “exigência”, embora seja difícil estabelecer uma linha divisória entre os dois usos, isto é, autoridade do falante e exigência.

*Students **must** pass an entrance examination to study at this school.*

Shawcross (1992, p. 94) afirma que *must* na linguagem técnica é geralmente usado para expressar “necessidade de que algo seja feito”:

*If the tire is deflated the mechanic **must** remove it.*

*When the demand for cooling air decreases, the Ram Air Outlet **must** be closed.*

Outro uso de *must* é empregado em função de um conhecimento a que se chegou por inferência ou raciocínio:

*This **must** be the right address! (That is necessary the case-no other explanation is possible).*

Esse uso de *must* normalmente não admite as formas negativa e interrogativa.

- **Shall**

Assim como *will*, *shall* tem a função de auxiliar modal e de auxiliar de futuro. Leech (1989) afirma que o uso de *shall* parece estar em declínio, sobretudo no inglês americano (IA), ocorrendo atualmente apenas em uns poucos contextos lingüísticos restritos.

Em declarativas com *I* ou *we* como sujeito, *shall* é um equivalente mais formal de *will*:

*We **shall** not disclose any customer-specific information to any third party unless we are required to do so by law.*

Shall pode também expressar intenção:

*Tomorrow I **shall** inform you news on it.*

Existe um uso volitivo, bastante infreqüente atualmente, na concessão de favores, com sujeito na segunda e na terceira pessoa nas declarativas:

*You **shall** stay with us for as long as you like.*

Em documentos legais, ou quase legais, *shall* aparece expressando obrigação. Nestes casos, *shall* poderia ser substituído por *must* (obrigação) ou pelo *will* (quase imperativo):

*The pilot **shall** advise ATC when weather deviation is no longer required, or when a weather deviation has been completed.*

*When open on the ground, the landing gear bay doors **shall** be safetied (locked by a safety pin or sleeve).*

*Non-approved lubricants **shall** not be used. (SHAWCROSS, 1992, p.94)*

- **Should**

O *should* é historicamente o passado de *shall*. Todavia, no inglês contemporâneo, essa relação aparece apenas no discurso indireto:

*John said, 'I shall go.'-John said that he **should** go.*

Um dos usos de *should* funciona como um atenuador de *must* para expressar necessidade lógica (uso epistêmico). O *should* epistêmico expressa apenas uma suposição. O *should*¹⁰⁰ diminui a força de *must* ao indicar que o falante tem dúvidas quanto à firmeza de sua conclusão. Esse uso pode ser equiparado ao de “probabilidade”.

*Our guests **must** be home by now. (I conclude that they are, in that they left half-an-hour ago, have a fast car, and live only a few miles away) (LEECH, 1989, p. 113)*

¹⁰⁰ Leech (1989) refere-se a ambos *should* e *ought to* como atenuadores de *must*. Neste trabalho *ought to* não será explorado, uma vez que não está entre os modais analisados e não está presente no corpus de estudo.

*Our guests **should** be home by now. (I conclude that they are, in that..., but whether my conclusion is right or not I don't know-perhaps they were hindered by some unforeseen occurrence such as a breakdown).*

*The use of higher-octane fuels in internal combustion engines **should** result in improved performance (SAGER et al., 1980, p. 211)*

Com relação à necessidade lógica *should* expressa uma atitude favorável em relação ao evento ou estado referido:

*Roses **should** grow well in this soil. (LEECH,1989, p. 113),* dessa forma, a sentença abaixo soaria estranha:

*Roses **should** grow badly in this soil.*

Ao ser usado em situações no passado, *should* tem geralmente uma conotação mais forte de “contrariamente ao fato”:

*She **should** have seen my car coming. (...but in fact she didn't.)*

O *should* é também um atenuante do *must* no eixo da obrigação (deôntico). Portanto, seu uso implica que o falante não tem a confiança de que a ação expressa pelo verbo principal será realizada, ou seja, ao usar *must*, o falante espera ser obedecido, mas no caso de *should* não há essa certeza:

*The floor **should** be washed at least once a week. (QUIRK et al., 1985, p. 227)*

Na sua forma mais forte, o *should* assume um significado de obrigação moral ou dever:

*You **should** do as he says. (QUIRK et al., 1985, p. 227)*

Na versão fraca, o *should* oferece um conselho (quando subjetivo):

*You **should** walk round the ramparts of the old city too. (COATES, 1983, p. 58)*

Ou descreve um procedimento correto, quando objetivo:

*I just insisted very firmly on calling her Miss Tillman, but one **should** really call her President. (COATES, 1983, p. 59)*

O uso referido acima de *should* é o mais encontrado no inglês técnico, que Shawcross (1992) e Sager et al. (1980) citam como expressando recomendação ou especificações:

*Chocks **should** be placed under the main gear wheels during a night stop. (SHAWCROSS, 1992, p. 94)*

*By 'critical' is meant a machine feature which, if it is incorrect, **should** be rectified before the programme is allowed to continue (SAGER et al., 1980, p. 211).*

Entretanto, Sager et al. (1980) e Trimble (1985) sugerem que *should* possa exceder uma mera recomendação e expressar uma instrução ou um comando. Trimble afirma ser esse o uso mais frequente (embora não o único) de *should* em material técnico instrucional. Esse uso equivale ao *must* (embora com um tom mais diplomático) ou ao imperativo, e pode substituí-lo ou ser usado em conjunto:

*The clamp **should** be fixed so that the middle portion is approximately 10mm below the surface of the bench top, and then secured with three 1 7/8" (40mm) coach bolts which are provided. (SAGER et al., 1980, p. 212).*

*Steel weld backing **should** be sufficiently thick so that the molten metal will not burn through the backing. In most cases the steel weld backing is fused and remains part of the weldment. (TRIMBLE, 1985, p. 120)*

Para Trimble, a situação acima pode ser parafraseada por:

*If the steel weld backing is not sufficiently thick, the molten metal **will** burn through the backing.*

Ou seja, não é uma simples recomendação, mas a única opção a ser realizada para evitar maiores danos, e poderia, assim, ser semanticamente substituído por *must*. Nesses casos, leitores (principalmente os não nativos de inglês) podem erroneamente entender a força da elocução pressupondo haver uma escolha, mas não existe essa escolha. Shawcross (1992) não apresenta esse uso, apesar de seu livro tratar do inglês técnico para manutenção de aeronaves. Investigarei, assim, se esse uso também está presente no corpus deste estudo, e, caso estiver, o resultado apontará a uma falha na descrição proposta por Shawcross.

O uso de *should* em uma condicional real¹⁰¹ no lugar do presente simples torna a condição mais experimental e acadêmica:

*If you **should** hear the news of them, please let me know. (LEECH, 1989, p. 130)*

Outra expressão de condicional real pode ser obtida através da omissão do *if* e pela inversão entre sujeito e o auxiliar *should*. Nesses casos *should* passa a ser uma conjunção, e não mais um VM:

***Should** you change your mind, there will still be a place for you.*

¹⁰¹ Uma condicional real é aquela em que o falante acredita que a situação poderá ocorrer.

Should the normal servovalve fail, the alternate servovalve will take over.
(SHAWCROSS, 1992, p. 95)

O VM poderá ser também usado em condições irrealis¹⁰²:

Should the container explode, there would almost certainly be widespread damage.

Em orações subordinadas com *that*, o *should* é comum como quase subjuntivo no inglês britânico (IB):

*It is most necessary that we **should** have the funeral bill.* (COATES, 1983:68)

Conforme Coates (1983), esse uso é encontrado em situações mais formais. Além disso, quando funciona como um quase subjuntivo puro, o *should* é semanticamente vazio.

Segundo Sager et al. (1980, p. 212), esse uso está presente nos textos técnicos:

*The basic requirement of the fuel pump is that it **should** pressurize the fuel sufficiently to raise the needle in the injector nozzle to deliver a fine spray into the cylinder.*

- **Will**

Will (assim como o *shall*) tem dupla função de auxiliar modal e de auxiliar de futuro. Conforme Leech (1989, p. 68), “as duas funções são tão intimamente ligadas que é difícil separá-las.” Uma das razões para a confusão que se faz entre o uso modal e futuro é que com relação aos acontecimentos futuros, não se pode ter a mesma certeza, que temos quanto aos eventos passados e presentes, e, por essa razão, mesmo o mais confiante dos prognósticos deve indicar algo da atitude do falante, ou seja, uma expressão de modalidade.

Leech (ibid) afirma que a palavra que mais efetivamente caracteriza o futuro de *will* é “previsão”. Dessa forma, o futuro de *will* é usado em uma grande variedade de contextos em que é apropriado fazer previsões:

*Tomorrow’s weather **will** be cold and cloudy.*

Sager et al. (1980, p. 210) sugerem que o sentido mais frequentemente encontrado no inglês específico é o de “previsão”, que, segundo os autores, indica que uma ação sempre, ou tipicamente acontece sob algumas circunstâncias:

¹⁰² Uma condição irreal, ou hipotética, implica uma suposição por parte do falante, de que o acontecimento não ocorreu, não ocorre ou não ocorrerá.

*Strongly oxidizing acids (nitric acid or concentrated sulfuric acid) **will** attack resins. Strong organic solvents, such as benzene, carbon tetrachloride or phenol **will** attack most polyesters or vinyl esters.*

Will com o significado de “previsão/previsibilidade” faz referência ao futuro, mas também existe nele uma espécie de previsão que se refere ao presente ou passado:

*By now they **will** be eating dinner.* (olhando-se para o relógio)

*That **will** be the electrician-I’m expecting him to call about some rewiring* (ao se ouvir a campainha da porta)

*They **will** have arrived home by now.*

Em *By now they **will** be eating dinner*, o falante faz uma “previsão quanto ao presente” com relação a um evento não observável diretamente. Da mesma forma, *That **will** be the electrician* prevê a identidade de alguém que não está visível no momento. Leech (1989) sugere que *will* (previsão) pertence a contextos similares a *must* (necessidade lógica) e que *must* poderia substituir *will* em todos os exemplos acima. O significado de volição também é relacionado a idéia futura de previsão. Leech (1989, p. 98) reconhece três tipos de volição: a disposição, a insistência e a intenção. Como exemplo de disposição, ou “volição fraca”, nos termos de Leech:

*I’**ll** lend you some money, if you like.*

O significado de insistência, ou volição forte, traz em si fortes mudanças emocionais:

*He **will** go swimming in dangerous water. (He insists on going swimming...)*

Entre disposição e a volição assertiva de insistência, situa-se a categoria intermediária de intenção. Nesses casos, no ato de fala foi tomada uma decisão e o cumprimento da intenção está garantido:

*I’**ll** write tomorrow.*

Will é particularmente comum na oração principal cujas orações subordinadas são introduzidas por conjunções condicionais e temporais, tais como *if*, *unless*, *as soon as*, *as* etc. Todavia, na condicional, a condição futura é expressa por um tempo presente comum:

*If you pull this lever the roof **will** slide back.*

Nas condicionais, *will* frequentemente expressa previsões mais gerais ou habituais. Esse uso é comum em orações declarativas de caráter científico ou quase-científico:

*If an aircraft stalls while yawing a spin **will** occur. (Whenever x happens, it is predictable that y happens)*

*In case of normal servovalve failure the alternate valve **will** take over. (SHAWCROSS 1992, p. 95)¹⁰³*

O *will* pode ser equivalente ao presente simples e similar ao seu uso na oração principal da condicional, onde *will* enfatiza a inevitabilidade de um evento que ocorre sob certas condições que são geralmente implicadas ou expressas por alguma outra oração na sentença no lugar de uma oração com *if* (SAGER et al., 1980, p. 211):

*Like untreated cotton, a hydrophilic fiber such as PET-PEO **will** swell when wet.*

*One also accepts that some liquids do not stay mixed, such as water and oil stirred together, and that they **will** separate on being left to reach an equilibrium state.*

Existe um tipo de sentença em que *will* indica disposição. Este uso de *will* assemelha-se ao sentido de “possibilidade” de *can*. Ele poderia, entretanto, ser considerado um tipo de “previsibilidade”, em que a oração condicional fica subentendida:

*The auditorium **will** seat 500. (One can seat 500 in the auditorium.) (If required)*

Ainda em sobreposição com o modal *can* (e *able* ou *capable*), Sager et al. (1980, p. 211) mostram que *will* pode expressar habilidade. Entretanto, esse uso só foi encontrado em orações descrevendo o desempenho ou qualidade de máquinas, materiais ou equipamentos:

*Measuring 203 by 450 mm, the table **will** support workpieces up to 76 mm thick which weight up to 138kg.*

*This machine (RT800 surface grinder) **will** handle work measuring up to 800 mm long by 550mm wide by 450 mm high.*

Há ainda um uso de *will* quase imperativo, encontrado com sujeitos em segunda e terceira pessoas:

*The Duty Officer **will** report for duty at 0700 hours.*

Esse *will* é um equivalente mais forte do *must*. Implica que o falante ou escritor tem tanta autoridade sobre aquele a quem se dirige que a não execução da ação prevista está fora de questão.

¹⁰³ Shawcross afirma que somente a forma básica da condicionais, isto é, *present simple+present simple (zero conditional)*, ou *present simple+future (first conditional)*, é empregada.

- **Would**

O *would* funciona como o passado de *will* e como um marcador hipotético geral, sendo o segundo uso muito mais importante. O futuro no passado descreve acontecimentos que estão no futuro com relação a um ponto avançado no passado. Esse tipo de construção é incomum e restringe-se largamente a um estilo literário de escrita histórica ou narrativa. O evento previsto não pode estar no futuro do ponto de vista presente do narrador:

*Twenty years later, Dick Whittington **would** be the richest man in London.* (LEECH, 1989, p. 64)

Would é usado como passado de *will* como previsibilidade, sendo bastante comum nas descrições históricas ou fictícias de caráter, comportamento típico, etc.

*In his last years, the king **would** spend whole days in morose solitude, speaking only to his immediate family and refusing all official audiences.* (LEECH, 1989, p. 108)

Ainda como passado, *would* substitui *will* no discurso indireto:

*The plan **will** fail.- I warned them that the plan **would** fail.*

Um dos principais usos de *would* é nas condições irrealis, ou hipotéticas. Elas são normalmente formadas pelo uso do passado (do indicativo ou do subjuntivo) na oração condicional e a construção *would* + infinitivo na oração principal. Implica uma suposição, por parte do falante, de que o acontecimento não ocorreu, não ocorre ou não ocorrerá. Esse comprometimento negativo com a verdade a distingue do significado fatural:

It's laughable that Septimus is in Love. (Yes, in fact he's in Love.)

Quando o hipotético seria:

*It **would** be laughable if Septimus were in love. (But actually he's not in Love)*
(LEECH, 1989, p. 131)

Nos acontecimentos imaginários, há um sentido de contrário à suposição, quando no presente:

If you really cared for your children, you'd look after them properly. (... but I assume you don't care for them...)

E um sentido de contrário à expectativa quando no futuro:

*If it were to snow tomorrow, the match **would** have to be cancelled. (...but I don't expect it will snow)* (LEECH, 1989, p. 136)

Ao se referirem a eventos passados imaginários, as formas hipotéticas normalmente têm o sentido de “contrário ao fato”, uma vez que eventos passados são facilmente conhecidos:

*If your father had caught us, he **would** have been furious. (... but in fact he didn't catch us.)*

Nas orações principais, uma forma verbal hipotética muitas vezes requer a presença de uma oração condicional, uma vez que *The Eiffel tower **would** fall down*, não é, por si só, uma sentença inglesa auto-suficiente, pois, a menos que seja acrescentado uma condição a ela, o ouvinte provavelmente ficará se perguntando *if what?* Entretanto, há circunstâncias em que a oração principal hipotética pode ser auto-suficiente, quando existe uma condição suprimida, como é freqüente na expressão de desejos:

I'd hate to live in a house like that. (...if I had to...) (LEECH, 1989, p. 133)

Isso provavelmente acontece porque existe uma certa indelicadeza na expressão explícita de desejos, ocasionando que uma declarativa de desejo possa ser interpretada como imperativa. Ou seja, *I'd like juice*, é mais polida que *I want juice*. Esse uso estabelece uma suposição diplomática de que o falante não será servido.

O uso cortês de *would*, em lugar de *will* (disposição) nas solicitações de segunda e terceira pessoas, mostra um outro exemplo do uso absoluto de uma oração hipotética com verbos expressando desejo:

***Would** you lend me fifty pence?* (LEECH, 1989, p.139)

Entretanto, a resposta típica de tal interrogativa exige o uso de um auxiliar não correspondente. Dessa forma, a resposta para a pergunta acima poderia ser:

*Yes, I **will**.*

Quadro 4.2.2.A: Síntese das formas e usos dos VMs

Uso	Forma
Condicional- Substituto de <i>If</i>	<i>Should</i>
Habilidade/capacidade	<i>can/could/ will</i>
Hipótese	<i>could/might/should/would</i>
Obrigação, ordem, comando	<i>may/must/shall/should/will</i>
Necessidade Epistêmica /Inferência	<i>can't/mustn't /must/ /will/should</i>
Pedir ou conceder permissão	<i>can/may/could/might/would</i>
Possibilidade Deontica/Epistêmica	<i>can/could/may/might/</i>
Proibição	<i>cannot (can't)/may not/must not</i>
Previsão/ volição/ intenção	<i>shall/will/would</i>
Quase subjuntivo/ subjuntivo	<i>Should</i>
Sugestão	<i>can/could//should/would</i>
Volição	<i>shall/will/would</i>

O quadro 4.2.2.A tem como objetivo representar e sintetizar a extensa rede de usos e formas dos VMs. O quadro mostra os diferentes usos associados a cada modal e vice-versa, isto é, os diferentes modais associados a cada uso. Por exemplo, *should* é relacionado a seis significados diferentes, como “condicional” (podendo substituir *if*), “hipótese”, “obrigação”, “necessidade epistêmica/inferência”, “quase subjuntivo” e “sugestão”. Se o quadro for abordado pelo outro lado, isto é, pelos usos, pode-se notar, por exemplo, que possibilidade deôntica/epistêmica, pode ser expressa por *can*, *could*, *may*, *might* ou *will*. Após a análise dos usos dos VMs no AC será desenhado um quadro com os seus usos e comparado a este do inglês geral.

Até este ponto, demonstrei os usos mais freqüentes em que os VMs são usados principalmente em contextos afirmativos. Como visto na seção sobre as formas dos VMs, esses não necessitam de outro verbo auxiliar para fazerem a negação. Entretanto, há particularidades com relação ao uso dos VMs em contextos negativos. Alguns deles negam o modal, outros, o verbo principal. Essa questão será detalhada a seguir.

4.2.3 A problemática da Negação

Conforme Palmer (1995, 2003), há uma particularidade referente à negação de verbos epistêmicos e deônticos de possibilidade e necessidade em inglês¹⁰⁴ (*can*, *may*, *must* e *need*). O sistema é por vezes irregular, havendo, assim, a necessidade de supleção. O motivo é que há uma falta de correspondência exata entre forma e significado, não sendo possível prever para uma certa forma qual será o seu significado, ou qual será a forma que irá expressar um significado modal específico. Se uma distinção fosse feita entre a modalidade e a proposição, deveria ser formalmente possível distinguir entre *not possible* e *possible not*, ou seja, entre *it's not possible for us to go* e *it's possible for us not to go*. Algumas línguas fazem essa distinção colocando a negação antes do VM ou antes do verbo principal, mas a língua inglesa não apresenta essa possibilidade. Dessa forma, o marcador negativo sempre segue o modal, como um clítico-*n't* ou como *not*.

A língua inglesa faz a supleção, com, por exemplo, *may* e *can* para possibilidade epistêmica (relação semelhante é obtida para *must* e *need* para necessidade deôntica) (PALMER, 2003, p. 9):

They may be in the Office.

They can't be in the office. (not possible)

¹⁰⁴ Assim como em outras línguas também, especialmente as germânicas (PALMER, 1995, p. 453).

*They **may not** be in the office. (possible not)*

Em segundo lugar, a língua inglesa pode fazer uso da lógica, uma vez que há equivalência lógica entre “*not possible*” e “*necessary not*” e entre “*possible not*” e “*not necessary*”, e, assim, usar uma forma logicamente equivalente. As formas abaixo são usadas como negativos de *must*.

*They **must** be in the office.*

*They **may not** be in the office. (possible not ou not necessary)*

*They **can't** be in the office. (not possible ou necessary not)*

O *can't* acima nega a modalidade – *it's not possible that they are in the Office*. Nesse sentido, funciona como negativo de *can*, mas também como negativo de *must*, indicando que é necessariamente o caso de que eles não estão no escritório. Um VM de possibilidade é utilizado aqui porque o *must* não possui outra forma de expressar *necessary not*, mas é enganoso afirmar que é o negativo de *must*; é formalmente o negativo de *can*, mas usado como supleção.

Entretanto, no IA, e em algumas variantes do IB, *must not* é utilizado quando *can't* é utilizado no IB. As únicas formas disponíveis em comum (para a modalidade epistêmica) são *may not* e *must not*- que mais naturalmente negam a proposição- *possible not* e *necessary not* respectivamente. Como há novamente somente duas formas, há a necessidade da supleção para suprir as formas que negam a modalidade. Com as equivalências lógicas, *not necessary* está sendo sinalizado por *may not* e *not possible* por *must not*. Conforme os exemplos de IB (b) e IA (c):

*They **must** be in the office.*

*They **may not** be in the office (possible not para not necessary)*

*They **must not** be in the office. (necessary not)*

Há ainda alguns contextos em que *mustn't* ou *must not* parecem possíveis mesmo no discurso do falante. Por exemplo:

*The restaurant is always empty. It **musn't** be very good.*

Can't seria possível, but *mustn't* também é. Entretanto, *mustn't* não seria possível nas seguintes situações:

A: *That woman over there, is it Mrs Wilson?*

B: *It **can't** be Mrs Wilson. She died two years ago.*

A: *Someone is knocking at the door. It **may** be Mary.*

B: *It **can't** be Mary. She went to a movie tonight.*

O contexto em que *musn't* parece possível é aquele em que é importante enfatizar *necessary not*, ao invés de *not possible-it must be the case that not*, ao invés de *it can't be the case that*. Dessa forma, a supleção nem sempre se aplica.

Com os modais deônticos, a situação é ligeiramente diferente. Com possibilidade que expressa permissão há a supleção com *needn't (not necessary)* para *possible for (not possible)* sendo expresso por *can't*, como em:

*You **can/may** come in.*

*You **needn't** come in (not necessary para not possible)*

*You **can't/may not** come in. (not possible)*

A primeira das formas negadas expressa *permission not to*, a segunda *no permission*. Contudo, parece haver uma distinção formal clara e regular entre os deônticos *not possible* e *necessary not*, uma vez que a negação da permissão com *can't* não é igual a expressar uma obrigação negativa com *musn't*. Há uma diferença clara entre:

*You **can't/may not** come in. (not possible-no permission)*

e

*You **mustn't** come in. (necessary not)*

Há ainda outra característica específica relacionada à negação e modalidade. Entende-se que a diferença entre o deôntico *needn't* e *musn't* é que o primeiro expressa *not necessary* e o segundo expressa *necessary not*. Há então, a negação da modalidade com *needn't* e a negação da proposição com *mustn't*, como em:

*You **needn't** come in. (it's not necessary for you to come)*

*You **mustn't** come in. (it's necessary for you not to come)*

Leech (1989) afirma que o fenômeno da reversibilidade¹⁰⁵ não se aplica da mesma forma nos contextos negativos, levando a uma equivalência entre VMs que, num contexto positivo, têm significados opostos:

*You **may not** smoke in here. (I don't permit you to smoke...)*

*You **mustn't** smoke in here. (I oblige you not to smoke...)*

¹⁰⁵Reversibilidade refere-se ao contraste de significado entre “permissão” e “obrigação” e entre “possibilidade” e “necessidade”. Isto é, “permissão” é o inverso de “obrigação” e “possibilidade” é o inverso de “necessidade”

Ambas as declarativas exprimem proibição, mas diferem entre si pelo fato de a segunda soar um pouco mais incisiva, definitivamente proibindo, em vez de apenas negar permissão. Essa equivalência acontece porque a oposição “inversa” entre os dois significados é anulada pelo contraste entre a negação do auxiliar e a do verbo principal.

4.2.4 Os atos de fala

Além dos estudos até aqui apresentados, os VMs são estudados indiretamente através de outra “porta de entrada” de estudos lingüísticos: os atos de fala. Mencionarei brevemente esse assunto uma vez que, como mencionado no capítulo anterior, os manuais os manuais são textos que expressam certos atos comunicativos, como, por exemplo, dar instruções.

A teoria dos atos de fala foi elaborada inicialmente por Austin (1962) e desenvolvida posteriormente por Searle (1969). Austin postula que determinadas sentenças são na verdade ações, na medida em que, ao proferir algo, uma ação estará sendo simultaneamente realizada. Alguns exemplos são dizer "sim" perante um juiz ou padre; "nos encontraremos amanhã pela tarde" para um colega; ou ainda, perguntar a um amigo: "você tem dez reais para me emprestar?".

Em cada uma dessas frases é realizada uma ação, embora seu sucesso não dependa apenas do sujeito que as profere, mas de uma série de condições. Por exemplo, a noiva pode dizer "não"; alguém pode, mesmo contra a sua vontade, faltar à reunião; um amigo pode não ter o dinheiro para emprestar ao outro. Isso, contudo, não significa que o que foi dito é falso, apenas que não teve sucesso, do mesmo modo que ocorre com outras ações, quando, por exemplo, quando alguém corre para pegar o ônibus, mas chega tarde demais. Tendo sucesso ou não, prometer, pedir, exigir, protestar, jurar etc. já são ações por si mesmas.

Portanto, é chamada de “ato de fala”, toda ação que é realizada através do dizer. As ações que se realizam através dos atos de fala podem ser muito diferentes e Austin distinguiu três dimensões diferentes. Por exemplo, na frase: "o senhor está pisando no meu pé", três atos de fala estão sendo realizados: o primeiro deles é o *ato locucionário*, ou seja, o ato de dizer a frase; o segundo ato é o que Austin chama de *ilocucionário*, o ato executado na fala, ou seja, a força que um enunciado produz. Nesse caso, ao dizer "o senhor está pisando no meu pé" a intenção não foi simplesmente a de constatar uma situação, mas a de protestar ou advertir para que a outra pessoa parasse de pisar no pé. Por fim, há ainda um terceiro ato, chamado de *perlocucionário*, que é o ato de provocar um efeito em outra pessoa através da locução,

influenciando em seus sentimentos ou pensamentos. Na situação acima, para que o outro tire o pé de cima. Assim o ato locucionário diz algo, o ato ilocucionário realiza uma ação ao ser dito e o perlocucionário tem a intenção de provocar nos ouvintes certos efeitos (convencer, levar a uma decisão etc.).

Alguns exemplos de atos de fala, ou atos ilocucionários utilizando os VMs são:

-fazer um pedido: *Could you pass me the salt?*

-dar uma ordem: *You must do it now.*

-prometer algo: *I promise I'll give it back*

-Pedir a atenção de um público: *Ladies and gentlemen, may I have your attention, please?"*

-Desafiar alguém: *Can you race with me to that building over there?*

Todas essas noções são retomadas e sistematizadas por Searle (1969). Searle distingue cinco grandes categorias de atos de fala:

1. representativos (mostram a crença do locutor quanto à verdade de uma proposição: afirmar, asseverar, dizer);
2. diretivos (tentam levar o alocutário a fazer algo: ordenar, pedir, mandar);
3. comissivos (comprometem o locutor com uma ação futura: prometer, garantir);
4. expressivos (expressam sentimentos: desculpar, agradecer, dar boas vindas);
5. declarativos (produzem uma situação externa nova: batizar, demitir, condenar).

Descrevo os diretivos por acreditar (consoante SAGER et al., 1980) serem os mais freqüentes no tipo de texto estudado nesta pesquisa. Sager et al. (1980, p. 27) contextualizam os atos de fala diretivos no universo da linguagem técnico-científica. Esses têm o objetivo de modificar um comportamento e, para especificar o comportamento, fazem uso da natureza referencial da língua. A intenção diretiva é orientada para o ouvinte. Se o ouvinte realiza o que lhe ele foi pedido, a comunicação obteve sucesso. Os pedidos, instruções, proibições ou ordens são atos diretivos, e suas formas lingüísticas distintas são geralmente óbvias. Esses atos podem aplicar-se a um número limitado de ouvintes, como por exemplo, as instruções em um manual. Os autores sugerem que através desses atos de fala é possível distinguir a linguagem da ciência da linguagem da tecnologia. A linguagem da ciência diz respeito à extensão do conhecimento, enquanto a linguagem da tecnologia está relacionada à aplicação do conhecimento. Ou seja, a tecnologia tende a ser expositória, educacional, instrucional. O ato de fala tem como foco o ouvinte e há sempre uma desigualdade de conhecimento.

Conforme Sager et al. (1980), os atos de fala diretivos estão virtualmente ausentes da linguagem da ciência, mas são bastante freqüentes na linguagem da tecnologia.

A Teoria da Polidez de Brown e Levinson (1987) também faz referência indireta aos VMs. Segundo os autores, a interação humana é governada por princípios de polidez que se refletem nos universais lingüísticos. Utilizando a noção de “face” de Goffman (1967)¹⁰⁶, Brown e Levinson (1987, p. 13) sugerem que essa consiste de dois tipos de desejos básicos, ou “desejos da face”, que acreditam que todas as pessoas possuem: o desejo de ser aprovado em certos aspectos, ou “face positiva” e o desejo de não ser impedido ou coagido em suas ações, ou “face negativa”. Há certos tipos de atos que, intrinsecamente, ameaçam os desejos da face. Os atos de ameaça à face são aqueles que por sua natureza ignoram ou contrariam os desejos da face do *ouvinte/leitor* e/ou do *falante/escritor*.

Alguns desses atos de fala, como por exemplo, ordens e comandos, atacam a face negativa do ouvinte, sendo assim considerados atos de ameaça à face. Por outro lado, instruções e procedimentos não são considerados atos de ameaça à face por serem realizados de forma a satisfazerem as necessidades dos ouvintes. Para que o ouvinte os interprete de forma mais precisa, é melhor produzir esses atos sem mitigação ou elementos lexicais que expressem polidez, da forma mais direta e clara possível. Ou seja, ambos falante e ouvinte concordam que a relevância da face pode ser suspensa tendo em vista os interesses de urgência ou eficiência. (BROWN e LEVINSON, 1987, p. 69). Os autores argumentam que, quando o foco da interação é orientado para a realização de uma tarefa, noções de polidez tornam-se irrelevantes. Leech salienta que (1983, p. 84), a “polidez é altamente irrelevante em funções colaborativas, tais como instruções”. Para Leech, grande parte dos discursos escritos entram nessa categoria.

Segundo Brown e Levinson (1987), uma forma através da qual o falante pode demonstrar que não quer atingir a face do ouvinte é através de diferentes opções sintáticas. Isto é, ao dar uma ordem, o uso da voz passiva, e a exclusão do pronome-sujeito *you*, pode ser uma dessas estratégias, como nos exemplos abaixo:

You must do it tomorrow.

This must be done tomorrow

¹⁰⁶ GOFFMAN, E. ‘On facework: an analysis of ritual elements in social interaction’, in Jaworski, A., and Coupland, N. (eds.) *The Discourse Reader*, London, Routledge, 1967, pp. 306-321.

Apesar de também expressar uma ordem, o segundo exemplo, não ameaça à face do ouvinte da mesma forma que o primeiro, sendo considerado mais polido. Entretanto, no discurso escrito, “a polidez resulta não somente do uso de estratégias individuais no nível da sentença. Mais precisamente, resulta de um número de atos ilocucionários complexamente expressos em conjunto em um texto” (PILEGAARD, 1997, p. 224).

4.2.5 Distribuição

As seções anteriores deste capítulo trataram sobre a sintaxe, a semântica e a pragmática dos VMs, temas freqüentemente presentes nos estudos sobre modalização. Nesta seção, discutirei o fator distribucional. A investigação da distribuição de diferentes itens lexicais em um corpus, assim como de um mesmo item lexical em diferentes corpora, é, sem dúvida, uma grande contribuição da LdC aos estudos lingüísticos em geral, pois, como já mencionado, as diferenças em freqüência não são aleatórias. Pelo contrário, a recorrência de traços lingüísticos peculiares aponta que a linguagem é padronizada (*patterned*) e motivada por diversos fatores. Segundo Biber et al. (1999), um dos mais importantes usos das investigações baseadas em corpus é precisamente “fornecer informações sobre a freqüência do uso” (BIBER et al., 1999, p. 8). Professores, acadêmicos, alunos e autores/produtores de materiais didáticos consideram útil saber quais padrões gramaticais são comuns, merecedores, portanto, de mais atenção. Até muito recentemente, essa informação era geralmente baseada na intuição dos falantes nativos de uma língua, que não possuem formas adequadas de constatar a freqüência e padrões recorrentes em uma língua. Entretanto, McEnery et al. (2006) sugerem que as intuições dos falantes nativos são freqüentemente incorretas, ou pelo menos inexatas, pois:

cada um de nós possui apenas um conhecimento parcial da própria língua, como nossas preferências e preconceitos, nossa memória é fraca, nossa imaginação é poderosa (assim podemos conceber contextos para os contextos mais implausíveis), e tendemos a prestar mais atenção nas palavras ou estruturas incomuns e não notar os mais comuns.

(McENERY et al., 2006, p. 147)

O quadro abaixo mostra as diferenças de freqüência de ocorrência dos VMs em sete diferentes corpora, conforme diferentes estudos (COATES, 1983; BIBER et al., 1999;

KENNEDY, 2002; GABRIELATOS e SARMENTO, 2006; DERNARDIN, 2008)¹⁰⁷. Os seis primeiros corpora apresentados no quadro, Survey, LOB, LSWE, FROWN, BNC e BROWN, são corpora de inglês geral. O OSC é um corpus de inglês específico composto de manuais de sistemas operacionais para computadores pessoais. Os VMs estão apresentados em ordem decrescente de ocorrência, ou seja, o VM mais freqüente naquele corpus aparece na posição 1 e o menos freqüente na posição 9 do quadro:

Quadro 4.2.5.A: Classificação da freqüência dos VMs entre diferentes corpora

Freqüência	Survey	Lancaster	LSWE	FROWN	BNC	BROWN	OSC
1	<i>Will</i>	<i>would</i>	<i>will</i>	<i>would</i>	<i>will</i>	<i>will/would</i>	<i>can</i>
2	<i>Can</i>	<i>will</i>	<i>would</i>	<i>can</i>	<i>would</i>		<i>will</i>
3	<i>Would</i>	<i>can</i>	<i>can</i>	<i>will</i>	<i>can</i>	<i>can</i>	<i>may</i>
4	<i>Could</i>	<i>could</i>	<i>could</i>	<i>could</i>	<i>could</i>	<i>could</i>	<i>might</i>
5	<i>Must</i>	<i>may</i>	<i>may</i>	<i>may</i>	<i>may</i>	<i>may</i>	<i>should</i>
6	<i>Should</i>	<i>should</i>	<i>should</i>	<i>should</i>	<i>should</i>	<i>must</i>	<i>would</i>
7	<i>May</i>	<i>must</i>	<i>must</i>	<i>must</i>	<i>must</i>	<i>should</i>	<i>must</i>
8	<i>Might</i>	<i>might</i>	<i>might</i>	<i>might</i>	<i>might</i>	<i>might</i>	<i>could</i>
9	<i>Shall</i>	<i>shall</i>	<i>shall</i>	<i>shall</i>	<i>shall</i>	<i>shall</i>	<i>shall</i>

Apesar de os seis corpora de inglês geral representarem diferentes variedades do idioma (IA e/ou IB, oral e/ou escrito, geral ou específico), e terem sido compilados em diferentes épocas (anos 60 a 90), pode-se notar algumas similaridades com relação à freqüência e distribuição dos VMs. Na primeira posição, encontram-se *will* e/ou *would*, sendo que *will* tem a preferência. Surpreendentemente, Dearnardin (2008) encontrou exatamente a mesma freqüência de ocorrência de *will* e *would* no corpus BROWN, com o valor absoluto de 2001 ocorrências para cada VM. O *can* ocupa a segunda ou terceira posição, sendo mais freqüente na terceira posição. O *could* é o quarto modal mais freqüente em todos os seis corpora. A quinta posição é ocupada por *may* em cinco e *must* em um dos corpora. Similarmente, a sexta posição é ocupada por *should* em seis corpora e *must* em um, enquanto

¹⁰⁷ O Survey utilizado por Coates (1983) continha 725 mil palavras de inglês escrito e oral. O LOB (Lancaster-Oslo/Bergen Corpus), também investigado por Coates possui um milhão de palavras de inglês escrito em 15 diferentes gêneros. O corpus LSWE (Longman Spoken and Written English Corpus) foi estudado por Biber Et al. (1999) contendo mais de 40 milhões de palavras. O FROWN, investigado por Gabrielatos e Sarmento, 2006) é um corpus de inglês americano escrito, composto de um milhão de palavras compilado nos anos 90. O BNC, analisado por Kennedy (2002) é um corpus de aproximadamente 100 milhões de palavras de IB escrito e falado, que foi disponibilizado em 1995. O BROWN consiste em um milhão de palavras de IA compilado no início dos anos 60. O OSC é um corpus de inglês escrito composto dos manuais dos sistemas operacionais para computadores pessoais: Windows e Linux. O primeiro consiste em 93.277 tokens e o segundo em 81.305, totalizando 174.582 tokens. O OSC (Operational System Corpus) foi compilado por Dearnardin (2008) e estudado pelo mesmo autor juntamente com o BROWN.

must possui a sétima posição em quatro dos corpora investigados. *Might* e *shall* ocupam respectivamente a oitava e a nona posições em todos os seis corpora de inglês geral.

Cotejando os dados dos corpora de inglês geral com o OSC, é possível observar uma diferença saliente entre as frequências de ocorrência. O *can* é o VM mais freqüente, diferentemente dos outros seis corpora. *Will* ocupa o segundo lugar, da mesma forma que o LOB. O *may* e o *might* são também mais usados no OSC, sendo que o *may* sai da quinta posição dos corpora de inglês geral para ocupar a terceira no OSC. *Might* ocupa apenas a oitava posição nos outros corpora e é o quinto no OSC. Em contrapartida, *would*, que é geralmente o primeiro ou o segundo em ordem de freqüência, é apenas o sexto no OSC. O mesmo pode ser dito sobre *could*, que é o quarto em todos os corpora de inglês geral sendo o oitavo no OSC. *Should*, *must* e *shall* não apresentam maiores variações.

Biber et al. (1999) investigaram também a distribuição dos VMs nos diferentes registros do LWSE, a saber, Conversação, Ficção, Notícias e Acadêmico. *Can* e *could* são relativamente freqüentes em todos os registros, sendo *can* extremamente comum em conversação e prosa acadêmica e *could* mais comum em ficção. *May* é extremamente comum em prosa acadêmica e raro em conversação. *Must* e *should* são relativamente comuns em prosa acadêmica.

Pode-se notar que a variação de ocorrência dos VMs nos corpora específicos é realmente diferente, o que é compreensível, uma vez que as necessidades funcionais das linguagens específicas, são, como sugere a categoria, específicas. É importante ressaltar que as análises acima levam em consideração somente a forma, ou seja, a entrada lexical da palavra ortográfica. Como foi visto na seção anterior, todos os VMs possuem vários significados e funções, dessa forma, não é possível saber quais são os usos mais recorrentes de cada VM. Nesse sentido, Coates (1983) classificou também a freqüência de uso dos VMs com relação ao seu significado ou função nos corpora de inglês geral e obteve o seguinte resultado, em ordem decrescente (COATES, 1983, p. 25):

Quadro 4.2.5.B: Frequência dos diferentes modais/uso

VM	Uso
<i>Will</i>	Previsão
<i>Would</i>	Previsão hipotética
<i>Can</i>	Possibilidade deôntica
<i>Will</i>	Intenção
<i>Could</i>	Hipotético
<i>Can</i>	Habilidade
<i>May</i>	Possibilidade epistêmica
<i>Must</i>	Obrigaçao
<i>Will</i>	Volição
<i>Must</i>	Inferência Epistêmica
<i>Might</i>	Possibilidade Epistêmica
<i>Should</i>	Obrigaçao
<i>Could</i>	Possibilidade deôntica no passado
<i>Could</i>	Habilidade no passado
<i>Shall</i>	Previsão
<i>Would</i>	Volição no passado
<i>Should</i>	Hipotético
<i>Would</i>	Previsão no passado
<i>Should</i>	Inferência Epistêmica
<i>Shall</i>	Intenção
<i>Can</i>	Permissão

Nessa hierarquia de uso (oposta a uma hierarquia das formas) é possível notar que todos os modais têm um significado que predomina quantitativamente. A exceção é o *must*, que possui dois significados com frequência semelhante de uso. Alguns usos frequentemente presentes em materiais didáticos, como por exemplo, o uso de *may* para pedir permissão, apresenta uma ocorrência tão baixa que nem mesmo aparece nos resultados¹⁰⁸. Para esse uso, o VM *can*, que é muitas vezes considerado “errado” por pais e educadores, aparece como mais freqüente.

Biber et al. (1999) verificaram que os VMs que expressam “volição” e “previsão”, assim como os que expressam “permissão” e “possibilidade” são consideravelmente mais freqüentes do que os que marcam “obrigação” e “necessidade”. Conforme os autores, isso se deve primeiramente ao fato de que há uma tendência geral em “evitar a força ameaçadora de face das expressões com o significado de obrigação” (IBID: 489). Em segundo lugar, alguns

¹⁰⁸ Coates (1983) não explicita qual a frequência mínima para inclusão na tabela de hierarquia de uso.

semi-modais¹⁰⁹ tornaram-se mais comuns nesse domínio semântico, sendo o *have to* o mais freqüente em todos os quatro registros.

Analisando o BNC, Kennedy (2002) conclui que os VMs associados com os significados deônticos de “disposição/intenção”, “habilidade”, “hábitos” e “hipóteses” e os usos epistêmicos de “previsão”, “certeza” e “possibilidade/probabilidade” respondem por 73% dos *tokens* modais no BNC. O uso dos VMs para expressar o significado deôntico de “obrigação/necessidade” ou “permissão” é relativamente infreqüente no BNC.

Tabela 4.2.5.C: Padrões sintáticos possíveis para os verbos modais em inglês

	Estrutura Modal	%
1	Modal sozinho	1,9
2	Modal + infinitivo (modal + voz ativa)	76,0
3	Modal + <i>be</i> + passado particípio (modal+ voz passiva)	14,7
4	Modal+ <i>be</i> +presente particípio (modal+progressivo)	1,5
5	Modal+ <i>have</i> +passado particípio (ou adjetivo) (modal+perfeito)	5.1
6	Modal+ <i>be+being</i> +passado particípio (ou adjetivo) (modal+ <i>be</i> +progressivo+passado particípio/adjetivo)	0,0
7	Modal+ <i>have+been</i> +passado particípio (Modal+voz passiva+perfeito)	0,7
8	Modal+ <i>have+been</i> +presente particípio (modal+voz passiva+progressivo)	0,1
9	Modal+ <i>have+been+being</i> +passado particípio (ou adjetivo) (modal+voz passiva+perfeito+progressivo)	0,0

(KENNEDY, 2002, p. 82)

Baseado nas descrições das gramáticas tradicionais, Kennedy (2002) estabeleceu as nove estruturas sintáticas em que os VMs ocorrem, que podem ser vistas na tabela 4.2.5.C. O pesquisador aponta que todas as estruturas, com exceção da número 1 e da número 5, possuem a cópula *be*, seja como parte da voz passiva, do aspecto progressivo, ou como o próprio verbo lexical. Salienta que a multifuncionalidade da cópula *be* é provavelmente um dos responsáveis pelas dificuldades apresentadas pelos aprendizes com relação aos VMs. Sem uma análise distribucional um aprendiz pode concluir que todas as nove estruturas ocorrem da

¹⁰⁹ Em Biber et al. (1999), *have to*, *had better*, *have got to*, *need to*, *be supposed to* foram considerados semi-modais.

mesma forma, uma vez que a maioria das gramáticas não fornece informações relativas às suas freqüências.

Kennedy (2002, P. 82) investigou como os VMs são distribuídos entre as nove estruturas no BNC e encontrou enormes diferenças na distribuição entre as diferentes estruturas. A estrutura número 2, modal+infinitivo, responde por 76% das ocorrências, seguida pela estrutura número 3, modal+voz passiva, com menos de 15%. A estrutura número 5, modal+*have*+passado particípio ocorre em aproximadamente 5% dos casos, enquanto as outras estão virtualmente ou totalmente ausentes. Ao verificar os diferentes subcorpora, Kennedy (ibid) constatou haver uma diferença ainda maior na distribuição dos VMs entre as estruturas, principalmente nos subcorpora “oral” e “prosa imaginativa”, com freqüências de 84,9% e 81,3% respectivamente na estrutura 2, conseqüentemente com menos ocorrências na estrutura 3 (voz passiva). Entretanto a estrutura com modal+passiva é proporcionalmente mais freqüente nos gêneros “Ciência” e “Comércio e Finanças”.

Biber et al. (1999) sugerem que enquanto a maioria dos VMs não ocorre com voz ou aspecto marcados, alguns modais mostram preferências distintas por essas combinações. Eles constatam que, por exemplo, enquanto *can*, *could*, *should* e *must* são relativamente comuns nas construções passivas, *may*, *might*, *should* e *must* são os mais freqüentemente usados com o aspecto perfeito. Quanto à voz, Biber et al. (ibid) verificaram que o uso da passiva junto com os VMs é bastante raro na conversação e na ficção, mas relativamente comum com alguns verbos na prosa acadêmica. Os VMs mais freqüentemente encontrados em orações passivas foram respectivamente, *can*, *should*, *could* e *must*. *Will*, *would* e *shall* são extremamente raros nessa construção, sendo *shall* o mais raro proporcionalmente. Segundo os pesquisadores, esses últimos verbos geralmente marcam “volição” e “previsão”, geralmente expressando abertamente o agente do verbo principal, como por exemplo:

We shall deal with chromatography in general in 6.3 of this chapter (ibid, p. 499)

Por outro lado, os VMs *can* e *could* ocorrem na voz passiva geralmente expressando “possibilidade”. Essas construções são de dois tipos: sem a presença do agente, ou com um processo nominalizado com a presença de *by*. Em ambos os casos, o uso de *can* e *could* com a voz passiva evita a identificação explícita do agente humano do verbo principal, ou seja, da pessoa em condições de realizar a ação sugerida. Como resultado, o significado entendido é de que a situação é logicamente possível:

*Each interpretation **can** be seen generally to flow through the abbreviated text as a whole.*

*The methods **could** be refined and made more accurate.*

*Its answer **can** be illustrated **by** considering again the action of a polarizer.*

Must e should apresentam também uma considerável frequência com a voz passiva na prosa acadêmica, geralmente expressando um tipo de “obrigação coletiva”. Assim, a voz passiva é usada para evitar a identificação explícita da pessoa que é obrigada a agir (Biber et al., 1999, p. 500):

*Care **must** be taken to ensure that the diffusion in the stator is kept to a reasonable level.*

*It **should** be noted that the following scenario is nothing more than one of many potential scenarios.*

Essa estrutura é bastante comum no discurso técnico (SAGER et al., 1980; TRIMBLE, 1985) principalmente orações que expressam instruções:

*Particularly for heavier weld sections the edges of the metal **must** be prepared for submerged arc welding. Various methods **may** be employed in preparing the edges for welding... Metal cutting or grooving **can** also be accomplished by the carbon-arc method. (TRIMBLE, 1985, p. 120)*

Além dos cenários distribucionais até aqui apresentados, Biber et al. (1999, p. 490) verificaram que alguns verbos lexicais ocorrem preferivelmente acompanhados de um VM, ao invés de ocorrerem em orações com marcação de tempo verbal, como geralmente acontece. Os seguintes verbos ocorrem na companhia de um VM em aproximadamente 60% de suas ocorrências:

abide, admit, afford, appeal, assure, cope (with), discern, fathom, guarantee, handle, imagine, interact, resist, settle for, suffice, survive, tolerate

Enquanto que os que seguem em 40% de suas ocorrências:

account for, advise, aid, believe, benefit, claim, compete, continue, contribute, count on, debate, deduce, endanger, end up, expect, exit, focus, forgive, get over, grumble, harm, insure, jeopardize, muster, outweigh, overwhelm, pause, predispose, reach, rely, respect, settle for, solve, tempt, withstand.

Para os autores, os verbos “mentais” apresentam a mais forte associação com os VMs. Esses verbos mentais geralmente expressam várias emoções, atitudes, ou estados cognitivos que são intrinsecamente pessoais, e, dessa forma, co-ocorrem com os modais expressando uma posição pessoal. Nos seus dados, os modais *can* e *could*, geralmente combinados com negação, são particularmente comuns com os verbos mentais:

*You **can imagine** what it's like.*

*I mean as an acquaintance I **could tolerate** her.*

*I **can't cope** with this.*

*I **can't abide** the dratted things.*

*The council **cannot afford** to maintain them.*

Kennedy (2002) acredita que as descrições lingüísticas baseadas em corpus são relevantes por adicionarem, entre outras coisas, uma dimensão distribucional à descrição lingüística. Embora as descrições gramaticais não sejam inteiramente previsíveis, seus usos na educação lingüística para o desenvolvimento de programas e ementas e novos materiais de ensino estão tornando-se cada vez mais comuns. Tendo em vista a análise de ocorrências dos VMs no BNC o autor sugere que os nove modais centrais continuam a ser importantes para o ensino da língua inglesa, apesar de o VM *shall* ser consideravelmente menos freqüente do que os outros. Todavia, levando-se em consideração os resultados obtidos do OSC (DENARDIN, 2008) e as constatações feitas por Sager et al. (1980) e Trimble (1985) pode-se concluir que, o ensino de ESP (English for Specific Purposes) deve levar em consideração as particularidades específicas de cada gênero, que só podem ser reveladas através de descrições detalhadas baseadas em *corpora* específicos.

Tendo em vista essa imbricada rede de significados, usos, formas, estruturas sintáticas, freqüências e distribuição que permeiam os VMs, que pretendo oferecer a minha contribuição para o conhecimento da linguagem da aviação, em especial para o gênero textual “manuais de aviação”. Acredito que através da verificação da freqüência de ocorrência de cada VM, sua distribuição nos três manuais aqui analisados, comparação com corpora de inglês geral e outros manuais, verificação das estruturas sintáticas, e investigação das colocações, tanto à esquerda, quanto à direita, seja possível estabelecer uma taxonomia de uso dos VMs neste cenário. Com essa taxonomia busco contribuir para o desenvolvimento de materiais didáticos que melhor reflitam o que realmente acontece nesses manuais.

5 OS CORPORA DE ANÁLISE E SUA SELEÇÃO

Neste capítulo descrevo os três corpora utilizados neste estudo, são eles:

- 1- Corpus de estudo (AC)
- 2- Corpus de referência (BNC)
- 3- Corpus de contraste (OSC)

5.1 CORPUS DE ESTUDO

É importante caracterizar os textos que são objetos de estudo. De uma forma geral, os manuais de aviação são obras com tamanhos e objetivos variados. Alguns são mais extensos, como os manuais de manutenção (MM) e o manual de operações (OM), enquanto outros são mais curtos e concisos, como por exemplo, o manual de referência rápida (QRH). Normalmente apresentam um prefácio dirigido aos leitores e, na seqüência, o sumário. Os capítulos dos manuais são subdivididos em itens. Cada item apresenta descrições e instruções, geralmente em forma de *check-lists*. Algumas partes dos manuais caracterizam-se por um uso maior de figuras; outros apresentam mais texto.

Os manuais estão restritos à área da aviação, contudo, existem diferentes classes de profissionais envolvidos. Os manuais são escritos por técnicos redatores e serão lidos por mecânicos (MM) e pilotos (OM, QRH); há uma relação de “obediência” entre os interlocutores, tendo-se em mente que os destinatários devem executar com precisão as diretrizes postas pelos redatores dos manuais. Pode-se dizer que os destinatários dos manuais aqui analisados formam duas comunidades discursivas (Swales, 1990) distintas e ocupam diferentes posições na escala hierárquica de uma empresa aérea. Os pilotos, destinatários do OM e do QRH, ocupam a posição mais alta na empresa. O treinamento de um piloto é altamente dispendioso, custando às empresas aéreas alguns milhares de dólares ao ano. Os pilotos tornam-se assim, funcionários importantes de serem mantidos em uma empresa. Além disso, eles podem ocupar posições gerenciais e participar de importantes decisões em uma empresa, como por exemplo, decidir qual tipo de aeronave comprar. Os mecânicos, que são os interlocutores do MM, não possuem poder semelhante. Hierarquicamente, estão mais próximos à base. Outra diferença importante diz respeito ao local de trabalho. Para os pilotos, seguir as instruções em um manual não é só uma questão de eficiência, mas literalmente de “vida ou morte”, uma vez que eles estão sempre a bordo das aeronaves. Os mecânicos

realizam seu trabalho no solo. Sendo uma ação erroneamente executada não colocará a sua vida em risco, mas sim dos passageiros e pilotos.

Os textos desses manuais são considerados portadores do conhecimento sobre as aeronaves, e não são passíveis de questionamentos por parte dos destinatários. Possuem um caráter altamente deôntico. São textos: descritivos, divididos em capítulos por partes da aeronave (hidráulica, trem de pouso, escadas, etc.) e/ou procedurais, divididos por tipo de procedimentos (decolagem, pouso, etc.). São ricos em informações não verbais, com várias tabelas, figuras e gráficos.

A escolha do corpus¹¹⁰ foi movida por alguns fatores, entre eles cito os dois principais: 1) o Boeing 737 é a aeronave mais utilizada pelas empresas aéreas brasileiras, dessa forma, seus manuais são os mais utilizados em cursos de treinamento no Brasil; e 2) é bastante difícil ter acesso a esse tipo de publicação, mas, por ser o 737 o tipo de aeronave utilizada (para operar, ou em treinamento) em instituições em que trabalhei, obtive acesso aos dados.

O AC compreende três manuais da aeronave Boeing 737: (i) o manual de manutenção (MM) destinado a mecânicos e técnicos de aeronaves com um total de 249.691 palavras (*tokens*); (ii) o manual de referência rápida (QRH), direcionado a pilotos com 29.929 palavras; e, finalmente (iii) o manual de operações (OM) também para pilotos, com 183.959 palavras. O número total de *tokens* nos referidos manuais é de 463.585. Berber Sardinha (2000) classifica um corpus com aproximadamente 500.000 *tokens* como sendo de tamanho médio, e acredita ser significativo, principalmente levando-se em consideração o grau de especificidade do AC. Descrevo cada um dos sobcorpora a seguir.

O MM¹¹¹ contém as informações necessárias para revisões, resoluções de problemas, verificações, reparos e substituições de todos os sistemas e equipamentos instalados nas aeronaves 737-300/400/500. Tais ações são geralmente realizadas quando as aeronaves estão nos hangares, mas podem também ser realizadas com os aviões na pista. Além disso, o MM contém informações relacionadas à inspeção e manutenção das estruturas das aeronaves. Segue alguns exemplos de concórdâncias do MM com o modal can:

¹¹⁰ Aqui me refero à decisão de ter escolhido os manuais da Boeing em detrimento, por exemplo, dos manuais da AIRBUS ou Embraer.

¹¹¹ Entre os três que compõem o corpus de estudo (AC), apenas o MM não foi utilizado na íntegra, pois obtive acesso somente a algumas partes dele. Além disso, o manual como um todo é muito maior do que os outros dois manuais para pilotos juntos, somando 98 capítulos. Os capítulos utilizados foram: CHAPTER 00 – INTRODUCTION; CHAPTER 05 – TIME LIMITS/MAINTENANCE CHECKS; CHAPTER 12 – SERVICING; CHAPTER 20 – STANDARD PRACTICES

N Concordance

1 THE OIL STAY ON YOUR SKIN. YOU CAN ABSORB POISONOUS
 2 YOUR SKIN FOR A LONG TIME. YOU CAN ABSORB POISONOUS
 3 OIL TOUCH YOUR SKIN. YOU CAN ABSORB TOXIC MATERIALS
 4 OIL TOUCH YOUR SKIN. YOU CAN ABSORB TOXIC MATERIALS
 5 SEALER FOR THE DECAL. YOU CAN BREAK THE EDGE SEAL WITH
 6 USE AN EXTENSION HANDLE. YOU CAN CAUSE DAMAGE TO SEALED
 7 THAN 55 PSI FUEL PRESSURE, YOU CAN . CAUSE DAMAGE TO THE
 8 IF YOU ARE NOT CAREFUL YOU CAN CAUSE DAMAGE TO THE
 9 IF YOU ARE NOT CAREFUL, YOU CAN CAUSE DAMAGE TO THE
 10 FROM SURFACES THAT MOVE. YOU CAN CAUSE FAILURE OF MOVING
 11 THE OTHER ENGINE PARTS. YOU CAN CAUSE INJURY TO YOURSELF IF
 12 S 683-060 C_A_U_T_I_O_N_: IF YOU CAN DRAIN THE FUEL FROM THE
 13 ONE THREAD . ENGAGED. YOU CAN COMPRESS THE UNION AND
 14 THE VACUUM BREAK VALVE, YOU CAN DAMAGE THE VSCF. (4) ON
 15 THE VACUUM BREAK VALVE, YOU CAN DAMAGE THE VSCF. (3) Remove
 16 THE VACUUM BREAK VALVE, YOU CAN DAMAGE THE VSCF. (a) ON

O OM fornece as informações relativas aos sistemas da aeronave, desempenho, procedimento, e limitações nas operações. Isto é, todas as informações que a tripulação de voo necessita saber para operar a aeronave 737 de forma eficiente e segura durante todas as operações previstas.

N Concordance

1 extends outboard of measuring stick 4 can be used as an indication of the
 2 Increased thrust lever activity can be expected when encountering
 3 disabled. 8 Speed Brake Lever Actuator Can be armed to raise ground spoilers
 4 enabled. 8 Speed Brake Lever Actuator Can be armed to raise ground spoilers
 5 total hydraulic power failure the ailerons can be mechanically positioned by
 6 realignment cycle before the airplane can be moved. April 28, 2000 737
 7 for the origin and destination airport can be selected for entry into the flight
 8 -1000 feet. Airspeed Entry Airspeeds can be entered into the FMC as
 9 windshear. Predictive windshear alerts can be enabled prior to takeoff by
 10 windshear. Predictive windshear alerts can be enabled prior to takeoff by

Finalmente, o QRH consiste em *checklists* para “Procedimentos Padrão” e “Procedimentos Não Padrão”. Os *checklists* de procedimentos Padrão são utilizados para verificar se determinados procedimentos essenciais foram realizados. *Checklists* Não-padrão,

por conseguinte, incluem informações que a tripulação de vôo deverá usar no caso de emergências.

N Concordance

- 1 . . ESTABLISH If smoke/fumes source can be determined: ELECTRICAL
- 2 be jammed: (continued) If rudder pedals can be centered: Accomplish the normal
- 3OFF If rudder pedals can be centered: Accomplish JAMMED
- 4 . . ESTABLISH If smoke/fumes source can be determined: ELECTRICAL
- 5 . . ESTABLISH If smoke/fumes source can be determined: ELECTRICAL
- 6 be jammed: (continued) If rudder pedals can be centered: Accomplish the normal
- 7OFF If rudder pedals can be centered: Accomplish JAMMED

É ainda importante ressaltar que as aeronaves 737 requerem dois tripulantes de cabine: um comandante e um co-piloto. Cada um dos membros lê, em voz alta, os itens dos *checklists* de acordo com sua área de responsabilidade enquanto o outro membro executa ou verifica a situação do item. No quadro 5.1.A resumo as principais informações de cada manual, incluindo a sua dimensão.

Quadro 5.1.A: Síntese do AC

MANUAL	DATA DE REVISÃO	DESTINATÁRIO	OBJETIVO	TOKENS
737-300/400/500 MAINTENANCE MANUAL MM	NOVEMBRO 1998	MECÂNICOS DE AERONAVES	Informações necessárias para a manutenção das aeronaves (montadas) ¹¹² nos hangares	249.691
737-7L9 OPERATIONS MANUAL OM	JULHO 2000	PILOTOS	Informações relativas aos sistemas da aeronave; para ser utilizado em treinamento ou em vôo	183.959
737-7J0 QUICK REFERENCE HANDBOOK QRH	ABRIL 2000	PILOTOS	<i>Checklists</i> para Procedimentos Padrão e Não-padrão; para ser utilizado em vôo	29.929
AVIATION CORPUS AC (MM+OM+QRH)				463.585

¹¹² Informações sobre a revisão ou manutenção de partes da aeronave que precisem ser removidas, são encontradas no manual 737 *Overhaul Manual/Component Maintenance Manual*.

Além das informações acima, caracterizo agora o corpus conforme a tipologia de Biber (1993) com vários aspectos referentes aos corpora que representam uma variedade específica da língua inglesa:

Quadro 5.1.B: Caracterização do corpus de estudo (AC)

Canal	Escrito
Formato	Publicado
Cenários	Institucional / Institucional
Receptor	Mecânicos e pilotos de aeronaves
Produtor	Instituição (autor não divulgado)
Factualidade	Informação Factual
Propósito	Informar / Instruir / Descrever

5.2 CORPUS DE REFERÊNCIA

Como já mencionado, somente através de comparações e contrastes é possível delimitar o que é verdadeiramente peculiar em um corpus de linguagem específica. Dessa forma, utilizei o BNC como o corpus de referência. Esse contraste tem por objetivo “fornecer uma norma com a qual se fará a comparação das frequências do corpus de estudo” (BERBER SARDINHA, 2004, p. 97). Para que um corpus seja considerado de referência, ele não pode conter o corpus de estudo e deve ser bem maior do que o corpus de estudo. Segundo Berber Sardinha (2004), “Os tamanhos críticos de corpora de referência são 2, 3 e 5 vezes o tamanho do corpus de estudo.” (p. 97). O BNC é, na verdade, aproximadamente 200 vezes maior do que o corpus de estudo. O BNC escrito, que foi a parte do corpus de referência utilizada para os contrastes, possui em torno de 90 milhões de palavras, ou seja, é aproximadamente 180 vezes maior do que o AC.

O *British National Corpus* (BNC) foi escolhido por ser um dos corpora de inglês geral mais representativos e balanceados. Os textos escritos foram selecionados através de três critérios: domínio (área), data e meio. O domínio refere-se ao tipo de conteúdo ou área do texto; a data refere-se ao período em que o texto foi escrito (o BNC é um corpus diacrônico), enquanto meio refere-se ao tipo de publicação, tal como livro, jornal, periódico, entre outros. McEnery et al. (2006, p. 17, 18) resumiram as informações provenientes do BNC escrito no quadro abaixo:

Quadro 5.2.A: BNC escrito

Domínio	%	Data	%	Meio	%
Imaginativo	21.91	1960-74	2.26	Livro	58.58
Artes	8.08	1975-93	89.23	Periódico	31.08
Crenças e pensamento	3.40	Não classificado	8.49	Miscelânea Publicada	4.38
Comércio/finanças	7.93			Miscelânea Não publicado	4.00
Lazer	11.13			Para-ser-falado	1.52
Ciência pura e natural	4.18			Não classificado	
Ciência aplicada	8.21				
Ciência Social	14.80				
Assuntos gerais	18.39				
Não classificado	1.93				

5.3 CORPUS DE CONTRASTE

O corpus de contraste utilizado é composto de dois manuais de sistemas operacionais para computadores pessoais, o OSC. Corpora de contraste refere-se a um corpus utilizado em comparação ao corpus de estudo, mas que não preenche as exigências (como tipo de linguagem e dimensão) para que seja considerado um corpus de referência. Essa tripla comparação (estudo x referência x contraste) é importante de forma a verificar o que é peculiar das linguagens especializadas e o que constitui especificidades somente do corpus de estudo, aqui caracterizado pelos manuais de aviação.

O *Operating System Corpus* (ou OSC, gentilmente cedido por Denardin) compreende dois guias de sistemas operacionais: Windows Vista e Linux. Esse corpus foi escolhido por também tratar-se de um manual técnico e escrito na língua inglesa. Dessa forma, faz parte do macrogênero manuais técnicos¹¹³. Os textos foram obtidos através dos websites oficiais dos sistemas.¹¹⁴ Um sistema operacional pode ser visto como um programa de grande complexidade que é responsável por todo o funcionamento de uma máquina, desde o software, a todo o hardware instalado na máquina. Todos os processos de um computador estão por de trás de uma programação complexa que comanda todas as funções que um

¹¹³ A noção de macrogênero textual foi descrita no capítulo 3.

¹¹⁴ *Windows Vista Product Guide*: <https://www.microsoft.com/downloads/details.aspx?FamilyID=bbc16ebf-4823-4a12-afe1-5b40b2ad3725&DisplayLang=en> e *Linux Guide*: <http://tldp.org/LDP/intro-linux/intro-linux.pdf>

utilizador impõe à máquina. Existem vários sistemas operacionais; entre eles, os mais utilizados no dia a dia, normalmente utilizados em computadores domésticos, são o *Windows* e *Linux*. Os manuais do OSC contém informações sobre a descrição dos sistemas e sua utilização. A principal diferença entre esse tipo de corpus (OSC) e o corpus de estudo (AC), é com relação ao tipo de destinatário. Os manuais do AC são escritos por especialistas (redatores técnicos) para especialistas (mecânicos e pilotos). O manuais do OSC são escritos por especialistas (redatores técnicos) mas o público alvo é constituídos por não-especialistas, isto é, os manuais destinam-se aos usuários leigos dos sistemas operacionais em questão. Por vezes, profissionais que realizam serviços de manutenção de software podem também utilizarem-se das informações desses manuais, mas os manuais não os têm como alvo, conforme os respectivos prefácios. Esses prestadores de serviço são pessoas que realizaram um curso técnico, geralmente de curta duração, não possuindo, ao meu ver, um alto grau de especialização. Além disso, a grande maioria dos usuários dos sistemas operacionais não consulta os manuais para utilizar os sistemas. As ferramentas e técnicas dos sistemas são aprendidas na prática. Outra diferença contextual entre o AC e o OSC é relacionada às consequências que ações má executadas podem ter. Talvez a pior consequência possível quanto à não conformidade às instruções dos manuais do OSC seja danificar um computador. Entretanto, os leitores do AC podem colocar muitas vidas em risco ao não realizar as instruções corretamente. Considero importante enfatizar essas diferenças contextuais uma vez que o objeto mais específico de análise desta tese são os verbos modais, que, como já visto no capítulo 4, são amplamente utilizados nos atos de fala relacionados a instruções. Dessa forma deve-se considerar a possibilidade de essas diferenças dos contextos de uso dos manuais influenciarem os redatores quanto a qual verbo modal utilizar para expressar determinados atos.

No quadro 5.3.A resumo as principais informações de cada manual do OSC, incluindo a sua dimensão.

Quadro 5.3.A: Síntese do OSC

MANUAL	DATA DE REVISÃO	DESTINATÁRIO	OBJETIVO	TOKENS
WINDOWS VISTA	2008	USUÁRIOS	Informações necessárias para a utilização do sistema	93.277
LINUX	2008	USUÁRIOS	Informações necessárias para a utilização do sistema	81.305
Operating System Corpus OSC Windows Vista + LINUX				174.582

Conforme a tipologia estabelecida por Biber (1993), o OSC é assim caracterizado:

Quadro 5.3.B: Caracterização do corpus de contraste

Canal	Escrito
Formato	Publicado
Cenários	Institucional / Pessoal
Receptor	Usuários leigos
Produtor	Instituição (autor não divulgado)
Factualidade	Informação Factual
Propósito	Informar / Instruir / Descrever /

O próximo capítulo deste trabalho apresenta a metodologia construída para conduzir a investigação da tese.

6 METODOLOGIA

“A observação contribui mais para a teoria do que a teoria contribui para a observação” (LEECH, 1992, p. 111). Seguindo nessa linha e diante da inexistência de um modelo específico para a análise de VMs em textos técnicos, procurei estabelecer os princípios que nortearam a escolha e a delimitação das unidades de análise, bem como a estruturação das várias fases da pesquisa. A metodologia tem o reflexo de estudos prévios, que contribuíram para a compreensão do emprego dos VMs nos manuais de aviação e para o amadurecimento do presente trabalho (SARMENTO, 2005a; SARMENTO, 2005b; SARMENTO 2007; GABRIELATOS e SARMENTO 2006).

Antes de iniciar o detalhamento sobre a metodologia, considero importante esclarecer a abordagem metodológica que é empregada neste trabalho. Uso a LdC como metodologia, na crença de que, pelo menos para o objetivo desta tese, que é conhecer a incidência e uso dos VMs em manuais de aviação de forma a fornecer subsídios para a elaboração de material didático, seja o método existente mais eficaz para conhecer a realidade desses manuais. Quando à dicotomia *corpus-based* e *corpus-driven*, tomo uma posição intermediária, consoante McEnery et al. (2006).

Considero também importante retomar as questões que norteiam o desenvolvimento deste estudo, cujo objetivo geral é analisar a incidência e a funcionalidade dos verbos modais no segmento da linguagem da aviação representado pelo texto do tipo manual técnico operacional e de manutenção de forma a colher subsídios para a elaboração futura de material didático.

Para alcançar esse objetivo, sigo o seguinte roteiro metodológico composto de quatro etapas:

- i) Contrasto as frequências de cada VM entre o corpus de estudo (AC), o corpus de referência (BNC) e o corpus de contraste (OSC), observando se há diferenças quantitativas significativas no uso dos VMs.
- ii) Contrasto as frequências das estruturas sintáticas em que os VMs se encontram entre o corpus de estudo (AC), o corpus de referência (BNC escrito) e o corpus de contraste (OSC), observando se há diferenças significativas no uso das estruturas.

- iii) Verifico as diferenças quanto à frequência de cada VM e de cada estrutura sintática em cada um dos três subcorpora que compõem o corpus de estudo.
- iv) Identifico as colocações de cada VM em cada um dos corpora.

Para desenvolver as etapas (i) e (ii) acima, acessei três corpora diferentes que foram detalhados no capítulo 5.

- O Corpus de Estudo, ou Aviation Corpus (AC), composto de três manuais técnicos da aeronave 737 da BOEING, sendo dois manuais para pilotos (OM e QRH) e um manual para mecânicos (MM). Os manuais utilizados já estavam em formato digital, formato pdf não protegido. Isso exigiu que os textos fossem transformados para arquivo.txt para que pudessem ser estudados através da ferramenta *WordSmith Tools*, pois a ferramenta somente aceita textos no formato .txt. Após a conversão, foi feita uma “limpeza” manual dos textos, pois, devido ao processo de conversão foram incluídos caracteres adicionais no lugar de figuras e tabelas.

- O Corpus de referência, ou British National Corpus (BNC escrito), composto de textos escritos do inglês geral produzidos até 1993. Esse corpus é apresentado em vários formatos. Utilizei o BNCWeb, que pode ser acessado pelo endereço <http://escorp.unizh.ch/>, mediante obtenção da licença de uso. Essa versão já possui as próprias ferramentas para estudar o corpus. As ferramentas do BNCWeb utilizadas nesta pesquisa são semelhantes às utilizadas nos outros dois corpora através do *WordSmith*.

- O Corpus de contraste, ou Operation Systems Corpus (OSC), composto de dois manuais de sistemas operacionais de computadores pessoais. Esse corpus já havia sido utilizado por Dernardin (2008) e estava pronto (convertido e formatado em .txt). Para o estudo do OSC foi também utilizado o *WordSmith Tools*.

Analiso três corpora primeiramente para verificar o que diferencia o inglês geral (BNC) do inglês específico (AC e OSC). Em segundo lugar, para verificar o que é específico dos manuais de aviação (AC) em comparação com outro tipo de manuais técnicos (OSC).

As etapas (iii) e (iv) serão desenvolvidas mediante o acesso ao corpus de estudo (AC) somente.

Como mencionado, para a análise dos corpora de estudo e de contraste, utilizei a ferramenta *WordSmith*, que é um conjunto de programas integrados destinados à análise linguística. O *WordSmith* foi desenvolvido em 1996 por Mike Scott e publicado pela Oxford

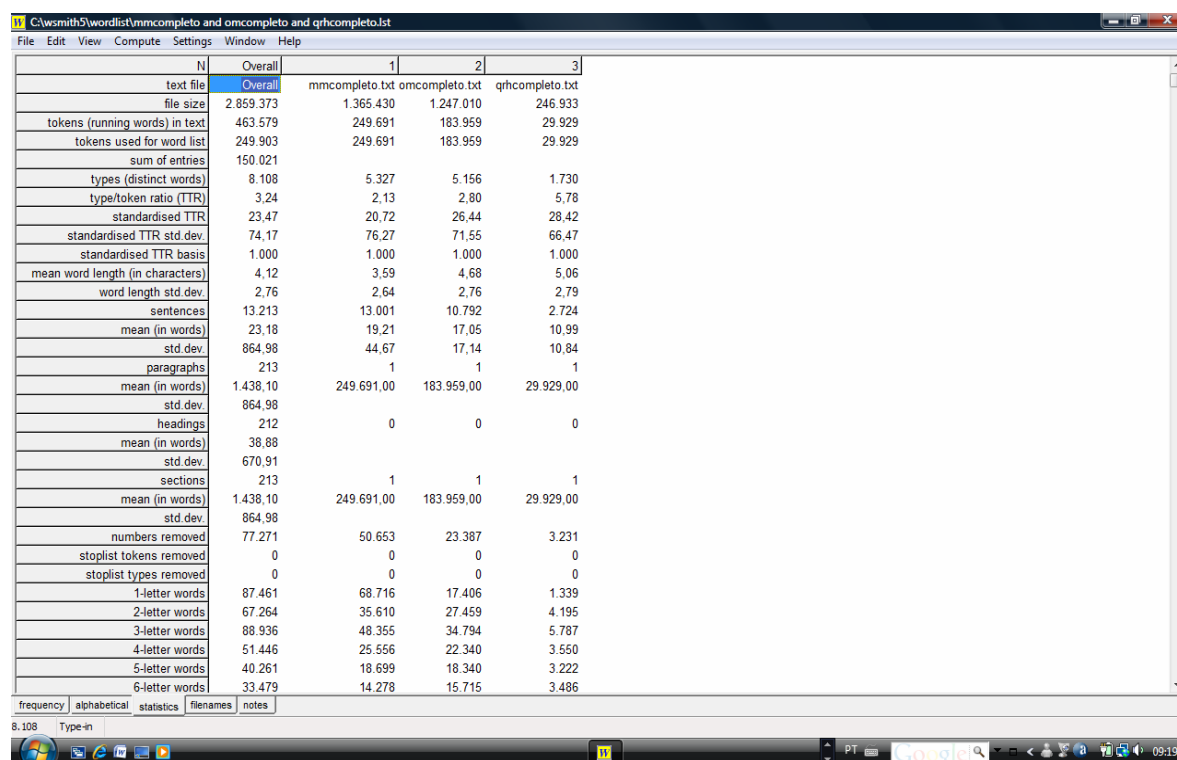
University Press e está atualmente em sua quinta versão. Neste trabalho faço uso da versão 5¹¹⁵.

O *WordSmith* possui as ferramentas- *WordList*, *Keywords*, *Concord*. Neste trabalho utilizei principalmente algumas funções do *Wordlist* e do *Concord*.

A seguir detalho as quatro etapas que compõem o roteiro metodológico desta tese.

i) Contraste das frequências de cada VM entre o corpus de estudo (AC), o corpus de referência (BNC) e o corpus de contraste (OSC), observando se há diferenças quantitativas significativas no uso dos VMs.

Num primeiro momento utilizei o *WordList* para verificar o número exato de *tokens* em cada um dos subcorpora que compõe o AC e o OSC:



N	Overall	1	2	3
text file	Overall	mmmcompleto.txt	omcompleto.txt	qhcompleto.txt
file size	2.859.373	1.365.430	1.247.010	246.933
tokens (running words) in text	463.679	249.691	183.959	29.929
tokens used for word list	249.903	249.691	183.959	29.929
sum of entries	150.021			
types (distinct words)	8.108	5.327	5.156	1.730
type/token ratio (TTR)	3.24	2.13	2.80	5.78
standardised TTR	23.47	20.72	26.44	28.42
standardised TTR std.dev.	74.17	76.27	71.55	66.47
standardised TTR basis	1.000	1.000	1.000	1.000
mean word length (in characters)	4.12	3.59	4.68	5.06
word length std.dev.	2.76	2.64	2.76	2.79
sentences	13.213	13.001	10.792	2.724
mean (in words)	23.18	19.21	17.05	10.99
std.dev.	864.98	44.67	17.14	10.84
paragraphs	213	1	1	1
mean (in words)	1.438,10	249.691,00	183.959,00	29.929,00
std.dev.	864,98			
headings	212	0	0	0
mean (in words)	38,88			
std.dev.	670,91			
sections	213	1	1	1
mean (in words)	1.438,10	249.691,00	183.959,00	29.929,00
std.dev.	864,98			
numbers removed	77.271	50.653	23.387	3.231
stoplist tokens removed	0	0	0	0
stoplist types removed	0	0	0	0
1-letter words	87.461	68.716	17.406	1.339
2-letter words	67.264	35.610	27.459	4.195
3-letter words	88.936	48.355	34.794	5.787
4-letter words	51.446	25.556	22.340	3.550
5-letter words	40.261	18.699	18.340	3.222
6-letter words	33.479	14.278	15.715	3.486

Figura 1: Tela com as estatísticas fornecidas pela *WordList*.

Num segundo momento, foi utilizada a ferramenta *Concord* para obter as concordâncias através de uma lista contendo todas as ocorrências de cada um dos VMs no contexto. Cada palavra investigada é destacada no centro de um fragmento do texto, conforme figura abaixo. Essa ferramenta foi utilizada para verificar a quantidade de cada um dos VMs

¹¹⁵ A licença para utilizar tal ferramenta pode ser obtida através do site: <http://www.lexically.net/wordsmith/>.

no AC e no OSC. Para a verificação do BNC, foi utilizada a ferramenta presente no próprio corpus.

N	Concordance	Set	Tag	Word #	t	os	#	os	t	#	os	File	%
1	all types of windshear. The flight crew must continue to rely on traditional	08.472	689	8%	0	9%	0	9%	0	9%	0	omcompleto.txt	99%
2	and A or B pressure lost. Wheel speed must be greater than 60 knots for	01.761	372	3%	0	6%	0	6%	0	6%	0	omcompleto.txt	96%
3	and A or B pressure lost. Wheel speed must be greater than 60 knots for	01.229	313	3%	0	6%	0	6%	0	6%	0	omcompleto.txt	96%
4	prior to takeoff the following conditions must exist: • airplane on the ground •	00.200	236	8%	0	5%	0	5%	0	5%	0	omcompleto.txt	95%
5	deceleration rate for landing • switch must be pulled out to select MAX	98.589	141	9%	0	4%	0	4%	0	4%	0	omcompleto.txt	94%
6	DISCONTINUITY The entered route must always form a continuous path of	87.903	565	0%	0	9%	0	9%	0	9%	0	omcompleto.txt	89%
7	an updated speed, the current speed must be deleted or a different Vref	85.548	457	3%	0	8%	0	8%	0	8%	0	omcompleto.txt	88%
8	mode is ECON PATH. The crew must select a speed descent mode. The	81.731	280	4%	0	6%	0	6%	0	6%	0	omcompleto.txt	87%
9	can be omitted • distance only entries must start with a / . ETA – displays the	81.108	245	6%	0	6%	0	6%	0	6%	0	omcompleto.txt	86%
10	VORTAC, or DME stations that must not be used for FMC position	80.799	226	3%	0	6%	0	6%	0	6%	0	omcompleto.txt	86%
11	system is invalid, the invalid data must be inhibited to prevent incorrect	80.741	224	9%	0	6%	0	6%	0	6%	0	omcompleto.txt	86%
12	runway data display page. Runways must be stored in the permanent	79.726	175	0%	0	5%	0	5%	0	5%	0	omcompleto.txt	86%
13	RW13R, and the proper airport identifier must be entered on the runway data	79.716	174	0%	0	5%	0	5%	0	5%	0	omcompleto.txt	86%
14	the waypoint method is used, entry must be in the form RW13R, and the	79.705	174	0%	0	5%	0	5%	0	5%	0	omcompleto.txt	86%
15	at DEF is depicted). The entry must be executed. -----	77.797	079	7%	0	5%	0	5%	0	5%	0	omcompleto.txt	85%
16	of which geographical location to use must be made before the waypoint can	73.847	89	3%	0	3%	0	3%	0	3%	0	omcompleto.txt	83%
17	shows a discontinuity. The airplane must fly direct from LAC01 to BANDR.	73.053	86	4%	0	2%	0	2%	0	2%	0	omcompleto.txt	83%
18	BEEZR, and ELN. The airplane must fly direct from LAC01 to BANDR.	72.329	83	4%	0	2%	0	2%	0	2%	0	omcompleto.txt	83%
19	whenever necessary. The new waypoint must first be placed into the CDU	71.303	78	8%	0	1%	0	1%	0	1%	0	omcompleto.txt	83%
20	by overwriting displayed data. Entry must be in one of the following forms: •	69.025	66	6%	0	0%	0	0%	0	0%	0	omcompleto.txt	82%
21	Waypoint (RTA WPT) Waypoint entry must be in flight plan or the CDU	68.494	64	9%	0	0%	0	0%	0	0%	0	omcompleto.txt	81%
22	prompt key is pushed, the route must be executed by pushing the	67.237	59	5%	0	0%	0	0%	0	0%	0	omcompleto.txt	81%
23	that provides a steeper climb angle must be manually selected. If a CLB 1 or	65.681	49	0%	0	9%	0	9%	0	9%	0	omcompleto.txt	80%
24	REF and DEPARTURES pages must be modified to agree. The	65.427	48	0%	0	9%	0	9%	0	9%	0	omcompleto.txt	80%
25	waypoint of the previous route segment must be the same as the beginning point	60.722	205	2%	0	6%	0	6%	0	6%	0	omcompleto.txt	78%
26	if the entry is valid. The route segment must contain the waypoint entered in the	60.705	204	8%	0	6%	0	6%	0	6%	0	omcompleto.txt	78%
27	a TO waypoint first. Valid airways must : • contain the fix entered in the TO	60.604	199	2%	0	6%	0	6%	0	6%	0	omcompleto.txt	78%
28	waypoint of the previous route segment must be the same as the beginning point	60.058	169	2%	0	6%	0	6%	0	6%	0	omcompleto.txt	78%
29	if the entry is valid. The route segment must contain the waypoint entered in the	60.041	168	8%	0	6%	0	6%	0	6%	0	omcompleto.txt	78%
30	a TO waypoint first. Valid airways must : • contain the fix entered in the TO	59.940	163	2%	0	6%	0	6%	0	6%	0	omcompleto.txt	78%
31	is required for route activation. Entries must be in the navigation database. July	59.597	139	9%	0	6%	0	6%	0	6%	0	omcompleto.txt	77%
32	for activation of the route. Entries must be in the navigation database. New	59.544	133	8%	0	6%	0	6%	0	6%	0	omcompleto.txt	77%
33	Navigation - FMC Preflight Valid entries must be in the navigation database.	59.308	113	1%	0	6%	0	6%	0	6%	0	omcompleto.txt	77%

Figura 2: Ferramenta Concord com linhas de concordância de *must* no AC.

É possível notar que o nódulo *must* está destacado em azul na coluna central. No final da tela, à esquerda, observa-se o número 172. Ele refere-se ao número total de ocorrências do nódulo *must* no OM (Operation Manual), um dos subcorpora que compõe o AC. Neste momento não foi necessária uma organização (*sort*) à esquerda ou à direita do nódulo, pois o objetivo era unicamente realizar um levantamento quantitativo do número de ocorrências de cada VM. O AC e o OSC foram investigados sem qualquer tipo de anotação (POS tag ou sintática), uma vez que as entradas ortográficas dos VMs não possuem qualquer tipo de homógrafos nesses dois corpora. Para o BNC, foi necessário realizar uma busca pelo *Lema type: verb* (POS tag), pois algumas entradas ortográficas, por exemplo, *may*, além de ser um VM, é também um mês. Dessa forma, se POS tag não fosse usada, não seria possível obter a quantidade precisa de *may* como VM. Um exemplo dessa estratificação pode ser visto na próxima figura 3.

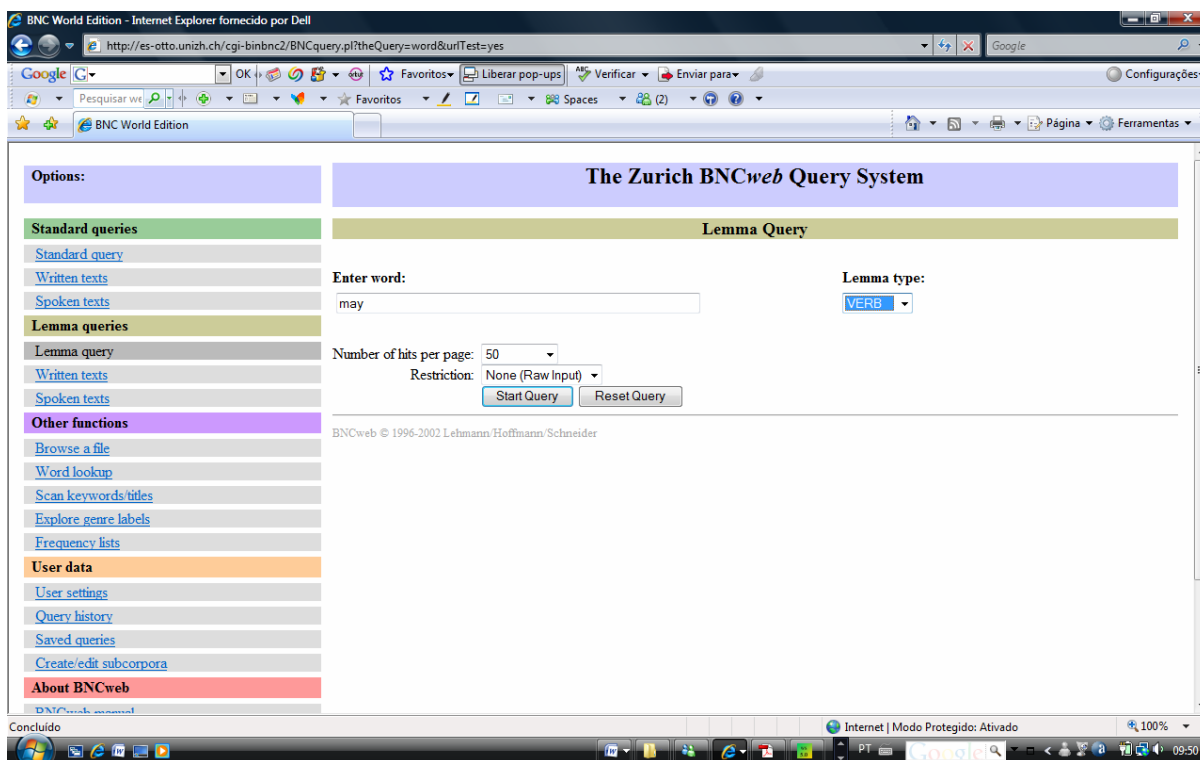


Figura 3: Busca por lema no BNCWeb

Após essa tela obtém-se o seguinte resultado, conforme figura 4:

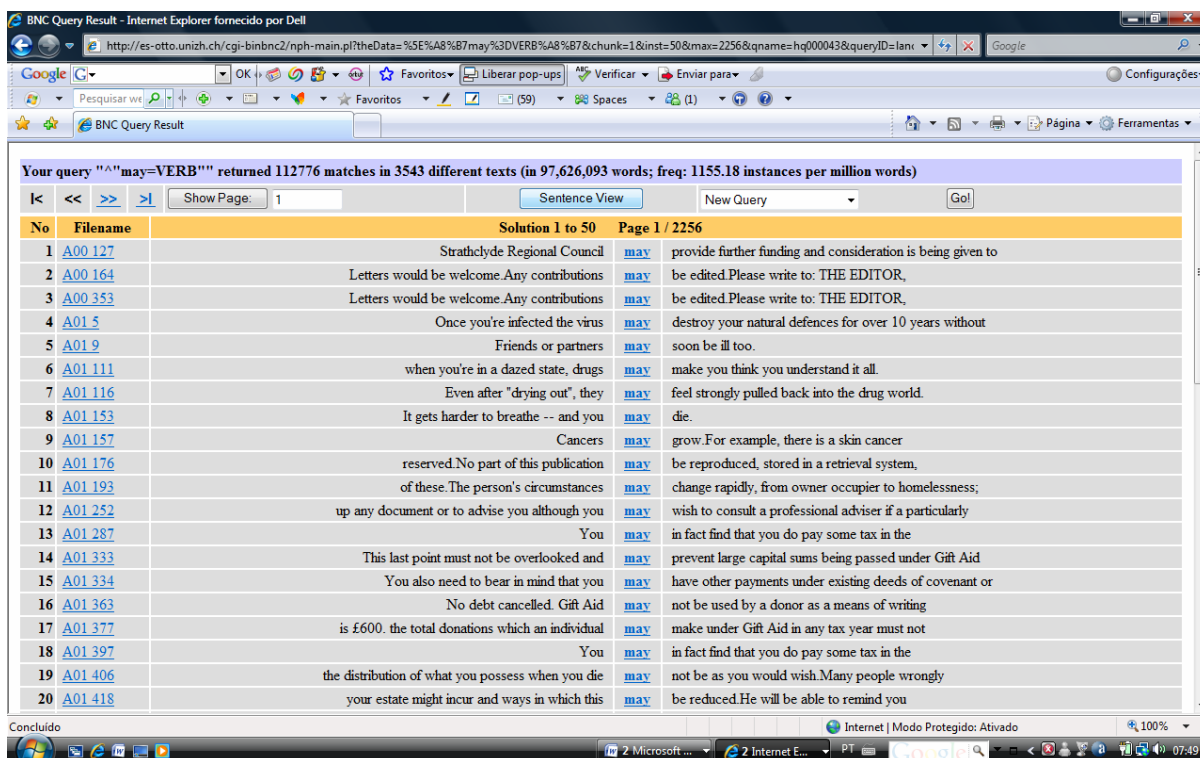


Figura 4: Linhas de concordância de *can-* BNC

No topo da tela, pode-se notar o tipo de busca (*may=VERB*), o número de unidades encontradas (112776), o número de textos nos quais *may* está presente (3543), o número total

de palavras, ou *tokens*, no corpus (97.626.093) e a quantidade de *may* a cada um milhão de palavras (1155,18).

Como os três corpora possuem dimensões diferentes, foi necessário normalizar os resultados de forma que pudessem ser comparados entre si. Para obter-se o valor normalizado, isto é, o número de ocorrências de um VM a cada mil palavras, divide-se o número total de ocorrências de um VM pelo número total de *tokens* no corpus e multiplica-se o resultado por mil, conforme exemplo abaixo:

2973= número total de ocorrências de VMs no AC

463.585= número total de *tokens* no AC

$2973/463.585 = 0,006413 \times 1000 = 6,4$ (valor normalizado)

De posse dos valores normalizados, foram realizados os contrastes: AC x BNC, AC x OSC e AC x BNC x OSC.

Os contrastes buscam verificar se a presença dos VMs no AC, em comparação com o BNC e o OSC é suficientemente significativa a ponto de caracterizar os VMs como uma peculiaridade do AC. Para verificar se a diferença normalizada é realmente significativa e não é aleatória, foi utilizado o teste estatístico *Log Likelihood* (LL). Conforme Rayson (2002), o teste estatístico com melhores resultados para a comparação da frequência de palavras ou expressões entre dois corpora é o LL. Se o resultado obtido após a aplicação do LL foi de 6,63 ou mais, a probabilidade de a diferença entre os dois corpora ter acontecido aleatoriamente é de menos de 1%. Dessa forma, o pesquisador pode estar 99% confiante de que o resultado é significativo. Esse resultado é geralmente expresso como $p < 0.01$. Para a obtenção do resultado do LL foi utilizado o “*Log Likelihood Calculator*” no endereço <http://ucrel.lancs.ac.uk/llwizard.html>. Vide exemplos abaixo:

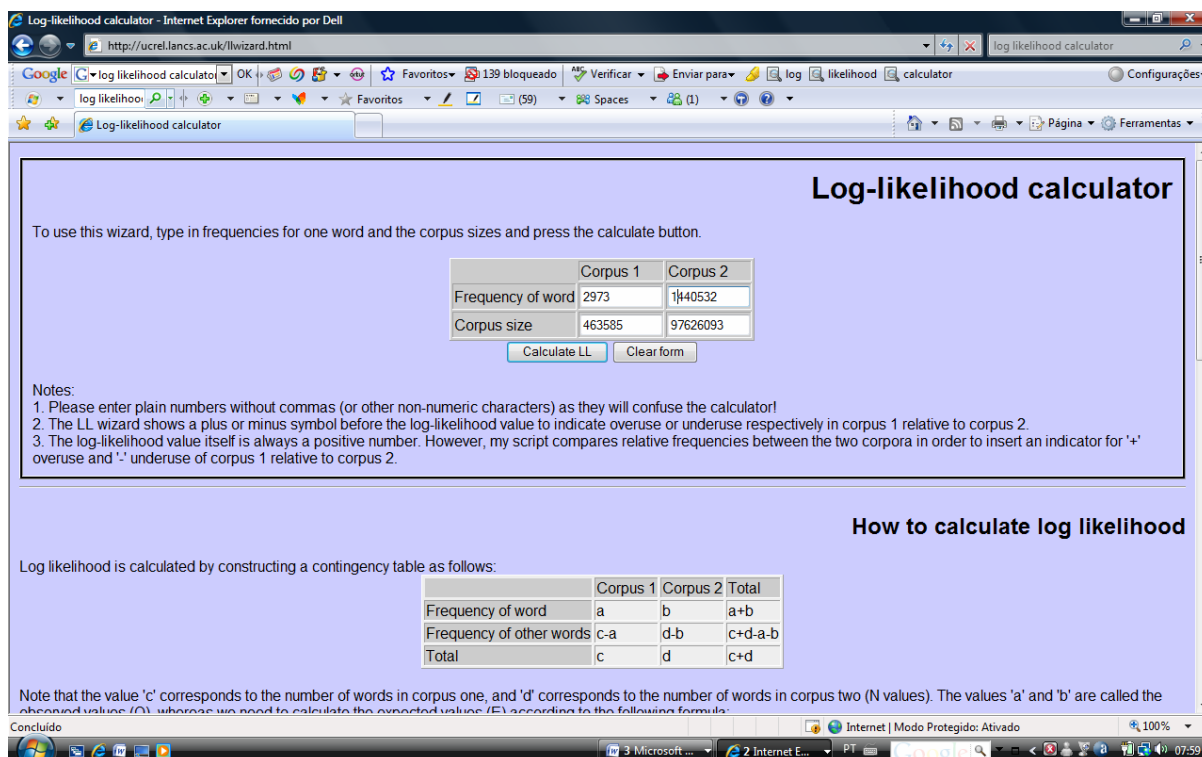


Figura 5: *Log Likelihood Calculator*- Tela inicial

No site mostrado acima, foi realizado o cálculo para verificar a diferença entre o total de VMs no AC e o total de VMs no BNC. Clicando em “*Calculate LL*”, obtém-se o seguinte resultado:

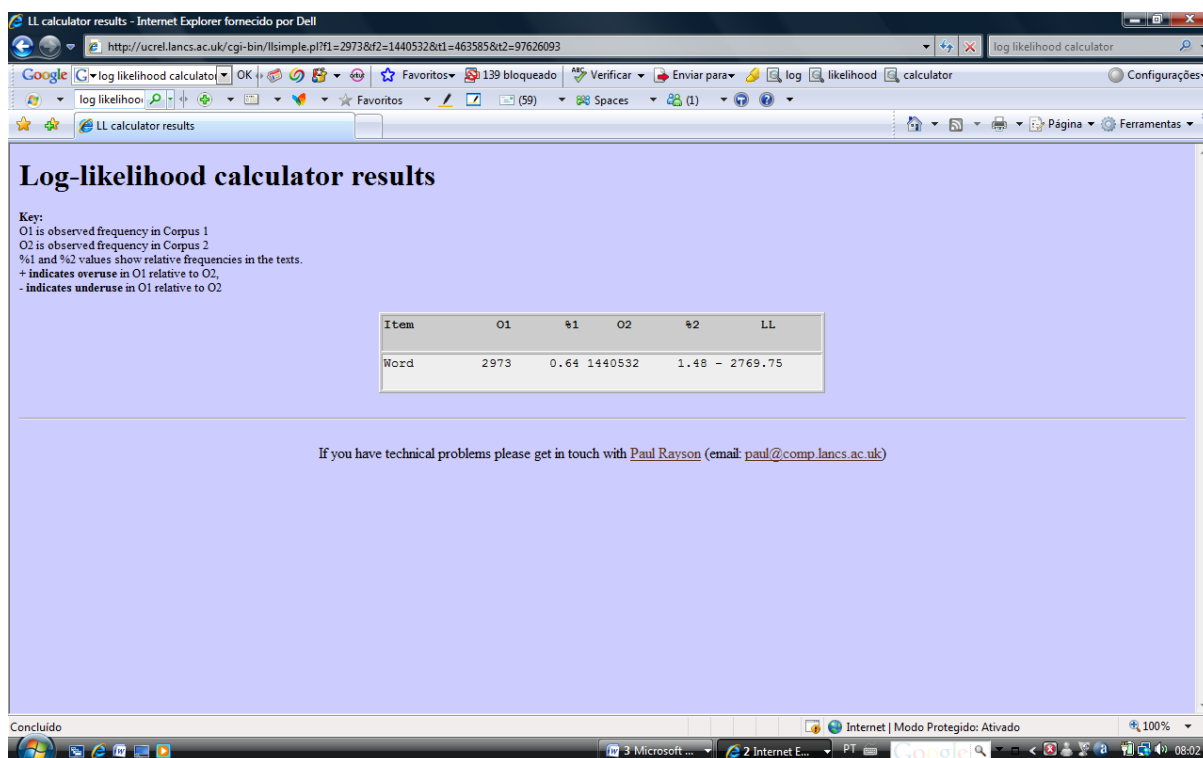


Figura 6: Tela com os resultados do *Log Likelihood*

Na coluna 01, aparece o valor total de VMs no corpus 1, no caso o AC. Na coluna %1 aparece o valor percentual de VMs no AC, neste caso, 6,4%. A coluna 2 mostra a quantidade total de VMs no corpus 2, o BNC. Em % 2, é possível ver o percentual de VMs no BNC. Finalmente, a última coluna mostra o valor LL (negativo) de -2769.75. O sinal (-) na frente do número significa que o corpus 1 faz menos uso da unidade lexical pesquisada do que o corpus 2. Como neste trabalho foi usado o ponto de corte de 6,63, nota-se que a diferença é significativa. Mostro, a seguir, a segunda etapa metodológica.

- ii) Contrasto as frequências das estruturas sintáticas em que os VMs se encontram entre o corpus de estudo (AC), o corpus de referência (BNC) e o corpus de contraste (OSC), observando se há diferenças significativas no uso das estruturas.

Para esta etapa, utilizei como ponto de partida as nove estruturas gramaticais sugeridas por Kennedy (2002), conforme o quadro abaixo:

Quadro 6.1.A: Estruturas sintáticas dos modais

	Estrutura	Exemplo
1	Modal sozinho	(Who will go?) I will.
2	Modal + infinitivo (modal + voz ativa)	Sam <i>can</i> swim./ She <i>must</i> be hungry./ You <i>should</i> have a rest.
3	Modal + <i>be</i> + passado particípio (modal+ voz passiva)	It <i>should</i> be replaced.
4	Modal+ <i>be</i> +presente particípio (modal+progressivo)	They <i>will</i> be arriving soon.
5	Modal+ <i>have</i> +passado particípio (ou adjetivo) (modal+perfeito)	He <i>might</i> have done it. He <i>must</i> have been hungry.
6	Modal+ <i>be</i> + <i>being</i> +passado particípio (ou adjetivo) (modal+ <i>be</i> +progressivo+passado particípio/adjetivo)	It <i>might</i> be being done tomorrow. He <i>could</i> be being awkward.
7	Modal+ <i>have</i> + <i>been</i> +passado particípio (Modal+voz passiva+perfeito)	It <i>should</i> have been fixed.
8	Modal+ <i>have</i> + <i>been</i> +presente particípio (modal+voz passiva+progressivo)	He <i>must</i> have been lying.
9	Modal+ <i>have</i> + <i>been</i> + <i>being</i> +passado particípio (ou adjetivo) (modal+voz passiva+perfeito+progressivo)	He <i>might</i> have been being blackmailed. They <i>must</i> have been being careless.

(KENNEDY, 2002, p. 82)

Foi utilizado o *Concord*, do *WordSmith* e feita uma análise manual de todas as linhas de concordância para cada VM no AC e no OSC. Para facilitar a busca, as linhas de concordância foram organizadas alfabeticamente à esquerda do nódulo conforme figura 7:

N	Concordance	Set	Tag	Word #	t.	# os.	%	# os.	%	# os.	%	File	%
1	can be omitted. distance only entries must start with a /, ETA - displays the	81.108	245	6%	0	6%	0	6%	0	6%	0	omcompleto.txt	86%
2	Normally, both detector loops must sense a fire or overheat condition	98.790	269	0%	0	7%	0	7%	0	7%	0	omcompleto.txt	50%
3	Normally, both detection loops must sense smoke to cause an alert.	99.926	325	0%	0	8%	0	8%	0	8%	0	omcompleto.txt	51%
4	mode is ECON PATH. The crew must select a speed descent mode. The	81.731	280	4%	0	6%	0	6%	0	6%	0	omcompleto.txt	87%
5	all cases, such revisions and changes must remain compatible with the	841	74	3%	0	0%	0	0%	0	0%	0	omcompleto.txt	1%
6	the ISDU keyboard. The airplane must remain stationary during alignment.	43.264	210	0%	0	8%	0	8%	0	8%	0	omcompleto.txt	69%
7	under these conditions, the airplane must remain within a distance equal to	51.963	686	8%	0	2%	0	2%	0	2%	0	omcompleto.txt	74%
8	under these conditions, the airplane must remain within a distance equal to	51.854	680	0%	0	2%	0	2%	0	2%	0	omcompleto.txt	74%
9	- Engine System Description The EEC must receive electrical power to supply	91.855	926	6%	0	4%	0	4%	0	4%	0	omcompleto.txt	47%
10	AOA. With this trim schedule the pilot must pull more aft column to stall the	06.556	616	0%	0	1%	0	1%	0	1%	0	omcompleto.txt	54%
11	Unit (ADIRU) ADIRU alignment must not be attempted at latitudes	7.351	339	4%	0	3%	0	3%	0	3%	0	omcompleto.txt	4%
12	VORTAC, or DME stations that must not be used for FMC position	80.799	226	3%	0	6%	0	6%	0	6%	0	omcompleto.txt	86%
13	to be zero. Random fuel imbalance must not exceed 453 kgs for taxi,	7.557	351	3%	0	4%	0	4%	0	4%	0	omcompleto.txt	4%
14	compass synchronization the crew must manually enter the initial magnetic	43.658	233	4%	0	8%	0	8%	0	8%	0	omcompleto.txt	70%
15	receivers are installed. The flight crew must manually tune the VOR on the	44.565	284	0%	0	9%	0	9%	0	9%	0	omcompleto.txt	70%
16	OPTIONS page 2/2. The flight crew must manually tune the DME on the	44.396	274	4%	0	9%	0	9%	0	9%	0	omcompleto.txt	70%
17	navigation control panel. The flight crew must manually tune the ILS for display	44.474	278	8%	0	9%	0	9%	0	9%	0	omcompleto.txt	70%
18	airplane. In this situation, the flight crew must, if necessary to avoid ground	74.422	164	3%	0	5%	0	5%	0	5%	0	omcompleto.txt	39%
19	BEEZR, and ELN. The airplane must fly direct from LAC01 to BANDR.	72.329	834	0%	0	2%	0	2%	0	2%	0	omcompleto.txt	83%
20	shows a discontinuity. The airplane must fly direct from LAC01 to BANDR.	73.053	864	0%	0	2%	0	2%	0	2%	0	omcompleto.txt	83%
21	whenever necessary. The new waypoint must first be placed into the CDU	71.303	788	2%	0	1%	0	1%	0	1%	0	omcompleto.txt	83%
22	open position. To close the window, it must first be unlocked. Pull forward on	49.474	042	0%	0	4%	0	4%	0	4%	0	omcompleto.txt	27%
23	prior to takeoff the following conditions must exist: - airplane on the ground -	00.200	236	8%	0	5%	0	5%	0	5%	0	omcompleto.txt	95%
24	plan complete with all performance data must exist before the required time of	33.554	229	0%	0	6%	0	6%	0	6%	0	omcompleto.txt	19%
25	N1 is reached, however, thrust increase must continue to a minimum of 80% N1	36.677	460	6%	0	7%	0	7%	0	7%	0	omcompleto.txt	21%
26	CO ROUTE line. CO ROUTE entries must correspond to a company defined	58.957	099	3%	0	6%	0	6%	0	6%	0	omcompleto.txt	77%
27	all types of windshear. The flight crew must continue to rely on traditional	08.472	689	8%	0	9%	0	9%	0	9%	0	omcompleto.txt	99%
28	a TO waypoint first. Valid airways must: - contain the fix entered in the TO	59.940	163	2%	0	6%	0	6%	0	6%	0	omcompleto.txt	78%
29	a TO waypoint first. Valid airways must: - contain the fix entered in the TO	60.604	199	2%	0	6%	0	6%	0	6%	0	omcompleto.txt	78%
30	if the entry is valid. The route segment must contain the waypoint entered in the	60.041	168	8%	0	6%	0	6%	0	6%	0	omcompleto.txt	78%
31	if the entry is valid. The route segment must contain the waypoint entered in the	60.705	204	8%	0	6%	0	6%	0	6%	0	omcompleto.txt	78%
32	approved AFM with which the operator must comply. In the event of conflict with	852	74	5%	0	0%	0	0%	0	0%	0	omcompleto.txt	1%
33	NAV, or ATT. If the IRS is turned off, it must complete a full realignment cycle	43.830	243	5%	0	8%	0	8%	0	8%	0	omcompleto.txt	70%

Figura 7: Colocados organizados alfabeticamente à direita do nódulo *must*.

Primeiramente, as linhas de concordância foram analisadas com vistas a classificá-las conforme as nove estruturas presentes na tabela proposta por Kennedy (2002). As linhas foram salvas em arquivos separados já classificadas nessas estruturas, pois, a partir deste ponto, as outras análises serão sempre feitas conforme as estruturas em que os VMs se encontram. Para o BNC, utilizei os resultados apresentados por Kennedy (2002), em que o autor apresenta o levantamento quantitativo dos modais e suas estruturas nos diferentes registros do BNC. Sigo para a próxima etapa:

- iii) Verifico as diferenças quanto à frequência de cada VM e de cada estrutura em cada um dos três subcorpora que compõem o corpus de estudo.

A partir desta etapa, investigo somente o corpus de estudo, ou seja, o Aviation Corpus (AC). Nesta etapa, foram utilizados as mesmas ferramentas e recursos das duas etapas anteriores para verificar: a quantidade de *tokens* em cada subcorpora (ou manual); a quantidade de VM em cada subcorpora; o valor normalizado de cada VM; a presença de cada estrutura sintática para cada VM em cada subcorpora; e os contrastes. As linhas de concordâncias são salvas separadamente para que possam ser utilizadas nas outras etapas.

- iv) Identifico as colocações de cada VM em cada um dos corpora.

A descrição do comportamento de cada VM é feita baseada nos resultados obtidos na seção anterior que apontou a estrutura sintática em que cada VM ocorre preferencialmente em cada um dos três manuais (MM, OM e QRH) que compõe o AC. Entretanto, a aplicação da metodologia de extração de colocados e de suas padronizações no QRH mostrou-se improdutivo. Isso se deve ao reduzido tamanho do subcorpus¹¹⁶ e a conseqüente baixa frequência dos itens pesquisados (VMs). Assim, analiso somente os dois corpora mais representativos, o MM e o OM. Ainda a mencionar é o fato de o QRH ser também direcionado aos pilotos, assim como o OM. Dessa forma, a linguagem dos manuais técnicos para pilotos ainda é contemplada.

Como visto, os modais podem ocorrer em nove estruturas sintáticas diferentes. Contudo, os resultados mostram que no Corpus de Aviação (AC), há duas estruturas sintáticas predominantes, conforme a tabela proposta por Kennedy (2002):

- a) Estrutura 2- modal +infinitivo (modal + voz ativa)

Exemplo:

*Airplanes **may** operate with uncured fay sealant or aerodynamic sealant on wing external surface.*

- b) Estrutura 3- modal + *be*+ passado particípio (modal+ voz passiva)

Exemplo:

*The following is a list of cleaners and polishes that **could be** required . during regular maintenance of the airplane.*

Dessa forma, os colocados¹¹⁷ para cada modal são mostrados seguindo a classificação realizada na seção 7.1. Optei por mostrar os resultados dessa forma por dois motivos:

(1) no caso da voz passiva, o verbo *be* faz parte da estrutura em si, não sendo, dessa forma, considerado um colocado do VM;

(2) para verificar se há relação entre a estrutura em que o modal se encontra e seu sentido. Isto é, se o fato de o modal estar na estrutura modal+voz ativa, ou, modal+voz passiva influencia no seu sentido.

¹¹⁶ O QRH apresenta apenas 29.929 *tokens*, comparados a 249.691 no MM e 183.959 no OM. O número absoluto de tokens modais no QRH é de 199.

¹¹⁷ A noção de colocados é explicada no Capítulo 2.

Esta etapa é dividida em duas sub-etapas que serão seguidas sempre que possível. Com base nas etapas (i), (ii) e (iii), foi possível verificar que alguns dos VMs estudados estão ausentes ou apresentam pouquíssimas ocorrências em um manual. Quando for este o caso, mostro todas as linhas de concordância do VM com poucas ocorrências e discuto o(s) seu(s) uso(s). A lista de colocados só pode ser obtida se alguns critérios forem cumpridos. Um desses critérios é que, pelo menos um dos colocados do nóculo apresente uma frequência mínima de quatro unidades em um horizonte de seis palavras, três à esquerda (L1, L2 e L3) e três à direita (R1, R2 e R3).

As duas sub-etapas são:

A: Colocados à esquerda

B: Colocados à direita

Descrevo e explico agora essas duas sub-etapas e mostro as categorias que foram usadas para classificar as ocorrências.

A: Colocados à esquerda (L1, L2, L3...)

Subdividida em duas fases:

1-Relação dos 10 colocados à esquerda com maior grau de atração.

Conforme o capítulo 2, uma lista de colocados fornece uma espécie de perfil semântico das palavras envolvidas. Essas listas de colocados foram feitas para cada modal em cada uma das estruturas encontradas na etapa (iii) realizada anteriormente. Para medir a força ou o grau de cada colocação, foi utilizado o teste estatístico *Log Likelihood*. Se o resultado obtido após a aplicação do LL foi de 6,63 ou mais, a probabilidade de a atração entre os dois itens lexicais ter acontecido aleatoriamente é de menos de 1%. Dessa forma, o pesquisador pode estar 99% certo de que o resultado é significativo. A função da estatística é apontar se os itens realmente formam colocações ou se são co-ocorrências espúrias. Os cálculos não foram realizados por mim, mas pelo próprio *WordSmith*, que apresenta também outras opções de testes estatísticos para medir a “força” de uma colocação (HUNSTON, 2002), conforme figura abaixo:

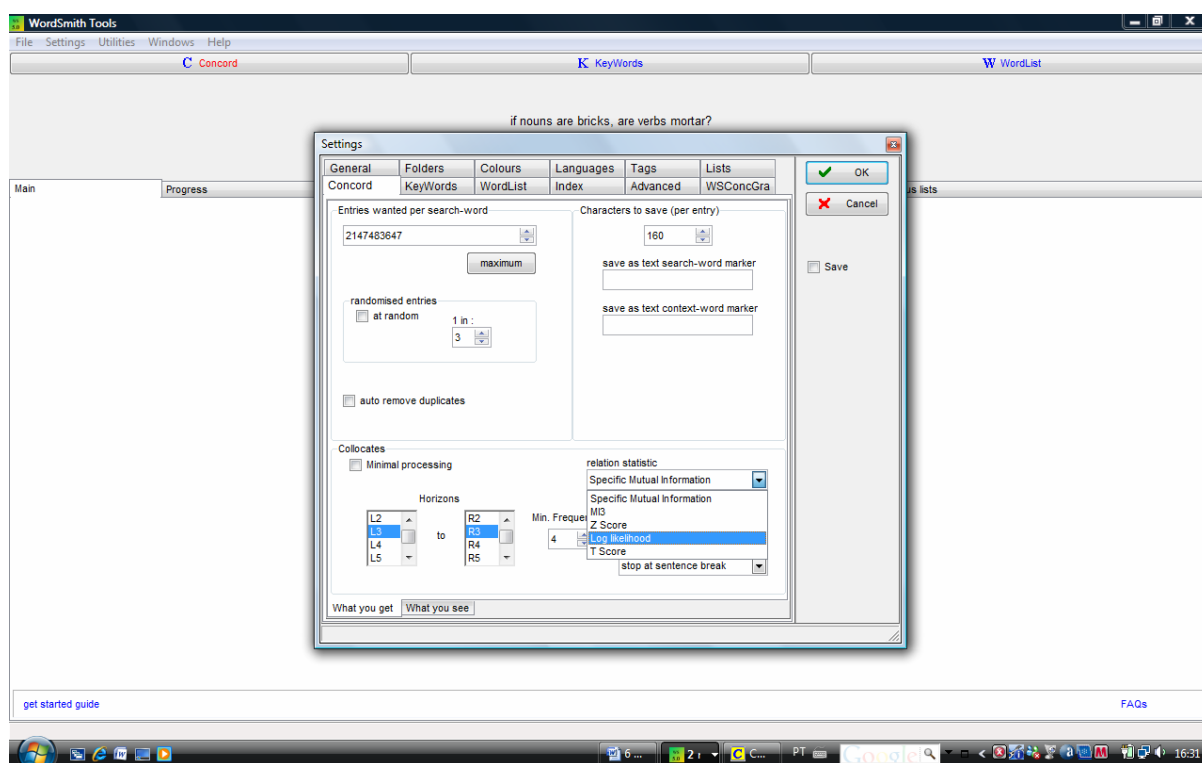


Figura 8: Tela das configurações do *Concord-WordSmith*

Mede-se o grau de associação entre duas palavras tendo em vista o número de vezes que essas duas palavras ocorrem juntas. Ou seja, quanto mais vezes dois elementos ocorrem simultaneamente e menos vezes ocorrem separadamente, mais forte será sua associação. Um elevado valor da medida de associação indica que as variáveis em causa estão associadas, isto é, a ocorrência de uma delas está associada à ocorrência da outra. Cada estatística tende a trazer resultados diferentes. De modo sucinto, a diferença entre as estatísticas é que, alguns testes (por exemplo, o *Mutual Information*, MI) capturam associações em que os itens aparecem juntos frequentemente em termos relativos, enquanto outros apontam associações em que os itens co-ocorrem frequentemente em termos absolutos. Isso significa dizer que o MI localiza pares de palavras que aparecem juntos muitas vezes em relação à frequência de um e de outro, mesmo que cada um ocorra poucas vezes. Já o LL dá preferência aos colocados que são, por si só, muito frequentes. Como o objetivo desta tese é descrever o que “é mais frequente” no AC para auxiliar na elaboração de material didático, considerei o LL mais apropriado para esse fim.

Dessa forma, são considerados colocados, os itens lexicais que se coloquem dentro de um horizonte de três palavras à esquerda (posições L1, L2 e L3) a três palavras à direita (posições R1, R2 e R3) do nóculo, com uma frequência mínima de quatro ocorrências, e uma relação LL de pelo menos 6,6. Através da lista de colocados é possível verificar se algum item

lexical se destaca em termos de frequência ou força de associação. Assim, nessa fase, já é possível verificar algumas características relativas ao comportamento do nóculo no corpus. A partir da verificação das listas de colocados à esquerda do nóculo, foi possível observar uma variação na classe de palavras que ocupam a posição L1, isto é, uma posição à esquerda do nóculo, que corresponderia teoricamente ao sujeito do VM. Conforme figura abaixo:

The screenshot shows a window titled 'mmcan-active-limpo.cnc' with a menu bar (File, Edit, View, Compute, Settings, Window, Help) and a table of concordance data. The table has columns: N, Word, Wwith, Relation, Texts, Total, tal, Left, al, Right, L3, L2, L1, Centre, R1, R2, R3. The first row is highlighted in blue and corresponds to the word 'CAUSE'. Below the table are tabs for 'concordance', 'collocates', 'plot', 'patterns', 'clusters', 'filenames', 'follow up', 'source text', and 'notes'. The status bar at the bottom shows '708 Set CAUSE'.

N	Word	Wwith	Relation	Texts	Total	tal	Left	al	Right	L3	L2	L1	Centre	R1	R2	R3
1	CAUSE	can 2.197,822		1 221	3 218	3	0	0	215	1	2					
2	YOU	can 1.292,390		1 200	180 20	1	0	179	0	0	15	5				
3	OCCUR	can 844,076		1 87	3 84	2	1	0	0	81	3	0				
4	TO	can 703,931		1 176	37 139	14	23	0	0	0	15	124				
5	DAMAGE	can 702,852		1 98	26 72	11	2	13	0	8	60	4				
6	THE	can 590,289		1 254	126 128	58	68	0	0	0	94	34				
7	INJURY	can 535,313		1 56	13 43	13	0	0	0	0	36	7				
8	USE	can 239,650		1 53	2 51	2	0	0	0	43	4	4				
9	BE	can 225,430		1 47	1 46	1	0	0	0	42	4	0				
10	WHICH	can 190,390		1 32	27 5	1	1	25	0	0	1	4				
11	THIS	can 190,312		1 45	34 11	2	6	26	0	0	10	1				
12	THAT	can 179,180		1 45	37 8	13	9	15	0	0	0	8				
13	OIL	can 168,686		1 43	33 10	3	6	24	0	0	3	7				
14	IT	can 152,536		1 35	26 9	2	1	23	0	0	6	3				
15	INJURIES	can 150,427		1 15	0 15	0	0	0	0	0	15	0				
16	COME	can 139,476		1 14	0 14	0	0	0	0	0	12	1				
17	GET	can 135,120		1 24	0 24	0	0	0	0	24	0	0				
18	BURN	can 134,380		1 13	0 13	0	0	0	0	0	13	0				
19	CORROSION	can 131,760		1 19	3 16	0	2	1	0	0	15	1				
20	ABSORB	can 131,221		1 12	0 12	0	0	0	0	0	12	0				
21	AND	can 127,113		1 62	31 31	8	9	14	0	0	11	20				
22	IF	can 122,661		1 37	7 30	0	7	0	0	0	12	18				
23	ALSO	can 113,533		1 19	2 17	0	1	1	0	17	0	0				
24	AGENTS	can 109,640		1 12	12 0	0	12	0	0	0	0	0				
25	HOT	can 109,625		1 17	16 1	1	6	9	0	0	1	0				
26	EXPLOSION	can 106,886		1 11	9 2	0	1	8	0	0	0	2				
27	OR	can 106,336		1 45	38 7	16	22	0	0	0	1	6				
28	MOVE	can 104,096		1 19	6 13	5	1	0	0	8	3	2				
29	AN	can 103,870		1 25	9 16	4	5	0	0	0	10	6				
30	FLAMMABLE	can 103,762		1 13	12 1	11	1	0	0	0	0	1				
31	THESE	can 103,058		1 23	19 4	4	14	1	0	0	1	3				
32	FREEZE	can 97,727		1 11	2 9	0	2	0	0	3	0	6				
33	MATERIALS	can 92,932		1 19	2 17	1	0	1	0	0	0	17				

Figura 9: Lista de colocados de *can-WordSmith*

Assim, foi necessário analisar todas as linhas de concordância do nóculo VM de forma a verificar que tipo de estrutura compunha essas posições à esquerda (L1, L2, L3...). A partir dessa verificação, foi possível classificar os colocados à esquerda em cinco grupos: (1) *You*; (2) anáforas; (3) unidades simples; (4) sintagmas nominais; (5) sintagmas verbais e preposicionais. Essas categorias surgiram dos dados e não foram baseadas em uma taxonomia pré-existente. Como esta tese visa a fornecer subsídios para a elaboração de material didático, considerei importante classificar os tipos de elementos lexicais que antecedem os VMs por acreditar que alguns deles impõem mais dificuldade na compreensão da função do modal, uma vez que apresentam o sujeito separado do VM por outros elementos lexicais intermediários. Esse tipo de construção deve merecer mais destaque nos materiais didáticos. O único dos grupos (nas posições L1, L2, L3...) constituído de um único item lexical é o grupo “*you*”. Esses grupos são descritos a seguir:

(1) *you*- *You* não foi classificado em um grupo juntamente com outros pronomes pessoais retos por apresentar uma função diferente nesses textos. Por tratar-se de um tipo de texto com características deônticas que estabelece o que pode ou não ser feito, e o que é ou não possível, *you* difere dos usos dos outros pronomes pessoais que nesses textos, ocupam a função de um elemento anafórico. Segue alguns exemplos ilustrativos do *you* +VM¹¹⁸:

N Concordance

1 will dry on the surface before **you can** flush it with water. EFFECTIVITY□.
 2 surface can dry the cleaners before **you can** flush them with water. The dried
 3 it has these properties: (1) **You can** fill the tanks with fuel more quickly
 4 IF THE OIL TANK IS COOL, **YOU CAN** FILL IT TOO MUCH AND CAUSE
 5 of the tanks at the same time. (3) **You can** fill only one tank if it is necessary.
 6 as far as the attached hardware. **You can** fully seal the attached hardware. If
 7 of the fuselage, section 48 that **you can** get access to. Look at the structure

(2) Anáforas - Neste grupo há sempre um elemento anafórico na posição L1, como por exemplo, *this*, *which*, *it*, *that*, *they*, *these* e *who*. Lembrando que elementos anafóricos são aqueles que retomam um elemento referencial. Vide os exemplos abaixo:

N Concordance

25 beam. N_O_T_E_: A device **which can find mercury** (mercury sniffer) is
 26 movement of items or equipment **which can hit the** airplane or injure persons.
 27 bearings, screens, screws, etc. **This can also include** capping of lines,
 28 in order to perform the task. **This can also include** uncapping of lines,
 29 TEMPERATURES CAN OCCUR. **THIS CAN CAUSE A FIRE OR EXPLOSION**

(3) Unidades simples- Foram consideradas unidades simples os casos em que há apenas um substantivo, precedido por um determinante ou não. Seguem alguns exemplos ilustrativos deste grupo:

N Concordance

2 with hot or cold **water immediately**. **Alcohol can** also be used. S 162-036 (5) Dry the
 3 BEFORE YOU START THE **APU. THE APU CAN** BE DAMAGED BY THE ICE OR SNOW
 4 Strike Zone (Zone 1, Fig. 201). **Damage can** also be found on the skin trailing edge
 5 the damaged area you **can see**. **Delamination can** be found by instrumental NDI procedures
 6 TIRE OR **WHEEL. AN EXPLOSION CAN** BE CAUSED AND INJURY TO

¹¹⁸ Os exemplos são linhas de concordância extraídas aleatoriamente entre todas as linhas de concordâncias. A numeração é aquela que constava ao lado da linha de concordância e não corresponde a qualquer classificação.

- (4) Sintagmas nominais- Foram incluídos nesse grupo os colocados à esquerda representados por sintagmas nominais com pelo menos dois elementos, como mostra os exemplos abaixo:

N Concordance

1 show more oil than there is. Expanded air can push the oil from the oil cooler into the
 2 (450jF). Heat guns and soldering guns also can contain electrical switches which can
 3 HIGH HAZARDOUS GAS CONTENT AREA CAN CAUSE AN EXPLOSION. THE
 4 the airplane. C. Many small leakage areas can add to create an excessive fuselage
 5 W_A_R_N_I_N_G: VOLCANIC ASH CAN IRRITATE YOU EYES AND BE A
 6 power in the battery. A fully charged battery can operate the refuel system for 15 to 20

- (5) Sintagmas verbais e preposicionais- Este grupo é constituído de dois tipos de colocados à esquerda: os sintagmas verbais, por exemplo: *oil system parts that are hot can*; e sintagmas preposicionais, por exemplo, *damage to the valve can occur*. A característica em comum neste grupo é o fato de o sujeito, ou referencial, do VM não estar diretamente (imediatamente) ligado a ele, conforme os exemplos abaixo:

N Concordance

1 impedance changes from 10mp to 100mp can possibly be a cable connector and are
 2 TO THE UNIT YOU ARE NOT SURE ABOUT CAN MELT AND LET SOME WATER AND
 3 of the external surfaces of the airplane can be below freezing, ice can occur and
 4 LOAD LIMITS. DAMAGE TO THE AIRPLANE CAN OCCUR. (a) Calculate the value that each
 5 AND SLATS ARE FAST ACTING AND CAN CAUSE SERIOUS INJURY TO
 6 OIL WILL PUT STAINS ON CLOTHES AND CAN MAKE PAINT SOFT. (3) Remove the filler

É importante ressaltar que, para a classificação dos grupos “*You*”, “anáforas” e “unidades simples” foram considerados somente os colocados na posição L1. Para os “sintagmas nominais” e “sintagmas verbais e preposicionais”, não foi estabelecido uma posição limite, uma vez que o sujeito pode ser um composto nominal de até quatro elementos, ou até uma oração completa. Outro fato a ser comentado é que com exceção do grupo “*you*”, os outros grupos são compostos por elementos inanimados na posição de sujeito.

B: Colocados à direita (R1, R2, R3)

Analiso e classifico os colocados à direita dos VMs em cada uma das estruturas a partir da lista de colocados e concordâncias em duas fases:

A primeira fase apenas aponta os 10 colocados com maior grau de atração ao nódulo (quando houver 10). Diferentemente dos colocados nas posições à esquerda, que podem pertencer a diferentes classes gramaticais, a posição R1 apresenta uma restrição gramatical quanto ao tipo de elemento lexical que pode seguir o VM. Dessa forma, o VM pode ser imediatamente seguido apenas por verbo ou advérbio. Nos casos de VM + voz passiva, o *be* não é considerado um colocado do VM por ser considerado elemento obrigatório para estabelecer a estrutura. Dessa forma, nos casos de voz passiva, a posição R1 é a primeira posição após o *be*.

Como os colocados à direita são principalmente verbos, é importante conhecer o tipo de verbo que segue o VM. A lista de colocados fornece apenas a frequência de palavras na posição L1, mas não as classifica por função ou sentido. Alguns VMs podem apresentar uma preferência por um tipo de verbo ou por outro, mas essa informação só será conhecida se os verbos forem semanticamente classificados. Assim, utilizo a taxonomia proposta por Biber et al. (1999) para classificar os verbos lexicais que acompanham os VMs. Essa taxonomia foi escolhida pelas seguintes razões: por ser semântica; por ter sido totalmente desenvolvida a partir de um corpus; e por ser suficientemente abrangente a ponto de incluir as várias funções encontradas no AC. Descrevo a taxonomia de forma sucinta e apresento as categorias de análise.

Embora a grande maioria dos verbos tenha mais de um significado, Biber et al (1999) classificam os verbos em sete domínios semânticos: (1) de atividade, (2) de comunicação, (3) mentais, (4) causativos, (5) de ocorrência única, (6) de existência ou relacionamento e de (7) aspecto¹¹⁹. Entretanto, os próprios autores mencionam problemas com essa taxonomia. Alguns verbos não possuem somente uma única classificação, uma vez que seu sentido básico pode ser considerado como pertencendo a uma ou outra categoria. Por exemplo, os verbos *hesitate*, *pretend*, *find* e *resist* podem ser considerados tanto como verbos de atividade, quanto mentais. Os verbos *read*, *deny*, *confirm* e *blame* podem ser classificados como atos comunicativos, atos mentais, ou estados. Ainda, alguns verbos podem ser usados com diferentes sentidos pertencendo, assim, a mais de um domínio semântico. Isso acontece principalmente com os verbos de atividade que, muitas vezes, possuem um sentido secundário. Por exemplo, os verbos *contact* e *raise* podem se referir a atividades físicas ou a atos comunicativos, enquanto os verbos *admit* e *consult* podem referir atividades mentais,

¹¹⁹ A listagem completa como sugerida por Biber et al (1999) está no anexo b.

comunicativas, ou físicas. Os verbos *follow*, *gather*, *face* e *overcome* podem ser físicos ou mentais; *change*, *rise*, e *open* podem referir tanto uma atividade física como uma ocorrência simples; *look* pode referir uma atividade mental, física ou um estado de existência (por exemplo, *you look happy*). Contudo, a maioria dos verbos possui um sentido nuclear e pertencem somente a um domínio semântico. Nos casos em que os verbos apresentaram mais de um sentido, foram categorizados pelos autores conforme o sentido mais freqüentemente atribuído pelos falantes, ou segundo seus usos mais freqüentes, sempre que o sentido atribuído pelo falante apresentou divergência quanto ao sentido típico. Explico sucintamente as categorias a seguir:

- (1) Os verbos de atividade denotam uma ação que pode ser associada com uma escolha e assim o sujeito tem o papel semântico de agente. Por exemplo: *bring*, *buy*, *carry*, *come*, *give*, *go*, *leave*, *move*, *open*, *run*, *take*, *work*. Alguns verbos de atividade podem ser usados para expressar ações não-volitivas, eventos, ou relações estáticas, por exemplo, (BIBER et al., 1999, p. 362):

*We'll **give** it to the dogs, they'll eat it.* (volitivo)

*This will **give** the electron the chance to get over the hump.* (não-volitivo)

- (2) Os verbos de comunicação são, na verdade, uma subcategoria dos verbos de atividade que envolve atividades de comunicação: *ask*, *announce*, *call*, *discuss*, *explain*, *say*, *shout*, *speak*, *state*, *suggest*, *talk*, *tell*, *write*. Como Biber (et al. 1999, p. 362) sugerem:

*You **said** you didn't have it.*

- (3) Os verbos mentais denotam uma variada gama de atividades e estados experienciados por humanos e que não necessariamente acarretam volição. O sujeito geralmente tem o papel de recipiente. Inclui os significados cognitivos (*think*, *know*), emocionais (*love*, *want*), percepção (*see*, *taste*), e receptores de comunicação (*read*, *hear*). Por exemplo (ibid, p. 362):

*I **think** it was Freddy Kruger.*

- (4) Os verbos causativos, tais como *allow*, *cause*, *enable*, *force*, *help*, *let*, *require* e *permit* indicam que uma pessoa, ou objeto inanimado causam um novo estado das coisas. Esses verbos geralmente ocorrem seguidos de um objeto direto nominalizado, por exemplo:

*Still other rules **cause** the deletion of elements from the structure.*

- (5) Os verbos de ocorrência (simples) primeiramente reportam eventos (tipicamente físicos) que ocorrem sem qualquer volição. O sujeito geralmente possui um papel semântico afetado: *become, change, happen, develop, grow, increase, e occur*. Por exemplo:

*The lights **changed**.*

- (6) Os verbos de existência ou relacionamento reportam um estado ou relacionamento que existe entre as entidades. Alguns dos verbos de existência ou relacionamento são *be, have, seem, appear, exist, live e stay*. Alguns exemplos de verbos de relacionamento são *contain, include, involve, e represent*.

*The problem **is** most accute in rural areas.*

*Well she **has** a day off.*

- (7) Os verbos de aspecto, tais como *begin, continue, finish, keep, start e stop*, caracterizam uma etapa em progressão ou algum outro evento ou atividade, tipicamente reportada em uma oração suplementar que segue o sintagma verbal. Por exemplo:

*After another day, he **began** to recover.*

Em alguns casos, verifiquei que o sentido dos verbos lexicais no AC não correspondiam ao sentido estabelecido por Biber et. al (1999). Nesses casos, os verbos foram enquadrados conforme seu sentido naquela ocorrência. Por exemplo, *make* é classificado pelos autores como um verbo de atividade, mas, quando acompanhado por algum outro elemento lexical, como no caso de *make sure* assume outro sentido. *Make sure* foi, então, considerado um verbo mental. Outros verbos freqüentes no AC não fazem parte da lista de verbos sugerida pelo autores, eles foram também classificados conforme o seu sentido no corpus, como por exemplo *damage*. Ele foi classificado como causativo, por possuir características semanticas semelhantes a *cause*.

Neste momento, são usados os resultados das etapas (iii) e (iv) para verificar o sentido de cada ocorrência. Parto das categorias e exemplos apresentados no capítulo 4 (seção 4.2.2) para analisar os sentidos, usos e peculiaridades dos VMs no AC. Entretanto, também verifico a possibilidade de o VM estar expressando uma função não antes descrita. Apresento exemplos retirados do corpus que expressam os usos mais salientes em cada um dos subcorpora.

A partir do próximo capítulo, descrevo os dados encontrados no corpus de estudo e os contrastes realizados com os outros dois corpora. A descrição de dados tem duas partes:

1- A primeira, basicamente quantitativa, privilegia o contraste do corpus de estudo com os outros dois corpora para averiguar se a presença de VMs é característica do corpus de estudo, formado pelos manuais de aviação ou não. Ainda nesta parte comparo e contrasto os três subcorpora que compõe o corpus de estudo para verificar se a incidência dos VMs é semelhante ou se apresenta diferenças significativas.

2- A segunda parte empreende a descrição das linhas de concordância para verificar as colocações e usos de cada VM em cada subcorpus.

7 DESCRIÇÃO E ANÁLISE DOS DADOS

Neste capítulo, apresento as análises das realizações lingüísticas dos VMs encontrados nos textos do corpus de estudo (AC), corpus de referência (BNC) e corpus de contraste (BNC). Na primeira etapa, contrasto o corpus de estudo com o corpus de referência. Na segunda, analiso quantitativamente o corpus de estudo e contrasto os três subcorpora que compõe o AC. Finalmente, na terceira etapa, com base nos levantamentos realizados nas duas primeiras etapas, descrevo, em detalhes, os usos de cada VM no corpus de estudo.

7.1 PRIMEIRA ETAPA: CORPUS DE ESTUDOS X CORPUS DE REFERÊNCIA X CORPUS DE CONTRASTE

Apresento agora os resultados, já numa relação contrastiva entre o corpus de estudo, o corpus de referência e o corpus de contraste. No final da seção, esboço algumas considerações com base na análise dos resultados.

7.1.1 Contrastes

Nesta seção, observarei os dados encontrados no corpus de referência (BNC) e no corpora de contraste (OSC), comparando-os com os dados encontrados no AC. Essa comparação tem por objetivo verificar se a presença dos VMs caracteriza o gênero textual “manual de aeronave”. Embora já tenham sido descritos anteriormente, retomo, aqui, o corpus de estudo, o corpus de referência e o corpus de contraste no quadro a seguir.

Quadro 7.1.1.A: Corpus de estudo, Corpus de referência e Corpus de contraste

Corpus	Especificação	Dimensão
AC (Aviation Corpus)	Três manuais da aeronave BOEING 737	463.585
BNC escrito (British National Corpus)	Corpus diacrônico de inglês britânico ¹²⁰	87.284.364
OSC (Operating System Corpus)	Dois manuais de sistema operacionais de computadores pessoais: Linux e Windows Vista	174.582

Divido esta seção em três itens que abordam o contraste entre o AC e os outros dois corpora:

- Item A: Corpus de estudo x Corpus de referência

¹²⁰ Nesta tese utilizo somente o subcorpus escrito uma vez que o corpus de estudo é composto somente de textos da linguagem escrita.

- Item B: Corpus de estudo x Corpus de contraste
- Item C: Corpus de estudo x Corpus de contraste x Corpus de referência

Nesta fase, o levantamento de dados se limitou a uma observação quantitativa. Por isso, não foram coletados os contextos em nenhum dos corpora. No final da seção, esboço algumas considerações parciais relacionadas aos três contrastes.

Item A:

O primeiro contraste focaliza o corpus de estudo (AC) e o corpus de referência (BNC). Primeiramente, contrastarei o grupo geral de VMs no AC com o BNC escrito, tendo em vista que o corpus de estudo é composto somente por textos da linguagem escrita. Após, contrastarei a frequência de ocorrência de cada VM no BNC e no AC. Em um terceiro momento, será observada a distribuição dos VMs entre as nove estruturas sintáticas (KENNEDY, 2002) nos dois corpora.

Os contrastes buscam verificar se a presença dos VMs no AC, em comparação com o BNC, é suficientemente significativa a ponto de caracterizar os VMs como uma peculiaridade do AC. Ou seja, como o BNC é composto de textos provenientes da língua geral, sem a presença de manuais de aviação, busco verificar em que medida o uso dos VMs é uma característica da linguagem técnica em relação à língua geral.

Para verificar os VMs no AC, foi primeiramente realizada uma busca geral de todos os VMs juntos, com o auxílio da ferramenta *concordance* do *WordSmith*. Para verificar os VMs no BNC, foi utilizado o concordanciador presente no próprio corpus¹²¹. O cálculo do valor *Log Likelihood* foi feito com a ajuda da ferramenta presente na página *Log Likelihood calculator*¹²².

Na primeira coluna, apresento o total bruto de ocorrências de VMs no AC, na segunda coluna o total bruto de VMs no BNC escrito, na terceira o valor normalizado¹²³ do AC e na quarta o valor normalizado do BNC. Finalmente, na quinta coluna, apresento o valor obtido no teste estatístico *Log likelihood*, lembrando que um valor de 6,6 ou maior é considerado

¹²¹ O BNC foi acessado através do site www.escorp.unizh.ch.

¹²² O site para calcular o *log likelihood* é <http://ucrel.lancs.ac.uk/llwizard.html>.

¹²³ O valor normalizado pode ser obtido multiplicando-se o número total de ocorrências do item pesquisado por mil e dividindo-se esse valor pelo número total de palavras no corpus, no caso, $VM \times 1000 = 2973000$, $2973000/463.585 = 6,4$. Através dos valores normalizados é possível ter uma idéia mais precisa da incidência proporcional de algum item lexical em um determinado corpus, uma vez que os corpora apresentam dimensões bastante diferentes.

como uma diferença estatisticamente significativa, ou seja, 99% de chance de a diferença não ser aleatória, mas acontecer por alguma razão específica.

Tabela 7.1.1.B Comparação entre o AC e o BNC escrito (freq. por 1.000 palavras)

AC	BNC escrito	AC (normalizado)	BNC escrito (normalizado)	LL
2973	1.227.289	6.4	14,06	- 2406.29

Através da observação dos valores apresentados nas terceira e quarta coluna, é possível constatar que o BNC escrito apresenta uma quantidade de VM bem acima do que o AC. O LL mostra que essa diferença é bastante considerável, apesar da importância do papel deôntico dos VM nos manuais. Não é possível afirmar com certeza, mas pode-se aventar a hipótese que a baixa incidência de VMs seja devido à ausência da modalização epistêmica nos manuais de aviação. Essa hipótese é verificada na seção 7.3, em que trato dos usos dos VM.

Passo agora para o contraste individual dos nove VM entre o AC e o BNC escrito. A partir deste momento, utilizarei somente os dados do corpus de referência escrito (BNC).

Tabela 7.1.1.C Comparação entre os VMs individuais: AC x BNC

VM	No. absoluto de tokens no AC	AC (normalizado)	No. absoluto de tokens no BNC escrito	BNC (normalizado)	LL
<i>Can</i>	1168	2,6	216.142	2,5	+ 0,34
<i>Could</i>	53	0,1	147.821	1,7	- 1174,87
<i>May</i>	485	1,0	108.071	1,2	- 14,49
<i>Might</i>	03	0,0	53.414	0,6	- 532,62
<i>Must</i>	507	1,1	65.836	0,8	+ 61,69
<i>Shall</i>	20	0,0	17.390	0,2	- 83,22
<i>Should</i>	165	0,4	99.376	1,1	- 340,58
<i>Will</i>	558	1,2	275.768	3,2	- 733,40
<i>Would</i>	14	0,1	243.471	2,8	- 2424,82

Os dados encontrados na tabela 7.1.1.C mostram que, com exceção de *can* e *must*, todos os outros VMs são muito mais frequentes no BNC do que no AC, como já era possível de ser deduzido, tendo em vista o contraste entre os VMs como grupo conforme a tabela 7.1.1.B. *Can* apresenta uma pequena diferença de uso, mas não estatisticamente significativa. Já *must* mostra uma diferença de + 61,69, acima do ponto de corte utilizado para considerar uma diferença como significativa (6,6). Individualmente, todos os outros VM são menos frequentes no AC em comparação ao BNC, sendo *would* (-2424,82) o verbo que apresenta a

maior diferença, seguido, respectivamente por *could* (-1174,87), *will* (- 733.40), *might* (- 532,62), *should* (-563.94). e *shall* (- 76.95) e *may* (- 14,49).

Na próxima tabela (7.1.1.D), sugerida por Kennedy (2002¹²⁴), apresento a distribuição dos VMs entre as estruturas sintáticas em que eles se encontram.

Tabela 7.1.1.D Padrões sintáticos possíveis para os verbos modais em inglês em %

		AC	BNC escrito
1	Modal sozinho	0,0	1,3
2	Modal + infinitivo (modal + voz ativa)	67,3	74,5
3	Modal + <i>be</i> + passado particípio (modal+ voz passiva)	32,4	16,6
4	Modal+ <i>be</i> +presente particípio (modal+progressivo)	0,2	1,5
5	Modal+ <i>have</i> +passado particípio (ou adjetivo) (modal+perfeito)	0,1	5,4
6	Modal+ <i>be</i> + <i>being</i> +passado particípio (ou adjetivo)	0,0	0,0
7	Modal+ <i>have</i> + <i>been</i> +passado particípio	0,0	0,7
8	Modal+ <i>have</i> + <i>been</i> +presente particípio	0,0	0,1
9	Modal+ <i>have</i> + <i>been</i> + <i>being</i> +passado particípio (ou adjetivo)	0,0	0,0

Pode-se dizer que, tanto no BNC quanto no AC, a estrutura dois (Modal + infinitivo), seguida pela estrutura três (3), mantém a preferência. Somando-as, obtém-se 99,7% das ocorrências no AC, e 91,1% no BNC escrito. A maior diferença entre os dois corpora está no número de ocorrências na estrutura três (3) (voz passiva); o AC apresenta aproximadamente o dobro de ocorrências do que o BNC. As outras estruturas estão praticamente ausentes do AC e muito pouco freqüentes no BNC. Conhecer as estruturas onde os VMs ocorrem em diferentes corpora é importante por várias razões. Primeiramente, para conhecer as especificidades sintáticas de cada corpus, e, desta maneira, poder caracterizá-los. Em segundo lugar, é um passo em direção ao reconhecimento do comportamento de cada VM, para auxiliar na descrição dos usos dessas unidades. Esse resultado é de especial importância neste trabalho, através dessa informação, pode-se sugerir exercícios didáticos que contemplem os contextos naturais em que os VMs realmente ocorrem. Por exemplo, conforme os resultados observados na tabela 7.1.1.D, a estrutura três (3) deve receber uma atenção mais especial ao desenvolver materiais didáticos focalizados ao inglês específico da aviação, uma vez que essa estrutura é duas vezes mais freqüente nesse corpus (AC) do que no corpus de referência (BNC).

¹²⁴ A tabela proposta por Kennedy (2002) já foi apresentada no capítulo sobre Modalização.

O objetivo das comparações entre o corpus de estudo e o corpus de referência não foi desenvolver um estudo exaustivo sobre o uso dos VMs no AC e no BNC, mas perceber algumas peculiaridades e diferenças no emprego dos VMs. Tendo em vista o grupo total dos nove VMs, nota-se que eles são bem mais freqüentes no corpus de referência do que no corpus de estudo. Da mesma forma, a maior parte dos VMs é muito mais freqüente no BNC do que no AC, sendo *would* o que apresenta a maior diferença, pois é o segundo VM mais freqüente no BNC e é o segundo VM menos freqüente no AC. A diferença entre os dois corpora obteve o resultado de - 2424.82 (LL), lembrando que o ponto de corte para considerar uma diferença como não aleatória é de 6,6. A ausência do *would* não apresenta uma surpresa, pois é freqüentemente usado para expressar uma condição irreal ou hipotética, que não se espera em um manual técnico. O VM *must* é o único que apresenta um uso estatisticamente mais alto no AC, com uma diferença de + 61,69, provavelmente devido ao seu caráter deontico. As estruturas sintáticas são também diferentemente distribuídas entre os dois corpora, sendo que a maior diferença concentra-se na estrutura três (3) (voz passiva). O AC apresenta proporcionalmente duas vezes mais modais nessa estrutura do que o BNC escrito.

A seguir, descrevo os dados encontrados na comparação entre o corpus de estudo e o corpus de contraste.

Item B:

Neste item, os dados entre o corpus de estudo (AC) e o corpus de contraste (OSC) são comparados. O corpus de contraste é composto de dois manuais para instalação e uso de sistemas operacionais para computadores pessoais. O público alvo desse tipo de texto não é necessariamente o especialista da área de informática, mas de usuários de computador que optam por um dos dois sistemas. Contudo, o especialista¹²⁵ até pode estar entre os leitores do texto dos manuais, mas o texto não é direcionado para esses especialistas. Esses manuais são também compostos por textos que descrevem os sistemas operacionais e fornecem instruções de como instalá-los e usá-los. Ou seja, são textos com algumas finalidades semelhantes aos encontrados no corpus de estudo. Portanto, precisa utilizar linguagem acessível e clara. O uso dos VMs nesses textos pode ser um recurso utilizado para mostrar ao leitor o que é esperado que ele faça.

¹²⁵ Os especialistas aqui referidos são profissionais que “prestam serviços” de instalação e/ou manutenção de softwares a usuários diversos. A maior parte desses profissionais não possui formação específica, e aprende o ofício através da prática ou cursos técnicos de curta duração.

Após o levantamento dos dados, feito com a ajuda da ferramenta *Concordance* do *WordSmith Tools*, foi possível comparar os resultados do OSC e do corpus de estudo (AC), conforme tabela 7.1.1.E a seguir:

Tabela 7.1.1.E Comparação entre os VMs no AC e no OSC

AC	OSC	AC (normalizado)	OSC (normalizado)	LL
2973	2481	6,4	14,2	- 729.23

Percebe-se uma grande diferença na quantidade bruta de VMs, com a frequência de apenas 6,4 VMs por mil palavras no AC comparadas a 14,2 no OSC. Dessa forma, o OSC emprega os VMs mais do que o dobro das vezes que o AC. Entretanto, para obter-se uma noção mais precisa sobre a incidência de uso dos VMs, faz-se necessário contrastar o uso de cada modal em cada um dos corpora.

Tabela 7.1.1.F Comparação entre os VMs individuais: AC x OSC

VM	No. absoluto de <i>tokens</i> no AC	AC (normalizado)	No. absoluto de <i>tokens</i> no OSC	OSC (normalizado)	LL
<i>Can</i>	1168	2,6	1415	8,1	- 795.21
<i>Could</i>	53	0,1	29	0,2	- 1.95
<i>May</i>	485	1,1	161	0,9	+ 3.63
<i>Might</i>	03	0,0	125	0,7	- 289.50
<i>Must</i>	507	1,1	43	0,2	+ 143.68
<i>Shall</i>	20	0,0	3	0,0	+ 3.05
<i>Should</i>	165	0,4	107	0,6	- 71.47
<i>Will</i>	558	1,3	527	3,0	- 199.43
<i>Would</i>	14	0,1	71	0,4	- 112.70

Apesar de o OSC também ser composto por manuais, percebe-se que há uma grande variação no uso dos VMs. A maior diferença concentra-se no *can*, que, apesar de também ser o VM mais freqüente no AC, é aproximadamente quatro vezes mais freqüente no OSC conforme os valores normalizados, com 2,6 e 8,1 *tokens* a cada mil, respectivamente. Na verdade, cabe destacar a grande quantidade de *can* no OSC, que é responsável por aproximadamente 57% do total geral dos VM nesse corpus. *Will*, que é o segundo VM mais freqüente em ambos os corpora, apresenta mais que o dobro de ocorrências proporcionais no OSC em relação ao AC. *Should*, *would* e *might* são também proporcionalmente mais freqüente no OSC, apesar de as diferenças entre *would* e *might* serem de alguma forma esperadas, pois esses VMs estão praticamente ausentes no AC. O resultado obtido com o VM

must pode ser considerado surpreendente, pois, como o OSC é um manual que fornece instruções sobre a instalação e procedimentos de um sistema, seria esperado uma maior incidência de *must*, bastante usado para expressar instruções e ordens. Entretanto, a sua incidência no OSC é mais baixa do que a incidência de *might* no mesmo corpus, sendo que *might* possui apenas três ocorrências no AC. Os modais *could*, *may* e *shall* não apresentaram diferenças estatisticamente significativas.

Contrasto, a seguir, a distribuição dos VMs entre as possíveis estruturas nos dois corpora. A intenção é verificar o uso dos VMs nas diferentes estruturas e se o contraste obtido entre o AC e o BNC está também presente em outras linguagens técnicas, mais precisamente manuais, ou se é peculiar do gênero textual manual de aviação.

Tabela 7.1.1.G Contraste entre padrões sintáticos: AC x OSC em %

		AC	OSC
1	Modal sozinho	0,0	0,0
2	Modal + infinitivo (modal + voz ativa)	67,3	76,05
3	Modal + <i>be</i> + passado particípio (modal+ voz passiva)	32,4	23,13
4	Modal+ <i>be</i> +presente particípio (modal+progressivo)	0,2	0,3
5	Modal+ <i>have</i> +passado particípio (ou adjetivo) (modal+perfeito)	0,1	0,5

Tanto no AC, quanto no OSC, a estrutura dois (2) (Modal + infinitivo) e a estrutura três (3) (Modal+*be*+passado particípio) mantêm a preferência. Somando-as, obtém-se 99,7% no AC, e 99,18% no OSC. O OSC apresenta algumas ocorrências a mais nas estruturas 4 (Modal+*be*+presente particípio) e 5 (Modal+*have*+passado particípio). Todavia, a frequência de ocorrência é também muito baixa.

Item C:

Passo agora ao contraste entre o corpus de estudo, o corpus de contraste e o corpus de referência. Nesse triplo contraste, visto a verificar se a incidência dos VM pode ser peculiar do gênero textual manual de aviação e se existe uma semelhança no uso dos VMs entre os dois corpora especializados quando comparados com o corpus de referência. Por permitirem uma visão mais real dos usos dos VMs nos corpora, serão utilizados somente valores normalizados (por mil palavras) ou porcentagens.

A tabela 7.1.1.H apresenta a frequência total dos VMs em cada um dos três corpora analisados:

Tabela 7.1.1.H Comparação entre o AC o OSC e o BNC escrito

AC (normalizado)	OSC (normalizado)	BNC escrito (normalizado)
6,4	14,2	14,6

Essa tabela síntese mostra que há muito menos VMs no corpus de estudo quando comparados aos outros dois corpora. Esse resultado aponta uma peculiaridade no corpus de estudo com relação à frequência de uso: de que é uma linguagem menos modalizada do que as outras duas analisadas. Ainda sem ter analisado as razões para a baixa incidência de VMs, pode-se especular que seja a ausência da modalização epistêmica. Para um diagnóstico mais claro a respeito da frequência de uso dos VMs, faz-se necessário observar e contrastar as ocorrências de cada um deles em cada um dos corpora. Para melhor visualização, cada resultado mais freqüente foi destacado na tabela 7.1.1.I a seguir:

Tabela 7.1.1.I Comparação entre os VMs individuais: AC x OSC x BNC

VM	AC (normalizado)	OSC (normalizado)	BNC escrito (normalizado)
<i>Can</i>	2,6	8,1	2,5
<i>Could</i>	0,1	0,2	1,7
<i>May</i>	1,1	0,9	1,2
<i>Might</i>	0,0	0,7	0,6
<i>Must</i>	1,1	0,2	0,8
<i>Shall</i>	0,0	0,0	0,2
<i>Should</i>	0,4	0,6	1,1
<i>Will</i>	1,3	3,0	3,2
<i>Would</i>	0,1	0,4	2,8

Analisando a incidência de cada modal separadamente, observa-se que seis entre os nove modais são mais freqüentes no BNC, sendo eles *could*, *may*, *shall*, *should*, *will* e *would*. Dois modais são mais freqüentes no OSC, *can* e *might* e apenas *must* é mais freqüente no AC. A diferença mais acentuada é relativa ao VM *would*. Ele é o segundo mais freqüente no BNC, mas apresenta uma incidência muito baixa nos outros dois corpora mais específicos.

Outro ponto a ser observado é relativo à distribuição dos VMs em cada um dos corpora. Principalmente o OSC, mas também o AC, mostram uma clara preferência por alguns VMs em detrimento dos outros, ao passo que o BNC apresenta uma distribuição mais uniforme entre todos os nove VMs. No OSC, *can* sozinho responde por 57% das ocorrências.

Somando-se às ocorrências de *will*, obtém-se o valor de 78% das ocorrências modais concentradas em apenas dois VMs, enquanto os outros sete respondem por menos de 22% das ocorrências. Assim, enquanto *can* apresenta uma frequência de 57%, *shall*, o menos freqüente, mostra uma incidência de pouco mais de 0,1%. Essa concentração principalmente no modal *can*, explica a grande quantidade de VMS por mil palavras no OSC, que é até maior do que no BNC. No AC, *can*, *may*, *must* e *will* correspondem a mais de 91% de todas as ocorrências, enquanto os outros cinco respondem por apenas 9% do total. *Can*, o VM mais freqüente responde por 39% das ocorrências, enquanto *might*, o menos freqüente, é responsável por 0,1%. No BNC, *will*, que é o modal mais freqüente, responde por 22% do total de ocorrências modais, enquanto *shall*, o menos freqüente, responde por 1,5%. Ou seja, a diferença de frequência entre o modal menos freqüente e o mais freqüente é bem menor no BNC do que nos outros dois corpora.

Apresento agora a tabela com as informações relativas à distribuição dos VMs entre as estruturas sintáticas com o objetivo de verificar se essa distribuição pode também caracterizar cada um dos corpora.

Tabela 7.1.1.J Contraste entre padrões sintáticos: AC x OSC x BNC em %

		AC	OSC	BNC
1	Modal sozinho	0,0	0,0	1,3
2	Modal + infinitivo (modal + voz ativa)	67,3	76,05	74,5
3	Modal + <i>be</i> + passado particípio (modal+ voz passiva)	32,4	23,13	16,6
4	Modal+ <i>be</i> +presente particípio (modal+progressivo)	0,2	0,3	1,5
5	Modal+ <i>have</i> +passado particípio (ou adjetivo) (modal+perfeito)	0,1	0,5	5,4
6	Modal+ <i>be+being</i> +passado particípio (ou adjetivo) (modal+ <i>be</i> +progressivo+passado particípio/adjetivo)	0,0	0,0	0,0
7	Modal+ <i>have+been</i> +passado particípio (Modal+voz passiva+perfeito)	0,0	0,0	0,7
8	Modal+ <i>have+been</i> +presente particípio (modal+voz passiva+progressivo)	0,0	0,0	0,1
9	Modal+ <i>have+been+being</i> +passado particípio (ou adjetivo) (modal+voz passiva+perfeito+progressivo)	0,0	0,0	0,0

Todos os três corpora apresentam uma clara preferência pelas estruturas 2 (voz ativa) e 3 (voz passiva). No AC, essas duas estruturas respondem por 99,7% de todas as ocorrências. No OSC elas correspondem a 99,18% das incidências e no BNC por 91,1%. Assim como em

relação à distribuição dos modais discutida acima, nota-se que o BNC apresenta uma distribuição um pouco mais equilibrada entre as estruturas quando comparado aos outros dois corpora específicos. A estrutura três (3) (voz passiva) parece realmente ser mais freqüente nas linguagens especializadas, apresentando aproximadamente o dobro de ocorrências no AC do que no BNC, e sendo também mais freqüente no OSC com relação ao BNC.

7.1.2 Considerações Parciais

O objetivo dessa etapa, como já mencionado, foi contrastar os dados encontrados no corpus de estudo com o corpus de referência e o corpus de contraste a fim de observar se a presença dos VMs pode ser considerada peculiar em um gênero textual específico, ou seja, manuais de aviação, ou se é comum ao macrogênero manuais.

O quadro abaixo sintetiza as preferências modais de cada manual. Os resultados são apresentados em ordem decrescente conforme a freqüência de ocorrência:

Quadro 7.1.2.A Corpus de estudo x Corpus de contraste x Corpus de referência

AC	OSC	BNC
<i>Can</i>	<i>Can</i>	<i>Will</i>
<i>Will</i>	<i>Will</i>	<i>Would</i>
<i>Must</i>	<i>May</i>	<i>Can</i>
<i>May</i>	<i>Might</i>	<i>Could</i>
<i>Should</i>	<i>Should</i>	<i>May</i>
<i>Could</i>	<i>Would</i>	<i>Should</i>
<i>Shall</i>	<i>Must</i>	<i>Must</i>
<i>Would</i>	<i>Could</i>	<i>Might</i>
<i>Might</i>	<i>Shall</i>	<i>Shall</i>

Conforme o quadro 7.1.2.A, é possível observar uma grande diferença com relação à incidência de cada VM em cada um dos manuais. Essas diferenças são provavelmente causadas pelas diferentes necessidades comunicativas dos textos de cada um dos corpora. O BNC escrito, por ser um corpus de inglês geral, representa uma grande variedade de textos de diferentes domínios e com diferentes propósitos. Essa variedade comunicativa é traduzida em uma distribuição mais igualitária quando comparada às distribuições dos VMs entre as estruturas nos outros dois corpora.

Os dois corpora mais específicos (AC e OSC) apresentaram algumas semelhanças entre si, que divergem do corpus de inglês geral: (a) uma concentração de uso (tanto em alguns modais quanto nas estruturas) condicionada pela restrição de propósitos comunicativos. Assim, pode-se dizer que esses dois corpora (AC e OSC) apresentam, a rigor,

apenas duas estruturas sintáticas, a modal+voz ativa (estrutura 2) e a modal+voz passiva (estrutura 3); (b) a baixa incidência de *could* e *would*, que são dois verbos bastante freqüentes no BNC, principalmente o *would*, que é o segundo mais freqüente no BNC; (c) o *shall* também é pouco freqüente no BNC, mas ainda menos freqüente no AC e no OSC; (d) o *should* é também bem menos usado nos dois corpora específicos.

Contudo, por possuírem públicos alvo distintos, isto é, o AC é escrito por especialistas (redatores técnicos) para especialistas (mecânicos de aeronave ou pilotos) e o OSC de especialistas (redatores técnicos) para não-especialistas (usuários em geral), os dois corpora apresentam algumas diferenças. De uma forma geral, o OSC pode ser classificado em uma posição intermediária entre o AC e o BNC com relação ao uso dos VMs. Isso pode ser observado na maior incidência de VMs como grupo, quando comparado ao AC, assemelhando-se bastante ao BNC, e na distribuição dos VMs entre as estruturas sintáticas, pois a estrutura dois (2) (voz ativa) apresenta uma freqüência de ocorrência bastante semelhante. A estrutura três (3) (voz passiva) apresenta o dobro de ocorrências no AC com relação ao BNC e 40% a mais no OSC com relação ao BNC, ou seja, apesar de a diferença entre o OSC e o BNC ser grande, a diferença entre o AC e o BNC é ainda mais substancial.

Uma peculiaridade em relação à quantidade dos VMs no AC é o *must*. Esse verbo é bem mais freqüente no AC do que nos outros dois corpora, incluindo o OSC, que também representa o macro-gênero manual, apontando uma provável preponderância da modalização deôntica forte no AC. Portanto, a significativa presença do modal *must* não é uma característica de manuais como um todo, mas, segundo o levantamento quantitativo realizado neste trabalho, é sem dúvida uma saliência do AC. Esse papel deôntico pode ter sido ocupado pelo *should* no OSC, pois é proporcionalmente mais freqüente do que *must* no AC. A escolha pelo *should* pode ter sido motivada pelo caráter menos impositivo expresso pelo modal, pois, por ser endereçado a um público menos especializado, portanto mais diversificado, pode preferir uma linguagem mais polida e mitigada. Outra opção pela força na expressão da modalização deôntica, representada pelo uso do *must*, pode ser motivada pela diferença de cenários onde os textos dos dois corpora específicos são lidos. Ao não entender a força de uma instrução em manuais de sistemas operacionais, o pior dano causado será o não-funcionamento do sistema ou, por vezes, do computador. Contudo, com relação aos manuais de aviação, a má interpretação da força de uma instrução pode causar problemas muito maiores e até perdas de vida, o que explicaria o uso por uma modalização mais forte e impositiva.

Em síntese, os dados parecem mostrar que, quanto mais específico for o texto, mais específico e peculiar será a incidência e o uso dos VMs, pois mais específicos serão seus contextos de uso e o seu público alvo. Acredito que essa seção tenha também mostrado a importância de uma comparação tripla para que se possa chegar a conclusões mais sensatas a respeito dos contrastes entre os diferentes corpora.

Na seção seguinte, passo à descrição detalhada do corpus de estudo.

7.2 SEGUNDA ETAPA: DESCRIÇÃO DO CORPUS DE ESTUDO

Nesta seção apresento os contrastes entre os três subcorpora que compõe o corpus de estudo. No final da seção, esboço algumas considerações preliminares com base na análise dos resultados.

7.2.1 Contrastes

Nesta segunda etapa da descrição dos dados, focalizo a presença dos VMs no corpus de estudo. Apesar de já terem sido descritos, retomo, aqui, o corpus de estudo, agora subdividido entre os três subcorpora, conforme quadro a seguir.

Quadro 7.2.1.A Manuais do AC

Manual	Público Alvo	Dimensão
MM (Maintenance Manual)	Mecânicos de aeronave	249.691
OM (Operation Manual)	Pilotos	183.959
QRH (Quick Reference Handbook)	Pilotos	29.929

Conforme o quadro acima, é possível notar que o maior subcorpus é o MM (249.691), seguido pelo OM (183.959). O QRH é bem menor do que os outros dois manuais. Dessa forma, o MM, por ser o maior dos três subcorpora, teve a maior representatividade nos contrastes realizados na seção anterior e o QRH, a menor.

O AC consiste em três manuais direcionados a diferentes públicos (mecânicos e pilotos) e a diferentes procedimentos e questões (manutenção e operações), sendo usados em diferentes ambientes (hangar, cabine de comando e centros de treinamento) e sob diferentes condições relativas ao tempo para realizar as instruções. É esperado que essas diferenças contextuais entre os manuais estudados sejam refletidas através da linguagem e, em particular, no uso dos modais.

Esta seção será dividida em dois itens que abordam os contrastes entre os três subcorpora:

- Item A: Contraste dos VMs nos três subcorpora
- Item B: Distribuição dos VMs nas estruturas sintáticas nos três subcorpora

Item A:

O primeiro contraste focaliza os dois públicos específicos para os quais os manuais se destinam. Assim, contraste, neste primeiro momento o MM com o OM+QRH, lembrando que o público do MM é constituído por mecânicos de aeronaves, e o público do OM+QRH é constituído por pilotos. A tabela 7.2.1 mostra a frequência normalizada dos modais de acordo com o público-alvo dos manuais.

Para identificar os VMs que mais se destacam em cada corpus, utilizei mais uma vez a ferramenta *concordance* do software *WordSmith*. Para verificar se a diferença entre os manuais é estatisticamente significativa foi utilizado o teste *Log Likelihood* com a ajuda da ferramenta presente na página *Log Likelihood calculator*.

O contraste entre os manuais para mecânicos e pilotos mostra uma diferença estatística significativa entre a frequência dos modais como grupo, com o MM apresentando 6,8 VMs a cada mil palavras e o OM+QRH apresentando 5,8 ocorrências a cada mil palavras. Contudo, há diferenças ainda mais marcantes no uso individual de cada modal. O MM contém *can* em quantidades bem maiores do que os manuais para pilotos (OM + QRH), enquanto os manuais para pilotos apresentam uma incidência bem mais alta de *may*. *Could*, *must*, *shall* e *will* também apresentam mais ocorrências no MM com diferenças estatisticamente significativas. Por outro lado, *should* é significativamente mais frequente no OM + QRH.

Tabela 7.2.1.B Comparação entre o manual para mecânicos x manuais para pilotos

VM	No. absoluto de tokens no MM	MM (normalizado)	No. absoluto de tokens no OM + QRH	OM + QRH (normalizado)	LL
<i>Can</i>	809	3,2	359	1,7	+ 115.35
<i>Could</i>	39	0,2	14	0,1	+ 8.73
<i>May</i>	110	0,4	375	1,8	- 196.88
<i>Might</i>	03	0,0	00	0,0	+ 3.71
<i>Must</i>	324	1,3	183	0,9	+ 20.97
<i>Shall</i>	18	0,1	02	0,0	+ 12.37
<i>Should</i>	66	0,3	99	0,5	- 12.73
<i>Will</i>	355	1,4	203	0,9	+ 21.77
<i>Would</i>	03	0,0	11	0,1	- 6.18
Total	1727	6,8	1246	5,8	+ 21.56

As diferenças observadas acima apontam à necessidade de comparar os dois manuais para pilotos (OM e QRH).

Tabela 7.2.1.C Comparação entre o OM x QRH

VM	No. absoluto de tokens no OM	OM (normalizado)	No. absoluto de tokens no QRH	QRH (normalizado)	LL
<i>Can</i>	327	1,8	32	1,1	+ 8,66
<i>Could</i>	12	0,0	2	0,0	0,00
<i>May</i>	288	1,6	87	2,9	- 22,76
<i>Might</i>	00	0,0	00	0,0	0,00
<i>Must</i>	172	0,9	11	0,4	+ 11.94
<i>Shall</i>	02	0,0	00	0,0	0,00
<i>Should</i>	77	0,4	22	0,7	- 4.86
<i>Will</i>	163	0,9	40	1,3	- 4.98
<i>Would</i>	06	0,0	05	0,2	- 6.32
Total	1047	5,7	199	6,6	- 3.90

Esses dois manuais compartilham o mesmo público, mas são utilizados em circunstâncias diferentes. O OM (Operation Manual) contém as descrições dos sistemas operacionais da aeronave 737 e informações sobre os procedimentos normais durante um voo. Este manual é primeiramente usado quando o piloto está em treinamento, antes de começar a pilotar a aeronave. Entretanto, conforme o regulamento estabelecido pela OACI¹²⁶, é obrigatória a presença desse manual a bordo da aeronave, sempre na língua em que foi escrita pelo fabricante, neste caso, a língua inglesa, sendo assim consultado sempre que os pilotos julgarem necessário durante um voo. O QRH é composto por *checklists* para a realização de operações normais, e por instruções para situações de emergência. É quase que unicamente consultado somente durante um voo. Com essas características em mente, busco observar se as diferenças contextuais são expressas linguisticamente através dos VMs.

A frequência do grupo dos modais é comparada entre os dois manuais para pilotos assim como a diferença entre cada modal individualmente. Como pode ser visto, as diferenças entre os dois manuais para pilotos são bem menores do que as diferenças encontradas entre o MM e o OM+QRH. Os modais que apresentam diferenças estatisticamente significativas são os VMs *can* e *must*, que são mais frequentes no OM e *may*, que é mais frequente no QRH.

¹²⁶ Organização da Aviação Civil Internacional

Tabela 7.2.1.D Distribuição dos Verbos Modais no Corpus de Aviação (AC) %

1	2	3	4	5	6
VM	tokens no AC	AC %	MM %	OM %	QRH %
<i>Can</i>	1168	39,28	46,84	31,23	16,09
<i>Could</i>	53	1,78	2,25	1,14	1,00
<i>May</i>	485	16,31	6,36	27,50	43,71
<i>Might</i>	03	0,10	0,17	0,00	0,00
<i>Must</i>	507	17,05	18,76	16,42	5,52
<i>Shall</i>	20	0,67	1,15	0,19	0,0
<i>Should</i>	165	5,54	3,82	7,35	11,05
<i>Will</i>	558	18,76	20,55	15,56	20,10
<i>Would</i>	14	0,47	0,17	0,57	2,51
Total %	100	100	100	100	100
tokens	2973	2973	1727	1047	199
%o		6,4	6,8	5,7	6,6

Na tabela 7.2.1.D, é possível ver a frequência relativa dos modais nos três manuais. A coluna 2 mostra que 2973 tokens no AC são modais. A coluna 3 mostra o percentual de cada modal entre o total de ocorrências dos VMs no AC. *Will*, *can*, *may* e *must* totalizam 91,40% de todos os modais, sendo *can* o mais freqüente com 39,28%. Em contrapartida, *could*, *might*, *shall* e *would* juntos respondem por apenas 3,02%, sendo *might* o menos freqüente.

As colunas 4, 5 e 6 mostram a parcela de participação que cada modal tem entre o total de ocorrência dos VMs em cada um dos corpora. *Can* sozinho representa mais de 46% das ocorrências modais, quando somado a *must* e a *will* obtém-se o valor de 86,15%. *Might* e *would* juntos correspondem a menos de 0,5% das ocorrências. *Could*, *may*, *shall* e *should* respondem por 13,58%. Como já ressaltado anteriormente, os corpora especializados parecem concentrar seus usos em alguns VMs, provavelmente devido ao número reduzido de atos de fala em comparação aos corpora de inglês geral, como por exemplo, o BNC.

No OM, *can*, a exemplo do MM, também é o mais freqüente, entretanto, responde por 31% das ocorrências dos VM, bem abaixo dos 46% que o *can* possui no MM. A concentração de usos parece ser um pouco mais distribuída no OM, apesar da ausência total de *might*. Quatro verbos estão entre os mais freqüentes no OM, *can*, *may*, *must* e *will*, que correspondem a 90,71% das ocorrências modais. *Should* é usado em 7,35% das ocorrências, enquanto *could*, *shall* e *would* apresentam pouquíssimas ocorrências.

May é o VM que mais apresenta tokens no QRH, com 43,71% do total de VMs no subcorpus, diferentemente do OM e do MM nos quais o *can* é o mais representativo. Como nos outro dois subcorpora, apenas quatro modais correspondem a quase a totalidade das ocorrências, que no caso do QRH são os VMs *may*, *can*, *should* e *will*.

Apesar das diferenças, pode-se notar algumas similaridades entre os três subcorpora relacionadas às frequências e distribuição dos VMs. Em primeiro lugar, há um VM que corresponde a pelo menos 1/3 das ocorrências modais em cada corpus, sendo *can* no OM e no MM e *may* no QRH. Em segundo lugar, *could*, *might*, *shall* e *would* apresentam pouquíssimas ocorrências nos três subcorpora.

Item B:

Analiso agora a distribuição dos VMs nas estruturas sintáticas possíveis de ocorrência. Essa etapa é importante para conhecer a estrutura prevalente em cada manual, assim como o “ambiente” real de ocorrência de cada modal em cada um dos manuais de forma a desenvolver materiais didáticos mais eficientes. Além disso, como já mencionado, a relação entre léxico e padrão gramatical é bem mais complexa do que lingüistas imaginavam até pouco tempo atrás. Essa complexidade inclui a associação de um dos sentidos de um item lexical com padrões sintáticos preferidos. Nesse sentido, para se chegar à padronização ou fraseologia de cada VM, parte-se primeiramente da estrutura para então chegar-se ao significado e ao uso de cada uma das unidades lexicais estudadas.

A tabela Tabela 7.2.1.E mostra que as estruturas um (1), seis (6), sete (7), oito (8) e nove (9) não apresentam ocorrências no AC. As estruturas quatro (4) e cinco (5) têm frequências extremamente baixas, somando apenas 0,3% dos *tokens* modais em todo o corpus. As estruturas dois (2) (voz ativa) e três (3) (voz passiva), por sua vez, respondem por 99,7% das ocorrências, com 56,90% e 42,02% respectivamente. A frequência da estrutura dois (2) no MM corresponde àquela do BNC, ou seja, é a mais freqüente. No QRH, a estrutura dois (2) e a estrutura três (3) aparecem praticamente a mesma quantidade de vezes, aproximadamente 50% cada uma. O OM apresenta uma maior ocorrência da estrutura três (3) (55,68%) do que da estrutura dois (2) (44,03%).

Tabela 7.2.1.E Estrutura sintática dos VMs nos três manuais (%)

	Estrutura	AC	MM	OM	QRH
1	Modal sozinho	0,0	0,0	0,0	0,0
2	Modal+voz ativa	67,3	83,51	44,03	49,75
3	Modal+voz passiva	32,4	16,32	55,68	49,25
4	Modal+ <i>be</i> +presente particípio	0,2	0,00	0,29	1,00
5	Modal+ <i>have</i> +passado particípio	0,1	0,17	0,0	0,0
6	Modal+ <i>be+being</i> + passado particípio	0,0	0,0	0,0	0,0
7	Modal+ <i>have+been</i> + passado particípio	0,0	0,0	0,0	0,0
8	Modal+ <i>have+been</i> +presente particípio	0,0	0,0	0,0	0,0
9	Modal+ <i>have+been+being</i> + passado particípio	0,0	0,0	0,0	0,0

No sentido de conhecer a padronização de cada VM, apresento agora a distribuição dos VMs nas estruturas sintáticas em que ocorrem. Esses dados serão mostrados separadamente por manual, uma vez que já foi visto que cada um deles apresenta diferentes preferências por VM e estrutura sintática, sendo, ao meu ver, infrutífero considerá-los como um único tipo de texto. Esses dados serão apresentados nas próximas três tabelas. Na tabela 7.2.1.F apresento as ocorrências do MM.

Tabela 7.2.1.F Distribuição dos VMs nas estruturas sintáticas no MM

	Estruturas	Can %	Could %	May %	Might %	Must %	Shall %	Should %	Will %	Would %
2	Modal+ infinitivo	89,37	92,30	74,54	33,33	77,47	83,33	48,48	84,79	66,67
3	Modal+be+ passado particípio	10,63	7,70	25,46	0,0	22,53	16,67	51,52	14,92	33,33
4	Modal+be+ presente particípio	0,0	0,0	0,0	0,0	0,86	0,0	0,0	0,29	0,0
5	Modal+have+ passado particípio	0,0	0,0	0,0	66,67	0,43	0,0	0,0	0,0	0,0

Como esperado (tendo em vista a tabela 7.2.1.E que mostra que 83,51% dos VMs ocorre na estrutura dois) oito, entre os nove modais estudados, preferem a estrutura dois (2) (voz ativa). Apenas *should* mostra uma pequena tendência à estrutura três (3), com 51,52% das ocorrências na estrutura três (3) comparado a 48,48% das ocorrências na estrutura dois (2). Em números reais, essa diferença percentual é traduzida em 32 ocorrências na estrutura dois e 34 na estrutura três, ou seja, apenas duas ocorrências, pois *should* está entre os VMs menos usados no MM. Já o uso das estruturas do modal *can* apresenta substanciais diferenças em números reais, pois há 809 *tokens* do modal *can*, sendo 723 ocorrências (89,37%) na estrutura dois comparada a apenas 86 ocorrências na estrutura três (10,63%). *Might* e *would* apresentam apenas três *tokens* cada. Assim, cada ocorrência corresponde a 33,3% do total. Em síntese, pode-se afirmar que o MM apresenta uma clara preferência pelo uso da estrutura dois (voz ativa), expressando, assim, as instruções de uma forma mais direta quando comparado aos manuais para pilotos.

A seguir, a tabela 7.2.1.G mostra as quantidades percentuais que cada modal ocorre em cada estrutura no OM. Diferentemente do MM, o OM mostra uma clara preferência pela estrutura três (3), voz passiva, com 55,68% das ocorrências. Os VMs *can*, *may*, *must* e *should* acontecem em maior quantidade nessa estrutura. Por outro lado, *could*, *will* e *would* apresentam mais ocorrências na estrutura dois (2), voz ativa. Entretanto, *could* e *would*

apresentam uma quantidade baixíssima de ocorrências (12 e seis respectivamente). Entre esses três modais, apenas *will* possui um total bruto significativo com 136 ocorrências na estrutura dois (2) e 27 ocorrências na estrutura três (3).

Tabela 7.2.1.G Distribuição dos VMs nas estruturas sintáticas no OM

	Estruturas	Can %	Could %	May %	Might %	Must %	Shall %	Should %	Will %	Would %
2	Modal+ infinitivo	25,99	91,66	39,93	0,0	47,09	50,0	33,77	83,43	100
3	Modal+be+ passado particípio	74,01	8,34	59,37	0,0	51,74	50,0	66,23	16,57	0,0
4	Modal+be+ presente particípio	0,0	0,0	0,7	0,0	1,17	0,0	0,0	0,0	0,0
5	Modal+have+ passado particípio	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0

Em seguida, mostro a distribuição dos VMs entre as estruturas no QRH, lembrando que as estruturas dois (2) e três (3) apresentam uma quantidade de ocorrências muito semelhante com 49,75% e 49,25% respectivamente. *Could*, *may*, *will* e *would* apresentam mais ocorrências na estrutura dois (2), ao passo que *can*, *must* e *should* preferem a estrutura três (3).

Tabela 7.2.1.H Distribuição dos VMs nas estruturas sintáticas no OM

	Estruturas	Can %	Could %	May %	Might %	Must %	Shall %	Should %	Will %	Would %
2	Modal+ infinitivo	34,37	100	55,17	0,0	18,18	0,0	18,18	72,50	60
3	Modal+be+ passado particípio	65,63	0,0	42,52	0,0	81,82	0,0	81,82	27,50	40
4	Modal+be+ presente particípio	0,0	0,0	2,31	0,0	0,0	0,0	0,0	0,0	0,0
5	Modal+have+ passado particípio	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0

7.2.2 Considerações Parciais

Nesta seção observei a incidência dos VMs em cada um dos manuais assim como a distribuição dos modais entre as estruturas sintáticas. Para uma melhor visualização, o quadro

abaixo apresenta os VMs em ordem decrescente de frequência de ocorrência em cada um dos manuais, sintetizando os resultados obtidos na seção anterior.

Quadro 7.2.2.A MM x OM x QRH

MM	OM	QRH
<i>Can</i>	<i>Can</i>	<i>May</i>
<i>Will</i>	<i>May</i>	<i>Will</i>
<i>Must</i>	<i>Must</i>	<i>Can</i>
<i>May</i>	<i>Will</i>	<i>Should</i>
<i>Should</i>	<i>Should</i>	<i>Must</i>
<i>Could</i>	<i>Could</i>	<i>Would</i>
<i>Shall</i>	<i>Would</i>	<i>Could</i>
<i>Might/Would</i>	<i>Shall</i>	

Com base nas informações obtidas até o momento, esboço algumas considerações parciais a respeito do corpus de estudo. Reitero que o presente trabalho tem por finalidade conhecer o funcionamento dos VMs visando ao desenvolvimento de material didático para pilotos e mecânicos. Dessa forma, baseado somente na análise quantitativa, é possível sugerir que os VMs *can*, *must* e *will* são merecedores de mais atenção e explicações mais detalhadas quando o público for constituído de mecânicos de aeronaves. Os VMs *may* e *should* devem ser também incluídos, uma vez que suas frequências são altas o suficiente. Por outro lado, quando o público for constituído de pilotos, o material didáticos deverá priorizar a presença de *can*, *may*, *must*, *will* e *should*.

Acredito que *must* mereça uma explicação especial tendo em vista sua baixa incidência no QRH, pois, por ser o manual que fornece instruções regulares e para lidar com emergências, o QRH seria o candidato mais provável para um uso maior das modalidades deônticas fortes. Entretanto ele é mais freqüente no MM e no OM. Esse fato pode ser explicado pela alta frequência de imperativos puros no QRH¹²⁷. Ou seja, as instruções são muitas vezes expressas sem a presença de um auxiliar modal.

Quanto às estruturas sintáticas, vimos que o MM apresenta uma maior incidência da estrutura dois (voz ativa) com 83,51% dos *tokens* modais. Essa incidência é ainda mais alta do que o BNC e o OSC, pois ambos apresentam aproximadamente 75% dos *tokens* modais nessa estrutura. Ou seja, a frequência de ocorrência da estrutura dois no MM é marcadamente alta. Uma das razões para tal preferência pode ser as diretrizes estabelecidas pelo *Simplified*

¹²⁷ Conforme Gabrielatos e Sarmiento (2006).

*English*¹²⁸, que instrui redatores técnicos a evitarem a estrutura modal+ voz passiva. Entretanto, devido a quase inexistência das outras estruturas, a estrutura três apresenta uma frequência bastante semelhante ao BNC.

Nos manuais para pilotos, principalmente no OM, observa-se o quadro inverso. A estrutura três, voz passiva, apresenta uma frequência de ocorrência bem mais alta do que os outros subcorpora ou do que os outros corpora estudados na seção anterior. No BNC, apenas 16,6% das ocorrências modais são nessa estrutura, e no OSC 23,13%. O OM apresenta 55,68% dos VMs na forma passiva, uma ocorrência 330% mais alta do que no BNC e 240% mais alta do que no OSC. Assim, é possível afirmar que modal+passiva é uma característica peculiar dos manuais para pilotos, ainda mais marcada no OM. Esse uso aponta, provavelmente, a uma tendência a mitigar as instruções nesses manuais, de forma que elas não pareçam tão diretas e impositivas, pois o uso da voz passiva é uma das estratégias utilizadas para esse fim.

No quadro abaixo sintetizo as preferências que os modais têm por cada estrutura nos três manuais. Como alguns modais são extremamente infreqüentes, apresento apenas aqueles com pelo menos 03% do total de ocorrências modais em cada subcorpus:

Quadro 7.2.2.B Contraste entre as estruturas sintáticas no MM x OM x QRH

Estrutura	MM	OM	QRH
Estrutura 2 (voz ativa)	<i>can, must, will, may</i>	<i>will</i>	<i>may, will</i>
Estrutura 3 (voz passiva)	<i>should</i>	<i>can, may, must, should</i>	<i>can, must, should</i>

Will aparece mais freqüentemente na estrutura dois (voz ativa) nos três corpora, enquanto *should* aparece na estrutura três (voz passiva) nos três corpora. Os outros VMs apresentam diferentes preferências em cada um deles. Com base nessas observações, sugiro que nos materiais didáticos para mecânicos, *can, may, must* e *will* apareçam mais na voz ativa, e *should* na voz passiva. Por outro lado, os materiais para pilotos deveriam apresentar *will* principalmente na voz ativa, *can, must* e *should* na voz passiva, e *may* em ambas, uma vez que é mais freqüente na voz ativa no QRH e mais freqüente na voz passiva no OM.

Nessas duas seções, realizei vários levantamentos quantitativos. Primeiramente contrastei o corpus de estudo, o corpus de referência e o corpus de contraste. O objetivo desse contraste triplo foi verificar quais os itens e estruturas lexicais que são comuns na linguagem

¹²⁸ O *Simplified English* é descrito em detalhes no capítulo 3.

geral, quais são próprios da linguagem técnica dos manuais, e, mais relevante para o presente trabalho, os itens peculiares na linguagem da aviação. Na segunda etapa, verifiquei os itens e estruturas salientes de cada corpus e sugeri os VMs e as estruturas que deveriam ser incluídas em materiais didáticos para os diferentes públicos: mecânicos e pilotos. Com este levantamento quantitativo, já é possível perceber os itens salientes de cada manual. Partindo dessas saliências, passo agora à análise qualitativa de cada modal em cada um dos manuais com vistas a estabelecer uma taxonomia de uso.

7.3 TERCEIRA ETAPA: PADRÕES E REGULARIDADES

Nesta seção, são apresentados os usos de cada VM em dois manuais que compõem o AC (corpus de estudo), o MM e o OM. Para essa análise, é necessária a observação de todos os contextos de ocorrência dos itens em questão, usando-se instrumentos como concordância e lista de colocados¹²⁹. O motivo para esse tipo de análise está na crença de que as palavras possuem padronização, escolhendo algumas outras palavras em detrimento de outras. É preciso ressaltar, neste momento, que cada pesquisa desenvolvida a partir de uma metodologia baseada em corpus é única. Sendo assim, os procedimentos adotados em cada pesquisa são diferentes e norteados pelos dados e pelas necessidades impostas pela pesquisa no decorrer da análise.

A descrição do comportamento de cada VM é feita baseada nos resultados obtidos na seção anterior que apontou a estrutura sintática em que cada VM ocorre preferencialmente em cada um dos dois manuais analisados. Apesar de os modais poderem ocorrer em nove estruturas sintáticas diferentes, os resultados mostraram que no Corpus de Aviação (AC), há duas estruturas sintáticas predominantes:

- a) Estrutura 2 -modal +infinitivo (modal + voz ativa)
- b) Estrutura 3- modal + *be*+ passado particípio (modal+ voz passiva)

Dessa forma, os colocados para cada modal serão mostrados, seguindo essa classificação prévia realizada na seção 7.2.

Esta seção é dividida em três etapas metodológicas para cada VM em cada manual (MM, OM) que serão, sempre que possível¹³⁰, seguidas:

A Colocados à esquerda

¹²⁹ Explicadas em detalhe no capítulo 3.

¹³⁰ Esse roteiro metodológico não poderá ser seguido em todos os casos devido ao baixo número de ocorrências de alguns modais em algumas das estruturas. Nesses casos, não é possível detectar regularidades ou colocações.

B Colocados à direita

C Usos e funções

Apesar de já terem sido descritas no capítulo 6 sobre metodologia, retomo essas noções a seguir.

A Colocados à esquerda

Analiso e classifico os colocados¹³¹ à esquerda dos VMs em cada uma das estruturas¹³² a partir da lista de colocados e concordâncias. Este item é subdividido em duas outras etapas. Em um primeiro momento são mostrados os 10 (quando houver 10) colocados com maior grau de atração ao nóculo. A seguir, classificarei os colocados à esquerda em cinco grupos: (1) *you*; (2) anáforas; (3) unidades simples; (4) sintagmas nominais; (5) sintagmas verbais e preposicionais. O único dos grupos (nas posições L1, L2, L3...) formado de um único item lexical é o grupo “*you*”.

B Colocados à direita

Analiso e classifico os colocados à direita dos VMs em cada uma das estruturas a partir da lista de colocados e concordâncias. Este item, a exemplo do Item A, também é subdividido em duas fases. A primeira etapa apenas aponta os 10 colocados com maior grau de atração ao nóculo (quando houver 10). Diferentemente dos colocados nas posições à esquerda, que podem pertencer a diferentes classes gramaticais, a posição R1 apresenta uma restrição gramatical quanto ao tipo de elemento lexical que pode seguir o VM. Dessa forma, o VM pode ser imediatamente seguido apenas por verbo ou advérbio. Nos casos de VM + voz passiva, o *be* não é considerado um colocado do VM por ser considerado elemento obrigatório para estabelecer a estrutura. Dessa forma, nos casos de voz passiva, a posição R1 é a primeira posição após o *be*.

Na segunda etapa, classifico semanticamente todos os verbos colocados à direita do VM nas posições R1, ou R2 (quando precedido de um advérbio), conforme a taxonomia proposta por Biber et al. (1999) (explicada em detalhes na seção sobre metodologia). Essa observação, seguida pela classificação foi obtida através da análise das linhas de concordância

¹³¹ Para medir a força de cada colocação, foi utilizado o teste estatístico *Log Likelihood*¹³¹. Se o resultado obtido após a aplicação do LL foi de 6,63 ou mais, a probabilidade de a atração entre os dois itens lexicais ter acontecido aleatoriamente é de menos de 1%. Ou seja, a função da estatística é apontar se os itens realmente formam colocações ou se são co-ocorrências espúrias.

¹³² Os resultados serão apresentados nesta ordem quando houver ocorrências dos VM nessas duas estruturas.

do VM nas duas estruturas, voz ativa e voz passiva. As sete categorias usadas para classificar os verbos lexicais que acompanham os VMs são: (1) de atividade, (2) de comunicação, (3) mentais, (4) causativos, (5) de ocorrência, (6) de existência e (7) de aspecto.

Item C: Usos e funções

Com base nos resultados obtidos nos itens A e B, analiso as linhas de concordância de cada VM, observo seus usos e peculiaridades e sugiro exemplos que expressem esses usos.

7.3.1 Maintenance Manual (MM)

Como já mencionado anteriormente, o MM é o subcorpus mais extenso do AC, com 249.691 *tokens*. Os destinatários deste manual são os mecânicos de aeronaves. Este é o único entre os três subcorpora do AC que possui um guia de redação, o *Simplified English* (SE, descrito no capítulo 3). Isto é, há regras no guia SE que influenciam, ou, devem influenciar a forma como este manual é redigido. Passo agora aos resultados.

7.3.1.1 Can

No MM, o modal *can*, que é o modal mais freqüente em todo o corpus (AC) apresenta um total de 809 ocorrências, sendo:

- 723 ocorrências na estrutura 2, voz ativa
- 86 ocorrências na estrutura 3, voz passiva

A Colocados à esquerda

O quadro abaixo mostra os 10 principais colocados à esquerda juntamente com a quantidade de ocorrências:

Tabela 7.3.1.1.A: Colocados à esquerda (L1, L2, L3...) de *can* no MM

Posição	Colocado <i>can</i> voz ativa	Quantidade	Relação LL	Colocado <i>can</i> voz passiva	Quantidade	Relação LL
1	YOU (L1)	191	1.292,390	SEALED (L2)	04	25,705
2	WHICH (L1)	27	190,390	FORCE (L3)	04	23,862
3	THIS (L1)	34	190,312	THEY (L1)	04	16,751
4	THAT (L1)	37	179,180	BEARINGS (L1)	04	11,007
5	OIL (L1)	33	168,686	GREASE (L3, L2)	04	8,607
6	IT (L1)	26	152,536	WHICH (L1)	04	7,645
7	AGENTS (L2)	12	109,640	-----	-----	-----
8	HOT (L1)	16	109,625	-----	-----	-----
9	EXPLOSION (L1)	9	106,886	-----	-----	-----
10	FLAMMABLE (L3)	12	103,762	-----	-----	-----

Destaca-se sobremaneira a frequência e a força de atração de *you can*¹³³. É também surpreendente, as diferentes classes de palavras que ocupam a posição L1 de *can*, com vários determinantes (*which*, *this*, e *that*), e um adjetivo (*hot*). Por tratar-se de um verbo, a expectativa era de haver substantivos, como no caso de *oil* e *explosion* ou pronomes pessoais retos, como *you*. A lista de colocado fornece “pistas” a respeito do nóculo. Entretanto, faz-se necessário conhecer e classificar todos os colocados à esquerda de *can*. Após a análise de todas as linhas de concordância de *can*, os resultados foram:

Tabela 7.3.1.1.B: Classificação das unidades lexicais à esquerda de *can* no MM

Colocados à esquerda <i>Can</i> - voz ativa	Número de ocorrências	Colocados à esquerda <i>Can</i> -voz passiva	Número de ocorrências
<i>You</i> (L1)	190/ 26,3%	<i>You</i> (L1)	-----
Anáforas (L1)	97/ 13,4%	Anáforas (L1)	13/ 15,1%
Unidades Simples (L1)	88/ 12,2%	Unidades Simples (L1)	15/ 17,4%
Sintagmas nominais	149/ 20,6 %	Sintagmas nominais	28/ 32,6%
Sintag. Verbais e prepos.	199/ 27,5%	Sintag. Verbais e prepos.	30/ 34,9%

¹³³ Lembro que a lista de colocados é baseada no LL para calcular a força da colocação, e não somente no fator frequência. Dessa forma, um item mais frequentemente encontrado na presença do nóculo, como por exemplo, *can get* (26 ocorrências), é considerado menos significativo, ou fraco, do que um colocado com menos ocorrências, como por exemplo *can come* (12). Isso se deve ao fato de que *get* aparece 185 vezes nesse corpus, enquanto *come* apresenta apenas 19 ocorrências em todo o corpus. Assim, a atração de *come* por *can* é proporcionalmente mais forte do que a de *get* por *can*.

Como pode ser visto, o *you* sozinho responde por 26, 3% das ocorrências de *can* na voz ativa no MM. Os outros grupos são formados por variados itens lexicais. Como era de se esperar, o *you* não está presente nos casos de *can* +voz passiva, pois os colocados à esquerda em uma oração passiva não são geralmente pronomes pessoais retos.

B Colocados à direita

Mostro agora os colocados à direita de *can*.

Tabela 7.3.1.1.C: Colocados à direita (R1, R2, R3) de *can* no MM

Posição	Colocado <i>can</i> voz ativa	Quantidade	Relação LL	Colocado <i>can</i> voz passiva	Quantidade	Relação LL
1	CAUSE (R1)	218	2.197,822	DAMAGED (R1)	11	56,666
2	OCCUR (R1)	84	844,076	USED (R1)	12	53,643
3	DAMAGE (R2)	72	702,852	BY (R2)	13	45,657
4	INJURY (R2)	43	535,313	FOUND (R1)	6	30,215
5	USE (R1)	52	239,650	THAT (R2)	4	7,606
6	BE (R1)	46	225,430	WITH (R2)	7	6,792
7	INJURIES (R2)	15	150,427	-----	-----	-----
8	COME (R1)	12	139,476	-----	-----	-----
9	GET (R1)	26	135,120	-----	-----	-----
10	BURN (R1)	13	134,380	-----	-----	-----

Nesta lista dos principais colocados de *can* à direita, nota-se que aparecem apenas verbos na posição R1, como esperado. É bastante saliente, entretanto, a frequência e atração do verbo lexical *cause* na voz ativa. As 218 ocorrências de *can cause* correspondem a mais de 30% de *can* (voz ativa) no MM. Os verbos *occur* e *use* também apresentam uma forte atração com o modal *can*. Destaco que o cópula *be*, entre os colocados de *can*+voz ativa, não é um verbo auxiliar como no caso das construções passivas, mas o verbo principal. Ele é o quarto verbo com maior atração a *can*.

Nas ocorrências de *can* + voz passiva, os verbos com maior grau de atração são *damaged* e *used*, que apresentam um grau de atração bem inferior a *cause* na voz ativa. A ser destacado, mas que na verdade não apresenta qualquer surpresa, é a presença da preposição

by, que é geralmente usada em construções passivas para apontar a entidade que realizou a ação. Além disso, destaca-se também a presença do verbo *found*.

Tabela 7.3.1.1.D: Verbos na posição R1 de *can* -MM

Tipos de Verbos	<i>Can</i> -voz ativa	<i>Can</i> -voz passiva
De atividade	202/28%	58/ 68%
De comunicação	-----	-----
Mentais	17/2,4 %	15/ 17%
Causativos	305/42%	13/ 15%
De ocorrência	130/18%	-----
De existência	69/9,6%	-----
De aspecto	-----	-----

Como era de se esperar, após a constatação da expressiva ocorrência de *cause*, os causativos são os colocados mais frequentes nos casos de voz ativa, seguidos pelos verbos de atividade e de ocorrência.

Há uma predominância dos verbos de atividade entre os colocados (R1) de *can* (voz passiva). Entretanto, cabe mencionar que não há um único item lexical que domine a maioria das ocorrências (como por exemplo, o *cause*, entre os casos de voz ativa que sozinho responde por mais de 30% dos colocados R1 de *can*), sendo o *used* o mais frequente. Assim, a maior parte dos verbos nesse grupo apresenta apenas uma ou duas ocorrências.

C Usos e funções

Foi constatado que:

- *Can* é usado com bastante mais frequência na voz ativa, na proporção de uma ocorrência na voz passiva a cada nove na voz ativa. É usado principalmente na forma afirmativa, na proporção de uma ocorrência negativa (seguidas pelo advérbio *not*) a cada dez afirmativas.
- Todas as ocorrências afirmativas de *can* expressam a idéia de possibilidade, podendo essa possibilidade ser relacionada a uma prosódia semântica¹³⁴ neutra ou negativa. Foi observado um contexto estendido¹³⁵, além somente da linha de concordância, sempre que necessário para classificar o uso expresso pelo VM.
- *Can*, quando acompanhado por um verbo causativo, ou de ocorrência (60% dos casos), é utilizado para expressar uma consequência negativa que pode

¹³⁴ A noção de prosódia semântica é explicada no capítulo 2.

¹³⁵ Sempre que necessário, foi observado um contexto estendido para que o sentido pudesse ser identificado. O contexto estendido é facilmente obtido clicando-se na linha de concordância que se quer observar. Através dessa ação tem-se acesso ao texto como um todo.

acontecer devido a uma ação mal realizada pelo mecânico. Em outras palavras, nesses casos *can* possui uma prosódia semântica negativa.

- Nas seções de *WARNING* (Aviso) ou *CAUTION* (Cuidado), é freqüente o uso de *can*, com 358 ocorrências nesta seção. Torna-se fácil observar se o VM ocorre em alguma dessas duas seções, uma vez que, é feito o uso de caixa alta (letras maiúsculas). O uso de *can* nestas seções é sempre vinculado a uma prosódia semântica negativa.
- *Can* acompanhado de outros tipos de verbos não apresenta a característica de expressar uma conseqüência negativa de forma tão marcante. Sua função geralmente expressa apenas uma possibilidade mais neutra, mostrando possíveis ações a serem realizadas.
- *Can* ocorre com freqüentemente em orações condicionais com a presença da conjunção *if* (74 ocorrências), principalmente na oração principal de uma condicional (60 ocorrências). Entretanto, *can* também aparece na oração subordinada (14 ocorrências).
- Nos contextos negativos também predomina a função de possibilidade, ou melhor, de ausência de possibilidade devido a uma causa externa.
- Apenas três ocorrências de *cannot* expressam a idéia de forte recomendação negativa; isto é, recomendando o mecânico a não usar certos dispositivos, por exemplo:

*C_A_U_T_I_O_N_: SOME OIL TYPES **CANNOT BE** MIXED WITH OTHER TYPES OF OIL. DO NOT MIX OIL TYPES, UNLESS YOU ARE SURE THAT THEY CAN BE MIXED OR ENGINE DAMAGE CAN OCCUR.*

Com base nos resultados apresentados nos itens A, B e C, ilustro as características e usos mais salientes de *can* nesse manual, o MM.

*1-C_A_U_T_I_O_N_: CAREFULLY MOVE ROPES OR FABRIC HOSES ON THE WING OR FUSELAGE. EQUIPMENT THAT IS INSTALLED ON THE SURFACE OF THE WING OR FUSELAGE **CAN BE** DAMAGED BY THE MOVEMENT OF THE ROPES OR FABRIC.*

*2- N_O_T_E_: If the fan (N1) **cannot** turn during an engine start because of ice, this **can** cause heavy damage.*

3- If you **can** see the lining after bolt removal, refer to the bench check instructions and examine the lining material.

4- W_A_R_N_I_N_G : INSTALL LEADING EDGE FLAP AND SLAT LOCKS TO PREVENT INADVERTENT OPERATION OF LEADING EDGE FLAPS AND SLATS. FLAPS AND SLATS ARE FAST ACTING AND **CAN** CAUSE SERIOUS INJURY TO PERSONNEL.

5- Immediately after refueling, the fuel quantity shown on the FMC **can** be less than the actual fuel quantity in the tank.

7.3.1.2 Could

No MM, o modal *could* apresenta um total de 39 ocorrências, sendo:

- 36 ocorrências na estrutura 2, voz ativa
- três ocorrências na estrutura 3, voz passiva

A Colocados à esquerda

Na voz ativa, *could* apresenta 12 colocados ao total, incluindo as duas palavras gramaticais *the* e *and*, que não serão incluídas na lista de colocados. Como o *could* na voz passiva apresenta somente três ocorrências, elas não serão analisadas seguindo todas as etapas metodológicas. Suas ocorrências serão mostradas e comentadas no Item C.

O quadro abaixo mostra os colocados à esquerda juntamente com a quantidade de ocorrências:

Tabela 7.3.1.2.A: Colocados à esquerda de *could* no MM

Posição	Colocado <i>Could</i> voz ativa	Quantidade	Relação LL
1	THIS (L1)	5	100,719
2	LONG (L3)	3	79,363
3	WHICH (L1)	4	64,087
4	LANDING (L1)	6	60,653

Devido à baixa frequência de *could*, não há dez itens lexicais que preenchem às exigências (mínimo de quatro ocorrências e relação LL de 6,6) para serem incluídos na lista de colocados. Nota-se que o determinante *this* é o colocado que apresenta maior grau de atração com o *could*, seguido por *long*, só que na posição L3. Os outros principais colocados na posição L1 são *which* e *landing*.

Na próxima tabela mostro a classificação dos colocados à esquerda:

Quadro 7.3.1.2.B: Colocados de *could*-voz ativa-MM

Colocados à esquerda	Número de ocorrências
<i>You</i> (L1)	-----
Anáforas (L1)	13/ 36%
Unidades simples (L1)	11/ 30,5%
Sintagmas nominais	05/ 13,9%
Sintag. verbais e preposic.	07/ 19,4%

Como é possível observar, o grupo com mais ocorrências é aquele representado pelas anáforas, sendo o grupo “sintagmas nominais” o menos representativo. Saliento a ausência do pronome *you*.

B Colocados à direita

Descrevo agora os contextos com base nos colocados à direita de *could*.

Tabela 7.3.1.2.C: Colocados à direita de *could* no MM

Posição	Colocados <i>Could</i> voz ativa	Quantidade	Relação LL
1	CAUSE (R1)	18	302,994
2	OCCUR (R1)	05	96,972
3	RETRACT (R2)	05	81,814
4	UNWANTED (R3)	03	79,363
5	DAMAGE (R3)	04	75,702
6	POSSIBLE (R2)	04	59,347

Entre os poucos colocados encontrados, nota-se que há uma semelhança aos colocados de *can*, no caso *cause*, *occur* e *damage*. O *cause* sozinho responde por exatamente 50% das ocorrências de *could* no MM. *Retract* também possui um grau de atração forte, sendo o terceiro mais freqüente. Os outros itens lexicais que aparecem à direita de *could* não estão na posição R1.

Como era de se esperar, após a constatação da expressiva ocorrência de *cause*, os causativos são os colocados mais freqüentes, seguidos pelos verbos que expressam atividade e de ocorrência simples.

Tabela 7.3.1.2.D: Verbos na posição R1 de *could* do MM

Tipos de Verbos	<i>Could</i> -voz ativa
De atividade	05 / 13%
De comunicação	-----
Mentais	-----
Causativos	19/ 55%
De ocorrência	7/ 19%
De existência	05/ 13%
De aspecto	-----

Há uma predominância dos verbos causativos entre os colocados (R1) de *could*, pois somente o *cause* apresenta 18 ocorrências. As outras ocorrências são mais uniformemente divididas entre os outros tipos de verbo.

Item C: Usos e funções

Como as ocorrências na voz passiva não puderam ser analisadas conforme o roteiro metodológico acima, apresento as linhas de concordância com as três ocorrências a seguir:

N Concordance

- 1 is a list of cleaners and polishes that **could be** required . during regular
- 2 of torque-to-tension of threaded fasteners **could be** changed from the molybdenum
- 3 3) When the friction of braking components **could be** decreased from the molybdenum

Com base nessas três ocorrências e nos itens A e B, pode-se concluir que:

- Todas as ocorrências de *could* são em contextos positivos e expressam a idéia de possibilidade.
- *Could*, acompanhado por um verbo “causativo”, ou “de ocorrência”, é a exemplo de *can*, utilizado para expressar uma consequência negativa que pode acontecer devido a uma ação mal realizada pelo mecânico. Como essas colocações representam 74% dos casos de *could* no MM, pode-se dizer que esse é o principal uso de *could*.
- *Could* é usado basicamente na voz ativa, na proporção de uma ocorrência na voz passiva a cada onze na voz ativa.
- *Could* é usado somente na forma afirmativa.
- *Could* tem como principais colocados à esquerda, as “Anáforas” e “Unidades simples”.
- *Could* aparece freqüentemente nas seções WARNING e CAUTION do MM, com 25 ocorrências, o que corresponde a 64% do total.

Dessa forma, acredito que os seguintes exemplos seriam representativos dessas características:

*1- IF THE OIL TOUCHES YOUR SKIN FOR A LONG TIME, IT **COULD** CAUSE DERMATITIS.*

*2-C_A_U_T_I_O_N_: MAKE SURE THE O-RINGS DO NOT HAVE PINCHES. MAKE SURE THE BOLTS AND FITTINGS ARE CORRECTLY ATTACHED. IF NOT, FAILURE **COULD** OCCUR.*

7.3.1.3 May

No MM, o modal *may* apresenta um total de 110 ocorrências, sendo:

- 82 ocorrências na estrutura 2, voz ativa
- 28 ocorrências na estrutura 3, voz passiva

A Colocados à esquerda

A tabela abaixo mostra os 10 principais colocados à esquerda juntamente com a quantidade de ocorrências:

Tabela 7.3.1.3.A: Colocados à esquerda (L1, L2, L3) de *may* no MM

Posição	Colocado <i>may</i> voz ativa	Quantidade	Relação LL	Colocado <i>may</i> voz passiva	Quantidade	Relação LL
1	HAZARDOUS (L3)	10	144,676	IT (L1)	04	15,689
2	SIGNALS (L2)	8	124,237	-----	-----	-----
3	SENSING (L3)	8	100,736	-----	-----	-----
4	MATERIALS (L2)	10	75,339	-----	-----	-----
5	WHICH (L1)	10	74,954	-----	-----	-----
6	SOLVENTS (L1)	10	74,954	-----	-----	-----
7	FLIGHT (L1)	7	49,522	-----	-----	-----
8	BEFORE (L2)	7	49,062	-----	-----	-----
9	PORTS (L3)	5	46,145	-----	-----	-----
10	VIBRATION (L1)	6	46,035	-----	-----	-----

Não há destaques com relação aos colocados à esquerda (quando comparados aos colocados de *can*, por exemplo). Na posição L1, os colocados mais fortes são *which* e

solvents, com dez ocorrências cada, seguidos de *flight* (sete ocorrências) e *vibration* (seis ocorrências). Cabe mencionar que, entre os dez colocados mais importantes nas posições L, três estão na posição L2 e três na posição L3, inclusive o que apresenta maior grau de atração conforme o LL, o adjetivo *hazardous*.

Mostro a classificação dos colocados à esquerda de *may* na tabela abaixo.

Tabela 7.3.1.3.B: Classificação das unidades lexicais à esquerda de *may* -MM

Colocados à esquerda <i>May</i> - voz ativa	Número de ocorrências	Colocados à esquerda <i>May</i> -voz passiva	Número de ocorrências
<i>You</i> (L1)	02/ 2,5%	<i>You</i> (L1)	-----
Anáforas (L1)	13/ 15,8%	Anáforas (L1)	05/ 17,8%
Unidades simples (L1)	16/ 19,5%	Unidades simples (L1)	07/ 25%
Sintagmas nominais	18/ 22%	Sintagmas nominais	05/ 17,8%
Sintag. verbais e prep.	33/ 40,2%	Sintag. verbais e prep.	11/ 39,4%

O grupo mais frequente é o composto de sintagmas verbais e preposicionais, tanto nos casos de voz ativa como nos casos de voz passiva. Em segundo lugar aparece o grupo de sintagmas nominais seguido por Unidades simples e anáforas. *You* apresenta apenas duas ocorrências na voz ativa.

B Colocados à direita

Descrevo agora os contextos com base nos colocados à direita de *may*.

Tabela 7.3.1.3.C: Colocados à direita (R1, R2, R3) de *may* no MM

Posição	Colocado <i>may</i> voz ativa	Quantidade	Relação LL	Colocado <i>may</i> voz passiva	Quantidade	Relação LL
1	BE (R1)	22	199,340	USED (R1)	07	47,494
2	FLAMMABLE (R2)	10	113,630	-----	-----	-----
3	ERRORS (R3)	08	111,348	-----	-----	-----
4	LOSS (R3)	08	103,515	-----	-----	-----
5	LEAD (R1)	08	100,736	-----	-----	-----
6	CAUSE (R1)	12	84,118	-----	-----	-----
7	LARGE (R2)	8	83,791	-----	-----	-----
8	PHASES (R3)	04	63,647	-----	-----	-----
9	REQUIRE (R1)	05	63,647	-----	-----	-----
10	NOISE (R3)	05	51,249	-----	-----	-----

O verbo *be* apresenta a maior força de colocação com *may*, lembrando que, nesse caso, o *be* não faz parte da estrutura (como nos casos de voz passiva). Os outros verbos com maior grau de atração são *lead*, *cause* e *require*. Os colocados principais na posição R2 são adjetivos, *flammable* e *large*, enquanto na posição R3 são substantivos, *errors*, *loss*, *phases* e *noise*.

Tabela 7.3.1.3.D: Verbos na posição R1 de *may* (voz ativa)-MM

Tipos de Verbos	<i>May</i> -voz ativa	<i>May</i> -voz passiva
De atividade	11/ 13,7%	21/ 75%
De comunicação	-----	02/ 7%
Mentais	4/ 5,4%	
Causativos	28/ 34%	03/ 11%
De ocorrência	6/ 6,8%	-----
De existência	33/ 40,1%	02/ 7%
De aspecto		-----

C Usos e funções

Foi constatado que:

- *May* é usado com bastante mais frequência na voz ativa, na proporção de uma ocorrência na voz passiva a cada três na voz ativa. É usado principalmente em contextos positivos, na proporção de uma ocorrência negativa a cada 22 positivas.
- Todas as ocorrências de *may* na forma afirmativa expressam a idéia de possibilidade, podendo essa possibilidade ser relacionada a contextos neutros ou contextos negativos, que serão explicados a seguir.
- *May*, quando acompanhado por um verbo “causativo”, ou “de ocorrência” (40% dos casos), é utilizado para expressar uma consequência negativa que pode acontecer devido á uma ação mal realizada pelo mecânico.
- Nas seções de *WARNING* (Aviso) ou *CAUTION* (Cuidado), é frequente o uso de *may* na voz ativa, com 34 ocorrências (41%) nestas seções. Nesses casos, *may* sempre introduz uma consequência negativa, isto é, possui uma prosódia semântica negativa.
- *May* acompanhado de outros tipos de verbos não apresenta a característica de expressar uma consequência negativa de forma tão marcante. Sua função geralmente expressa apenas uma possibilidade mais neutra, mostrando possíveis ações a serem realizadas.

- Quando *may* é apresentado na sua forma negativa (seguido por *not*) também predomina a função de possibilidade, ou melhor, de ausência de possibilidade devido a uma causa externa.

Com base nos resultados apresentados nos itens A, B e C, ilustro as características e usos mais salientes de *may* no MM através dos seguintes exemplos:

- 1- *W_A_R_N_I_N_G_ : FAILURE TO REMOVE BARRICADE TAPE AND VINYL ADHESIVE TAPE FROM THE STATIC PORTS BEFORE FLIGHT MAY CAUSE LARGE ERRORS IN AIRSPEED-SENSING AND ALTITUDE-SENSING SIGNALS, WHICH MAY LEAD TO LOSS OF SAFE FLIGHT.*
- 2- *SOLVENTS ARE HAZARDOUS MATERIALS. SOLVENTS MAY BE FLAMMABLE OR HARMFUL TO THE ENVIRONMENT.*
- 3- *Off-airplane data loading equipment **may be** used to load several software loadable LRUs.*

7.3.1.4 Might

No MM, o modal *might* apresenta apenas três ocorrências.

- uma ocorrência na estrutura 2, voz ativa

*N_O_T_E_ : When twisting wire, do not use any tool which **might** nick the wire.*

- duas ocorrências na estrutura 7, voz passiva+ perfeito

*Any areas from step (h) or (i) and any joint which **might** have been touched by mercury must be X-rayed.*

*4) Any areas from step (f) 3) and any joint which **might** have been touched by mercury must be X-rayed.*

Nota-se que as três ocorrências de *might* possuem um elemento anafórico CNA posição L1. Todas as três apresentam o sentido de possibilidade. Os dois exemplos de *might* +voz passiva+ perfeito são as únicas ocorrências de VMs nesse corpus (AC).

7.3.1.5 Must

No MM, o modal *must*, apresenta um total de 323 ocorrências, sendo:

- 250 ocorrências na estrutura 2, voz ativa
- 73 ocorrências na estrutura 3, voz passiva

A Colocados à esquerda

O quadro abaixo mostra os 10 principais colocados à esquerda juntamente com a quantidade de ocorrências:

Tabela 7.3.1.5.A: Colocados à esquerda (L1, L2, L3) de *must* no MM

Posição	Colocado <i>must</i> voz ativa	Quantidade	Relação LL	Colocado <i>must</i> voz passiva	Quantidade	Relação LL
1	YOU (L1)	159	1.281,429	ACTUATION (L2)	04	36,502
2	LIST (L2)	21	166,930	SLAT (L3)	05	24,871
3	THIS (L3)	24	141,216	VALVE (L1)	05	13,624
4	TACKY (L2)	4	106,555	-----	-----	
5	BUT (L1)	6	76,613	-----	-----	
6	IF (L2)	11	39,088	-----	-----	
7	WATER (L2)	8	37,105	-----	-----	
8	FITTINGS (L2)	6	31,775	-----	-----	
9	PERSONS (L1)	5	26,892	-----	-----	
10	AIRPLANE (L1)	4	25,828	-----	-----	

Destaca-se a força de atração de *you must*. É também notória a presença da conjunção *but* na posição L1, onde se tem a expectativa de encontrar substantivos, e pronomes pessoais retos.

Classifico as unidades lexicais à esquerda de *must* na tabela abaixo.

Tabela 7.3.1.5.B: Classificação do tipo de unidades lexicais à esquerda de *must* -MM

Colocados à esquerda <i>Must</i> - voz ativa	Número de ocorrências	Colocados à esquerda <i>Must</i> -voz passiva	Número de ocorrências
<i>You</i> (L1)	159	<i>You</i> (L1)	-----
Anáforas (L1)	05	Anáforas (L1)	01
Unidades simples (L1)	41	Unidades simples (L1)	19
Sintagmas nominais	27	Sintagmas nominais	32
Sint. verbais e prepos.	18	Sint. verbais e prepos.	21

De uma forma geral, o grupo *you*, com 159 ocorrências, é o mais freqüente na posição de sujeito, empregado quase três vezes mais do que o grupo de sintagmas nominais com 59 ocorrências (voz ativa + voz passiva juntas). O grupo anáforas apresenta apenas seis ocorrências ao total, não constituindo, dessa forma, um grupo representativo.

Item B: Colocados à direita

Mostro agora os colocados à direita de *must*.

Tabela 7.3.1.5.C: Colocados à direita (R1, R2, R3) de *must*- no MM

Posição	Colocado <i>must</i> voz ativa	Quantidade	Relação LL	Colocado <i>must</i> voz passiva	Quantidade	Relação LL
1	BE (R1)	44	338,691	DONE (R1)	10	111,229
2	INDIVIDUAL (R3)	21	237,345	REMOVED (R1)	06	36,280
3	IDENTIFY (R2)	22	207,736	BEFORE (R2)	06	27,263
4	EACH (R2)	23	170,256			
5	NOT (R1)	27	113,888			
6	HAVE (R1)	14	90,972			
7	MORE (R3)	08	65,671			
8	DO (R1)	16	59,308			
9	FULLY (R1)	05	56,876			
10	USE (R1)	15	56,249			

Na tabela acima se observa os principais colocados de *must* no MM, voz ativa (250 ocorrências) e voz passiva (73 ocorrências). Entre os dez primeiros colocados de *must* na voz ativa conforme a relação LL destaca-se o verbo *be*, na posição R1, com 44 casos. Destaca-se a presença dos advérbios *not* e *fully* na posição R1 entre os principais colocados, pois esta posição é mais frequentemente ocupada por verbos. Na próxima tabela mostro os verbos que ocupam as posições à esquerda classificados em sete grupos semânticos:

Tabela 7.3.1.5.D: Verbos na posição R1 de *must*-MM

Tipos de Verbos	<i>Must</i> -voz ativa	<i>Must</i> -voz passiva
De atividade	133/ 53,2 %	70/ 95,9%
De comunicação	-----	-----
Mentais	39/ 15,6%	03/ 4,1%
Causativos	06/ 2,4%	-----
De ocorrência	06/ 2,4%	-----
De existência	65/ 26%	-----
De aspecto	01/ 0,4%	-----

Nota-se que a maior parte das ocorrências de *must* é acompanhada por um verbo de atividade. Isso acontece em 63 % dos casos. O grupo existência ocupa o segundo lugar (voz ativa) devido, principalmente, a alta ocorrência de *be*¹³⁶.

C Usos e funções

Foi constatado que:

- Todas as ocorrências de *must* expressam a idéia de necessidade de que algo seja feito.
- *Must* é usado com bastante mais freqüência na voz ativa, na proporção aproximada de uma ocorrência na voz passiva a cada três na voz ativa. Considerando os casos em que *must* é usado com verbos de atividade, essa proporção diminui para um caso na voz passiva para cada dois na voz ativa.
- *Must* é usado principalmente em contextos positivos, na proporção de uma ocorrência negativa a cada nove positivas. Entretanto, o *not* é considerado um colocado forte de *must*.
- *Must* é principalmente acompanhado de verbos de atividade, de existência e mentais.
- *Must* ocorre freqüentemente em orações condicionais com a presença da conjunção *if* (18 ocorrências), principalmente na oração subordinada de uma condicional (11 ocorrências). Entretanto, *must* também aparece na oração principal (7 ocorrências).
- A maioria dos casos em que *you* é o sujeito, o verbo é de atividade, em 125 ocorrências. Há mais dois casos em que o sujeito é *persons*. Todas as outras ocorrências de *must* possuem um sujeito inanimado.

Com base nos resultados apresentados nos itens A, B e C, tomo os seguintes exemplos para ilustrar as características e usos mais salientes de *must* nesse manual, o MM.

1- *You must move fuel from a different tank if you **must** refuel the center tank when its indicators do not operate.*

2- *THE WATER PRESSURE **MUST NOT BE MORE THAN 55 PSI.***

3- *Cargo restraining equipment **must** be replaced if damaged by mercury.*

¹³⁶ Lembrando que *be* não é considerado colocado dos VM na voz passiva por ser elemento necessário à constituição da estrutura.

4- *The flight safety of the airplane, together with the safety of health and the prevention of bad smells, **must** be your first consideration when you complete this task.*

7.3.1.6 Shall

No MM, o modal *shall*, apresenta um total de 18 ocorrências, sendo:

- 15 ocorrências na estrutura 2, voz ativa
- 3 ocorrências na estrutura 3, voz passiva

A Colocados à esquerda

Devido à baixa quantidade de *shall* no MM, esse não apresenta colocados à esquerda conforme os critérios estabelecidos neste trabalho (mínimo de quatro ocorrências em um horizonte de três palavras à esquerda e três à direita e um LL de 6,63 ou maior). Todavia, é possível classificar os itens lexicais à esquerda nas cinco categorias. Vide tabela abaixo:

Tabela 7.3.1.6.A: Classificação do tipo de unidades lexicais à esquerda de *shall* -MM

Colocados à esquerda <i>Shall</i> - voz ativa	Número de ocorrências	Colocados à esquerda <i>Shall</i> -voz passiva	Número de ocorrências
<i>You</i> (L1)	-----	<i>You</i> (L1)	-----
Anáforas (L1)	-----	Anáforas (L1)	02/ 67%
Unidades simples (L1)	-----	Unidades simples (L1)	-----
Sintagmas nominais	05/ 33%	Sintagmas nominais	-----
Sint. verbais e prepos.	10/ 67%	Sint. verbais e prepos.	01/ 33%

De uma forma geral, o grupo composto por sintagmas verbais e preposicionais é o mais freqüente, com 11 ocorrências.

B Colocados à direita

Shall apresenta apenas um colocado à direita, o verbo *be*, com oito ocorrências na posição R1 (voz passiva) e um LL de 73,290. Apesar da baixa ocorrência, classifico os casos entre as categorias verbais:

Tabela 7.3.1.6.B: Verbos na posição R1 de *shall* (voz ativa)-MM

Tipos de Verbos	Shall-voz ativa	Shall-voz passiva
De atividade	05/ 31,2%	-----
De comunicação	-----	-----
Mentais	-----	01/ 33%
Causativos	-----	-----
De ocorrência	-----	-----
De existência	9/ 62,5%	02/67%
De aspecto	01/ 6,3%	-----

Nota-se que a maior parte das ocorrências de *shall* é acompanhada por um verbo de existência, ocorrendo em 12 entre o total de 18 casos.

C Usos e funções

Foi constatado que:

- *Shall* no MM expressa a idéia de necessidade ou obrigação.
- *Shall* é acompanhado à esquerda principalmente por sintagmas verbais e preposicionais.
- *Shall* é principalmente acompanhado de verbos de atividade e de existência.

Com base nos resultados apresentados nos itens A, B e C, sugiro o seguinte exemplo para ilustrar as características e usos mais salientes de *shall* nesse manual, o MM.

1- For multiple fasteners spaced less than four inches apart, the maximum number which can be safety wired together shall be the number than can be wired with a 24-inch length of wire.

7.3.1.7 Should

No MM, o modal *should* apresenta um total de 66 ocorrências, sendo:

- 32 ocorrências na estrutura 2, voz ativa
- 34 ocorrências na estrutura 3, voz passiva

A Colocados à esquerda

O quadro abaixo mostra os principais colocados à esquerda juntamente com a quantidade de ocorrências:

Tabela 7.3.1.7.A: Colocados à esquerda (L1, L2, L3) de *should* no MM

Posição	Colocado <i>should</i> voz ativa	Quantidade	Relação LL	Colocado <i>should</i> voz passiva	Quantidade	Relação LL
1	BLUE (L2)	04	31,767	ENGINES (L1)	06	59,240
2	WATER (L1)	04	22,599	OIL (L2)	04	22,673

Não há destaque quanto aos colocados. Há apenas quatro, dois para a voz ativa e dois para a voz passiva. E apenas dois na posição L1. Investigo agora todos os itens lexicais nas posições à esquerda do nóculo *should*. Para o nóculo *should*, foi necessário incluir outra categoria para classificar esses itens lexicais, conforme tabela abaixo:

7.3.1.7.B: Classificação das unidades lexicais à esquerda de *should* no MM

Colocados à esquerda <i>Should</i> - voz ativa	Número de ocorrências	Colocados à esquerda <i>Should</i> - voz passiva	Número de ocorrências
<i>You</i> (L1)	01/ 3,1%	<i>You</i> (L1)	-----
Anáforas (L1)	03/ 9,4%	Anáforas (L1)	
Unidades simples (L1)	04/ 12,5%	Unidades simples (L1)	11/32,4%
Sintagmas nominais	8/ 25%	Sintagmas nominais	16/ 47%
Sintag. Verbais e prepos.	13/ 40,6%	Sintag. Verbais e prepos.	07/ 20,6%
Início da oração	03/ 9,4%		

À exemplo da tabela dos colocados, não há uma grande predominância de um ou outro grupo. Os destaques são os sintagmas, tanto os nominais quanto os verbais e preposicionais, na posição de sujeito. Esses respondem por 66% das ocorrências. Outro destaque é *should* sem um sujeito, que acontece em três casos. Esses casos são aqueles nos quais *should* é usado como condicional, no início da oração.

B Colocados à direita

Mostro agora os contextos com base nos colocados à direita de *should*.

Tabela 7.3.1.7.C: Colocados à direita (R1, R2, R3) de *should* no MM

Pos.	Colocados <i>should</i> voz ativa	Quantidade	Relação LL	Colocados <i>should</i> voz passiva	Quantidade	Relação LL
1	BE (R1)	09	57,912	OPERATED (R2)	07	84,626
2	-----	-----	-----	NOT (R1)	08	39,683
3	-----	-----	-----	USED (R1)	05	35,663

Nesta lista dos principais colocados de *should* o verbo que mais se destaca é o *be* na voz ativa (lembrando que na voz passiva *be* é considerado parte da estrutura). O *be* sozinho responde por aproximadamente 30 % entre os acompanhantes de *should* nessa estrutura. Entre os casos de *be* na voz passiva, há dois verbos, *operated* e *used* e o advérbio *not*, com oito ocorrências.

Tabela 7.3.1.7.D: Verbos na posição R1 de *should* no MM

Tipos de Verbos	<i>Should</i> -voz ativa	<i>Should</i> -voz passiva
De atividade	14/ 43,8%	33/ 93%
De comunicação	-----	-----
Mentais	-----	-----
Causativos	02/ 6,3%	01/ 3%
De ocorrência	03/ 9,4%	-----
De existência	11/ 34,2	-----
De aspecto	02/ 6,3%	-----

Nos casos de voz passiva, os verbos que mais freqüentemente acompanham *should* são os de atividade. Na voz passiva eles também são mais numerosos, mas há uma distribuição um pouco mais ampla entre os outros tipos de verbos, com os de existência em segundo lugar.

C Usos e funções

Foi constatado que:

- Todos os casos de *should* na voz passiva expressam uma instrução ou comando, ou seja, possuem um uso deôntico.
- Na voz ativa, *should* expressa três funções diferentes:

(1) Como instrução, isto é, uso deôntico, em 20 ocorrências. É importante ressaltar que esse uso de *should* não é o de apenas recomendar, ou sugerir uma ação, mas sim o de estabelecer a única regra a ser seguida. Ainda sobre esse uso, o sujeito é, na maioria das vezes, inanimado com um verbo intransitivo. Dessa forma, esse uso fica muito próximo do uso como probabilidade (que será explicado a seguir). É, assim, difícil interpretar sua real função. Como por exemplo:

*Fueling procedures **should** stop if strong wind conditions are present.*

(2) Como probabilidade, isto é, expressando uma suposição, uma necessidade lógica atenuada. Esse uso ocorre nove vezes:

*N_O_T_E_: Most of the blue water **should** drain out of the airplane through the floor drains and the waste water system.*

(3) Em três situações, ocupando o lugar do *if* em uma condicional¹³⁷.

Entretanto, dicionários e livros texto, por exemplo por exemplo:

***Should** operators elect not to follow this recommended procedure, they should develop alternate procedures or establish conditions adequately protecting the personnel and equipment involved.*

Nota-se que há duas ocorrências de *should*, e cada uma das ocorrências expressa algo diferente: a primeira uma condição e a segunda uma instrução.

¹³⁷ Nesses casos *should* muda sua classe gramatical, passando a ser considerado como conjunção. Entretanto, essa mudança de uso não é geralmente retratada nas gramáticas ou livros, como por exemplo, Shawcross (1992). A opção de incluir esse uso de *should* nesta tese deve-se ao fato de que estou aqui considerando a palavra ortográfica como item lexical pesquisado, uma vez que é a palavra que será lida pelo destinatário do manual.

- *Should* é, na maioria das vezes, seguido por um verbo de atividade. Os sujeitos são principalmente sintagmas nominais, verbais ou preposicionais.

Com base nos resultados apresentados nos itens A, B e C, sugiro os seguintes exemplos para ilustrar as características e usos mais salientes de *should* neste manual, o MM.

- 1- *The ambient temperatures at time of measurements are taken **should be** within 20°F when you do this step.*
- 2- *N_O_T_E_: A record of a blue stain area can help to find the cause of a possible new leak of blue water, **should** a leak occur in the future.*
- 3- *Valves with P/N 10-60704-1 thru 10-60704-6 **should** close at differential pressures of 2.0 to 2.8 psi.*
- 4- *ENGINES **SHOULD** NOT BE OPERATED WITH COVERS IN PLACE BECAUSE THE COVERS CAN COME OFF AND DAMAGE THE ENGINES.*

7.3.1.8 Will

No MM, o modal *will* apresenta um total de 355 ocorrências, sendo:

- 301 ocorrências na estrutura 2, voz ativa
- 53 ocorrências na estrutura 3, voz passiva
- Uma ocorrência na estrutura 5, perfeito (positiva)

A Colocados à esquerda

O quadro abaixo mostra os 10 principais colocados à esquerda juntamente com a quantidade de ocorrências:

Tabela 7.3.1.8.A: Colocados à esquerda (L1, L2, L3...) de *will* no MM

Posição	Coloc. voz ativa	Quantidade	LL	Coloc.voz passiva	Quantidade	LL
1	YOU (L1)	52	327,104	WHICH (L1)	16	97,451
2	THIS (L1)	36	193,728	SURFACES (L2)	12	64,534
3	THAT (L2)	33	163,299	FROM (L3)	12	43,987
4	WHERE (L2)	20	163,078	REVISION (L1)	04	24,268
5	OIL (L1)	13	97,646	-----	-----	-----
6	WHICH (L1)	13	81,387	-----	-----	-----
7	IT (L1)	10	57,359	-----	-----	-----
8	SURE (L3)	07	47,342	-----	-----	-----
9	SYSTEMS (L3)	06	38,728	-----	-----	-----
10	GROMMET (L1)	05	36,705	-----	-----	-----

Destaca-se *you* como colocado de *will*. É também notória a frequência de elementos anafóricos nas posições L1 e também L2, tanto nos casos de voz ativa quanto nos casos de voz passiva.

Tabela 7.3.1.8.B: Classificação das unidades lexicais à esquerda de *will* no MM

Colocados à esquerda <i>Will</i> - voz ativa	Número de ocorrências	Colocados à esquerda <i>Will</i> -voz passiva	Número de ocorrências
<i>You</i> (L1)	52/ 17,3%	<i>You</i> (L1)	-----
Anáforas (L1)	52/ 17,3%	Anáforas (L1)	17/ 32,1%
Unidades simples (L1)	65/ 21,5%	Unidades simples (L1)	7/ 13,2%
Sintagmas nominais	60/ 20,0%	Sintagmas nominais	19/ 35,8%
Sintag. Verbais e prepos.	72/ 23,9%	Sintag. Verbais e prepos.	10/ 18,9%

Como pode ser visto, o *you* sozinho responde por 17,3% das ocorrências de *will* na voz ativa no MM, exatamente o mesmo número de ocorrências que o grupo dos elementos anafóricos. Entretanto, somando todas as ocorrências de *will* na voz ativa e na voz passiva, nota-se que o grupo mais freqüente é o de sintagmas verbais e preposicionais (82), seguido pelo grupo de sintagmas nominais (79). O grupo de apenas um substantivo é também representativo (72).

B Colocados à direita

Mostro agora os colocados à direita de *will*.

Tabela 7.3.1.8.C: Colocados à direita (R1, R2, R3) de *will* no MM

Posição	Colocados <i>will</i> voz ativa	Quantidade	Relação LL	Colocados <i>will</i> voz passiva	Quantidade	Relação LL
1	NOT (R1)	32	146,600	LUBRICATED (R1)	12	119,553
2	FLOW (R3)	15	134,941	IDENTIFIED (R1)	04	24,745
3	BE (R1)	24	127,501	USED (R1)	04	13,498
4	OIL (R2)	10	97,646	-----	-----	-----
5	DECREASE (R1)	10	91,713	-----	-----	-----
6	HAVE (R1)	14	88,220	-----	-----	-----
7	PREVENT (R1)	11	87,141	-----	-----	-----
8	WHEN (R3)	16	84,203	-----	-----	-----
9	CAUSE (R1)	16	83,509	-----	-----	-----
10	LET (R1)	10	65,231	-----	-----	-----

Nesta lista dos principais colocados de *will* à direita, nota-se que, além de verbos, a presença do advérbio *not* na posição R1 como o colocado mais significativo de *will*. Destaco o cópula *be*, entre os colocados de *can*+voz ativa, sendo o terceiro verbo com maior atração a *can*.

Tabela 7.3.1.8.D: Verbos na posição R1 de *will* -MM

Tipos de Verbos	<i>Will</i> -voz ativa	<i>Will</i> -voz passiva
De atividade	141/ 46,8%	39/ 73,5%
De comunicação	2/ 0,7%	-----
Mentais	16/5,3%	7/ 13,2%
Causativos	60/ 20%	-----
De ocorrência	28/ 9,3%	01/ 1,9%
De existência	44/ 14,6%	5/ 9,5%
De aspecto	10/ 3,3%	01/ 1,9%

É notória a preferência que *will* apresenta pelos verbos de atividade, com os causativos em segundo lugar, seguidos pelos de existência, em terceiro.

C Usos e funções

Will na estrutura 5 (modal+perfeito) apresenta apenas uma ocorrência:

N_O_T_E_: *A heat-resistant coating is used on the brake mounting flanges, on the axles in the areas of brake mounting flanges and on the sleeves. The hydraulic fluid-resistant paint finish **will** have burned away in most cases.*

Com base nos itens A, B e C, e na ocorrência acima, foi constatado que:

- *Will* é usado com bastante mais freqüência na voz ativa, na proporção de uma ocorrência na voz passiva a cada seis na voz ativa. É usado principalmente na forma afirmativa, na proporção de uma ocorrência negativa a cada nove afirmativas.
- As ocorrências de *will* na voz passiva expressam a idéia de previsão futura, indicando que uma ação/conseqüência sempre ou tipicamente acontecerá.
- A maior parte das ocorrências de *will* na voz ativa também expressa uma previsão. Entretanto, *will* é também usado como quase imperativo, expressando instruções. Esse uso ocorre 12 vezes, como por exemplo:

*YOU **WILL** MOVE THE SEAL PLATES IN AN OUTWARD DIRECTION.*

- *Will* ocorre em orações condicionais com a presença da conjunção *if* tanto na oração principal de uma condicional (8 ocorrências) quanto na subordinada (7 ocorrências). Esse uso de *will* na oração subordinada é raro e considerado não-gramatical na maioria dos materiais didáticos quando tratam de condicionais, como por exemplo:

*If you **will** paint the adjacent area, leave the premask in position.* □

Após analisar todas as ocorrências de *will* nessa posição, foi possível aventar a hipótese de que, nesses casos, parece haver uma insistência em realizar uma ação mesmo tendo sido sugerido que a ação implica riscos, ou que não precisa ser realizada, como é o caso do exemplo acima.

Quando o sujeito é inanimado, esse uso de *will* na oração subordinada, parece sugerir que a ação expressa pelo verbo lexical que segue o *will* não era prevista de acontecer, mas se ela “insistir” em acontecer, realize a ação sugerida na oração principal. Vide exemplo abaixo:

Do the steps that follow if the cabin temperature will decrease below the freezing point.

Em síntese, esse uso parece expressar um fato “contrário à expectativa”.

- Quando *will* é seguido por verbos causativos ou de ocorrência, ele anuncia uma consequência negativa fruto de uma ação mal realizada pelo mecânico, isto é, o VM possui uma prosódia semântica negativa.

Com base nos resultados apresentados nos itens A, B e C, mostro os seguintes exemplos para ilustrar as características e usos mais salientes de *will* nesse manual, o MM.

- 1- *C_A_U_T_I_O_N_: DO NOT PUT TOO MUCH OIL INTO THE GENERATOR DRIVE. TOO MUCH OIL **WILL** CAUSE THE GENERATOR DRIVE TO BECOME TOO HOT AND CAUSE DAMAGE TO THE GENERATOR DRIVE.*
- 2- *If you **will** clean the flaps, extend them to the fully down position.*
- 3- *A fill valve in the open position **will** not prevent high pressure in the pneumatic system when the tank is full.*
- 4- *The extra oil **will** be blown overboard through the engine vent system.*
- 5- *You **will** use an adhesive to get a smooth surface that will give the placard a strong bond.*

7.3.1.9 *Would*

No MM, o modal *would*, apresenta apenas as três ocorrências .

- duas ocorrências na estrutura 2, voz ativa

(b) *If a task applies to several areas in a zone, then the location zone may be listed with a 00 zone. For example, a procedure on passenger seats **would** have the following zone listed:*

*G. Install and twist the safety wire so the loop around the head stays down. (If the loop tended to come up over the bolt head there **would** be a slack loop.)*

- uma ocorrência na estrutura 3, voz passiva

*Example: A step involving hydraulic tubing and electrical wiring are separate sub-tasks and **would** be structured as two separate steps.*

A princípio, *would* parece expressar uma hipótese irreal. Entretanto, devido à baixíssima frequência de *would* nesse texto e à característica de um manual técnico, parece-me pouco provável que se trate de uma hipótese irreal. Assim, acredito que a elocução tenha sido usada para expressar uma possibilidade ou uma consequência e que *would* poderia ter sido substituído pelo VM *may*, ou pelo modal *will* (no segundo exemplo). Para o segundo exemplo de *would* na voz ativa, poderia ter sido escolhido a primeira condicional e o *would* seria substituído por *will*. Vide sugestão abaixo:

*G. Install and twist the safety wire so the loop around the head stays down. (If the loop **comes** up over the bolt head there **will** be a slack loop.)*

7.3.2 Operations Manual (OM)

O OM é o segundo subcorpus mais extenso do AC, com 183.959 *tokens*. Os destinatários deste manual são os pilotos. Esse manual é usado principalmente em solo quando o piloto está em algum tipo de treinamento. Entretanto, a presença do OM a bordo é obrigatória para os casos em que os pilotos necessitam resolver um problema não referido no QRH. Passo agora aos resultados.

7.3.2.1 *Can*

No OM, o modal *can*, que é o modal mais freqüente apresentando um total de 327 ocorrências, sendo:

- 85 ocorrências na estrutura 2, voz ativa

- 242 ocorrências na estrutura 3, voz passiva

A Colocados à esquerda

O quadro abaixo mostra os 10 principais colocados esquerda juntamente com a quantidade de ocorrências:

Tabela 7.3.2.1.A: Colocados á esquerda (L1, L2, L3...) de *can* no OM

Posição	Colocado <i>can</i> voz ativa	Quantidade	Relação LL	Colocado <i>can</i> voz passiva	Quantidade	Relação LL
1	IT (L1)	4	19,428	BUT (L2)	4	41,343
2	FMC (L3)	5	8,331	MODE (L1)	12	36,833
3	APU (L1)	4	7,501	THRUST (L2)	10	36,460
4	-----	-----	-----	MODES (L1)	7	36,418
5	-----	-----	-----	IT (L1)	6	34,049
6	-----	-----	-----	WAYPOINTS (L1)	6	31,782
7	-----	-----	-----	ENTRY (L1)	8	31,397
8	-----	-----	-----	THIS (L2)	7	30,750
9	-----	-----	-----	INFORMATION (L1)	7	30,357
10	-----	-----	-----	EXECUTED (L2)	4	28,591

Não há um item lexical que apresente uma preferência que se destaque. Nota-se, por outro lado, a ausência do pronome *you* e a baixa ocorrência de elementos anafóricos.

Tabela 7.3.2.1.B: Classificação das unidades lexicais à esquerda de *can* no OM

Colocados a esquerda <i>Can</i> - voz ativa	Número de ocorrências	Colocados a esquerda <i>Can</i> -voz passiva	Número de ocorrências
<i>You</i> (L1)	-----	<i>You</i> (L1)	-----
Anáforas (L1)	04/ 4,7%	Anáforas (L1)	13/ 5,4%
Simples (L1)	21/ 24,7%	Simples (L1)	54/ 22,31%
Sintagmas nominais	30/ 35,3%	Sintagmas nominais	126/ 52,1%
Sintag. Verbais e prepos.	30/ 35,3%	Sintag. Verbais e prepos.	49/ 20,2%

Como pode ser visto, o grupo dos sintagmas nominais é preferido por *can* na voz ativa e empata com o grupo de sintagmas verbais e preposicionais na voz passiva. Os elementos anafóricos apresentam poucas ocorrências nas duas estruturas.

B Colocados à direita

Mostro agora os colocados à direita de *can*.

Tabela 7.3.2.1.C: Colocados à direita (R1, R2, R3) de *can* no OM

Posição	Colocado <i>can</i> voz ativa	Quantidade	Relação LL	Colocado <i>can</i> voz passiva	Quantidade	Relação LL
1	CAUSE (R1)	8	61,130	SELECTED (R1)	20	128,946
2	INCLUDE (R1)	7	58,679	BY (R2)	34	125,537
3	SUPPLY (R1)	7	49,692	USED (R1)	17	108,128
4	ALSO (R1)	7	39,925	REENGAGED (R1)	5	59,708
5	POWER (R2)	7	32,594	ALSO (R1)	9	56,011
6	OPERATE (R1)	5	32,329	FROM (R2)	12	55,289
7	OCCUR (R1)	5	28,880	OPENED (R1)	5	54,536
8	PROVIDE (R1)	5	28,313	ON (R2)	19	51,752
9	EITHER (R2)	4	20,890	MADE (R1)	7	49,970
10	BE (R1)	7	17,651	ENGAGED (R1)	6	47,499

Nesta lista dos principais colocados de *can* à direita, nota-se que os verbos que ocorrem na voz ativa não se repetem na voz passiva, sendo o *cause* o mais freqüente na voz ativa e *selected* na passiva. A preposição *by* é um colocado forte de *can* (voz passiva), ocorrendo em aproximadamente 11% dos casos. O advérbio *also* é o único item lexical que aparece nas duas estruturas.

Tabela 7.3.2.1.D: Verbos na posição R1 de *can* -OM

Tipos de Verbos	<i>Can</i> -voz ativa	<i>Can</i> -voz passiva
De atividade	52/ 61%	181/ 74,8%
De comunicação	2/ 2,4%	1/ 0,4%
Mentais	1/ 1,2%	17/7%
Causativos	12/ 14,1%	1/ 0,4%
De ocorrência	1/ 1,2%	15/ 6,2%
De existência	16/ 18,9%	4/ 1,7%
De aspecto	1/ 1,2%	23/ 9,0%

Há uma predominância dos verbos de atividade entre os colocados (R1) de *can* (voz ativa e voz passiva). Entretanto, cabe mencionar que não há um único item lexical que domine a maioria das ocorrências.

Na observação manual das linhas de concordância, foi possível verificar que há considerável incidência de advérbios na posição R1, tanto na voz ativa quanto na voz passiva. Na voz ativa, esses advérbios (14 ao todo, *automatically, still, only, manually e also*) ocorrem logo após o VM. Na voz passiva eles podem ocorrer entre o nódulo e o auxiliar *be* (15 ocorrências, *also, only, then therefore*) como entre o nódulo e o auxiliar *be* (nove ocorrências, *automatically, manually, mechanically*).

C Usos e funções

Foi constatado que:

- *Can* é usado com bastante mais frequência na voz passiva, na proporção de uma ocorrência na voz ativa a cada três na voz passiva. É usado principalmente na forma afirmativa, na proporção de uma ocorrência negativa a cada cinco afirmativas.
- Todas as ocorrências de *can* positivos expressam possibilidade.
- *Can*, quando acompanhado por um verbo causativo, ou de ocorrência expressa uma consequência negativa que pode acontecer devido a uma ação mal realizada pelo piloto ou por uma mal funcionamento da aeronave.
- A função do *by* nas ocorrências de *can* na voz passiva, indicar um agente animado, neste caso, pilotos (*flight crew*) em apenas quatro casos, ou 12%. Nos outros casos, o *by* indica o procedimento que pode gerar a ação já descrita à esquerda do *can*.

Com base nos resultados apresentados nos itens A, B e C, ilustro as características e usos mais salientes de *can* nesse manual, o OM.

*1-The engine indications **can** also be manually selected to either the Captain's or First Officer's inboard DU, or the lower DU, using the respective display select panel.*

*2- Air for the bleed air system **can** be supplied by the engines, APU, or an external air cart/source..*

*3- After localizer and glideslope capture during a dual autopilot approach, CWS **cannot** be engaged by manually overriding pitch and roll..*

*4- A spinning tire with a loose tread must be stopped prior to entering the wheel well or it can **cause** damage to wheel well components.*

7.3.2.2 *Could*

No OM, o modal *could* apresenta um total de 12 ocorrências, sendo:

- 11 ocorrências na estrutura 2, voz ativa
- uma ocorrência na estrutura 3, voz passiva

A Colocados à esquerda

Na voz ativa, *could* apresenta três colocados ao total. Como o *could* na voz passiva apresenta somente uma ocorrência, ela não é analisada seguindo todas as etapas metodológicas.

O quadro abaixo mostra os colocados à esquerda juntamente com a quantidade de ocorrências:

Tabela 7.3.2.2.A: Colocados à esquerda de *could* no OM

Posição	Colocados voz ativa	Quantidade	Relação LL
1	SUBSTANTIAL (L2)	4	78,613
2	DAMAGE (L1)	4	60,064

Devido à baixa frequência de *could*, não há dez itens lexicais que preencham às exigências (mínimo de quatro ocorrências e relação LL de 6,6) para serem incluídos na lista de colocados. Nota-se que o adjetivo *substantial* e o substantivo *damage* são os únicos colocados à esquerda.

Na próxima tabela mostro a classificação dos colocados à esquerda:

Tabela 7.3.2.2.B: Unidades à esquerda de *could*-voz ativa-OM

Colocados à esquerda	Número de ocorrências
<i>You</i> (L1)	-----
Anáforas (L1)	1/ 9,1%
Unidades simples (L1)	1/ 9,1%
Sintagmas nominais	7/ 63,4%
Sintag. verbais e preposic.	2/ 18,2%

Como é possível observar, o grupo com mais ocorrências é aquele representado pelos sintagmas nominais. Saliento a ausência do pronome *you*.

B Colocados à direita

Descrevo agora os colocados à direita de *could*.

Tabela 7.3.2.2.C: Colocados à esquerda de *could* no OM

Posição	Colocados <i>Could</i> voz ativa	Quantidade	Relação LL
1	RESULT (R1)	4	53,739

Há apenas um colocado, o verbo *result*. Na próxima tabela classifiquemos os outros verbos à direita de *could*.

Tabela 7.3.2.2.D: Verbos na posição R1 de *could* do OM

Tipos de Verbos	<i>Could</i> -voz ativa
De atividade	-----
De comunicação	-----
Mentais	-----
Causativos	2
De ocorrência	7
De existência	2
De aspecto	----

Há uma predominância dos verbos de ocorrência simples entre os elementos verbos lexicais na posição (R1) de *could*.

Item C: Usos e funções

Como a ocorrência na voz passiva não pode ser analisada conforme o roteiro metodológico acima, apresento-a a seguir:

*The desired lateral offset distance is entered on line 2L In the example, the 10.0 nm offset left of course **could** be entered L10.0, L10, 10.0L, or 10L.*

Com base nessa ocorrência e nos itens A e B, pode-se concluir que:

- Todas as ocorrências de *could* expressam a idéia de possibilidade.
- *Could*, acompanhado por um verbo “causativo”, ou “de ocorrência”, é, a exemplo de *can*, utilizado para expressar uma consequência negativa que pode acontecer devido a uma ação mal realizada pelo piloto. Esses casos representam 75% das ocorrências de *could* no OM. Assim, pode-se concluir que *could*, na maioria dos casos, apresenta uma prosódia semântica negativa.
- *Could* é usado basicamente na voz ativa, na proporção de uma ocorrência na voz passiva a cada onze na voz ativa.
- *Could* é usado somente na forma afirmativa.

- *Could* tem como principais colocados à esquerda, as “anáforas” e “Unidades simples”.

Dessa forma, acredito que s seguinte exemplo representa essas características:

1- CAUTION: Airstair handrail extensions must be stowed or substantial damage could result.

7.3.2.3 May

No OM, o modal *may* apresenta um total de 288 ocorrências, sendo:

- 115 ocorrências na estrutura 2, voz ativa
- 171 ocorrências na estrutura 3, voz passiva
- Uma ocorrência na estrutura 4, progressivo
- Uma ocorrência na estrutura 5, perfeito

Por apresentar baixa frequência, *may* nas estruturas 4 e 5 é apresentado somente no final da seção.

A Colocados à esquerda

A tabela abaixo mostra os 10 principais colocados à esquerda juntamente com a quantidade de ocorrências:

Tabela 7.3.2.3.A: Colocados à esquerda (L1, L2, L3) de *may* no OM

Posição	Colocados voz ativa	Quantidade	LL	Colocados voz passiva	Quantidade	LL
1	CYCLE (L1)	4	27,705	ITEM (L1)	18	212,843
2	SHIFT (L1)	4	26,263	ON (L3)	19	83,056
3	NOTE (L2)	5	22,484	MANUALLY (L3)	4	54,394
4	HOLDING (L2)	4	17,983	ENTRY (L1)	7	25,100
5	ROLL (L2)	4	17,146	LNAV (L1)	5	23,717
6	OIL (L3)	4	16,162	ENTRIES (L1)	4	19,630
7	CREW (L1)	4	15,145	DESCENT (L1)	6	15,712
8	THAT (L1)	4	14,663	DATA (L1)	6	15,017
9	MAP (L2)	5	13,809	LANDING (L2)	5	12,739
10	PRESSURE (L1)	4	13,594	CRUISE (L3)	4	11,418

Não há destaques com relação aos colocados à esquerda na voz ativa. Na posição L1, os colocados mais fortes são *cycle e shift*, ambos com quatro ocorrências. Entretanto, *item* destaca-se na posição L1, com 18 ocorrências, o que corresponde a mais de 10%. Mostro a classificação dos colocados à esquerda de *may* na tabela abaixo.

Tabela 7.3.2.3.B Classificação do tipo de colocados de *may* nas posições L-OM

Colocados à esquerda <i>May</i> - voz ativa	Número de ocorrências	Colocados à esquerda <i>May</i> -voz passiva	Número de ocorrências
<i>You</i> (L1)		<i>You</i> (L1)	-----
Anáforas (L1)	13/ 11,3 %	Anáforas (L1)	4/ 2,3 %
Unidades simples (L1)	16/13,9 %	Unidades simples (L1)	50/29,2 %
Sintagmas nominais	53/46,1%	Sintagmas nominais	69/ 40 %
Sintag. verbais e prep.	33/28,7 %	Sintag. verbais e prep.	48/ 28,5 %

O grupo mais freqüente é o composto de sintagmas nominais, tanto nos casos de voz ativa como nos casos de voz passiva. Em segundo lugar aparece o grupo de sintagmas verbais e preposicionais na voz ativa e pelo grupo composto por unidades simples na voz passiva.

B Colocados à direita

Descrevo agora os colocados à direita de *may*.

Tabela 7.3.2.3.C: Colocados à direita (R1, R2, R3) de *may* no OM

Posição	Colocado <i>may</i> voz ativa	Quantidade	Relação LL	Colocado <i>may</i> voz passiva	Quantidade	Relação LL
1	RESULT (R1)	10	87,739	FOUND (R1)	19	211,607
2	CAUSE (R1)	8	60,445	ENTERED (R1)	13	66,056
3	OCCUR (R1)	7	58,618	MADE (R1)	8	58,734
4	EXCEED (R1)	5	46,461	MANUALLY (R2)	7	54,394
5	NOT (R1)	11	36,046	ACCOMPLISHED (R1)	7	52,160
6	TAKE (R1)	4	35,711	USED (R1)	10	51,734
7	ALSO (R1)	6	31,833	BY (R2)	14	49,084
8	INCREASE (R1)	4	25,837	ALSO (R1)	8	47,151
9	TO (R3)	14	20,660	SELECTED (R1)	12	43,197
10	IN (R2)	10	17,173	OPERATED (R1)	4	38,477

O verbo *result* apresenta a maior força de colocação com *may* na voz ativa, enquanto *found* é o principal colocado de *may* na voz passiva. Como já havia ocorrido com as colocações dos outros VMs que expressam possibilidade, *cause* e *occur* estão entre os principais colocados na voz ativa. O advérbio *also* aparece em ambas as estruturas.

7.3.2.3.D: Verbos na posição R1 de *may* (voz ativa)-MM

Tipos de Verbos	<i>May</i> -voz ativa	<i>May</i> -voz passiva
De atividade	39/ 33,9	113/ 66,1
De comunicação	-----	7/ 4,1
Mentais	01/0,9	31/ 18,1
Causativos	10/8,7	10/ 5,8
De ocorrência	28/ 24,3	1/ 0,6
De existência	35/ 30,4	-----
De aspecto	2/ 1,8	9/ 5,3

C Usos e funções

Apresento as duas ocorrências de *may* nas estruturas 4 e 5:

Although the F/D, A/P and A/T may be performing as previously described, severe windshear may exceed the performance capability of the system and/or the airplane.

Foi constatado que:

- *May* é usado com mais frequência na voz passiva, na proporção de uma ocorrência na voz ativa a cada 1,5 na voz passiva. É usado principalmente em contextos positivos, na proporção de uma ocorrência negativa a cada 19 positivas.
- Todas as ocorrências de *may* expressam a idéia de possibilidade, podendo essa possibilidade ser relacionada a contextos neutros ou contextos negativos, que serão explicados a seguir.
- *May*, quando acompanhado por um verbo “causativo”, ou “de ocorrência” apresenta uma prosódia semântica negativa, ou seja, as probabilidades expressas nesses casos são negativas.
- *May* acompanhado de outros tipos de verbos não apresenta a característica de expressar uma possibilidade negativa. Sua função geralmente expressa apenas uma possibilidade mais neutra, mostrando possíveis ações a serem realizadas.

- *May* é geralmente precedido por um sintagma (nominal, verbal ou preposicional 53% das ocorrências) e seguido por um verbo de atividade (71% das ocorrências).

Com base nos resultados apresentados nos itens A, B e C, ilustrar as características e usos mais salientes de *may* no MM através dos seguintes exemplos:

1-A fault in the parking brake system may cause the ANTISKID INOP light to illuminate.

2-Normal turn construction may not be possible under certain combinations of airspeed, short leg length, and a significant change in leg direction.

3-The emergency exit lights may also be illuminated by a switch on the aft attendants panel.

7.3.2.4 Must

No OM, o modal *must*, apresenta um total de 168 ocorrências, sendo:

- 79 ocorrências na estrutura 2, voz ativa
- 89 ocorrências na estrutura 3, voz passiva

A Colocados à esquerda

O quadro abaixo mostra os 10 principais colocados à esquerda juntamente com a quantidade de ocorrências:

Tabela 7.3.2.4.A: Colocados à esquerda (L1, L2, L3) de *must* no OM

Posição	Colocado voz ativa	Quantidade	LL	Colocado voz passiva	Quantidade	LL
1	CREW (L1)	9	57,827	EXTENSIONS (L1)	5	48,021
2	SEGMENT (L1)	5	39,557	HANDRAIL (L2)	5	45,368
3	AIRPLANE (L1)	8	38,347	NAV (L2)	5	24,909
4	ENTRIES (L1)	5	31,891	AIRSTAIR (L3)	4	24,900
5	ROUTE (L2)	6	28,504	VHF (L3)	4	22,724
6	BATTERY (L2)	4	25,803	PATH (L2)	4	18,296
7	ANTI (L2)	5	23,059	SWITCH (L1)	5	13,465
8	ICE (L1)	5	22,299	SPEED (L1)	4	7,901
9	FLIGHT (L2)	7	15,803	-----	-----	-----
10	ENGINE (L3)	6	15,295	-----	-----	-----

Apesar de *must* ocorrer mais na voz passiva do que na voz ativa nesse manual, é possível observar que, na voz ativa, os colocados possuem maior atração e são mais numerosos. Isso significa que na voz passiva há um maior número de itens lexicais diferentes que precedem *must*.

Entre os colocados à esquerda na voz ativa, *crew* é o principal, observando que refere um sujeito animado, isto é, a quem deve desempenhar a ação. Nota-se também que, diferentemente dos outros VMs até agora descritos, não há elementos anafóricos entre os colocados (considerando colocados as unidades com no mínimo de quatro ocorrências e LL 6,63 ou maior).

Classifico as unidades lexicais à esquerda de *must* na tabela abaixo.

Tabela 7.3.2.4.B: Classificação das unidades lexicais à esquerda de *must* -OM

Colocados à esquerda <i>Must</i> - voz ativa	Número de ocorrências	Colocados à esquerda <i>Must</i> -voz passiva	Número de ocorrências
<i>You</i> (L1)	-----	<i>You</i> (L1)	-----
Anáforas (L1)	01	Anáforas (L1)	03
Unidades simples (L1)	20	Unidades simples (L1)	26
Sintagmas nominais	51	Sintagmas nominais	52
Sint. verbais e prepos.	7	Sint. verbais e prepos.	8

De uma forma geral, o grupo dos sintagmas nominais prevalece perante os demais. Há poucas ocorrências de elementos anafóricos e *you* está ausente. No grupo formado por unidades simples na voz ativa, além da presença de *crew* já relatada acima, há uma ocorrência de *pilot* na posição L1, que também refere a quem deve realizar a ação.

Item B: Colocados à direita

Mostro agora os colocados à direita de *must*.

Tabela 7.3.2.4.C: Colocados à esquerda (R1, R2, R3) de *must*- no OM

Posição	Colocado voz ativa	Quantidade	LL	Colocado voz passiva	Quantidade	LL
1	BE (R1)	46	308,829	VERIFIED (R1)	4	43,537
2	ON (R2)	17	53,933	STOWED (R1)	5	41,554
3	CONTAIN (R1)	4	37,921	ENTERED (R1)	8	40,977
4	IN (R2)	12	35,851	TUNED (R1)	4	29,546
5	REMAIN (R1)	4	31,832	BEFORE (R2)	4	23,782
6	MANUALLY (R1)	4	21,649	IN (R2)	7	11,641
7	THAN (R3)	4	16,998	SELECTED (R2)	4	10,024
8	DURING (R3)	4	15,296	-----	-----	-----

Entre os dez primeiros colocados de *must* na voz ativa conforme a relação LL destaca-se o verbo *be*, na posição R1, com 46 casos. Além de *be*, só outros dois verbos são colocados significativos na voz ativa e cinco verbos na voz passiva. Os outros itens lexicais são preposições ou advérbios. Na próxima tabela mostro os verbos que ocupam as posições à esquerda classificados em sete grupos semânticos:

Tabela 7.3.2.4.D: Verbos na posição R1 de *must* OM

Tipos de Verbos	<i>Must</i> -voz ativa	<i>Must</i> -voz passiva
De atividade	14/ 17,7	68/ 76,4
De comunicação	-----	01/ 1,1
Mentais	2/ 2,3	12/ 1,3
Causativos	1/ 1,3	2/ 2,2
De ocorrência	-----	-----
De existência	57/ 72,4	02/ 2,2
De aspecto	5/ 6,3	04/ 4,4

Nota-se que a maior parte das ocorrências de *must* na voz passiva é acompanhada por um verbo de atividade (76,4). Entretanto, nos casos de voz ativa, é possível observar quase a mesma quantidade proporcional de verbos de existência (72,4). A diferença é que, com relação aos verbos de atividade, o grupo é composto de variados verbos, enquanto no caso dos verbos de existência, o grupo é dominado por *be*, que tem mais de 80% da posição L1 de *must* voz ativa. O grupo composto pelos verbos mentais é o segundo mais representativo nos casos de voz passiva. Isso se deve à presença dos verbos *observe* e *verify*.

C Usos e funções

Foi constatado que:

- Todas as ocorrências de *must* expressam a idéia de necessidade de que uma ação seja realizada. Em sua forma negativa, *must not* expressa proibição.
- *Must* é usado com mais freqüência na voz passiva, mas a diferença entre a freqüência das duas estruturas (ativa e passiva) é pequena, 79 e 89 ocorrências respectivamente.
- *Must* é usado principalmente na forma afirmativa, com apenas duas ocorrências de *must not* na voz passiva.

Com base nos resultados apresentados nos itens A, B e C, tomo os seguintes exemplos para ilustrar as características e usos mais salientes de *must* nesse manual, o OM.

1- The minimums reference selector **must** be in the BARO position to adjust the reference altitude marker..

2- The flight crew **must** manually tune the ILS for display on CDS.

3- CAUTION: Airstair handrail extensions **must** be stowed or substantial damage could result.

4- ADIRU alignment **must** not be attempted at latitudes greater than 78 degrees 15 minutes.

7.3.2.5 Shall

No MM, o modal *shall*, apresenta um total de duas ocorrências, sendo:

- uma ocorrência na estrutura 2, voz ativa

In the event of conflict with the AFM, the AFM shall supersede.

- uma ocorrência na estrutura 3, voz passiva

Quadrant shall be determined by the resulting inbound course.

Os dois usos de *shall* acima indicam que uma ação sempre ou tipicamente ocorre sob algumas circunstâncias. Indica uma previsão e pode ser substituído por *will*.

7.3.2.6 Should

No OM, o modal *should* apresenta um total de 77 ocorrências, sendo:

- 26 ocorrências na estrutura 2, voz ativa
- 51 ocorrências na estrutura 3, voz passiva

A Colocados à esquerda

O quadro abaixo mostra os principais colocados à esquerda juntamente com a quantidade de ocorrências:

Tabela 7.3.2.6.A: Colocados à esquerda (L1, L2, L3) de *should* no OM

Posição	Colocado <i>should</i> voz ativa	Quantidade	Relação LL	Colocado <i>should</i> voz passiva	Quantidade	Relação LL
1	CREW (L1)	04	25,512	-----	-----	-----
2	FLIGHT (L2)	05	12,834	-----	-----	-----

Devido à baixa frequência de *should* no OM, há apenas dois colocados na voz ativa e nenhuma na voz passiva. Acredito que os dois colocados formem o sintagma *flight crew*.

Investigo agora todos os itens lexicais nas posições à esquerda do nóculo *should* para verificar se há algum grupo dominante. Para o nóculo *should*, foi necessário incluir outra categoria para classificar esses itens lexicais, conforme tabela abaixo:

Tabela 7.3.2.6.B: Classificação das unidades à esquerda de *should* no OM

Colocados à esquerda <i>Should</i> - voz ativa	Número de ocorrências	Colocados à esquerda <i>Should</i> - voz passiva	Número de ocorrências
<i>You</i> (L1)	1/ 3,8%	<i>You</i> (L1)	-----
Anáforas (L1)	-----	Anáforas (L1)	1/ 2,0 %
Unidades simples (L1)	2/ 7,7%	Unidades simples (L1)	15/ 29, 4 %
Sintagmas nominais	13/ 50%	Sintagmas nominais	22/ 43, 1%
Sintag. Verbais e prepos.	7/ 26,9%	Sintag. Verbais e prepos.	13/ 25,5%
Condicional	05/ 11,6%	Condicional	

De uma forma geral, *should* tende a ser precedido por um sintagma (nominal, verbal ou preposicional). Nota-se a presença de *you* que não aparece como sujeito dos outros VMs. Na verdade, essa é a única ocorrência de *you* no em todo o OM. Destaco as ocorrências de *should* sem um sujeito, que acontece em três casos. Esses casos são aqueles nos quais *should* é usado como condicional, no início da oração.

B Colocados à direita

Mostro agora os colocados à direita de *should*.

Tabela 7.3.2.6.C: Colocados à direita (R1, R2, R3) de *should* no OM

Posição	Colocados voz ativa	Quantidade	Relação LL	Colocados voz passiva	Quantidade	Relação LL
1	-----	-----	-----	USED (R1)	04	23,788
2	-----	-----	-----	NOT (R1)	04	14,882
3	-----	-----	-----	IN (R2)	04	8,302

Should na voz ativa não apresenta colocados à direita. Na voz passiva, o verbo que mais se destaca é *used*.

Tabela 7.3.2.6.D: Verbos na posição R1 de *should* no MM

Tipos de Verbos	<i>Should</i> -voz ativa	<i>Should</i> -voz passiva
De atividade	13/ 50%	37/ 72,4%
De comunicação	1/ 3,8%	-----
Mentais	-----	8/ 15,6%
Causativos	02/ 7,6%	01/ 2,0%
De ocorrência	03/ 11,5%	-----
De existência	5/ 19,2%	02/ 4,0%
De aspecto	02/ 7,6%	03/ 6,0%

Os verbos que mais freqüentemente acompanham *should* são os de atividade. Na voz passiva os verbos mentais são os segundo mais freqüentes, devido, principalmente à presença de *observe* e *monitor* (no sentido de observar). Os verbos de existência se tornam mais freqüentes na voz ativa do que na voz passiva, devido aos verbos *be* e *remain*.

C Usos e funções

Foi constatado que:

- Todos os casos de *should* na voz passiva expressam uma instrução ou comando, ou seja, possuem um uso deôntico.
- Oito ocorrências de *should* na voz ativa são seguidas de sujeitos que correspondem aos agentes que devem exercer a ação, cinco vezes *flight crew* (s) e três vezes *the pilot not flying*.
- Na voz ativa, *should* expressa quatro funções diferentes:

-Como instrução, isto é, uso deôntico, em 18 ocorrências. É importante ressaltar que esse uso de *should* não é o de apenas recomendar, ou sugerir uma ação, mas sim o de estabelecer a única regra a ser seguida. Como por exemplo:

*The pilot not flying **should** call out any deviations from normal.*

-Como probabilidade, isto é, expressando uma suposição, uma necessidade lógica atenuada. Esse uso ocorre uma vez:

*Selecting the alternate fan **should** restore airflow and extinguish the OFF light within approximately 5 seconds.*

-Em duas situações, *should* é usado no lugar do presente simples em uma condicional real, esse uso assemelha-se ao uso descrito abaixo, mas apresenta uma estrutura sintática diferente.

*If the flaps **should** stop, the flap lever should be placed immediately in the same position as indicated.*

-Em cinco situações, ocupando o lugar do *if* em uma condicional, sendo três vezes no início da oração, por exemplo:

***Should** airspeed fall below the trim airspeed, unusual control column forces may be required to maintain the desired pitch attitude.*

- *Should* é, na maioria das vezes, seguido por um verbo de atividade. Os sujeitos são principalmente sintagmas, nominais, verbais ou preposicionais.

Com base nos resultados apresentados nos itens A, B e C, sugiro, além dos casos apresentados acima, os seguintes exemplos para ilustrar as características e usos mais salientes de *should* neste manual, o OM.

1-Should this occur, place the “B” system electric pump ON and retract the flaps.

*2-If cooling air is available from an outside source, the supply **should** be plugged in immediately after engine shutdown and **should** not be removed until just prior to engine start.*

*3-After the flaps are up, wing anti-ice **should** be used to melt any accumulation of slush.*

*4-The flight crew **should** search for any clues to the presence of windshear along the intended flight path.*

7.3.2.7 Will

No OM, o modal *will* apresenta um total de 163 ocorrências, sendo:

- 136 ocorrências na estrutura 2, voz ativa
- 27 ocorrências na estrutura 3, voz passiva
- Uma ocorrência na estrutura 5, perfeito (positiva)

A Colocados à esquerda

O quadro abaixo mostra os 10 principais colocados à esquerda juntamente com a quantidade de ocorrências:

Tabela 7.3.2.7.A: Colocados à esquerda (L1, L2, L3...) de *will* no OM

Posição	Colocado voz ativa	Quantidade	LL	Colocado voz passiva	Quantidade	LL
1	VNAV (L1)	18	129,489	-----	-----	-----
2	FMC (L1)	11	63,751	-----	-----	-----
3	LIGHT (L1)	12	58,037	-----	-----	-----
4	OFFSET (L1)	5	35,291	-----	-----	-----
5	BUGS (L1)	4	33,404	-----	-----	-----
6	WHICH (L1)	4	28,583	-----	-----	-----
7	DESCENT (L1)	5	27,828	-----	-----	-----
8	MODE (L1)	7	26,341	-----	-----	-----
9	THIS (L1)	4	23,773	-----	-----	-----
10	LNAV (L3)	4	21,656	-----	-----	-----

Nota-se que a sigla VNAV é o colocado mais forte de *will*, com 18 ocorrências. Além dessa sigla, FMC, em segundo lugar, e LNAV, em décimo, também estão presentes entre os colocados. Há dois elementos anafóricos, *which* e *this*, com quatro ocorrências cada. Os outros colocados são substantivos. O *will* na voz passiva não apresenta colocados.

Através da classificação dos elementos lexicais à esquerda do nóculo, é possível obter mais informações a respeito do comportamento de *will* no OM. Vide tabela abaixo:

Tabela 7.3.2.7.B: Classificação das unidades lexicais à esquerda de *will* no OM

Colocados à esquerda Will- voz ativa	Número de ocorrências	Colocados à esquerda Will-voz passiva	Número de ocorrências
<i>You</i> (L1)	-----	<i>You</i> (L1)	-----
Anáforas (L1)	8/ 5,9	Anáforas (L1)	1/ 3,7
Unidades simples (L1)	41/30,1	Unidades simples (L1)	4/ 14,8
Sintagmas nominais	59/ 43,4	Sintagmas nominais	13/48,1
Sintag. Verbais e prepos.	28/20,6	Sintag. Verbais e prepos.	9/33,4

Nota-se que o grupo mais freqüente é o de sintagmas nominais, com 72 ocorrências ao total (59 + 13). O grupo formado por unidades simples é o segundo mais freqüente na voz ativa, enquanto o grupo dos sintagmas verbais e preposicionais é o segundo mais freqüente na voz passiva.

B Colocados à direita

Mostro agora os colocados à direita de *will*.

Tabela 7.3.2.7.C: Colocados à direita (R1, R2, R3) de *will* no OM

Posição	Colocados <i>will</i> voz ativa	Quantidade	Relação LL	Colocados <i>will</i> voz passiva	Quantidade	Relação LL
1	BE (R1)	20	99,789	DISPLAYED (R1)	9	41,791
2	ILLUMINATE (R1)	10	72,372	IN (R3)	5	8,998
3	DISPLAY (R1)	10	63,134	-----	-----	-----
4	APPEAR (R1)	6	57,607	-----	-----	-----
5	REMAIN (R1)	6	53,728	-----	-----	-----
6	NOT (R1)	12	52,994	-----	-----	-----
7	OCCUR (R1)	5	43,821	-----	-----	-----
8	DISENGAGE (R1)	6	43,377	-----	-----	-----
9	RESULT (R1)	5	41,786	-----	-----	-----
10	CHANGE (R1)	5	33,038	-----	-----	-----

Nesta lista dos principais colocados de *will* à direita, nota-se que, além de verbos, a presença do advérbio *not* na posição R1 como um colocado significativo de *will*. Destaco o cópula *be*, entre os colocados de *will*+voz ativa, como o verbo com maior atração a *will*.

Tabela 7.3.2.7.D: Verbos na posição R1 de *will* OM

Tipos de Verbos	Will-voz ativa	Will-voz passiva
De atividade	60/44,1 %	22/81,5 %
De comunicação	2/ 1,4 %	-----
Mentais	1/0,7 %	1/3,7%
Causativos	11/ 8,2 %	1/3,7%
De ocorrência	19/ 14,0 %	1/3,7%
De existência	34/25,0 %	1/3,7%
De aspecto	9/6,6 %	1/3,7%

É notória a preferência que *will* apresenta pelos verbos de atividade, com os de existência em segundo lugar, mais uma vez devido ao alto número de ocorrências de *be*.

C Usos e funções

Com base nos itens A, B e C, foi constatado que:

- *Will* é usado com bastante mais frequência na voz ativa, na proporção de uma ocorrência na voz passiva a cada cinco na voz ativa. É usado principalmente na forma afirmativa, na proporção de uma ocorrência negativa a cada nove afirmativas.
- A maior parte das ocorrências de *will* na voz ativa expressa a idéia de previsão futura indicando que uma ação/conseqüência sempre ou tipicamente acontecerá
- *Will* ocorre em orações condicionais com a presença da conjunção *if* tanto na oração principal de uma condicional (17 ocorrências) quanto na subordinada (5 ocorrências). Esse uso de *will* na oração subordinada é raro e considerado não-gramatical nos materiais didáticos quando tratam de condicionais. Todas as ocorrências de *will* na oração subordinada ocorrem na voz passiva, como por exemplo:

*If the potable water tank **will** not be refilled immediately after the system is emptied, open the following circuit breakers and attach DO–NOT–CLOSE tags:*

P18–3 circuit breaker panel

- *LAVATORY WATER HEATER A*
- *LAVATORY WATER HEATER D*
- *LAVATORY WATER HEATER E*

À exemplo do que já foi relatado na seção 7.3.1.8, esse uso parece expressar algo contrário à expectativa. No caso acima, a expectativa é de que o tanque de água potável seja preenchido assim que esteja vazio, mas, caso isso não seja realizado, abra os disjuntores e ...

Com base nos resultados apresentados nos itens A, B e C, mostro os seguintes exemplos para ilustrar as características e usos mais salientes de *will* nesse manual, o OM.

*1-LNAV andVNAV **will** disengage if autopilot B is in use (can be reengaged if autopilot A is selected).*

*2-If runway is limited for the planned takeoff flap setting, consider using the next greater flap position with improved climb performance. This **will** provide additional stall margins with minimum performance penalties*

*3-If an airspeed restriction exists at the waypoint, it **will** be displayed in large font; otherwise the predicted speed will be displayed in small font.*

4- If the battery **will** be exposed to temperatures below -18°C (0°F), the battery should be removed and stored in an area warmer than -18°C (0°F), but below 40°C (104°F).

7.3.2.8 Would

No OM, o modal *would*, apresenta seis ocorrências, todas na voz ativa. Vide as linhas de concordância abaixo:

N Concordance

1 selects the highest climb thrust available **which would not result** in a throttle increase, when the aircraft
 2 position of the engine bleed air switches and isolation **valve would permit possible** backpressure of the APU. Therefore,
 3 the 9th stage modulating and shutoff valve. **This would cause the** 9th stage valve to close. DUAL BLEED Light
 4 the switch being extinguished. When engagement of a **mode would conflict with** current AFS operation, pushing the mode
 5 Description The above illustrated method of **departure would probably be** the easiest for most crewmembers. This
 6 the vertical angle for the glidepath of an ILS **approach would typically be** 3 degrees. This angle is displayed on the

A princípio, o uso parece expressar uma situação hipotética, ou o resultado de um evento imaginário. Entretanto, uma análise mais detalhada e o fato de ser um manual técnico, mostra que as situações não são hipotéticas ou irreais, mas expressam possibilidade. Assim, o *would* poderia ser substituído por outro VM mais freqüente no manual, ou pelo presente simples, transformando a oração em uma condicional real. Vide o exemplo e a sugestão abaixo:

*When engagement of a mode **would** conflict with current AFS operation, pushing the mode selector switch has no effect.*

Nesse caso o *would* poderia ser substituído por presente simples:

*When engagement of a mode **conclicts** with current AFS operation, pushing the mode selector switch has no effect.*

8 SÍNTESE DESCRITIVA E CONSIDERAÇÕES

Este capítulo visa a sintetizar os resultados descritos no capítulo anterior. Primeiro são resumidos os dados do *corpus* de estudo. O procedimento se justifica, na medida em que as observações foram variadas e incidiram sobre diferentes pontos.

8.1 SÍNTESE DO MM

Nesta seção resumo os dados do MM, ou manual de manutenção de aeronaves, e explico algumas reflexões a respeito dos resultados. Em seguida, comparo os resultados com as recomendações estabelecidas para a redação do MM no guia *Simplified English*. Logo após, comparo os resultados com o capítulo do livro de Shawcross (1992) que trata sobre os VMs em manuais de manutenção de aeronaves.

O quadro 8.1.A mostra a síntese dos colocados à esquerda dos VMs na voz ativa.

Quadro 8.1.A Síntese dos colocados à esquerda na voz ativa MM

<i>can</i>	<i>could</i>	<i>may</i>	<i>must</i>	<i>should</i>	<i>will</i>
YOU (L1)	THIS (L1)	HAZARDOUS (L3)	YOU (L1)	BLUE (L2)	YOU (L1)
WHICH (L1)	LONG (L3)	SIGNALS (L2)	LIST (L2)	WATER (L1)	THIS (L1)
THIS (L1)	WHICH (L1)	SENSING (L3)	THIS (L3)	-----	THAT (L2)
THAT (L1)	LANDING (L1)	MATERIALS (L2)	TACKY (L2)	-----	WHERE (L2)
OIL (L1)	-----	WHICH (L1)	BUT (L1)	-----	OIL (L1)
IT (L1)	-----	SOLVENTS (L1)	IF (L2)	-----	WHICH (L1)
AGENTS (L2)	-----	FLIGHT (L1)	WATER (L2)	-----	IT (L1)
HOT (L1)	-----	BEFORE (L2)	FITTINGS (L2)	-----	SURE (L3)
EXPLOSION (L1)	-----	PORTS (L3)	PERSONS (L1)	-----	SYSTEMS (L3)
FLAMMABLE (L3)	-----	VIBRATION (L1)	AIRPLANE (L1)	-----	GROMMET (L1)

Considerando somente os colocados à esquerda dos VMs na voz ativa no MM, observa-se que alguns itens lexicais são os preferidos na posição L1. Listo os itens e o número de ocorrências brutas dos principais colocados:

- *You* 405
- *This* 99
- *That* 70

- *Oil* 47
- *Which* 45
- *It* 42

Assim, é possível verificar que *you* destaca-se sobremaneira como item lexical que acompanha os VMs na voz ativa no MM, sobretudo *can*, *must* e *will*. Nota-se ainda que os elementos anafóricos são também muito freqüentes na posição L1, principalmente o *this*. A palavra lexical mais freqüente é *oil*. Diante disso, realizei uma busca de *oil* no MM e obtive 781 ocorrências. Esse número de ocorrências pode ser considerado bastante alto, ou seja, é uma palavra muito freqüente no corpus (a vigésima mais freqüente). Tendo em mente que o MM não é analisado em sua totalidade (apenas 4 capítulos), como foi explicado no capítulo 5, considere importante verificar se a palavra *oil* é igualmente distribuída entre todos os quatro capítulos ou se é peculiar de um capítulo. Para tanto, verifiquei a distribuição da palavra *oil* no MM. Essa informação pode ser obtida através do aplicativo *plot*, que mostra a concentração/distribuição de uso de uma palavra em um corpus. Vide tela abaixo:

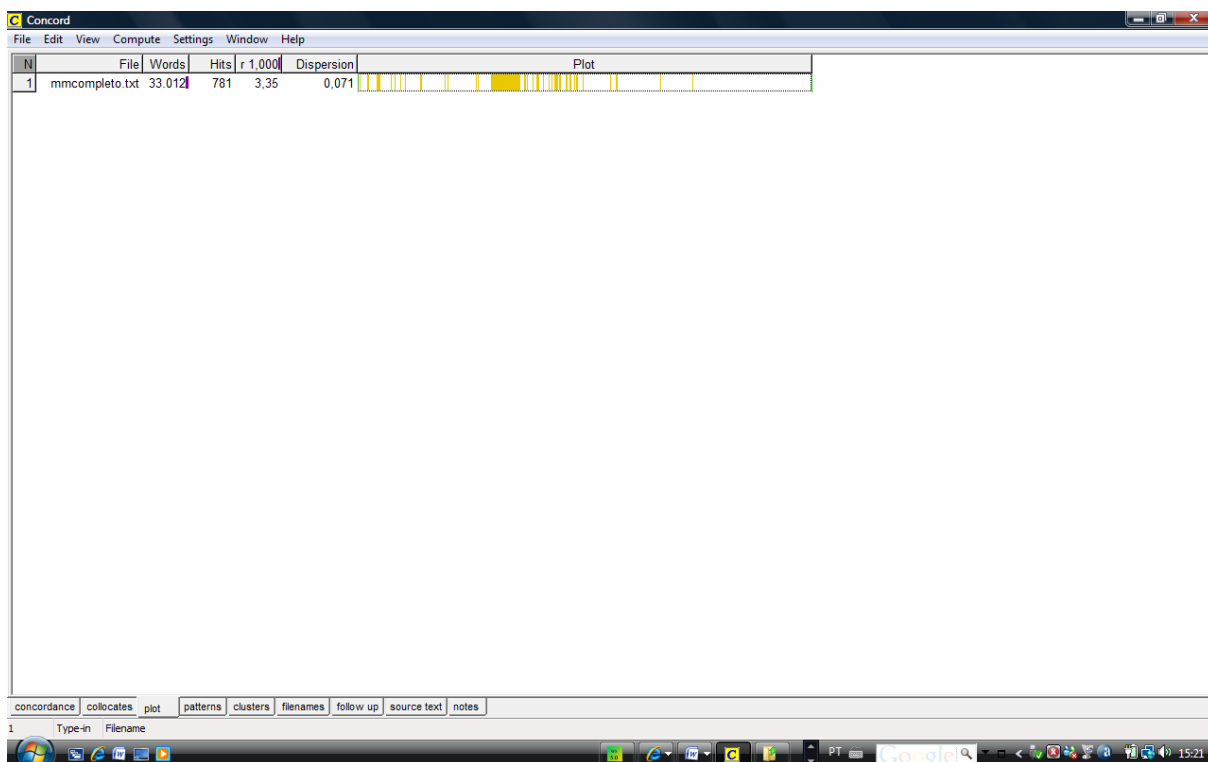


Figura 10: Distribuição da palavra *oil* no MM

Como pode ser visto, há uma densa presença de ocorrências de *oil* em uma parte do corpus, ou seja, em um dos capítulos. Isso significa dizer que, tivesse se eu tivesse analisado outros capítulos, há a possibilidade de que *oil* não estaria entre os itens lexicais mais

recorrentes com os VMs. Para contrastar, mostro a distribuição do colocado *you* no corpus, que é a décima segunda palavra mais freqüente no MM, com 1167 ocorrências.

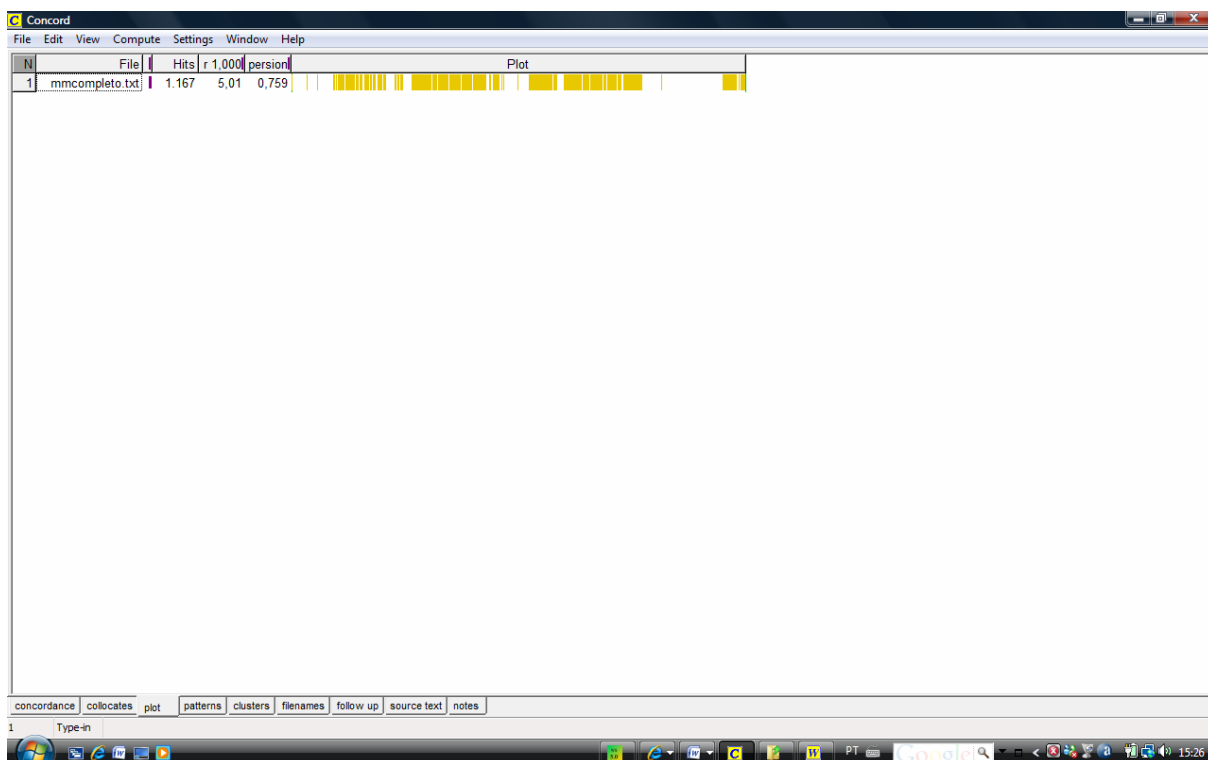


Figura 11: Distribuição da palavra *you* no MM

Pode-se verificar que *you* é distribuído de forma mais uniforme que *oil*. Isso significa que *you* não é peculiar de um capítulo somente, mas de todos os quatro capítulos analisados. Sendo assim, pode-se aventar a hipótese de que mesmo que outros capítulos tivessem sido analisados, o *you* seria ainda um colocado significativos dos VMs.

Observo agora o quadro síntese dos colocados à esquerda dos VMs na voz passiva no MM. Lembro que nesse manual, apenas 16,3 % das ocorrências modais estão na voz passiva. Desse modo, é natural que o número de colocados seja menor.

Quadro 8.1.B Síntese dos colocados à esquerda na voz passiva MM

<i>can</i>	<i>may</i>	<i>must</i>	<i>should</i>	<i>will</i>
SEALED (L2)	IT (L1)	ACTUATION (L2)	ENGINES (L1)	WHICH (L1)
FORCE (L3)	-----	SLAT (L3)	OIL (L2)	SURFACES (L2)
THEY (L1)	-----	VALVE (L1)	-----	FROM (L3)
BEARINGS (L1)	-----	-----	-----	REVISION (L1)
GREASE (L3, L2)	-----	-----	-----	-----
WHICH (L1)	-----	-----	-----	-----

A primeira observação a ser feita é de que nenhum dos VMs apresenta o número de dez colocados (mínimo de quatro ocorrências e LL de 6,63 ou mais). Outro fato a ser observado é que apenas *which* aparece como colocado de dois VMs, o *can* e o *will*. O *which* aparece como colocado dos VMs no MM em 22 ocorrências. Como esperado, devido à baixa frequência de VMs nessa estrutura, não há itens lexicais individuais que se destacam de forma significativa. Nota-se que *oil* também está presente entre os colocados dos VMs na voz passiva.

Observo agora os cinco (ou seis, considerando *should*) agrupamentos que representam os colocados à esquerda dos VMs. Essa observação é útil principalmente para os casos da voz passiva, que não apresentam um único item lexical de destaque. Assim, pode-se ter uma idéia do tipo de unidades lexicais que precedem os VMs.

Lembro que a taxonomia usada para categorizar as unidades lexicais que precedem os VMs surgiu da observação dos dados. Essa classificação foi realizada devido à crença de que algumas estruturas sintáticas são mais complexas, podendo gerar problemas de compreensão. Considero as categorias formadas pelo pronome *you* e por unidades simples, as de mais fácil compreensão, uma vez que o sujeito encontra-se junto ao VM. As outras categorias exigem que o leitor identifique primeiramente a quem, ou ao que o modal se refere, para então conseguir entender a função expressa pelo VM. Por exemplo, uma oração assim:

*YOU **MUST** OPEN THE VALVES SLOWLY OR HIGH TEMPERATURES CAN OCCUR.*

Será, provavelmente, mais facilmente compreendida do que uma oração assim:

*When persons are in the airplane during this test, they **must** be in good physical condition.*

Ou assim:

*If you know the loaded inflation pressure of the tire, the sum of all the air pressures added to the tire **must** not be more than the maximum given in Fig. 302.*

No MM, os agrupamentos lexicais mais recorrentes à esquerda dos VMs na voz ativa são:

Tabela 8.1.C: Unidades lexicais à esquerda dos VMs-voz ativa

Unidades lexicais à esquerda VM- voz ativa	Número de ocorrências
<i>You</i> (L1)	405/ 28,1%
Anáforas (L1)	183/ 12,7%
Unidades Simples (L1)	225/ 15,6%
Sintagmas nominais	272/ 18,8%
Sintag. Verbais e prepos.	355/ 24,6%
Condicional	03/ 0,2%

Pode-se notar que o grupo do *you* é o mais recorrente, com 28% das ocorrências. Esse grupo é seguido pelo grupo composto de sintagmas verbais e preposicionais e pelo grupo dos sintagmas nominais, com 24,6% e 18,8% respectivamente. O grupo composto pelos elementos anafóricos ocorre em 12,7% dos casos. Assim, apesar de *you* ser o agrupamento mais numeroso, a maioria das ocorrências dos VMs na voz ativa no MM ocorre em orações sintaticamente complexas. Analiso agora as unidades lexicais à esquerda:

Tabela 8.1.D: Unidades lexicais à esquerda dos VMs-voz passiva

Unidades lexicais à esquerda VM- voz passiva	Número de ocorrências
<i>You</i> (L1)	-----
Anáforas (L1)	38/ 13,7%
Unidades Simples (L1)	59/ 21,3%
Sintagmas nominais	100/ 36,1%
Sintag. Verbais e prepos.	80/ 28,9%
Início da oração	-----

Na voz passiva, o grupo mais freqüente é o composto pelos sintagmas nominais (36,1%) seguido pelo grupo dos sintagmas preposicionais (28,9%). Assim como na voz ativa, os elementos lexicais à esquerda na voz passiva também tendem a ser sintaticamente complexos, sendo que uma estrutura passiva já é considerada sintaticamente mais complexa.

Como foi possível observar, os modais no MM são geralmente inseridos em orações sintaticamente complexas. Essa informação é de extrema relevância uma vez que o destinatário desse manual são mecânicos de aeronaves provenientes de vários países e falantes das mais variadas línguas maternas. Os materiais didáticos destinados a esse público, deveriam, necessariamente, considerar esses dados ao elaborar explicações e exercícios a respeito do assunto.

Passo agora a analisar os colocados e as unidades lexicais à direita dos VMs no MM. O primeiro quadro mostra os itens lexicais que se colocam com os VMs à direita na posição R1.

Quadro 8.1.E Colocados à direita dos VMs na voz ativa-MM

<i>can</i>	<i>could</i>	<i>may</i>	<i>must</i>	<i>should</i>	<i>will</i>
CAUSE (R1)	CAUSE (R1)	BE (R1)	BE (R1)	BE (R1)	NOT (R1)
OCCUR (R1)	OCCUR (R1)	FLAMMABLE (R2)	INDIVIDUAL (R3)	-----	FLOW (R3)
DAMAGE (R2)	RETRACT (R2)	ERRORS (R3)	IDENTIFY (R2)	-----	BE (R1)
INJURY (R2)	UNWANTED (R3)	LOSS (R3)	EACH (R2)	-----	OIL (R2)
USE (R1)	DAMAGE (R3)	LEAD (R1)	NOT (R1)	-----	DECREASE (R1)
BE (R1)	POSSIBLE (R2)	CAUSE (R1)	HAVE (R1)	-----	HAVE (R1)
INJURIES (R2)	-----	LARGE (R2)	MORE (R3)	-----	PREVENT (R1)
COME (R1)	-----	PHASES (R3)	DO (R1)	-----	WHEN (R3)
GET (R1)	-----	REQUIRE (R1)	FULLY (R1)	-----	CAUSE (R1)
BURN (R1)	-----	NOISE (R3)	USE (R1)	-----	LET (R1)

Observando-se o quadro acima, nota-se que alguns verbos aparecem como colocados de mais de um VM. Listo abaixo os verbos preferidos dos VMs no MM e mostro com quais VMs eles ocorrem. Nessa listagem, incluo também os resultados dos VMs que, devido a sua baixa ocorrência, não estão inseridos no quadro acima:

- *Cause* 264 4VMs
- *Be* 157 8VMs
- *Occur* 95 4VMs
- *Damage* 79 3VMs
- *Use* 71 4VMs
- *Have* 33 5VMs

Nota-se que *cause* é, sem dúvida, o preferido. Ele ocorre com os três VMs que expressam possibilidade (*can*, *could* e *may*) e com o VM que expressa previsão (*will*). Esse fato parece apontar que, de alguma forma, o sentido, ou função desses VMs é similar. O mesmo acontece com o VM *occur*. O *damage* ocorre com dois VMs que expressam possibilidade (*can* e *could*) e com *will* (previsão).

A seguir verifico os colocados à direita dos VMs na voz passiva.

Quadro 8.1.F Colocados à direita dos VMs na voz passiva-MM

<i>can</i>	<i>may</i>	<i>must</i>	<i>should</i>	<i>will</i>
DAMAGED (R1)	USED (R1)	DONE (R1)	OPERATED (R2)	LUBRICATED (R1)
USED (R1)	-----	REMOVED (R1)	NOT (R1)	IDENTIFIED (R1)
BY (R2)	-----	BEFORE (R2)	USED (R1)	USED (R1)
FOUND (R1)	-----	-----	-----	-----
THAT (R2)	-----	-----	-----	-----
WITH (R2)	-----	-----	-----	-----

O verbo *used* aparece como colocado de quatro VMs, com um total de 28 casos. O verbo *damaged*, que é um importante colocado dos VMs na voz ativa, aparece também na voz passiva, com 11 ocorrências nessa estrutura. É necessário lembrar que *be* não é considerado colocado dos VMs na voz passiva por ser elemento obrigatório da estrutura. Entretanto, o *be* está lá, acompanhando um VM em 434 ocorrências (277 na voz passiva e 157 na voz ativa), sendo assim, o item lexical com maior grau de atração aos VMs. Os leitores dos manuais devem estar consciente a respeito desse fato e saber distinguir entre *be* como verbo auxiliar:

*This inspection procedure **must be** done when the airplane tail was dragged (touched the runway) when the airplane landed, or during a takeoff.*

E *be* como verbo principal, por exemplo:

*The flight safety of the airplane, together with the safety of health and the prevention of bad smells, **must be** your first consideration when you complete this task.*

Chamo também a atenção para o fato de a preposição *by* ser um colocado somente de um VM na voz passiva, o *can*, com apenas 13 ocorrências. Os livros didáticos de inglês tendem a ensinar a voz passiva com a presença da preposição *by* para indicar o agente da ação expressa pelo verbo principal.

Na próxima tabela mostro os tipos de verbo que seguem os VMs na voz ativa.

Tabela 8.1.G: Tipos de verbo na voz ativa-MM

Tipos de Verbos	Will	Should	Shall	Must	May	Could	Can
De atividade	141/ 46,8%	14/43,8%	05/31,3%	133/53,2%	11/ 13,7%	05 / 13%	202/28%
De comunicação	2/ 0,7%	-----	-----	-----	-----	-----	-----
Mentais	16/5,3%	-----	-----	39/ 15,6%	4/ 5,4%	-----	17/2,4 %
Causativos	60/ 20%	02/6,3%	-----	06/ 2,4%	28/ 34%	19/ 55%	305/42%
De ocorrência	28/ 9,3%	03/9,4%	-----	06/ 2,4%	6/ 6,8%	7/ 19%	130/18%
De existência	44/ 14,6%	11/34,4%	10/62,5%	65/ 26%	33/ 40,1%	05/ 13%	69/9,6%
De aspecto	10/ 3,3%	02/6,3%	01/6,3%	01/ 0,4%		-----	-----

Os totais gerais dos verbos colocados à direita dos VMs no MM na voz ativa são:

- De atividade 511
- Causativos 420
- De existência 237
- De ocorrência 180
- Mentais 76
- De aspecto 14
- De comunicação 02

Na tabela que segue apresento os verbos que seguem os VMs na voz passiva.

Tabela 8.1.H: tipos de verbo na voz passiva-MM

Tipos de Verbos	Will	Should	Must	May	Can
De atividade	39/ 73,5%	33/97%	70/ 95,9%	21/ 75%	58/ 68%
De comunicação	-----	-----	-----	02/ 7%	-----
Mentais	7/ 13,2%	-----	03/ 4,1%		15/ 17%
Causativos	-----	01/3%	-----	03/ 11%	13/ 15%
De ocorrência	01/ 1,9%	-----	-----	-----	-----
De existência	5/ 9,5%	-----	-----	02/ 7%	-----
De aspecto	01/ 1,9%	-----	-----	-----	-----

Os totais gerais dos verbos colocados à direita dos VMs no MM na voz passiva são:

- De atividade 221
- Mentais 25
- Causativos 17
- De aspecto 14
- De existência 07
- De comunicação 02
- De ocorrência 01

Nota-se que os verbos de atividade detêm a preferência tanto na voz ativa quanto na voz passiva. Lembro que a voz passiva não é uma característica marcante entre as estruturas modais neste manual, com somente 16,32% das ocorrências, comparadas com 83,51 % das ocorrências na voz ativa. Os verbos de atividade são encontrados como colocados de todos os VMs. Uma das razões para isso é que, devido á característica deôntica do texto estudado, o MM, é natural que os modais sejam seguidos por verbos que expressem alguma ação a ser realizada.

Por outro lado, principalmente os causativos na voz ativa, mas também os de ocorrência em menor grau, mostram uma clara preferência pelos VMs que estão no eixo de possibilidade ou previsibilidade. *Will* (previsão) é acompanhado por um causativo em 20% de suas ocorrências na voz ativa. *Can, could e may* (possibilidade) são acompanhados por um causativo em 42%, 55% e 34% dos casos respectivamente, ao passo que *must e should* apresentam um causativo à direita em apenas 2,4% e 6,3% das ocorrências. Esse fato pode ser explicado com base no sentido e na prosódia semântica dos causativos. Hunston (2002), ao analisar o uso de *cause* em um corpus de inglês geral, mostra que o verbo *cause* é, quase em sua totalidade seguido por uma situação negativa, como por exemplo, *cause a disaster, cause an accident*. Situação semelhante é observada no MM. O verbo *cause* ocorre 328 vezes no MM, e em 288 vezes acompanhando um VM. Os principais colocados à direita de *cause* são *damage* com 69 ocorrências e *injury/injuries* com 43 ocorrências. Entretanto o principal colocado de *cause* é o VM *can*, em 214 ocorrências (voz ativa e voz passiva). Vide exemplo abaixo:

W_A_R_N_I_N_G_ : DO NOT GET HYDRAULIC FLUID ON YOU. HYDRAULIC FLUID, BMS 3-11 CAN CAUSE INJURY TO PERSONS.

As colocações dos VMs de possibilidade com verbos lexicais de prosódia semântica negativa também explicam a alta frequência desses VMs nas seções de CAUTION e WARNING. Essas seções são graficamente (caixa alta) destacadas devido a sua importância, isto é, o mecânico deve prestar mais atenção a essas seções. Como os VMs de possibilidade anunciam conseqüências negativas e perigosas, seu “habitat natural” no VM é sem dúvida, essas seções.

Com base nesses resultados pode-se concluir que *cause* possui uma prosódia semântica negativa no MM e que essa prosódia semântica negativa é extrapolada aos VMs com os quais ele ocorre. Além de *cause*, os VMs de possibilidade colocam-se com outros

verbos lexicais que também apresentam uma semântica negativa, são eles, *damage*, *occur*, entre outros. Torna-se, assim, fácil de entender a razão pela qual os verbos que expressam obrigação (seja ela forte ou fraca), como é o caso de *must* ou *should*, não se colocam com os causativos. Não seria lógico instruir que o mecânico causasse um dano á aeronave, mas é muito mais lógico instruir ao mecânico que opere um motor, como no exemplo abaixo:

*YOU **MUST** OPERATE THE ENGINE AT AN IDLE TO GET AN OIL LEVEL INDICATION THAT IS ACCURATE.*

Os verbos de existência estão em terceiro lugar na voz ativa, lembrando que na voz passiva *be* não é considerado colocado do modal por fazer parte da estrutura, sendo assim considerado nóculo juntamente com o VM. O mais freqüente verbo desta categoria é *be*. Os verbos de existência ou de relacionamento (mesma categoria, conforme apresentado no capítulo 6) expressam o estado que a situação está ou deve estar para que outra ação possa ser realizada, como por exemplo:

*1) The brakes **must be** cool (you can put your hand on the housing) before you apply deicing/anti-icing fluid to the wheel area.*

Os verbos mentais ocupam o quinto lugar na voz ativa e o segundo na voz passiva. Os principais verbos desta categoria encontrados no MM são *see*, *examine* e *make sure*, vide exemplo abaixo:

N_O_T_E_: *Damage you **cannot see**, for example delamination, can extend to the areas around the damaged area you **can see**.*

De uma forma geral, os VMs apresentam menos usos quando comparados aos vários usos e sentidos possíveis de cada VM (conforme capítulo 4). Apresento sucintamente os VMs e seus sentidos no quadro abaixo.

Quadro 8.1.I: Sentidos dos VMs no MM

VM	SENTIDOS
<i>Can</i>	Possibilidade (806), proibição (forma negativa, <i>cannot</i> 3)
<i>Could</i>	Possibilidade (39)
<i>May</i>	Possibilidade (110)
<i>Might</i>	Possibilidade (3)
<i>Must</i>	Obrigação (323)
<i>Shall</i>	Obrigação (18)
<i>Should</i>	Obrigação (20), probabilidade (9), Condicional (3)
<i>Will</i>	Previsão (342), obrigação (12)
<i>Would</i>	Possibilidade (2), previsão (1)

Conforme o quadro acima, pode-se constatar que o principal uso dos VMs no MM é o de expressar possibilidade, com 963 ocorrências. O segundo uso é o de obrigação, com 373 ocorrências. O sentido de previsão vem em seguida com 343 casos, probabilidade com nove e condicional com 3. Pode-se assim dizer que os usos dos modais neste manual são realmente reduzidos quando comparados ao inglês geral.

Ainda que de fora assistemática, foi possível observar que, há a recorrência de modais em uma mesma oração ou ato comunicativos, vide exemplo:

*C_A_U_T_I_O_N_: YOU **MUST** LUBRICATE ALL THE BEARINGS AND JOINTS IN THE AREA YOU CLEANED. THE LUBRICANT **WILL** REMOVE THE UNWANTED FLUIDS WHICH **COULD** FREEZE OR CAUSE CORROSION TO THE BEARING OR JOINT. IF YOU DO NOT LUBRICATE THE BEARINGS AND JOINTS, DAMAGE TO THE COMPONENTS CAN OCCUR.*

Na mesma seção CAUTION nota-se a presença de quatro VMs, sendo *will* e *could* na mesma oração. Como mencionado, não foi realizado um levantamento para verificar quantas vezes isso ocorre, mas pode-se sugerir como mais uma característica peculiar do uso dos VMs no MM e deveria, assim, ser inserido nos materiais didáticos.

Na próxima seção comparo os resultados do MM com as recomendações estabelecidas para a redação deste mesmo manual através do guia *Simplified English* (explicado no capítulo 3).

8.1.1 Maintenance Manual e o Simplified English

No capítulo 3 apresentei o guia *Simplified English* e as regras estabelecidas pela AECMA para redigir manuais de manutenção de aeronaves. Lembro que o *Simplified English* não consiste em um “estudo” sobre manuais de manutenção, mas em uma proposta prescritiva de guiar os redatores dos manuais quanto ao tipo de linguagem que o manual deve conter. A justificativa para o estabelecimento dessas regras está no fato de que esses manuais são obrigatoriamente escritos na língua inglesa e não podem ser traduzidos, uma vez que os fabricantes não se responsabilizam pelas traduções. Como consequência, profissionais ao redor do mundo que trabalham na manutenção da aeronave precisam ler e compreender esses manuais escritos na língua inglesa. Entretanto, foi constatado pela AECMA¹³⁸ (informação que corroboro) que, diferentemente dos pilotos, os mecânicos (de empresas aéreas), de uma

¹³⁸ The European Association of Aerospace Industries;

forma geral, não possuem proficiência na língua inglesa adequada. Assim, além de instruir as empresas aéreas a investir no ensino da língua inglesa para esses profissionais, a AECMA decidiu investir na padronização da linguagem de forma a simplificar a compreensão dos manuais.

Nesta seção retomo essas regras e verifico até que ponto elas são seguidas no manual estudado. Lembro que as regras estabelecidas pelo SE são gerais e abrangentes, com a pretensão de cobrir todas as áreas lingüísticas que perfazem um manual de manutenção de aeronaves. Contudo, limito-me a observar somente as questões que possuem algum tipo de relação com os VMs ou seus usos. A observação quanto ao cumprimento ou não das regras se justifica, na medida em que, ao elaborar material didático para os mecânicos é necessário saber se as regras são efetivamente seguidas, pois, quando e se não forem, essa informação deve constar no material didático. Listo assim as regras e as discuto à luz dos resultados obtidos neste trabalho.

(i) univocidade, ou seja, para um significado existe uma palavra, da mesma forma, para uma palavra existe um significado e ainda, para uma palavra existe somente uma função.

Parece haver uma real intenção em manter essa regra. Como visto no capítulo 4, os VMs possuem quase que uma infinidade de usos e sentidos, mas eles são consideravelmente reduzidos no MM. Entretanto, alguns VMs, leia-se *can*, *should*, *will* e *would* apresentam mais de um sentido no manual, apesar de haver sempre um usos quantitativamente predominante. Por outro lado, o mesmo sentido é expresso por mais de um VM. Vide os exemplos abaixo de *may*, *can* e *could*:

W_A_R_N_I_N_G_: DO NOT BREATHE THE VAPORS OF THE SOLVENTS OR ETCH SOLUTIONS. DO NOT GET THEM IN YOUR EYES, ON YOUR SKIN OR ON YOUR CLOTHES. INJURY TO PERSONS **MAY** OCCUR.

ou

W_A_R_N_I_N_G_: DO NOT GO NEAR THE LANDING GEAR AREA FOR 1 HOUR AFTER THE AIRPLANE HAS MADE A HIGH ENERGY STOP. INJURY TO PERSONS **CAN** OCCUR.

*C_A_U_T_I_O_N_: REPLACE THE TUBE WITH A TEMPORARY REPAIR WITH A NEW SECTION OF TUBE WITHIN A MAXIMUM OF 250 FLIGHT HOURS. IF YOU DO NOT REPLACE THE TUBE, DAMAGE **COULD** OCCUR.*

Ou

*C_A_U_T_I_O_N_: DO NOT PERFORM THE FIELD HARDNESS TEST ON THE VERTICAL FACES OF THE BRAKE MOUNTING FLANGES OR OTHER PARTS. IF YOU DO NOT OBEY THESE INSTRUCTIONS, DAMAGE TO THE AIRPLANE **CAN** OCCUR.*

A diferença entre os VMs *can*, *could* e *may* é geralmente bastante tênue ou difícil de ser identificada. Coates (1995) sugere que *may* envolve subjetividade, com uma expressão simultânea a respeito do que está sendo falado, como por exemplo, *I'm not sure*, enquanto *can* envolve objetividade. A autora também comenta que *may* seria o correspondente formal de *can* para expressar possibilidade. A diferença entre *can* e *could* estaria no grau de possibilidade. Essas distinções não parecem se aplicar nos contextos aqui estudados. Não há traços de subjetividade ou de formalidade e o grau de possibilidade ou probabilidade é o mesmo. Todos os quatro usos acima parecem expressar o mesmo uso, uma possibilidade fatural (Leech, 1989), ou, nos termos de Shawcross (1992, p. 94), de que um evento ou uma ação podem ocorrer. Em outras palavras, não há diferenças semânticas nos usos de *can*, *could* e *may* acima e o mesmo modal poderia ter sido utilizado em qualquer uma das quatro situações. Assim, a regra relacionada à univocidade parece não estar sendo cumprida na íntegra¹³⁹.

As duas próximas regras do SE são relacionadas ao mesmo tema, uso da voz passiva, assim apresento e discuto as duas regras concomitantemente:

(ii) uso somente da voz ativa, apenas um entre dez usos pode ser na voz passiva.

(iii) Não usar o passado particípio de um verbo juntamente com o verbo modal (chamado de *helping verb* no SE), de forma a não tornar a frase mais complexa. O SE fornece os seguintes exemplos:

WRITE: ...you **can** adjust

NOT: ...**can** be adjusted

¹³⁹ Essa discussão também está relacionada com o próximo assunto especificamente relacionado aos VM, em que é sugerido o uso de *can* em todos esses contextos.

A voz passiva não foi observada em separado, somente nas estruturas modais, mas considerando as estruturas em que os VMs se encontram, há na verdade um caso de voz passiva a cada 5,2 casos de voz ativa. Dessa forma, essa regra não está sendo seguida. Considerando ainda a regra (iii), que também é relacionada ao uso da voz passiva, mas dessa vez diretamente relacionada aos VM, a regra é ainda mais violada. Por outro lado, na seção 7.2 contrastei a frequência de uso da estrutura ativa e passiva nos três manuais do AC, o MM, o OM e o QRH. No MM a estrutura VM+passiva é usada em somente 16,32% das ocorrências, comparando a 55,68% no OM e 49,25% no QRH. Assim, apesar de a regra (iii) não estar sendo fielmente seguida, ela parece estar realmente influenciando a forma como o MM é redigido.

Retomo agora o quadro com as recomendações específicas sobre o uso dos VMs. Lembro que o SE apresenta as palavras que devem ser usadas no manual em letras maiúsculas. Para facilitar a apresentação, apresento e discuto cada linha do quadro separadamente:

Keyword (part of speech)	Assigned meaning/USE	APPROVED EXAMPLE	Not Acceptable
CAN (v) CAN, COULD	Helping verb that “means” to be “able” to	YOU CAN CLEAN THE DRAIN HOLES WITH THE CLEARING TOOL.	

Essa regra parece ser seguida. Entretanto, não fica claro se o sentido de can é o de habilidade ou o de possibilidade. Leech (1989) sugere que um implica o outro, ou seja, habilidade implica possibilidade.

Keyword (part of speech)	Assigned meaning/USE	APPROVED EXAMPLE	Not Acceptable
CANNOT	“Can not”	YOU CANNOT FILL THE CONTAINERS WHILE THEY ARE INSTALLED IN THE AIRCRAFT.	

O exemplo acima não deixa claro qual o uso de *cannot* deve ser usado:

-o de proibição ou recomendação forte para não fazer algo, isto é, “não encha os *containers* senão algo vai acontecer”;

-ou o de falta de possibilidade, isto é, “você não terá a possibilidade de encher os *containers* quando eles estão instalados na aeronave”.

Assim, não é possível afirmar se a regra foi cumprida ou não, apesar de os dois usos estarem presentes no manual.

Keyword (part of speech)	Assigned meaning/USE	APPROVED EXAMPLE	Not Acceptable
May (v)	CAN, POSSIBLY (adv)	IF YOU USE INCORRECT EQUIPMENT, YOU CAN CAUSE DAMAGE TO THE VANES	The vanes may be damaged by using incorrect equipment.

Essa regra não é cumprida. O *may* apresenta 110 ocorrências no MM, todas com o uso similar a *can*.

Keyword (part of speech)	Assigned meaning/USE	APPROVED EXAMPLE	Not Acceptable
MUST (v) NOTE: no other form of this verb	Helping verb that shows “obligation”	THE “BAT 1” SWITCH MUST BE IN THE “ON” POSITION	

Essa regra é cumprida. Todas as ocorrências de *must* expressam obrigação.

Keyword (part of speech)	Assigned meaning/USE	APPROVED EXAMPLE	Not Acceptable
shall (V)	MUST	HOLES MUST NOT HAVE SHARP EDGES.	Holes shall not have sharp edges.

Essa regra não é cumprida, há 18 ocorrências de *shall* que poderiam ter sido substituídas por *must*.

Keyword (part of speech)	Assigned meaning/USE	APPROVED EXAMPLE	Not Acceptable
Should (v)	Must, if (CON)	PERSONNEL MUST PUT ON PROTECTIVE CLOTHING. IF A FAILURE OCCURS, STOP THE TEST.	Personnel should wear protective clothing. Should a test occur, stop the test.

Essa regra não é cumprida. *Should* é usado 66 vezes no MM, em três sentidos diferentes. Há casos em que *should* apresenta dois sentidos em uma mesma oração, como no exemplo abaixo:

Should operators elect not to follow this recommended procedure, they should develop alternate procedures or establish conditions adequately protecting the personnel and equipment involved.

O primeiro *should* poderia ter sido substituído por *if* e o segundo por *must* como dita a regra.

Keyword (part of speech)	Assigned meaning/USE	APPROVED EXAMPLE	Not Acceptable
WILL (v) NOTE: no other forms of this verb	Helping verb that shows simple future tense	THE TEST ALTITUDES WILL BE HIGHER.	

Essa regra não é seguida em sua totalidade. Primeiramente, cabe ressaltar uma construção sintática idiossincrática de *will* no MM quando o VM aparece na oração subordinada de uma condicional, vide exemplo abaixo:

*If you **will** remove the starter, do not service the starter with oil.*

Como já mencionado, *will* nessa posição parece enfatizar uma ação a ser realizada contrária as expectativas. Entretanto, tendo em vista os esforços para padronizar e simplificar a linguagem dos manuais de manutenção de aeronaves, considero essa ênfase desnecessária. A oração “If you remove the starter, do not service the starter with oil” seria de mais fácil compreensão principalmente para os milhares de leitores não falantes nativos de inglês.

Will também expressa obrigação em 12 ocorrências.

Keyword (part of speech)	Assigned meaning/USE	APPROVED EXAMPLE	Not Acceptable
Would (v)	CAN	SOLVENTS THAT STAY ON THE PART CAN CAUSE CORROSION.	Solvents left on the part would corrode the part.

Essa regra também não é cumprida. *Would* ocorre em três situações que poderiam ter sido substituídas por *can* ou *will*.

As comparações acima mostram que nem todas as regras são seguidas. Entretanto, nota-se que no que diz respeito aos VMs, a grande maioria das ocorrências está realmente de acordo com as regras. Assim, ao elaborar materiais didáticos de inglês com fins específicos, apesar de dar ênfase especial ao que é mais recorrente e em consonância com as regras, não é suficiente se ater somente ao que diz a regra, mas sim ao que o MM efetivamente apresenta.

Na próxima seção discuto e comparo o English for Aircraft com os resultados deste trabalho.

8.1.2 English for Aircraft

O *English for Aircraft – Documentation Handbook* (Shawcross, 1992) é um livro destinado a mecânicos de aeronaves que já possuem um conhecimento básico da língua inglesa e de manutenção de aeronaves. Considero esse livro de importância incontestável, pois foi a primeira obra didática para esse fim. Conforme o autor, o livro pode ser usado em aula ou par estudo individual. O livro é dividido em 26 módulos ou capítulos. Cada módulo trata de um assunto diferente. Os módulos consistem em explicações e exercícios práticos retirados de manuais de aeronaves. Um desses módulos é denominado “*Possibility, probability, necessity, conditions*”, e trata sobre os verbos modais. Nesta seção, viso a verificar até que ponto as informações e exercícios apresentados por Shawcross são realmente representativos da linguagem dos manuais de manutenção de aeronaves.

Os dados analisados no capítulo 7 e sintetizados na seção 8.1 mostram que apesar de haver uma redução considerável nos sentidos que os modais expressam no MM, há diferentes nuances nesses sentidos e diferentes prosódias semânticas na oração como um todo. Além disso, apesar de haver um esforço na tentativa de padronizar a linguagem da documentação técnica de manutenção, os VMs tendem a ocorrer em estruturas sintáticas bastante complexas e por vezes até consideradas não-gramaticais nos livros didáticos e gramáticas.

Alguns exemplos já foram apresentados no capítulo 4. Começo a análise com *can*.

- ***Can***

Shawcross (1992) apresenta os usos de *can* sob o título de “*Possibility*” (p. 93):

Use Can + Infinitive to express possibility. (It’s usually better than “it’s possible to/that...”.)

Can expresses physical or technical possibility, capacity etc.

Após essa breve explicação, Shawcross oferece os seguintes exemplos:

*The A320 **can** carry 172 passengers.*

*Coaxial cables **can** transmit numerous messages.*

*In the event of an engine failure, the remaining engine **can** power all the hydraulic systems through the PTU.*

*The APU **can** be started up to 25,000 feet.*

O autor denomina o seguinte uso de *can* de “*capacity*”:

*The outer tank **can** hold 3,500kg.* (ibid, p. 95)

Após explicar os usos dos outros VMs nos manuais de manutenção, há um exercício que consiste em frases com lacunas a serem preenchidas com o VM correto. A instrução para o exercício é a seguinte:

*Choose **can, may, must, should or need to**¹⁴⁰ complete these sentences, as in the example:*

The handpump can pressurize the yellow system.

Há 25 frases no exercício. De acordo com a chave de respostas, *can* é a resposta certa em seis frases e em outra o autor sugere a resposta certa como *can* ou *may*, vide abaixo:

Bird strikes can/may damage the fan blades.

Com base nos exemplos e exercícios, observei as seguintes características:

- *Can* é apresentado somente na forma afirmativa, na voz ativa e na voz passiva.
- O sentido é o de possibilidade ou capacidade.
- Os sujeitos são unidades simples ou sintagmas nominais.
- *Can* é apresentado na oração principal de uma condicional.

Observei a ausência das seguintes características salientes conforme os resultados deste estudo. No *English for Aircraft* os exemplos de *can* não apresentam:

- *You*, anáforas, sintagmas verbais ou preposicionais como sujeito. Essas unidades lexicais representam mais de 60% no MM. Na verdade, esse tipo de sujeito não é apresentado com nenhum dos VMs.

- A forma negativa
- Prosódia semântica negativa, e verbos lexicais causativos ou de ocorrência
- Informações sobre a grande frequência do VM nas seções WARNING e CAUTION

- ***Could***

Shawcross apresenta apenas um exemplo de *could* na unidade

*Hot oil **could** cause injury.* (SHAWCROSS, 1992, p. 94)

Observei a ausência das seguintes características salientes nos exemplos de *could*:

- Prosódia semântica negativa, e verbos lexicais causativos ou de ocorrência
- Anáforas como sujeitos

¹⁴⁰ Como *need to* não foi estudado nesta tese, não apresentarei seus exemplos.

- Informações sobre a grande frequência do VM nas seções WARNING e CAUTION

- **May**

Shawcross (1992) sugere apenas que *may* é usado para indicar que um evento ou uma ação podem ocorrer, ou acontecer. Semelhantemente a *could*, o autor sugere a idéia de possibilidade expressando incidentes técnicos não planejados e fenômenos naturais para o uso de *may* (ibid, p. 94):

*In dense weather, ice **may** form on the wings.*

*If there is a lot of traffic on arrival, the aircraft **may** have to be hold before landing.*

*The assembly **may** be repaired before overhaul.*

*Water **may** accumulate in the lower fuselage. (SHAWCROSS, 1992, p.95)*

Shawcross declara que em inglês geral, *can* e *may* são intercambiáveis, mas, que o *Simplified English* recomenda *can*.

Entre os 25 itens do exercício, *may* é considerado a única resposta certa (conforme a chave de resposta fornecida pelo autor) em seis frases. Além desses seis casos, o autor sugere tanto *may* quanto *can* como certos. No MM *can* apresenta cerca de 800 ocorrências, enquanto *may* apresenta pouco mais de cem. Apesar da maior importância quantitativa de *can*, o autor enfatiza ambos os VMs da mesma forma.

Com base nos exemplos e exercícios, observei as seguintes características:

- *May* é apresentado somente na forma afirmativa, na voz ativa e na voz passiva.

- O sentido é o de possibilidade implicando que fatos e incidentes não planejados podem ocorrer.

- Os sujeitos são unidades simples ou sintagmas nominais.

- *May* é apresentado na oração principal de uma condicional.

- *May* aparece na mesma oração que outro modal, o *should*, vide exemplo abaixo:

*Only original parts **should** be used, but approved equivalent parts **may** be installed.*

Foi possível notar a ausência das seguintes características salientes do uso de *may* no MM:

- Anáforas, sintagmas verbais ou preposicionais como sujeito. Essas unidades lexicais representam quase de 60% dos sujeitos de *may* no MM.

- A forma negativa, que corresponde a 1/3 das ocorrências.

- Prosódia semântica negativa, e verbos lexicais causativos ou de ocorrência

- Informações sobre a grande frequência do VM nas seções WARNING e CAUTION

- ***Might***

Might não é apresentado por Shawcross. Lembro que no MM *might* é realmente raro, com apenas três ocorrências.

- ***Must***

Shawcross (1992, p. 94) afirma que *must* na linguagem técnica é geralmente usado para expressar “necessidade de que algo seja feito”:

*If the tire is deflated the mechanic **must** remove it.*

*When the demand for cooling air decreases, the Ram Air Outlet **must** be closed.*

Entre os 25 itens do exercício, *must* é a resposta certa em oito.

Com base nos exemplos e exercícios, observei as seguintes características:

- *Must* é apresentado somente na forma afirmativa, na voz ativa e na voz passiva.
- O sentido é o de obrigação, ou seja, de forte recomendação de que a instrução expressa pelo verbo principal seja seguida.
- Os sujeitos são unidades simples ou sintagmas nominais, um caso de sintagma preposicional e um caso de *you*.

Foi possível notar a ausência das seguintes características salientes do uso de *must* no MM:

- A forma negativa, pois o advérbio *not* é o quinto colocado mais importante de *must* na voz ativa.
- Exemplos de *must* em orações condicionais.

Apesar dessas ausências, o *must* parece estar bem representado no livro de Shawcross.

- ***Shall***

Conforme Shawcross, na linguagem técnica, *shall* aparece expressando obrigação. Nestes casos, *shall* poderia ser substituído por *must*:

*When open on the ground, the landing gear bay doors **shall** be safetied (locked by a safety pin or sleeve).*

*Non-approved lubricants **shall** not be used.* (SHAWCROSS, 1992, p.94)

Shall não é incluído no exercício. Com base nos exemplos, é possível constatar que:

- *Shall* é apresentado nas formas afirmativa e negativa e somente na voz passiva.
- O sentido é o de obrigação, ou seja, de forte recomendação.
- Os sujeitos são sintagmas nominais.

Foi possível notar a ausência das seguintes características salientes do uso de *shall* no MM:

- A voz passiva. Apesar de *shall* apresentar poucas ocorrências, as ocorrências ativas são cinco vezes mais frequentes do que as ocorrências passivas.
- Sintagmas verbais ou preposicionais como sujeito. Essas unidades lexicais representam mais de 60% dos sujeitos de *shall* no MM.

- ***Should***

Shawcross sugere que *should* expressa uma recomendação e não uma obrigação ou necessidade.

Chocks should be placed under the main gear wheels during a night stop.
(SHAWCROSS, 1992, p. 94)

Shawcross expõe outro uso de *should* em uma condicional real no lugar de *if*:

Should the normal servovalve fail, the alternate servovalve will take over.
(SHAWCROSS, 1992, p. 95)

Entre os 25 itens do exercício, *should* é a resposta sugerida em um.

Com base nos exemplos e exercícios, observei as seguintes características:

- *Should* é apresentado somente na forma afirmativa, na voz passiva e no lugar de *if* em uma condicional.
- O sentido é o de recomendação ou de condicional.
- Os sujeitos são unidades simples ou sintagmas nominais.

Foi possível notar a ausência das seguintes características salientes do uso de *should* no MM:

- Os usos de *should* como obrigação e probabilidade. Na verdade, o uso como mera recomendação não foi encontrado no MM, mas é o uso sugerido como principal por Shawcross.

- Exemplos de *should* com sintagmas verbais e preposicionais como sujeitos.

- **Will**

Shawcross não introduz ou explica o uso de *will*, apenas coloca que na maior parte dos documentos técnicos apenas a forma básica das condicionais são usadas, isto é, presente simples+presente simples, ou presente simples+futuro, etc.

*In case of normal servovalve failure the alternate valve **will** take over.*

Assim, apenas o uso de *will* na oração principal de um condicional expressando previsão é mostrado.

Foi possível notar a ausência das seguintes características salientes do uso de *should* no MM:

- *Will* na voz passiva e na forma negativa.
- As ocorrências de *will* na voz passiva expressam a idéia de previsão futura, indicando que uma ação/conseqüência sempre ou tipicamente acontecerá.
- *Will* expressando instruções ou obrigação.
- *Will* na oração subordinada de uma condicional, principalmente por ser raro. (7 ocorrências).
- *You*, Anáforas, unidades simples, sintagmas verbais ou preposicionais como sujeito. Essas unidades lexicais representam quase de 80% dos sujeitos de *will* no MM.
- O uso de *will* com verbos causativos ou de ocorrência anunciando uma conseqüência negativa.

- **Would**

Would não é apresentado por Shawcross e é realmente raro no MM, com somente três ocorrências.

Como visto acima, apesar de Shawcross apresentar vários usos, há outros que não são contemplados no capítulo sobre modalização. Os sentidos também não parecem ser enfatizados da forma necessária, como, por exemplo com os VMs *can*, *may* e *could*, que apresentam um importante uso de apontar sérias conseqüências no MM, mas são apresentados por Shawcross como expressando simples possibilidade.

Nota-se também o tipo de oração que o autor apresenta. As orações tendem a ser simples, com sujeito (simples ou sintagma nominal)+VM+verbo lexical. Por outro lado, as orações com modais no MM tendem a ser muito complexas, com sintagmas verbais, sintagmas preposicionais e elementos anafóricos como sujeito. Esse tipo de construção sintática impõe uma dificuldade extra na compreensão da elocução e deve ser explicitamente ensinado. Além disso, os modais fazem partes de eventos comunicativos maiores, necessitando de um contexto maior para sua compreensão, como é o caso dos *WARNING* e *CAUTION*, por exemplo:

*W_A_R_N_I_N_G_ : DO NOT USE TOO MUCH FRESH WATER TO REMOVE THE BLUE STAIN BECAUSE ADDITIONAL WATER **COULD** CAUSE POSSIBLE DAMAGE TO ELECTRICAL EQUIPMENT. THIS **CAN** RESULT IN AN UNWANTED EFFECT ON THE FLIGHT SAFETY OF THE AIRPLANE.*

O *English for Aircraft* foi escrito em 1992, quando as pesquisas com corpora eram ainda incipientes e rudimentares. Assim, é possível constatar a importância do tipo de pesquisa descritiva realizada neste trabalho para poder oferecer subsídios mais reais aos escritores de materiais didáticos com fins específicos.

8.2 SÍNTESE DO OM

Nesta seção resumo os dados do OM, ou manual de operações de aeronaves, e expresso algumas reflexões a respeito dos resultados.

O quadro 8.2.A mostra a síntese dos colocados à esquerda dos VM na voz ativa.

Quadro 8.2.A: Síntese dos colocados à esquerda na voz ativa -OM

<i>Can</i>	<i>Could</i>	<i>may</i>	<i>must</i>	<i>should</i>	<i>will</i>
IT (L1)	SUBSTANTIAL (L1)	CYCLE (L1)	CREW (L1)	CREW (L1)	VNAV (L1)
FMC (L1)	DAMAGE (L3)	SHIFT (L1)	SEGMENT (L2)	FLIGHT (L2)	FMC (L1)
APU (L1)	-----	NOTE (L2)	AIRPLANE (L1)	-----	LIGHT (L2)
-----	-----	HOLDING (L2)	ENTRIES (L1)	-----	OFFSET (L1)
-----	-----	ROLL (L2)	ROUTE (L2)	-----	BUGS (L1)
-----	-----	OIL (L3)	BATTERY (L2)	-----	WHICH (L1)
-----	-----	CREW (L1)	ANTI (L2)	-----	DESCENT (L1)
-----	-----	THAT (L1)	ICE (L1)	-----	MODE (L1)
-----	-----	MAP (L2)	FLIGHT (L2)	-----	THIS (L1)
-----	-----	PRESSURE (L1)	ENGINE (L3)	-----	LNAV (L1)

Os VMs no OM preferem a voz passiva (55,68%), assim, os VMs, com exceção de *will* apresentam mais ocorrências na voz passiva do que na voz ativa. Considerando somente os colocados à esquerda dos VMs na voz ativa no OM, observa-se que apenas poucos itens lexicais se repetem como colocados de mais de um VM. São eles:

- *Crew* 17
- *Flight (L2)* 11
- *FMC* 16

Assim, é possível verificar que não há um destaque individual. Destacam-se as siglas ou acrônimos na posição L1, como APU, FMC, LNAV e VNAV. O colocado *crew*, que em 11 situações forma o sintagma *flight crew*, parece ter função semelhante a *you* no MM. *Flight crew* refere-se aos pilotos, destinatários dos manuais e responsáveis por executar as ações, vide exemplo abaixo:

*The attitude mode can also provide heading information, but to establish compass synchronization the **crew must** manually enter the initial magnetic heading.*

Ou seja, *the crew* mencionado acima poderia ter sido substituído por *you*, pois o pronome *you* é tanto a segunda pessoa do singular quanto a segunda pessoa do plural, podendo assim se referir somente a um dos pilotos (comandante ou co-piloto ou a qualquer um dos dois tripulantes).

Oil, que é a palavra lexical mais freqüente entre os colocados à esquerda no MM, também está presente entre os colocados de *may* na posição L3 com quatro ocorrências. Como não há um único colocado com um número de ocorrências significativo, constata-se que *oil* é uma palavra significativa no vocabulário da aviação. Somente a título de curiosidade, *oil* tem 134 ocorrências no OM.

Observo agora o quadro síntese dos colocados à esquerda dos VMs na voz passiva no OM. Lembro que nesse manual, 55,68% das ocorrências modais estão na voz passiva.

Quadro 8.2.B: Síntese dos colocados à esquerda na voz passiva OM

<i>can</i>	<i>may</i>	<i>must</i>	<i>should</i>	<i>will</i>
BUT (L2)	ITEM (L1)	EXTENSIONS (L1)	-----	-----
MODE (L1)	ON (L3)	HANDRAIL (L2)	-----	-----
THRUST (L2)	MANUALLY (L3)	NAV (L2)	-----	-----
MODES (L1)	ENTRY (L1)	AIRSTAIR (L3)	-----	-----
IT (L1)	LNAV (L1)	VHF (L3)	-----	-----
WAYPOINTS (L1)	ENTRIES (L1)	PATH (L2)	-----	-----
ENTRY (L1)	DESCENT (L1)	SWITCH (L1)	-----	-----
THIS (L2)	DATA (L1)	SPEED (L1)	-----	-----
INFORMATION (L1)	LANDING (L2)	-----	-----	-----
EXECUTED (L2)	CRUISE (L3)	-----	-----	-----

A primeira observação a ser feita é de que *should* e *will* não apresentam colocados à esquerda na voz passiva. Há apenas um colocado comum a dois VMa, *entry*: oito ocorrências com *can* e sete com *may*. Nota-se a presença de siglas, como *LNAV*, *NAV* e *VHF*, a primeira também considerada colocado dos VMs na voz ativa.

Observo agora os cinco (ou seis, considerando *should*) agrupamentos que representam os colocados à esquerda dos VMs. Essa observação é útil uma vez que não há um único colocado predominante no OM, nem na voz ativa nem na voz passiva. Através da classificação dos itens lexicais à esquerda pode-se ter uma melhor visão das preferências dos VMs neste manual.

No OM, os agrupamentos lexicais mais recorrentes à esquerda dos VMs na voz ativa são:

Tabela 8.2.C: Unidades lexicais à esquerda dos VMs no OM -voz ativa

Unidades lexicais à esquerda VM- voz ativa	Número de ocorrências
<i>You</i> (L1)	1/0,2%
Anáforas (L1)	27/6%
Unidades Simples (L1)	101/ 22,3%
Sintagmas nominais	213/ 47,0%
Sintag. Verbais e prepos.	107/ 23,7%
Condicional	5/ 0,8%

Pode-se notar que o grupo do *you* apresenta apenas uma ocorrência, que é, na verdade, a única ocorrência de *you* em todo o OM. O grupo denominado “condicional” refere-se aos casos em que *should* ocupa o lugar de *if*, e possui somente cinco ocorrências. O grupo dos elementos anafóricos também não é muito freqüente como sujeito dos VMs no OM, correspondendo a apenas 6% das ocorrências. As unidades simples e os sintagmas verbais e preposicionais possuem aproximadamente o mesmo número de ocorrências, pouco mais de 20% cada. Os sintagmas nominais são, sem dúvida, o grupo mais representativo, com quase metade de todas as ocorrências na voz ativa. Os sintagmas nominais são considerados característicos da linguagem técnica, como já reportado por Sager et al. (1980). Analiso agora as unidades lexicais à esquerda na voz passiva:

Tabela 8.2.D: Unidades lexicais à esquerda dos VMs no OM- voz passiva

Unidades lexicais à esquerda VM- voz passiva	Número de ocorrências
You (L1)	-----
Anáforas (L1)	22/37,9%
Unidades Simples (L1)	149/25,7%
Sintagmas nominais	282/48,6%
Sintag. Verbais e prepos.	127/21,9%
Condiciona	-----

Na voz passiva, o grupo mais freqüente é o composto pelos sintagmas nominais (48,6%), seguido pelas unidades simples (25,7%) e pelo grupo dos sintagmas verbais e preposicionais (21,9%). Assim como na voz ativa, os elementos lexicais à esquerda na voz passiva também tendem a ser sintaticamente complexos, sendo que uma estrutura passiva já por si só sintaticamente mais complexa. Também de forma semelhante à voz ativa, os sintagmas nominais ocupam a posição de sujeito em quase a metade das ocorrências verbais.

Como foi possível observar, os modais no OM, a exemplo do MM, freqüentemente ocorrem em orações sintaticamente complexas. Assim sendo, essa característica sintática deve ser ressaltada nos materiais didáticos para pilotos.

Passo agora a analisar os colocados e as unidades lexicais à direita dos VMs no OM. O primeiro quadro mostra os colocados dos VMs na voz ativa.

Quadro 8.2.E: Colocados à direita dos VMs na voz ativa-OM

<i>can</i>	<i>could</i>	<i>may</i>	<i>must</i>	<i>will</i>
CAUSE (R1)	RESULT (R1)	RESULT (R1)	BE (R1)	BE (R1)
INCLUDE (R1)	-----	CAUSE (R1)	ON (R2)	ILLUMINATE (R1)
SUPPLY (R1)	-----	OCCUR (R1)	CONTAIN (R1)	DISPLAY (R1)
ALSO (R1)	-----	EXCEED (R1)	IN (R2)	APPEAR (R1)
POWER (R2)	-----	NOT (R1)	REMAIN (R1)	REMAIN (R1)
OPERATE (R1)	-----	TAKE (R1)	MANUALLY (R1)	NOT (R1)
OCCUR (R1)	-----	ALSO (R1)	THAN (R3)	OCCUR (R1)
PROVIDE (R1)	-----	INCREASE (R1)	DURING (R3)	DISENGAGE (R1)
EITHER (R2)	-----	TO (R3)	-----	RESULT (R1)
BE (R1)	-----	IN (R2)	-----	CHANGE (R1)

Observando-se o quadro acima, nota-se que alguns verbos aparecem como colocados de mais de um VM. Listo abaixo os verbos preferidos dos VMs no OM e mostro com quantos VMs eles ocorrem. Nesta listagem, incluo também os resultados dos VMs que, devido a sua baixa ocorrência, não estão inseridos no quadro acima:

- *Be* 94 6VMs
- *Result* 19 3VMs
- *Cause* 16 2VMs
- *Occur* 16 3VMs

Nota-se que *be* é o preferido. Ele ocorre com seis VMs. Os outros três verbos recorrentes com mais de um modal são causativos (*cause*), ou de ocorrência (*result e occur*) e são colocados dos VMs que expressam possibilidade (*can, could e may*) e com o VM que expressa previsão (*will*). Esse fato pode apontar que, de alguma forma, o sentido, ou função desses VMs é similar.

A seguir verifico os colocados à direita dos VMs na voz passiva.

Quadro8.2.F Colocados à direita dos VMs na voz passiva-OM

<i>can</i>	<i>may</i>	<i>must</i>	<i>should</i>	<i>will</i>
SELECTED (R1)	FOUND (R1)	VERIFIED (R1)	USED (R1)	DISPLAYED (R1)
BY (R2)	ENTERED (R1)	STOWED (R1)	NOT (R1)	IN (R3)
USED (R1)	MADE (R1)	ENTERED (R1)	IN (R2)	
REENGAGED (R1)	MANUALLY (R2)	TUNED (R1)	-----	-----
ALSO (R1)	ACCOMPLISHED (R1)	BEFORE (R2)	-----	-----
FROM (R2)	USED (R1)	IN (R2)	-----	-----
OPENED (R1)	BY (R2)	SELECTED (R2)	-----	-----
ON (R2)	ALSO (R1)	-----	-----	-----
MADE (R1)	SELECTED (R1)	-----	-----	-----
ENGAGED (R1)	OPERATED (R1)	-----	-----	-----

O verbo *used* aparece como colocado de três VMs, com um total de 31 casos. O verbo *selected* também aparece como colocado de três VMs, com um total de 36 ocorrências. Nota-se também a presença da preposição *by* como colocado de dois VMs, *may* e *can*, com 48 ocorrências. A preposição *by* coloca-se com cerca de 10% das ocorrências modais na voz passiva no OM. A presença de *by* parece indicar uma preferência por mostrar o agente da ação de forma menos direta do que seria mostrado através da opção da voz ativa. Lembro que *be* não é considerado colocado dos VMs na voz passiva por ser elemento obrigatório da estrutura. Entretanto, somando-se as ocorrências de *be* na voz ativa (94) com as ocorrências da voz passiva (575), constata-se que *be* acompanha um VM em 669 ocorrências ou 64% das ocorrências, sendo o item lexical com maior grau de atração aos VMs no OM. Os pilotos, destinatários desse manual, devem ser explicitamente ensinados a respeito desse fato para saber distinguir entre *be* como verbo auxiliar:

*Landing gear doors **should be** free of snow or ice.*

E *be* como verbo principal, por exemplo:

*The flap position indicator and leading edge devices annunciator panel **should be** closely observed for positive movement.*

Na próxima tabela mostro os tipos de verbo que seguem os VMs na voz ativa.

Tabela 8.2.G verbos lexicais na voz ativa-OM

Tipos de Verbos	<i>Will</i>	<i>Should</i>	<i>Must</i>	<i>May</i>	<i>Could</i>	<i>Can</i>
De atividade	60/44,1 %	13/ 50%	14/ 17,7	39/ 33,9	-----	52/ 61%
De comunicação	2/ 1,4 %	1/ 3,8%	-----	-----	-----	2/ 2,4%
Mentais	1/0,7 %	-----	2/ 2,3	01/0,9	-----	1/ 1,2%
Causativos	11/ 8,2 %	02/ 7,6%	1/ 1,3	10/8,7	2/18,2%	12/ 14,1%
De ocorrência	19/ 14,0 %	03/ 11,5%	-----	28/ 24,3	7/63,6%	1/ 1,2%
De existência	34/25,0 %	5/ 19,2%	57/ 72,4	35/ 30,4	2/18,2%	16/ 18,9%
De aspecto	9/6,6 %	02/ 7,6%	5/ 6,3	2/ 1,8	-----	1/ 1,2%

Os totais gerais dos verbos colocados à direita dos VMs no MM na voz ativa são:

- De atividade 178
- De existência 149
- De ocorrência 58
- Causativos 38
- De aspecto 19
- Mentais 05
- De comunicação 05

Na tabela que segue apresento os verbos que seguem os VMs na voz passiva.

Tabela 8.2.H Verbos lexicais na voz passiva-OM

Tipos de Verbos	<i>Will</i>	<i>Should</i>	<i>Must</i>	<i>May</i>	<i>Can</i>
De atividade	22/81,5 %	37/ 72,4%	68/ 76,4	113/ 66,1	181/ 74,8%
De comunicação	-----	-----	01/ 1,1	7/ 4,1	1/ 0,4%
Mentais	1/3,7%	8/ 15,6%	12/ 1,3	31/ 18,1	17/7%
Causativos	1/3,7%	01/ 2,0%	2/ 2,2	10/ 5,8	1/ 0,4%
De ocorrência	1/3,7%	-----	-----	1/ 0,6	15/ 6,2%
De existência	1/3,7%	02/ 4,0%	02/ 2,2	-----	4/ 1,7%
De aspecto	1/3,7%	03/ 6,0%	04/ 4,4	9/ 5,3	23/ 9,0%

Os totais gerais dos verbos colocados à direita dos VMs no OM na voz passiva são:

- De atividade 421
- Mentais 69
- De aspecto 40
- De ocorrência 17
- Causativos 15
- De existência 09
- De comunicação 09

Nota-se que os verbos de atividade detêm a preferência tanto na voz ativa quanto na voz passiva. Lembro que a voz passiva é a estrutura predominante no OM, com mais de 55% das ocorrências. Os verbos de atividade são encontrados como colocados de todos os VMs. Uma das razões para isso é que, devido à característica deôntica do texto estudado, isto é, um manual técnico que fornece instruções sobre como operar uma aeronave. Dessa forma, é natural que os modais sejam seguidos por verbos que indiquem alguma ação a ser realizada.

Por outro lado, principalmente os causativos na voz ativa, mas também os de ocorrência em menor grau, mostram uma clara preferência pelos VMs que estão no eixo de possibilidade ou previsibilidade. *May* (possibilidade) é acompanhado por um verbo de ocorrência em 24% dos casos na voz ativa. Esse fato pode ser explicado com base no sentido e na prosódia semântica dos causativos e os verbos de ocorrência, que é negativa. Vide exemplo abaixo:

Above 30,000 feet, without fuel pump pressure, thrust deterioration or engine flameout may occur.

Os verbos de existência estão em segundo lugar na voz ativa, cujo exemplar mais freqüente é *be*. Os verbos de existência ou de relacionamento (mesma categoria, conforme apresentado no capítulo 6) expressam o estado que a situação está ou deve estar para que outra ação possa ser realizada, como por exemplo:

*Relative bearings indicated by pointers **may be** correct if the receiver is operating..*

Os verbos mentais ocupam o segundo lugar na voz passiva. Os principais verbos desta categoria encontrados no OM são *see*, *find* e *hear*, vide exemplo abaixo:

*A tone may **be heard** through a headset plugged into the headset jack.*

Os verbos de aspecto estão em terceiro lugar nos casos de voz passiva. Os verbos mais freqüentes são *finish* e *stop*.

De uma forma geral, os VMs apresentam menos usos quando comparados aos vários usos e sentidos possíveis de cada VM (conforme capítulo 4). Apresento sucintamente os VMs e seus sentidos no OM no quadro abaixo.

Quadro 8.2.I Síntese dos Sentidos-OM

VM	SENTIDOS
<i>Can</i>	Possibilidade (327)
<i>Could</i>	Possibilidade (12)
<i>May</i>	Possibilidade (288)
<i>Might</i>	-----
<i>Must</i>	Obrigaç�o (165), Proibiç�o (03)
<i>Shall</i>	Previs�o (02)
<i>Should</i>	Obrigaç�o (18), Probab (1), Cond (substituindo if) (5), Presente simples (2)
<i>Will</i>	Previs�o (163)
<i>Would</i>	Possibilidade (6)

Pode-se constatar que o principal uso dos VMs no OM   o de expressar possibilidade, com 633 ocorr ncias. O segundo uso   o de obrigaç o, com 183 ocorr ncias. O sentido de previs o vem em seguida com 165 casos, probabilidade com um e condicional com cinco. H  ainda o uso de *should* que praticamente n o tem valor sem ntico, com duas ocorr ncias. Pode-se assim dizer que os usos dos modais neste manual s o realmente reduzidos quando comparados ao ingl s geral.

Ainda que de fora assistem tica, foi poss vel observar que h  a recorr ncia de modais em uma mesma oraç o ou ato comunicativos, vide exemplo:

*Brake temperature levels **may** be reached which **can** cause the wheel fuse plugs to melt and deflate the tires*

*Each class of fire calls for specialized action. Using the wrong extinguisher **may** do more harm than good. For your own protection, you **should** know these basic types, how to use them, and why.*

Ambas as situaç es acima apresentam dois modais, sendo que a primeira apresenta dois VMs na mesma oraç o. Como mencionado, n o foi realizado um levantamento sistem tico para verificar quantas vezes isso ocorre, mas pode-se sugerir como sendo mais uma caracter stica peculiar do uso dos VMs no MM e deveria, assim, ser inserido nos materiais did ticos.

9 CONCLUSÕES

No final desta tese retomo os objetivos e apresento as respostas que pude depreender das observações do corpus. Logo após retomo a hipótese inicial do trabalho.

O objetivo geral desta pesquisa foi:

- **Analisar a incidência e a funcionalidade dos verbos modais no segmento da linguagem da aviação representado pelo texto do tipo manual técnico operacional e de manutenção para oferecer subsídios para a elaboração de material didático.**

Para alcançar esse objetivo, segui o seguinte roteiro metodológico:

- Contraste das frequências de cada VM entre o corpus de estudo (AC), o corpus de referência (BNC) e o corpus de contraste (OSC), observando se há diferenças significativas no uso dos VMs**

Nesta seção foi feito um contraste quantitativo entre os três corpora AC x BNC x OSC. Foi verificado que **há uma diferença significativa com relação à incidência de cada VM em cada corpus**. O BNC, por ser um corpus de inglês geral, apresenta uma distribuição mais uniforme dos VMs. Já os dois corpora especializados, AC (manuais de aviação) e OSC (manuais de sistemas operacionais para computadores pessoais), apresentam alguns VMs com incidência bem mais alta, enquanto outros estão virtualmente ausentes. Apresento agora uma tabela síntese mostrando quais modais são mais usados em cada corpus e seu valor normalizado, isto é, o número de *tokens* a cada mil palavras.

Tabela 9.A: Frequência dos VMs por ordem decrescente nos três corpora

AC	Tokens ‰	OSC	Tokens ‰	BNC escrito	Tokens ‰
<i>Can</i>	2,6	<i>Can</i>	8,1	<i>Will</i>	3,2
<i>Will</i>	1,3	<i>Will</i>	3,0	<i>Would</i>	2,8
<i>Must</i>	1,1	<i>May</i>	0,9	<i>Can</i>	2,5
<i>May</i>	1,1	<i>Might</i>	0,7	<i>Could</i>	1,7
<i>Should</i>	0,4	<i>Should</i>	0,6	<i>May</i>	1,2
<i>Could</i>	0,1	<i>Would</i>	0,4	<i>Should</i>	1,1
<i>Shall</i>	0,1	<i>Must</i>	0,7	<i>Must</i>	0,8
<i>Would</i>	0,1	<i>Could</i>	0,2	<i>Might</i>	0,6
<i>Might</i>	0,0	<i>Shall</i>	0,0	<i>Shall</i>	0,2

Analisando o quadro, pode-se notar que o AC e o OSC são mais semelhantes entre si quando comparados ao BNC. Com relação aos modais individuais, o *would*, que é o segundo

mais freqüente no BNC, com uma incidência de 2,8 a cada mil palavras apresenta um índice normalizado de apenas 0,4 no OSC e 0,1 no AC. Conclui-se, a partir disso, que o *would* não representa a linguagem da aviação. O *can* desponta como o mais freqüente entre os dois corpora especializados. Esse resultado sugere que a principal função dos VMs em um texto de manual técnico é a de mostrar possibilidades, provavelmente fruto de ações tomadas por quem está operando ou consertando uma máquina, ou por problemas não esperados.

ii) Contraste das freqüências das estruturas sintáticas em que os VMs se encontram entre o corpus de estudo (AC), o corpus de referência (BNC) e o corpus de contraste (OSC), observando se há diferenças significativas no uso das estruturas.

Foi verificado que **há diferenças nas preferências das estruturas**. Essas diferenças podem ser observadas através da freqüência de ocorrência da estrutura modal+voz passiva. Retomo a tabela que mostra os contrastes entre os três corpora. Para melhor visualização, apresento apenas as duas estruturas mais freqüentes.

Tabela 9.B: Estruturas Modais AC x OSC x BNC

		AC	OSC	BNC escrito
2	Modal + infinitivo (modal + voz ativa)	67,3 %	76,05%	74,5%
3	Modal + <i>be</i> + passado particípio (modal+ voz passiva)	32,4%	23,13%	16,6%

Nota-se que o BNC apresenta menos ocorrências na voz ativa do que o OSC, mas como o BNC apresenta também ocorrências nas outras estruturas¹⁴¹ (apresentadas na seção 7.2), o OSC também apresenta mais ocorrências na voz passiva. O AC é o corpus que mais apresenta ocorrências na voz passiva, 32,4%, e menos ocorrências na voz ativa, 67,3%. Constata-se, assim, que a elevada incidência dos VMs na voz passiva pode ser considerada uma característica peculiar da linguagem técnica dos manuais de aviação.

iii) Verificação das diferenças quanto à freqüência de cada VM e de cada estrutura em cada um dos três manuais que compõe o corpus de estudo.

Foi constatado que **os três manuais apresentam diferenças entre si**. Lembro que o MM é o manual de manutenção destinado a mecânicos de aeronave, enquanto o OM e o QRH

¹⁴¹ Na seção 7.2 foi apresentada uma tabela com nove estruturas modais nas quais os VM podem ocorrer e o BNC mostrou a distribuição mais uniforme entre as estruturas.

são destinados aos pilotos. O *can* é o VM mais freqüente no MM e no OM e *may* o mais freqüente no QRH. O *may* apresenta uma incidência relativa bem mais baixa no MM do que nos outros dois corpora, com apenas 6,36% das ocorrências modais. No OM, *may* é o segundo mais freqüente com 27,5% e no QRH, *may* é o mais freqüente, com 43,71% das ocorrências modais. O *must* é mais freqüente no MM em relação aos outros dois corpora, com 18,76% no MM, 16,42% no OM e 5,52% no QRH.

Quanto às estruturas sintáticas, apresento a tabela abaixo:

Tabela 9.C: Estruturas Modais no AC

	Estrutura	MM	OM	QRH
2	Modal + infinitivo	83,51	44,03	49,75
3	Modal + be + past participle (Modal with Passive)	16,32	55,68	49,25

Nota-se aqui uma grande diferença entre as freqüências das estruturas sintáticas. O MM apresenta cinco vezes mais ocorrências dos VMs na voz ativa do que na voz passiva, enquanto o OM apresenta mais ocorrências na voz passiva do que na voz ativa. O QRH apresenta praticamente o mesmo número de ocorrências na voz ativa e na voz passiva. Esse resultado mostra que o valor obtido na etapa metodológica acima (tabela 9.B), com 32,4% de ocorrências de modais na voz passiva no AC, não representa a realidade de nenhum dos três manuais, pelo menos no que diz respeito aos modais e suas estruturas sintáticas e mostra a relevância de estudar os três manuais separadamente.

iv) Identificação das colocações de cada VM em cada um dos corpora.

A partir desta etapa foram analisados somente dois manuais, o MM e o OM. A decisão de analisar somente esses dois manuais recaiu sobre a dimensão dos mesmos. A metodologia da LdC mostra-se produtiva quando é possível extrair os colocados (4 ocorrências e LL de 6,63 ou mais). Como o QRH é bem menor do que os outros dois corpora, aproximadamente 10% do tamanho, esse tipo de análise não se mostrou produtiva. Além disso, a linguagem dos manuais técnicos para pilotos já está representada no OM.

Lembro que o MM mostra uma preferência pela estrutura modal+voz ativa e que o OM mostra uma preferência por modal+ voz passiva. Foram levantadas as colocações à esquerda e à direita de cada um dos VMs.

Foi observada uma grande diferença na quantidade de colocados em cada um dos subcorpora na voz ativa. Os colocados¹⁴² no MM apresentam uma frequência, ou força de colocação bem superior ao OM, por exemplo, o pronome *you* apresenta 190 ocorrências como colocado de can, ao passo que o colocado mais forte de can no OM apresenta somente 12 ocorrências (*mode*).

No MM, as ocorrências mais marcantes são o pronome *you*, os elementos anafóricos e a palavra lexical *oil*. No OM notou-se a presença de *flight crew*, que se aplica à segunda pessoa do plural, ou seja, qualquer um dos dois pilotos, equivalendo assim ao pronome *you* no MM. Foi também observada a grande incidência de siglas e acrônimos no OM como colocados dos VMs na posição L1: FMC, LNAV, APU, entre outras.

Na próxima tabela mostro os grupos que mais ocupam a posição de sujeito nos dois manuais na voz ativa.

Tabela 9.D: Unidades lexicais à esquerda-Voz ativa

Unidades lexicais à esquerda VM- voz ativa	MM	OM
<i>You</i> (L1)	405/ 28,1%	1/0,2%
Anáforas (L1)	183/ 12,7%	27/6%
Unidades Simples (L1)	225/ 15,6%	101/ 22,3%
Sintagmas nominais	272/ 18,8%	213/ 47,0%
Sintag. Verbais e prepos.	355/ 24,6%	107/ 23,7%
Condicional	03/ 0,2%	5/ 0,8%

A principal diferença entre os dois corpora é a presença marcante do pronome *you* como o principal colocado à esquerda no MM e a quase ausência desse pronome no OM. Há também uma maior incidência de anáforas na posição L1 no MM, o que é de certa forma surpreendente, uma vez que existe um esforço oficial através do SE da AECMA em tornar a linguagem dos manuais técnicos de manutenção mais fáceis de serem compreendidos. Um elemento anafórico na posição L1 não é de forma alguma fácil de ser compreendido. Após essas duas constatações, torna-se natural que outros grupos sejam mais frequentes no OM. O que mais se destaca são os sintagmas nominais, que ocupam a posição de sujeito em aproximadamente 50% das ocorrências na voz ativa no OM.

Apresento agora os agrupamentos à esquerda na voz passiva.

¹⁴² Lembro que são considerados colocados os itens lexicais localizados em um horizonte de três casas à direita e três à esquerda, com frequência mínima de quatro e valor LL de 6,63 ou maior.

Tabela 9.E: Unidades lexicais à esquerda-Voz passiva

Unidades lexicais à esquerda VM- voz passiva	MM	OM
<i>You</i> (L1)	-----	-----
Anáforas (L1)	38/ 13,7%	22/3,8%
Unidades Simples (L1)	59/ 21,3%	149/25,7%
Sintagmas nominais	100/ 36,1%	282/48,6%
Sintag. Verbais e prepos.	80/ 28,9%	127/21,9%
Condicional	-----	-----

Entre as ocorrências modais na voz passiva, nota-se que a situação é semelhante à voz ativa. Os sintagmas nominais também são os preferidos no OM e os elementos anafóricos também são mais freqüentes no MM em relação ao OM.

De uma forma geral, conclui-se que os VMs encontram-se freqüentemente em estruturas sintáticas bastante complexas. Essa constatação é muito relevante tendo em vista o objetivo final desta pesquisa que é o de obter subsídios para a elaboração de materiais didáticos para fins específicos. Não é suficiente apenas apresentar o VM e apontar o seu uso, isto é, se o modal expressa uma possibilidade, uma proibição ou uma obrigação. Um material didático de qualidade deve apresentar o modal dentro do seu contexto real de uso para que os alunos os possam compreender quando lerem os manuais. Assim, a apresentação de exemplos somente com unidades simples, ou *you* como sujeitos, não irá de forma alguma capacitar os alunos a entenderem as orações com modais nos manuais, já que não representam a maioria dos casos.

Quanto aos colocados à direita foi possível notar que alguns verbos se repetem nos dois manuais, como *be*, *cause* e *result*¹⁴³. Isso sugere que os VMs que têm esses verbos como colocados apresentam sentidos semelhantes. Como já visto, os VMs que se colocam com os verbos *cause* e *occur*, não expressam uma mera possibilidade, mas anunciam uma consequência negativa fruto de uma ação mal realizada ou de um problema no funcionamento da aeronave.

Apesar de *by* aparecer como colocado dos VMs na voz passiva em ambos os subcorpora, nota-se que a grande maioria das ocorrências de modal+voz passiva acontece sem a presença do agente que executa ou deve executar a ação expressa pelo verbo principal. Essa informação é importante uma vez que muitos livros-texto e gramáticas da língua inglesa

¹⁴³ Apesar de haver uma incidência muito mais alta desses verbos no OM.

apresentam a voz passiva geralmente acompanhada pela preposição *by*. Entretanto, esse não é o caso no AC e tais resultados devem ser considerados nos exemplos e exercícios.

Apresento a seguir os verbos principais classificados conforme seu domínio semântico, primeiro os verbos na voz ativa e depois os verbos na voz passiva.

Tabela 9.F: Unidades lexicais à direita-Voz ativa

Tipo de verbo MM	Ocorrências	Tipo de verbo OM	Ocorrências
De atividade	511	De atividade	178
Causativos	420	De existência	149
De existência	237	De ocorrência	58
De ocorrência	180	Causativos	38
Mentais	76	De aspecto	19
De aspecto	14	Mentais	05
De comunicação	02	De comunicação	05

Tabela 9.G: Unidades lexicais à esquerda-Voz passiva

Tipo de verbo MM	Ocorrências	Tipo de verbo MM	Ocorrências
De atividade	221	De atividade	421
Mentais	25	Mentais	69
Causativos	17	De aspecto	40
De aspecto	14	De ocorrência	17
De existência	07	Causativos	15
De comunicação	02	De existência	09
De ocorrência	01	De comunicação	09

É possível observar que os verbos de atividade são os mais frequentes colocados dos VMs. Na verdade, Biber et al. (1999) constataram que os verbos de atividade são os mais frequentes no inglês geral também. Neste corpus (AC), a presença marcante dos verbos de atividade justifica-se devido ao caráter deôntico do texto e também dos VMs. Não há ocorrências de modais com sentido epistêmico, que nos termos de Lyons (1977) significa “a opinião ou atitude com relação à proposição que a sentença expressa ou a situação que a proposição descreve” (p. 452). Em um manual técnico não parece haver lugar para esse tipo de modalização epistêmica.

A presença de VM + *be* deve ser destacada. Como já mencionado, essa é a associação mais frequente entre todas as associações no AC, com 434 ocorrências no MM e 669 ocorrências no OM. Na voz passiva, *be* não foi considerado colocado, por fazer parte da estrutura, mas ele está lá, ao lado do VM, de forma similar aos casos de voz ativa. É de

extrema importância que essa característica seja explicitamente apresentada nos materiais didáticos. Pilotos e mecânicos devem ser ensinados a distinguir entre a estrutura modal+*be* (voz ativa) e modal + *be* (voz passiva) para, ao lerem os manuais (de manutenção ou operação), saibam exatamente o que está sendo instruído ou indicado através do modal.

As categorias de verbos principais que ocorrem com mais frequência na estrutura modal + voz ativa, são diferentes da estrutura modal + voz passiva. Esse fato aponta e justifica a análise separada das estruturas, pois, assim, pode-se elaborar materiais que contemplem os verbos que realmente ocorrem em uma estrutura com um sentido específico de forma representativa.

Os verbos causativos e os de ocorrência são muito mais frequentes no MM do que no OM. Uma possível explicação para essa diferente incidência pode ser atribuída às condições de trabalho e local de consulta dos manuais. Sem ter o propósito de avaliar o grau de profissionalização de qualquer um dos destinatários dos manuais, sejam eles mecânicos ou pilotos, é fundamental lembrar que uma ação mal realizada pelo mecânico, pode colocar muitas vidas em risco, mas não necessariamente a dele, ao contrário dos pilotos, cujas ações têm reflexo na sua própria vida. Assim, as possíveis trágicas consequências de uma execução ineficaz, devem ser explicitadas de forma mais marcante no MM. Além disso, notou-se que os causativos e os verbos de ocorrência são muito recorrentes nas seções *WARNING* e *CAUTION* no MM. Essas seções recebem um tratamento diferenciado entre os dois manuais, como pode ser visto nos exemplos abaixo:

No OM, para pilotos, uma seção denominada *WARNING* ou *CAUTION*, aparece graficamente assim:

*WARNING: If engine No. 1 was started using a ground air source, to minimize the hazard to ground personnel, the external air **should** be disconnected and engine No. 2 started using the Engine Crossbleed Start procedure.*

É possível observar que a seção não é destacada através de caixa alta (ou letra maiúscula) nem recebe qualquer outro tipo de destaque. Esse tipo de destaque é geralmente considerado importante nos manuais técnicos¹⁴⁴ (BLAKE e BLY, 1993). Da mesma forma, não há uma tendência dos VMs de possibilidade, como *can*, *may* ou *could* ocorrerem nessas seções.

¹⁴⁴ Vide normas para redação de documentação técnica na seção 3.2.2.2.

Por outro lado, no MM, as seções *WARNING* e *CAUTION* aparecem da seguinte forma:

W_A_R_N_I_N_G : MAKE SURE EACH TUBE AND THE PORT FITTINGS HAVE TAGS TO IDENTIFY THE CORRECT INSTALLATION LOCATION. IF YOU DO NOT PUT TAGS ON THE TUBES AND PORT FITTINGS, CROSS-CONNECTION OF THE TUBES CAN OCCUR DURING INSTALLATION. IF THE TUBES ARE CROSS-CONNECTED, UNINTENDED OPERATION OR MALFUCTION OF AIRPLANE SYSTEMS CAN RESULT AND CAUSE INJURY TO PERSONS AND DAMAGE TO EQUIPMENT.

Pode-se notar que essas seções são bastante destacadas e tendem a ser longas, explicando os procedimentos e as conseqüências mais detalhadamente. Outra explicação, que não necessariamente anula a anterior, é a de que são necessários muitos anos de treinamento teórico e prático para que um piloto possa voar um BOEING. Dessa forma, esse profissional conhece melhor as causas e as conseqüências de ações e problemas com a aeronave. Os mecânicos fazem um curso técnico bem direcionado e detalhado, mas em menos de um ano já estão capacitados a trabalhar na manutenção das aeronaves.

Foi também visto que o MM é escrito baseado no guia SE, que estabelece regras explícitas a respeito da redação do manual. Essas regras são apenas parcialmente seguidas, mas resultam em grandes diferenças com relação ao OM em alguns aspectos. Os mais marcantes parecem ser a redução acentuada de casos de VM+voz passiva com a conseqüente inclusão do pronome pessoal *you* como sujeito na voz ativa. O OM apresenta apenas uma ocorrência do pronome pessoal *you* e inclui, de forma marcante, as estruturas VM+voz passiva. Essa constatação não é somente baseada no contraste MM x OM com relação à voz passiva, mas também com outros corpora. O OM apresenta 55% dos VMs na estrutura VM+voz passiva. Nenhum outro corpus pesquisado apresenta resultado igual ou semelhante. O OSC, que também é um manual técnico, tem 23% dos VMs nessa estrutura, ao passo que o BNC apresenta apenas 16,6%. Observa-se ainda a maior freqüência relativa do VM *may* no OM em relação ao MM, com 27,5% e 6,36% respectivamente. O *may* é considerado um correspondente mais formal de *can* para expressar possibilidade (COATES, 1983). Dessa forma, conforme Brown e Levinson (1987), pode-se aventar a hipótese de haver uma tendência a mitigar as instruções e as possibilidades quando os destinatários dos manuais são os pilotos. Para Pilegaard, 1997, no discurso escrito, “a polidez resulta não somente do uso de estratégias individuais no nível da sentença. Mais precisamente, resulta de um número de atos ilocucionários complexamente expressos em conjunto em um texto” (p. 224).

Retomo agora o quadro apresentado no Capítulo 4 que sintetiza os usos dos VMs conforme a revisão da literatura apresentada naquele capítulo. Contrasto essa síntese com os usos dos VMs no AC:

Quadro 9.H: Síntese das formas e usos dos VMs- Geral x AC

Uso	Forma- Inglês Geral	Forma- AC
Condicional Substituto de <i>If</i>	<i>should</i>	<i>Should</i>
Habilidade/capacidade	<i>can/could/ will</i>	-----
Hipótese/probabilidade	<i>could/might/should/would</i>	<i>Should</i>
Obrigaç�o, ordem, comando	<i>may/must/shall/should/will</i>	<i>Must/shall/should/will</i>
Necessidade Epist�mica /Infer�ncia	<i>can 't/mustn 't /must/ /will/should</i>	-----
Pedir ou conceder permiss�o	<i>can/may/could/might/would</i>	-----
Possibilidade De�ntica/Epist�mica	<i>can/could/may/might/</i>	<i>Can/could/may/might/</i>
Proibi�o	<i>cannot (can 't)/may not/must not</i>	<i>Cannot/must not</i>
Previs�o/ voli�o/ inten�o	<i>shall/will/would</i>	<i>Shall/ will/would</i>
Quase subjuntivo/ subjuntivo	<i>Should</i>	<i>Should</i>
Sugest�o	<i>can/could//should/would</i>	-----
Voli�o	<i>shall/will/would</i>	-----

Como esperado, o AC apresenta uma simplifica o nos usos dos modais quando comparados aos usos da l ngua geral. V rios usos n o s o contemplados no AC, como habilidade, infer ncias, permiss o, sugest o ou voli o. Ou seja, a l ngua t cnica dos manuais de avia o   simplificada em compara o ao ingl s geral. A maioria dos modais apresenta apenas um uso, ou pelo menos um uso predominante.

Na se o 4.2, Coates (1983) classifica a freq ncia de uso dos VMs com rela o ao seu significado ou fun o nos corpora de ingl s geral e mostra o resultado, em ordem decrescente (COATES, 1983, p. 25). Os usos mais freq entes s o *will* (previs o), *would* (previs o hipot tica) e *can* (possibilidade). A autora apresenta ainda outros 18 usos e seus modais correspondentes.

Apresento agora um quadro semelhante com rela o aos VMs e seus usos no AC:

Quadro 9.I: Freqüência dos diferentes modais/usos no AC

VM	Uso
<i>Can</i>	Possibilidade (1133)
<i>Will</i>	Previsão (505)
<i>Must</i>	Obrigaç�o (467)
<i>May</i>	Possibilidade (398)
<i>Could</i>	Possibilidade (51)
<i>Should</i>	Obrigaç�o (18)
<i>Must</i>	Proibiç�o (24)
<i>Shall</i>	Obrigaç�o (18)
<i>Will</i>	Obrigaç�o (12)
<i>Should</i>	Probabilidade (9)
<i>Should</i>	Condicional (8)
<i>Would</i>	Possibilidade (8)
<i>Might</i>	Possibilidade (3)
<i>Cannot</i>	Proibiç�o (3)
<i>Shall</i>	Previs�o (2)

Comparando-se o quadro acima com o proposto por Coates (1993) apresentado no cap tulo 4, nota-se que h  uma reduç o de usos e formas. Mesmo assim, e apesar de esforç os padronizadores, n o h  uma univocidade total. Um uso   expresso por mais de um modal, e um mesmo modal pode expressar mais de um uso. Dessa forma, os materiais did ticos devem apresentar esses diferentes usos, ainda que raros.

Como j  mencionado anteriormente, um manual t cnico   um texto de ntico, no qual instruç es s o repassadas e devem ser seguidas sem questionamento. Entretanto, o uso mais freq ente dos VMs no AC n o   o de obrigaç o ou forte recomendaç o, mas o de possibilidade. No AC esse uso   expresso por *can*, *may*, *could*, *would* e *might*, com 1593 ocorr ncias. Obrigaç o (ou instruç o/forte recomendaç o)   o segundo uso mais freq ente, expresso atrav s de *must*, *should*, *shall* e *will*. Logo ap s vem previs o, atrav s de *will* e *shall*. Proibiç o tem 27 ocorr ncias com *must not* e *cannot*.

As instruç es podem ser expressas atrav s de v rios recursos ling sticos, al m dos modais de nticos. Uma dessas formas   o imperativo. Isso provavelmente explica a maior incid ncia relativa do uso de possibilidade, pois o sentido de possibilidade, especialmente na linguagem escrita, acontece principalmente atrav s dos auxiliares modais, como *may* e *can* (CRYSTAL, 1985; LYONS, 1977; GABRIELATOS & MCENERY, 2005).

Passo agora   hip tese de investigaç o, lanç ada no começo deste trabalho.

O uso dos VMs é peculiar no corpus de estudo, tanto em relação à frequência de uso e distribuição quanto em relação aos sentidos e funções.

Parcialmente confirmada. As frequências de uso e distribuição são realmente diferenciadas tanto em relação ao BNC que é um corpus representativo de inglês geral, quanto ao OSC, que é um corpus composto de manuais de sistemas operacionais de computadores pessoais.

Quanto aos usos e funções, foi constatado que o AC apresenta uma quantidade menor de usos, que cada uso tende a ser expresso por menos VMs e que cada VM expressa menos funções do que na língua geral. Assim, essas constatações podem, de certa forma, ser consideradas peculiares da linguagem técnica dos manuais de aviação.

Os resultados obtidos nesta pesquisa mostram que somente uma abordagem fraseológica pode desvendar a riqueza de detalhes necessária para a melhor compreensão de um item lexical. Conforme Hoffmann¹⁴⁵ (1998[2007]), é importante escolher e delimitar o material de ensino de forma que o aluno possa aprender, em menor tempo, a quantidade máxima possível de conhecimentos. O autor considera que esse *minimum* de conteúdo, necessita conter as ocorrências lexicais e gramaticais mais úteis e que o desvendar desse *minimum* constitui o cerne das pesquisas voltadas ao ensino de línguas estrangeiras com propósitos específicos.

Como perspectivas para pesquisas futuras, julgo necessário estudar, em outros manuais da BOEING e também em manuais de outros fabricantes, as unidades aqui analisadas, de modo a conhecer com mais precisão seu funcionamento nos manuais de aviação. Igualmente necessários são estudos que tenham como objeto de análise outras unidades lexicais no mesmo corpus como forma de colher subsídios para a elaboração de materiais didáticos representativos da linguagem da aviação. Há muitos outros fatores que podem ser observados.

A continuidade deste trabalho contribuirá de diferentes formas. Cooperará para que os escritores de material didático dêem mais atenção e dediquem maior cuidado com o estudo das unidades lexicais a serem apresentadas nos livros didáticos com propósitos específicos. Auxiliará os professores de ESP, e principalmente os que ensinam inglês para aviação a

¹⁴⁵ Capítulo 2, seção 2.8.

conhecerem os possíveis padrões de uso dos VMs para que saibam como ensiná-los. E, conseqüentemente, beneficiará os futuros alunos que utilizarão esses materiais.

10 REFERÊNCIAS

- AECMA:SIMPLIFIED ENGLISH. A guide for the preparation of Aircraft Maintenance Documentation in the International Aerospace Maintenance Language, 2004.
- ALTENBERG, B. *On the Phraseology of Spoken English: The Evidence of Recurrent Word-Combinations. Phraseology*. Ed. A.P.Cowie. Oxford: Clarendon Press. p. 101, 1998.
- AUSTIN, J. L. *Quando Dizer é fazer: palavras e ação*. (Título original: *How to do things with words*. Oxford: OUP, 1962) Porto Alegre: Artes Médicas, 1990.
- AUGER, P. (1993) Pour un modèle variationniste de l'implantation terminologique dans les entreprises au Québec. A: *Colloque sur la problématique de l'aménagement linguistique: Actes du Colloque*. Quebec: OLF, 483-493, 1994.
- AZEREDO, S. *Expressões Enunciadoras de Paráfrase em Manuais Acadêmicos de Química: Um Estudo Baseado em Corpus*. Dissertação de Mestrado. Porto Alegre: Universidade Federal do Rio Grande do Sul, 2007.
- BEAUGRANDE, R. & DRESSLER, W. (1981[ed.rev. 2002]) *Introduction to Text Linguistics*. In: <http://beaugrande.com/introduction_to_text_linguistics.htm>. Acessado em: 17/05/2006
- BERBER SARDINHA, A.P. *Linguística de Corpus: histórico e Problemática*. D.E.L.T.A., Vol.16 N 2 :323-367, 2000.
- BERBER SARDINHA, A.P. *Linguística de Corpus*. Barueri: Manole, 2004.
- BEVILACQUA, C. R. *Unidades Fraseológicas Especializadas: Novas Perspectivas para sua Identificação e tratamento*. In: KRIEGER, M. G.; BECKER, A. M. M. *Temas de terminologia*. São Paulo: FFCH/USP, 2001.
- BIBER, D. Representativeness in corpus design. *Literacy and Linguistic Computing* 8, 243-257, 1993.
- BIBER, D., CONRAD, S.; REPPEN, R. *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge University Press, 1998.
- BIBER, D.; JOHANSSON, S.; LEECH, G.; CONRAD, S.; FINEGAN, E. *Longman Grammar of Spoken and Written English*. London: Longman, 1999.
- BLAKE, G. E BLY, R. W. *The Elements of Technical Writing*. The essential guide to writing clear, concise proposals, reports, manuals, letters, memos, and other documents in every technical field. New York, Macmillan General Reference, 1993.
- BOULANGER, J-C. *Compte-rendu. Terminogramme*. 1991.
- BROWN, P; LEVINSON S. C. *Politeness. Some universals in language usage*. Cambridge: Cambridge University Press, 1987.
- BYBEE, J. The Semantic Development of Past Tense Modals in English. In: Bybee, Joan; Fleischman, Suzanne. *Modality in Grammar and Discourse*. Philadelphia: John Benjamins, 503-517, 1995.

CARTER, R.; McCARTY, M. Grammar and the Spoken Language. *Applied Linguistics*, 16, 141-158, 1995.

CABRÉ, M. T.; FREIXA, J.; LORENTE, M.; TEBE, C. La terminología hoy: replanteamiento o diversificación. *Terminologia e integração. Organon*, v.12, n.26:33-41, 1998.

CIAPUSCIO, G. E. *La terminología en la descripción y tipificación del discurso especializado*. Simposio de RiTerm (2000). Disponível em: <<http://www.riterm.net/actes/7simposio/ciapuscio.htm>>. Acesso em 14 de julho de 2008.

COATES, J. *The Semantics of the Modal Auxiliaries*. London & Canberra: Croom Helm, 1983.

COATES, J. The Expression of Root and Epistemic Possibility in English. In: Bybee, Joan; Fleischman, Suzanne. *Modality in Grammar and Discourse*. Philadelphia: John Benjamins, 55-66, 1995.

CORTELAZZO, M.A. *Lingue Speciale. La Dimensione Verticale*. Padova: Unipress, 1994.

CRYSTAL, D. *A dictionary of linguistics and phonetics*. 2nd edition. New York: Basil Blackwell. 1985

CRYSTAL, D. The past, present and future of English parsing. *English Today* 3, p. 21-5, 1988

DENISON, D. *English Historical Syntax*. London: Longman, 1993.

DENARDIN, F. W. *Central Modals in an Operating System Corpus: Frequency, Distribution and Teaching Suggestions*. Monografia de Especialização. Porto Alegre, PUC-RS, 2008.

DUBOIS, J et al. *Dicionário de Lingüística*. São Paulo: Cultrix, 1978.

DUDLEY-EVANS, T. E ST. JOHN M. *Developments in English for Special Purposes*. Cambridge: CUP, 1998

ELLIS, S E GERIGHTY, T. *English for Aviation*. Oxford: OUP, 2008

EMERY, H. e ROBERTS, A. *Aviation English*. Oxford: Macmillan, 2008.

FARRINGTON, G. *Simplified English: The State of the Art*. International Aviation English Association Newsletter, 1993.

FAULSTICH, ENILDE. *A socioterminologia na comunicação científica e técnica*. Ciência e Cultura. abr./jun. 2006, vol.58, no.2, p.27-31. Disponível em: <http://cienciaecultura.bvs.br/scielo.php?script=sci_arttext&pid=S000967252006000200012&lng=es&nrm=iso>. Acessado em 25 de janeiro de 2008.

FILLMORE, C. 'Corpus Linguistics' or 'Computer-aided armchair linguistics' In: Svartvik, J. (ed.) *Directions in Corpus Linguistics. Proceedings of Nobel Symposium 82, Stockholm, 4-8 August 1991*. Berlin: De Gruyter. p.35-60, 1992.

FINATTO, M.J.B. *Definições Terminológicas: Fundamentos teórico-metodológicos para sua descrição e explicação*. Tese de Doutorado. Porto Alegre: UFRGS, 2001

FINATTO, M.J.B. Termos, textos e textos com termos: novos enfoques dos estudos terminológicos de perspectiva lingüística. In: Isquierdo, A. e Krieger, M. G. (Orgs) *As Ciências do Léxico II*. Campo Grande: UFMS, p. 341-357, 2004.

FINATTO, M. J. B.; EICHLER, M. L.; DEL PINO, J. C. *Sujeitos e agentes de poder e dever em textos sobre equilíbrio químico: aspectos lingüístico-terminológicos e aspectos conceituais da enunciação científica e o ensino aprendizagem de química*. Organon 16 (32-33), 83-104, 2003.

FRANCIS, W.N. e KUCERA, H. *Brown Corpus Manual: MANUAL OF INFORMATION to accompany A Standard Corpus of Present-Day Edited American English, for use with Digital Computers, 1964*. Internet: <<http://khnt.hit.uib.no/icame/manuals/brown/INDEX.HTM>> Acessado em 10/2/2008.

GABRIELATOS, C., McENERY, T. *Epistemic Modality in MA Dissertations*. Lengua y Sociedad: Aportaciones recientes en Linguística Cognitiva, lenguas en Contacto, Lenguajes de especialidad Y Linguística del Corpus. Valladolid, Universidad de Valladolid, 2005.

GABRIELATOS, C, SARMENTO, S. Central Modals in an Aviation Corpus: Frequency and distribution. *Letras de Hoje*, Vol 41, N.2. Porto Alegre: EDIPUCRS, 2006.

GAUDIN, F. *Pour une socioterminologie*. Rouen: Publications Université de Rouen, 1993.

GAVIOLI, L. *Exploring Corpora for ESP Learning*. Amsterdam: Jonh Benjamins, 2005.

GLÄSER, R. *The Stylistic Potential of Phraselological Units in the Light of Genre Analysis*. Phraselogy. Ed. A.P.Cowie. Oxford:Clarendon Press. 1998.

GÓMEZ, R. *Variability and Detection of Invariant Structure*. Psychological Sciences, Vol.3. No 5, 431-436, 2002.

GOTTI, M. *I Linguaggi Specialistici*. Firenze, La Nuova Italia 1991.

HALLIDAY, M. A. K. *On The Language of Physical Science*. In Ghadessy, M: Registers of Written English , London: Pinter, 1988.

HALLIDAY, M.A.K *Current Ideas in Systemic Practice and Theory*, London : Pinter, 1991

HOFFMANN, L. *Conceitos Básicos da Lingüística das Linguagens Especializadas* (Título Original: Grundbegriffe der Frachsprachenlinguistik, 1988). Disponível em: Cadernos de Tradução, Porto Alegre, nº 17, outubro-dezembro, p. 79-90, 2004.

HOFFMANN,L. Texts and Texts types in LSP. In: Schröder, H. *Subject-oriented Texts: Language for Special Purposes and Text Theory*, Berlin: Walter de Gruyter, p. 158-166, 1991

HOFFMANN, L. Possibilidades de aplicação e a aplicação atual de métodos estatísticos na pesquisa de linguagens especializadas (Título Original: anwendungsmöglichkeiten und bisherige Anwendung von Statistischen Methoden in der Fachsprachenforschung, 1998). Disponível em: *Cadernos de Tradução*, Porto Alegre, nº 20, janeiro-junho, p. 61-76, 2007

HOYE, L. *Adverbs and Modality in English*. Essex: Longman,1997.

HUDDLESTON, R.; PULLUM, G.K. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press, 2002.

- HUNSTON, S. *Corpora in Applied Linguistics*. London: Cambridge University Press, 2002.
- HYMES, D. *Foundations in sociolinguistics: an ethnographic approach*. Philadelphia: University of Pennsylvania Press, 1974.
- HYMES, D. H. *On Communicative Competence*. Disponível em: BRUMFIT, C. J. & JOHNSON, K. *The Communicative Approach to Language Teaching*. Oxford: Oxford University Press, 1979.
- ICAO. *Manual on the Implementation of ICAO Language Proficiency Requirements*. International Civil Aviation Organization, 2004.
- KENNEDY, G. D. *An introduction to corpus linguistics*. Nova York : Longman, 1998.
- KENNEDY, G. D. *Variation in the distribution of modal verbs in the British National Corpus*. In: Reppen, R.; Fitzmaurice, S.M., Biber, D. *Using Corpora to Explore Linguistic Variation*. Amsterdam: John Benjamins: 73-91, 2002.
- KILIAN, C. *A retomada de unidades de significação especializada em textos em língua alemão sobre gestão de resíduos: uma contribuição para a tradução técnico-científica*. Tese (Doutorado em Letras) - Universidade Federal do Rio Grande do Sul, 2007.
- KITTREDGE, R. *Variation and Homogeneity of Sublanguages*. In: KITTREDGE, R. E LEHRBERGER, J. *Sublanguages: Studies of Language in Restricted Domain*. Walter de Gruyer, NY, 1982
- KRIEGER, M. G. *A Face Lingüística da terminologia*. In: Krieger. M. da G.; Maciel, A.M.B. *Temas de Terminologia*. Porto Alegre/São Paulo: Ed. da Universidade/Humanitas, 2001.
- KRIEGER, M.G.; FINATTO, M.J.B. *Introdução a Terminologia - Teoria e Prática*, São Paulo: Contexto, 2004.
- LAKOFF, G. Hedges: A Study in Meaning Criteria and the Logic of Fuzzy Concepts', In P.M. Peranteau, J. N. Levi and G.C. Phares (eds.), *Papers from the Eighth Regional Meeting*, Chicago Linguistics Society (CLS 8), pp 183-228, 1972
- LAKOFF, R. *The Pragmatics of Modality*, In P.M. Peranteau, J.N. Levi and G.C. Phares (eds) *Papers from the Eight Regional Meeting*, Chicago Linguistic Society. Chicago: Chicago Linguistic Society, 229-246, 1972.
- LEECH, G. N. *Meaning and the English verb*. Essex, U.K.: Longman Group Limited, 1971.
- LEECH, G. N. *Principle of Pragmatics*. London: Longman, 1983.
- LEECH, G. N. *O significado do verbo em inglês*. São Paulo: Ática.1989
- LEECH, G. N. *The State of Art in Corpus Linguistics*. London: Longman, 1991.
- LEECH, G. N. *Corpora and Theories of Linguistic Performance*. Berlin: Mouton de Gruyter, 1992.
- LEECH, G. N. *Introducing corpus annotation*. In Roger Garside, Geoffrey Leech e Anthony McEnery, eds. *Corpus Annotation. Linguistic information from computer text corpora*. London, Longman, pp. 1-18. 1997

- LEHRBERGER, J. *Automatic Translation and the Concept of Sublanguage*. In: KITTREDGE, R. E LEHRBERGER, J. *sublanguages: Studies of Language in Restricted Domain*. 81-106, NY: Walter de Gruyter, 1982
- LYONS, J. *Semantics*. Cambridge: Cambridge University Press, 1977.
- MACIEL, A.M.B. *Para o reconhecimento da especificidade do termo jurídico*. Tese de Doutorado. Porto Alegre: Universidade Federal do Rio Grande do Sul, 2001a
- MACIEL, A.M.B. Para o reconhecimento da especificidade do termo jurídico. In: Krieger. M. da G.; Maciel, A.M.B. *Temas de Terminologia*. Porto Alegre/São Paulo: Ed. da Universidade/Humanitas: 436-438, 2001b
- McCARTEN. *Teaching Vocabulary-Lessons from the Corpus, Lessons for the Classroom*, CUP: Cambridge 2007.
- McENERY, T.; GABRIELATOS, C. *English corpus linguistics*. In B. Aarts & A. McMahon (eds.), *The Handbook of English Linguistics* (pp. 33-71), Oxford: Blackwell, 2006
- McENERY, T. e WILSON, A. *Corpus Linguistics*. Edinburgh: Edinburgh University Press, 1996.
- McENERY, T; XIAO, R. e TONO, Y. *Corpus-based Language Studies: an advanced resource book*, Oxon: Routledge, 2006
- MYHILL, JOHN E SMITH, LAURA. *The Discourse and Interactive Functions of Obligation Expressions*. In: Bybee, Joan; Fleischman, Suzanne. *Modality in Grammar and Discourse*. Philadelphia: John Benjamins, 239-292, 1995.
- NEVES, M. H. de M. *Texto e Gramática*. São Paulo: Contexto, 2006.
- OLIVEIRA, FÁTIMA. *Modalidade e modo* in MATEUS ET AL, *Gramática da Língua Portuguesa*. 5ª edição revista e aumentada. Lisboa: Caminho, 245-247, 2003.
- OPPIZZI, A. *Compounding in Specialized Languages*. Trento: Editrice Uni Service, 2006
- PALMER, F.R. *The English verb*. London: Longman, 1974.
- PALMER, F.R. *Mood and Modality*. Cambridge: Cambridge University Press, 1986.
- PALMER, F.R. *Modality and the English Modals*. 2nd Edition, London: Longman, 1990.
- PALMER, F.R. Negation and the Modals of Possibility and Necessity In: Bybee, Joan; Fleischman, Suzanne. *Modality in Grammar and Discourse*. Philadelphia: John Benjamins, 453-472, 1995.
- PALMER F.R. Modality in English: Theoretical, descriptive and typological issues. In; Facchinetti, Roberta; Krug, Manfred; Palmer, Frank: *A corpus based study of grammaticalization: Modality in Contemporary English*. Cambridge: Cambridge University Press 2003.
- QUIRK, R.; GREENBAUM, S.; LEECH, G.; SVARTIK, J. *A Comprehensive Grammar of the English Language*. London: Longman, 1985.

- RAVIN, Y.; LEACOCK, C. *Polysemy: An Overview*. In: Ravin, Y.; Leacock, C. (eds). *Polysemy- Theoretical and Computational Approaches*. Oxford: Oxford Linguistics, 2000.
- RAYSON, P. *Matrix: A statistical method and software tool for linguistic analysis through corpus comparison*. Tese de doutorado. Universidade de Lancaster, 2002
- ROBERTSON, F. A. *Airpeak: Radiotelephony Communication for Pilots*. Oxford: Prentice Hall, 1988.
- ROSCH, H. E. *Universals in color naming and memory*. *Journal of Experimental Psychology*, 93, 10-20, 1972.
- SAGER, J. C.; DUNGWORTH, D.; MCDONALD P.F.. *English Special Languages*. Principles and Practice in Science and Technology. Wiesbaden, Brandstetter Verlag, 1980.
- SARMENTO, S. *Distribution of Modal Verbs in an Aviation Corpus*. Proceedings of the Corpus Linguistics Conference, 2005a. Disponível em <<http://www.corpus.bham.ac.uk/pclc>> Acessado em 10 de abril de 2007.
- SARMENTO, S. *A Pragmatic Account of Aviation Modals*. In: ESP-WORLD. Issue 3(11), Volume 4, 2005b. Disponível em < <http://www.esp-world.info/contents.htm>> Acessado em 15 de abril de 2007.
- SARMENTO, S. The Case of “Can”. *Revista Virtual de Estudos da Linguagem – ReVEL*. V. 5, n. 8, 2007. Disponível em: <<http://www.revel.inf.br>>. Acessado em 14 de julho de 2008.
- SCHLATTER, M., GARCEZ, P. M. & SCARAMUCCI, M. O papel da interação na pesquisa sobre aquisição e uso de língua estrangeira: implicações para o ensino e para avaliação. *Letras de Hoje*, 39 (3), 345-378, 2004.
- SCOTT, M. *WordSmith Tools*. (1996) Oxford: Oxford University Press. Versão 5, 2008.
- SEARLE, J.R. *Speech Acts – An essay in the philosophy of language*. Cambridge: Cambridge University Press, 1969.
- SHAWCROSS, P. *English for Aircraft Maintenance*. Paris: Belin, 1992.
- SILVA-CORVALAN, C. *Contextual Conditions for the Interpretation of ‘poder’ and ‘deber’ in Spanish*. In: Bybee, J; Fleischman, S. (eds) *Modality in Grammar and Discourse*. Amsterdam/Philadelphia: John Benjamins, 67-106, 1995.
- SINCLAIR, J. *Corpus, Concordance, Collocation*. Oxford: OUP, 1991.
- SINCLAIR, J. *Paper Presented at XI Encontro da Associação Portuguesa de Lingüística*. Lisboa, 1995.
- SINCLAIR, J. A Way With Words. In H. Hasselgard and S. Oksefjell (eds.). *Out of Corpora: Studies in Honor of Stig Johansson*. Amsterdam: Rodopi, 1999.
- SINCLAIR, J. *Trust the Text: Language, Corpus and Discourse*. London: Routledge, 2004.
- STUBBS, M. Grammar, text and ideology: Computer Assisted Methods in the linguistic of representation. *Applied Linguistics*, 15, 201-233, 1994.
- STUBBS, M. *Corpus and Text Analysis*. Oxford: Blackwell, 1996.

- STUBBS, M. *Words and Phrases*. Oxford: Blackwell, 2001.
- SWALES, J.M. *Genre Analysis*. Cambridge: Cambridge University Press, 1990.
- SWEETSER, E. *From Etymology to Pragmatics*. Cambridge: CUP, 1990.
- TALMY, L. *Force Dynamics in language and cognition* In *Cognitive Science*, 12, 1, 49-100, 1988.
- TALMY, L. *Toward a cognitive semantics*, vol.1: concepts structuring systems. Cambridge, Mass.: MIT Press, 2000.
- TEMMERMAN, R. *Une théorie réaliste de la Terminologie: Le Sociocognitivism*. 2000.
- TOGNINI-BONELLI, E. *Corpus Linguistics at Work*. Amsterdam: John Benjamins, 2001.
- TRIMBLE, LOUIS. *English for Science and Technology. A discourse approach*. Cambridge: C.U.P., 1985.
- WITTGENSTEIN, L. (1953) *Philosophical Investigations*. In: Aarts, B.; Denison, D.; Keizer, E.; Popova, G. *Fuzzy Grammar: A Reader*. Oxford: Oxford University Press, 2004.

ANEXO A- FÓRMULA DO *LOG LIKELIHOOD* PARA COMPARAR CORPORA DIFERENTES

O LL para esse tipo de análise utiliza a seguinte tabela de contingência:

	<i>Corpus 1</i>	<i>Corpus 2</i>	Total
Frequência da palavra	a	b	a+b
Frequência das outras palavras	c-a	d-b	c+d-a-b
Total	c	d	c+d

O valor 'c' corresponde ao número de palavras no primeiro corpus, e 'd' corresponde ao número de palavras no segundo corpus (valores N). Os valores 'a' e 'b' correspondem aos valores observados (O), enquanto se faz necessário calcular os valores esperados (E) de acordo com a seguinte fórmula:

$$E_i = \frac{N_i \sum_j O_j}{\sum_i N_i}$$

No caso $N_1 = c$, e $N_2 = d$, então, para essa palavra, $E_1 = c*(a+b) / (c+d)$ e $E_2 = d*(a+b) / (c+d)$. O cálculo para os valores esperados leva os tamanhos dos dois *corpora* em consideração. Dessa forma, não é necessário normalizar os resultados antes de aplicar a fórmula. É possível calcular os valores LL de acordo com a seguinte fórmula:

$$-2 \ln \lambda = 2 \sum_i O_i \ln \left(\frac{O_i}{E_i} \right)$$

Quanto mais alto for o valor G2, mais significativa será a diferença entre os resultados. Para essas tabelas, um G2 de 3.8 ou maior é significativo no nível de $p < 0.05$ e um G2 de 6.6 ou maior é significativo a < 0.01 .

- percentil 95; nível 5%; $p < 0.05$; valor crítico = 3,84
- percentil 99; nível 1%; $p < 0.01$; valor crítico = 6,63
- percentil 99,9; nível 0.1%; $p < 0.001$; valor crítico = 10,83
- percentil 99,99; nível 0.01%; $p < 0.0001$; valor crítico = 15,13

ANEXO B- PRINCIPAIS DOMÍNIOS SEMÂNTICOS DOS VERBOS

Atividade	Comunicação	Mentais	Causativos	Ocorrência	Existência	Aspecto
Accompany	Accuse	Accept	Affect	Arise	Appear	Begin
Acquire	Acknowledge	Afford	Allow	Be	Concern	Cease
Advance	Address	Agree	Assist	Become	Constitute	Continue
Add	Admit	Appreciate	Cause	Change	Contain	Complete
Apply	Advise	Approve	Enable	Develop	Define	End
Arrange	Announce	Assess	Ensure	Die	Derive	Finish
Beat	Answer	Assume	Force	Disappear	Deserve	Keep
Behave	Appeal	Blame	Guarantee	Emerge	Exist	Stop
Borrow	argue	Bear	Help	Fall	Fit	Start
Bring	Ask	Believe	Influence	Flow	Illustrate	
Burn	Assure	Bother	Let	Grow	Imply	
Buy	Call	Calculate	Permit	Happen	Include	
Carry	Challenge	Care	Prevent	Have	Indicate	
Catch	Claim	Celebrate	Require	Increase	Involve	
Check	Complain	Choose		Last	Lack	
Clean	Consult	Compare		Occur	Live	
Climb	Convince	Conclude		Rise	Locate	
Combine	Declare	Confirm		Shine	Look	
Come	Demand	Consider		Sink	Matter	
Control	Deny	Count		Slip	Owe	
Cover	Describe	Dare			Own	
Defend	Discuss	Decide			Reflect	
Deliver	Emphasise	Deserve			Relate	
Dig	Encourage	Detect			Remain	
Divide	Excuse	Determine			Represent	
Earn	Explain	Discover			Reveal	
Eat	Express	Dismiss			Seem	
Encounter	Inform	Distinguish			Sound	
Engage	Insist	Doubt			Stand	
Exercise	Invite	Enjoy			Stay	
Expand	Mention	Estimate			tend	
Explore	Note	Examine				
Extend	Offer	Expect				
Fix	Persuade	Experience				
Form	Phone	Face				
Follow	Pray	Fear				
Get	Promise	Feel				
Give	Propose	Find				
GO	Publish	Forget				
Hang	Question	Forgive				
Hold	Quote	Guess				
Join	Recommend	Hate				
Leave	Remark	Hear				
Lie	Reply	Hope				
Lose	Report	Identify				
Meet	Respond	Ignore				
Move	Say	Imagine				

Obtain	Shout	Impress				
Open	Sign	Interpret				
Pass	Sing	Intend				
Pay	Speak	Judge				
Play	Specify	Justify				
Produce	State	Know				
Provide	Suggest	Learn				
Pick	Swear	Like				
Put	Talk	Listen				
Reach	Teach	Love				
Reduce	Tell	Mean				
Run	Thank	Mind				
Sell	Threaten	Miss				
Send	Urge	Need				
Shake	Warn	Notice				
Sit	Welcome	Observe				
Smile	Whisper	Perceive				
Spend	Write	Plan				
Show		Predict				
Stare		Pretend				
Take		Prefer				
Turn		Prove				
Try		Realize				
Use		Read				
Wait		Reckon				
Walk		Recall				
Watch		Recognize				
Wear		Regard				
Win		Remember				
Work		Remind				
		Satisfy				
		See				
		Solve				
		Study				
		Suffer				
		Suppose				
		Suspect				
		Taste				
		Think				
		Trust				
		Understand				
		Want				
		Wish				
		Wonder				
		Worry				