UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, LETRAS E CIÊNCIAS HUMANAS
DEPARTAMENTO DE LETRAS MODERNAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS LINGUÍSTICOS E LITERÁRIOS EM INGLÊS

MALILA CARVALHO DE ALMEIDA PRADO

Levantamento dos padrões léxico-gramaticais do inglês para aviação: um estudo vetorado pela Linguística de *Corpus*

VERSÃO CORRIGIDA

São Paulo
2015

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, LETRAS E CIÊNCIAS HUMANAS
DEPARTAMENTO DE LETRAS MODERNAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS LINGUÍSTICOS E LITERÁRIOS EM INGLÊS

MALILA CARVALHO DE ALMEIDA PRADO

Levantamento dos padrões léxico-gramaticais do inglês para aviação: um estudo vetorado pela Linguística de *Corpus*

VERSÃO CORRIGIDA

Dissertação apresentada à Faculdade de Filosofia, Letras e Ciências Humanas, Departamento de Letras Modernas da Universidade de São Paulo para a obtenção do título de Mestre em Letras no Programa de Estudos Linguísticos e Literários em Inglês.

Orientadora: Profa. Dra. Stella Esther Ortweiler Tagnin

São Paulo
2015

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Catalogação na Publicação
Serviço de Biblioteca e Documentação
Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo

P896l

Prado, Malila Carvalho de Almeida
Levantamento de padrões léxico-gramaticais do inglês para aviação: um estudo vetorado pela Linguística de Corpus / Malila Carvalho de Almeida Prado ; orientadora Stella Esther Ortweiler Tagnin. - São Paulo, 2015.
133 f.

Dissertação (Mestrado)- Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo. Departamento de Letras Modernas. Área de concentração: Estudos Linguísticos e Literários em Inglês.

1. Linguística de Corpus. 2. Ensino de língua estrangeira. 3. Língua inglesa. I. Tagnin, Stella Esther Ortweiler, orient. II. Título.

FOLHA DE APROVAÇÃO

Malila Carvalho de Almeida Prado

Levantamento dos padrões léxico-gramaticais do inglês para aviação: um estudo vetorado pela Linguística de *Corpus*

Dissertação apresentada à Faculdade de Filosofia, Letras e Ciências Humanas, Departamento de Letras Modernas da Universidade de São Paulo para a obtenção do título de Mestre em Letras no Programa de Estudos Linguísticos e Literários em Inglês.

Aprovada em: ____ / ____ / ____

Banca examinadora

Prof. Dr.: _____

Instituição: _____ Assinatura: _____

Prof. Dr.: _____

Instituição: _____ Assinatura: _____

Prof. Dr.: _____

Instituição: _____ Assinatura: _____

Dedicatória

À Walkyria, minha mãe

AGRADECIMENTOS

À Professora Stella, meus sinceros agradecimentos pelo apoio, pela paciência, pela motivação e pela inspiração.

Ao Edimilson, meu parceiro de vida e meu consultor operacional.

A minha família, principalmente ao meu irmão e à Roberta, pela paciência e pela compreensão da minha ausência.

Aos meus afilhados e enteadas, que perderam a madrinha e a madrasta por alguns longos momentos.

Aos meus amigos, que me apoiaram sempre, especialmente a Andrea, Sueli e Fabrícia.

Ao pessoal do COMET, principalmente Joacyr, Danilo e Marlene. Obrigada pela companhia, pelo apoio e pela amizade!

À Rozane, que tão cuidadosamente revisou esta dissertação.

Aos meus alunos queridos, por me darem o motivo que me levou até aqui.

Aos colegas da aviação, que me auxiliaram com a revisão dos textos que compõem o *corpus* de estudo: Henry Emery, Dennis Gliddon, Christopher Matthiessen, Rami Soueidi, Rodrigo Farina, Thiago Vilas Boas, André Danita, Fernando Fernandes, Damerel Neto e, principalmente, Eduardo Venna e Ricardo Aleixo.

SUMÁRIO

ÍNDICE DAS FIGURAS ix
ÍNDICE DOS GRÁFICOS xi
ÍNDICE DE QUADROS xii
ÍNDICE DE ABREVIATURAS xiii
RESUMO xiv
ABSTRACT xv
INTRODUÇÃO 1
CAPÍTULO 1 – INGLÊS PARA AVIAÇÃO 5
1.1 A aviação e o inglês 5
1.2 A Fraseologia Aeronáutica 6
1.3 O conceito de Proficiência Linguística no inglês para aviação 10
CAPÍTULO 2 – FUNDAMENTAÇÃO TEÓRICA 15
2.1 O conteúdo do ensino de línguas 15
2.2 Os padrões fraseológicos 17
2.3 A Linguística de *Corpus* 20
2.4 *Corpora* orais 21
2.4.1 A Teoria da Língua em Ato 24
CAPÍTULO 3 – METODOLOGIA 28
3.1 O *corpus* de estudo 28
3.1.1 O texto 33
3.1.2 Os falantes 36
3.1.3 Armazenamento do *corpus* 38
3.2 Procedimentos de análise para a exploração do *corpus* 42
3.2.1 Lista de palavras 43
3.2.2 A lista de palavras-chave 48
3.2.3 As linhas de concordância 50
3.2.4 Os *clusters* 51
3.2.5 Os colocados 53
CAPÍTULO 4 – ANÁLISE 56
4.1 As estruturas léxico-gramaticais do inglês para aviação 56
4.1.1 *Runway* 61

4.1.2 *Emergency* 75
4.1.3 *Aircraft* 81
4.2.4 *Engine* 86
4.1.5 *Fuel* 92
4.2 *The book is on the table* ou *the plane is on the runway*? Um estudo comparativo com o Glossário de Estruturas Básicas 97
CONSIDERAÇÕES FINAIS 106
REFERÊNCIAS BIBLIOGRÁFICAS 110
ANEXO 117

ÍNDICE DAS FIGURAS

Figura 1: Fases do voo. 7
Figura 2: Exemplificação retirada do DOC 9432, Manual de Fraseologia Aeronáutica. 8
Figura 3: Comparações entre o que livros de gramática informam e o que o *corpus* mostra. . 16
Figura 4: *Site* LiveATC.net, com a página *Interesting Recordings*. 32
Figura 5: Organização do *corpus* em pastas. 39
Figura 6: Tela do WST com a primeira lista de palavras gerada do *corpus* de estudo. 44
Figura 7: Parte da *stoplist*. 45
Figura 8: Lista de palavras geradas com a *stoplist*. 46
Figura 9: Tela com a lista de palavras gerada com a *stoplist* e lematizada. 48
Figura 10: *Website* do MICASE, com a ferramenta de busca e possíveis seleções que possam restringir os campos de análise. 49
Figura 11: Tela das palavras-chave do *corpus* de inglês aeronáutico. 50
Figura 12: Tela do WST da ferramenta Concord, com linhas de concordância da palavra *runway*. 51
Figura 13: Telas de *clusters* de *runway* do WST, comparando dois processos. 52
Figura 14: Tela de linhas de concordância com o *cluster 'on the runway'* no núcleo. 53
Figura 15: Tela de colocados da palavra *runway*. 54
Figura 16: Porcentagem de acidentes aéreos por fase de voo, de 1999 a 2008. 62
Figura 17: Modelo de diálogo retirado do DOC 9432, Manual de Fraseologia Aeronáutica. . 75
Figura 18: Extrato de documento oficial sobre categoria de mensagens de emergência. 76
Figura 19: Extrato do Anexo 10 da Convenção de Chicago sobre comunicações de emergência. 80
Figura 20: Excerto retirado de documentos de Fraseologia Aeronáutica sobre procedimentos de emergência para ATCOs. 86
Figura 21: Exemplo de uma declaração de emergência com uso de conjunção, nesse caso, *due to*. 91
Figura 22: Ilustração retirada do DOC 9432 p. 11-2. 93
Figura 23: Diálogo sobre o procedimento de alijamento de combustível em progresso. 95
Figura 24: Busca por linhas de concordância com pronomes interrogativos. 99
Figura 25: Seleção das linhas de concordância com pronomes em frases interrogativas. 99
Figura 26: Tela do *Concord* com busca por várias palavras, inseridas por arquivo txt. 99

ÍNDICE DAS TABELAS

Tabela 1: Termos Taxonômicos de Categoria de Ocorrência. 30
Tabela 2: Etiquetas utilizadas no *corpus.* 36
Tabela 3: Vinte primeiras palavras de dois sub*corpora*: 'nativos' e 'não-nativos'. 37
Tabela 4: Descrição do *corpus.* 42
Tabela 5: As 50 palavras mais frequentes do *corpus*, após *stoplist* e lematização. 59
Tabela 6: Lista das 50 palavras-chave mais frequentes do *corpus* de estudo. 60
Tabela 7: Lista das palavras candidatas a campos lexicais. 61
Tabela 8: Fraseologia Aeronáutica padronizada, com a palavra *runway* 63
Tabela 9: Linhas de concordância com '*on * runway*'. 65
Tabela 10: *Clusters* com a palavra *runway*. 66
Tabela 11: Linhas de concordância com o *cluster 'off the runway'* ao centro. 67
Tabela 12: Linhas de concordância com *runway is.* 67
Tabela 13: Posições de *is* à esquerda de *runway*. 68
Tabela 14: Linhas de concordância com o núcleo *runway* e *is* como colocado à esquerda. 69
Tabela 15: Linhas de concordância com o verbo *´re*, sendo *runway* o nódulo. 70
Tabela 16: Linhas de concordância com o 8º colocado de *runway*, o lema *TO BE*. 72
Tabela 17: Seleção dos vinte primeiros *clusters* de 3 palavras com ideia de futuro e com modais. 73
Tabela 18: Lista de colocados da palavra *emergency*. 77
Tabela 19: Agrupamentos lexicais com a palavra *emergency*. 77
Tabela 20: Vinte primeiros colocados da palavra *aircraft.* 82
Tabela 21: Linhas de concordância da palavra *aircraft* com *to* no entorno. 84
Tabela 22: Linhas de concordância com a expressão *of * aircraft*. 85
Tabela 23: *Patterns* mais frequentes da palavra *engine*. 87
Tabela 24: Linhas de concordância com a palavra *engine* e variações de *have*. 88
Tabela 25: *Clusters* com a palavra *fuel.* 92
Tabela 26: Linhas de concordância com a palavra *fuel* e *on.* 94
Tabela 27: Tela de *patterns* do *Concord* com a palavra *fuel* ao centro. 96
Tabela 28: Comparação entre o Glossário de Estruturas Básicas e suas ocorrências no *corpus* de estudo. 103

ÍNDICE DOS GRÁFICOS

Gráfico 1: Comparação entre as três listas geradas, cada uma com seu filtro..... 57
Gráfico 2: O uso de conjunções no *corpus* de estudo. 91

ÍNDICE DE QUADROS

Quadro 1: Parte da lista de estruturas básicas determinadas pelo *Eurocontrol Institute of Air Navigation Services.* 13
Quadro 2: Parte da lista de estruturas complexas determinadas pelo *Eurocontrol Institute of Air Navigation Services*..... 14
Quadro 3: Elocução extraída do C-Oral Brasil..... 23
Quadro 4: Diálogo transcrito de uma apresentação do *campus* universitário a novos alunos nos Estados Unidos 26
Quadro 5: Elocução extraída do *corpus* MICASE..... 33
Quadro 6: Exemplo de medida estatística da riqueza lexical. 33
Quadro 7: Exemplos de quebras prosódicas e mudanças de turno retirados do *corpus* de estudo. 35
Quadro 8: Exemplos de hesitações e suas grafias, retiradas do *corpus* de estudo. 36
Quadro 9: Cabeçalho das transcrições do *corpus* de estudo. 40
Quadro 10: Contrações presentes nas transcrições e suas lematizações. 47

ÍNDICE DE ABREVIATURAS

ANAC – Agência Nacional de Aviação Civil

ATCO – *Air Traffic Control Operator*

CEFR – *Common European Framework of Reference*

DOC – Documento (oficial publicado pela OACI)

FAA – *Federal Aviation Administration*

LC – Linguística de *Corpus*

OACI – Organização de Aviação Civil Internacional

PRICESG – *Proficiency Requirements in Common English Study Group*

PTT – *push to talk*

SARP – *Standards and Recommended Practices*

TLA – Teoria da Língua em Ato

MICASE – *Michigan Corpus of Academic Spoken English*

BNC – *British National Corpus*

COLT – *Corpus of London Teenage Speech*

C-ORAL-ROM – *Corpora of Spontaneous Spoken Language of the Main Romance*

RESUMO

PRADO, M. **Levantamento dos padrões léxico-gramaticais do inglês para aviação: um estudo vetorado pela Linguística de *Corpus***. 2015. 133 f. Dissertação (Mestrado em Estudos Linguísticos e Literários em Inglês)–Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2015.

A OACI (Organização de Aviação Civil Internacional), órgão que estabelece regulamentações na aviação civil em nível mundial, define o inglês para aviação como uma junção entre a Fraseologia Aeronáutica – sublinguagem empregada por pilotos e controladores de tráfego aéreo em comunicações radiotelefônicas em situações rotineiras – e o inglês comum (*plain English*), utilizado quando a Fraseologia Aeronáutica não se mostra suficiente. Após averiguar que acidentes aéreos foram agravados por falhas de comunicação, a OACI passou a exigir dos profissionais supracitados uma licença para operações internacionais. Para tal licenciamento, é requerida uma avaliação de proficiência linguística, prescrita no Manual de Implementação de Proficiência Linguística, ou DOC 9835, publicado pela OACI em 2004. A partir de então, o inglês aeronáutico teve seu grau de importância elevado e começou a ser avaliado em países ao redor do mundo, por governos e/ou por instituições internacionais, e, nos últimos anos, foram publicados materiais didáticos que se propõem a atender a essa nova demanda. Identificando a falta de autenticidade nesses materiais, e um descompasso entre a definição proposta pela OACI e o Glossário de Estruturas Básicas e Complexas adicionado à segunda edição do DOC 9835 (ICAO, 2010), buscamos no aporte teórico da Linguística de *Corpus* uma metodologia para a descrição do *plain English* utilizado em contextos aeronáuticos em uso real. Assim, o objetivo desta pesquisa é apresentar um estudo descritivo da linguagem utilizada por controladores de tráfego aéreo e pilotos em comunicações radiotelefônicas em situações anormais; para atingi-lo, compilamos um *corpus* falado extraído desse contexto. A partir desse *corpus*, analisamos, primeiramente, os padrões fraseológicos no entorno de cinco itens lexicais – *runway* [pista], *aircraft* [aeronave], *emergency* [emergência], *fuel* [combustível] e *engine* [motor] –, que foram evidenciados como as palavras de conteúdo mais frequentes no *corpus* de estudo. Contrastamos, em seguida, nossos resultados com o Glossário de Estruturas do próprio DOC 9835. Nossos resultados apontam que, assim como recomendado pela OACI, o inglês para aviação é simples e objetivo, não correspondendo ao tipo de linguagem sugerida no referido anexo nem ao conteúdo programático dos materiais didáticos utilizados atualmente.

PALAVRAS-CHAVE: inglês para aviação, Linguística de *Corpus*, *corpora* orais, ensino de língua de especialidade.

ABSTRACT

PRADO, M. **Analysis of lexical-grammar patterns of aviation English: a study vectored by Corpus Linguistics**. 2015. 133 p. Thesis. São Paulo: Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2015.

ICAO (International Civil Aviation Organization), a specialized agency which regulates civil aviation worldwide, defines aviation English as a combination of Aeronautical Phraseology – a sublanguage used by air traffic controllers and pilots in radiotelephony communications in routine situations – and plain English, used when Phraseology does not suffice. After verifying that air crashes were aggravated by miscommunications, ICAO started requesting a proficiency level accredited in the licenses of these two professionals when operating internationally. To obtain this license, applicants are required to take a proficiency exam, prescribed in the Manual of Implementation of Proficiency Requirements, or DOC 9835, published by ICAO in 2004. Since then, aviation English has reached a higher level of importance, and started being evaluated by governments and institutions throughout the world and, in the last years, coursebooks have been published with the objective of meeting those needs. Having identified a lack of authenticity in such materials, as well as an imbalance between the definition of aviation English in DOC 9835 and the Glossary of Basic and Complex Structures added to its second edition (ICAO, 2010), we used Corpus Linguistics as a methodology to investigate this language for special purposes in its real environment. Therefore, this research presents a description of the language used by air traffic controllers and pilots in radiotelephony communications when in abnormal situations. To achieve our goals, we compiled a spoken corpus extracted from this context. For this study, we primarily investigated the phraseological patterns in the surroundings of five lexical items – ‘runway’, ‘aircraft’, ‘emergency’, ‘fuel’ and ‘engine’ –, chosen because of their high frequency status in our corpus. We then contrasted our results with the Glossary attached to DOC 9835, which suggests a list of grammar structures on which the assessment and training of aviation English should be based. This research concludes with a demonstration of the simplicity, objectiveness and clarity of the plain English identified in our corpus, which does not correlate with the structures recommended in the Glossary nor with the language explored in coursebooks published to this end.

KEYWORDS: aviation English, Corpus Linguistics, spoken corpora, English for specific purposes

INTRODUÇÃO

Da experiência da sala de aula vieram os conflitos que motivaram esta pesquisa. Foram vários os relatos de alunos de que não havia relação entre o que era ensinado no curso e o que realmente acontecia na vida profissional. Muitos professores ouvem isso de seus alunos. Porém, quando se trata de aeronautas, o quadro se torna um pouco mais preocupante.

O ensino do inglês para aviação não é recente; sua propagação, no entanto, aumentou nos últimos anos, após o órgão internacional regulatório – a Organização de Aviação Civil Internacional (doravante OACI) – publicar, em 2004, um documento exigindo comprovação de proficiência em língua inglesa para licenciamento de aeronautas e controladores de tráfego aéreo (doravante ATCO, como em *Air Traffic Control Operator*)[1] que operarem internacionalmente. O prazo para os governos signatários da OACI se ajustarem à nova regulamentação se estendeu de março de 2008 a março de 2011.

A partir de então, a proficiência linguística adquiriu um papel extremamente importante, pois se tornou mais uma "licença para pilotar". Pilotos brasileiros, por exemplo, só podem voar a países vizinhos, na América do Sul, se comprovarem nível mínimo de inglês. Recentemente, uma empresa aérea no norte do Brasil cancelou seus voos entre o Acre e cidades no Peru por não ter em seu quadro pilotos com tal licença válida.[2] Algumas empresas aéreas exigem essa licença para promoverem seus pilotos a comandantes. Esses, entre outros fatores, contribuem para o aumento da necessidade do ensino da língua inglesa.

O documento que rege tanto o ensino quanto a avaliação do idioma é o Manual de Implementação de Proficiência Linguística, ou DOC 9835, cuja primeira edição foi publicada em 2004 e revisada em 2010. Muitos materiais foram publicados internacionalmente baseados na primeira edição do documento, que trazia concepções muito generalizadas sobre o inglês para aviação. Agora, as avaliações estão sendo atualizadas tendo por base a segunda edição, cuja definição de inglês para aviação é mais assertiva. Com o surgimento das avaliações ocorre também um *washback effect*, ou um efeito retroativo (BAILEY, 1999; FULCHER e DAVIDSON, 2006; entre outros)[3], que é o impacto dos testes no ensino e na aprendizagem,

[1] Disponível em: <http://www.decea.gov.br/sirius/atco-controlador-de-trafego-aereo/>. Acesso em: 06 dez. 2014.
[2] Notícia vinculada em: <http://g1.globo.com/ac/acre/noticia/2014/10/sem-pilotos-que-falem-ingles-empresa-suspende-voos-do-ac-para-o-peru.html>. Acesso: em 02 nov. 2014.
[3] Ver também Scaramucci (2004), que traça um "estado da arte" do efeito retroativo.

por vezes negativo, em que o foco do conteúdo ensinado é somente aquele que será avaliado, e o ensino torna-se um 'treinamento para a prova'.

Para a OACI, o inglês para aviação é "a comunicação da radiotelefonia aeronáutica [...] que corresponde à porção limitada de usos de linguagem de somente dois profissionais – ATCOs e tripulação de voo. Inclui Fraseologia Aeronáutica e o uso de *plain language* (inglês de uso comum)"[4] (ICAO, 2010, p. 3-2, tradução nossa). Essa é a definição que está na segunda edição do documento, em que também há um elenco de estruturas gramaticais, denominado Glossário de Estruturas Básicas e Complexas (Apêndice B, ICAO, 2010, p. B-13), que serve como conteúdo programático tanto para o ensino quanto para a avaliação. Essa lista, porém, contradiz a linguagem preconizada no corpo do documento, que advoga simplicidade e objetividade, como a Fraseologia Aeronáutica, uma sublinguagem utilizada pelos profissionais da área em situações rotineiras (ICAO, 2004, p.2-3), que será discutida no Capítulo 1 desta dissertação.

Essa contradição entre o real e o ensinado, no entanto, ainda não foi devidamente estudada. Na realidade, há uma carência de estudos descritivos da linguagem utilizada por pilotos e ATCOs (BOCORNY, 2011). Este trabalho propõe-se a preencher essa lacuna. Para tanto, valeu-se da Linguística de *Corpus* (LC), que permite a descrição e a análise da linguagem em uso. A fim de atingirmos nossos objetivos, foi compilado um *corpus* composto por situações anormais de voo para investigarmos o uso do inglês de uso comum nas comunicações radiotelefônicas e, assim, respondermos às questões principais deste trabalho:

1. Quais são os padrões léxico-gramaticais mais utilizados por pilotos e controladores de tráfego aéreo em situações anormais, não previstos na Fraseologia Aeronáutica?
2. Esses padrões correspondem às **estruturas básicas** constantes do Glossário de Estruturas Básicas e Complexas sugerido no Apêndice B do DOC 9835?[5]

Nesta pesquisa, o inglês aeronáutico é visto como uma linguagem compartilhada em uma área profissional específica e, como tal, é utilizado por essa comunidade como uma

[4] *"[...] aeronautical radiotelephony communications, a specialized subcategory of aviation language corresponding to a limited portion of the language uses of only two aviation professions – controllers and flight crews. It includes ICAO standardized phraseology and the use of plain language."* [todas as traduções a partir daqui são nossas, salvo indicação em contrário].

[5] O escopo deste trabalho é o estudo das estruturas básicas do inglês para aviação, razão pela qual não foram contempladas as estruturas complexas.

lingua franca, com um núcleo central. Essa concepção de língua, defendida por Brumfit (2001, p. 116) como inglês internacional, é considerada como pertencente[6] às pessoas que a utilizam, independente de sua nacionalidade, ou do contexto em que vivem (monolíngue ou multilíngue). A ideia de propriedade linguística mencionada aqui é crucial para o entendimento do ensino de inglês nos moldes atuais: a "necessidade de redefinir critérios de aceitabilidade à luz do conceito de comunicabilidade" (RODRIGUES, 2005, p. 29) provocou um desligamento do inglês com as nações em que essa língua é oficial, e um repasse da propriedade à comunidade que a utiliza. Isso se deve ao crescente uso da língua por falantes 'não nativos' e, portanto, à inevitabilidade da mudança dos modelos para o ensino (Cf. SILVA, 2012; CRYSTAL, 2003; GRADDOL, 1997, 2006; RAJAGOPALAN, 1997; KACHRU, 1992;).

Partindo desse conceito, buscamos, com o uso da LC, uma descrição do núcleo de linguagem compartilhado por pilotos e ATCOs em nível internacional, por meio do levantamento de suas estruturas léxico-gramaticais. Por léxico-gramática entendemos a relação existente "entre as palavras e a gramática adjacente, ou entre as estruturas gramaticais e seus léxicos circundantes"[7] (BIBER; CONRAD; RAPPEN, 1998, p.84). É o recorte que priorizamos neste trabalho, facilitado pelas ferramentas de busca que a LC nos oferece.

O *corpus* coletado é um *corpus* falado, ou seja, uma coletânea de textos transcritos de comunicações radiotelefônicas de pilotos e ATCOs quando em situações anormais ou de emergência. As transcrições se fazem necessárias devido às ferramentas de análise utilizadas neste trabalho, que somente processam a palavra escrita. Porém, visando a reusabilidade do *corpus*, todas as transcrições estão acompanhadas dos respectivos áudios, o que pode facilitar futuros estudos fonéticos, pragmáticos, entre outros. O *corpus*, que para este estudo é composto de 92 textos e 70.323 palavras, buscou contemplar as categorias de ocorrências sugeridas pela OACI para a padronização de relatórios de acidentes e incidentes[8] por todo o mundo (ICAO, 2006).

---

[6] No original, *ownership*..

[7] *"Between words and their grammatical environments, or between grammatical structures and their lexical environments"*.

[8] A fim de classificar ocorrências na aviação, a distinção entre acidentes e incidentes é, segundo a OACI, a seriedade das consequências: caso haja ferimentos, mortes ou danos graves à aeronave, a ocorrência é classificada como acidente; se houver iminência do acidente, no entanto sem danos materiais graves ou ferimentos, é considerada como incidente. (ICAO, Annex 13, 1994, disponível em: < http://www.iprr.org/manuals/Annex13.html>. Acesso em: 08 nov. 2014)

Consideram-se, portanto, as características de linguagem oral, espontânea, produzida no momento da ação. Por isso, adotamos como base de transcrição da fala a Teoria da Fala em Ato (CRESTI, 2000, *apud* RASO, 2012, p. 92), que considera a linguagem oral como composta por enunciados e não por frases, como na linguagem escrita. Para nossa investigação, a divisão em enunciados nos possibilitou analisar estruturas léxico-gramaticais que não estariam completas caso fossem vistas como frases, mas que ainda assim possuem unidade de sentido.

Esta dissertação está assim organizada:

No Capítulo 1, esboçamos a necessidade da língua inglesa na aviação, descrevemos brevemente a Fraseologia Aeronáutica padronizada pela OACI, e abordamos os conceitos de língua, ensino e avaliação expostos no DOC 9835.

O Capítulo 2 traz a fundamentação teórica em que nos respaldamos – a Linguística de *Corpus* –, com enfoque em *corpora* orais, e os estudos descritivos por meio de padrões léxico-gramaticais.

No Capítulo 3, apresentamos a Metodologia utilizada, descrevendo o *corpus* de estudo, o *corpus* de referência e os procedimentos usados para a análise.

O Capítulo 4, por sua vez, delineia duas análises: uma tendo por base cinco palavras-chave e seus padrões léxico-gramaticais, demonstrando o uso do inglês geral nos contextos em que não há linguagem suficiente na Fraseologia Aeronáutica; a outra, com um contraste entre o *corpus* de estudo e o Apêndice B do DOC 9835, que apresenta a lista de estruturas gramaticais simples.

As Considerações Finais trazem as conclusões a que chegamos, e que, esperamos, possam contribuir para os Estudos Descritivos do inglês para aviação e, consequentemente, para o seu ensino.

O texto se completa com a Bibliografia utilizada para esta pesquisa e os anexos.

CAPÍTULO 1 – INGLÊS PARA AVIAÇÃO

We want the air to unite the peoples, and not to divide them.

Lord Swinton

Neste capítulo, contextualizamos nossa pesquisa, discutindo os seguintes tópicos:

- a importância do inglês para aviação;
- a Fraseologia Aeronáutica;
- a criação de um documento oficial para a regulamentação do inglês para aviação.

1.1 A aviação e o inglês

Como meio de transporte utilizado por pessoas do mundo inteiro, a aviação cruza fronteiras rápida e facilmente. O número de voos diários, domésticos e internacionais, aumenta consideravelmente a cada ano. Somente em 2013, foram 3.1 bilhões de passageiros que cruzaram os ares, 32 milhões de decolagens e aproximadamente 1.500 aeronaves comerciais novas entregues. Estudos indicam que o tráfego aéreo duplica a cada quinze anos.[9]

Desde o início da aviação comercial houve necessidade de regularização dos voos internacionais. Com esse intuito, foi criada, em 1944, a OACI (Organização de Aviação Civil Internacional), uma agência especializada das Nações Unidas (ONU), por meio da assinatura da Convenção de Chicago (*Chicago Convention*), realizada por vários países – o Brasil entre eles. A OACI tem a função de apresentar a documentação necessária por meio de SARPs (*Standards and Recommended Practices*) para que os governos signatários tenham embasamento para o desenvolvimento e regularização de suas operações aeronáuticas e para que, assim, haja uma padronização em nível global. Essas SARPs estão descritas em dezenove Anexos à Convenção de Chicago. É também responsabilidade da OACI a

[9] Dados disponíveis em: <http://www.icao.int/annual-report-2013/Documents/Appendix_1_en.pdf>. Acesso em: 02 nov. 2014.

assistência aos governos em relação às implementações de regulamentos, além de auditorias na execução dos procedimentos.

A comunicação entre os controladores de tráfego aéreo (ATCOs) e pilotos de diferentes nacionalidades exigiu, desde o início da aviação comercial, uma língua em comum; e o inglês, sendo a língua dos líderes da OACI, a língua dos principais fabricantes de aeronaves de então, e o principal idioma da maioria dos pilotos (militares pós-guerra), foi sugerida pela OACI, em 1951, como língua oficial da aviação (CRYSTAL, 2003).

Hoje a OACI reconhece outros idiomas como oficiais e permite o uso da língua local dentro das fronteiras dos países. Assim, dentro do Brasil, por exemplo, é possível para ATCOs e pilotos se comunicarem em português. Porém, ao operarem em aeroportos ou centros de controle de espaço aéreo internacionais, pilotos e ATCOs devem se comunicar somente em inglês. Tal exigência entrou em vigor após investigações de acidentes aéreos terem comprovado que o mau ou nenhum uso do inglês foi uma das causas contribuintes. A OACI passou, então, a oficializar a necessidade do uso da Fraseologia Aeronáutica, uma sublinguagem específica utilizada por ATCOs e pilotos e, no início deste século, a requerer a averbação de uma nova licença de operação: a proficiência linguística.

O Anexo 1 da Convenção de Chicago trata do licenciamento de pessoal. Todas as licenças mínimas exigidas para profissionais que lidam diretamente com o voo estão descritas nesse documento. Uma delas está relacionada à proficiência linguística que pilotos e ATCOs devem demonstrar para operações internacionais. Assim como todas as licenças, a avaliação da capacidade dos profissionais é de responsabilidade dos governos signatários. Por isso, atualmente cada país é responsável pela implementação dessa licença, guiado por um documento oficial, o Manual de Implementação de Proficiência Linguística, ou DOC 9835.

1.2 A Fraseologia Aeronáutica

A Fraseologia Aeronáutica é uma sublinguagem específica utilizada por ATCOs e pilotos do início ao término de cada voo. É composta por um conjunto de enunciados emitidos por rádio e cobre todas as situações rotineiras, ou seja, procedimentos operacionais. Com

vocabulário reduzido, tem cerca de 400 palavras com significado ligado ao domínio da aviação, e é caracterizada pela ausência de determinantes, auxiliares e pronomes (PHILPS, 1991).

Um voo sem ocorrências começa com informações gravadas transmitidas via rádio sobre o tempo e informações sobre o aeroporto (como velocidade do vento, temperatura, pista em uso), que devem ser anotadas pelos pilotos e inseridas nos instrumentos/computadores de bordo e consideradas no momento da decolagem. Em seguida, recebem confirmação do plano de voo e possíveis alterações. Os pilotos, então, pedem autorização para o *push-back* (procedimento em que o avião é rebocado até uma área em que possa iniciar seus motores com segurança) ou instruções para o taxiamento da aeronave (procedimento em que o piloto conduz a aeronave, em solo, até a pista de decolagem). Os momentos que se seguem são a decolagem, a subida e o nivelamento da aeronave na altitude atribuída pelo plano de voo ou ATCO, até o início da descida, passando às fases finais do voo, em que os pilotos cumprem com os procedimentos de aproximação ao aeroporto (curvas, reduções de velocidade e outras sequências de instruções), pousam e taxiam a aeronave até o terminal de passageiros. A Figura 1, elaborada pela agência de aviação americana, a *Federal Aviation Administration* (FAA), representa as fases do voo:



Figura 1: Fases do voo.[10]

Alguns excertos da Fraseologia Aeronáutica para esses momentos do voo são:

Cleared to land – permissão para o pouso

[10] Disponível em: <http://www.faa.gov/nextgen/nvs/overview/capabilities/?print=go>. Acesso em: 04 nov. 2014.

Say rate of climb – informe razão de subida (‘a quantos pés por minuto a aeronave está subindo’)

Climb to flight level 350 – instrução para a subida a 35.000 pés de altura

Ready to descend – enunciado pelo piloto quando pronto para os procedimentos de descida

A comunicação usando a Fraseologia Aeronáutica é feita por rádio e, portanto, depende da ação do pressionamento do botão (PTT, ou *push-to-talk*). Sem os elementos indicativos de passagem de turno conversacional das interações face-a-face, depende de alguns procedimentos em sua operação; é iniciada pela identificação, denominada *call sign*, da estação (ATCO) ou da aeronave (piloto), como na figura a seguir:

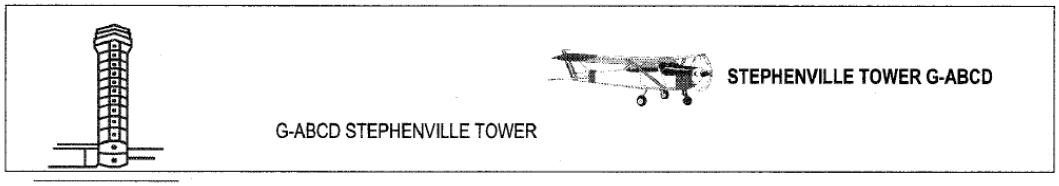


Figura 2: Exemplificação retirada do DOC 9432, Manual de Fraseologia Aeronáutica.

O exemplo acima é a representação de como a comunicação entre a aeronave e a estação deve ser iniciada. *Stephenville Tower* é o nome da estação – Torre de Controle; *G-ABCD* é o *call sign* da aeronave (neste caso, sua matrícula, e no caso da aviação comercial, o nome da empresa acompanhado do número do voo).

Todas as instruções emitidas pelo ATCO devem ser cotejadas (*readback*), ou seja, repetidas pelo piloto. Dessa forma, caso a autorização para a decolagem, por exemplo, seja emitida a uma determinada aeronave (*G-ABCD cleared for take-off runway 17*), o piloto deve responder/cotejar (*cleared for take-off runway 17 G-ABCD*) com o *call sign* ao final da emissão. Além disso, números devem ser pronunciados um por um (*flight level 350* deve ser pronunciado como *three five zero*), as letras do alfabeto são emitidas como *alpha, bravo, charlie, delta* etc., e as respostas ‘sim’ e ‘não’ são locucionadas como *affirm* e *negative.*

Essas normas, estabelecidas em prol da segurança de voo nas comunicações, junto à descrição da Fraseologia Aeronáutica, estão em documentos oficiais[11], a saber:

i) Anexo 1 – Licenças de Pessoal;

ii) Anexo 6 – Operação de Aeronaves;

iii) Anexo 10 – Telecomunicações Aeronáuticas;

iv) Anexo 11 – Serviços de Tráfego Aéreo;

v) Doc 9432 – Manual de Radiotelefonia.

Apesar de tantos documentos descrevendo e reforçando a importância da Fraseologia Aeronáutica, ela não parece ser usada de forma tão clara e transparente quanto requisitado pela OACI. Há, inclusive, estudos sobre problemas de comunicação resultantes do mau uso da Fraseologia: Cookson (2009), por exemplo, analisa fatores linguísticos que resultam de momentos de *stress* tais como a mudança de código linguístico; Cushing (1994 e 1995) estuda ambiguidade e homofonia nas transcrições de acidentes e incidentes; Drury e Ma (2002) investigam especificamente erros na Fraseologia Aeronáutica em uma pesquisa financiada pelo FAA dos Estados Unidos, e constroem um banco de dados com tais erros; mais recentemente, Barshi e Farris (2013) pesquisaram que tipo de erro pode causar mal-entendidos na Fraseologia.

O DOC 9835 aponta o uso impróprio da Fraseologia Aeronáutica como fator que contribui para acidentes aéreos (ICAO, 2010, p.1-1). O maior acidente ocorrido na história da aviação poderia ter sido evitado caso a padronização da Fraseologia Aeronáutica fosse respeitada: o piloto emitiu a frase *we´re at take-off roll* [estamos na corrida para a decolagem] ao iniciar sua decolagem, em vez do padrão *cleared for take-off* [livre decolagem]. Como não foi impedido pelo ATCO, que não entendeu a mensagem transmitida, colidiu com outra aeronave que estava na pista de decolagem, resultando na morte de quase 600 pessoas em Tenerife, em 1977 (ICAO, 1978).

[11] Informação retirada do Apêndice A (*ICAO STANDARDS AND RECOMMENDED PRACTICES* (SARPS)) do DOC 9835 (ICAO, 2010, pp. A1-6).

1.3 O conceito de Proficiência Linguística no inglês para aviação

Em 1998, um grupo de profissionais relacionados à aviação começou a se reunir para discutir propostas que resolvessem problemas de comunicação que resultavam em – ou contribuíam para – acidentes. Denominado PRICESG – *Proficiency Requirements in Common English Study Group* –, esse grupo de estudos apresentou ao secretariado da OACI, em 2004, a primeira edição do Manual de Implementação de Proficiência Linguística DOC 9835. O objetivo era impor aos Estados signatários a obrigatoriedade de assegurar ao seu pessoal proficiência "ao conduzir e compreender comunicações radiotelefônicas na língua inglesa"[12] (ICAO, 2010, p. vii), tendo como prazo para aderir aos novos critérios o período de março de 2008 a março de 2011. No Brasil, a Agência Nacional de Aviação Civil (ANAC) estabeleceu inicialmente o prazo de março de 2008, postergando-o para março de 2009 (Resolução nº 100 de 13 de maio de 2009 – DOU 91, 15/05/2009 alterado pelo DOU 97, de 25/05/2009).

O DOC 9835 delineia sua concepção de inglês para aviação considerando ensino e avaliação, e procedimentos para a implementação de provas para fins de licenciamento de pessoal, a cargo dos Estados. O documento apresenta recomendações para a elaboração e a aplicação de avaliações, enfatizando práticas voltadas à oralidade, à especificidade da língua, à visão comunicativa da linguagem e ao ensino de língua para adultos.

A primeira edição do documento generalizou a definição de inglês para aviação, estendendo-a a todos os profissionais envolvidos com a aviação, mas não especificando o uso da Fraseologia Aeronáutica, gerando, assim, equívocos na confecção de materiais didáticos e avaliações. Há, por exemplo, atividades de compreensão oral que utilizam conversas informais em corredores de empresa, em vez de focarem em conversas radiotelefônicas; há, também, desconsideração do contexto aeronáutico em exercícios que promovem trabalho pedagógico sobre a função comunicativa *giving instructions* usando um mapa ilustrativo de um bairro e requerendo instruções de como chegar à farmácia. Há ainda um desprezo ao conceito de brevidade e objetividade, em uma atividade em que há comunicação prolixa entre os dois pilotos na cabine de comando e uma ATCO divagando sobre um helicóptero próximo à pista durante a decolagem, momento crucial do voo em que os pilotos devem falar tão minimamente quanto possível (fase denominada *sterile cockpit*).

---

[12] *"[...] in conducting and comprehending radiotelephony communications in the English language."*

Após críticas a esses materiais[13], foi publicada a segunda edição, em 2010, mais assertiva, determinando que o único interesse da avaliação e do ensino do inglês para aviação fosse a comunicação entre pilotos e ATCOs, e que se considerasse a Fraseologia Aeronáutica além do inglês de uso comum (ICAO, 2010, p. 3-2). Por inglês de uso comum, ou *plain English*, entende-se:

> [...] o uso espontâneo, criativo e não codificado de uma linguagem natural, embora limitado pelas funções e tópicos (relacionados à aviação ou não), que são exigidos tanto nas comunicações radiotelefônicas como também pelos requisitos que são críticos à segurança em relação à inteligibilidade, linguagem direta, adequação, não ambiguidade e concisão[14]. (ICAO, 2010, p. 3-5)

O grupo também propôs novos parâmetros para a avaliação linguística, refutando o uso de quadros já desenvolvidos, como o *Common European Framework of Reference* (CEFR)[15], destacando que a especificidade da linguagem aeronáutica faz com que o conhecimento linguístico se sobreponha à competência comunicativa, sugestão dos CEFRs (ICAO, 2010, p. 2-4). Esses novos parâmetros são descritos na Escala de Proficiência Linguística, exposta no "*Attachment A*" do Anexo 1 da Convenção de Chicago (Licenciamento de Pessoal), e no Apêndice A do DOC 9835 (nas duas edições), dividida em seis áreas linguísticas (pronúncia, estrutura, vocabulário, fluência, compreensão e interação), consideradas em seis diferentes níveis de proficiência (1-pré-elementar, 2-elementar, 3-pré-operacional, 4-operacional, 5-avançado, 6-*expert*). O nível mínimo exigido aos profissionais para a averbação da proficiência linguística em suas licenças é 4-operacional; cada área linguística recebe uma nota de 1 a 6, e a proficiência do candidato é classificada pela menor

---

[13] Comunicação oral de Sergey Melnichenko, membro do PRICESG, no evento *One-to-Go*, promovido pela ICAO em Roma, Itália, em março de 2010.

[14] *Plain language in aeronautical radiotelephony communications means the spontaneous, creative and noncoded use of a given natural language, although constrained by the functions and topics (aviation and non-aviation) that are required by aeronautical radiotelephony communications, as well as by specific safety-critical requirements for intelligibility, directness, appropriacy, non-ambiguity and concision.*

[15] O CEFR, ou *Common European Framework of Reference*, é um documento de referência para o ensino e avaliação de línguas estrangeiras que estabelece uma base comum, com objetivos, conteúdos e princípios metodológicos, a ser utilizado internacionalmente. Disponível em: <http://www.coe.int/t/dg4/education/elp/elp-reg/cefr_EN.asp>. Acesso em: 03 nov. 2014.

nota. Por exemplo, se um candidato recebe nota 5 nas áreas linguísticas de pronúncia, estrutura, vocabulário, fluência e compreensão, mas 4 em interação, sua nota final é 4.

Devido ao escopo deste trabalho, concentramo-nos brevemente na exigência mínima – nível operacional – das áreas de estrutura e vocabulário. A Escala completa encontra-se no Anexo 1 desta dissertação.

Em estrutura, temos a seguinte descrição:

> Aspectos gramaticais básicos e padrões estruturais das frases são usados com criatividade e geralmente são bem controlados. Erros podem ser cometidos, especialmente em circunstâncias não usuais ou inesperadas, mas raramente interferem no significado.[16] (ICAO, 2010, p. 4-10; ICAO, 2004, p. A-8)

E em vocabulário:

> A riqueza e a precisão do vocabulário geralmente são suficientes para permitir a comunicação eficaz sobre tópicos comuns, concretos e relacionados ao trabalho. [O candidato] geralmente consegue valer-se de paráfrase quando faltar vocabulário em circunstâncias não usuais ou inesperadas.[17] (ICAO, 2010, p. 4-10; ICAO, 2004, p. A-8)

Nesses exemplos, a falta de precisão ou de dados objetivos, juntamente com a escassez de estudos descritivos desse tipo de linguagem (BOCORNY, 2011), pode fazer com que aqueles envolvidos com o ensino e a avaliação dependam de sua intuição e/ou experiência como especialista de línguas ao avaliar a proficiência de um candidato.

O mesmo documento, publicado em 2010, oferece uma lista que define algumas estruturas como sendo básicas e/ou complexas (Apêndice B Parte IV). Essa lista foi proposta por profissionais de língua no *Eurocontrol Institute of Air Navigation*, o centro de treinamento

[16] "*Basic grammatical structures and sentence patterns are used creatively and are usually well controlled. Errors may occur, particularly in unusual or unexpected circumstances, but rarely interfere with meaning*".
[17] "*Vocabulary range and accuracy are usually sufficient to communicate effectively on common, concrete and work-related topics. Can often paraphrase successfully when lacking vocabulary in unusual or unexpected circumstances*".

de uma organização europeia responsável pelo gerenciamento de tráfego aéreo na Europa. Alguns exemplos de estruturas básicas (Quadro 1):

Basic structures
Articles
• Adverbs of frequency
Always, Generally, Usually, Often, Sometimes, Seldom, Never, etc.
• Comparison of adjectives
• Discourse markers
Actually, Basically, Anyway, (and) yeah (more and more frequent), Listen, I mean, Let's see/Let me see, Like, Oh, Now, Okay, So, Well, You know, You see, You know what I mean, It is true, Of course, But, Still, (and) by the way, Besides, Another thing is, On top of that, So, Then, First(ly), Second(ly), etc., First of all, In the first/second place, Finally, In the end, In short
• Modal verbs
Can, May, Must, Have Got to, Should, Ought to, Would, Could, Might, Needn't, Don't have to, Mustn't
• Numbers (cardinal and ordinal)
• Passive voice
Simple present
Simple past
• Position of direct and indirect objects:
Bob sent some flowers to his girlfriend.
Bob sent his girlfriend some flowers.

Quadro 1: Parte da lista de estruturas básicas determinadas pelo *Eurocontrol Institute of Air Navigation Services.*

Em seguida, sugestões de estruturas complexas, segundo o Eurocontrol (Quadro 2):

Complex Structures
• Adjectives
Gradable and ungradable adjectives
Fairly angry (gradable)
Totally amazed (ungradable)
Prepositions after adjectives
Angry about, afraid of, etc.
Adjectives + that clause or to + infinitive
Enough, sufficiently, too + adjective
The sooner the better, etc.
• Adverbs and conjunctions
Comment adverbs
apparently, frankly, rightly
Viewpoint adverbs
biologically, ideologically, morally
Adverbial clauses of time
before, until, after, as soon as, before, when, while, hardly, no sooner, scarcely
Giving reasons
seeing that, since, in as much as, due to, owing to, with so many people ill

Quadro 2: Parte da lista de estruturas complexas determinadas pelo *Eurocontrol Institute of Air Navigation Services*

Desconsiderando o destaque feito pelo PRICESG sobre a necessidade de criação de uma escala de proficiência para a aviação devido a sua especificidade, as listas anteriores (Quadros 1 e 2) mostram um movimento contrário, trazendo amostras de língua não pertencentes ao domínio de inglês para aviação (“*Bob sent some flowers to his girlfriend”, “with so many people ill”*). Essas listas servem de parâmetros ao ensino e avaliação e causam esse distanciamento entre a prática pedagógica e o uso real da língua, já esboçado no início desta dissertação. Por esse motivo, é uma das referências ao estudo proposto aqui. O Glossário de Estruturas Básicas está apresentado no Capítulo 4 desta dissertação, junto às análises do *corpus.*

CAPÍTULO 2 – FUNDAMENTAÇÃO TEÓRICA

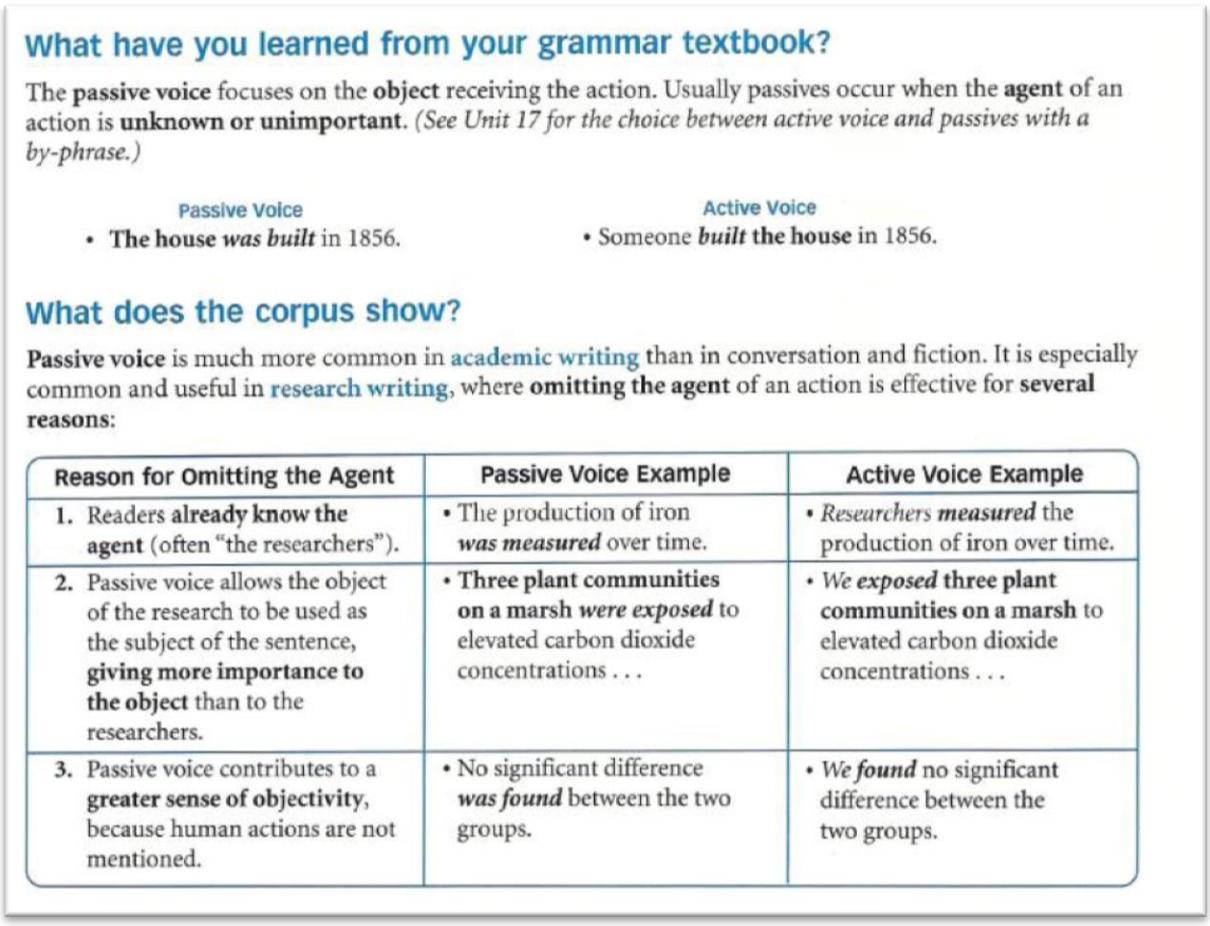
> O estudo da linguagem [...] subjaz à prática em sala de aula: é preciso entender o funcionamento de uma língua estrangeira para que seja possível facilitar a compreensão de como os alunos podem expressar suas ideias na mesma. (VIANA, 2010, p. 25)

Ocupando-nos, nesta pesquisa, com a inquietação de que há lacunas entre o que é ensinado e o que acontece na vida real no âmbito da aviação, utilizamos a Linguística de *Corpus* como metodologia empírica, pois ela possibilita a análise de língua autêntica. Neste capítulo, evidenciamos como estudos em *corpora* podem auxiliar no ensino de idiomas, tratamos dos padrões fraseológicos e da Linguística de *Corpus*, que permitem estudos descritivos por meio da busca por padrões léxico-gramaticais, e detemo-nos brevemente em *corpora* orais, interesse de nosso estudo.

2.1 O conteúdo do ensino de línguas

A dissonância entre o que é ensinado e o que acontece na vida real não é exclusividade do inglês para aviação. A falta de autenticidade dos livros didáticos tem sido alvo de críticas mesmo após o advento da abordagem comunicativa, que preconiza a contextualização da vida real nas atividades em sala de aula (Cf. RICHARDS, 2001; PEREIRA, 2004; XAVIER; SOUSA, 2008). O conteúdo dos materiais didáticos apoia-se, muitas vezes, na intuição do produtor, e estudos em *corpora* (coletâneas de textos autênticos) apresentam evidências de que nossas impressões podem ser falsas, principalmente em relação à semântica e à gramática (O´KEEFE; McCARTHY; CARTER, 2007). Vários trabalhos com esse foco comparativo têm surgido para comprovar tal afirmação (Cf. BIBER, 2006; CONRAD; BIBER, 2009; SARMENTO, 2009; CARTER; McCARTHY, 2006; SANTOS, 2011). Há diversos exemplos de como as análises feitas em *corpora* refutam conceitos antigos expostos em livros de gramática que ainda são adotados. A título de ilustração, apresentamos, na Figura 3, um

extrato do livro *Real Grammar*, de Conrad e Biber (2009, p.46), comparando o que é ensinado comumente em livros de gramática e o que é demonstrado pelo *corpus*:



What have you learned from your grammar textbook?

The **passive voice** focuses on the **object** receiving the action. Usually passives occur when the **agent** of an action is **unknown or unimportant**. *(See Unit 17 for the choice between active voice and passives with a by-phrase.)*

Passive Voice	Active Voice
• **The house *was built*** in 1856.	• Someone ***built* the house** in 1856.

What does the corpus show?

Passive voice is much more common in academic writing than in conversation and fiction. It is especially common and useful in research writing, where **omitting the agent** of an action is effective for **several reasons:**

Reason for Omitting the Agent	Passive Voice Example	Active Voice Example
1. Readers **already know the agent** (often "the researchers").	• The production of iron ***was measured*** over time.	• *Researchers* ***measured*** the production of iron over time.
2. Passive voice allows the object of the research to be used as the subject of the sentence, **giving more importance to the object** than to the researchers.	• **Three plant communities on a marsh *were exposed*** to elevated carbon dioxide concentrations . . .	• *We* ***exposed* three plant communities on a marsh** to elevated carbon dioxide concentrations . . .
3. Passive voice contributes to a **greater sense of objectivity**, because human actions are not mentioned.	• No significant difference ***was found*** between the two groups.	• *We* ***found*** no significant difference between the two groups.

Figura 3: Comparações entre o que livros de gramática informam e o que o *corpus* mostra.

Observa-se, na Figura 3, que o *corpus* expõe como a língua se concretiza entre os falantes, considerando não somente sua estrutura, mas também seus diferentes gêneros (linguagem acadêmica, oral etc.). Vemos, também, análises de 'quando' o falante opta por omitir o sujeito, não somente 'como' o faz, exemplificando a língua em uso real.

No ensino de línguas de especialidade, a investigação da linguagem a ser ensinada por meio de *corpora* é especialmente útil. É exigido do professor um conhecimento da área muito peculiar, o que o compele muitas vezes a requisitar auxílio de especialistas. Tal necessidade pode ser reduzida por meio do uso de *corpora*, que não só revelam a léxico-gramática como também aspectos textuais típicos da esfera em que se pretende trabalhar (GAVIOLI, 2005).

No âmbito do inglês para aviação, os trabalhos de Bocorny (2008) e Sarmento (2008) utilizam um *corpus* compilado a partir de manuais técnicos de operação e manutenção de aeronaves para o levantamento de padrões recorrentes nesse tipo de texto. Bocorny (2008) estuda os sintagmas nominais e Sarmento (2008), os verbos modais. Como extensão da pesquisa, Sarmento (2009) compara seus resultados a um livro didático voltado a mecânicos de aeronaves, identificando discrepâncias.

Quanto ao ensino da língua falada, há pesquisas que evidenciam uma transferência da linguagem escrita a exercícios que deveriam contemplar a linguagem falada, em materiais didáticos (O´KEEFE; McCARTHY; CARTER, 2007). Mais recentemente, porém, encontram-se estudos que, por meio de *corpora* orais e escritos coletados no âmbito acadêmico, revelam diferenças na forma de organização textual entre os dois modos, e também no uso léxico-gramatical da prosa acadêmica falada e escrita – o que tem contribuído para a produção de materiais didáticos específicos na área (BIBER, 2006). Já há materiais que incluem aspectos da linguagem conversacional – retirados de *corpora* orais – que não eram outrora evidenciados (McCARTHY; McCARTEN; SANDIFORD, 2005).

2.2 Os padrões fraseológicos

Os estudos direcionados por *corpora* têm, em seu cerne, a autenticidade da língua, ou seja, a língua em uso real. Tais estudos trouxeram uma nova visão às investigações linguísticas: a de probabilidade em oposição à possibilidade, ou seja, nem tudo que é possível ser dito é convencional. A pesquisa em LC se vale de estudos estatísticos que evidenciam padrões recorrentes produzidos socialmente. Por isso, é fundamental que a análise seja executada em coletâneas de textos autênticos.

A questão da probabilidade é também apreendida em dois conceitos fundamentais de Sinclair (1991): o princípio idiomático e o princípio da livre escolha.[18] O princípio idiomático pressupõe que o falante tenha à disposição uma série de escolhas já construídas socialmente, ou, nas palavras de Sinclair (1991, p. 109), "[...] um modo de ver o texto como o resultado de

[18] "*The idiom principle and the open-choice principle*".

um grande número de escolhas complexas".[19] O significado, ou o sentido, estaria embutido na expressão como um todo, e não nas palavras individuais que a constituem. Quando o usuário da língua não dispõe dessas expressões, pode então fazer uso do princípio da 'livre-escolha', que é o conceito que prevê liberdade de seleção por parte do falante, ou seja, a escolha de quais verbos, construções, adjetivos etc. serão usados em determinadas estruturas (HUNSTON, 2002, p. 143).

O entendimento de que "[...] o usuário recorre a estruturas com as quais teve contato anterior, seja pela fala ou pela escrita de outrem" (VIANA, 2010, p. 40) leva-nos à investigação desses padrões recorrentes, ou fraseologismos. Trata-se de uma "[...] tendência fraseológica, em que as palavras tendem a caminhar juntas e formar significado por meio de suas combinações"[20] (SINCLAIR, 2004, p. 29). Esses padrões, analisados no contexto de produção, materializam-se em unidades pré-fabricadas, ou blocos de linguagem, constituídas de escolhas que um falante tem à disposição (SINCLAIR, 1991, p. 110). Além disso, confirmam o princípio idiomático na medida em que são utilizados da mesma forma por vários falantes, e "[...] é melhor vê-los como evidência de escolhas linguísticas únicas do que como pedaços organizados no momento da fala. Fazem com que a fluência se torne uma realidade"[21] (O´KEEFE; McCARTHY; CARTER, 2007, p. 75).

As colocações, entendidas como o "[...] uso estatisticamente significativo de uma palavra ocorrer em companhia de outra" (VIANA, 2010), são exemplos de padrões recorrentes que devem apresentar uma unidade de sentido. Assim, a detecção da unidade de sentido depende da seleção manual do pesquisador, que procede à análise qualitativa dos dados levantados quantitativamente. A título de ilustração, citamos alguns exemplos extraídos de Tagnin (2013), que contrastam duas línguas – inglês e português – e o que falantes proficientes usariam em cada uma delas:

Outside chance / hipótese remota

Main course / prato principal

Make a joke / contar uma piada

---

[19] *"[...] a way of seeing language text as a result of a very large number of complex choices."*
[20] *"[...] phraseological tendency, where words tend to go together and make meanings by their combinations."*
[21] *"[...] are best viewed as being evidence of single linguistic choices rather than assembled piece by piece at the moment of speaking. They make fluency a reality."*

Come into force / entrar em vigor

Cada um desses itens é uma colocação, pois é uma combinação consagrada, evidenciada a partir de análises probabilísticas.

É também o olhar do analista que irá reconhecer outro aspecto fundamental no estudo pragmático: a prosódia semântica, ou o efeito que certo item lexical, em seu padrão fraseológico, causa no interlocutor (HUNSTON, 2002). Como exemplo, mencionamos a expressão '*impending doom / disaster / horror*' (ruína, desastre, horror iminente), cujas coocorrências sugerem uma prosódia negativa da palavra *impending* (TAGNIN, 2013, p.53). O'Keefe, McCarthy e Carter (2007) apresentam outro exemplo: a expressão *peace and quiet*, utilizada somente em contextos em que há comparações com um momento turbulento; portanto, o uso de tal expressão simplesmente para enfatizar uma necessidade de paz, sem um contraste negativo anterior, não seria considerado natural.

Concebendo a língua dessa forma, portanto, adotamos a visão de que gramática e léxico não são entidades separadas, ou, nas palavras de Sinclair (1991, p. 6), "cada significado pode ser associado com um padrão formal distinto".[22] Dessa maneira, o uso de uma palavra está condicionada ao seu entorno. Um sentido possui "ramificações lexicais e gramaticais"[23], segundo Sinclair (1991, p. 8), e a escolha por um padrão que o represente se faz independente de os itens que o compõem serem gramaticais ou lexicais.

A não diferenciação entre palavras lexicais e gramaticais pode ser verificada em estudos que evidenciam esse fato; a expressão *come to a head*, por exemplo, em que o artigo '*a*' não pode ser substituído por outro artigo, como '*the'*, sendo parte dessa unidade pré-fabricada, analisada por Sinclair (1999, apud HUNSTON, 2002, p. 150), mostra que "[...] palavras gramaticais participam do princípio idiomático exatamente como as palavras lexicais"[24]. Em pesquisa sobre os verbos modais *may* e *must*, verifica-se a possibilidade de investigação desses verbos por meio de seus significados da mesma forma que verbos lexicais (HUNSTON, 2002). Diferentes significados podem ser atribuídos a determinada palavra dependendo de sua conjugação ou transitividade, assim como constatado em um estudo de

[22] *"[...] each meaning can be associated with a distinctive formal patterning".*
[23] *"Lexical and grammatical ramifications".*
[24] *"Grammatical words such as '*a*' participate in the idiom principle just as lexical words do."*

Sinclair (1991, p. 53-66), que explora variados significados do verbo *yield* dentro de contextos – e conjugações – distintos.

2.3 A Linguística de *Corpus*

A Linguística de *Corpus* (LC) se ocupa dos estudos que investigam os padrões linguísticos convencionalizados, por meio de *corpora*. Assumimos, aqui, a definição de *corpus* como "[...] uma coletânea de textos, necessariamente em formato eletrônico, compilados e organizados segundo critérios ditados pelo objetivo de pesquisa a que se destina" (TAGNIN, 2013, p. 26). Estudos gerados pela LC possibilitam descobertas em diversas áreas, tais como Linguística Descritiva, Linguística comparativa, Terminologia, Literatura, Tradução e Ensino.

À LC é atribuído um caráter científico; para Sinclair (2004, p. 9), é necessário um "[...] equilíbrio entre especulação e fato"[25] nos estudos linguísticos, pois há aspectos da língua que podem passar despercebidos a olho nu. Além disso, a busca por padrões dificilmente é bem sucedida quando nos apoiamos apenas em estudos realizados manualmente. O acesso às ferramentas computacionais permite uma busca maior de dados, que, por sua vez, contribui para a qualidade da investigação linguística (SINCLAIR, 1991).

Reforçando a característica empírica da LC, Hunston (2002, p. 3) defende seu uso para a descrição de língua afirmando que "[...] um *corpus* não contém informações novas sobre a linguagem, mas as ferramentas de busca oferecem novas perspectivas sobre aquilo que parece familiar"[26]. Tognini-Bonelli (2001, p. 18), ao exemplificar usos da LC, afirma que, "[...] sendo um *corpus* uma coleção de textos, o objetivo da LC tem sido visto, com razão, como a análise e a descrição do uso da língua, conforme produzido em texto(s)".[27] Hunston (2002, p. 137), também abordando a necessidade da descrição de língua para o ensino de línguas estrangeiras, sugere a busca por meio da LC pelos fraseologismos, que são "a tendência de

[25] "*[...] balance between speculation and fact.*"
[26] "*A corpus does not contain new information about language, but the software offers us a new perspective on the familiar.* "
[27] "*Given that a corpus is a collection of texts, the aim of corpus linguistics has rightly been seen as the analysis and description of language use, as realised in text(s).*"

palavras coocorrerem, não aleatoriamente, ou até mesmo apenas de acordo com regras gramaticais, mas em sequências preferidas"[28].

Os estudos mencionados na seção 2.1 deste capítulo foram todos viabilizados pela LC. A extração linguística é norteada por dados estatísticos, que revelam os padrões e as probabilidades de ocorrências – e coocorrências. A partir daí, a análise manual permite identificações de fraseologismos, e assim, da convencionalidade característica da língua investigada.

2.4 *Corpora* orais

Para que a análise seja realmente representativa da amostra de linguagem que se pretende estudar, o *corpus* precisa ser muito bem definido e criteriosamente compilado. A pergunta de pesquisa deve nortear seu escopo, e a construção do *corpus* deve ser feita com textos autênticos.

Há vasta literatura e pesquisa sobre *corpora* escritos. Como critérios para compilação, sugeriu-se considerar (cf. BERBER-SARDINHA, 2004):

- o modo (escrito ou falado)
- o tempo (momento de produção)
- a seleção (escopo da pesquisa)
- o conteúdo (generalizado, específico, regional etc.)
- o tamanho ou representatividade (número de palavras / textos)
- o balanceamento (quantas palavras por textos / número de textos)

[28] "*The tendency of words to occur, not randomly, or even in accordance with grammatical rules only, but in preferred sequences.*"

Porém, em relação a *corpora* orais, há falta de padronização entre os pesquisadores a respeito de quesitos como transcrição, tamanho, representatividade, balanceamento etc. Como exigem maior investimento de tempo e dinheiro, existem em menor número se comparados a *corpora* escritos (McCARTHY; O´KEEFE, 2009). Ademais, há obstáculos para a seleção de textos, para a obtenção de permissão dos falantes, e para a própria definição do que é um texto (ATKINS; CLEAR; OSTLER, 1992). Diferentemente do que se entende por texto escrito – organizado dentro de um diagrama estabelecido socialmente – o texto oral é espontâneo, com uma organização estabelecida durante a execução, e altamente dependente do contexto de produção. A alta frequência de dêiticos em *corpora* orais é um bom exemplo do vínculo com o momento de produção da fala. Essa característica momentânea e dialógica impossibilita uma definição do que é começo, meio e fim em textos orais produzidos espontaneamente, por exemplo. Há trabalhos que apresentam dados produzidos com controle da fala, principalmente aqueles cujo objetivo é o treinamento de ferramentas (como *softwares* de reconhecimento de fala), porém são restritos a esses fins.

Para a investigação de uma gama ampla de eventos da fala, o contexto natural da produção é essencial. Essa é outra dificuldade, apresentada por Čermák (2009): a espontaneidade da língua pode esbarrar em questões como atribuição de direitos autorais. Porém, ao informarmos o falante sobre a gravação de sua produção linguística, incorremos na perda da espontaneidade. É, portanto, o próprio objetivo da pesquisa que define a arquitetura de cada *corpus*, em cada quesito. Temos como exemplo o segmento oral do *British National Corpus* (BNC) que, para investigações de variedades da fala, foi compilado considerando diversas faixas etárias, regiões, classes sociais etc. Já o *Corpus of London Teenage speech* (COLT) representa apenas uma faixa etária – adolescentes –, e só pode conter informantes pertencentes a ela (McCARTHY; O´KEEFE, 2009, p. 1013).

A investigação de *corpora* também possibilita a análise das diferenças entre a linguagem oral e a escrita, além do uso de contrações, repetições, registro formal ou informal etc. Muitos pesquisadores defendem a ideia de que as duas linguagens têm estruturas completamente diferentes (BIBER et al, 1999; BRAZIL, 1995; McCARTHY; CARTER, 2001). A estrutura frasal, segundo eles, tem que ser reexaminada (McCARTHY; CARTER, 2001, p. 51), já que a língua oral é composta por "enunciados factuais" (*factual utterances*) em vez de "frases corretas" (*correct sentences*). Portanto, é necessário um cuidado extra com a transcrição para que as análises dessas diferenças, caso existam, possam ser reveladas

(McCARTHY; O´KEEFE, 2009). Um exemplo é o enunciado abaixo, que é uma produção de linguagem autêntica, com *false starts*, repetições, hesitações, interrupções, contrapondo-se a uma frase com sujeito, verbo e complemento:

e então tinha muito texto do tio Carlos então ele falava ah ele é tio da minha tia

Quadro 3: Elocução extraída do C-Oral Brasil.

A elocução acima, extraída do C-Oral Brasil, um *corpus* oral de português brasileiro (RASO, 2012, p. 93), exemplifica um enunciado emitido como uma única unidade tonal, verificável por meio do áudio ou por uma representação que privilegie características pragmáticas. Transcrições feitas sem a consideração de aspectos importantes da fala podem trazer algumas consequências indesejáveis para a pesquisa, visto que, nesse caso, somente alguns aspectos – como o gramatical e o lexical – podem ser investigados, não contemplando, portanto, análises importantes tais como a prosódica e a fonológica, entre outras (LEECH, 2000, p. 678). Além disso, na escrita é possível dividir orações, frases, sintagmas. Já na fala, 30% das unidades autônomas não contém verbos (MONEGLIA; CRESTI, 2006; RASO, 2012). A transcrição, portanto, deve manter as características interacionais da linguagem, como mudança de turno, entonação etc., mas também precisa ser analisável pelas ferramentas de busca (que se utilizam de palavras escritas ortograficamente). Se a compilação de um *corpus* oral deve considerar todo o comportamento verbal da situação (ADOLPHS; KNIGHT, 2010; RASO, 2012; SINCLAIR, 2004), como, então, evidenciar essas informações durante a transcrição?

Uma possibilidade está no alinhamento dos *corpora* orais, que consiste no uso de *softwares* especializados para disponibilizar, lado a lado, transcrição e áudio e, assim, viabilizar a análise pragmática, ou fonológica, por exemplo (cf. MELLO, 2012; McENERY; HARDIE, 2012; RASO, 2012). Outra está em buscar um modelo de transcrição que permita a análise da linguagem oral como ela se concretiza, com interrupções, elipses, dêiticos etc. O modelo escolhido foi a Teoria da Língua em Ato, explicada a seguir.

2.4.1 A Teoria da Língua em Ato

A Teoria da Língua em Ato (doravante TLA) (CRESTI, 2000, apud RASO, 2012, p. 92), que "[...] identifica a unidade de referência da fala na dimensão do enunciado", surge como uma solução plausível para o problema identificado na seção anterior. O C-ORAL-ROM (*Corpora of Spontaneous Spoken Language of the Main Romance Languages*), editado por Cresti e Moneglia (2005), é o maior exemplo de um *corpus* oral compilado levando-se em consideração a segmentação do fluxo da fala em unidades. Contém produções orais espontâneas, em contextos formais e informais, em quatro línguas, a saber, italiano, espanhol, português europeu e francês (BLANCHE-BENVENISTE, 2006), e uma extensão para o português brasileiro, nomeado C-ORAL BRASIL (RASO; MELLO, 2012), compilado na Universidade Federal de Minas Gerais.

A TLA considera "a identificação das unidades de referência [como] a principal informação adicionada para a análise linguística de um *corpus* falado"[29] (MONEGLIA; CRESTI, 2006, p. 91). Austin (1962) abordou a questão do enunciado (*utterance*) como unidade de significado mínima para a fala, com um significado pragmático visto além da gramática tradicional, e que

> [...] inclui o enunciado de certos ruídos, o enunciado de certas palavras em certa construção, e o seu enunciado com certo 'significado' no sentido filosófico preferido daquela palavra, i.e. com certo sentido e com certa referência.[30] (AUSTIN, 1962, p. 94)

Definida de acordo com essa tradição pragmática de Austin (1962), a unidade de referência, para a TLA, "[...] é a entidade linguística mínima que pode ser interpretada pragmaticamente" (MONEGLIA; CRESTI, 2006, p. 91).

---

[29] "*The identification of the units of reference [is] the main added information for the linguistic analysis of a spoken corpus.*"
[30] "*[…] includes the utterance of certain noises, the utterance of certain words in a certain construction, and the utterance of them with a certain 'meaning' in the favourite philosophical sense of that word, i.e. with a certain sense and with a certain reference.*"

Para que a análise das unidades de referência seja possível, as quebras prosódicas precisam ser bem definidas na transcrição. Mello (2012, p. 37) exemplifica a importância das quebras no segmento transcrito abaixo, e afirma que "a interpretabilidade de (1) está totalmente condicionada a sua contrapartida prosódica, ou seja, acústica". Há três diferentes exemplos de como a frase (1) poderia ter sido enunciada (a, b e c):

1. aquele ali aquele ali já deu altura

a) aquele // ali // aquele ali já deu altura

b) aquele ali // aquele // ali já deu altura

c) aquele ali // aquele ali já // deu altura

Sinais de pontuação não conseguiriam representar as pausas, as interrupções e a pragmática do enunciado acima, e a análise linguística não poderia abranger a riqueza da produção, pois ignoraria aspectos relevantes à fala informal investigada no C-ORAL-BRASIL. Aqui, "cada sequência que se conclui com uma quebra terminal no fluxo da fala é considerada um enunciado" (RASO, 2012, p. 77).

Comparemos com o exemplo no Quadro 4, retirado do *Michigan Corpus of Academic Spoken English* (MICASE)[31], em que os enunciados são tratados como frases com alguns indicativos (hifens, reticências, colchetes, barras, mudança de cor na fonte). Conseguimos facilmente perceber a linguagem oral, principalmente pelas contrações, mas os enunciados extremamente longos e incompletos (característicos da linguagem oral) causam certa estranheza para a análise:

[31] Disponível em: <http://quod.lib.umich.edu/m/micase/>. Acesso em: 02 nov. 2014.

Title: Freshman Orientation Tour
Transcript ID: TOU999JU030
Academic Division: Not Applicable/Other
Publisher: Michigan Corpus of Academic Spoken English, English Language Institute, University of Michigan
Interactivity Rating: Mostly interactive
Number of Participants:
Students: 15
Speakers: 13
Recording Date: August 3, 1998
Recording Duration: 77 minutes
Word Count: 12995

S1: okay. so. i got a couple quick_ please listen to me please please please listen to me please listen to me, thank you. i got a couple quick administrative things to go through real quick. not only did i have to ask you about this English Language Institute thing, but i also have to ask_ guys move in come on come on come on. this is ridiculous. i'm not gonna shoot you guys i promise i'm not gonna shoot anybody. now, couple quick things, first thing that the c- campus tour is my favorite part of the entire orientation thing. this is where i get off the most, i absolutely get a big kick out of this because i think i give a pretty decent tour, and i've been doing tours actually for the last two years. so anyway, basically i have a really good time giving a tour. if you do not want to be here, i don't want you to be here and the reason is, because, i don't want you to cramp my style i don't want you to cramp anybody else that wants to actually be here, so if you don't want to be here that's fine. i won't hold it against you it's just like that English Lang- Language Institute thing, that's okay if you don't wanna be here cuz you got something else to do you got another agenda please go. because... hey, if you're gonna be like all pissed off that we made you come, it's not a big deal i don't really care. so honest to God if you guys don't wanna be here i'm not gonna be pissed at you. so if you wanna leave, be back at eight o'clock at the same, um auditorium that we were at this uh morning, for our_ for the comedy troupe thing.

S2: can you tell us why you have a water gun?

S1: i can tell you why i have a water gun. now the reason i have a water gun is because not so that i can shoot you guys [SU-f: no no no] i promise see this is this is definitely group love this is spade love okay? definitely group love all the other groups pretty much suck okay? family at U-of-M talked about that a lot this morning but all the other groups they suck. now okay, we're divided, okay, so we're gonna be yelling a lot of, semi-obscene things at the other groups, we're gonna be shooting them with water

S3: you have like the smallest water gun though *<LAUGH>*

SU-f: yeah you do.

S3: you really do

S1: really?

S3: um yes.

S1: you like testing [S5: no] *<SS: LAUGH>* the water gun? you wanna see really what the water gun can do? no i didn't think so. i didn't think so... it's leaking... trust me.

S3: okay.

Quadro 4: Diálogo transcrito de uma apresentação do *campus* universitário a novos alunos nos Estados Unidos.

Frases assim descritas dificultam a análise, pois se considerarmos a busca pelas fraseologias da linguagem produzida, há a possibilidade de conectarmos palavras que pertençam a outras estruturas fraseológicas. A segmentação prosódica sugerida pela TLA, que

identifica as unidades tonais como já imbuídas de significado, facilita o estudo das estruturas léxico-gramaticais proposto nesta pesquisa. Considerando a unidade tonal dentro dos enunciados, podemos também identificar aquelas com unidade de significado e, assim, identificar padrões fraseológicos pertinentes à língua falada.

CAPÍTULO 3 – METODOLOGIA

> *The beginning of any corpus study is the creation of the corpus itself. The decisions that are taken about what is to be in the corpus, and how the selection is to be organized, control almost everything that happens subsequently. The results are only as good as the corpus. (SINCLAIR, 1991, p. 13)*[32]

Para investigarmos os padrões léxico-gramaticais do inglês para aviação em contextos não rotineiros, objetivo desta pesquisa, compilamos um *corpus* a ser analisado sob a luz da LC. Neste capítulo, apresentamos o *corpus* de estudo e sua arquitetura, delimitações, balanceamento, transcrições e armazenamento. Também propomos o método de análise utilizado para o levantamento dos fraseologismos compartilhados pela comunidade aeronáutica.

3.1 O *corpus* de estudo

Para a investigação do inglês comum utilizado por pilotos e ATCOs que ocorre nas situações anormais, compilamos o *corpus* aqui apresentado. São comunicações radiotelefônicas com problemas técnico-operacionais, entre ATCOs e pilotos, assim como sugerido pela definição de inglês para aviação da OACI, a retomar:

> [...] o único objeto dos requerimentos de proficiência linguística da OACI é a comunicação da **radiotelefonia** aeronáutica [...] que **corresponde à porção limitada de usos de linguagem de somente dois profissionais** – controladores de tráfego aéreo e tripulação de voo. Inclui Fraseologia Padrão

[32] "O início de qualquer estudo de *corpus* é a criação do próprio *corpus*. As decisões que são tomadas sobre o que deve estar contido no *corpus*, e como a seleção deve ser organizada, controlam quase tudo o que acontece subsequentemente. Os resultados são tão bons quanto o *corpus*".

> da OACI e o uso de inglês de uso comum[33]. (ICAO, 2010, p. 3-2, grifo nosso)

No início da compilação, análises prévias indicavam uma tendência de problemas com motores; assim, buscamos outra forma, menos aleatória, de selecionar os textos. Adotamos outro documento da OACI, que padroniza o registro de relatórios de acidentes e incidentes aéreos no mundo, como solução para a questão do balanceamento do *corpus*. O documento de padronização de categorias de ocorrências aéreas que nos permitiu essa decisão é denominado Taxonomia de Categoria de Ocorrência (*Occurence Category Taxonomy*)[34], que consiste em uma lista de ocorrências que podem acontecer no meio aeronáutico, em diferentes fases do voo, envolvendo falhas na aeronave, problemas meteorológicos ou fatores humanos (falha humana, ferimentos a bordo, ameaça de bomba, entre outros). A seguir, a tabela com a lista de taxonomias, com seus nomes em inglês (devido à padronização internacional das siglas) e suas respectivas traduções (nossas) em colchetes:

	Termo	Sigla
Primary (15)	***A**bnormal **r**unway **c**ontact* [contato anormal com a pista]	ARC
	***Bird**strike* [colisão com pássaro]	BIRD
	***C**ontrolled **f**light **i**nto or toward **t**errain* [Colisão controlada contra o solo]	CFIT
	***C**ollision with obstacle(s) during **t**ake-**o**ff and **l**anding* [colisão com obstáculo durante a corrida na pista]	CTOL
	***F**ire/smoke (**n**on-**i**mpact)* / [Fogo ou fumaça – impacto não é a causa]	F-NI
	***G**round **Col**lision* [Colisão em solo entre aeronaves ou aeronave com veículos]	GCOL
	***L**oss **o**f **c**ontrol – inflight* [Perda de controle em voo]	LOC-I
	*Airprox/ ACAS alert/ loss of separation/ (near) **mid**air **c**ollisions* [Quase-colisão em voo / Alerta de TCAS]	MAC
	Ground Handling [Serviços de solo] (a sigla RAMP se deve à área de	RAMP

[33] "*The sole object of ICAO language proficiency requirements is aeronautical radiotelephony communications, a specialized subcategory of aviation language corresponding to a limited portion of the language uses of only two aviation professions — controllers and flight crews. It includes ICAO standardized phraseology and the use of plain language.* "

[34] Esse documento faz parte do Anexo 13 da Convenção de Chicago, por sua vez denominado Investigações de Acidentes com Aeronaves (*Aircraft Accident Investigation)*, e foi desenvolvido pela OACI por uma equipe nomeada *Commercial Aviation Safety Team* (CAST), formada por oficiais de governo e líderes da aviação. Disponível em: <http://www.intlaviationstandards.org/>. Acesso em: 08 nov. 2014.

	movimentação de aeronaves, chamada de *ramp*, ou *apron* – pátio, em português)	
	***R**unway **e**xcursion* [Excursão de pista]	RE
	***R**unway **i**ncursion - wildlife presence* [Incursão de pista por animais]	RI-A
	***R**unway **i**ncursion* - **v**ehicle, **a**ircraft or **p**erson [Incursão de pista]	RI-VAP
	***S**ystem/**c**omponent **f**ailure or malfunction (**n**on-**p**owerplant)* [Pane de sistema que não o motor]	SCF-NP
	***P**owerplant failure or malfunction* [Pane de motor]	SCF-PP
	***U**ndershoot/**o**vershoot* [Não alcance à pista / ultrapassa o limite da pista]	USOS
Secondary (18)	*Communication Navigation Surveillance/**A**ir **T**raffic **M**anagement* [Gerenciamento de Tráfego Aéreo, comunicações ou assistência à navegação]	ATM
	***L**oss **o**f **c**ontrol – **g**round* [Perda de controle no solo]	LOC-G
	***Turb**ulence encounter* [Turbulência]	TURB
	***Fuel** related* [Problemas de combustível]	FUEL
	***A**ero**dr**o**m**e* [Problemas com a estrutura aeroportuária]	ADRM
	***L**ow **alt**itude operations* [Operações em baixa altitude]	LALT
	***F**ire/smoke* (**post**-impact) / [Fogo ou fumaça pós-impacto]	F-POST
	***W**ind**s**hear or **t**hunderstorm* [Tesoura de vento ou tempestade]	WSTRW
	***Ic**ing* [Gelo]	ICE
	***Evac**uation* [Evacuação]	EVAC
	***Sec**urity related* [Atos contra a segurança]	SEC
	***Cabin** safety events* [Eventos de cabine]	CABIN
	***A**brupt **man**oeuvre* [Manobras abruptas]	AMAN
	***L**oss **o**f **li**fting conditions en-route* [Perda de sustentação em rota]	LOLI
	U**nintended flight in **IMC – instrument meteorological conditions [Voo IMC – condições meteorológicas por instrument – não intencional]	UIMC
	***G**lider **tow**ing related events* [Eventos relacionados ao reboque de planador]	GTOW
	***Ext**ernal **l**oad related occurrences* [Ocorrências com o procedimento de carga]	EXTL
	***Unk**nown or undetermined* [Desconhecido ou indeterminado]	UNK

Tabela 1: Termos Taxonômicos de Categoria de Ocorrência.

Para o balanceamento do *corpus* foi determinado que houvesse três ou quatro áudios para cada categoria da Tabela 1. Nesta pesquisa, foram desconsideradas três categorias, pois são relacionadas a eventos em que não há comunicação entre ATCOs e pilotos (como reboque de planador e procedimento de carga, por exemplo), além da categoria *Unknown or undetermined*, uma vez que todos os áudios pertencem a categorias conhecidas. Há, ainda, áudios que apresentam problemas pertinentes a mais de uma categoria; para casos assim, as escolhas foram baseadas considerando a melhor distribuição do *corpus*.

Os áudios foram retirados da internet, principalmente de um *website*, o *LiveATC.net*, que disponibiliza frequências transmitidas via rádio, ao vivo, de diferentes aeroportos e órgãos de controle do mundo todo. Essas frequências são, então, transformadas em arquivos de áudio segmentados pela duração de 30 minutos e armazenados no banco de dados do próprio *site*, onde podem ser acessados por até 30 dias.

Dave Pascoe, fundador e moderador do *website*, e outras pessoas afiliadas ao *LiveATC.net*, são voluntários que residem nas proximidades de aeroportos ao redor do mundo e possuem equipamentos que captam frequências radiotelefônicas. Os equipamentos devem ter alta capacidade, tanto para a captação dos áudios quanto para a sua disponibilização no *site*. É exigência para ser *feeder* (membro voluntário que disponibiliza o acesso às frequências) possuir *scanner* (aparelho que captura a frequência), Internet conectada 24 horas por dia, antena, cabo de áudio, computador de alta capacidade e *software* fornecido pelo próprio *website*.[35]

Embora o acesso às comunicações tenha sido facilitado pelo *website*, a grande maioria delas é rotineira, ou seja, não apresenta problemas técnico-operacionais, condição estabelecida nesta pesquisa para a coleta dos áudios. Mais uma vez, porém, o próprio *site* auxilia no processo: entusiastas do mundo inteiro escutam as comunicações e buscam gravações assim que algum acidente ou incidente é transmitido pela mídia; além disso, os próprios *feeders* ficam atentos às comunicações captadas pelos seus *scanners*. Quando algo incomum ocorre, os ouvintes disponibilizam, em uma área restrita do site, acessada somente por membros registrados gratuitamente, o arquivo em que tal situação ocorre. Abaixo, a tela do *website* em que se encontram os arquivos de áudio:

[35] Informações de como tornar-se um voluntário retiradas de: <http://www.liveatc.net/about.php>. Acesso em: 09 nov. 2014.

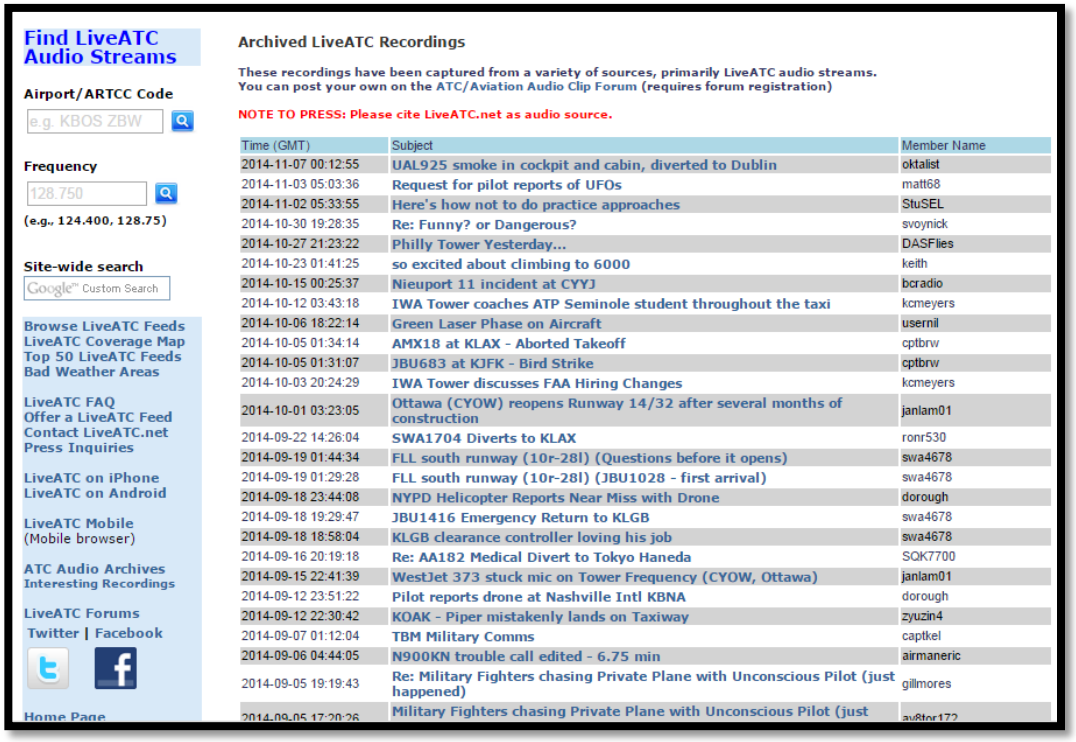


Figura 4: *Site* LiveATC.net, com a página *Interesting Recordings*.

Na Figura 4, pode-se ver, à esquerda, o último item – *Interesting Recordings* –, e ao centro, o início da lista com os arquivos de áudio.

Outros áudios foram retirados de bancos de dados oficiais em que se encontram relatórios de acidentes ocorridos. Por esse motivo, cada transcrição tem sua fonte identificada, junto a outras informações, no cabeçalho do texto, a ser explicado na seção 3.1.3.

Limitado a 92 textos e 70.323 palavras para este estudo, o *corpus* do inglês aeronáutico em situações anormais foi compilado em quase quatro anos. Devido à dificuldade de se estabelecer um tamanho para *corpora* orais (ČERMÁK, 2009), durante a compilação precisamos definir critérios que norteassem a melhor representatividade do tipo de linguagem que buscávamos investigar. Um dos fatores que direcionou a compilação foi a constante análise da riqueza lexical, que oscilou entre 3% e 4% desde que o *corpus* contabilizava 40.000 palavras. A riqueza lexical é medida estatisticamente por meio da razão forma/item (*type/token ratio)* do *corpus*, ou seja, calculando o número total de palavras (*tokens*) em um texto e relacionando-o com a quantidade de palavras não repetidas (*types*). A fim de exemplificação, vejamos a seguinte elocução:

i think yeah that was good experience **yeah yeah i think it's a- i think it's** really rich, mhm **i** mean just as **a**, kind of educational opportunity looking around different, programs and stuff on campus.

Quadro 5: elocução extraída do *corpus* MICASE.

Há 38 palavras (*tokens*) no enunciado do Quadro 5. Porém, algumas são repetidas (em negrito), o que equivaleria a dizer que há menos tipos de palavras diferentes (*types*), pois a palavra *we*, por exemplo, aparece três vezes. Assim, há 28 palavras diferentes (*types*) nesse excerto. A medida feita é a multiplicação de *types* por 100. O resultado, então, é dividido pelo número de *tokens*:

Riqueza lexical (%) =	*types* x 100	28 x 100 = 73,68%
	tokens	38

Quadro 6: Exemplo de medida estatística da riqueza lexical.

A elocução do Quadro 6 tem, portanto, 73,68% de riqueza lexical. Quanto maior o número de *types*, maior é a riqueza lexical do texto ou *corpus* (BERBER-SARDINHA, 2004, p. 94). Essa análise – feita automaticamente pelas ferramentas computacionais que auxiliam na análise de *corpus* – foi essencial para esta pesquisa, pois a alta repetição de palavras indicou que o balanceamento do *corpus* deveria ser feito de forma distinta, e não baseado em tamanho, ou número de palavras. A razão *type/token* do estado atual do *corpus* é de 3,68%.

3.1.1 O texto

Cada áudio foi transcrito manualmente; em seguida, as transcrições foram voluntariamente revisadas por pilotos, ATCOs e professores de inglês para aviação. As

dificuldades nesse processo foram inúmeras: além da seleção baseada na Taxonomia de Categorias – o que dificultou o processo de compilação, já que deveríamos buscar ocorrências muitas vezes raras –, a transcrição foi laboriosa (cada minuto de áudio representou cerca de 30-40 minutos de transcrição) e a revisão, demorada. Com um total de quase 100 textos, foi desafiador encontrarmos profissionais relacionados à área dispostos a se desprender de suas tarefas diárias para se dedicarem à revisão textual.

Os áudios, por serem oriundos dos mais diversos locais, produzidos por falantes de diferentes nacionalidades, via rádio, captados por *scanners* via antena, são majoritariamente de baixa qualidade. Acrescente-se a isso o fato de serem extremamente específicos, pois cada um traz nomes de fixos (antenas balizando aerovias para auxílio à navegação) estabelecidos para aquelas áreas, nomes de empresas, nomes de cartas[36] (mapas que orientam os pilotos em solo ou em voo) de decolagem e de pouso e nomes de aeroportos que são, em muitos momentos, incompreensíveis até para ouvidos experientes. Nas situações em que não houve compreensão da fala, registramos a ocorrência como <unreadable>, ou seja, ininteligível (conforme jargão aeronáutico), e entre os sinais de maior e menor para que as ferramentas de busca desprezem essa palavra, já que não faz parte do texto verbal.

Os textos não têm limites pré-determinados, conforme mencionado no Capítulo 2 desta dissertação; a condição é que contenham a porção designada ao problema, que é o interesse desta pesquisa. Alguns áudios têm menos de 1 minuto, outros mais de 10. Os áudios são, muitas vezes, editados, com períodos de silêncio removidos, para facilitar o trabalho da transcrição.

A transcrição foi feita manualmente utilizando-se o *software SoundScriber*[37], um leitor de mídia que possibilita a utilização de teclas de atalho para funções como *play* e *pause* sem que seja necessária a saída da tela do processador de texto. O *software* mencionado foi desenvolvido por Eric Breck e é utilizado pela equipe do MICASE[38] (*The Michigan Corpus of Academic Spoken English*), um *corpus* oral (vide seção 3.2.2) compilado pela Universidade de Michigan. Houve tentativa de utilizar *software* de reconhecimento de voz, mas as dificuldades com o excesso de ruído e variedade de vozes impossibilitaram esse trabalho.

[36] Cartas aeroportuárias, de saída ou aproximação, são documentos oficiais em que se registram a diagramação dos aeroportos e procedimentos de decolagem e aproximação.
[37] Disponível em: <http://www-personal.umich.edu/~ebreck/code/sscriber/>. Acesso em: 02 nov.14.
[38] Disponível em: <http://quod.lib.umich.edu/m/micase/>. Acesso em: 02 nov. 14.

A transcrição preserva as quebras prosódicas sugeridas pela TLA. Vale ressaltar, aqui, que a troca de turnos entre falantes na radiotelefonia é concretizada de maneira distinta à comunicação face a face: o falante passa o turno ao interlocutor no momento em que solta o botão do rádio (cujo nome é PTT – *push-to-talk*). Portanto, não há interrupções de turno, mas sim da própria elocução. Veremos, a seguir, um exemplo da preservação das quebras prosódicas e da mudança de turno:

<ATC> American nine sixty-three heavy / runway one seven right / line up and wait // <Pilot> Line up and wait one seven right / American nine six three heavy // <Pilot> Center / we got an engine fire / right engine / will call you back <unreadable> // start dumping fuel //

Quadro 7: Exemplos de quebras prosódicas e mudanças de turno extraídos do *corpus* de estudo.

Nesse extrato, mesmo que as instruções sejam dadas sem interrupções ou pausas na fala, é possível definir quais são as unidades tonais. Os dois primeiros enunciados são contínuos e a quebra prosódica terminal se faz na mudança de interlocutor. Porém, na última elocução há quebras terminais dentro do mesmo enunciado, emitido pelo mesmo falante.

Considerando elementos não verbais, tais como risos ou interrupções de enunciados, adotamos o etiquetamento de suas descrições. As etiquetas servem à LC de várias formas, e podem ter naturezas distintas; podem ser, por exemplo, morfossintáticas, ou podem identificar partes do texto (cf. TAGNIN; TEIXEIRA, 2004). No *corpus* aqui descrito, a etiquetagem identifica elementos não constituintes do texto oral e é representada pela palavra ou expressão descritiva desse elemento entre os sinais de maior ou menor para, assim, ser ignorada pelas ferramentas de busca. São apenas informações extratextuais que julgamos pertinente expor durante o processo de transcrição, levando em consideração a reusabilidade do *corpus*.

Estas são as etiquetas utilizadas no *corpus*:

<laugh>	Risos
<break>	Interrupção da fala do meio do enunciado
<phew>	Interjeição de alívio
<pause>	Pausas mais longas no meio do enunciado
<sigh>	Suspiro

<unreadable>	Palavra retirada da Fraseologia Aeronáutica, que significa ‘ininteligível’

Tabela 2: Etiquetas utilizadas no *corpus.*

A hesitação e outros ruídos foram descritos na transcrição como *uh, er, um*, dependendo de como foram entendidos:

1. <ATCO> // Jesus / JetBlue / **arrgh** // Clearance four zero
2. <pilot> Hey **uh** ground this is Lufthansa
3. <ATCO> **Um** / aircraft calling / please advise again with the tail number / so I can advise them //

Quadro 8: Exemplos de hesitações e suas grafias, extraídos do *corpus* de estudo.

O entendimento claro dessas hesitações é, por vezes, árduo. Em casos assim, optamos pela padronização utilizando apenas *uh*. O registro dessa hesitação é importante, pois participa do texto oral e pode fomentar futuros estudos sobre fluência, por exemplo.

3.1.2 Os falantes

São ATCOs e pilotos, por vezes pessoal de assistência à emergência, que vivenciam a situação anormal. Houve preocupação em preferenciar produções orais de nacionalidades diferentes. Porém, devido à maior disponibilidade de áudios nos Estados Unidos da América, onde a captação e a propagação das frequências radiotelefônicas são permitidas, a maioria dos falantes é de origem norte-americana. Vale enfatizar que essa foi outra dificuldade desta pesquisa: alguns governos não permitem a captação das frequências radiotelefônicas, o que restringiu nosso campo de estudo.

Há duas observações a serem expostas aqui: a primeira é que a intenção deste projeto não é o estudo do inglês americano, mas sim o estudo do inglês compartilhado pela classe de ATCOs e pilotos em situações não rotineiras. Grande parte dos áudios advém de aeroportos internacionais, e há comunicações entre diversas nacionalidades (estamos baseando-nos na nacionalidade da companhia aérea, já que é impossível saber qual é a nacionalidade do piloto que emitiu a mensagem). Com receio de eventualmente analisarmos o inglês americano em detrimento do núcleo aeronáutico internacional, investigamos dois pequenos *corpora*, oriundos do *corpus* de estudo, um com cinco textos produzidos totalmente por nativos americanos, e outro, também com cinco textos, com pilotos mexicanos, portugueses, chineses e franceses. O resultado, exposto na Tabela 3, trouxe-nos confiança de que a linguagem produzida pela comunidade aeronáutica independe da nacionalidade:

Corpus nativo: TTR 17,48%

N	Word	Freq.
1	THE	95
2	YOU	74
3	TO	67
4	RIGHT	48
5	RUNWAY	46
6	WE	45
7	I	40
8	AND	37
9	IT	37
10	UH	37
11	OF	36
12	IS	33
13	ON	31
14	THAT	30
15	AIRPORT	27
16	A	25
17	AT	24
18	IN	23
19	JUST	22
20	RE	20

Corpus não-nativo: TTR 17,07%

N	Word	Freq.
1	TO	62
2	YOU	54
3	THE	50
4	RIGHT	45
5	UH	42
6	AND	38
7	WE	38
8	OKAY	27
9	IS	23
10	LEFT	23
11	RUNWAY	23
12	ARE	21
13	AT	20
14	HEAVY	19
15	I	19
16	A	18
17	RE	18
18	FOR	17
19	ON	17
20	#	16

Tabela 3: Vinte primeiras palavras de dois sub*corpora*: ‘nativos’ e ‘não nativos’.

Na comparação entre as duas listas de palavras, vemos que a riqueza lexical entre elas é muito próxima (um pouco além de 17%) e as primeiras palavras são, em sua maioria, as

mesmas. Destacadas em vermelho estão as palavras que são distintas nas duas colunas: palavras gramaticais, ou relacionadas à aviação, como *airport* [aeroporto] e *heavy* [pesado – utilizado para descrever jatos de grande porte].

A segunda ponderação é a de que, entendendo a nossa pesquisa como representando um núcleo mínimo de proficiência linguística, optamos por selecionar os áudios após 2008, data em que a OACI passou a exigir a certificação de proficiência linguística (o prazo final foi em 2011). Seguindo esse critério, acreditamos que os países signatários cumpriram suas obrigações e somente permitiram a operação internacional de seus pilotos e ATCOs quando licenciados para tanto.

3.1.3 Armazenamento do *corpus*

Os arquivos, todos em formato txt, são nomeados conforme as categorias apontadas na seção 3.1.1, e armazenados em pastas nomeadas com as siglas de cada ocorrência. Para sua identificação, os arquivos foram nomeados com a sigla da categoria de ocorrência (Tabela 1) e o *call sign* da aeronave (a maneira como a aeronave se identifica, que pode ser o nome da empresa com o número do voo ou o registro da matrícula no caso de aviões particulares). BIRD_AA1193, por exemplo, é o arquivo com a transcrição da comunicação em que o voo 1193 da American Airlines sofreu uma colisão com pássaro. Esse arquivo está na pasta BIRD, nome da categoria de tal ocorrência.

As pastas são todas duplicadas – um conjunto armazena os arquivos de áudios, e o outro, os arquivos transcritos em txt, nomeados da mesma forma:

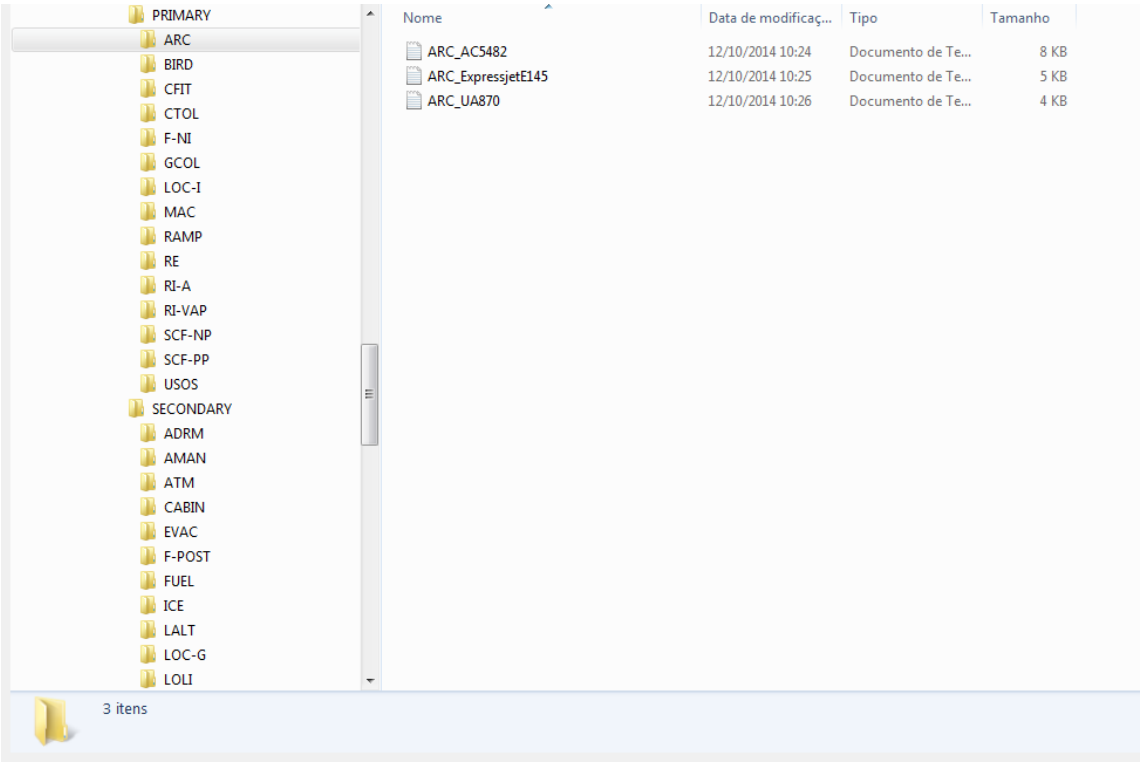


Figura 5: Organização do *corpus* em pastas.

Cada transcrição recebe um cabeçalho com informações que a contextualiza e a vincula com o momento da produção da comunicação, assim como a fonte de onde o áudio foi extraído. Assim, acrescentamos dados sobre onde e quando cada comunicação ocorreu, o tipo de ocorrência (acidente ou incidente), a categoria da ocorrência (conforme a Tabela 1), os *call signs* das aeronaves envolvidas (já inseridas no título) e, sempre que possível, *links* para os *websites* em que constam os relatórios da ocorrência.

O cabeçalho é aberto com a etiqueta <header> e fechado com outra etiqueta </header>, desta vez com a barra invertida. Essas etiquetas são interpretadas pela ferramenta computacional e possibilitam-nos selecionar ou rejeitar o conteúdo entre as etiquetas nas análises. Observamos o modelo de cabeçalho do *corpus* de estudo no Quadro 9:

<header> ‣ início do cabeçalho
    <title> ‣ título da transcrição (nome do arquivo)
    <sourceText> ‣ fonte de onde o áudio foi retirado
    <place> ‣ lugar onde a situação ocorreu
    <type> ‣ acidente ou incidente
    <OCT> ‣ ocorrência segundo a Taxonomia
    <callSign> ‣ matrícula ou número de voo da aeronave envolvida na situação
    <textClarification> ‣ explicação extra sobre o problema
    <date> ‣ data em que o problema aconteceu
    <otherInfo> ‣ qualquer outra informação pertinente
</header> ‣ fechamento do cabeçalho

Quadro 9: Cabeçalho das transcrições do *corpus* de estudo.

Apresentamos, a seguir, uma tabela com os nomes de todos os arquivos, duração do áudio e número de *tokens* de cada um, além da soma total por categoria de ocorrência. Ao término, a contagem final de *tokens* (Tabela 4):

Nome do arquivo	Tempo de áudio	Número de tokens	Número de tokens por categoria
ARC_ExpressJet145	0:03:08	411	
ARC_AC5482	0:02:50	666	
ARC_UA870	0:02:02	333	1410
BIRD _Delta1073	0:07:46	1593	
BIRD_AWE739	0:04:37	747	
BIRD_CACTUS1549	0:03:08	463	2803
CFIT_Asiana	0:07:19	549	
CFIT_UPS	0:05:31	621	
CFIT_Cessna	0:03:38	650	1820
CTOL_ N9409E	0:08:33	1766	
CTOL_PHLengine cowl	0:01:58	383	
CTOL_Pinnacle Airlines	0:12:08	2365	4514
F-NI_AAL963	0:01:36	298	
F-NI_ACA101	0:08:20	586	
F-NI_Acey 5828	0:03:59	918	
F-NI_Delta30	0:04:39	611	2414
GCOL_bos	0:06:40	1276	
GCOL_JFK Air France	0:02:34	334	
GCOL_KAL	0:05:11	830	2440

LOC-I_EZYSCM	0:05:46	889	
LOC-I_Kalitta66	0:04:31	467	
LOC-I_KingAir	0:09:23	2032	
LOC-I_XAUSD	0:02:37	402	3790
MAC_Aerogal	0:03:57	567	
MAC_Cactus	0:02:59	415	
MAC_UA889	0:03:56	576	1558
RAMP_HAL8002	0:01:58	1777	
RAMP_KPHL	0:01:26	180	
RAMP_KBOS	0:01:11	219	2176
RE_AAL4069	0:05:40	1118	
RE_Cactus1702	0:02:15	304	
RE_Delta385	0:03:53	701	2123
RI-A_Cactus1819	0:00:41	118	
RI-A_DOGONRWY	0:02:51	426	
RI-A_KMYFcoyote	0:26:00	67	610
RIVAP_Berlin7450	0:02:13	420	
RIVAP_EGF4125	0:01:30	271	
RIVAP_Mooney3277F	0:02:55	479	1170
SCF-NP_ACA 001	0:06:47	1415	
SCF-NP_ASQ7097	0:10:55	1948	
SCF-NP_BWI8535	0:09:16	1478	4841
SCF-PP_AirPortugal	0:04:73	459	
SCF-PP_SWA1226	0:08:20	1134	
SCF-PP_Swiss885	0:05:02	749	2342
USOS_BA038	0:05:39	372	
USOS_Comanche	0:04:07	623	
USOS_KBWI	0:03:29	533	1528
ADRM_BADDAY	0:10:51	2073	
ADRM_Comair238	0:01:26	173	
ADRM_Etihad101	0:02:17	522	2768
AMAN_AM1640	0:02:37	355	
AMAN_Delta30	0:02:22	440	
AMAN_SW2534	0:02:12	399	1194
ATM_Air China981	0:01:41	289	
ATM_Alitalia lost	0:03:25	436	
ATM_Stuck_mic	0:02:27	365	1090
CABIN_Air France	0:03:37	448	
CABIN_United224	0:10:37	1757	
CABIN_United479	0:03:20	660	2865
EVAC_ASH7262	0:08:49	1412	
EVAC_Delta987	0:07:35	1288	
EVAC_SW4695	0:01:41	290	2990
F-POST_AAL41	0:01:24	297	
F-POST_Challenger	0:06:40	990	
F-POST_Cirrus7850P	0:03:02	491	1778
FUEL_AA180	0:13:33	2525	

FUEL_JAL787	0:03:51	595	
FUEL_Palo Alto	0:03:16	754	3874
ICE_EGF4187	0:03:26	684	
ICE_MU2	0:05:09	717	
ICE_N233R	0:05:41	1103	2504
LALT_Delta534	0:01:53	406	
LALT_tall vessel	0:01:16	198	
LALT_TAM 8082	0:01:46	231	835
LOC-G _cactus271	0:06:44	1121	
LOC-G _JFKsnow	0:04:16	747	
LOC-G_Com568	0:01:25	263	2131
LOLI_Alitalia61G	0:08:35	1225	
LOLI_DHL188	0:04:32	685	
LOLI_Pipeline21	0:04:30	922	2832
SEC_AAL24	0:02:55	445	
SEC_Chautauqua6132	0:03:03	495	
SEC_JetBlue 923	0:01:07	228	1168
TURB_Comair3049	0:06:44	1119	
TURB_USAIR	0:04:23	678	
TURB_wake	0:00:52	179	1976
UIMC_N223CD	0:08:33	782	
UIMC_n360jh	0:13:24	2395	
UIMC_Piper79N	0:02:23	415	3592
WSTRW_AAL002	0:03:57	631	
WSTRW_JetBlue	0:08:03	1418	
WSTRW_United15	0:05:29	1138	3187
TOTAL	**7:26:56**		**70.323**

Tabela 4: Descrição do *corpus*.

Com o *corpus* assim ordenado, cuidadosamente compilado e revisado com a contribuição de detentores de conhecimento técnico-operacional, acreditamos dispor de material linguístico que permite a identificação das estruturas léxico-gramaticais do inglês para aviação. Na próxima seção, serão explicadas as análises sugeridas para essa investigação.

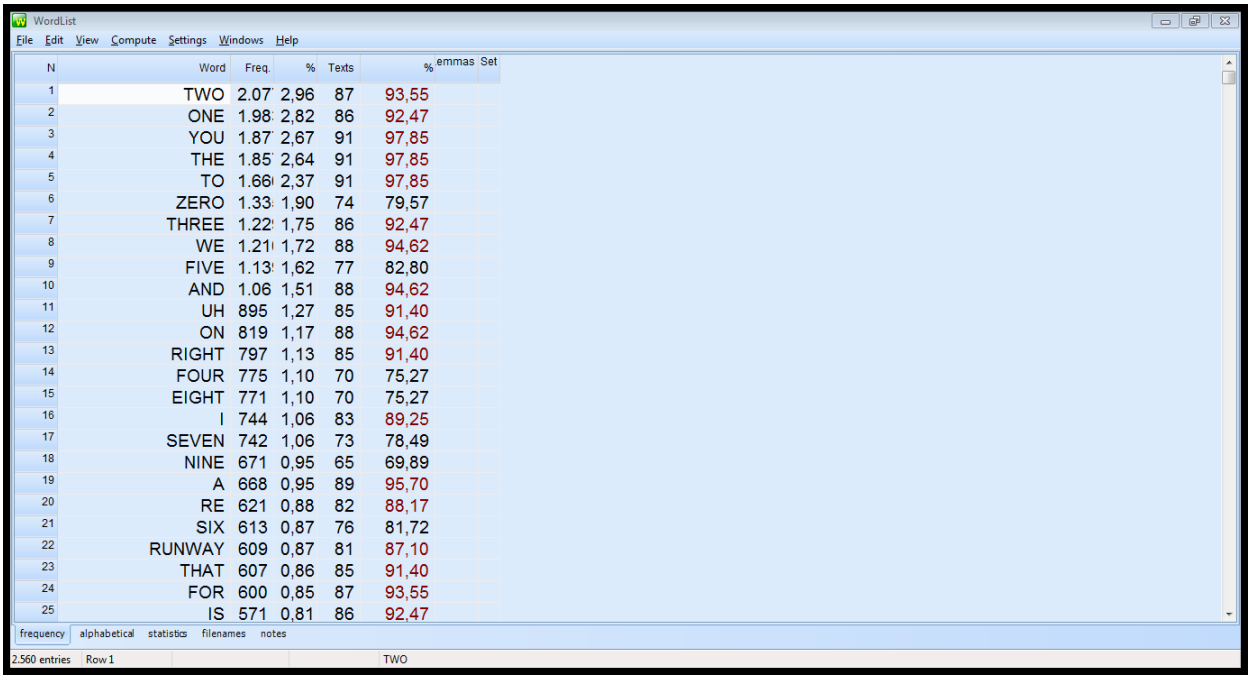
3.2 Procedimentos de análise para a exploração do *corpus*

São várias as possibilidades de ferramentas para a análise que a LC propicia; a escolha é norteada pelo objetivo da pesquisa. Apresentamos, neste capítulo, aquelas que favoreceram os estudos do nosso projeto.

Para a busca de unidades léxico-gramaticais, intuito desta pesquisa, utilizamos o *software WordSmith Tools* (WST) versão 6 (SCOTT, 2012). O WST é largamente difundido entre pesquisadores de LC pela facilidade de uso e pela combinação de recursos exploratórios em um único programa. A ferramenta traz diversas possibilidades de investigação linguística quantitativa, cabendo ao analista a pesquisa qualitativa.

3.2.1 Lista de palavras

A ferramenta *WordList* extrai uma lista de palavras dos textos selecionados – ou do *corpus* todo – considerando a frequência de cada palavra. Assim, a palavra mais recorrente aparecerá no topo da lista. Palavras gramaticais, ou seja, pronomes, artigos, preposições etc. são, normalmente, as primeiras da lista, pois são as mais frequentemente utilizadas. É também a lista de palavras que nos fornece candidatos ao estudo, ou seja, palavras que merecem, pela sua alta frequência, um estudo mais aprofundado. Em nosso caso, buscamos itens lexicais que possam abarcar o campo de análise e a busca pelas unidades léxico-gramaticais do inglês para aviação. Para tanto, investigamos palavras de conteúdo, ou palavras lexicalizadas, que tenham alta frequência no *corpus*, para a análise de seus entornos. Abaixo, a tela do WST com a primeira lista de palavras gerada a partir do *corpus* de comunicações aeronáuticas em situações anormais (Figura 6):



N	Word	Freq.	%	Texts	%	emmas	Set
1	TWO	2.07	2,96	87	93,55		
2	ONE	1.98	2,82	86	92,47		
3	YOU	1.87	2,67	91	97,85		
4	THE	1.85	2,64	91	97,85		
5	TO	1.66	2,37	91	97,85		
6	ZERO	1.33	1,90	74	79,57		
7	THREE	1.22	1,75	86	92,47		
8	WE	1.21	1,72	88	94,62		
9	FIVE	1.13	1,62	77	82,80		
10	AND	1.06	1,51	88	94,62		
11	UH	895	1,27	85	91,40		
12	ON	819	1,17	88	94,62		
13	RIGHT	797	1,13	85	91,40		
14	FOUR	775	1,10	70	75,27		
15	EIGHT	771	1,10	70	75,27		
16	I	744	1,06	83	89,25		
17	SEVEN	742	1,06	73	78,49		
18	NINE	671	0,95	65	69,89		
19	A	668	0,95	89	95,70		
20	RE	621	0,88	82	88,17		
21	SIX	613	0,87	76	81,72		
22	RUNWAY	609	0,87	81	87,10		
23	THAT	607	0,86	85	91,40		
24	FOR	600	0,85	87	93,55		
25	IS	571	0,81	86	92,47		

Figura 6: Tela do WST com a primeira lista de palavras gerada do *corpus* de estudo.

Propomos, aqui, uma abordagem direcionada pelo *corpus*, ou *corpus-driven approach* (TOGNINI-BONELLI, 2001), que parte do pressuposto de que as análises devem ser orientadas pelo que o *corpus* "diz"; é a premissa de *Trust the Text* (Sinclair, 2004). Como o foco de nosso estudo é o inglês comum, e, nas interações há uma mescla de inglês comum e Fraseologia Aeronáutica, mostrou-se importante a análise da Fraseologia Aeronáutica para excluí-la da nossa análise. [39] Esse foi um questionamento que persistiu durante toda a fase de coleta do *corpus*: como remover tais itens. Há elementos que são muito específicos das situações de rotina e, mesmo em uma emergência, são emitidos sempre de forma padronizada: 'vetorações' (instruções dadas por ATCOs quando assumem a navegação da aeronave, dando informações como velocidade, altura e direção),[40] ou autorizações e informações para o pouso, por exemplo. Por esse motivo, elaboramos um arquivo de texto sem formatação, com palavras muito específicas da operação aeronáutica, que as elimina da *WordList* ao ser adicionado à ferramenta; essa lista é denominada *stoplist*. A ela, adicionamos nomes de aeroportos, cidades, unidades de medida (*knots, feet, miles*), números, e o alfabeto fonético,

[39] Convém lembrar que há um estudo feito, também com uso da LC, sobre as situações de rotina (OLIVEIRA, 2007).

[40] No título desta pesquisa há uma alusão ao termo vetoração, pois comparamos esse momento em que o voo está sob a responsabilidade total do ATCO ao conceito *corpus-driven*, ou estudos direcionados pelo *corpus*.

cujo uso é exigido na comunicação radiotelefônica para desambiguar a pronúncia das letras (em vez de a, b, c, d..., diz-se *alpha, bravo, charlie, delta*...).

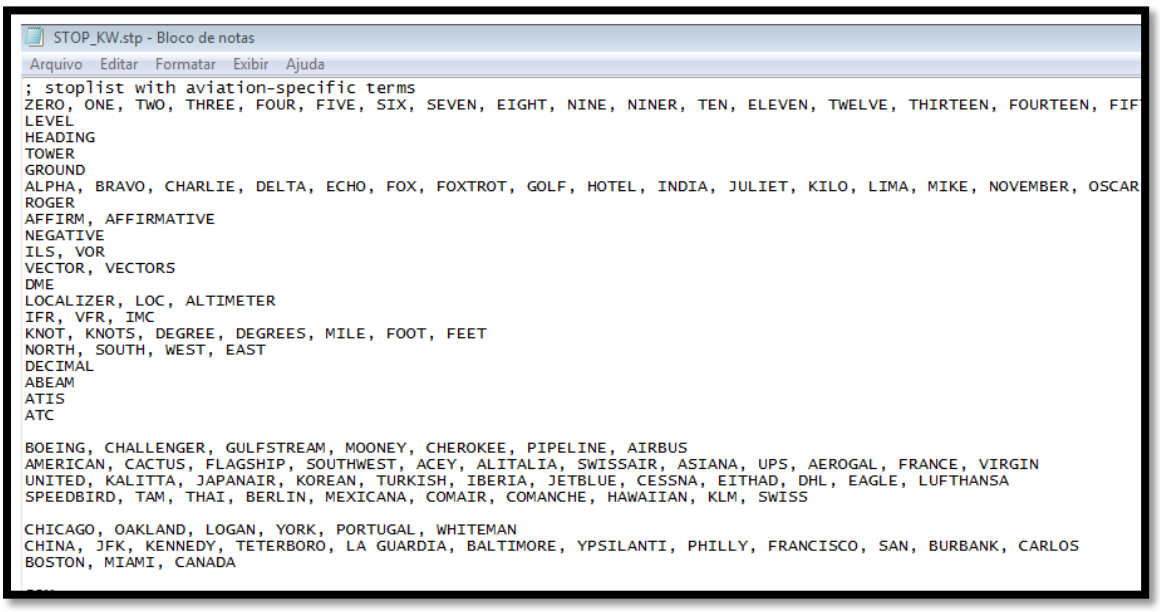
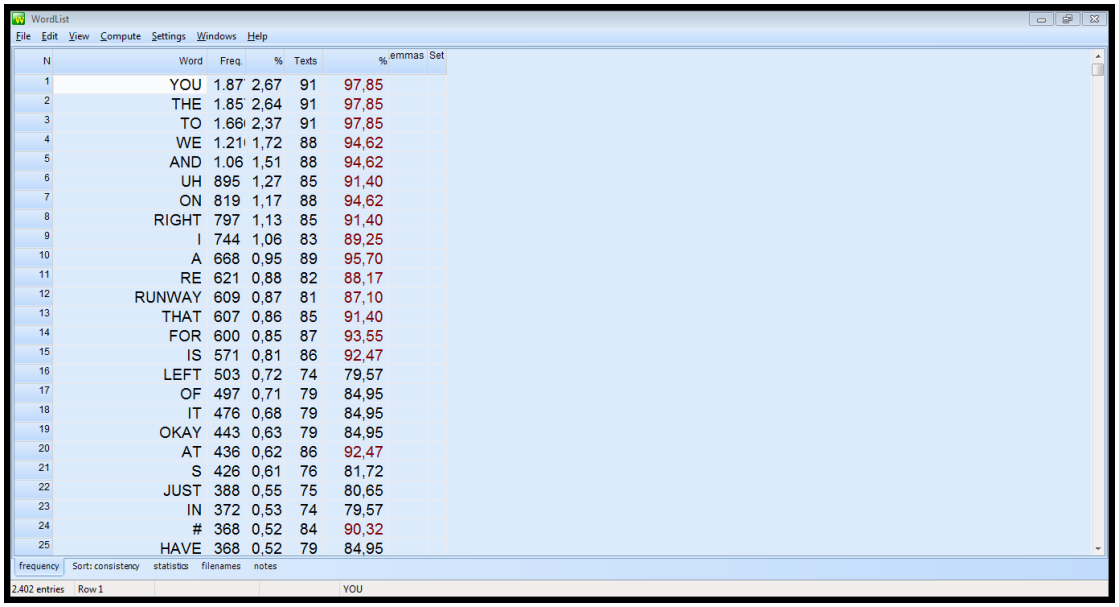


Figura 7: Parte da *stoplist*.

Com a remoção das palavras específicas da Fraseologia Aeronáutica, geramos uma nova lista de palavras, que se aproxima mais do tipo de linguagem que pretendemos estudar. A Figura 8 apresenta as 25 primeiras palavras desse levantamento:



N	Word	Freq.	%	Texts	% emmas Set
1	YOU	1.87	2,67	91	97,85
2	THE	1.85	2,64	91	97,85
3	TO	1.66	2,37	91	97,85
4	WE	1.21	1,72	88	94,62
5	AND	1.06	1,51	88	94,62
6	UH	895	1,27	85	91,40
7	ON	819	1,17	88	94,62
8	RIGHT	797	1,13	85	91,40
9	I	744	1,06	83	89,25
10	A	668	0,95	89	95,70
11	RE	621	0,88	82	88,17
12	RUNWAY	609	0,87	81	87,10
13	THAT	607	0,86	85	91,40
14	FOR	600	0,85	87	93,55
15	IS	571	0,81	86	92,47
16	LEFT	503	0,72	74	79,57
17	OF	497	0,71	79	84,95
18	IT	476	0,68	79	84,95
19	OKAY	443	0,63	79	84,95
20	AT	436	0,62	86	92,47
21	S	426	0,61	76	81,72
22	JUST	388	0,55	75	80,65
23	IN	372	0,53	74	79,57
24	#	368	0,52	84	90,32
25	HAVE	368	0,52	79	84,95

Figura 8: Lista de palavras geradas com a *stoplist*.

Ainda assim, as palavras mais frequentes são verbos comumente utilizados nas instruções e no cotejamento (repetição obrigatória das instruções do ATCO feita pelo piloto), em detrimento de outras ações que, devido a suas formas conjugadas, aparecem com baixa frequência;[41] a solução encontrada foi a da lematização da lista de palavras.

A lematização consiste em agrupar todas as formas de uma mesma palavra. Assim, ao ser lematizado com suas conjugações *tries* (uma ocorrência), *trying* (26 ocorrências) e *tried* (2 ocorrências), o verbo *to try*, que em nossa lista aparece em 206º lugar, com 39 ocorrências, passa a ocupar o 129º lugar, sendo, então, considerado um verbo altamente frequente no *corpus*. A forma lematizada, quando referida neste trabalho, é grafada em versalete: *TO TRY*.

Além disso, a junção de contrações salienta os auxiliares e os verbos de ligação na lista de palavras; portanto, consideramos a lematização de tais itens da seguinte forma:

[41] As conjugações verbais são importantes, pois podem trazer à luz alguns fatos como o uso preferencial socialmente compartilhado de certos verbos ocorrerem somente em determinado tempo verbal (SINCLAIR, 1991); porém, tal estudo pode ser feito em outro momento.

't, nt, t ▸ ´NT
has, having, had, 've, ve ▸ TO HAVE
am, are, is, was, were, being, been, 'm, m ▸ TO BE
'll, ll ▸ WILL

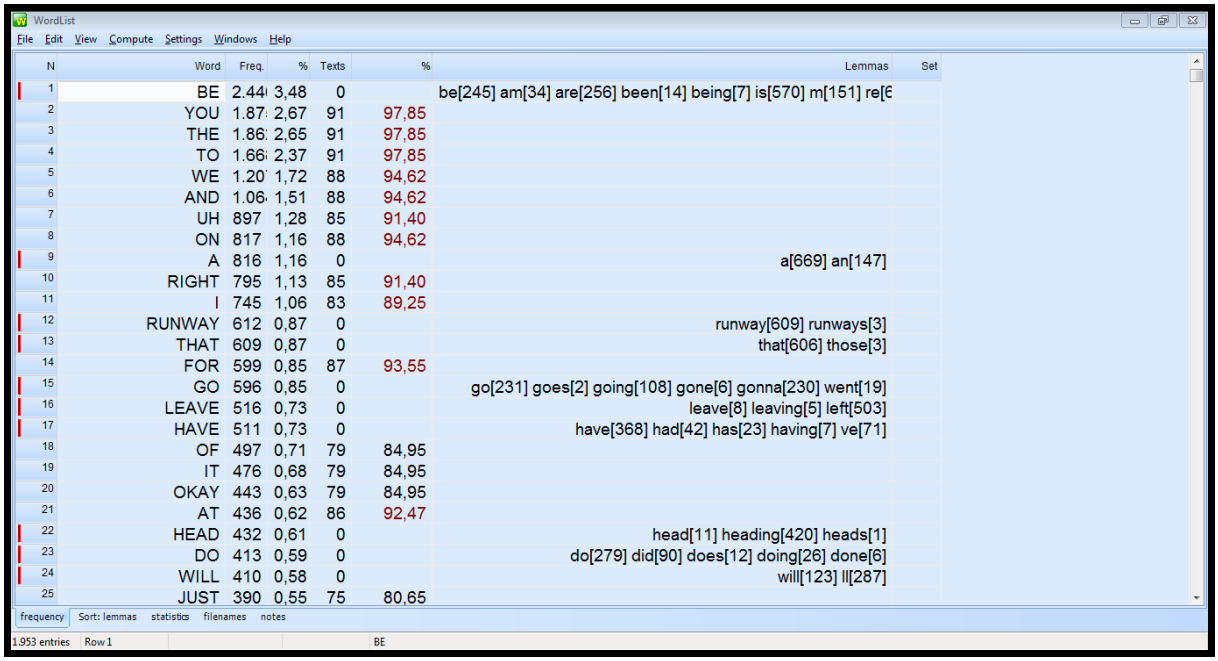
Quadro 10: Contrações presentes nas transcrições e suas lematizações.

Como a contração *´d* pode corresponder tanto ao modal *would* como ao auxiliar *had*, não foi considerada na lematização.

No próprio *website* do WST[42], encontramos uma lista de lemas, preparada por uma colaboradora, Yasumasa Someya, em 1998, que contém 40.659 palavras em 14.762 grupos de lemas (segundo informações do *site*). Para nossa análise, no entanto, promovemos alterações nessa lista: a palavra *ground* [solo], altamente frequente por se referir a momentos em que a aeronave está em solo, aparecia na lista de lemas como passado de *grind* [moer], sem referências em nosso *corpus*; foi, portanto, descartada. Além dessa, acrescentamos as contrações *WANNA* e *GONNA* para representar *want to* e *going to*, respectivamente.

Posteriormente, adicionando as duas listas (lema e *stoplist*) à ferramenta *WordList*, geramos a última lista de palavras (Figura 9), que parecia se adequar ao nosso estudo:

[42] Disponível em: <http://www.lexically.net/wordsmith/support/extras.html>. Acesso em: 07 nov. 2014.



N	Word	Freq.	%	Texts	%	Lemmas	Set
1	BE	2.44(	3,48	0		be[245] am[34] are[256] been[14] being[7] is[570] m[151] re[6	
2	YOU	1.87:	2,67	91	97,85		
3	THE	1.86:	2,65	91	97,85		
4	TO	1.66(	2,37	91	97,85		
5	WE	1.20'	1,72	88	94,62		
6	AND	1.06.	1,51	88	94,62		
7	UH	897	1,28	85	91,40		
8	ON	817	1,16	88	94,62		
9	A	816	1,16	0		a[669] an[147]	
10	RIGHT	795	1,13	85	91,40		
11	I	745	1,06	83	89,25		
12	RUNWAY	612	0,87	0		runway[609] runways[3]	
13	THAT	609	0,87	0		that[606] those[3]	
14	FOR	599	0,85	87	93,55		
15	GO	596	0,85	0		go[231] goes[2] going[108] gone[6] gonna[230] went[19]	
16	LEAVE	516	0,73	0		leave[8] leaving[5] left[503]	
17	HAVE	511	0,73	0		have[368] had[42] has[23] having[7] ve[71]	
18	OF	497	0,71	79	84,95		
19	IT	476	0,68	79	84,95		
20	OKAY	443	0,63	79	84,95		
21	AT	436	0,62	86	92,47		
22	HEAD	432	0,61	0		head[11] heading[420] heads[1]	
23	DO	413	0,59	0		do[279] did[90] does[12] doing[26] done[6]	
24	WILL	410	0,58	0		will[123] ll[287]	
25	JUST	390	0,55	75	80,65		

Figura 9: Tela com a lista de palavras gerada com a *stoplist* e lematizada.

Acima, a tela da lista de palavras já lematizadas, com itens pertencentes à Fraseologia Aeronáutica removidos.

3.2.2 A lista de palavras-chave

Apesar de o tamanho pequeno do *corpus* possibilitar uma busca manual, por tratarmos de uma linguagem extremamente específica, preferimos utilizar a ferramenta *KeyWordList* do WST para que as fórmulas estatísticas nos indicassem quais termos seriam realmente característicos desse domínio.

Para essa análise, a ferramenta necessita de um *corpus* de referência (também denominado *corpus* de controle), ou seja, um *corpus* maior, no nosso caso, falado, que sirva de base para a comparação estatística feita pelo programa. Esse *corpus* deve conter amostras de um tipo de linguagem que possibilite a extração do que é representativo para a análise do

corpus de estudo (BERBER-SARDINHA, 2004). Pela facilidade de acesso[43], escolhemos o *corpus* oral da Universidade de Michigan, o MICASE (*The Michigan Corpus of Academic Spoken English)* para esse propósito.

Utilizamos o MICASE pelas seguintes características: (i) é um *corpus* de língua inglesa; e (ii) é um *corpus* de transcrições de gravações orais, portanto, falado.

O MICASE é um *corpus* compilado em meio acadêmico; contém, contudo, extratos de conversas informais, como reuniões de grupos de alunos, passeios guiados pela universidade, serviços ao universitário etc. É explorado por meio de uma ferramenta de extração disponibilizada pela própria Universidade de Michigan em seu *website*, e permite a seleção de linhas de concordância e de textos por idade, nacionalidade e gênero do falante, entre outros. A Figura 10 apresenta a tela do MICASE, em que visualizamos as possibilidades de extração de dados:

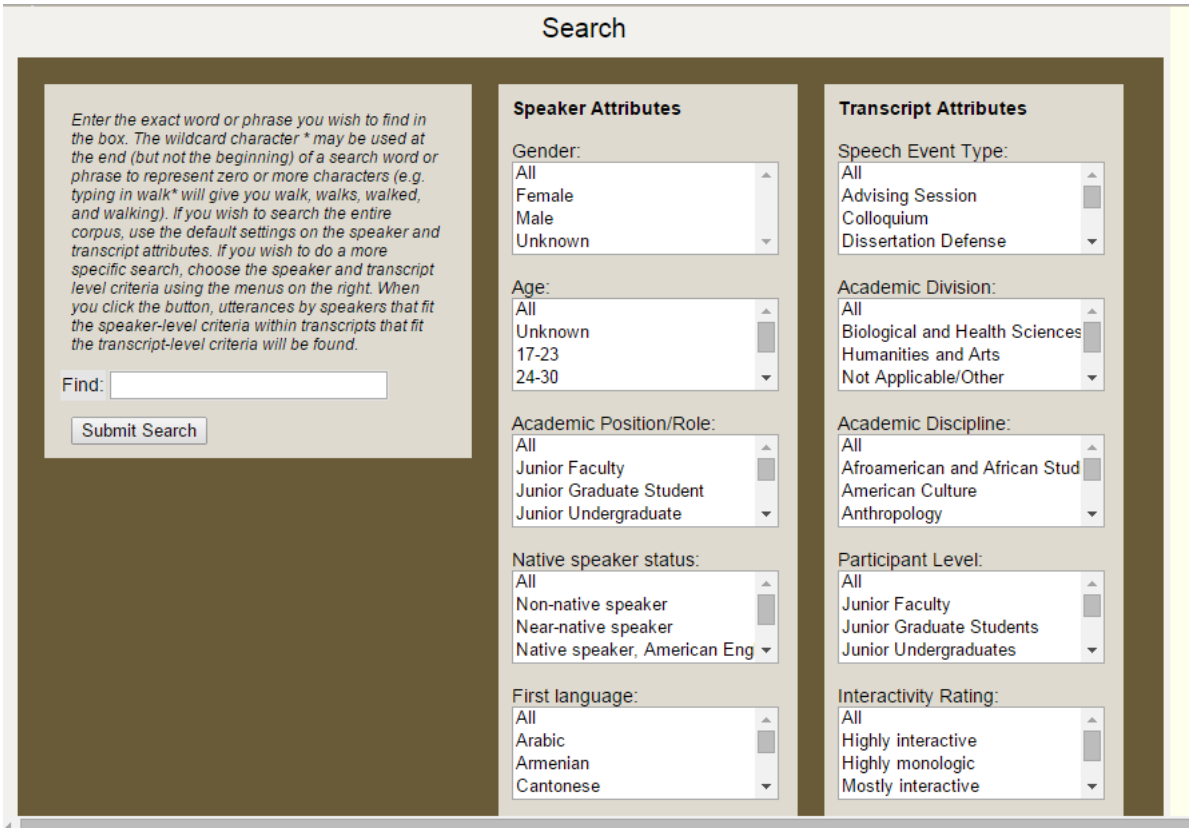


Figura 10: *Website* do MICASE, com a ferramenta de busca e possíveis seleções que possam restringir os campos de análise.

[43] Para acessar a lista de palavras do MICASE e, então, gerar a *KeyWord List* deste estudo, seria necessário adquirir uma licença paga. Consignamos, aqui, nossos agradecimentos à colega Luciana Ginezi, que disponibilizou sua cópia, possibilitando nosso estudo.

Ao gerarmos uma lista de palavras do MICASE e a compararmos, por meio da ferramenta *KeyWordList*, com a lista do *corpus* de estudo – já lematizada e limpa –, que evidencia palavras da Fraseologia Aeronáutica, salientamos aquilo que é realmente típico da linguagem que nos propomos estudar – as palavras-chave –, como apresentado na figura a seguir (Figura 11):

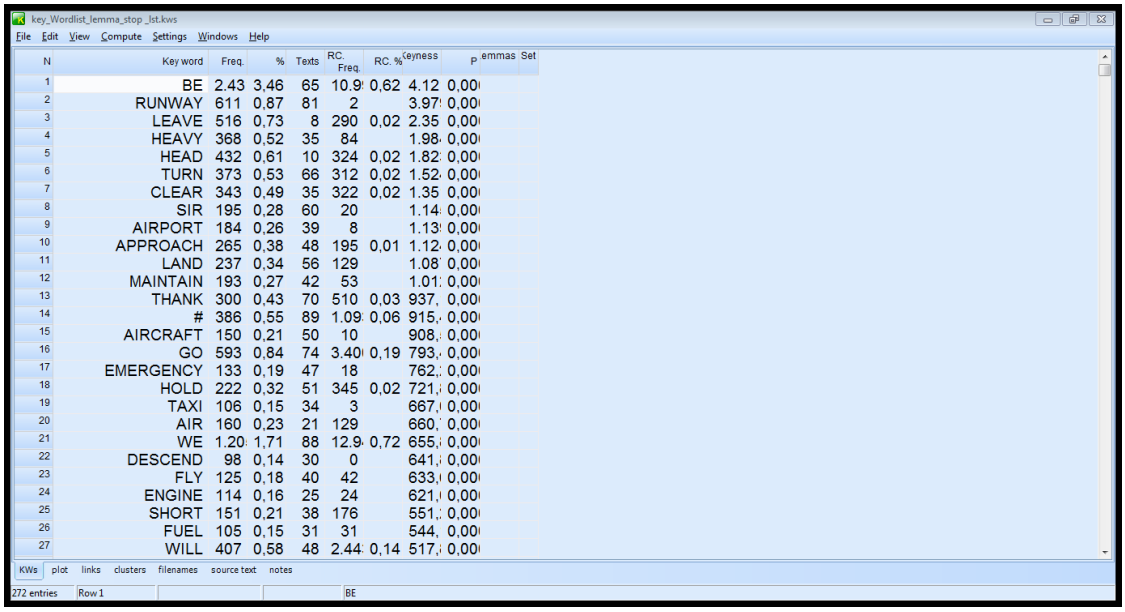


Figura 11: Tela das palavras-chave do *corpus* de inglês aeronáutico.

Procedemos, então, a uma busca manual por substantivos – palavras de conteúdo com teor lexical que possibilitam a análise de seus usos e colocados –, que resultou em uma lista de itens lexicais que guiou as análises seguintes.

3.2.3 As linhas de concordância

A palavra de conteúdo escolhida é, então, inserida em outra ferramenta do WST – o *Concord*, ou concordanciador. Trata-se de uma ferramenta que expõe linhas de concordância, ou seja, linhas extraídas do texto em que a palavra de busca é o centro da investigação

linguística. Por meio das linhas de concordância é possível analisar quais padrões fraseológicos rodeiam a palavra de busca, ou nódulo (*node word*), e buscar a convencionalidade da linguagem (cf. VIANA, 2010). A Figura 12 exemplifica as linhas de concordância com a palavra *runway* centralizada:

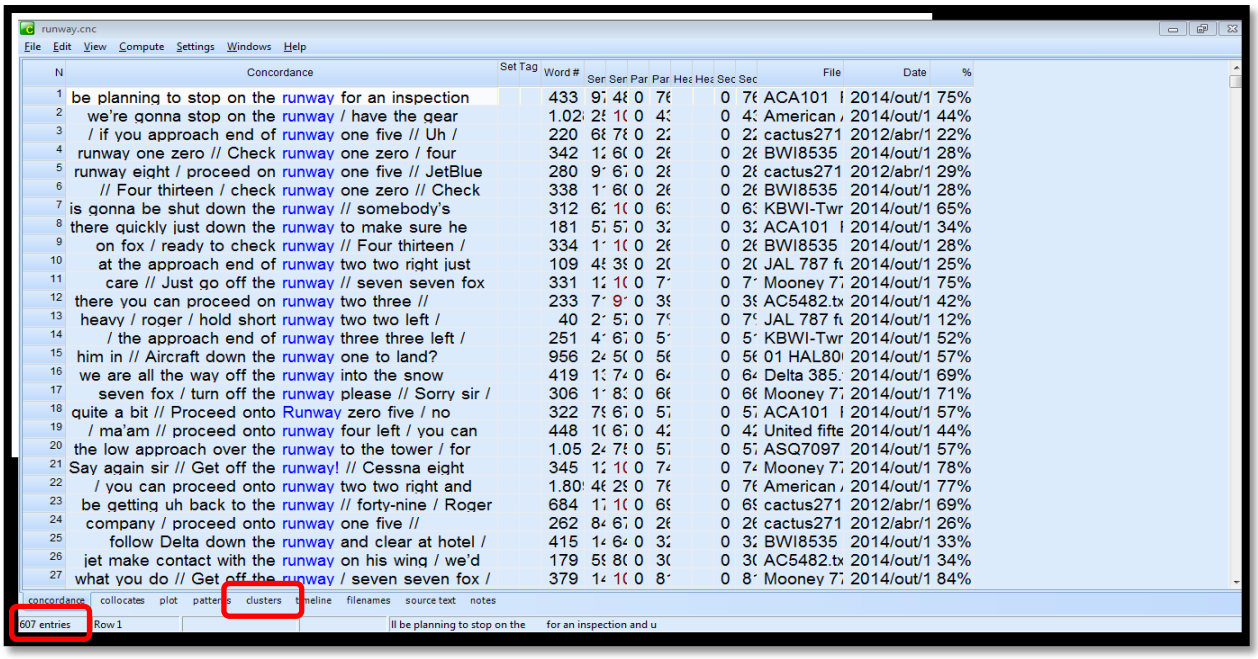


Figura 12: Tela do WST da ferramenta Concord, com linhas de concordância da palavra *runway*.

A alta quantidade de linhas de concordância (no quadro acima, 607) pode dificultar a análise qualitativa, ou seja, a análise manual feita pelo pesquisador. Para facilitar o trabalho, o uso de outros utilitários, a partir do concordanciador, possibilita outras investigações.

3.2.4 Os *clusters*

No concordanciador, há uma aba denominada *clusters* (a quinta da esquerda para a direita, em destaque), ou agrupamentos lexicais, que são "[...] lista(s) de sequências fixas de palavras recorrentes na concordância" (BERBER-SARDINHA, 2004, p. 111). Esses *clusters* são gerados por bases estatísticas, e determinamos como ponto de corte a frequência mínima

de cinco, ou seja, esses agrupamentos precisam aparecer mais de cinco vezes no *corpus* para serem significativos neste estudo.

Os *clusters* facilitam a investigação linguística na medida em que já salientam estruturas recorrentes, retiradas estatisticamente do *corpus* tendo a palavra pesquisada como núcleo. Porém, o simples levantamento dessas combinações pode resultar em muito ruído, pois a base estatística é aleatória – somente agrupa as palavras que ocorrem próximas uma(s) à(s) outra(s), sem critérios qualitativos. Portanto, a limpeza manual de *clusters* sem unidade de significado é necessária, além da eliminação de *clusters* gerados que não farão parte da análise, como números de voo ou elocuções da Fraseologia Aeronáutica. Demonstramos esse processo a seguir, nas figuras abaixo, que, quando comparadas entre si, revelam a limpeza manual, com a preservação dos *clusters* com unidade de sentido e característicos do inglês de uso comum (Figura 13):

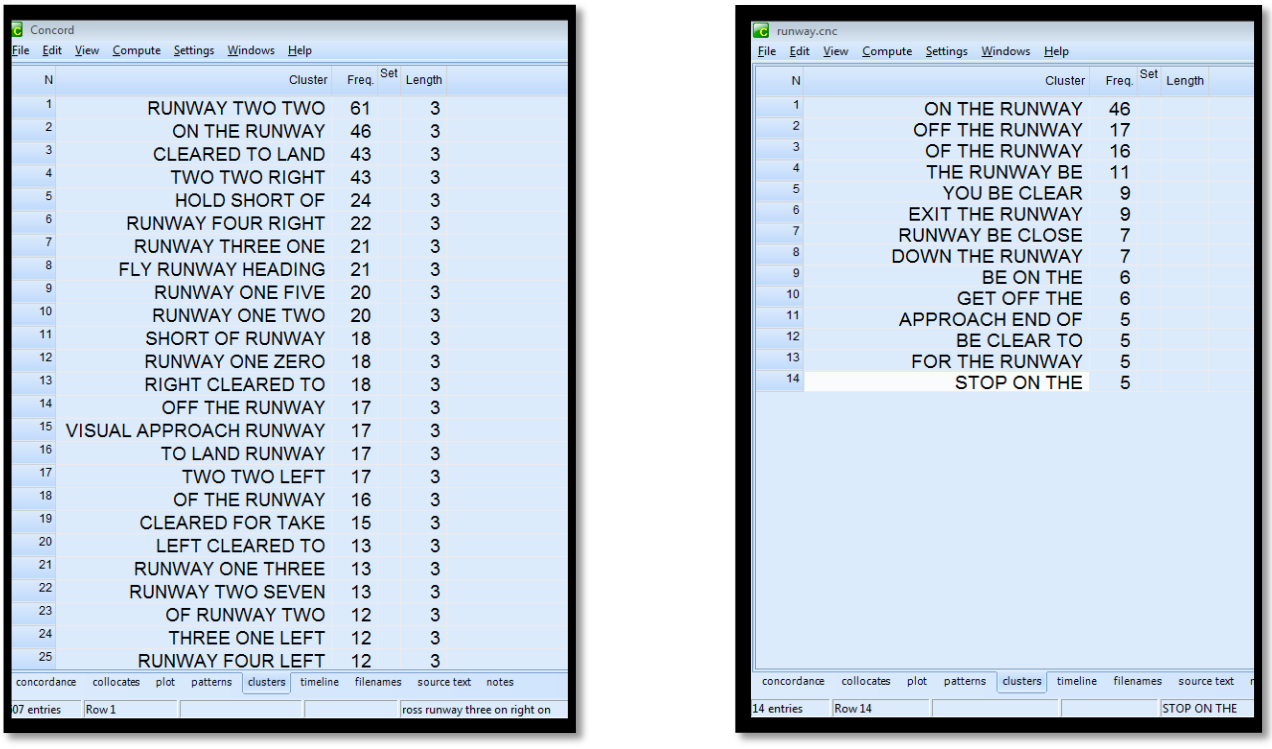


Figura 13: Telas de *clusters* de *runway* do WST, comparando dois processos.

Na segunda lista gerada, encontramos o verbo *TO BE* na forma infinitiva; isso acontece devido à lematização anterior. Consideramos o verbo conjugado para a preservação do

cluster: *the runway be*, por exemplo, representa *the runway is*, e foi mantido para a pesquisa. Com essa lista, voltamos ao *Concord* com a expressão completa – o *cluster* – como insumo da busca, gerando, assim, outras linhas de concordância para a análise qualitativa. A Figura 14 mostra o *cluster on the runway* no centro das linhas de concordância:

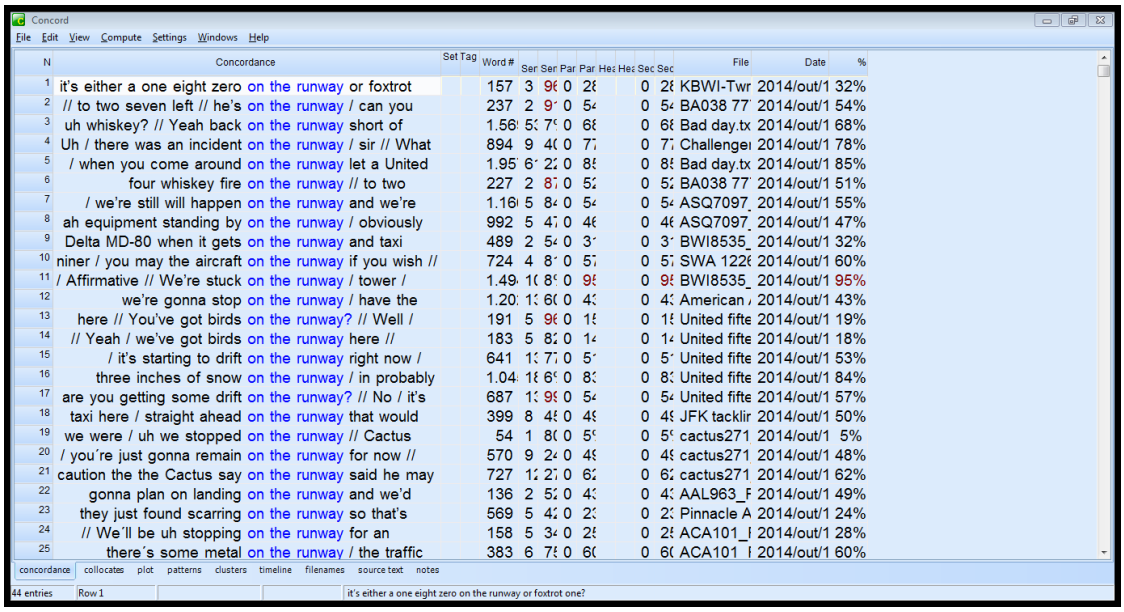
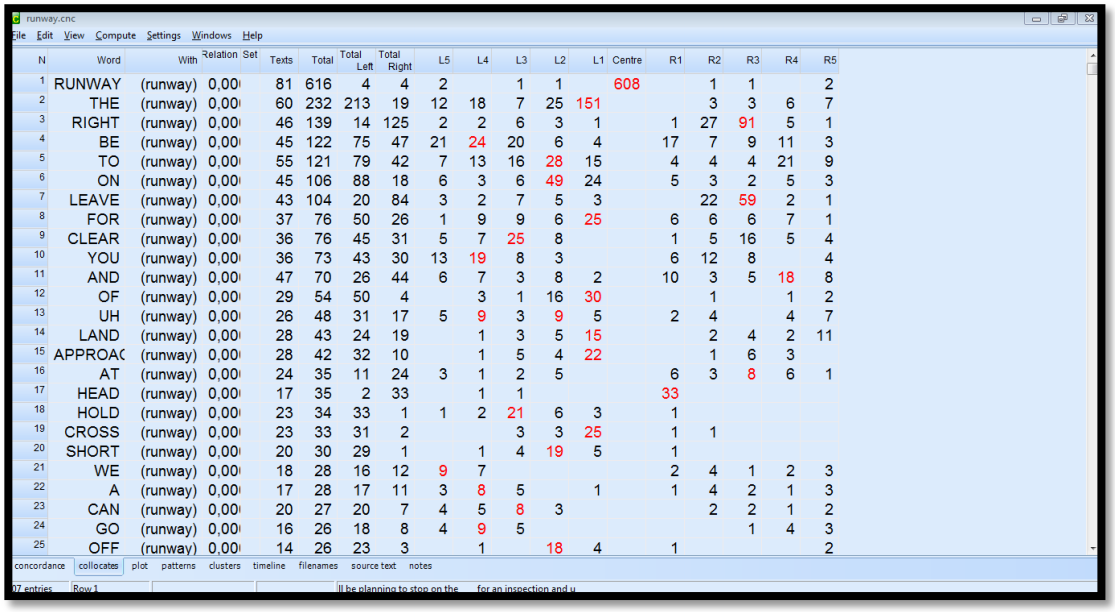


Figura 14: Tela de linhas de concordância com o *cluster 'on the runway'* no núcleo.

3.2.5 Os colocados

Os *clusters* podem ser investigados de duas maneiras: uma com o *cluster* completo, na forma fixa em que foi gerado, e outra em que podemos investigar outras combinações possíveis, pois o *cluster* pode ser variável. Por exemplo, o *cluster* '*on the runway'* pode, também, ser '*on that runway'*, '*on this runway'* ou alguma outra expressão semelhante. Para isso, a ferramenta interpreta buscas como *on * runway*, em que (*) é uma *wildcard*, ou um curinga, como expressões que trazem outras palavras entre *on* e *runway*. Outra possibilidade para esse mesmo tipo de busca é a ferramenta *Collocates*, que traz os colocados mais frequentes da palavra investigada, além da quantidade de instâncias em seu entorno, representadas por L quando a palavra se refere a um colocado do lado esquerdo do nódulo

(*node word*) ou R para os colocados à direita. Sua posição também é destacada, por números; assim, L1 é o primeiro colocado à esquerda, L2 o segundo e assim por diante.



N	Word	With	Relation	Set	Texts	Total	Total Left	Total Right	L5	L4	L3	L2	L1	Centre	R1	R2	R3	R4	R5
1	RUNWAY	(runway)	0,00		81	616	4	4	2		1	1		608		1	1		2
2	THE	(runway)	0,00		60	232	213	19	12	18	7	25	151			3	3	6	7
3	RIGHT	(runway)	0,00		46	139	14	125	2	2	6	3	1		1	27	91	5	1
4	BE	(runway)	0,00		45	122	75	47	21	24	20	6	4		17	7	9	11	3
5	TO	(runway)	0,00		55	121	79	42	7	13	16	28	15		4	4	4	21	9
6	ON	(runway)	0,00		45	106	88	18	6	3	6	49	24		5	3	2	5	3
7	LEAVE	(runway)	0,00		43	104	20	84	3	2	7	5	3			22	59	2	1
8	FOR	(runway)	0,00		37	76	50	26	1	9	9	6	25		6	6	6	7	1
9	CLEAR	(runway)	0,00		36	76	45	31	5	7	25	8			1	5	16	5	4
10	YOU	(runway)	0,00		36	73	43	30	13	19	8	3			6	12	8		4
11	AND	(runway)	0,00		47	70	26	44	6	7	3	8	2		10	3	5	18	8
12	OF	(runway)	0,00		29	54	50	4		3	1	16	30			1		1	2
13	UH	(runway)	0,00		26	48	31	17	5	9	3	9	5		2	4		4	7
14	LAND	(runway)	0,00		28	43	24	19		1	3	5	15			2	4	2	11
15	APPROAC	(runway)	0,00		28	42	32	10		1	5	4	22			1	6	3	
16	AT	(runway)	0,00		24	35	11	24	3	1	2	5			6	3	8	6	1
17	HEAD	(runway)	0,00		17	35	2	33		1	1				33				
18	HOLD	(runway)	0,00		23	34	33	1	1	2	21	6	3		1				
19	CROSS	(runway)	0,00		23	33	31	2			3	3	25		1	1			
20	SHORT	(runway)	0,00		20	30	29	1		1	4	19	5		1				
21	WE	(runway)	0,00		18	28	16	12	9	7					2	4	1	2	3
22	A	(runway)	0,00		17	28	17	11	3	8	5		1		1	4	2	1	3
23	CAN	(runway)	0,00		20	27	20	7	4	5	8	3				2	2	1	2
24	GO	(runway)	0,00		16	26	18	8	4	9	5						1	4	3
25	OFF	(runway)	0,00		14	26	23	3		1		18	4		1				2

Figura 15: Tela de colocados da palavra *runway*.

Na Figura 15, podemos visualizar, à esquerda da tela, os colocados mais comuns de *runway* e, ao centro da tela, a quantidade de instâncias e suas posições, sendo as mais frequentes destacadas em vermelho. Assim, *the* é o colocado mais frequente de *runway*, principalmente na primeira posição à esquerda (L1), em que ocorre 151 vezes. Ao clicarmos nesse último número, a ferramenta seleciona todas as linhas de concordância em que *the* aparece como primeiro colocado à esquerda de *runway*.

As ferramentas descritas nesta seção possibilitam uma extração automática dos dados. Porém, é o olhar do pesquisador que irá determinar a qualidade do que é investigado. Para a busca de fraseologismos, ou estruturas léxico-gramaticais, tendo aqui o item lexical – neste caso, o substantivo – como núcleo temático, adotamos a sugestão de Hunston (2002) como método de observação de significado e padrões linguísticos.

Por meio das linhas de concordância, Hunston (2002, p. 46-56) propõe a análise do 'cotexto'[44], ou o entorno fraseológico da palavra. Dessa forma, é possível analisar, primeiramente, o significado da palavra, pois está ligado ao contexto em que se insere. Por meio dos colocados – inclusive gramaticais (preposições, conjunções, pronomes) – é possível gerar grupos de linhas de concordância e verificar o uso da palavra e seu entorno para, então, definir as associações por significado. Para ilustrar, em estudo com a palavra *condem*, Hunston verifica que (2002, p. 47):

- com significado de criticar, é sempre um verbo transitivo direto ou um verbo transitivo direto acompanhado de *as* (*...condemned the incident as the dangerous...)*;

- com significado de sentenciar, ou de fazer com que algo ruim aconteça, é sempre acompanhado do infinitivo com *to (... people who a few days before had condemned him to death...)*.

Para simplificar o caminho a ser percorrido, O´Keefe, McCarthy e Carter (2007, p. 14)[45] sugerem, para a busca do que denominam "'perfis léxico-gramaticais' de uma palavra e seus contextos de uso",[46] as seguintes perguntas:

1. Colocados: que palavra(s) ocorre(m) com mais frequência com o nódulo?
2. *Clusters*: a palavra compõe *clusters* recorrentes?
3. Restrições sintáticas: existem padrões sintáticos que restringem a palavra (preposições, posições na oração, tempos verbais)?
4. Restrições semânticas: há alguma (por exemplo, nunca é utilizada com intensificador, por exemplo)?
5. Prosódia semântica: em quais contextos (positivo ou negativo, por exemplo) a palavra ocorre?

No capítulo 4, apresentamos nossa análise e o levantamento dos padrões léxico-gramaticais do inglês para aviação.

---

[44] Viana (2010, p. 71) explica a palavra 'cotexto', utilizada pela LC, como um termo empregado "para se referir ao ambiente linguístico, ou seja, aos itens que são utilizados à esquerda e à direita de uma dada palavra de busca."

[45] Hoey (2003) já sugeria perguntas similares: *(i) What does the word mean?; (ii) What words does it associate with?; (iii) What meanings does it associate with?; (iv) What grammatical functions does it associate with?; (v) What positions in the text does the word favour?*

[46] *'Lexico-grammatical profiles' of a word and its contexts of use.*

CAPÍTULO 4 – ANÁLISE

> *As increased air safety is the motivating factor beneath any initiative, including the establishment of provisions for language proficiency in civil aviation, it is important that language-training programs address appropriate needs within the domain of aviation operational communications* (MATHEWS, 2004)[47].

Para o levantamento das unidades léxico-gramaticais do inglês para aviação, compilamos um *corpus* de comunicações entre ATCOs e pilotos em situações anormais, descrito no Capítulo 3. Demonstramos, neste capítulo, o estudo dos fraseologismos mais recorrente nesse *corpus*, por meio das ferramentas computacionais da LC. Essa análise está dividida em duas partes: a primeira expõe, por meio de cinco itens lexicais, uma investigação dos fraseologismos típicos de cada um. A segunda compara tais buscas com o Glossário de Estruturas Gramaticais, explicitado no Capítulo 1 desta dissertação.

4.1 As estruturas léxico-gramaticais do inglês para aviação

Buscamos, nesta seção, responder à primeira pergunta desta pesquisa:

> Quais são as estruturas léxico-gramaticais mais utilizadas por pilotos e controladores de tráfego aéreo em situações anormais, não previstas na Fraseologia Aeronáutica?

Para essa investigação, selecionamos alguns itens lexicais que pudessem encabeçar as análises e identificar quais fraseologias são normalmente formadas com esses itens. Definimos que tais itens deveriam apresentar alta frequência no *corpus*, e ser relevantes ao inglês para aviação. Para esse levantamento, foi gerada uma lista de palavras, utilizando uma *stoplist*, elaborada no decorrer da compilação do *corpus* para a eliminação das palavras típicas

[47] "Como a melhoria na segurança de voo é o fator motivador sob qualquer iniciativa, incluindo o estabelecimento da implantação da proficiência linguística na aviação civil, é importante que programas de treinamento de línguas atendam apropriadamente às necessidades do domínio das comunicações operacionais aeronáuticas".

da Fraseologia Aeronáutica, e lematizada, para aumentar a frequência de verbos lexicais que não estavam sendo evidenciados nos primeiros estudos por nós realizados.

O motivo pelo qual esse filtro foi necessário se deve à alta frequência de palavras funcionais, ou gramaticais, com pouco conteúdo lexical, extraídas pela lista de palavras (O´KEEFE; MCCARTY; CARTER, 2007). Além dessas, algumas palavras lexicais, porém específicas das operações rotineiras da aviação, também ocuparam os primeiros lugares na lista. Pode-se perceber, pelo Gráfico 1, como a lista foi exaurida no processo de limpeza, para facilitar a seleção de itens lexicais:

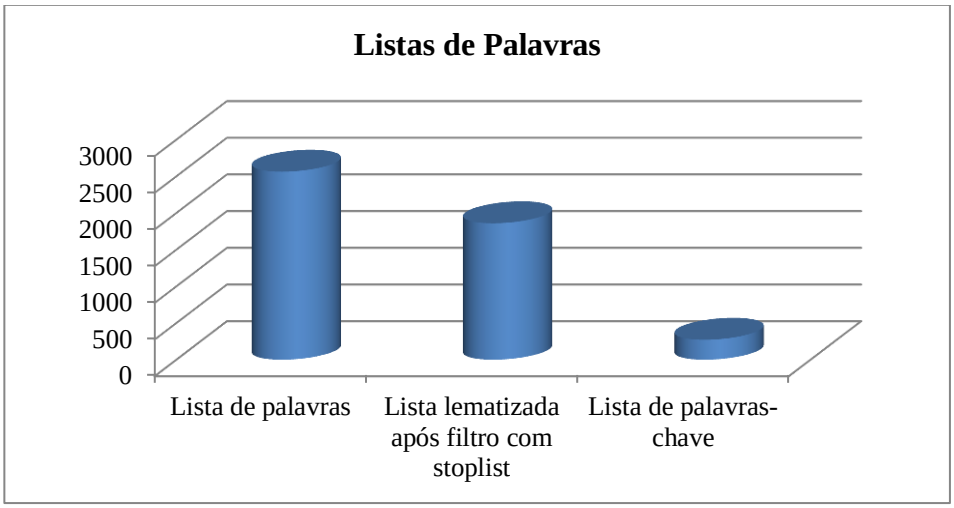


Gráfico 1: Comparação entre as três listas geradas, cada uma com seu filtro.

A primeira lista contém 2.563 palavras; a segunda, gerada com a *stoplist* e lematizada, levantou 1.860 palavras; a última, resultante da comparação com o *corpus* de referência MICASE (vide Capítulo 3), extraiu 272 entradas, o que simplificou a averiguação dos dados. Para ilustrar o processo utilizado, apresentamos (Tabela 5), primeiramente, a lista das cinquenta primeiras palavras resultantes do primeiro refinamento, com a *stoplist* e a lematização:

N	Word	Freq.	%	Texts	%	Lemmas
1	BE	2446	3,48	0		be[245] am[34] are[256] been[14] being[7] is[570] m[151] re[623] s[426] was[88] were[32]
2	YOU	1875	2,67	91	97,85	
3	THE	1862	2,65	91	97,85	
4	TO	1668	2,37	91	97,85	
5	WE	1207	1,72	88	94,62	
6	AND	1064	1,51	88	94,62	
7	UH	897	1,28	85	91,40	
8	ON	817	1,16	88	94,62	
9	A	816	1,16	0		a[669] an[147]
10	RIGHT	795	1,13	85	91,40	
11	I	745	1,06	83	89,25	
12	RUNWAY	612	0,87	0		runway[609] runways[3]
13	THAT	609	0,87	0		that[606] those[3]
14	FOR	599	0,85	87	93,55	
15	GO	596	0,85	0		go[231] goes[2] going[108] gone[6] gonna[230] went[19]
16	LEAVE	516	0,73	0		leave[8] leaving[5] left[503]
17	HAVE	511	0,73	0		have[368] had[42] has[23] having[7] ve[71]
18	OF	497	0,71	79	84,95	
19	IT	476	0,68	79	84,95	
20	OKAY	443	0,63	79	84,95	
21	AT	436	0,62	86	92,47	
22	HEAD	432	0,61	0		head[11] heading[420] heads[1]
23	DO	413	0,59	0		do[279] did[90] does[12] doing[26] done[6]
24	WILL	410	0,58	0		will[123] ll[287]
25	JUST	390	0,55	75	80,65	
26	TURN	373	0,53	0		turn[338] turned[7] turning[27] turns[1]
27	IN	372	0,53	74	79,57	
28	#	371	0,53	84	90,32	
29	HEAVY	368	0,52	35	37,63	
30	GET	352	0,50	0		get[192] gets[3] getting[22] got[133] gotten[2]
31	CLEAR	343	0,49	0		clear[51] cleared[284] clearing[7] clears[1]
32	CAN	334	0,47	78	83,87	
33	NOW	302	0,43	76	81,72	
34	THANK	300	0,43	0		thank[196] thanks[104]
35	THERE	285	0,41	69	74,19	
36	APPROACH	264	0,38	0		approach[253] approaches[1] approaching[10]
37	YOUR	260	0,37	71	76,34	
38	IF	258	0,37	67	72,04	
39	LAND	236	0,34	0		land[170] landed[7] landing[55] lands[4]
40	HOLD	222	0,32	0		hold[187] held[2] holding[33]
41	POINT	213	0,30	56	60,22	
42	UNITE	202	0,29	0		unite[0] united[202]
43	NEED	200	0,28	0		need[197] needed[2] needs[1]
44	DOWN	199	0,28	0		down[198] downed[1]
45	SIR	197	0,28	60	64,52	
46	WITH	194	0,28	60	64,52	
47	MAINTAIN	193	0,27	0		maintain[187] maintaining[6]
48	AIRPORT	184	0,26	0		airport[183] airports[1]
49	OUT	179	0,25	59	63,44	

50	'NT	178	0,25	0	'nt[0] t[178]

Tabela 5: As 50 palavras mais frequentes do *corpus*, após *stoplist* e lematização.

Embora a lista da Tabela 5 evidencie muitas palavras típicas do inglês para aviação, ainda apresenta muitos pronomes, determinantes, verbos suporte etc. Como, neste estudo, pretendemos identificar itens lexicais típicos do inglês de uso comum que nos permitam realizar as análises de seu entorno, geramos uma nova lista de palavras-chave utilizando o MICASE como *corpus* de referência. A lista gerada evidenciou muitas ações relacionadas à Fraseologia Aeronáutica (*leave, head, turn, clear*) e, em busca de palavras que remetam às situações anormais – e ao inglês comum –, julgamos que o levantamento de substantivos serviria melhor à análise, pois estariam também relacionados com as categorias de ocorrências de acidentes e incidentes. A seguir, a lista com as primeiras palavras-chave extraídas da comparação entre o *corpus* de estudo e o MICASE:

N	Key word	Freq.	Keyness
1	BE	2431	4121,26
2	RUNWAY	611	3979,45
3	LEAVE	516	2351,71
4	HEAVY	368	1984,07
5	HEAD	432	1823,76
6	TURN	373	1524,24
7	CLEAR	343	1351,22
8	SIR	195	1145,92
9	AIRPORT	184	1139,47
10	APPROACH	265	1124,28
11	LAND	237	1087,62
12	MAINTAIN	193	1012,05
13	THANK	300	937,12
14	#	386	915,48
15	AIRCRAFT	150	908,52
16	GO	593	793,41
17	EMERGENCY	133	762,23
18	HOLD	222	721,83
19	TAXI	106	667,04
20	AIR	160	660,72
21	WE	1205	655,86
22	DESCEND	98	641,87
23	FLY	125	633,63
24	ENGINE	114	621,02

25	SHORT	151	551,27
26	FUEL	105	544,12
27	WILL	407	517,88
28	ADVISE	83	505,61
29	TRAFFIC	97	487,93
30	DEPARTURE	86	487,27
31	CONTACT	159	482,61
32	RIGHT	792	481,94
33	PROCEED	81	464,00
34	CLIMB	82	454,67
35	AH	138	449,90
36	REQUEST	72	424,80
37	WIND	91	416,50
38	VEHICLE	72	386,06
39	SIGHT	64	343,52
40	FLIGHT	70	338,59
41	RESCUE	59	316,12
42	ON	816	311,50
43	SPEED	78	305,95
44	YES	234	301,52
45	FOLLOW	96	296,45
46	STAND	79	288,43
47	POSITION	94	282,59
48	TAXIWAY	43	281,61
49	GEAR	51	280,35
50	VISUAL	72	269,86

Tabela 6: Lista das 50 palavras-chave mais frequentes do *corpus* de estudo.

Partindo dessa última lista, que apresenta as palavras tipicamente utilizadas no contexto aeronáutico, foi possível buscar manualmente os itens lexicais que direcionam as análises. Definimos, assim, os substantivos, agrupando alguns pertencentes ao mesmo contexto semântico, como por exemplo, *vehicles* [veículos], *rescue* [resgate], *truck* [caminhão], *equipment* [equipamento], *assistance* [assistência] e *tug* [trator], dentro da categoria de veículos de assistência; *pilot* [piloto] e *captain* [comandante]; e *gear* [trem de pouso], *wing* [asa], *tire* [pneu], *nose* [nariz], *tail* [cauda], *wheel* [roda] *e APU* [*Auxiliary Power Unit* – unidade auxiliar de força], referentes a partes da aeronave. Vale ressaltar que optamos por individualizar o termo *engine* na lista, não o agrupando com partes da aeronave, devido a sua alta frequência e à separação dessa categoria no próprio documento da OACI, que identifica problemas com motores na categoria SCF-PP – ***S****ystem/****c****omponent* ***f****ailure or malfunction: PowerPlant Failure* e problemas em outros sistemas – e partes da aeronave – como SCF-NP – ***S****ystem/****c****omponent* ***f****ailure or malfunction: Non-PowerPlant* (Tabela 7):

Item lexical	Frequência
RUNWAY	611
AIRPORT	184
AIRCRAFT	150
EMERGENCY	133
ENGINE	114
FUEL	105
VEHICLE + rescue, truck, equipment, assistance, tug	252
FIRE	64
GEAR + wing, tire, nose, tail, wheel, APU	146
SMOKE	43
PILOT + captain	48

Tabela 7: Lista das palavras candidatas a campos lexicais.

Esses termos dão centralidade às análises que apresentamos a seguir. Para este estudo, investigamos os fraseologismos com as palavras *runway* [pista], *aircraft* [aeronave], *emergency* [emergência], *engine* [motor] e *fuel* [combustível], por terem apresentado resultados mais significativos. Apesar de a palavra *airport* [aeroporto] ter sido a segunda palavra de conteúdo, evidenciou muita linguagem relacionada à Fraseologia Aeronáutica e, portanto, foi descartada nesta pesquisa.

4.1.1 *Runway*

De acordo com um estudo[48] feito por uma das maiores fabricantes de aeronaves do mundo, a Boeing, a maioria dos acidentes e incidentes aéreos ocorre próximo à pista, na

[48] Disponível em: <http://www.boeing.com/news/techissues/pdf/statsum.pdf>. Acesso em: 09 nov. 2014.

decolagem e na aterrisagem. Na figura 16, em um diagrama elaborado pelo *site* IG[49], podemos visualizar a porcentagem de acidentes ocorridos entre 1999 e 2008, por fase de voo:

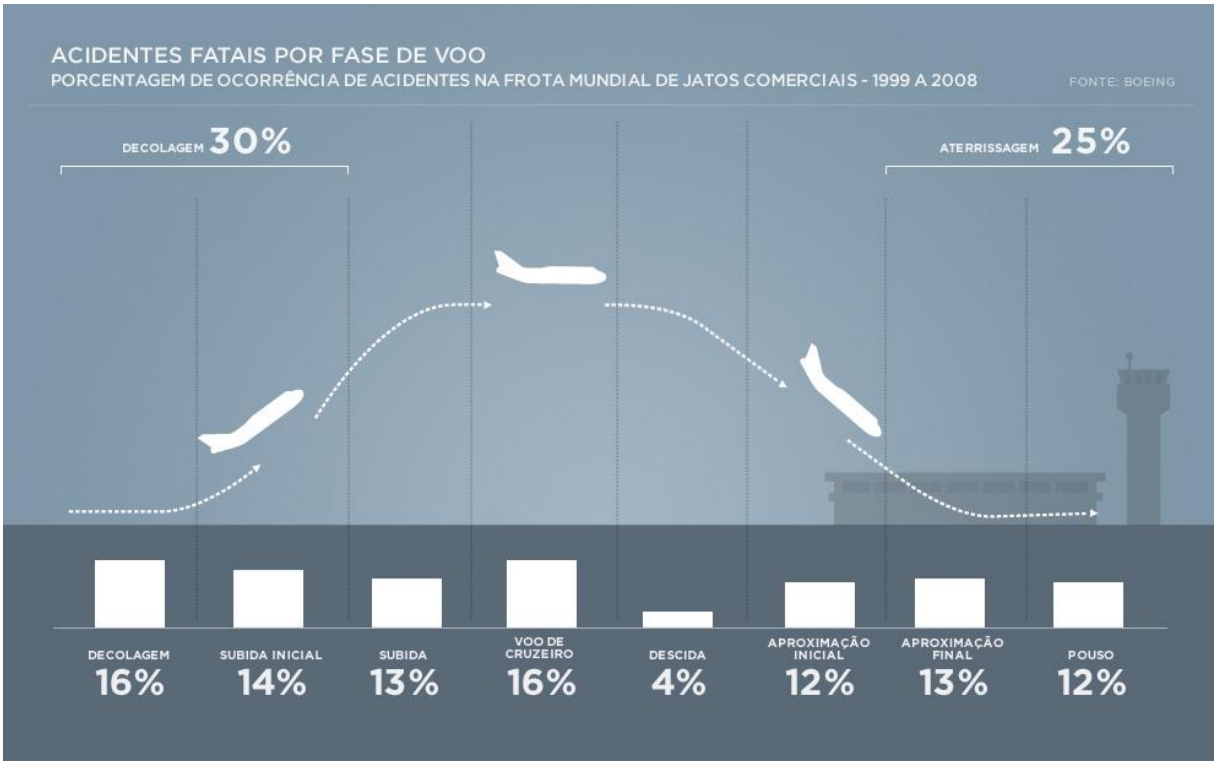


Figura 16: Porcentagem de acidentes aéreos por fase de voo, de 1999 a 2008.

A título de ilustração, esboçaremos alguns problemas que podem ocorrer na – ou próximo à – pista (*runway*): a decolagem pode ser interrompida quando a aeronave já está em velocidade na pista por motivo de falhas nos sistemas, ou pelo fato de algum veículo ou aeronave ter entrado na pista sem autorização (na Taxonomia de Ocorrências da OACI, tal problema corresponde à Categoria RI – incursão de pista). Outras ocorrências da lista de categorias da OACI também se referem ao uso da pista: ARC, ou contato anormal com a pista; CTOL, colisão com obstáculo durante a corrida na pista; RE, excursão de pista (saída involuntária da pista, sem controle dos pilotos); USOS, não alcance à pista ou ultrapassagem dos limites da pista, entre outras.

[49] Disponível em: <http://ultimosegundo.ig.com.br/desastresaereos/decolagem+e+pouso+sao+as+fases+mais+perigosas+do+voo/n1237717482604.html>. Acesso em: 09 nov. 2014.

A remoção das palavras da Fraseologia Aeronáutica por meio da *stoplist* é limitada às listas de palavras, não se estendendo às linhas de concordância. Portanto, nos levantamentos das linhas de concordância, encontram-se muitos exemplos de elocuções de Fraseologia Aeronáutica, pois são parte das transcrições em que há instruções e autorizações previstas nos documentos (vide 3.2.1). A fim de eliminar tais eventos, comparamos o uso da palavra em documentos oficiais que descrevem a Fraseologia Aeronáutica com as linhas de concordância do *corpus* de estudo; aquilo que revelou usos rotineiros foi descartado. Exemplificamos a seguir.

Na Fraseologia Aeronáutica, o uso da palavra *runway* é acompanhado pelas seguintes expressões (Tabela 8):

Expressões retiradas de documentos oficiais	Tradução livre
Anexo 10, p. 5-3: *runway transmitted as* 27 runway two seven 30 runway three zero	 [instruções sobre] pista transmitidas como: 27 pista dois sete 30 pista três zero
DOC 9432, p.p. 4-8, 4-9 ATCO: Fastair 345 runway 09 cleared for take-off Pilot: Cleared for take-off runway 09 Fastair 345	 ATCO: Fastair 345 pista 09 livre decolagem Piloto: Livre decolagem pista 09 Fastair 345
ATCO: G-CD take off immediately or hold short of runway Pilot: Holding short G-CD	ATCO: G-CD decolagem imediata ou mantenha posição próximo à pista Piloto: Mantendo posição [próximo à pista] G-CD
ATCO: G-CD take off immediately or vacate runway Pilot: Taking off G-CD	ATCO: G-CD decolagem imediata ou libere a pista Piloto: Decolando G-CD
ATCO: G-CD Cancel take-off vehicle on runway Pilot: Holding G-CD	ATCO: G-CD cancele decolagem veículo na pista Piloto: Mantendo posição [próximo à pista] G-CD

Tabela 8: Fraseologia Aeronáutica padronizada, com a palavra *runway*.

As expressões apresentadas na Tabela 8 evidenciam como a palavra *runway* ocorre na Fraseologia Aeronáutica. Tais expressões não nos interessam, pois estão relacionadas a situações de rotina. Realizamos, então, uma busca pela palavra *runway* no *corpus* de estudo, que salientou um alto número de ocorrências nas linhas de concordância (607), dificultando a

investigação. Analisamos, portanto, a tela de colocados da palavra *runway*, em que as palavras *on* e *the* aparecem nos dois primeiros lugares. Analisando novamente as linhas de concordância, percebe-se uma tendência de a expressão fixa *on runway*, sem um determinante antes do substantivo, ser acompanhada pelo número da pista, assim como previsto na Fraseologia Aeronáutica (Tabela 8). Essa expressão, portanto, foi descartada de nossas análises, pois não evidenciou material linguístico que justificasse o uso do inglês fora da rotina. Já o agrupamento lexical com determinante antes do substantivo – '*on * runway*' –, gerou 47 linhas de concordância, das quais somente uma (destacada em vermelho na Tabela 9) descreve uma situação de rotina:

N	Concordance	Set	Tag	Word #
1	re // You've got birds on the runway? // Well / we've got on	174	41	100%
2	Yeah / we've got birds on the runway here // You've got bird	167	39	88%
3	you getting some drift on the runway? // No / it's starting	593	147	100%
4	it's starting to drift on the runway right now / and couple	556	136	80%
5	e were / uh we stopped on the runway // Cactus Two seventy-o	43	21	100%
6	eventeen / the uh is on the runway / if you approach end o	214	67	100%
7	/ give me minimum time on the runway // Mayday mayday mayday	52	32	100%
8	is something going on on the runway here // Brickyard thirt	112	45	90%
9	thwest and along there on the runway center you can go over	279	77	48%
10	e three inches of snow on the runway / in probably eighty pe	893	224	100%
11	if he´ll be deplaning on the runway ? // Uh We'll be planni	423	95	97%
12	ll be planning to stop on the runway for an inspection and u	433	97	48%
13	port One / once you're on the runway if you could just do a	256	77	33%
14	rmative // We're stuck on the runway / tower / World eighty-	1.238	381	100%
15	gonna plan on landing on the runway and we'd like the uh fi	118	39	28%
16	ation we're gonna stop on the runway / have the gear pinned	1.028	257	100%
17	there was an incident on the runway / sir // What kind of a	742	224	100%
18	/ We´ll be uh stopping on the runway for an inspection might	142	43	53%
19	tly there´s some metal on the runway / the traffic is about	341	82	100%
20	ither a one eight zero on the runway or foxtrot one? / That is	152	19	65%
21	/ when you come around on the runway let a United go in firs	1.602	499	57%
22	equipment standing by on the runway / obviously we're not g	851	208	100%
23	mber four whiskey fire on the runway // to two seven left //	187	64	100%
24	/ you may the aircraft on the runway if you wish // Oakland	593	205	73%
25	u're gonna be landing on that runway you should be coming he	247	90	63%
26	two seven left // he's on the runway / can you confirm chute	195	68	100%
27	h no signs of anything on the runway? // Uh no sign of anyth	335	102	100%

28	roger // Why you were on the runway there / did you lose an	404	155	88%
29	re / and are you still on the runway? // Seventeen oh two /	146	56	100%
30	here / straight ahead on the runway that would be much easi	343	91	24%
31	u want to stay stopped on the runway or do you think you'll	330	85	55%
32	heavy / any preference on the runway today? // either one is	44	20	86%
33	and we´re gonna stay on the runway? Cactus two seventy-one	695	183	77%
34	ion the the Cactus say on the runway said he may have gone o	614	164	47%
35	whiskey? // Yeah back on the runway short of whiskey / Alit	1.285	399	67%
36	h you responded to // on this runway setup when you need to	111	33	31%
37	u´re just gonna remain on the runway for now // For now we´d	475	139	80%
38	art / the dog is again on the runway / it seems // Holding p	217	53	100%
39	ebris is a metal sharp on the runway / there is quite a bit	314	76	100%
40	lta MD-80 when it gets on the runway and taxi down the runwa	401	140	65%
41	you need anything else on the runway // Cactus seven thirty-	328	113	100%
42	ling and from us it is on the runway // Okay / thank you //	166	50	100%
43	re sitting // It's not on the runway though? // Well it's ha	103	30	86%
44	y-five / are you still on the runway? // Delta three eighty-	307	110	100%
45	ey just found scarring on the runway so that's what makes th	486	122	31%
46	e're still will happen on the runway and we're gonna call ou	999	242	28%

Tabela 9: Linhas de concordância com '*on * runway*'.

As análises das linhas de concordância da Tabela 9 revelaram problemas na pista – há sempre algum evento na pista, ou algo bloqueando a pista. Animais (*dog*, *birds*), veículos de assistência aeroportuária (*equipment*), pedaços de aeronave (*debris*), contaminação (*drift, snow*) ou ainda uma aeronave (*stopped*, *stuck*, *deplaning*) podem interferir na operação segura da pista.

Para estreitar as buscas, recorremos à extração de *clusters*[50], que podem indicar outros tipos de padrões. Assim, confirmamos a primeira impressão, baseada em uma análise manual das linhas de concordância, de que *on the runway* era um *cluster* frequente. Observando, então, os outros *clusters*, constatamos a alta presença do verbo *TO BE*[51], e outras palavras que julgamos pertinente investigar, como *exit* e *get off* (pois a palavra que representa o ato de sair da pista é *vacate*, segundo a Fraseologia Aeronáutica) e *of the runway*. Veremos, a seguir, a tabela dos agrupamentos lexicais da palavra *runway*:

[50] Utilizaremos *cluster* e agrupamento lexical alternadamente.

N	Cluster	Freq.
1	ON THE RUNWAY	46
2	OFF THE RUNWAY	16
3	OF THE RUNWAY	16
4	THE RUNWAY BE	11
5	YOU BE CLEAR	9
6	EXIT THE RUNWAY	9
7	RUNWAY BE CLOSE	7
8	DOWN THE RUNWAY	7
9	BE ON THE	6
10	GET OFF THE	6
11	APPROACH END OF	5
12	BE CLEAR TO	5
13	FOR THE RUNWAY	5
14	STOP ON THE	5

Tabela 10: *Clusters* com a palavra *runway*.

Como o primeiro *cluster* já havia sido analisado, partimos para o segundo agrupamento lexical mais frequente, *off the runway*, com dezesseis ocorrências nas linhas de concordância. A análise dessas linhas revelou que *off the runway* pode acompanhar o verbo TO BE, ou verbos de suporte (*GET*, *GO*, *COME*) que dependem de *off* para apresentar o significado de liberar a pista, o que poderia ser rotineiro caso não identificássemos um caráter de urgência nas instruções de liberação da pista nessas linhas de concordância. Reconhecemos essa prosódia semântica ao analisarmos essas linhas em contexto ampliado, que é a visualização de cada linha dentro do próprio texto em que cada uma ocorre. Atentando-se à pragmática de o*ff the runway*, podem-se observar outras situações em que tal agrupamento lexical é utilizado. A seguir, na Tabela 11, notam-se três situações distintas: a já notada urgência em liberar a pista; a incapacidade de liberação da pista devido a algum problema técnico na aeronave; e a saída involuntária da pista (excursão de pista):

N	Concordance	Categoria
1	at we are all the way off the runway into the snow covered t	Saída de pista
2	Say again sir // Get off the runway! // Cessna eight nine d	Urgência

3	don’t care // Just go off the runway // seven seven fox / ex	Urgência
4	even seven fox / turn off the runway please // Sorry sir / t	Urgência
5	able to get ourselves off the runway without one and I can’t	Inabilidade em sair
6	exit the runway / get off the runway / I don’t care what you	Urgência
7	t to verify / you are off the runway on the departure end /	Saída de pista
8	y-six heavy uh turning off of runway two two right on Juliet	Urgência
9	re what you do // Get off the runway / seven seven fox // Al	Urgência
10	not gonna be getting off the runway as we land / so that ru	Inabilidade em sair
11	eighty-five / are you off the runway? // Delta three eighty-	Saída de pista
12	he corporate jet came off the runway and we couldn’t / we ha	Saída de pista
13	said he may have gone off the runway just a little bit so on	Saída de pista
14	port one two plus two off the runway // Airport one two plus	Urgência
15	o says that can’t get off the runway // can you give me the	Inabilidade em sair
16	n alpha / report when off the runway traffic on a four mile	Urgência

Tabela 11: Linhas de concordância com o *cluster off the runway* ao centro.

A presença marcante do verbo *TO BE* na lista de *clusters* leva-nos à próxima busca: *runway* e as formas do verbo, como *am*, *´m*, *are*, *´re*, *is*, *´s*, *was*, *were* e *be*, este último como forma infinitiva com verbos modais, por exemplo. A seguir, na Tabela 12, vemos as linhas de concordância geradas com *runway* e o verbo *TO BE*, conjugado na terceira pessoa do singular (*is*):

N	**Concordance**
1	/ QNH one zero two one / the **runway is** blocked // Going aro
2	ve x-ray / please confirm the **runway is** vacated on whiskey /
3	t // two seven left runway // **runway is** closed // yeah / / t
4	hank you for the report / the **runway is** shut down obviously
5	land / earliest exit from the **runway is** appreciated / the wi
6	LS // Affirm you can use that **runway is** clear / back open //
7	inspection to ensure that the **runway is** serviceable // I und
8	s when he comes back but your **runway is** looking pretty obscu
9	e uh runway is closed? // The **runway is** closed sir // Novemb
10	// Currently right now / the **runway is** closed due to an inc
11	ear / back open // Roger that **runway is** open // Tower / Airp
12	d One / Roger // Check if the **runway is** open // Airport ten
13	/ / two seven is closed / // **runway is** closed / aircraft cr
14	o the runup area // So the uh **runway is** closed? // The runway

Tabela 12: Linhas de concordância com *runway is*.

Essas linhas mostram, em sua maioria, descrições da condição da pista em uso. O *cluster* é normalmente formado por *the runway is*, mas também pode ser precedido pelo determinante *that*. As condições podem ser favoráveis, como *open* e *clear*, ou adversas, impossibilitando o uso da pista, identificadas por *closed*, *blocked* ou *shut down* [fechada ou bloqueada]. Resumindo esses exemplos:

The runway is clear.
open.
closed.
blocked.
shut down.

Essa foi uma primeira análise com a forma verbal *is* ocorrendo à direita de *runway*; porém, *is* aparece com mais frequência à esquerda de *runway* (37 ocorrências), de forma mais distribuída. Observam-se, na tabela 13, os números de ocorrências de *is* à esquerda de *runway*, em vermelho:

N	Word	With	Set	Texts	Total	Total Left	Total Right	L5	L4	L3	L2	Centre
3	IS	runway	29	73	36	37	7	10	13	5	1	

Tabela 13: Posições de *is* à esquerda de *runway*.

Portanto, gerando linhas de concordância com *is* à esquerda, constatamos descrições de algo sobre (*on*) ou na direção (*for, to*) da pista. É comum o uso das preposições (em destaque na tabela) que reforçam essas ideias (Tabela 14):

N	Concordance
1	the position **is** the threshold **runway** two seven left / aircra

2	ling and from us it **is** on the **runway** // Okay / thank you //
3	/ that traffic **is** taking the **runway** now // Alright / AC fif
4	traffic **is** uh **is** inbound for **runway** two two right // Okay /
5	emergency **is** inbound now for **runway** four right / it might b
6	art / the dog **is** again on the **runway** / it seems // Holding p
7	ify everybody **is** north of the **runway** / I will be // Tower ai
8	**is** an emergency inbound to **runway** two two right / if he d
9	there **is** an engine cowling on **runway** two seven right prior t
10	tly there´**s** some metal on the **runway** / the traffic is about
11	two seven left // he'**s** on the **runway** / can you confirm chute
12	eah / we know / he'**s** stuck on **runway** heading / he's uh only
13	re sitting // It'**s** not on the **runway** though? // Well it's ha
14	scar Mike / what'**s** the active **runway** at Kelowna? // Okay / i
15	nd that´**s** probably landing on **runway** two two right / if he c
16	d / what'**s** that with crossing **runway** six? // Uh tower / Cact
17	he runway / he'**s** crossing the **runway** now // Darwin zero one
18	ty-five heavy // what **is** your **runway** request? One zero // Wo
19	? // Which one **is** the closest **runway**? It actually appears to
20	roger // Airport One / **is** the **runway** back open yet? // Runwa

Tabela 14: Linhas de concordância com o núcleo *runway* e *is* como colocado à esquerda.

Percebem-se salientadas nessa busca, especificamente nas linhas 4, 5, 8, 9 e 15 da Tabela 14, as poucas ocorrências em que as combinações com *runway* e números (sem um determinante precedendo o substantivo) representam situações anormais. As descrições de algo sobre a pista acompanham a preposição *on*; aeronaves (*traffic*, *emergency*) se dirigindo à pista são seguidas de *for* ou *to*. Há, também, ocorrências de frases interrogativas (linhas 14, 16, 18, 19 e 20) e presente contínuo (3, 15, 16 e 17). Ao ampliarmos as linhas de concordância no próprio WST (e avaliarmos as linhas no contexto ampliado), notamos a forte presença do verbo existencial *THERE TO BE*, como nas linhas 8, 9 e 10 da Tabela 14. Vejamos alguns exemplos:

1. *debris is a metal sharp on the runway / **there is** quite a bit*
2. *apparently **there´s** some metal on the runway*
3. ***there is** something going on on the runway here*
4. ***there was** an incident on the runway*
5. *south of the runway **there are** power lines uh a lot of houses around the area*

Para as descrições sobre algo na pista (ou próximo a ela), portanto, utiliza-se o verbo *is* como verbo de ligação, ou o verbo *THERE TO BE*, combinado com as preposições *on* (para

descrição de algo na pista), *of* (para determinar o local da pista onde o problema ocorreu) e *to* (com a ideia de movimento para a pista).

Selecionamos, em seguida, o quarto colocado mais frequente com *runway*, outra forma do verbo *TO BE*: *´re* (Tabela 15):

N	Concordance
1	ation we'**re** gonna stop on the **runway** / have the gear pinned
2	will be / we'**re** short of the **runway** shortly after the alert
3	and we´**re** gonna stay on the **runway**? / Cactus two seventy-one
4	fifty // We´**re** uh exiting the **runway** at this moment here / we
5	e sure you'**re** the only on the **runway** one two oh four three m
6	ey // You'**re** lined up for the **runway** so altitude at your dis
7	/ You'**re** looking at the right **runway** / That's Runway six at
8	port One / once you'**re** on the **runway** if you could just do a
9	o you'**re** gonna have plenty of **runway** to work with / and we g
10	/ It's a twelve thousand foot **runway** so you'**re** gonna have pl
11	Okay / you'**re** still okay for **runway** twenty-four right? / sw
12	Okay / sir / we'**re** working on **runway** one five right and I go
13	ten / ah we copy that / sir / **runway** will be / we'**re** short o
14	o six while you'**re** gonna taxi **runway** one two via alpha // Ok
15	rmative // We'**re** stuck on the **runway** / tower / World eighty-
16	man / while you'**re** gonna taxi **runway** one two via alpha // Ae

Tabela 15: Linhas de concordância com o verbo *´re*, sendo *runway* o nódulo.

We e *you* são pronomes de cunho conversacional, utilizados nesse *corpus* nas referências às aeronaves – as linhas em que *we* aparece são de elocuções emitidas por pilotos; as linhas com *you* são enunciadas por ATCOs. Portanto, confirmando as análises anteriores, há relatos da posição de algo na pista (sendo, aqui, aeronaves, representadas por *we* ou *you*) ou próximo à pista (*short of* significa, no jargão aeronáutico, 'pouco antes de'). A forte presença de *gonna* nas linhas 1, 3, 9, 14 e 16 da Tabela 15 indica ações futuras em relação ao uso da pista. Essas ações também são expressas por advérbios de tempo (*once you´re on the runway*, *while you´re gonna taxi to runway one two*).

Em um estudo sobre a linguagem utilizada por ATCOs (PHILPS, 1991), verificam-se transformações no nível frasal em elocuções emitidas por ATCOs ocasionadas pela deleção de determinantes, de verbos de ligação, de preposições e do advérbio *when*. Como ilustração

de deleção de algumas dessas palavras, pode-se considerar a frase *wheels appear up* [o trem de pouso parece recolhido] como uma forma curta de *wheels appear* (*to be) up* [o trem de pouso parece (estar) recolhido], *resume own navigation* [inicie própria navegação] como *resume (your) own navigation* [inicie (sua) própria navegação] (BARSHI; FERRIS, 2013, p. 51), e ainda a frase *speed at own discretion* [velocidade a seu critério] como uma forma abreviada de (*your) speed is at (your) own discretion* [a sua velocidade está ao (seu) critério]. Em uma pesquisa posterior, investigou-se se essas deleções causavam ambiguidade ou mau entendimento por parte dos pilotos; concluiu-se que a falta de verbos auxiliares e de ligação não afetava a compreensão eficaz do piloto (BARSHI; FERRIS, 2013).

Em oposição à omissão do verbo de ligação *TO BE* na Fraseologia Aeronáutica, em nossas análises do inglês utilizado fora da rotina, nota-se o uso frequente desse verbo e de verbos modais: o verbo *TO BE*, lematizado, é a palavra mais frequente do *corpus* (vide Tabela 5) e continua sendo a primeira na lista de palavras-chave (vide Tabela 6) – ou seja, quando contrastado com um *corpus* de referência, o verbo *TO BE* é muito mais recorrente no *corpus* de estudo do que no *corpus* de inglês acadêmico.

Dando sequência à investigação do verbo *TO BE*, observa-se que, em sua forma infinitiva, *be* aparece como 8º colocado, totalizando 25 ocorrências. Vale a pena gerarmos as linhas de concordância para um estudo mais aprofundado (Tabela 16):

N	Concordance
1	/ We´**ll be** uh stopping on the **runway** for an inspection / might
2	u're **gonna be** landing on that **runway** / you should be coming he
3	they're asking is if you think **runway** two seven **will be** accep
4	e **should be** able to clear the **runway** on our own // Okay chec
5	emergency is inbound now for **runway** four right / it **might be**
6	e five / the next arrival for **runway** four right **will be** him
7	**will be** holding their position **runway** one five right // Clear
8	ee Thousand / **will be** vectors **runway** two seven approach // R
9	e runway as we land / so that **runway will be** down / and ah w
10	ify everybody is north of the **runway** / I **will be** // Tower ai
11	ou to taxi and then cross the **runway** and then **be** in the sout
12	**d like to be** vectored to left **runway** downwind for runway one
13	// That's affirmative **will be runway** zero five alright / che
14	losest runway / that **would be runway** three three left // Uh
15	proceed direct to the closest **runway** / that **would b**e runway
16	ward and kinda go towards the **runway** that **would be** bad / so
17	/ Delta thirty heavy uh which **runway will be** your request on

18	e **will be** vectors for Toronto **runway** five and above to one o
19	if he´**ll be** deplaning on the **runway** ? // Uh We'll be planni
20	here / straight ahead on the **runway** that **would be** much easi
21	rescue six **will be** doing the **runway** inspection / Airport el
22	ten / ah we copy that / sir / **runway will be** / we're short o
23	ding / **you will be** overflying **runway** two two right / but I w
24	not **gonna be** getting off the **runway** as we land / so that ru
25	nine delta whiskey / it'**ll be runway** six / It's a twelve tho

Tabela 16: Linhas de concordância com o 8º colocado de *runway*, o lema *TO BE*.

A presença de verbos modais confirma a ideia de ações futuras em relação ao uso da pista (*will*, *going to*); adiciona-se a ideia de possibilidade, como vemos nos exemplos a seguir:

- sobre o bloqueio da pista:

*We´**ll be uh stopping** on the runway for an inspection*
*so that runway **will be down***
***should be able to clea**r the runway on our own*
*he´**ll be deplaning** on the runway?*
*rescue six **will be doing** the runway inspection*
***not gonna be getting off** the runway as we land*

- sobre qual pista usar:

*[you]'re **gonna be landing** on that runway*
*they're asking is if you think runway two seven **will be accep[table]***
*[I´]**d like to be vectored** to left runway downwind for runway one*
*That's affirmative **will be** runway zero five alright*
***that would be** runway…*

Em estudos descritivos sobre o inglês comum, os verbos modais carregam sentidos de certeza ou de necessidade, enfatizando ou amenizando tais sentidos ao serem combinados com os verbos lexicais que os seguem; uma lista de verbos modais é apresentada por Carter e McCarthy (2006), que adicionam a ela palavras que expressam a mesma função modal: (i) verbos lexicais como *look* e *seem*; (ii) adjetivos como *possible* e *certain*; e (iii) advérbios como *maybe*, *probably* e *definitely*. Por isso, baseados na alta frequência desses itens

indicadores de modalidade, há estudos que sugerem abordagens pedagógicas voltadas a eles desde os níveis iniciais em cursos de inglês (O´KEEFE; McCARTHY; CARTER, 2007, p. 37). Interessa-nos, portanto, confirmar se o uso de modais se aplica ao nosso *corpus* de estudo. Partindo da *Wordlist* lematizada no início de nossa investigação (Tabela 5), geramos três outras listas – de agrupamentos lexicais de três, quatro e cinco palavras – em que pudemos notar a alta frequência dos modais já evidenciados pelas nossas análises de *runway*, agrupadas abaixo (Tabela 17):

N	Word	Freq.
1	WE **D LIKE TO**	23
2	DO YOU **WANT TO**	15
3	I **NEED** YOU **TO**	13
4	ARE YOU **ABLE TO**	12
5	RE GONNA **HAVE TO**	12
6	WE RE **GONNA** HAVE	12
7	WE RE **GOING TO**	11
8	WE RE **GONNA NEED**	11
9	YOU **WANT** US **TO**	10
10	DO YOU **NEED TO**	9
11	I **D LIKE TO**	8
12	WE **LL** GET BACK	8
13	WE **LL** GIVE YOU	8
14	WE RE **GONNA** DO	8
15	IT S **GONNA** BE	7
16	**LL** GIVE YOU A	7
17	SO WE RE **GONNA**	7
18	WE RE **GONNA** BE	7
19	BE **ABLE TO** MAKE	6
20	I **LL** TELL YOU	6

Tabela 17: Seleção dos vinte primeiros *clusters* de três palavras com ideia de futuro e com modais.

Nas análises desses *clusters*, nota-se que o futuro é expresso tanto com o modal *will* quanto com a combinação *be going to* – contraído como *gonna*. Em investigações utilizando os *clusters* da Tabela 17, confirmamos que muitas das ocorrências são com a palavra *runway*. A modalização, confirmada também pela lista de *clusters* de três palavras, é um elemento constante também no inglês comum utilizado por ATCOs e pilotos.

Para organizarmos as investigações encabeçadas por *runway*, agrupamos as conclusões às quais chegamos:

- descrições de algo na pista:
  the dog is again on the runway
  there was an incident on the runway
  there´s some metal debris on the runway
  he´ll be deplaning on the runway
  we´re gonna stop on the runway
  we´ve got birds on the runway here
- caráter de urgência ao requisitar a liberação da pista:
  get off the runway
  go off the runway
  turn off the runway
  report when off the runway
- excursão de pista:
  the corporate jet came off the runway
  he may have gone off the runway just a little bit
- incapacidade em liberar a pista:
  who says that can´t get off the runway
  we´re not gonna be getting off the runway as we land
- descrições da condição da pista:
  the runway is closed
  the runway is blocked
  the runway is shut down
  the runway is clear
  the runway is open
- existência de algo na ou próximo à pista:
  apparently there´s some metal on the runway
  there is something going on on the runway
  south of the runway there are power lines
- sobre o bloqueio da pista:

We´ll be uh stopping on the runway for an inspection
so that runway will be down
should be able to clear the runway on our own

- sobre qual pista usar:
  you're gonna be landing on that runway
  they're asking is if you think runway two seven will be accep[table]
  [I´]d like to be vectored to left runway downwind for runway one

4.1.2 *Emergency*

Em um *corpus* de comunicações aeronáuticas em situações anormais, não deveria ser surpreendente observar a palavra *emergency* em segundo lugar na lista de palavras de conteúdo; no entanto, na busca pela palavra *emergency* nos documentos de Fraseologia Aeronáutica, encontra-se o seguinte modelo de diálogo entre piloto e ATCO (Figura 17):

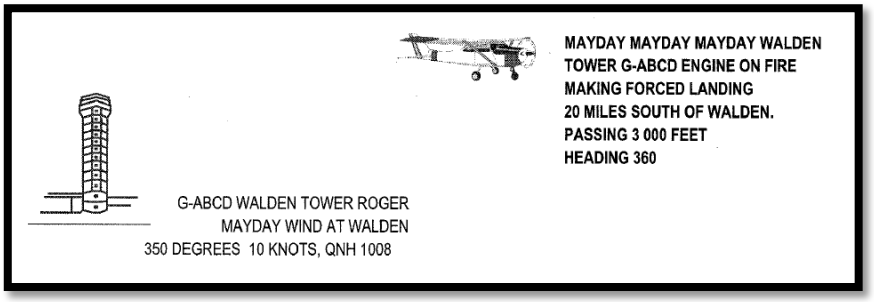
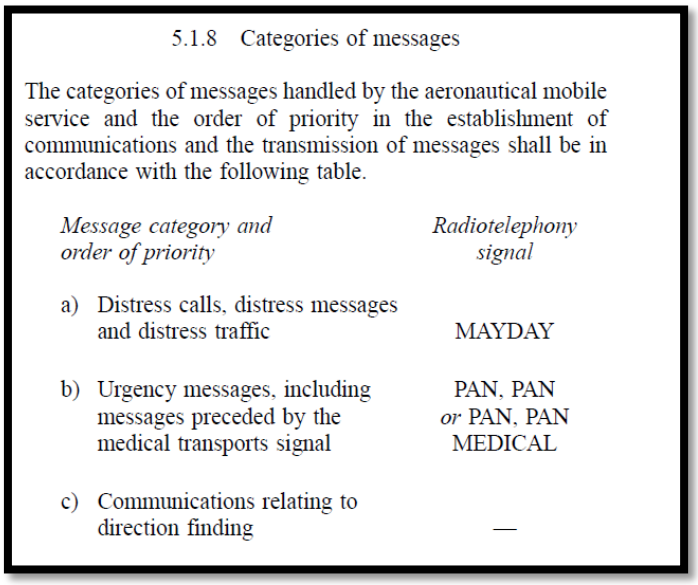


Figura 17: Modelo de diálogo retirado do DOC 9432, Manual de Fraseologia Aeronáutica.

Na Figura 17, pode-se observar que não há menção à palavra *emergency* em um diálogo entre piloto e ATCO simulando uma emergência com o motor e a necessidade de um pouso forçado, ou seja, um pouso fora do aeroporto. A busca pela palavra *emergency* nos documentos de Fraseologia Aeronáutica não retornou nenhum resultado; em situações de emergência, o piloto deve relatar a ocorrência ao ATCO de duas maneiras, conforme excerto extraído do documento da OACI (2001, p. 5-1), apresentado na Figura 18:



5.1.8 Categories of messages

The categories of messages handled by the aeronautical mobile service and the order of priority in the establishment of communications and the transmission of messages shall be in accordance with the following table.

Message category and order of priority	*Radiotelephony signal*
a) Distress calls, distress messages and distress traffic	MAYDAY
b) Urgency messages, including messages preceded by the medical transports signal	PAN, PAN *or* PAN, PAN MEDICAL
c) Communications relating to direction finding	—

Figura 18: Extrato de documento oficial sobre categoria de mensagens de emergência.

Na Figura 18, há duas categorias distintas, definidas pelo grau de urgência requerida. A emergência grave, em que a aeronave está em sério risco, é relatada pela palavra *Mayday*; já aquela em que o piloto necessita de atenção – mas com a situação ainda sob controle – é reportada seguida de *Pan, Pan*. Isso é o que 'deveria' ocorrer. No entanto, no *corpus* de estudo a palavra *emergency* ocorre 133 vezes, enquanto a palavra *mayday* tem dezesseis ocorrências, e *pan*, dezessete.

Ao gerarmos as linhas de concordância com a palavra *emergency*, percebemos claramente um padrão recorrente que pode ser demonstrado pela lista de *patterns*, um utilitário da ferramenta *Concord*. Nele, visualizamos a palavra de busca – *emergency* – ao centro, e a janela (*span*) de cinco palavras à esquerda e à direita que ocorrem com a palavra de busca. Direcionaremos a análise por essa janela, exibida na Tabela 18:

N	L5	L4	L3	L2	L1	Centre	R1	R2	R3	R4	R5
1	THE	**WE**	**RE**	**DECLARING**	**AN**	EMERGENCY	AIRCRAFT	**THIS**	**TIME**	TURN	RIGHT

2	WE	THE	THERE	TO	THE	WE	TO	THE	TWO	TWO
3	GO	ONE	THE	**DECLARE**	OF	INBOUND	YOU	TO	AND	WE
4			I	IS		**AT**	AT	S	THE	TO
5			AND	**HAVE**		AND	I	EIGHT		ON
6			ARE	OF		VEHICLES	IS	**NOW**		
7			FOUR	NATURE		EQUIPMENT		ON		

Tabela 18: Lista de colocados da palavra *emergency*.

Verificam-se à esquerda, na posição L2, os verbos *declare* e *have*. Há, também, a ocorrência do verbo de ligação *TO BE*. São, portanto, três verbos utilizados com *emergency*. Na posição L3, temos a palavra *there*, que, combinada com o verbo *is* na posição L2, forma o verbo *there is*. A presença de *´re* e *are* à esquerda dos verbos, precedidos do pronome *we*, indica um possível padrão com *we´re/are declaring*. À direita, notam-se os substantivos *aircraft, vehicles* e *equipment*, em colocações com a palavra *emergency*. Outras ocorrências que chamam a atenção são a palavra *now* e um possível agrupamento com *at this time* à direita do nódulo.

Os agrupamentos sugeridos pela primeira análise são confirmados na tela de *clusters* (Tabela 19):

N	Cluster	Freq.
1	DECLARING AN EMERGENCY	16
2	AT THIS TIME	7
3	EMERGENCY AT THIS	6
4	DECLARE AN EMERGENCY	6
5	WE RE DECLARING	5
6	AN EMERGENCY INBOUND	5

Tabela 19: Agrupamentos lexicais com a palavra *emergency*.

Na Tabela 19, notam-se agrupamentos que se complementam: o *cluster* da linha 3, por exemplo, é claramente o início de '*at this time'*, o *cluster* na linha 2; o da linha 5 (*we´re*

declaring) se completa com o primeiro (*declaring an emergency*). Isso demonstra uma forte repetição dos fraseologismos com a palavra *emergency*. Assim, identificamos, também, nas linhas de concordância, o imediatismo típico das situações de emergência na aviação (contexto de construção do *corpus*), evidenciado pelos advérbios de tempo *now, right now* e *at this time,* além do uso do presente contínuo, principalmente evidenciado por *declaring an emergency*.

A observação das linhas de concordância de *emergency* também revelou outros fraseologismos: um deles, com o verbo existencial *there is*, que também se mostrou recorrente assim como no entorno de *runway* (como visto a seguir, com exemplos retirados do contexto ampliado):

1. *we're checking around / there's an emergency with an engine fire potentially number four//*
2. *they always do that when there's an emergency*
3. *stay where you are and hold / there is an emergency inbound*
4. *continue the approach / there is an emergency inbound to runway two*
5. *we're surrounded by emergency vehicles / there's a reason for this //*

Outro colocado, *nature*, direcionou nosso olhar para a brevidade das elocuções e a necessidade de simplificação:

1. *When you have a chance / nature of the emergency //*
2. *say the <break> say the nature of your emergency /*
3. *I missed the nature of emergency and also the uh souls on board*
4. *state nature of emergency*
5. *and you said the nature of emergency was a bird strike? /*

A expressão *nature of emergency*, ou 'natureza da emergência', é emitida quando o ATCO necessita saber que problema ocorreu com a aeronave. Em vez de frases interrogativas, o que vemos é uma simplificação com o uso de verbos relacionados à ação de dizer. Há ainda, nessas últimas dez linhas, uma possível subordinação a ser investigada. Tal análise foi realizada de forma mais completa na seção 4.1.5, com a palavra *fuel*, que apresentou mais possibilidade de pesquisa. Concentramo-nos, aqui, na observação do verbo *declare*, que demonstrou possibilidade de conjugação além do presente contínuo visto na lista de *clusters*. No contexto expandido, notam-se os seguintes padrões no entorno de *emergency*:

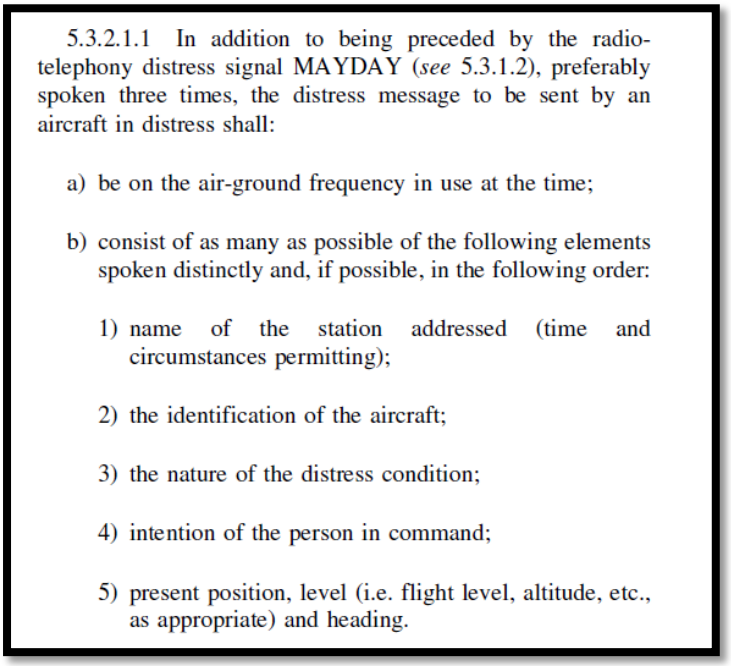
1. *we are declaring an emergency / we did <break> we have called the fire department*
2. *we had to declare an emergency*
3. *And just to confirm we did declare an emergency and have the equipment standing by*
4. *remove everybody from our way / We've declared an emergency /*
5. *declaring an emergency / we had to go around*
6. *we have declared an emergency / engine failure on the right engine*

Os exemplos acima representam eventos que foram executados em um passado recente – conjugados no passado simples ou no presente perfeito – com consequências no momento da fala. Essas consequências são evidenciadas por elocuções simples (*we have called the fire department*, *we had to go around*, *we have declared an emergency*), sem a presença de conjunções ou advérbios de tempo. Percebe-se, aqui, a dependência de conhecimento do contexto por parte do interlocutor, inclusive sobre que tipo de informação deve esperar do emissor da mensagem. Esse tipo de mensagem lacônica se deve à brevidade exigida pelo momento, pois, quando uma aeronave está em uma situação de emergência, a carga de trabalho dos pilotos aumenta demasiadamente. A comunicação, portanto, deve ser feita de forma objetiva, informando ao ATCO qual é a natureza da emergência e quais as intenções dos pilotos. É recomendado ao ATCO que aja de forma a prestar assistência ao piloto de forma concisa:

> Uma estação respondendo (ou originando uma resposta) a uma aeronave em emergência deve prover tantas recomendações, informações e instruções quanto sejam necessárias para assistir o piloto. Transmissões supérfluas podem distrair os pilotos em um momento em que já têm muito a fazer.[52] (ICAO, 2007, p. 9-2).

Para evitar o supérfluo, assim está descrito no Anexo 10 da Convenção de Chicago (ICAO, 2001, p. 5-19) sobre situações de emergência (Figura 19):

[52] *"A station replying (or originating a reply) to an aircraft in distress or urgency should provide such advice, information and instructions as is necessary to assist the pilot. Superfluous transmissions may be distracting at a time when the pilot's hands are already full."*



5.3.2.1.1 In addition to being preceded by the radiotelephony distress signal MAYDAY (*see* 5.3.1.2), preferably spoken three times, the distress message to be sent by an aircraft in distress shall:

a) be on the air-ground frequency in use at the time;

b) consist of as many as possible of the following elements spoken distinctly and, if possible, in the following order:

1) name of the station addressed (time and circumstances permitting);

2) the identification of the aircraft;

3) the nature of the distress condition;

4) intention of the person in command;

5) present position, level (i.e. flight level, altitude, etc., as appropriate) and heading.

Figura 19: Extrato do Anexo 10 da Convenção de Chicago sobre comunicações de emergência.

Nos exemplos apresentados anteriormente, os pilotos dependem que o ATCO reconheça os procedimentos de emergência a serem seguidos na cabine de comando. Essa é uma característica particular da comunicação radiotelefônica, segundo a OACI (2010) – ser altamente dependente do contexto[53]. Pilotos e ATCOs necessitam dominar as operações para que possam gerenciar bem as situações problemáticas. Por esse motivo, os pilotos são estimulados a reportar seu problema e sua intenção ao ATCO, para assim evitar comunicações prolixas. Daí, justificam-se a falta de conjunções e o foco em elocuções simples.

Uma última aplicação da palavra *emergency* é vista na sua colocação imediatamente à esquerda de alguns substantivos: *equipment*, *vehicles* e *aircraft*, referindo-se ora aos veículos de assistência à emergência, ora à própria aeronave em emergência. Esse último uso também é percebido na colocação *emergency inbound*, em que há omissão da palavra *aircraft* (*emergency aircraft inbound* – aeronave em emergência aproximando).

A análise das linhas de concordância da palavra *emergency* pode ser assim resumida:

[53] *High context-dependent.*

- Há alta frequência de palavras relacionadas ao momento presente (*now, right now, at this time, immediately*).
- Além do presente contínuo (*we´re declaring an emergency*), verbos indicadores de posse (*I, we have, ´ve got*) e o verbo existencial *THERE TO BE* são utilizados com o *cluster* '*an emergency'*.
- A característica sucinta da comunicação aeronáutica elimina o uso de conjunções.
- Há priorização do passado ou do presente perfeito para o esclarecimento de causas e consequências.
- Pode formar uma colocação nominal, como em *emergency equipment*, *emergency vehicles* ou *emergency aircraft*; no uso combinado com *inbound*, omite-se a palavra *aircraft* (*an emergency aircraft inbound*).

4.1.3 *Aircraft*

Ao gerarmos as linhas de concordância com determinada palavra de busca, a ferramenta *Concord* as apresenta, numa primeira instância, com a primeira palavra à esquerda em ordem alfabética. O pesquisador pode, no entanto, formatar a ferramenta para destacar outros colocados, à direita e à esquerda.

Uma concordância com a palavra *aircraft* ao centro, ordenada alfabeticamente pela L1, sinalizou uma alta frequência de determinantes nessa posição. A ferramenta *Concord* também disponibiliza ao pesquisador a possibilidade de organização manual das linhas de concordância: cada linha recebe, na coluna SET, uma letra correspondente à classificação desejada. No nosso caso, classificamos as linhas pela função gramatical da palavra na posição L1 e atribuímos D aos determinantes, A aos adjetivos e P às preposições. *The aircraft* ocorre 66 vezes nas linhas de concordância; *that aircraft* aparece nove vezes. Como já observado, a presença de determinantes é muito superior às outras classes gramaticais selecionadas imediatamente ao lado esquerdo de *aircraft*, enquanto *your* ocorre nove vezes; *this*, seis; *all*, seis; *an*, três; *any*, uma; *another*, uma; e *other*, uma, totalizando 102 ocorrências com determinantes em 148 linhas geradas com *aircraft*. Esses resultados contestam, mais uma vez, o estudo de Philps (1991), que alega haver deleção de determinantes na comunicação

aeronáutica rotineira; esse estudo já havia sido parcialmente refutado por Barshi e Ferris (2013), que verificaram o uso de determinantes em um *corpus* compilado em situações de rotina.

Em seguida, apresentamos, na Tabela 20, os colocados de *aircraft*, suas posições à esquerda e à direita em uma janela de três e, em vermelho, a posição mais frequente de cada colocado:

N	Word	Total	Total Left	Total Right	L3	L2	L1	Centre	R1	R2	R3
1	AIRCRAFT	159	4	4		3	1	151	1	3	
2	THE	98	83	15	2	16	65		3	6	6
3	IS	31	8	23	2	6			17	2	4
4	TO	30	22	8	9	13			4	1	3
5	OF	22	17	5	1	14	2			2	3
6	THAT	21	12	9	1	1	10		5	2	2
7	ON	20	6	14	2	4			5	6	3
8	UH	18	11	7	3	3	5		4	2	1
9	YOU	15	6	9	5	1			1	7	1
10	AND	12	6	6	2	4			6		
11	YOUR	11	10	1	1		9			1	
12	EMERGENCY	10	9	1			9		1		
13	WE	9	1	8	1				2	6	
14	DOWN	8	1	7	1				6		1
15	NOW	8	0	8					2	5	1
16	IN	8	2	6		1	1		6		
17	THERE	7	0	7					1	2	4
18	ARE	7	3	4	3				1		3
19	AN	7	6	1		2	4		1		
20	RIGHT	6	4	2	4				2		

Tabela 20: Vinte primeiros colocados da palavra *aircraft*.

Os colocados de *aircraft* na tela *Collocates* da ferramenta *Concord* podem ser investigados da seguinte forma: ao clicar no número, podem-se agrupar somente as linhas de concordância que apresentam aquele colocado naquela determinada posição (à direita ou à esquerda) ou em qualquer posição ao selecionar a coluna *Total*. O primeiro colocado cuja frequência total nos interessou foi, novamente, *is*. Em 31 linhas de concordância, pode-se observar o uso de *is* para determinar a posição da aeronave (*aircraft is on the turn to papa*; *emergency aircraft is on the runway*; *is the aircraft on final?)*, o tipo da aeronave (*aircraft type is a triple seven*; *what type of aircraft is United ten following?*), e a situação em que a aeronave se encontra (*the aircraft is under control*; *the aircraft is stable at this time*).

A preposição *to*, o quarto colocado mais frequente de *aircraft*, é extraída nas linhas de concordância com três funções: entre as linhas 1 e 7, com ideia de movimento; das linhas 8 a 13, como complemento gramatical de verbos de elocução (intransitivos, em inglês) ou da regência da palavra *damage* (dano); entre as linhas 14 e 27, complementando o infinitivo de verbos (Tabela 21):

N	Concordance	Set
1	g on runway four right **to** the **aircraft** // Acey fifty-eight t	Movimento
2	a to get uh back down **to** that **aircraft**? // Rescue unit / you	Movimento
3	e gonna go ahead and have the **aircraft** tow back **to** the gate	Movimento
4	Striker proceed **to** the downed **aircraft** // And tower for Beec	Movimento
5	’m gonna proceed over **to** that **aircraft** who says that can’t g	Movimento
6	onto runway four right **to** the **aircraft** // Rescue six five fi	Movimento
7	fire engine truck out **to** your **aircraft** / to make sure everyt	Movimento
8	tower did the **talking to** that **aircraft** // were you talking t	Dizer
9	e // were you **talking to** that **aircraft** that just went into S	Dizer
10	re **responding to** the alerting **aircraft** / we’d like to procee	Dizer
11	ure there´s no **damage to** the **aircraft** and if there is and s	Dano
12	is there any **damage to** your **aircraft**? // Uhh not that I ca	Dano
13	hat there´s no **damage to** the **aircraft** // let me find out /	Dano
14	we’re just checking under the **aircraft to** make sure no fuel	Ação
15	ort ten / they are the second **aircraft to** land after Southwe	Ação
16	s rescue six please advise if **aircraft** is next **to** land // Re	Ação
17	ing uh a tow truck **to** tow the **aircraft** off back to the runwa	Ação
18	that you need **to** go to // the **aircraft** in question is turnin	Ação
19	na have someone **to** circle the **aircraft** / uh take a visual /	Ação
20	two / Kennedy Ground // The **aircraft** appears **to** be normal	Ação
21	rstand you are able **to** fly an **aircraft**? // Single engine/ I	Ação

22	something in here **to** tow the **aircraft** back on the pavement	Ação
23	ll be people approaching your **aircraft to** inspect it // Okay	Ação
24	ur position // Prefer **to** land **aircraft** at destination airport	Ação
25	ou're not able **to** control the **aircraft** / Is that correct? //	Ação
26	w / you guys need **to** shut the **aircraft** down // air shuttle s	Ação
27	r we're going **to** evacuate the **aircraft** // you got sixty seco	Ação

Tabela 21: Linhas de concordância da palavra *aircraft* com *to* no entorno.

A ideia de movimento da preposição *to* evidenciada nas primeiras linhas quase sempre é emitida a veículos que prestam serviços de emergência às aeronaves; a única exceção é a linha 7, em que o ATCO avisa à aeronave que lhe enviará um veículo de assistência (*we´re gonna send a fire engine truck out to your aircraft*). Essa ideia de movimento também ocorre com o agrupamento *the runway*, visto na Seção 4.1.1. Pela ocorrência das linhas em que *to* aparece como constituinte de infinitivo, podemos categorizá-las dentro das seguintes funções:

- marcando finalidade (*a tow truck to tow the aircraft off back to the runway, have someone to circle the aircraft*)
- seguindo o verbo ou expressão modal (*need to, able to, the aircraft appears to be normal*)
- complementando expressões com ideia de futuro (*we´re going to evacuate the aircraft*).

Retomando as análises das preposições que coocorrem com *aircraft*, *of* aparece como 5º colocado mais frequente. Isolamos, então, as linhas de concordância em que *aircraft* ocorre com *of* na posição L2, a mais frequente (Tabela 22):

N	**Concordance**
1	on nine one one report **of** an **aircraft** down // I couldn't ge
2	'm gonna run a bunch **of** other **aircraft** until you're set out
3	re in the tug in front **of** the **aircraft** here // Who called? /
4	pilling from the back **of** your **aircraft** / and we do have the
5	ble damage to any part **of** the **aircraft** / So we're slowly / e
6	nybody in the vicinity **of** our **aircraft** who can have a look a
7	o / I have the captain **of** the **aircraft** / he'd like to go out
8	u saw somebody get out **of** the **aircraft**? // Roger he's steppe

9	steps to the back door **of** the **aircraft** on your side? // Affi
10	ge to the uh left wing **of** the **aircraft** uh fire fighter earli
11	ke is it like uh parts **of** the **aircraft** or is it just parts o
12	ting off the left side **of** the **aircraft** main door // There ar
13	are on the right side **of** the **aircraft** / the thinking now is

Tabela 22: Linhas de concordância com a expressão *of * aircraft*.

Em sua maioria, as linhas de concordância acima (Tabela 22) demonstram partes ou posição da aeronave, uso já visto nas buscas com o colocado *TO BE*; em uma das ocorrências, *of* é parte de um quantificador (linha 2, com *a bunch of*, agrupamento que ocorre somente mais uma vez no *corpus*); em outra ocorrência, *of* participa de um grupo nominal – o comandante da aeronave (*I have the captain of the aircraft*) em uma conversa entre o ATCO e o bombeiro.

A última preposição para nossa análise é *in*, evidenciada tanto no levantamento de *clusters* quanto nos colocados. As linhas de concordância revelaram dois *clusters* diferentes: *the aircraft in question* (duas ocorrências) e *follow the aircraft in front of you* (quatro ocorrências). Além de estar abaixo do ponto de corte determinado para este projeto (cinco), a instrução *follow the aircraft in front of you* é rotineira, fugindo ao nosso escopo; *the aircraft in question* é um equivalente ao agrupamento lexical já visto, *emergency aircraft*, e muito mais recorrente.

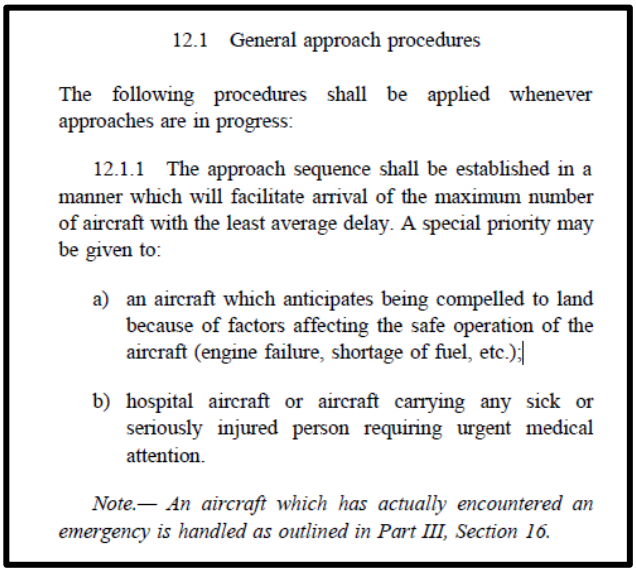
Os estudos dessas últimas preposições corroboraram as análises de *TO BE*. Reunindo essa última análise com as funções esboçadas anteriormente, temos os seguintes usos com a palavra *aircraft*:

- descrevendo posição da aeronave (*aircraft is on the turn to papa, emergency aircraft is on the runway, is the aircraft on final?)*
- descrevendo posição na aeronave (*the back of your aircraft, any part of the aircraft, the back door of the aircraft, the left wing of the aircraft*)
- descrevendo posição de algo no entorno da aeronave (*the tug in front of the aircraft, anybody in the vicinity of our aircraft, on the right side of the aircraft)*
- referindo-se ao tipo da aeronave (*aircraft type is a triple seven, what type of aircraft is United ten following?*)
- referindo-se à situação da aeronave (*the aircraft is under control, the aircraft is stable at this time*)

- indicando modalizações (*The aircraft appears to be normal, understand you are able to fly an aircraft, Prefer to land aircraft at destination*)
- explicitando finalidades (*we're just checking under the aircraft to make sure no fuel, a tow truck to tow the aircraft off back to the runway, have someone to circle the aircraft, people approaching your aircraft to inspect it*)

4.2.4 *Engine*

Problemas com o motor são relatados nos documentos de Fraseologia Aeronáutica como um dos poucos motivos relevantes para a não execução de procedimentos pré-estabelecidos, ou seja, como uma exceção. A título de ilustração, vemos abaixo, na Figura 20, um excerto do DOC 4444 (ICAO, 2008, p. 4-5), em que se explicam procedimentos de aproximação a aeroportos, considerando como prioridade aeronaves com problemas de motor ou de combustível, e aeronaves requerendo assistência médica a passageiros:



12.1 General approach procedures

The following procedures shall be applied whenever approaches are in progress:

12.1.1 The approach sequence shall be established in a manner which will facilitate arrival of the maximum number of aircraft with the least average delay. A special priority may be given to:

a) an aircraft which anticipates being compelled to land because of factors affecting the safe operation of the aircraft (engine failure, shortage of fuel, etc.);

b) hospital aircraft or aircraft carrying any sick or seriously injured person requiring urgent medical attention.

Note.— An aircraft which has actually encountered an emergency is handled as outlined in Part III, Section 16.

Figura 20: Excerto extraído de documentos de Fraseologia Aeronáutica sobre procedimentos de emergência para ATCOs.

Outra referência a motores aparece na Figura 17, que simula um diálogo em que o piloto vivencia um problema com o motor. Tais problemas também podem ser detectados no

corpus de estudo por meio da tela de *patterns* do *Concord* (Tabela 23), ao buscarmos linhas de concordância com a palavra *engine*:

N	L5	L4	L3	L2	L1	Centre	R1	R2	R3	R4	R5
1	WE	THE	THE	NUMBER	THE	**ENGINE**	**FAILURE**	WE	THE	ONE	THE
2	TO	WE	WE	THE	AN		IS	THE	TO	RIGHT	TO
3	THE	HEAVY	HAVE	A	ONE		AND	ON	SIX		WE
4	ENGINE	ONE	I	HAVE	TWO		UH	TO			ZERO
5	OF			ON	LEFT		**FIRE**	**OUT**			OF
6					RIGHT		SIX				ENGINE
7					**ROUGH**		BUT				
8							I				
9							COWLING				

Tabela 23: *Patterns* mais frequentes da palavra *engine*.

À direita, na posição R1, notamos os colocados mais frequentes de *engine* que descrevem problemas: *failure* e *fire*; há também a palavra *out* na posição R2, que pode se referir ao motor desligado (*engine is out*), ou ao sucesso de procedimento de extinção do fogo no motor (*fire out*). A investigação revelou que os *clusters engine out* ou *engine is out* referem-se realmente à falha de motor e não ao fogo; porém, sua frequência é baixa, com três ocorrências – as outras linhas evidenciam usos com verbos frasais (*check out, send out*), cuja baixa frequência tampouco justifica seu levantamento para esta pesquisa. Ainda sobre problemas no motor, à esquerda, na primeira posição (L1), temos *rough*, como em *rough engine* [motor com falha], também com três ocorrências. Logo, a preferência de uso está nos agrupamentos *engine failure*, que ocorre onze vezes, e *engine fire*, que ocorre sete.

Os determinantes – *the, an, one, two* – são altamente frequentes no lado direito do nódulo. A fim de clarificação, os motores são identificados como 'motor 1', ou esquerdo, e 'motor 2', ou direito, em aviões bimotores; justifica-se, assim, o fato de *one, two, left* e *right* serem frequentes na primeira posição à direita (R1), sendo a palavra *number* a mais recorrente na posição R2.

O verbo *TO HAVE* que, assim como o verbo *TO BE*, é altamente recorrente, levou-nos a outra investigação, especificamente pelas ocorrências de *engine* com conjugações de *TO HAVE*. As linhas geradas por essa busca foram as seguintes (Tabela 24):

N	Concordance
1	ty-one // We **had** a number one **engine** failure / Cactus uh / w
2	personnel **h ave** inspected the **engine** and the left main // th
3	left // the Delta **has** got an **engine** down / he may turn righ
4	from / we **have** shut down the **engine** and it stopped the fuel
5	e the equipment standing by / **engine** number four **has** a fire
6	one // two one / **have** a rough **engine** / I need to find a very
7	ghty-eight heavy / we **have** an **engine** failure / we're declare
8	runway heading / we **have** one **engine** failure // okay six one
9	avy / right now we **have** ah an **engine** failure / we **had** to shu
10	and uh pan pan pan we **have** an **engine** fire // Air Canada one
11	e got a / we'**ve** lost the left **engine** and what we're gonna do
12	er / after take-off we **had** an **engine** fire uh indication uh /
13	ta ten sixty-three **has had** an **engine** failure on the right en
14	**have** declared an emergency / **engine** failure on the right en
15	t five x-ray / we'**ve** lost one **engine** after take-off / reques
16	**have** the left we **have** the left **engine** running / the other eng
17	to Lauderdale? // We **have** an **engine** failure // American ten
18	know / maybe he's out for the **engine** / could **have** hit it but
19	rean shut down / they **have** an **engine** shut down // I'll tell

Tabela 24: Linhas de concordância com a palavra *engine* e variações de *have*.

Atentamo-nos ao uso do verbo *TO HAVE* com o sentido de posse e obtenção, e também ao uso do mesmo verbo como auxiliar do tempo composto (*Perfect Tenses*), como nas linhas 2, 5, 12, 14, 15, 16 e 19. Tal uso não serve somente para a descrição do problema com o motor, mas também para as combinações com a palavra *emergency* e com ações decorrentes do problema. A partir dessa constatação, retomamos a lista de concordâncias inicialmente gerada com a palavra *engine* e buscamos a lista de colocados. Nessa lista, observamos duas conjunções utilizadas no entorno da palavra *engine*: *and* (dez ocorrências) e *but* (três ocorrências). Apesar de não serem significativos, decidimos investigá-los, pois há indícios de uma possível subordinação ou coordenação aliada à presença dessas conjunções.

Essa percepção já havia surgido nas análises do entorno da palavra *emergency*. Uma forma de analisar as orações coordenadas e subordinadas é por meio de conjunções, segundo Raso e Mittman (2012, p. 206). Porém, a ausência de conjunções no *corpus* de estudo

dificultou a investigação. Já segundo McCarthy (1998), na investigação de *corpora* orais, é comum percebermos

> [...] a ausência de frases bem formadas, com orações principais e subordinadas. Ao contrário, é comum encontrarmos apenas locuções, orações incompletas, orações que se assemelham a orações subordinadas, mas que não parecem estar conectadas a alguma oração principal[54] (McCARTHY, 1998, p. 79, 80).

Ao analisarmos as linhas de concordância com a palavra *engine* em contexto ampliado, observamos que as elocuções poderiam levar a ideias de coordenação (por meio de orações adversativas) e subordinação (orações causais e consecutivas) na linguagem escrita. Porém, como afirmado por McCarthy em citação anterior, na maior parte das ocorrências o uso é direto, sem conjunções, ou com *and*, e, mais raramente, com *but*, *so*, e *due to*. Assim como preconizado pela OACI, observamos que, nas elocuções de emergência (ver Figura 17), anuncia-se o problema com o motor, a intenção do piloto em comando e, às vezes, o motivo pelo qual o motor falhou, sem qualquer conexão entre as elocuções por meio de conjunções.

A seguir, elocuções selecionadas para fim de exemplificação:

- sem conjunções, mas com possibilidade de coordenação ou subordinação:

1. *Center / we got an engine fire / right engine / will call you back //*
2. *Mayday mayday mayday / Alitalia six one golf / maintaining runway heading / we have one engine failure //*
3. *DHL one eighty-eight heavy / we have an engine failure / we're declaring an emergency at this time / ah // if it's ok with you / we'd like vectors back around to the appropriate runway for landing //*
4. *Southwest ah twenty six declaring an emergency / we need immediate back to the airport / we got an engine fire //*

- com conjunções coordenativas e subordinativas:

[54] *"[...] the absence of well-formed 'sentences' with main and subordinate clauses. Instead we often find turns that are just phrases, incomplete clauses, clauses that look like subordinate clauses but which seem not to be attached to any main clause etc."*

5. *left engine is still running / right engine is secure / **and** we´ll stand by for an inspection from the ground crew //*

6. *Eh / extremely rough engine // I'm going to try to hang it around // I may try to head back to Weiser or something / **but due to** the rain here I won't be able to make it to Bush //*

7. *and just to confirm we did declare an emergency **and** have the equipment standing by / engine number four has a fire indication **but** it appears to be out //*

8. *Delta ten sixty-three has had an engine failure on the right engine / declaring an emergency **due to** a bird strike //*

Tais exemplos corroboram a afirmação de McCarthy sobre frases incompletas, ou ausência de um conector e uma oração principal. É o caso do primeiro exemplo, onde se observa a ausência da conjunção. Nesse caso, a compreensão depende de um conhecimento extralinguístico por parte do interlocutor. Assim como também visto nas análises sobre *emergency*, o piloto que emitiu a mensagem parece esperar que o ATCO divida com ele o conhecimento de que a presença de 'fogo no motor' exige certos procedimentos que incluem o pouso imediato do avião – por isso, sua única informação sobre intenções foi '*will call you back*'. Pode, também, significar que o piloto está extremamente ocupado com seus afazeres e não consegue fornecer informações mais precisas naquele momento. O mesmo ocorre com os outros exemplos, em que as informações necessárias ao ATCO estão explicitadas, porém sem uso de conjunções que marquem a ligação de ideias.

Ainda sobre o uso de conjunções na comunicação radiotelefônica, temos o seguinte caso previsto no DOC 9432 (ICAO, 2007, p. 9-5), demonstrado na Figura 21:

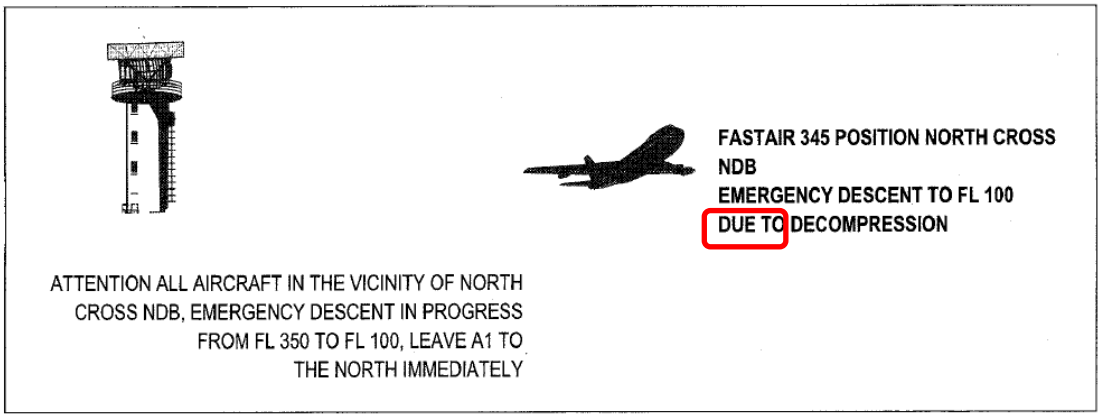


Figura 21: Exemplo de uma declaração de emergência com uso de conjunção, nesse caso, *due to*.

No exemplo da Fraseologia Aeronáutica (Figura 21), há o uso de *due to* (devido a) como justificativa a uma ação, assim como já visto em linhas extraídas do *corpus* de estudo. Buscamos, então, uma conjunção sinônima – a saber, *because of* – na lista de palavras, para verificar se a função representada por *due to* também se concretiza linguisticamente por meio de outra conjunção. A baixa frequência dessas palavras (*because* e *due* possuem, cada uma, dez ocorrências) motivou-nos a uma busca por conjunções, na lista de palavras (Gráfico 2):

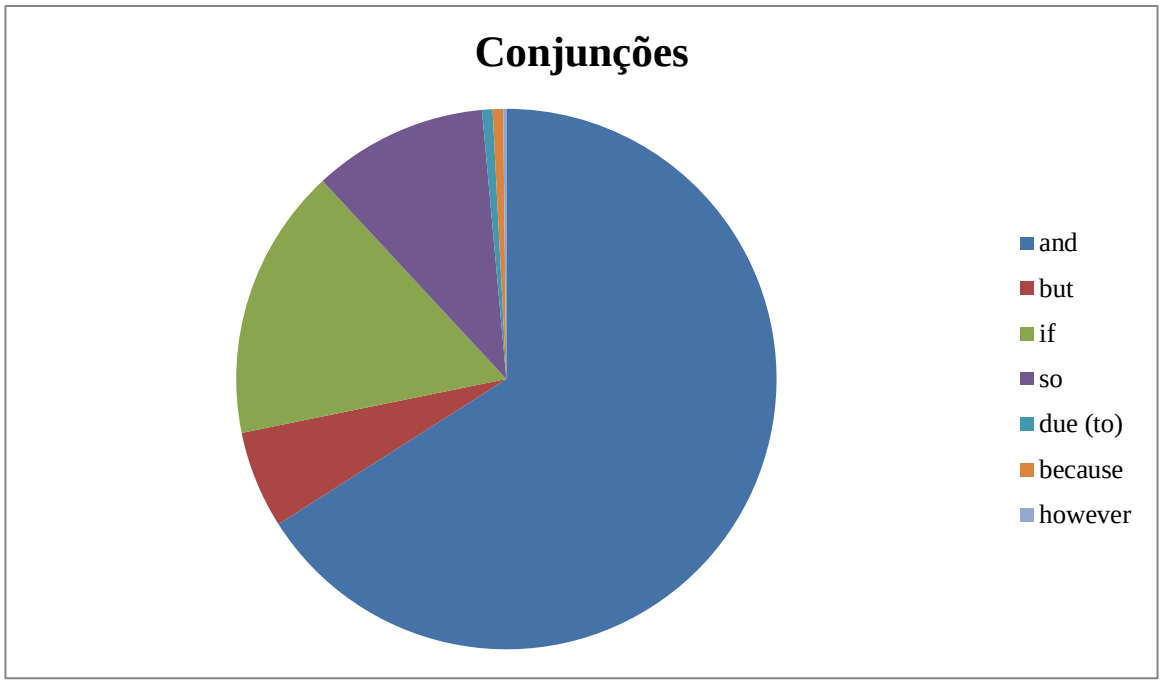


Gráfico 2: O uso de conjunções no *corpus* de estudo.

As conjunções utilizadas foram sete, sendo as coordenativas as mais frequentes (*and* com 1053 ocorrências, *but*, com 93 e *however*, com três ocorrências). As subordinativas são *if* (260 ocorrências), *so* (com 167), *due* e *because* (cada uma com dez ocorrências). Utilizamos a frequência 3 como ponto de corte, ou seja, a palavra precisa ocorrer pelo menos três vezes no *corpus* para ser incluída nesta análise. Esse fato dificultou a busca de orações coordenadas e subordinadas por meio de conjunções, como sugerida por Raso e Mittman (2012). Podemos considerar que a necessidade da comunicação e a rapidez intrínseca ao momento tornam as

orações simples, reduzindo-as a elocuções simples, encadeadas no mesmo enunciado. Além disso, a dependência do contexto faz com que pilotos e ATCOs compartilhem os passos que a mensagem deve seguir, como no exemplo a seguir:

Pilot: *Engine fire / Air Canada one zero one //*
ATCO: *Air Canada one zero one / what are your intentions? //*

Nesse exemplo, o piloto, extremamente sucinto, emitiu a mensagem somente com a informação sobre o problema. Na expectativa dos próximos passos, visto que não foram explicitados pelo piloto, o ATCO o interpela, a fim de se preparar para os procedimentos. Vimos, assim, que a simplicidade das comunicações radiotelefônicas se concretiza, também, na brevidade dos enunciados.

4.1.5 *Fuel*

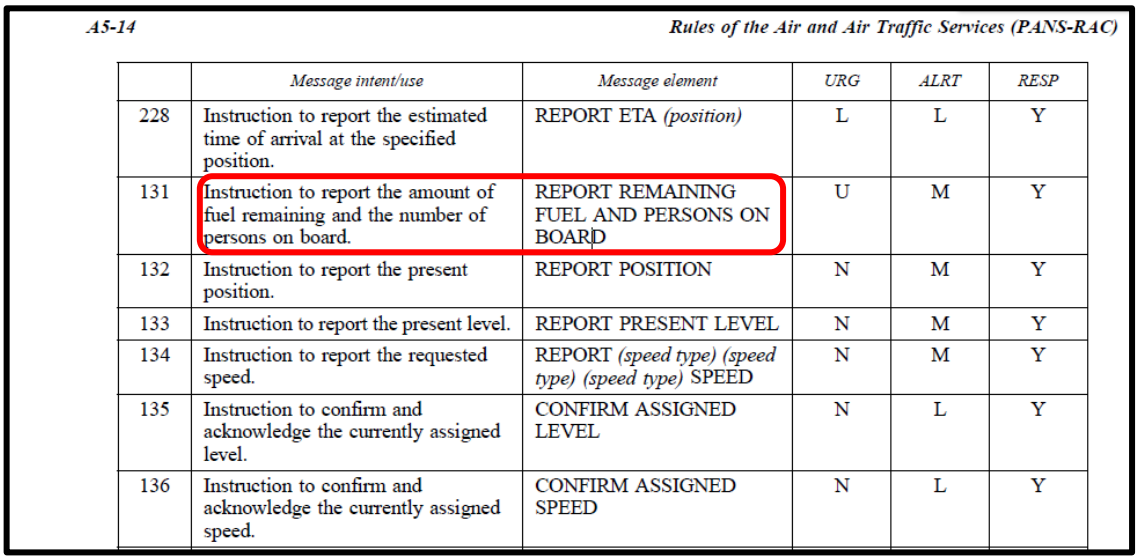
Fuel (combustível) ocorre no *corpus* de duas maneiras distintas: a primeira, mais recorrente, nas elocuções com informações sobre a quantidade de combustível que a aeronave carrega no momento em que o evento anormal acontece; a outra, sobre problemas com o combustível. Elaboraremos mais sobre cada análise a seguir.

No início da investigação da palavra *fuel*, ao gerarmos as 74 linhas de concordância, buscamos os agrupamentos lexicais (Tabela 25):

N	Cluster	Freq.
1	SOULS ON BOARD	11
2	ON BOARD AND	11
3	BOARD AND FUEL	8
4	POUNDS OF FUEL	6
5	AND SOULS ON	6
6	FUEL ON BOARD	5

Tabela 25: *Clusters* com a palavra *fuel*.

A repetição da palavra *fuel* e da expressão *on board* remeteu-nos à expressão *report remaining fuel and persons on board*. Essa expressão aparece no Manual de Regras do Espaço Aéreo, ou DOC 4444, também da OACI (ICAO, 2008, p. A5-14), na Figura 22:



A5-14 *Rules of the Air and Air Traffic Services (PANS-RAC)*

	Message intent/use	*Message element*	*URG*	*ALRT*	*RESP*
228	Instruction to report the estimated time of arrival at the specified position.	REPORT ETA *(position)*	L	L	Y
131	Instruction to report the amount of fuel remaining and the number of persons on board.	REPORT REMAINING FUEL AND PERSONS ON BOARD	U	M	Y
132	Instruction to report the present position.	REPORT POSITION	N	M	Y
133	Instruction to report the present level.	REPORT PRESENT LEVEL	N	M	Y
134	Instruction to report the requested speed.	REPORT *(speed type) (speed type) (speed type)* SPEED	N	M	Y
135	Instruction to confirm and acknowledge the currently assigned level.	CONFIRM ASSIGNED LEVEL	N	L	Y
136	Instruction to confirm and acknowledge the currently assigned speed.	CONFIRM ASSIGNED SPEED	N	L	Y

Figura 22: Ilustração extraída do DOC 9432, p. 11-2.

O ATCO solicita à aeronave que informe o número de passageiros e a quantidade de combustível para a preparação da equipe que assistirá à emergência, quando em solo. Para compararmos a forma como a OACI retrata a transmissão com o *corpus* de estudo, selecionamos os colocados *board* e *on*, sequencialmente, e obtivemos as seguintes linhas (Tabela 26):

N	Concordance
1	n board and souls on board // **Fuel** on board is three zero de
2	e / say ah souls on board and **fuel** remaining // One fifty-tw
3	to have the souls on board / **fuel** remaining in minutes / an
4	ou can uh people on board and **fuel** remaining I guess // stan
5	board // Sir / I'm gonna need **fuel** on board and souls on boa
6	United two twenty-four / say **fuel** and souls on board // Uni
7	ousand four hundred on the ah **fuel** // Eleven thousand four h
8	e / one eighty / we're low on **fuel** / we're just gonna make a

9	the aircraft to make sure no **fuel** is on your brakes // Japa
10	wenty-six / when able request **fuel** and souls on board // Giv
11	d / advise souls on board and **fuel** remaining // DHL one eigh
12	eight zero souls on board and **fuel** about two and a half hour
13	ne fifty-six souls on board / **fuel** is eight seven six zero p
14	t uh three souls on board and **fuel** is uh one six decimal thr
15	g strike / souls on board and **fuel** in pounds // Stand by //
16	ight thousand two / roger and **fuel** on board what was what wa
17	and I need souls on board and **fuel** remaining // Oh we have t
18	ada Four eleven // Say the uh **fuel** on board and souls on boa
19	many passengers on board and **fuel** available? // passengers
20	y-five thirty-five // Okay // **Fuel** on board and souls on boa
21	aiting you / and you'll be uh **fuel** dumping with them and uh
22	personnel on board as well as uh **fuel** in pounds // Okay / right
23	r uh forty thousand pounds of **fuel** // Forty thousand forty t
24	three and a half hours of uh **fuel** and eighty souls on board

Tabela 26: Linhas de concordância com a palavra *fuel* e *on*.

Não há nenhuma instância como a descrita pela OACI, que utiliza as palavras *report* e *persons*. Ressaltamos o fato de a palavra *souls* (33 ocorrências), comum nos agrupamentos com *fuel*, ser muito mais comum no *corpus* do que *persons* (duas ocorrências); no entanto, *people* ocorre vinte vezes. Nas linhas da Tabela 26, observamos a expressão *fuel on board* e suas variações (*fuel in pounds*, *fuel remaining*) como elocuções simples, sem outros colocados, ou precedidas pelos verbos *say*, *need*, *request* e *advise*. Tal fato, já registrado nas análises com *emergency*, levou-nos à indagação de como as frases interrogativas são elaboradas. Para identificarmos sua presença e, caso existam, investigarmos como são formadas, selecionamos os seguintes exemplos:

1. *can you **say** the personnel on board as well as uh fuel in pounds*
2. *and fuel on board what was **what was that again**?*
3. ***do you know** how many passengers on board and fuel available?*
4. *hree Juliet hotel **how is your fuel situation**?*
5. ***do you have** the souls and fuel yet?*
6. ***can you give me** the fuel in pounds by any chance?*

Confirmando a investigação feita no entorno da palavra *runway*, as frases interrogativas são normalmente indiretas, ou seja, dependem de uma oração principal, introdutória à pergunta. Das seis linhas selecionadas, três são interrogativas diretas (2, 4 e 5),

mas a linha 2 é um pedido de repetição da informação dada anteriormente, e a 5 é a repetição da pergunta já feita anteriormente (a palavra *yet* confirma essa afirmação). Observando todas as linhas em que há questionamento sobre a quantidade de combustível a bordo, percebe-se o uso de palavras relacionadas ao verbo 'dizer' (*say, state, report*), ou com agrupamentos lexicais introdutórios de sentenças (*Can you give me, I need to know, I´d like, I´m gonna need*), o último já visualizado na Tabela 17 deste capítulo. Essa forma de interrogar suaviza a elocução, algumas vezes emitida de forma direta (*souls and fuel on board*), mas, em sua maior parte, enunciada com um verbo ou agrupamento lexical que introduz o pedido.

Além do uso da palavra *fuel* no questionamento sobre a quantidade de combustível, encontram-se outras combinações, que nos levam à segunda fase da análise, e que remetem a um problema de vazamento (*fuel leak*) ou a procedimentos relacionados ao esvaziamento de combustível em voo (*fuel dumping* e *fuel burning*).

Sobre *fuel dumping*, há uma menção no DOC 9432, apresentada na Figura 23:

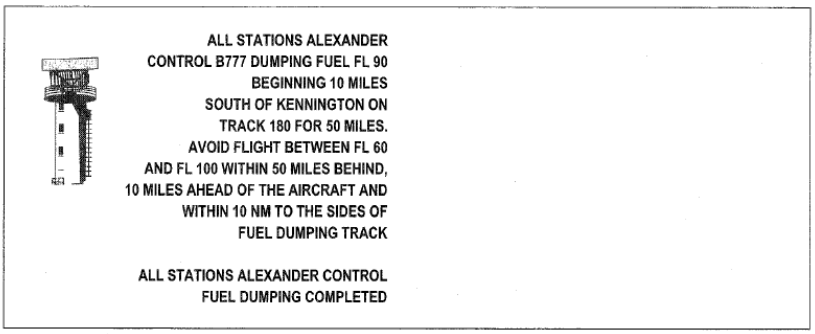


Figura 23: Diálogo sobre o procedimento de alijamento de combustível em progresso.

O procedimento de alijamento de combustível é executado quando a aeronave apresenta algum problema técnico e precisa eliminar combustível em uma área restrita para evitar um pouso com peso acima do limite – o que pode causar dano estrutural à aeronave. É necessário determinar o local do alijamento para que outras aeronaves não voem próximas à aeronave em emergência. Tal procedimento também consta na lista de *patterns* do Concord com a palavra *fuel*, na Tabela 27:

N	L5	L4	L3	L2	L1	Centre	R1	R2	R3	R4	**R5**
1	AND	SOULS	ON	BOARD	OF	FUEL	AND	BOARD	AND	AND	THE
2	UH	TO	THE	**POUNDS**	AND		**REMAINING**	YOU	YOUR	WE	I
3	YOU	YOU	THOUSAND	THE	THE		ON	UH	ONE	SOULS	ZERO
4	THREE	SIX	AND	**AMOUNT**	UH		**LEAK**	FROM	TO	SIX	WE
5	EIGHTY	A	TO	SOULS	YOUR		IN	IN	SEVEN		EIGHT
6	FOUR		I	ON	A		IS	**POUNDS**	**POUNDS**		ON
7					BOARD		**DUMPING**	SOULS	RE		**POUNDS**

Tabela 27: Tela de *patterns* do *Concord* com a palavra *fuel* ao centro.

Na Tabela 27, verificamos, em vermelho, palavras pertencentes ao agrupamento já estudado e, em verde, pronomes que possam servir de pista para novas investigações. Em negrito, observamos termos referentes à quantificação do combustível (*pounds* ou *amount of fuel*, ou *fuel remaining in pounds* ou *fuel in pounds)*, e os colocados *leak* (vazamento) e *dumping* (alijamento).

Gerando uma nova busca pelos pronomes realçados em verde na Tabela 27 no entorno de *fuel*, identifica-se a característica conversacional já vista nas investigações com *runway*. Mais uma vez, *you* e *we* (e variantes, como *your*, *our*) referem-se às aeronaves em situações anormais e são emitidas por ATCOs como *you*, e por pilotos como *we*. Os pronomes pessoais acompanham as ações, e observamos uma alta ocorrência de ações no presente contínuo (*we´re getting the final fuel here*, *we´re starting the fuel dump*, *we´re dumping fuel now*), no futuro (*you´ll be dumping fuel*, *we´re gonna need to dump fuel*) ou modalizadas (*there appears to be a fuel leak from your left wing*). Como tais estruturas já foram evidenciadas em outras análises, podemos afirmar que são padrões próprios do inglês comum na aviação.

Ao pensarmos na pragmática, de modo bastante geral podemos verificar algumas funções comunicativas típicas que acompanham o inglês comum nas situações anormais. Quando há algum evento não rotineiro, é necessário que o piloto reporte o problema e sua intenção; o problema é reportado com o uso de presente contínuo ou com *have* + problema; já a intenção é emitida pelo uso do futuro, tanto com *will* quanto com *going to / gonna*, ou ainda modalizada, representando a possibilidade das ações. A volição também ocorre nas ofertas ou pedidos relacionados à pista e à assistência. As interrogativas são, em sua maioria, reportadas

com o uso de *confirm, say, I need*. Já as orações complexas são emitidas com poucas conjunções e, quando ocorrem, são pouco variadas.

Considerando as análises apresentadas até aqui, podemos confirmar a definição da OACI de que o inglês para aviação deve ser simples e objetivo, como é a Fraseologia Aeronáutica. Pela razão forma/item, explicitada no início deste capítulo, verificamos que o *corpus* não mostra muita variação lexical; os padrões investigados demonstraram ser altamente recorrentes; as orações são simples e topicais, sem muita presença de subordinação (e, quando há, utilizam-se as mesmas conjunções).

Para respondermos a nossa segunda pergunta, comparamos o *corpus* de estudo ao Glossário de Estruturas Básicas do DOC 9835, ao qual nos referimos na justificativa desta pesquisa, para finalmente averiguarmos se tal elenco está realmente de acordo com a definição do inglês para aviação, segundo a OACI.

4.2 *The book is on the table* ou *the plane is on the runway*? Um estudo comparativo com o Glossário de Estruturas Básicas

Nesta seção, pretendemos comparar o *corpus* de estudo com o Glossário de Estruturas Básicas apresentado no Anexo B do DOC 9835, documento que rege o inglês para aviação, para, então, respondermos à segunda pergunta desta de pesquisa, a saber:

> Essas estruturas [léxico-gramaticais] correspondem ao Glossário de Estruturas Básicas e Complexas sugerido no Apêndice B do DOC 9835?

Como já visto na Introdução desta dissertação, há discrepâncias entre a definição de inglês para aviação no corpo do documento e o Glossário proposto em seus anexos. Sendo o inglês comum (*plain English*) aquele utilizado em situações anormais, para as quais não há Fraseologia Aeronáutica, mas com a obrigatoriedade de ser emitido de forma clara e objetiva, percebe-se uma desconformidade ao se explorar a lista de estruturas propostas.

Após o término das primeiras análises, realizadas de maneira a serem direcionadas – ou 'vetoradas' – pelo *corpus*, metodologia também conhecida como *corpus-driven*, foi

decidido confrontá-las com as estruturas elencadas no Anexo B do DOC 9835. Algumas estruturas não surgiram nas nossas análises iniciais, que foram norteadas pela alta frequência dos agrupamentos lexicais – o que não significa que estejam ausentes no *corpus* de estudo. Para essa verificação, utilizamos, aqui, outra abordagem de pesquisa, denominada 'baseada em *corpus*', ou *corpus-based*, cuja premissa é a investigação de hipóteses pré-concebidas.

Utilizamos a *WordList* lematizada para buscarmos as palavras individuais, como os verbos modais e os advérbios de frequência. Algumas palavras, porém, prolongaram nosso trabalho, já que necessitamos buscá-las nas linhas de concordância a fim de verificarmos suas funções: *like*, por exemplo, foi uma delas, pois suas 158 ocorrências no *corpus* poderiam ser como verbo, ou como marcador de discurso oral (o que somente ocorreu uma vez). Expressões fixas, por sua vez, foram todas buscadas no concordanciador; *you know*, por exemplo, foi um caso em que as linhas de concordância auxiliaram na observação de que tal agrupamento lexical é, realmente, um marcador de discurso, e não parte de uma pergunta.

Outras buscas foram dificultadas pelo fato de o *corpus* de estudo não ser *parsed*, ou seja, etiquetado gramaticalmente. Um *corpus* etiquetado assim traz consigo anotações referentes à classe gramatical de cada palavra. Em nosso caso, optamos por não fazê-lo pelo fato de nosso estudo ser direcionado pelo *corpus*, e não por nossas hipóteses. Assim, buscas por comparativos de adjetivos, por exemplo, foram conduzidas por meio da ferramenta *Concord*, com palavras de busca como **er* ou *the **est* para adjetivos curtos, seguidas de uma limpeza manual. Abaixo, um exemplo da busca de palavras referentes a pronomes interrogativos (*what*, *who*, *which*, *why*, *where*, *how*) (Figura 24), assim como sugeridas pelo Glossário, e a verificação manual para averiguar se são realmente frases interrogativas (Figura 25):

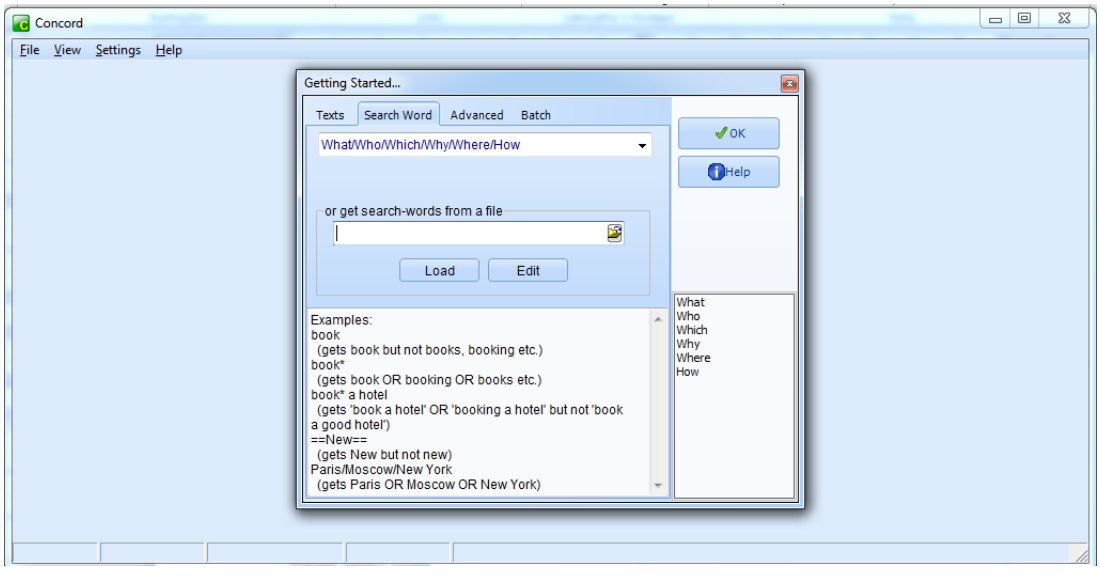


Figura 24: Busca por linhas de concordância com pronomes interrogativos.

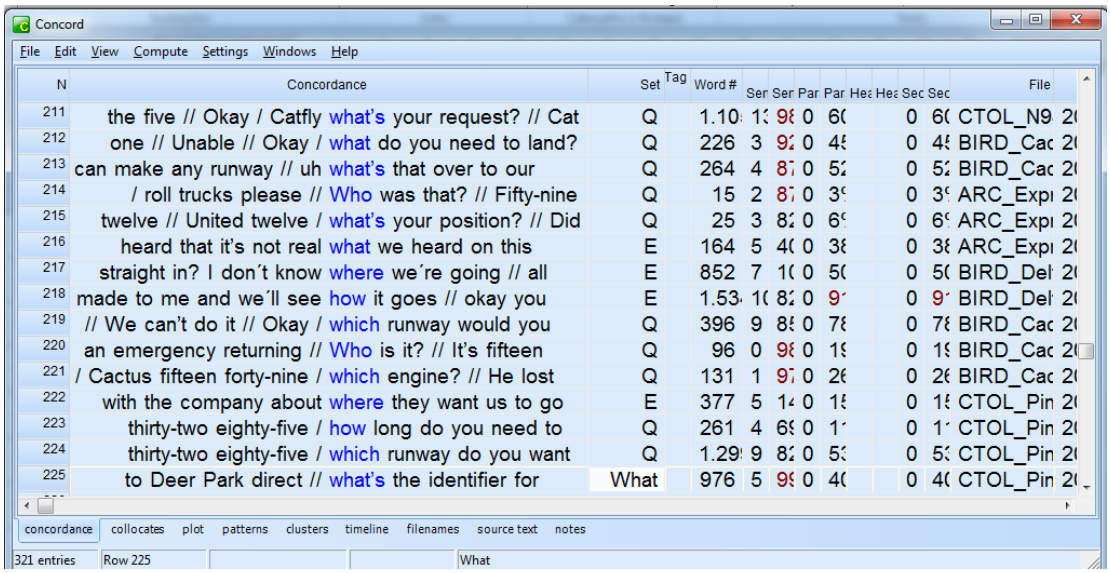


Figura 25: Seleção das linhas de concordância com pronomes em frases interrogativas.

Outra questão que causou dificuldade foi a investigação de verbos com dois objetos. Para resolvê-la, compilamos manualmente uma lista de verbos retirada de várias gramáticas, adicionamos suas conjugações de passado e a utilizamos na ferramenta *Concord*, que possibilita a adição de uma lista maior de busca por meio de um arquivo txt, assim como demonstrado na Figura 26:

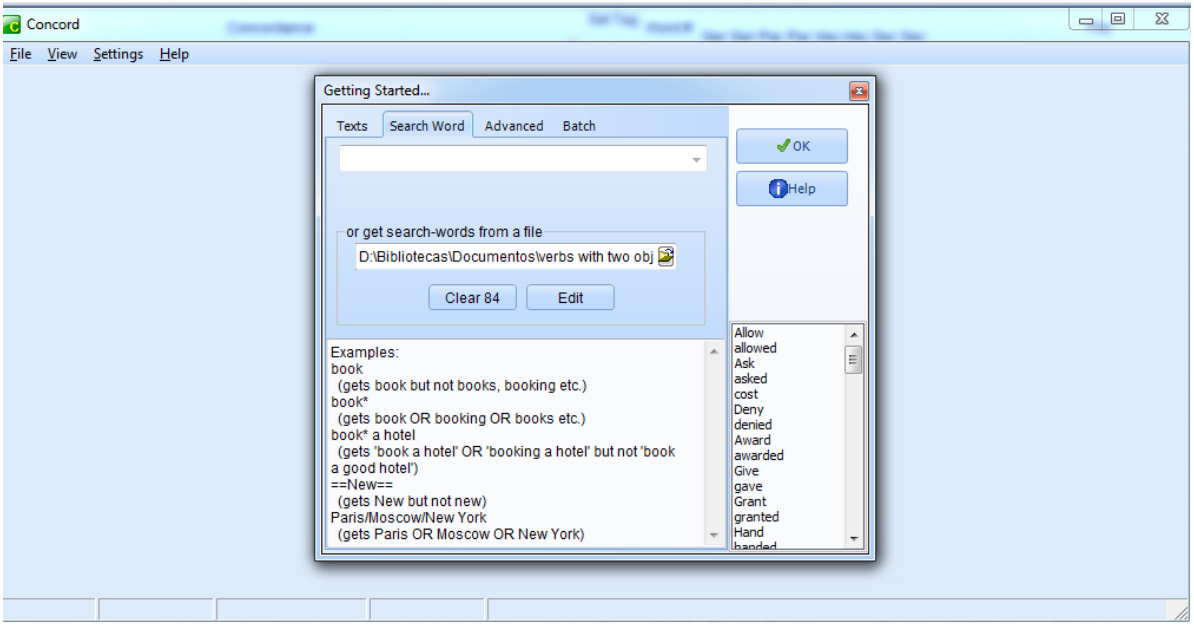


Figura 26: Tela do *Concord* com busca por várias palavras, inseridas por arquivo txt.

Em seguida, foi analisada cada linha para verificar se tais ocorrências eram realmente verbos com dois objetos.

A seguir, na Tabela 28, vemos o resultado da nossa investigação apresentado ao lado do elenco sugerido pela OACI. Destacamos, em nossa comparação, os itens recorrentes em verde, e os que não ocorreram, em rosa:

Estruturas do Glossário	*Corpus* de estudo
Articles	
	the – 1835
	a – 806
	an – 147
Adverbs of frequency:	
Always	Always – 3
Generally	Generally – sem ocorrência
Usually	Usually – sem ocorrência
Often	Often – 1
Sometimes	Sometimes – sem ocorrência
Seldom	Seldom – sem ocorrência
Never	Never – 2
Comparison of adjectives	Comparativos curtos – 63
	Comparativos longos – 3
	Superlativos longos – 2
	Superlativos curtos – 16
Discourse markers:	
Actually	Actually – 20
Basically	Basically – 2
Anyway	Anyway – 2
(and) yeah (more and more frequent)	Yeah – 141
Listen	Listen – 1

I mean	I mean – 3
Let’s see	Let´s see – 0
Let me see	Let me see – 3
Like	Like – 3
Oh	Oh – 91
Now	Now – 300
Okay	Okay – 442
So	So – 167
Well	Well – 50
You know	You know – 18
You see	You see – 14
You know what I mean	You know what I mean – 0
It is true	It is true – 0
Of course	Of course – 0
But	But – 93
Still	Still –57
By the way	By the way – 1
Besides	Besides – 1
Another thing is	Another thing is – 0
On top of that	On top of that – 0
Then	Then – 72
First(ly)	First – 1
Second(ly)	Second – 0
First of all	First of all – 0
In the first place	In the first place – 0
In the second place	In the second place – 0
Finally	Finally – 2
In the end	In the end – 0
In short	In short – 0
Modal verbs:	
Can	Can – 334
May	May – 29
Must	Must – 2

Have Got to	Have got to – 2
Should	Should – 31
Ought to	Ought to – 1
Would	Would – 133
Could	Could – 33
Might	Might – 21
Needn't	Needn´t – 0
Don't have to	Don´t have to – 1
Mustn't	Mustn´t – 0
Passive voice:	
Simple present	Simple present – 126
Simple past	Simple past – 7
Position of direct and indirect objects:	107
Question words for describing people and things and for requesting information:	
What	What – 81
Who	Who – 21
Which	Which – 11
Why	Why – 8
Where	Where – 40
How	How – 47
Relative pronouns:	
Who	Who – 9
Which	Which – 6
Whose	Whose – 1
Tenses	
Present simple	Present simple
I do	I do - *
Present continuous	Present continuous
I am doing	I am doing – 463
Past simple	Past simple
I did	I did - *
Past continuous	Past continuous

I was doing	I was doing – 36
Present perfect simple	Present perfect simple
I have done	I have done – 32
Present perfect continuous	Present perfect continuous
I have been doing	I have been doing – 2
Simple future tense	Simple future tense
Will	Will – 410
Future	Future
Going to	Going to – 267
There to be:	
Present	Present – 88
Past	Past – 2
Future	Future – 7

Tabela 28: Comparação entre o Glossário de Estruturas Básicas e suas ocorrências no *corpus* de estudo.

A ausência de etiquetagem gramatical no *corpus* impossibilitou o estudo do presente e passado simples; contudo, tais estruturas foram reveladas nas análises *corpus-driven* e são frequentes.

Resumindo, como já visto nas análises direcionadas pelo *corpus*, a frequência de determinantes é alta, opondo-se a estudos voltados à Fraseologia Aeronáutica que atestam a omissão dessa classe gramatical. Porém, os **determinantes** extraídos em nossas análises não eram somente os artigos, como proposto no Glossário, mas também os pronomes pessoais *my*, *your* e *our*, devido à característica conversacional do *corpus*. Os **advérbios de frequência** listados no Glossário são muito pouco utilizados, ocorrendo apenas seis vezes no total, em oposição aos advérbios de tempo *now*, *right now* e *immediately*, que demonstram o imediatismo das situações anormais. Já os **comparativos** curtos são mais recorrentes, e apontamos aqui um possível estudo futuro para comparações com manuais aeronáuticos, pois a maioria das instâncias pode se referir a frases típicas de *checklists* de emergência (*longest runway*, *nearest airport*).

Os **marcadores de discurso** listados no Glossário são vários, englobando, além dos marcadores de discurso propriamente ditos, diferentes categorias gramaticais tais como conjunções e advérbios de tempo. Dentre os marcadores propriamente ditos, poucos são altamente frequentes: *yeah*, *oh*, *okay*, *well*, *actually*, *you know* e *you see*; são marcadores que preenchem hesitações e identificam o aspecto interacional da linguagem. Os advérbios de tempo *now* e *then* e as conjunções *so* e *but* são recorrentes tanto no Glossário quanto no *corpus* de estudo. No entanto, a maioria dos marcadores apontados pelo Glossário ocorre com baixa frequência ou não ocorre na linguagem autêntica.

Os **verbos modais** também foram evidenciados em nossa investigação anterior, porém não na variedade apresentada pelo Glossário. Em nosso *corpus* de estudo, são recorrentes os modais *can* e *would*, além de *may*, *should*, *could* e *might*. Os outros não ocorreram ou não apresentaram frequência significativa.

Os **pronomes interrogativos** são recorrentes; vale lembrar que a maioria segue verbos de elocução, por uma provável suavização da linguagem (vide 4.1.5). Já os pronomes relativos recorreram dezesseis vezes, reforçando a questão da pouca instância de orações subordinadas vista nas primeiras análises (vide 4.1.5).

Os **tempos verbais** também foram investigados por meio dos pronomes pessoais e de alguns verbos recorrentes (como *have* e *declare*) no entorno dos itens lexicais. Assim, é possível confirmar a alta frequência de verbos no presente simples e no passado simples, já evidenciada nas análises anteriores; tal investigação não foi possível nesta segunda fase do estudo devido à falta de etiquetagem gramatical do *corpus*. A busca pelos outros tempos verbais sugeridos pelo Glossário foi viabilizada pela presença de auxiliares na forma afirmativa, como o verbo de ligação *TO BE* e o auxiliar *TO HAVE* (e suas conjugações), junto à forma *–ing*, dos tempos contínuos e *–ed*, dos tempos perfeitos e forma passiva. A alta ocorrência do presente contínuo e do verbo existencial *THERE TO BE*, principalmente no presente, já havia sido identificada na primeira parte do nosso estudo. O mesmo ocorre com a alta instância de elocuções no futuro. Vale ressaltar que os tempos verbais do futuro são mais comuns que os do passado.

Como se pode notar, a discrepância que já era evidenciada na sala de aula entre o que acontece e o que é ensinado foi confirmada. Relembramos que esse Glossário é utilizado para a confecção de materiais didáticos voltados ao ensino de inglês para aviação. Ao

contrapormos o Glossário com o *corpus*, totalmente extraído de situações reais, identificamos pouca similaridade entre os dois, o que faz com o que o aluno não veja correspondência entre o real e o ensinado. Face ao exposto, acreditamos que o ensino do inglês para aviação obteria melhores resultados se privilegiasse os elementos que foram mais recorrentes no nosso *corpus*.

CONSIDERAÇÕES FINAIS

Esta pesquisa teve por objetivo a descrição do núcleo de língua compartilhado entre dois profissionais – ATCOs e pilotos – em comunicações radiotelefônicas, quando em situações anormais. Os materiais didáticos disponíveis no mercado editorial seguem parâmetros estabelecidos pelo Manual de Implementação de Proficiência Linguística da OACI, conhecido como DOC 9835, e os currículos programáticos são norteados especificamente pelo Anexo B do manual, por meio de um elenco de estruturas gramaticais denominado Glossário de Estruturas Básicas e Complexas. Entretanto, verificou-se que esse material não contempla o contexto aeronáutico em situações anormais, gerando o descompasso entre o que é ensinado e o que acontece na vida real. Para atingirmos nosso objetivo, guiamo-nos por duas perguntas: (i) quais são os padrões léxico-gramaticais mais recorrentes do inglês para aviação e, (ii) como se comparam àqueles preconizados pelo Glossário.

A compilação do *corpus* que propiciou essa investigação foi laboriosa. Os critérios para a arquitetura do *corpus* foram sendo definidos à medida que a compilação era realizada, sendo submetido a ensaios constantes para a verificação de dados, tais como a riqueza lexical. Por fim, elegemos como norteador do balanceamento do *corpus* um documento de padronização de eventos intitulado Taxonomia de Ocorrências – um manual elaborado com o intuito de uniformizar relatórios redigidos no mundo inteiro. Essa taxonomia possibilitou-nos selecionar áudios pautados pelo tipo de ocorrência, desconsiderando outros critérios como duração do áudio, ou quantidade de palavras de cada transcrição. Assim, o *corpus* é representativo da língua em estudo e foi balanceado pelo tipo de ocorrência, com três ou quatro áudios para cada uma das 30 categorias selecionadas a partir da Taxonomia de Ocorrências.

A transcrição dos áudios foi feita com base na TLA (Teoria da Língua em Ato), que descreve a língua falada como enunciados pragmáticos, diferente da visão sintagmática (sujeito + verbo + complemento) da língua escrita. Para a TLA, os enunciados são divididos em unidades de referência, cuja definição propiciou um estudo pelas unidades de sentido e fraseologismos. Por fraseologismo, entendemos uma unidade em que a palavra investigada forma certos padrões com palavras no seu entorno. Nossa análise buscou, portanto, os

substantivos mais recorrentes – em nosso caso, *runway*, *aircraft*, *emergency*, *fuel* e *engine* –, que encabeçaram a investigação pelos seus entornos, ou por seus padrões. São as palavras de conteúdo mais frequentes do *corpus*, extraídas por meio de lematização da lista de palavras, seguida da remoção de palavras típicas da Fraseologia Aeronáutica, e comparadas a um *corpus* de referência – o MICASE. Durante as análises, as situações rotineiras foram descartadas, pois nosso objeto de estudo era o *plain English* utilizado nas situações anormais; para essa detecção, os documentos descritivos da Fraseologia Aeronáutica foram amplamente utilizados – e devidamente referenciados – neste trabalho.

Por meio dessas investigações observamos padrões recorrentes referentes a combinações com preposições e uso de determinantes, em contraposição ao estudo de Philps (1991), que analisa a omissão desses itens nas situações de rotina. O uso do verbo de ligação *TO BE* é altamente frequente, inclusive conjugado com *you* e *we*, pronomes que atestam a característica conversacional da comunicação radiotelefônica. Além disso, quando em sua forma infinitiva, o *to be* demonstra ocorrências com verbos modais, quando em referência a ações futuras, a possibilidades, a necessidades ou a volição.

O caráter objetivo exigido pelo momento das emergências é evidenciado por uma linguagem simples. Afirmamos isso tendo por base três caminhos percorridos nas análises, a saber: (i) a razão forma/item do *corpus*, com 3,68% de variação de riqueza lexical, que demonstra alta repetição de palavras; (ii) as orações emitidas em enunciados simples, ou em unidades de referência, sem conjunções ou com a repetição de apenas seis conjunções (*and*, *but*, *if*, *due to*, *because*, *so*); e (iii) as frases interrogativas, cuja maior frequência está nas formas indiretas utilizando verbos de elocução ou *clusters* relacionados ao ato de fornecer informação (*let me know*, *I need to know*, *can you give me*).

Os verbos frequentes são ou de suporte ou típicos da Fraseologia Aeronáutica, e conjugados principalmente no presente contínuo e no presente simples, ou, menos frequentemente, no passado e no presente perfeito; essas duas últimas conjugações se referem a ações recentes. O futuro é transmitido tanto por *will* quanto por *going to* (*gonna* detém a preferência).

A lista final de palavras de conteúdo possui apenas onze itens lexicais, o que nos leva a afirmar que são pouquíssimos os substantivos relevantes no *corpus*. Pelo cunho repetitivo do *corpus*, percebe-se uma preferência pelas mesmas palavras ou estruturas por alguns

possíveis motivos, dos quais salientamos dois, a saber, (i) a condição radiotelefônica das comunicações não contém sinais visuais (da linguagem corporal, por exemplo), aspecto que restringe a comunicação, além de acontecer em canais com baixa qualidade sonora; e (ii) a característica internacional da comunidade provoca uma adesão à acomodação (simplificação) da linguagem, que precisa ser eficaz e atingir um interlocutor que eventualmente tenha um nível de proficiência abaixo do previsto (ICAO, 2010).

A metodologia direcionada pelo *corpus* evidenciou resultados que contradizem hipóteses elencadas no Glossário de Estruturas Básicas do próprio documento da OACI – contraditório em si mesmo –, baseadas em concepções de ensino de línguas que não preveem o uso real do inglês utilizado por pilotos e ATCOs. Sem o uso da LC, que tem como um de seus princípios investigar a língua autêntica, tal estudo seria inviável. As ferramentas propiciaram a busca pelos padrões fraseológicos utilizando critérios estatísticos que atestam o caráter científico defendido por Sinclair (2004).

Por meio deste estudo, julgamos ter identificado quais padrões léxico-gramaticais devem ser priorizados no ensino do inglês para aviação. Nossa intenção, aqui, não é determinar o único conteúdo a ser trabalhado em sala de aula, mas sim apresentar uma descrição do *plain English* utilizado por ATCOs e pilotos em situações anormais, em que a Fraseologia Aeronáutica não é suficiente. Vale a ressalva de que a Fraseologia Aeronáutica deva ser a única linguagem utilizada nas situações de rotina, e, quando algo errado acontece, a linguagem é emitida de forma simples e objetiva, como a identificada nesta pesquisa. A importância do inglês para aviação não está somente na licença obtida pelos profissionais aprovados nas avaliações de proficiência linguística – está na segurança de voo. Afinal, "ATCOs e pilotos não se podem dar o luxo de atribuir significado às palavras; fazê-lo levaria a resultados catastróficos"[55] (PRINZO; BRITTON, 1993).

Sugerimos, por fim, alguns desdobramentos: o primeiro seria o de avançar os estudos iniciados por este projeto analisando os outros itens lexicais apresentados na lista final de substantivos; outro seria o de etiquetar o *corpus* gramaticalmente para a comparação com o Glossário de Estruturas Complexas do Anexo B do DOC 9835. Além disso, há outras áreas linguísticas sugeridas pela OACI a serem exploradas, como pronúncia, fluência, e interação, uma vez que este trabalho se ocupou somente da léxico-gramática. Acreditamos, também, que

[55] "*Controllers and pilots do not have the luxury of assigning meaning to words. To do so could have catastrophic results*".

a transcrição feita segundo critérios da TLA possibilitará estudos pragmáticos e prosódicos da língua oral. O *corpus*, pela sua peculiaridade e riqueza de conteúdo, pode beneficiar pesquisas relacionadas ao reconhecimento de fala nos contextos aeronáuticos.

REFERÊNCIAS BIBLIOGRÁFICAS[56]

ADOLPHS, S.; KNIGHT, D. Building a spoken corpus: What are the basics? In O´KEEFE, A.; McCARTHY, M. **The Routledge Handbook of *Corpus* Linguistics**. London: Routledge, 2010.

ATKINS, S.; CLEAR, J.; OSTLER, N. Corpus Designing Criteria. **Literary and Linguistic Computing**, Oxford, v. 7, n. 1, p. 1-16, 1992.

AUSTIN, J. **How to Do Things with Words.** Oxford: Oxford University Press, 1962.

BAILEY, K. **Washback in language testing**. New Jersey: Educational Testing Service, 1999.

BARSHI, I; FARRIS, C. **Misunderstandings in ATC communication**: language, cognition and experimental methodology. Burlington: Ashgate, 2013.

BERBER-SARDINHA, T. **Lingüística de *Corpus***. São Paulo: Manole, 2004.

BIBER, D. **University language**: a corpus-based study on spoken and written registers. Amsterdam/Philadelphia: John Benjamins, 2006.

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**. Harlow: Pearson Education, 1999.

BLANCHE-BENVENISTE, C. Linguistic Analysis of Spoken Language: The Case of French Language. In KAWAGUCHI, K.; ZAIMA, S.; TAKAGAKI, T. **Spoken Language Corpus and Linguistic Informatics**. Amsterdam/Philadelphia: John Benjamins, 2006.

BOCORNY, A. **Descrição das unidades especializadas poliléxicas nominais no âmbito da aviação**: subsídios para o ensino de inglês para fins específicos (ESP). 2008. 230 f. Tese (Doutorado em Estudos Linguísticos)–Faculdade de Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2008. Disponível em: <http://www.lume.ufrgs.br/bitstream/handle/10183/15548/000678237.pdf?sequence=1>. Acesso em: 01 dez. 2012.

[56] De acordo com a Associação Brasileira de Normas Técnicas. NBR 6023.

_____. Panorama dos estudos sobre a linguagem da aviação. **Revista brasileira de linguística aplicada**, Belo Horizonte, v. 11, n. 4, p. 963-986, 2011.

BRAZIL, D. **A grammar of speech**. Oxford: Oxford University Press, 1995.

BRUMFIT, C. **Individual freedom in language teaching**. Oxford: Oxford University Press, 2001.

CARTER, R.; McCARTHY, M. **Cambridge Grammar of English**: A comprehensive Guide. Spoken and Written English Grammar and Usage. Cambridge: Cambridge University Press, 2006.

ČERMÁK, F. Spoken Corpora Design: their constitutive parameters. **International Journal of Corpus Linguistics**, Amsterdam, v. 14, n. 1, p. 113-123, 2009.

CONRAD, S.; BIBER, D. **Real Grammar**: a corpus-based approach to English. White Plains: Pearson Education, 2009.

COOKSON, G. Zagreb and Tenerife: Airline Accidents Involving Linguistic Factors. **Australian Review of Applied Linguistics** [online], v. 32, n. 3, p. 22.1-22.14, 2009. Disponível em: < http://www.nla.gov.au/openpublish/index.php/aral/article/viewFile/2029/2412>. Acesso em: 9 maio 2013.

CRESTI, E. *Corpus* di Italiano parlato. v. 1. Firenze: Accademia dela Crusca, 2000. 2 v.
CRESTI, E.; MONEGLIA, M. (eds) **C-ORAL-ROM**: Integrated Reference Corpora for Spoken Romance Languages. Amsterdam/Philadelphia: John Benjamins, 2005.

CRYSTAL, D. **English as a global language**. 2 ed. Cambridge: Cambridge University Press, 2003.

CUSHING, S. **Fatal words:** communication clashes and aircraft crashes. Chicago: University of Chicago Press, 1994.

_____. It's not (only) what you say, it's how you say it. **Flight Safety Digest**, Virginia, v. 14, n. 7, p. 1-10, 1995. Disponível em: < http://flightsafety.org/fsd/fsd_jul95.pdf>. Acesso em: 08 dez. 2014.

DRURY, C.; MA, J. Report on literature of aviation language errors and analysis of error databases. **Federal Aviation Administration** (Relatório FAA #2002-G-025). 2002. 35p.

FULCHER, G.; DAVIDSON, F. **Language Testing and Assessment**: an advanced resource book. London and New York: Routledge, 2006.

GAVIOLI, L. **Exploring corpora for ESP learning**. Amsterdam/Philadelphia: John Benjamins, 2005.

GRADDOL, D. **The future of English**. London: British Council, 1997. Disponível em: < http://www.britishcouncil.org/learning-elt-future.pdf>. Acesso em: 03 dez. 2012.

GRADDOL, D. **English Next**. London: British Council, 2006. Disponível em: < http://www.britishcouncil.org/learning-research-english-next.pdf>. Acesso em: 03 dez. 2012.

HOEY, M. What´s in a word? **English Teaching Professional** [online], issue 27, April 2003. Disponível em: < http://www.onestopenglish.com/support/methodology/grammar-vocabulary-and-skills/whats-in-a-word/155130.article>. Acesso em: 08 dez 2012.

HUNSTON, S. **Corpora in Applied Linguistics**. Cambridge: Cambridge University Press, 2002.

INTERNATIONAL Civil Aviation Organization. **KLM, B-747, PH-BUF and Pan Am, B-747, N736PA, collision at Tenerife Airport, Spain, on 27 March, 1977**. Montreal, 1978. 48p. (Relatório Técnico Circular 153-AN/56).

_____. **Annex 13 to the Convention on International Civil Aviation:** Accident reports. Montreal, 1994. Disponível em < http://www.iprr.org/manuals/Annex13.html>. Acesso em: 08 nov. 2014.

_____. **Annex 10 to the Convention on International Civil Aviation**: Aeronautical Telecommunications – Volume II. Montreal, 2001.

_____. **Doc 9835**: Manual of Implementation of Language Proficiency. Montreal, 2004.

_____. **Annex 1 to the Convention on International Civil Aviation:** Personnel Licensing. Montreal, 2006.

_____. Aviation Occurrence Categories: Definitions and Usage Notes. Montreal: ICAO, 2006.

_____. **Doc 9432**: Manual of Radiotelephony. Montreal, 2007.

_____. **Doc 4444**: Air Traffic Management. Montreal, 2008.

_____. **Doc 9835**: Manual of Implementation of Language Proficiency. Montreal, 2010.

KACHRU, B. (ed.). **The Other Tongue**. Urbana /Chicago: University of Illinois Press, 1992.

LEECH, G. Grammars of Spoken English: New Outcomes of Corpus-Oriented Research. **Language Learning**, Michigan, v. 50, n. 4, p. 675–724, 2000.

MATHEWS, E. New provisions for English language proficiency are expected to improve aviation safety. **ICAO Journal**, v. 59, n. 1, p. 4–6, 2004.

McCARTHY, M. **Spoken Languages and applied linguistics**. Cambridge: Cambridge University Press, 1998.

McCARTHY, M.; CARTER, R. Ten criteria for a spoken grammar. In HINKEL, E.; FOTOS, S. (eds). **New Perspectives on Grammar Teaching in Second Language Classrooms**. Mahwah, NJ: Lawrence Erlbaum Associates, 2001. p. 51-75.

McCARTHY, M.; McCARTEN, J.; SANDIFORD, H. **Touchstone 1**. Cambridge: Cambridge University Press, 2005.

McCARTHY, M.; O´KEEFE, A. Corpora and spoken language. In LÜDELING, A.; KYTÖ, M. (eds.) **Corpus Linguistics**: An International Handbook. v. 2. Berlin/New York: Walter de Gruyter, 2009. 2 v.

McENERY, T.; HARDIE, A. **Corpus Linguistics**: Method, Theory and Practice. Cambridge: Cambridge University Press, 2012.

MELLO, H. Os *Corpora* Orais e o C-ORAL-BRASIL. In RASO, T.; MELLO, H. (eds.) **O C-ORAL-BRASIL I**: *Corpus* de referência do português brasileiro falado informal. Belo

Horizonte: Editora UFMG, 2012.

MONEGLIA, M.; CRESTI, E. C-ORAL-ROM: Prosodic Boundaries for Spontaneous Speech Analysis. In KAWAGUCHI, Y; ZAIMA, S.; TAKAGAKI, T. (eds) **Spoken Language Corpus and Linguistic Informatics**. Amsterdam /Philadelphia: John Benjamins, 2006. p. 89-114.

O´KEEFE, A.; McCARTHY, M.; CARTER, R. **From Corpus to Classroom**: language use and language teaching. Cambridge: Cambridge University Press, 2007.

OLIVEIRA, E. **Da Torre de Babel à torre de controle**: Desmistificando a linguagem dos céus. Um estudo descritivo da língua franca utilizada na comunicação piloto-controlador. 2007. 146f. Dissertação (Mestrado em Educação)–Faculdade de Educação, Universidade de São Paulo, São Paulo, 2007. Disponível em: <http://www.teses.usp.br/teses/disponiveis/48/48134/tde-05102007-111948/pt-br.php>. Acesso em: 03 dez. 2012.

PEREIRA, C. As várias faces do livro didático de língua estrangeira. In SARMENTO, S.; MÜLLER, V. (org.) **O ensino do inglês como língua estrangeira**: estudos e reflexões. Porto Alegre: APIRS, 2004. p. 195-210.

PHILPS, D. Linguistic Security in the Syntactic Structures of Air Traffic Control English. In **English World-Wide**, Amsterdam, v. 12, n. 1, p. 103-124, 1991.

PRINZO, O.; BRITTON, T. ATC/Pilot Voice Communications: A Survey of the Literature. **Federal Aviation Administration**. (NITS No. DOT/FAA/AM-93/20), 1993.

RAJAGOPALAN, K. Linguistics and the myth of nativity: comments on the controversy over ‘new/non-native Englishes’. **Journal of Pragmatics**, v. 27, n. 2, p. 225-231, 1997.

RASO, T.; MELLO, H. (org.). **C-ORAL-BRASIL I.** *Corpus* de referência do português brasileiro falado informal. Belo Horizonte: Editora UFMG, 2012.

RASO, T. O C-ORAL-BRASIL e a Teoria da Língua em Ato. In RASO, T.; MELLO, H. (eds.) **O C-ORAL-BRASIL I**: *Corpus* de referência do português brasileiro falado informal. Belo Horizonte: Editora UFMG, 2012.

RASO, T.; MITTMAN, M. As principais medidas da fala. In RASO, T.; MELLO, H. (eds.) **O C-ORAL-BRASIL I**: *Corpus* de referência do português brasileiro falado informal. Belo Horizonte: Editora UFMG, 2012.

RICHARDS, J. **Curriculum development in language teaching**. Cambridge: Cambridge University Press, 2001.

RODRIGUES, L.A.D. **Dos fios, das tramas e dos nós: a tessitura da rede de crenças, pressupostos e conhecimentos de professores de inglês que atuam no Ciclo I do Ensino Fundamental**. 2005. Tese (Doutorado em Educação)–Faculdade de Educação da Universidade de São Paulo, São Paulo, 2005.

SANTOS, A. Working closely with corpus: análise de colocações adverbiais em inglês para negócios. In VIANA, V.; TAGNIN, S. (eds). ***Corpora* no ensino de línguas estrangeiras**. São Paulo: HUB Editorial, 2011.

SARMENTO, S. **O uso dos verbos modais em manuais de aviação em inglês:** um estudo baseado em *corpus.* 2008. 262 f. Tese (Doutorado em Estudos Linguísticos)–Faculdade de Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2008. Disponível em: <http://www.lume.ufrgs.br/ bitstream/handle/10183/15568/000685529.pdf?sequence=1>. Acesso em: 01 dez. 2012.

SARMENTO, S. Linguística de *Corpus* e o desenvolvimento de material didático para inglês com propósitos específicos. In SARMENTO, S.; FREITAS, A. (orgs.) **O ensino do inglês como língua estrangeira**: estudos e reflexões II. Porto Alegre: EDIPUCRS, 2009. p. 259-290.

SCARAMUCCI, Matilde V. R. Efeito retroativo da avaliação no ensino/aprendizagem de línguas: o estado da arte. **Trabalhos em linguística aplicada** [online], v.43, n. 2, p. 203-226, 2004. Disponível em: < http://revistas.iel.unicamp.br/index.php/tla/article/view/2202/1708>. Acesso em: 02 nov. 2014.

SCOTT, M. **WordSmith Tools** version 6. Liverpool: Lexical Analysis Software, 2012.

SILVA, J. **Implicações culturais e didáticas do inglês como língua internacional**: o livro didático. 2012. 226f. Dissertação (Mestrado em Educação)–Faculdade de Educação, Universidade de São Paulo, 2012. Disponível em: <http://www.teses.usp.br/teses/disponiveis/48/48134/tde-19072012-150035/pt-br.php>. Acesso em: 12 nov. 2014.

SINCLAIR, J. **Corpus, concordance, collocation**: Describing English language. Oxford: Oxford University Press, 1991.

_____. A way with common words. In SINCLAIR, J.; HOEY, M.; FOX, G. (eds) **Out of corpora**: studies in honour of Stig Johansson. Amsterdam: Rodopi, 1999. p. 157-179

_____. **Trust the text**: language, corpus and discourse. London and New York: Routledge, 2004.

TAGNIN, S; TEIXEIRA, E. Lingüística de *Corpus* e Tradução Técnica: Relato da Montagem de um *Corpus* Multivarietal de Culinária. **TradTerm**, São Paulo, v. 10, p. 313-358, 2004.

TAGNIN, S. **O jeito que a gente diz**: combinações consagradas em inglês e português. Barueri: Disal, 2013.

TOGNINI-BONELLI, E. **Corpus Linguistics at Work.** Amsterdam: John Benjamins, 2001.

VIANA, V. Linguística de *Corpus*: conceitos, técnicas e análises. In VIANA, V.; TAGNIN, S. ***Corpora*** **no ensino de línguas estrangeiras**. São Paulo: HUB Editorial, 2010.

XAVIER, R.; SOUZA, D. O que os alunos pensam sobre o livro didático de inglês? **Trabalhos em Linguística Aplicada**, v. 47, n. 1, p. 65-89, jan./jun. 2008. Disponível em: < http://www.scielo.br/pdf/tla/v47n1/v47n1a04.pdf>. Acesso em: 14 nov. 2014.

ANEXO

ESCALA DE NÍVEIS DE PROFICIÊNCIA LINGUÍSTICA DA ICAO

Níveis Expert, Avançado e Operacional

Nível	**PRONÚNCIA**	**ESTRUTURA**	**VOCABULÁRIO**	**FLUÊNCIA**	**COMPREENSÃO**	**INTERAÇÕES**
Expert 6	Pronúncia, intensidade, ritmo e entonação, embora possivelmente influenciados pela língua materna ou variação regional, quase nunca interferem na facilidade de compreensão	Aspectos gramaticais básicos e complexos e padrões estruturais das frases são constantemente bem controlados	A riqueza e precisão do vocabulário são suficientes para comunicar-se efetivamente em uma grande variedade de tópicos familiares e desconhecidos. O vocabulário é idiomático, suficiente para expressar sutilezas e apropriado ao contexto	Capaz de falar por um longo período em ritmo natural e sem esforço. Varia o ritmo da fala para efeitos estilísticos, ex: para enfatizar um assunto. Usa marcadores de discurso e conectores apropriados espontaneamente.	A compreensão é constantemente precisa em quase todos os contextos e inclui o entendimento de sutilezas lingüísticas e culturais.	Interage com facilidade em quase todas as situações. É sensível à pistas verbais e não verbais, e as responde apropriadamente.
Avançado 5	Pronúncia, intensidade, ritmo e entonação, embora influenciados pela língua materna ou variação regional, raramente interferem na facilidade de compreensão	Aspectos gramaticais básicos e padrões estruturais das frases são constantemente bem controlados. Esforça-se para usar estruturas complexas mas com erros, o que algumas vezes interfere no significado	A riqueza e precisão do vocabulário são suficientes para comunicar-se efetivamente em tópicos comuns, concretos e relacionados ao trabalho. Parafraseia constantemente e com sucesso. O vocabulário é algumas vezes idiomático	Capaz de falar por um longo período com uma facilidade relativa em tópicos familiares, mas pode não variar o ritmo da fala para efeitos estilísticos. Pode usar marcadores de discurso e conectores apropriados.	A compreensão é precisa em tópicos comuns, concretos e relacionados ao trabalho e na maioria das vezes precisa quando o falante é confrontado com uma complicação lingüística ou situacional ou com eventos inesperados. É capaz de compreender uma gama de variedades do discurso (dialetos e/ou sotaques) ou registros	As respostas são imediatas, apropriadas e informativas. Gerencia a relação falante/ouvinte efetivamente
Operacional 4	Pronúncia, intensidade, ritmo e entonação são influenciados pela língua materna ou variação regional, mas apenas algumas vezes interferem na facilidade de compreensão.	Aspectos gramaticais básicos e padrões estruturais das frases são usados com criatividade e são geralmente bem controlados. Erros podem acontecer, particularmente em circunstâncias não usuais ou não esperadas, mas raramente interferem no significado	A riqueza e precisão do vocabulário são geralmente suficientes para comunicar-se efetivamente em tópicos comuns, concretos e relacionados ao trabalho. Pode geralmente parafrasear com sucesso quando faltar vocabulário em circunstâncias não usuais ou não esperadas.	Produz segmentos lingüísticos em cadência apropriada. Pode ocorrer perda de fluência ocasional na transição da fala ensaiada ou formatada para a interação espontânea, mas isto não impede a comunicação efetiva. Pode fazer uso limitado de marcadores de discurso ou conectores. Os marcadores conversacionais não	A compreensão é na maioria das vezes precisa em tópicos comuns, concretos e relacionados ao trabalho quando o sotaque ou variedade usada é suficientemente inteligível para uma comunidade internacional de usuários. Quando o falante é confrontado com uma complicação lingüística ou situacional ou com eventos inesperados, a compreensão pode ser mais vagarosa ou	As respostas são geralmente imediatas, apropriadas e informativas. Inicia e mantém trocas mesmo quando lidando com eventos inesperados. Sabe lidar adequadamente com mal entendidos aparentes, checando, confirmando ou esclarecendo

				comprometem a compreensão da mensagem.	requerer estratégias de clarificação.	
Pré-operacional 3	Pronúncia, intensidade, ritmo e entonação são influenciados pela língua materna ou variação regional e frequentemente interferem na facilidade de compreensão.	Aspectos gramaticais básicos e padrões estruturais das frases associados com situações previsíveis não são sempre bem controlados. Erros frequentemente interferem no significado.	A variação e exatidão do vocabulário são frequentemente suficientes para comunicar-se em tópicos comuns, concretos e relacionados ao trabalho, mas a variedade é limitada e a escolha de palavras frequentemente inapropriada. É muitas vezes incapaz de parafrasear com sucesso quando falta vocabulário.	Produz segmentos lingüísticos, mas a construção das frases e pausas são geralmente inapropriadas. Hesitações ou lentidão no processamento da língua pode impedir a comunicação efetiva. Os marcadores conversacionais algumas vezes comprometem a compreensão da mensagem.	A compreensão é freqüentemente precisa em tópicos comuns, concretos e relacionados ao trabalho quando o sotaque ou variedade usada é suficientemente inteligível para uma comunidade internacional de usuários. Pode falhar em entender uma complicação lingüística ou situacional ou eventos inesperados	As respostas são algumas vezes imediatas, apropriadas e informativas. Pode iniciar e manter trocas com relativa facilidade em tópicos familiares e em situações previsíveis. Geralmente inadequado quando lidando com eventos inesperados.
Elementar 2	Pronúncia, intensidade, ritmo e entonação são fortemente influenciados pela língua materna ou variação regional e geralmente interferem na facilidade de compreensão	Demonstra apenas controle limitado de poucas estruturas gramaticais e padrões estruturais das frases simples e memorizados.	Variedade de vocabulário limitada consistindo apenas de palavras isoladas e frases memorizadas	Pode produzir elocuções muito pequenas, isoladas e memorizadas com pausas freqüentes e um uso de marcadores conversacionais que desviam a atenção a fim de procurar expressões e articular palavras menos familiares.	A compreensão é limitada a frases isoladas e memorizadas quando são articuladas com cuidado e vagarosamente.	O tempo de resposta é lento, e freqüentemente inapropriado. Interação é limitada a trocas de rotina simples.
Pré-elementar 3	Desempenha em um nível abaixo do Nível Elementar.	Desempenha em um nível abaixo do Nível Elementar.	Desempenha em um nível abaixo do Nível Elementar.	Desempenha em um nível abaixo do Nível Elementar.	Desempenha em um nível abaixo do Nível Elementar.	Desempenha em um nível abaixo do Nível Elementar.

Disponível em: <http://www2.anac.gov.br/arquivos/escaladeniveisdaOACI.pdf>. Acesso em: 03 dez 2014.