

## Desenvolvendo o conteúdo programático de um curso de inglês para mecânicos de aeronaves com base em um *corpus DIY*<sup>1</sup>: um estudo de caso

Dilso Corrêa de ALMEIDA<sup>2</sup>

Malila Carvalho de Almeida PRADO<sup>3</sup>

**RESUMO:** *Cursos de Inglês Instrumental são geralmente desenvolvidos de acordo com as necessidades dos alunos. Essas necessidades são definidas com base em levantamentos junto aos aprendizes ou pela análise, por parte do professor/desenvolvedor do material, sobre o uso da linguagem (acadêmica, ocupacional ou vocacional) que os aprendizes precisarão desenvolver. Apresentamos, neste artigo, o processo de desenvolvimento do conteúdo programático de um curso de leitura de manuais de manutenção de aeronaves em inglês, tendo a Linguística de Corpus como balizadora da escolha dos aspectos léxico-semântico-gramaticais a serem trabalhados durante o curso.*

**PALAVRAS-CHAVE:** *ESP; Inglês Instrumental; inglês para aviação; Linguística de Corpus; manutenção de aeronaves.*

**ABSTRACT:** *ESP courses are generally developed based upon the students' needs. Those needs are raised through surveys applied to learners or through analysis, carried out by the teacher or materials designer, of the use of language (academic, occupational or vocational) which learners need to develop. In this article, we present the development process of a syllabus for a course of English aimed at reading aircraft maintenance manuals, in which Corpus Linguistics was used to define the lexico-semantic-grammatical aspects to be covered during the course.*

**KEYWORDS:** *Aircraft maintenance; Aviation English; Corpus Linguistics; ESP*

### 1 Introdução

No início de 2011, uma empresa de transporte aéreo nos consultou sobre a possibilidade de desenvolvermos um curso de inglês para mecânicos de aeronaves, que os habilitasse a ler a documentação técnica em inglês publicada pelos fabricantes. Para desenvolvermos esse

---

<sup>1</sup> *DIY web corpus, (Do It Yourself)*, ou “faça você mesmo” (Tagnin, 2007; Zanettin, 2002).

<sup>2</sup> Doutor em Linguística Aplicada; Mestre em Linguística Aplicada; Graduado em Letras; Especialista em Controle de Tráfego Aéreo. Interesses de pesquisa: Linguagem e cognição; Ensino e aprendizagem de línguas; Simulação e aprendizagem de línguas.

<sup>3</sup> Mestranda em Estudos Linguísticos e Literários em Inglês; Psicopedagoga; Graduada em Letras. Pesquisa a aplicação da Linguística de *Corpus* no ensino de inglês para aviação.

empreendimento, foi necessário, primeiro, conhecer com mais profundidade a atividade desses profissionais e como a linguagem é usada nas suas tarefas. Um período de visitas a centros de manutenção, conversas com mecânicos e o contato direto com a documentação alvo do trabalho forneceram os subsídios iniciais para o delineamento do curso. Neste artigo, apresentaremos a trajetória percorrida no desenvolvimento do conteúdo programático do curso.

Mecânicos de aeronaves são profissionais cujo trabalho se reflete diretamente na segurança de voo, pois são responsáveis pelo estado físico e pelas condições operacionais de aeronaves. São também encarregados das inspeções rotineiras que garantem o bom funcionamento do equipamento e podem intervir no processo decisório que define se uma aeronave pode ou não decolar.

Para que os procedimentos de inspeção e manutenção sejam seguidos de forma padronizada e com o rigor necessário exigido pela sua função, os mecânicos de aeronaves precisam ler manuais e documentos diversos em inglês. As traduções dos manuais são inviabilizadas por vários motivos, tais como a constante necessidade de atualizações enviadas pelos fabricantes de aeronaves e a falta de tradutores especializados no mercado (Drury; Ma, 2003, 2004).

A Joint Aviation Authority - JAA, agência de aviação europeia, regulamenta a necessidade de proficiência linguística para mecânicos.

Equipes certificadas devem ser capazes de ler, escrever e se comunicar em um nível compreensível na(s) língua(s) em que a documentação técnica e os procedimentos organizacionais necessários para dar suporte à emissão do certificado de liberação do serviço estejam escritos. (JAR-66, p. 4, tradução nossa).<sup>4</sup>

No Brasil, não há qualquer exigência, por parte da Agência Nacional de Aviação Civil –ANAC, a respeito da proficiência em língua inglesa para mecânicos. Há, sim, requisito para que o mecânico fale e leia perfeitamente a língua portuguesa<sup>5</sup>. Porém, como já mencionado anteriormente, a maioria dos manuais e documentos são escritos em inglês. Fica evidente, portanto, a lacuna existente entre o que as normas exigem do mecânico, em termos de proficiência linguística, e o que ele realmente precisa para desempenhar com segurança a sua atividade profissional.

Pesquisas recentes (Bocorny, 2008; Sarmiento, 2008), envolvendo o uso da linguagem técnica nos manuais de manutenção aeronáutica, abordam a questão da leitura de manuais e apresentam sugestões de ensino específicas para mecânicos, a partir de levantamentos feitos por meio da Linguística de *Corpus* (doravante L. C.). Segundo as autoras, através da L. C. é possível identificar

---

<sup>4</sup> “Certifying staff must be able to read, write and communicate to an understandable level in the language(s) in which the technical documentation and organisation procedures necessary to support the issue of the certificate of release to service are written.”

<sup>5</sup> Dispositivo constante do Regulamento Brasileiro de Homologação Aeronáutica nº 65. Disponível em: <<http://www.anac.gov.br/biblioteca/rbha/rbha065.pdf>>. Acesso em 14 DEZ 2011.

características comuns presentes nesse tipo de gênero textual, com o objetivo de orientar o desenvolvimento do material didático a ser usado em sala de aula.

No trabalho descrito neste artigo, devido a restrições de tempo impostas pelo regime de trabalho dos profissionais, o esforço de instrução foi planejado para contemplar um módulo introdutório de 50 horas/aula, abordando os aspectos operacionalmente mais críticos da leitura de manuais, conforme já apontados por Bocorny (2008) e Sarmiento (2008), bem como os aspectos linguísticos apontados por um *corpus* compilado a partir de manuais de manutenção utilizados pelos mecânicos, com base na sua frequência de ocorrência. Nossa intenção era que as fases de controle e avaliação da implantação do projeto fornecessem os subsídios necessários para o desenvolvimento de módulos complementares mais abrangentes.

## 2 Metodologia

O ensino de línguas estrangeiras é geralmente amparado no ensino das quatro habilidades – produção e compreensão orais, leitura e produção escrita. No caso exemplificado neste artigo, o objetivo é somente o ensino de leitura em inglês, a ser concretizado pela abordagem de Inglês Instrumental, ou *ESP (English for Specific Purposes)*. De acordo com Hutchinson e Waters, ESP é “[...] uma abordagem para o ensino de línguas em que todas as decisões quanto ao conteúdo e metodologia são baseadas nos motivos pelos quais o aprendiz quer aprender uma língua.” (Hutchinson; Waters, 1994, p. 19, tradução nossa).<sup>6</sup> Sendo assim, como objetivos gerais para o curso, definimos o desenvolvimento de estratégias de leitura (Braown, 2007), as quais, convém lembrar, compreendem o conhecimento léxico-sistêmico da língua alvo.

Ao ter definido a metodologia e os objetivos, foi necessário definir quais conteúdos deveriam compor o currículo. Quais textos, quais estruturas gramaticais e quais campos lexicais e semânticos priorizar em um espaço curto de tempo para que o aluno pudesse aplicar seus novos conhecimentos em seu dia a dia o mais rapidamente possível? Almejava-se que essas perguntas fossem respondidas pela L. C. Para tanto, compilamos um *corpus* com os textos retirados dos manuais de manutenção de aeronaves utilizados pelos mecânicos. Para Viana, a L. C. é

[...] uma compilação eletrônica e criteriosa de (amostras de) textos que ocorrem naturalmente com o objetivo de representar uma dada língua ou algum de seus aspectos mais pontuais de forma a possibilitar uma análise linguística previamente delineada. (Viana, 2008, p. 31).

Temos como objetivo demonstrar como um *corpus DIY*, ou “Faça Você Mesmo” (Tagnin, 2007; Zanettin, 2002), pode auxiliar no desenvolvimento do conteúdo programático de um curso de inglês

---

<sup>6</sup> “[...] an approach to language teaching in which all decisions as to content and method are based on the learner’s reason for learning.”

instrumental. Um *corpus* dessa natureza é compilado pelo próprio pesquisador, conforme seus interesses e necessidades (Tagnin, 2007).

### 3 Conteúdo programático

O conteúdo programático de um curso é “uma lista sequencial de tópicos, situações, habilidades e maneiras de ensinar” (Brown, 2007, p. 156). É uma projeção do que o aluno irá aprender ao longo de um curso e reflete os objetivos gerais e específicos que o professor pretende que seus alunos atinjam ao término do curso.

Os conteúdos programáticos têm sido definidos, há décadas, por meio de gradação e seleção de estruturas gramaticais e de vocabulário. Assim, estruturas consideradas mais básicas, como frases simples com o verbo de ligação *BE*, ocupam as primeiras posições na sequência, em detrimento daquelas consideradas mais complexas. A sintaxe de tais estruturas é considerada nesses casos - quanto mais simples essa sintaxe (sujeito + verbo / sujeito + verbo + complemento), mais inicial é a posição que ela ocupa no conteúdo programático. A seleção de vocabulário foi, por muitos anos, feita de maneira intuitiva, de acordo com o que desenvolvedores de materiais julgavam ser mais necessário. Mais recentemente, com o advento dos estudos linguísticos computacionais, a crença de que a frequência, tanto de estruturas gramaticais quanto de vocabulário, deve ser levada em consideração começa a perpassar conteúdos programáticos de vários cursos, principalmente os voltados para propósitos específicos (Richards, 2001).

Para o curso aqui mencionado, o foco foi a escolha e o sequenciamento do conteúdo, processo que teve a L.C. como principal eixo orientador. Richards (2001) afirma que estudos relacionados à frequência de ocorrência de palavras têm sido bastante prolíferos desde a década de 20 do século passado, tendo recebido um substancial impulso com o surgimento e a disseminação dos computadores pessoais. Gavioli (2005) sugere que o uso de *corpora* no desenvolvimento de conteúdos programáticos em ESP pode ser um excelente recurso e cita, como exemplo, Flowerdew (1993), que isolou aproximadamente 1.000 itens que ocorreram mais de 10 vezes em um *corpus* de 100.000 palavras para desenvolver o conteúdo programático de um curso para biólogos. Richards (2001), entretanto, alerta que a frequência de ocorrência de uma palavra, por si só, não indica, necessariamente, sua utilidade para fins de desenvolvimento de conteúdos programáticos. O autor recomenda que os textos escolhidos para compor o *corpus* sejam relevantes para os aprendizes e que a alta frequência ocorra em uma variada gama de amostras de linguagem. Isso, segundo o autor, assegura que a frequência de ocorrência de palavras em um *corpus* corresponda à importância de sua utilização pelos aprendizes.

Considerando que o conteúdo linguístico do curso devia atender apenas e tão somente a necessidade de leitura, descartamos a busca por um conteúdo fonético, uma vez que o quesito pronúncia não estava relacionado aos objetivos do curso. Portanto, a análise do *corpus* compilado foi direcionada aos aspectos léxico-semântico-gramaticais.

#### **4 Compilação do *corpus***

Com o propósito de compilar um *corpus DIY*, por sua praticidade de acesso e rapidez, decidimos buscar manuais disponibilizados na Internet. Sempre existe, entretanto, a possibilidade de o professor conseguir com seus alunos materiais já digitalizados e, então, montar o seu *corpus*.

A compilação de um *corpus* deve ser feita respeitando critérios bem definidos, tais como gênero, contexto de produção, público (idade, gênero, grau de escolaridade) e objetivo, entre outros. No caso em questão, coletamos manuais de manutenção, manuais operacionais, listas de tarefas e listas de equipamentos mínimos (*minimum equipment lists – MEL*) de aeronaves dos fabricantes Airbus, Boeing, Cessna e Beechcraft. Esses documentos foram escolhidos por serem aqueles que os mecânicos mais manuseiam, de acordo com informações prestadas pelos próprios profissionais. A escolha por aeronaves desses fabricantes se justifica por serem aquelas para as quais os principais centros de manutenção existentes no Brasil oferecem manutenção, tanto na aviação comercial quanto na aviação executiva.

Os arquivos originais são, via de regra, disponibilizados em formato PDF. O *software* utilizado no trabalho, contudo, reconhece apenas arquivos em formato txt. Foi necessária, portanto, a transformação dos arquivos originais em PDF para arquivos txt, dando origem a um *corpus* de 25.126.172 palavras, que pode ser processado pelo *software*. A compilação de um *corpus DIY* nos permitiu a utilização de procedimentos menos rigorosos de coleta de dados, possibilitando, assim, a rapidez necessária ao desafio que se apresentava.

#### **5 Análise léxico-semântico-gramatical do *corpus***

Um dos componentes do conteúdo programático de um curso de línguas é o campo léxico-semântico a ser explorado. Para esse curso, a precisão no levantamento recebeu atenção especial, devido ao caráter de segurança de voo inerente à profissão e espelhado nas tarefas executadas por meio do uso da linguagem. Baseados no *corpus* mencionado anteriormente, apresentamos, aqui, o caminho percorrido por nós para o delineamento do conteúdo programático sugerido. Tendo em vista que o objetivo deste artigo é apresentar a L. C. como balizadora da definição de conteúdos, é

importante ressaltar que não nos preocuparemos, portanto, com a descrição de aspectos pedagógicos, tais como as estratégias de leitura em si, avaliação etc.

Para essa análise, utilizamos o *WordSmith Tools*, versão 5.0 (Scott, 2008), um *software* proprietário. Há, entretanto, a possibilidade de se usar *softwares* livres, facilmente encontrados na Internet. Pretendíamos, com essa análise, buscar os campos lexicais e semânticos fundamentais para um curso introdutório de leitura, assim como detectar as estruturas gramaticais mais importantes para serem trabalhadas com aprendizes em níveis iniciais. Isso foi feito por meio das ferramentas disponíveis no *WordSmith*. Primeiramente, foi feito um levantamento da lista de palavras (*WordList*). A *WordList* gera uma lista com a ordem de frequência das palavras que compõem o *corpus*. Segundo Carter, McCarthy e O’Keefe (2007), a seleção de vocabulário para níveis básicos, como o sugerido neste artigo, pode ser realizada levando em conta os primeiros 1.500-2.000 itens mais frequentes do *corpus*, palavras estas potencialmente necessárias para a comunicação diária ou, no caso deste estudo, para a realização das tarefas diárias dos aprendizes. Ainda de acordo com esses autores,

“[...] armado com a informação complexa que uma lista de frequência pode fornecer, o professor ou o autor do material pode elaborar uma pedagogia cujo vocabulário é mais centrado no uso e prover itens linguísticos úteis e utilizáveis inclusive para aprendizes em níveis básicos” (Carter; McCarthy; O’Keefe, 2007, p. 47, tradução nossa)<sup>7</sup>.

A *WordList* apresentou várias palavras gramaticais, tais como conjunções (*and, if, or, when*), determinantes (*an, the, this, these,*) e os verbos de ligação *is* e *are*, que ocupam a 9ª e 19ª posições, respectivamente. As preposições (*for, from, in, of, on, to* e *with*) estão entre as 22 primeiras palavras da lista, o que sugere grande necessidade de trabalho com esse item gramatical. Uma primeira análise nesta lista nos incitou a curiosidade – a palavra *do*, em 32º lugar, nos fez indagar se sua aparição era mais frequente como verbo principal ou como verbo auxiliar. Outra palavra cuja alta frequência nos intrigou, devido a sua polissemia, é *get*, que apareceu em 147º lugar, com 22.571 ocorrências. Há movimentos distintos, tais como o *Simplified English* e o *Plain English*, que propõem o uso de um inglês mais simplificado e, portanto, mais objetivo e menos ambíguo, em textos que atingem um público internacional, o que, claramente, é o caso dos manuais de manutenção de aeronaves.

Para sanar as dúvidas sobre tais verbos, buscamos-os por meio do concordanciador do *WordSmith (Concord)*, gerando linhas de concordância. A partir daí, foi possível analisar suas primeiras instâncias em contexto.<sup>8</sup> Na TAB. 1, temos as 10 primeiras linhas de concordância para a palavra *do*. Pode-se logo perceber que a maior ocorrência desta palavra é como verbo auxiliar,

---

<sup>7</sup> “Armed with the complex information a frequency list can give, the teacher, syllabus designer or materials writer can elaborate a more use-centred vocabulary pedagogy at the elementary level and provide useful and usable language items even to very low level learners.”

<sup>8</sup> Seguindo a sugestão de Carter, McCarthy e O’Keefe (2007), adotamos um ponto de corte de 1500 itens. Por limitação de espaço, entretanto, os resultados aqui apresentados limitam-se aos mais relevantes.

ocorrendo, em nosso *corpus*, com *not*. Com exceção das linhas 8 e 9, a alta ocorrência do imperativo negativo, reforçado em momentos por letras maiúsculas, sugere esse *cluster* como possível candidato na construção do conteúdo programático proposto.

TABELA 1  
Linhas de concordância - *do*

N		CONCORDANCE	
1	taxiways, or runways. CAUTION:	Do	not operate engine or wing
2	temperatures less than -40C,	do	not start or motor the eng
3	t or are anticipated. WARNING:	Do	not rely on airframe visua
4	ather operations in general:	Do	not use assumed temperatur
5	QFE altimeter settings. Note:	Do	not use LNAV or VNAV below
6	the fix radial and/or distance	do	not intersect the route.
7	Note: Step climb evaluations	do	not consider buffet margin
8	ing the wrong extinguisher may	do	more harm than good. For
9	EXTINGUISHER ON A FIRE COULD	DO	MORE HARM THAN GOOD. FOR
10	WITH SMOKE NOTE: Some symbols	do	not apply to all configurations

Para verificarmos a veracidade de tal constatação, recorreremos à contagem dos primeiros colocados à direita da palavra *do*. O próprio *Concord* oferece uma ferramenta para tal busca. O total de ocorrências de *do* com a palavra *not* é de 32.505. Este fato valida as primeiras análises feitas com as linhas de concordância, o que pode facilitar na busca de outras palavras. A segunda palavra mais frequente coocorrendo com *do* é *the*, com um total de 16.678, ou seja, um pouco a mais da metade. As próximas palavras com as quais *do* coocorre são outros determinantes como *a*, *an* e *this*. Nesses casos, *do* ocorre como verbo principal. Verificando tais combinações (*do the*, *do a* etc) no concordanciador, encontramos os seguintes colocados: *do the test*, *do a check*, *do an inspection*, *do this procedure*. Outras combinações apresentam frequência baixa, o que não justificaria sua presença em um curso introdutório.

Usuários da língua inglesa conhecem a polivalência do verbo *get*. Em vários dicionários, encontramos diversos significados, muitas vezes apresentados em mais de uma página. Esse aspecto polissêmico pode levar a confusões na interpretação dos enunciados. Consideramos que a prevenção de ambiguidades é fator determinante no ensino de inglês com propósitos específicos à manutenção de aeronaves, devido aos níveis de segurança exigidos nessa atividade. Essa preocupação nos levou a buscar, através de linhas de concordância, os principais significados com que *get* aparece no *corpus*. Expomos, na TAB. 2, as linhas de concordância geradas com a palavra *get* e traçamos o mesmo caminho percorrido com a palavra *do*.

TABELA 2  
Linhas de concordância - *get*

N	CONCORDANCE		
1	WARNING: DO NOT LET OBJECTS	GET	IN THE HOUSING ASSEMBLY O
2	in and released frequently to	get	the flow of sealant.
3	in and released frequently to	get	sufficient sealant.
4	away from the door	get	access through the door
5	Use shims as necessary to	get	a maximum clearance of 0.
6	ortant to shim the assembly to	get	hinge point control. The
7	lower stud on the rear spar to	get	the necessary alignment.
8	shim that is necessary to	get	the dimension 0.003 in.
9	WARNING: DO NOT LET OBJECTS	GET	IN THE HOUSING ASSEMBLY O
10	Use a 10 ft platform to	get	access to the wing Leadin

Nas linhas de concordância com *get*, já podemos ver uma tendência pelo significado de entrar ou obstruir, no sentido de “deixar entrar”, em alguns casos. As expressões *get in* e *get access* na TAB. 2 também sugerem isso. A busca pelos colocados nos trouxe *access* como o colocado mais frequente à direita de *get*, com 10.990 ocorrências. O segundo colocado foi *the*, com 2.903 ocorrências. A palavra *in*, apesar de estar nas primeiras linhas de concordância, pois também coocorre em outras posições, só ocupa a primeira posição seguinte a *get* 44 vezes. Consideramos que essa não é uma frequência tão significativa para um curso introdutório. As outras linhas de concordância da TAB. 2 nos mostram um significado em comum: o sentido de obter. Chegamos, então, à conclusão de que *get* pode ser trabalhado sozinho, com o sentido de obter, ou junto ao colocado *access*, com o sentido de entrar.

A análise da *Wordlist* em um *corpus* com milhões de palavras exige um tempo longo e chama a atenção do desenvolvedor do conteúdo programático para questões tais como as vistas acima. Seria impraticável fazer a análise de palavra por palavra no concordanciador para avaliarmos e detectarmos quais estruturas ou palavras devem ser selecionadas. Isso sugere que uma *WordList* por si só pode não ser suficiente na obtenção de vocabulário para a construção de um conteúdo programático, inclusive porque, em nosso caso, a alta incidência de palavras gramaticais prevalece sobre a incidência de palavras de conteúdo. A *WordList* serve, assim, como ponto de partida para análises mais específicas. As palavras que de fato trazem dúvidas importantes ao analista podem ser avaliadas por meio do concordanciador, nas linhas de concordância ou nos colocados frequentes, que podem ocorrer tanto à direita do nóculo (*node word*) quanto a sua esquerda.

Até aqui, fizemos importantes descobertas: algumas palavras gramaticais podem ser de extrema importância para o aprendiz alvo do curso em desenvolvimento. Palavras que poderiam ser ambíguas foram contrastadas com suas funções e significados nas linhas de concordância. O próximo



passo tomado foi gerar uma *KeyWordList*, ou seja, uma lista com as palavras-chave do *corpus*. A TAB. 3 apresenta as 50 primeiras palavras-chave geradas por essa busca.

TABELA 3  
Lista das 50 primeiras palavras-chave do *corpus*

N	Palavra-chave	N	Palavra-chave	N	Palavra-chave
1	TASK	18	AIRCRAFT	35	FLIGHT
2	SUBTASK	19	VALVE	36	NOTE
3	REMOVE	20	REMOVAL	37	ENERGIZE
4	SURE	21	SAFETY	38	MATERIAL
5	POSITION	22	ENGINE	39	PLATFORM
6	ACCESS	23	SWITCH	40	CONFIG
7	PANEL	24	SPECIFIC	41	GROUND
8	INSTALL	25	FUEL	42	NUT
9	HYDRAULIC	26	PROCEDURE	43	BREAKER
10	WARNING	27	CONTROL	44	AFT
11	DESIGNATION	28	APU	45	FLAP
12	MAKE	29	PUT	46	COMPARTMENT
13	SYSTEM	30	MAINTENANCE	47	UNIT
14	DOOR	31	NO	48	ON
15	ELECTRICAL	32	CIRCUIT	49	CLOSE
16	INSTALATIONS	33	TORQUE	50	CONFIGURATION
17	TEST	34	PUSHBUTTON		

Uma avaliação da lista de palavras-chave nos traz alguns campos semânticos importantes, tais como partes da aeronave (*door, engine, APU, flap*), sistemas (*hydraulic, electrical, fuel*), interruptores ou conectores (*switch, pushbutton, breaker*) e ações envolvidas com a operação de máquinas ou computadores (*remove, install, test, close*), entre outros. Algumas dessas palavras, entretanto, podem promover ambiguidades, principalmente em relação à classe gramatical. *Switch* e *test*, por exemplo, podem ser verbos ou substantivos. Além disso, algumas dessas palavras são, reconhecidamente, expressões fixas, tais como *make sure, access panel, access door, circuit breaker*, entre outras. Para encontrar tais expressões, geramos, então, outra lista, desta vez de *clusters*, também denominados n-gramas, primeiramente com expressões contendo duas palavras, para, em seguida, gerar uma lista de três e quatro palavras. Biber, Conrad e Leech (2002) alegam que sequências de duas palavras (bigramas) são curtas demais para demonstrarem algum interesse. Ainda de acordo com esses autores, combinações de cinco ou seis palavras, apesar de existirem, são bastante incomuns. No nosso caso, a própria análise da lista de palavras-chave sugeriu que algumas delas possivelmente formassem bigramas. Como podemos observar na TAB. 4, algumas de nossas suspeitas em torno desses bigramas

foram confirmadas. Expressões fixas como *make sure* e *circuit breaker*, por exemplo, aparecem na lista dos 10 primeiros colocados. A presença do bigrama *do not* corroborou a nossa análise da palavra *do* como verbo auxiliar.

TABELA 4  
Lista dos 10 bigramas mais frequentes no *corpus*

N	BIGRAMA	FREQUÊNCIA
1	MAKE SURE	116.045
2	FOR DETAILS	63.667
3	AIRCRAFT MAINTENANCE	38.556
4	IN POSITION	38.025
5	NO SPECIFIC	37.453
6	SET UP	33.106
7	DO NOT	32.506
8	LANDING GEAR	24.346
9	ACCESS PLATFORM	23.540
10	CIRCUIT BREAKER	23.010

Confirmando a pesquisa de Bocorny (2008), há, na TAB. 4, várias ocorrências de sintagmas nominais, tais como *aircraft maintenance*, *landing gear*, *access platform*. A autora enfatiza a importância do ensino desse grupo léxico-gramatical, justificando:

[...] a capacidade de entendimento destas unidades pelos aprendizes é afetada pela dificuldade que os mesmos têm em identificar as unidades dentro do corpo do texto especializado, em entender as relações semânticas estabelecidas entre os pré-modificadores e entre estes e a base nominal e em dar sentido à unidade, a partir da interpretação do conceito representado por elas. (Bocorny, 2008, p. 16).

A presença desse grupo léxico-gramatical é, também, facilmente percebida na busca por trigramas. É nessa análise que encontramos outros aspectos sintáticos importantes, que nos auxiliaram na definição de estruturas léxico-gramaticais em determinados campos semânticos. Na lista de trigramas (TAB. 5), percebemos a presença de orientações de operação de maquinário no imperativo, caracterizando o gênero instruções, muito comum às tarefas dos mecânicos. Dentro do número de corte definido, ainda encontramos instâncias de voz passiva, tais como *item is illustrated*, *replacements are permitted*, expressões preposicionadas indicando posição na aeronave, como *in the cockpit*, *of the engine*, *on the ECAM* e, ainda, a presença de marcadores de discurso como *comes into view*, *in case of*, *in order to*.

TABELA 5  
Lista dos 10 trigramas mais frequentes no *corpus*

N	TRIGRAMA	FREQUÊNCIA
1	MAKE SURE THAT	91.460
2	JOB SET UP	32.853
3	SET UP INFORMATION	13.991
4	FOR THE JOB	13.749
5	ENERGIZE THE AIRCRAFT	13.621
6	AIRCRAFT ELECTRICAL CIRCUITS	13.458
7	AREA IS CLEAN	12.676
8	TELL PERSONS NOT	12.064
9	EQUIPMENT REFERENCE QTY	11.120
10	HAZARDOUS MATERIAL WARNINGS	11.110

Com o intuito de verificar a pertinência de alguns campos semânticos e estruturas que surgiram nas análises anteriores, buscamos, também, *clusters* de 4 palavras. Os resultados apontaram, novamente, para o uso de verbos imperativos (*remove all the tools, clean the component interface, put the aircraft back*), partes da aeronave (*aft cargo compartment door, leading edge and slats*), expressões preposicionadas (*on the center pedestal, on the overhead panel, on the landing gear, on the lower ECAM*) e para o uso da voz passiva (*breakers are not opened, material warnings are listed*).

Os dados também apontam que algumas formas gramaticais predominam em certos campos semânticos. Os verbos referentes à operação de maquinário, por exemplo, aparecem frequentemente no imperativo (*energize the aircraft, put a warning, remove the warning, put the access platform, put the safety cable*). Já o campo relacionado à condição de partes da aeronave, botões e componentes eletrônicos é geralmente representado com a voz passiva (*breakers are not opened, electrical cables are installed*). O funcionamento de sistemas é frequentemente descrito com o presente simples (*the engine bleed air valve closes, flap position indicator displays, standby hydraulic system powers*). É possível verificar tais exemplos buscando-nos nas listas de n-gramas ou em linhas de concordância por meio das palavras constantes das listas de palavras, tanto a *WordList* quanto a *KeyWordList*.

Sarmiento (2008) fez um estudo sobre verbos modais em manuais de manutenção aeronáutica, cuja alta frequência também ocorre no *corpus DIY* aqui apresentado. A autora constata que a recorrência de verbos modais, principalmente de *should* e *must*, é bem mais frequente nos manuais de manutenção de aeronaves, se comparada a outros *corpora*. Em nossa análise, a maior frequência recai

sobre o verbo modal *can*, em colocados com *can cause damage* ou *can cause injuries*. Essa constatação corrobora um alerta da autora:

“[...] com relação aos manuais de aviação, a má interpretação da força de uma instrução pode **causar problemas muito maiores** e até **perdas de vida**, o que explicaria o uso por uma modalização mais forte e impositiva.” (Sarmiento, 2008, p. 156, grifo nosso).

Expressões como as citadas no parágrafo anterior são incluídas, no conteúdo programático, no campo semântico referente a ameaças e avisos (*threats and warnings*).

O caminho por nós percorrido reflete a visão de Sinclair (1991) que rejeita a separação entre léxico e gramática. Para ele,

[...] cada significado pode ser associado com um padrão formal distinto. Essa regularidade é tamanha que, brevemente, podemos ver padrões formais sendo usados claramente como critério para análises de significados, o que se configura como uma posição mais segura e menos excêntrica para uma disciplina que aspira à seriedade científica. (Sinclair, 1991, p. 6-7, tradução nossa)<sup>9</sup>

Assim, as palavras gramaticais encontradas na primeira lista gerada por meio da *WordList* podem ser encontradas nas expressões surgidas nas buscas por bigramas, trigramas e poligramas, fazendo parte da fraseologia que permeia o gênero a ser estudado por mecânicos de aeronaves.

Os resultados das buscas realizadas no *corpus* sugerem que alguns dos procedimentos utilizados no desenvolvimento do conteúdo programático do curso podem, também, ser usados para orientar a seleção de tópicos e textos a serem trabalhados nas atividades didáticas. Uma busca por n-gramas, por exemplo, nos traz as linhas de concordância e, conseqüentemente, os textos onde elas se encontram, tornando essa seleção um processo mais dinâmico.

O QUADRO 1 apresenta alguns exemplos dos campos semânticos e estruturas gramaticais revelados na análise do *corpus*.

---

<sup>9</sup> “[...] each meaning can be associated with a distinctive formal patterning. So regular is this that in due course we see may see formal patterns being used overtly as criteria for analyzing meaning, which is a more secure and less eccentric position for a discipline which aspires to scientific seriousness.”

QUADRO 1

Campos semânticos e estruturas léxico-gramaticais

Campo semântico	Estruturas léxico-gramaticais
Parts of the aircraft	<p>Words:</p> <ul style="list-style-type: none"> <li>▪ door, engine, APU, flap, panels, cockpit</li> </ul> <p>Expressions:</p> <ul style="list-style-type: none"> <li>▪ landing gear, aft cargo compartment, on the center pedestal, on the overhead panel, on the lower ECAM, near the flight controls</li> </ul>
Systems	<p>Words:</p> <ul style="list-style-type: none"> <li>▪ hydraulic, electrical, fuel, trim, avionics, generator</li> </ul> <p>Expressions:</p> <ul style="list-style-type: none"> <li>▪ bleed air, bleed air valve closes, flap position, position indicator displays, standby hydraulic system powers</li> </ul>
Tools	<p>Words:</p> <ul style="list-style-type: none"> <li>▪ torque, nut, pin, bolts, screws, washers, tool, drain, clip</li> </ul> <p>Expressions:</p> <ul style="list-style-type: none"> <li>▪ install the bolts, specific torque range, ground safety locks, locks are in position, put blanking caps on, discard the cotter pin</li> </ul>
Electrical switches or connectors	<p>Words:</p> <ul style="list-style-type: none"> <li>▪ valve, switch, lever, pushbutton, connector</li> </ul> <p>Expressions:</p> <ul style="list-style-type: none"> <li>▪ circuit breakers, disconnected electrical connector, safety the circuit breakers, breakers are not opened</li> </ul>
Operation of machinery or computers	<p>Words:</p> <ul style="list-style-type: none"> <li>▪ check, do, close, open, set, disconnect, connect, seal</li> </ul> <p>Expressions:</p> <ul style="list-style-type: none"> <li>▪ remove all the tools, remove the ground support, push the line key, put the aircraft back</li> </ul>
Threats and warnings	<p>Words:</p> <ul style="list-style-type: none"> <li>▪ can, damage, injuries, warning, poisonous, flammable</li> </ul> <p>Expressions:</p> <ul style="list-style-type: none"> <li>▪ warning notice, hazardous material, make sure that, not breathe the gas, tell persons not to pressurize, can cause injury to</li> </ul>
The work area	<p>Words:</p> <ul style="list-style-type: none"> <li>▪ platform, clean, external, discard</li> </ul> <p>Expressions:</p> <ul style="list-style-type: none"> <li>▪ adjacent area, access platform, with cleaning agents material, discard the lockwire, lint free cotton cloth, put on protective clothing</li> </ul>

## 6 Conclusão

Neste artigo, apresentamos o caminho percorrido para definir o conteúdo programático de um curso de inglês para mecânicos de aeronaves, com foco na habilidade de leitura, com base em um *corpus DIY*, ou “faça você mesmo”. Partindo de documentos de manutenção disponibilizados na

Internet, compilamos um *corpus* de 25.126.172 palavras, o que nos propiciou uma amostra bastante abrangente para o fim a que nos propusemos. A partir de listas de palavras mais frequentes, linhas de concordância e listas de n-gramas, foi possível definir não apenas o conteúdo lexical, mas, também, os campos semânticos e as estruturas gramaticais componentes de um conteúdo programático que atendesse as necessidades de aprendizes nos níveis iniciais de competência na língua alvo. Ao término do processo, podemos afirmar que a Linguística de *Corpus* se mostrou uma ferramenta deveras eficaz na definição do conteúdo proposto.

Além do objetivo atingido, vislumbramos, também, outras possibilidades para o uso de um *corpus DIY* no desenvolvimento de cursos de línguas com propósitos específicos. À guisa de exemplo, esse tipo de *corpus* pode auxiliar na escolha dos textos a serem trabalhados, orientar o desenho de atividades didáticas e iluminar o estabelecimento de relações entre os diferentes tópicos definidos para o curso, permitindo, assim, o delineamento e posterior encadeamento das unidades didáticas. Esses processos, evidentemente, se mostram campos férteis para futuras pesquisas sobre o engajamento da Linguística de *Corpus*, tanto no desenvolvimento de materiais quanto no ensino de línguas estrangeiras.

## Referências

BIBER, D.; CONRAD, S.; LEECH, G. *Longman Grammar of Spoken and Written English*. London: Longman, 2002.

BOCORNHY, 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 da Linguagem) – Instituto de Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2008.

BROWN, H. D. *Teaching by Principles: An Interactive Approach to Language Pedagogy*. New York: Pearson Longman, 2007.

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

DRURY, C. G.; MA, J. *Language Errors in Aviation Maintenance: Year 1 Interim Report, Reports to William J. Hughes Technical Center – Federal Aviation Administration*. 2003. Disponível em: <[http://www.hf.faa.gov/docs/508/docs/maint\\_LanguageYr1rpt.pdf](http://www.hf.faa.gov/docs/508/docs/maint_LanguageYr1rpt.pdf)> Acesso em: 05 JAN 2011.

DRURY, C. G.; MA, J. *Language Errors in Aviation Maintenance: Year 2 Interim Report, Reports to William J. Hughes Technical Center – Federal Aviation Administration*. 2004. Disponível em: <[https://hfskyway.faa.gov/\(A\(fjvOpyVSyGekAAAANDQ4MmY4MWEtYWU1Zi00NTM0LThjM2U](https://hfskyway.faa.gov/(A(fjvOpyVSyGekAAAANDQ4MmY4MWEtYWU1Zi00NTM0LThjM2U)

tOTgyNzQwY2I4YjFiw9dzs9fQS97xprE\_ZpeRIq31O9U1))/HFTest/Bibliography%20of%20Publications/Human%20Factor%20Maintenance/MaintLanguageYr2.pdf> Acesso em: 05 JAN 2011.

FLOWERDEW, J. Concordancing as a tool in course design. *System*, 21, 1993. p. 231-243.

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

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

HUTCHINSON, T.; WATERS, A. *English for specific purposes*. Cambridge: CUP, 1994.

RICHARDS, J. *Curriculum Development in Language Teaching*. Cambridge: CUP, 2001.

SARMENTO, S. *O uso dos verbos modais em manuais de aviação em inglês: um estudo baseado em corpus*. 2008. 283 f. Tese (Doutorado em Estudos da Linguagem) – Instituto de Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2008.

SCOTT, M. *WordSmith Tools*. Versão 5. Liverpool: Lexical Analysis Software, 2008.

SINCLAIR, J. *Corpus, concordance, collocation*. Oxford: OUP, 1991.

TAGNIN, S. A Identificação de equivalentes tradutórios em *corpora* comparáveis. 2007. Disponível em: <[http://www.fflch.usp.br/dlm/comet/Novo/Stella\\_Abrapui%202007\\_artigo.pdf](http://www.fflch.usp.br/dlm/comet/Novo/Stella_Abrapui%202007_artigo.pdf)> Acesso em: 14 DEZ 2011.

VIANA, V. *Verbos modais em contraste: análise de corpus da escrita de universitários em inglês*. 2008. 230 f. Dissertação (Mestrado em Letras) – Departamento de Letras, PUC RJ. Rio de Janeiro, 2008.

ZANETTIN, F. DIY corpora: the WWW and the translator. In: MAIA, B.; HALLER, J.; ULRICH, M. (Eds.). *Training the Language Services Provider for the New Millennium*. Porto: Faculdade de Letras, Universidade do Porto, 2002. p. 239-248.