



ENSINO APLICADO DE INGLÊS NO IFMT CAMPUS PRIMAVERA DO LESTE: PROPOSTA A PARTIR DA LINGUÍSTICA DE CORPUS

Wesley Alves Siqueira; Eduardo Walter da Silva; Rafael Silva Smaniotto; Rosana Bueno de Sousa

Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso / E-mail: wesley.siqueria@pdl.ifmt.edu.br, edu.ado.5@hotmail.com, rafaelismaniotto18@gmail.com, rosana.sousa@bag.ifmt.edu.br

Resumo: O presente trabalho objetiva identificar as construções mais utilizadas em Língua Inglesa nos manuais técnicos utilizados nas indústrias que compõem o parque industrial da cidade de Primavera do Leste a fim da elaboração de materiais e métodos de ensino e autoaprendizagem. Metodologicamente o trabalho se organizou na definição dos parâmetros do ESP (English for Specific Purposes), seguido da escolha das amostras de texto para coleta de dados para análise. Para análise, ferramentas computacionais como Manyeyes e Word Smith Tools foram utilizadas a fim de quantificar e apontar relações de uso entre as palavras. Nosso trabalho conta com uma análise de corpora de 17.778 palavras, com nível de escolaridade College Graduate. O tempo médio de leitura do corpus é de 65 minutos e traz a palavra water como a palavra-chave mais frequente.

Palavras-Chave: Linguística de Corpus, ESP, Ensino, Currículo.

INTRODUÇÃO

O título do artigo é “Ser/Estar professor de inglês no cenário da escola pública: em busca de um contexto eficaz de ensino/aprendizagem” e data, por mais atual que pareça, de 2002. Passaram-se treze anos e pouco avançamos na aragem do solo para construção de espaços, públicos, férteis em que o ensino de língua inglesa se apresente enquanto possibilidade concreta de se realizar. É das autoras, ainda, Cox e Assis-Perterson (2002), a assertiva de que “falar de ensino de língua inglesa, como de qualquer outra língua estrangeira (LE), no cenário das escolas públicas [...] é, de um modo geral, falar de fracasso, de insucesso, de malogro, de frustração, de mal-estar”. É nesta perspectiva que pensamos o presente trabalho.

Quando falamos do ensino-aprendizagem do inglês pensamos nos resultados destes processos. Num estudo acerca do grau de proficiência no uso da língua inglesa, por profissionais da América Latina, o Pagegroup, empresa de recrutamento de executivos, profissionais técnicos e analistas, hoje, presente em 34 países, revelou o pouco avanço do setor do Agronegócio neste quesito.



Gráfico 1 : Proficiência por setores no Brasil (Profissionais Técnicos e Analistas).

Fonte: Dados do PageGroup – Page Personnel

O estudo ainda revelou que o Brasil em relação à Argentina, Chile e México é o país que apresenta menor grau de proficiência, fluente e avançado, juntos, sendo 64%, 68%, 68% e 86%, levando em conta as respostas de Gerentes e Diretores recrutados nestes países e, 56%, 56%, 49% e 60% para Profissionais Técnicos e Analistas, por país respectivamente.

Nosso trabalho foi desenvolvido no município de Primavera do Leste – MT. Cidade potencializada principalmente pelas atividades ligadas ao agronegócio. Grande partes dos estudantes vinculados a nossa escola tem como prospecção de futuro o investimento do trabalho neste mercado.

LISTAS PARA ENSINO OU ENSINO DE LISTAS

É muito comum se deparar na Internet com listas que elencam as palavras mais utilizadas em inglês como, por exemplo, os sites [englishspeak.com](http://www.englishspeak.com)¹, [englishexperts.com](http://www.englishexperts.com)², [duolingo.com](https://www.duolingo.com)³, [inglesnapontadalingua.com.br](http://www.inglesnapontadalingua.com.br)⁴, entre outros domínios que replicam as mesmas listas ou trazem outras com mais ou menos palavras. Associada a essa grande difusão de listas está a compreensão, também difundida na rede mundial de computadores, da concepção de fluência no idioma, no caso

¹ <http://www.englishspeak.com/pt/english-words.cfm>

² <http://www.englishexperts.com.br/forum/1000-palavras-mais-usadas-em-ingles-t4469.html>

³ <https://www.duolingo.com/comment/1445914>

⁴ <http://www.inglesnapontadalingua.com.br/2007/10/500-palavras-mais-usadas-na-Ingua.html>



III CONEDU

CONGRESSO NACIONAL DE
E D U C A Ç Ã O

da Língua Inglesa, enquanto decorrida do conhecimento destas palavras mais utilizadas (entre 1500 a 2000 palavras). Reflexo imediato, as pessoas estão decorando listas. Livros trazem listas⁵.
Professores ensinam listas

A pergunta inicial que relaciona esse movimento à nossa pesquisa é: de que material, de que *corpus* são extraídas essas informações que irão compor as listas dos livros? Para além desta indagação, tomando como o pressuposto de que estas são extraídas de materiais autênticos e que constituem materiais *corpora* de relevância, são estas, as palavras mais utilizadas, as que mais são recorrentes também nos textos técnicos? E para responder a esta pergunta propomos esta investigação de Linguística de *Corpus*.

Assim, a presente pesquisa está alinhada à área das pesquisas Linguísticas de *Corpus*. Berber Sardinha (2004) define *corpus* enquanto “conjunto de dados linguísticos (pertencentes ao uso oral ou escrito da língua, ou a ambos), sistematizados segundo determinados critérios”. Nas investigações linguísticas, faz-se de essencial importância a seleção de documentos que sejam autênticos e completos para análise, sob o risco de produção de análises extremamente superficiais. Para o linguista, estes dados devem ser:

[...] suficientemente extensos em amplitude e profundidade, de maneira que sejam representativos, da totalidade do uso linguístico ou de algum de seus âmbitos, dispostos de tal modo que possam ser processados por computador, com a finalidade de propiciar resultados vários e úteis para descrição e análise. (BEBER SARDINHA, 2004)

Neste sentido, Berber Sardinha⁶ aponta que os estudos em Linguística de *corpus* ocupam-se “da coleta e exploração de corpora, ou conjuntos de dados linguísticos textuais coletados criteriosamente, com o propósito de servirem para a pesquisa de uma língua ou variedade linguística.”

Exemplo de trabalho, a partir da Linguística de *Corpus*, com aceite da comunidade linguística mundial, é o trabalho da Wordcount.com que trabalha com dados da British National Corpus⁷, que tem um grande número de sentenças escritas e faladas de diferentes fontes.

Este trabalho pensa a construção de curso, *course design for language teaching*, a partir do investimento de uma pesquisa em *corpora* específica: em conjunto de textos de manuais técnicos de máquinas utilizadas nos pátios industriais da cidade Primavera do Leste. De modo a recolocar o

⁵ A título de exemplificação, a coleção Take Over, da autora Denise Santos, aprovada no Programa Nacional do Livro Didático. SANTOS, Denise. Take Over. São Paulo: Editora Lafonte Ltda, 2010. Vale destacar, que tratamos aqui da lista inicial apresentada no início do volume 1 da coleção com as palavras mais frequentes da língua inglesa e não o trabalho de linguística de *corpus* que geralmente compõe o final das obras em formato de glossário.

⁶ Idem.

⁷ Banco de dados linguísticos da Inglaterra.



III CONEDU

CONGRESSO NACIONAL DE
E D U C A Ç Ã O

sentido da aprendizagem enquanto ferramenta de auxílio ao desenvolvimento de atividades cotidianas que, conseqüentemente, culminarão para o desenvolvimento dos arranjos produtivos, econômicos e sociais local.

METODOLOGIA

O presente trabalho se insere na área de pesquisa da Linguística de *Corpus*. Hoje, um dos maiores pesquisadores na área de Linguística de *Corpus* no Brasil é Tony Beber Sardinha, pesquisador do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e professor do Departamento de Linguística e do Programa de Pós-Graduação em Linguística Aplicada e Estudos da Linguagem da Pontifícia Universidade Católica de São Paulo (PUCSP). Para o especialista, “a Linguística de Corpus se ocupa de quase todas as áreas de investigação linguística” (SARDINHA, 2004:1). Ainda para o autor, dentre as áreas de investigação, o léxico é a que mais recebe atenção. Há áreas mais recentes de estudo de *corpus*, como o ensino de línguas estrangeiras (SARDINHA, 2004).

Dos produtos resultantes do estudo de *corpora*, podemos citar os corretores ortográficos, resumidores, sintetizadores de voz, tradutores, digitadores, concordanciadores e extratores de frequência, sendo o último importante para análise do corpus utilizado para a elaboração de um material-base para o ensino de língua estrangeira (SARDINHA, 2004).

Um dos textos-base do projeto que dá sustentação a este artigo é o texto “O Uso de *Corpora* na Análise Linguística”, de Guilherme Fromm, que tem por objetivo a definição do termo *corpus* e a apresentação de maneiras de estudo de *corpus*. O trabalho foi feito com base nos textos de Baker (1995); Bidermann (2001); Fromm (2002); Houaiss (2001); Sardinha (1999).

Para o autor:

[...] corpus constitui um conjunto homogêneo de amostras da língua de qualquer tipo (orais, escritos, literários, coloquiais, etc.). Tais amostras foram escolhidas como modelo de um estado ou nível de língua predeterminado. A análise dos dados linguísticos de um corpus deve permitir ampliar o conhecimento das estruturas linguísticas da língua que eles representam (BIDERMANN apud FROMM, 2003).

Trata o *corpus*, portanto de uma seleção usos da língua, com a finalidade de compreensão desta.

Fromm recorre a Baker para pensar os critérios de seleção de *corpus*. Para o último, são os critérios:

Corpora are generally designed on the basis of a number of selection criteria, the most important of which are:
(i) general language vs. restricted domain
(ii) written vs. spoken language



III CONEDU

CONGRESSO NACIONAL DE
E D U C A Ç Ã O

- (iii) synchronic vs. diachronic
- (iv) typicality in terms of range of sources (writers/speakers) and genres (e.g. newspaper editorials, radio interviews, fiction, journal articles, court hearings)
- (v) geographical limits, e.g. British vs. American English
- (vi) monolingual vs. bilingual or multilingual.

Entre esses critérios, utilizamos um domínio restrito do uso da linguagem (i), em textos escritos (ii), a partir de uma perspectiva sincrônica (iii) de um gênero textual específico (iv): o manual de instrução, a fim de fazer a análise da frequência das palavras mais comuns, das classes gramaticais, construção de glossários e materiais para o ensino de língua estrangeira.

Uma das ferramentas que utilizamos para a análise do corpus é o WordSmith Tools, uma compilação de softwares que permite fazer a análises de frequência e co-ocorrência de palavras, além de permitir pré-processar os arquivos do *corpus* antes da análise. O programa tem a intenção de servir como uma ferramenta para consecução de tarefas relacionadas às análises de *corpora*, disponibilizando diversas ferramentas para o mesmo. (SARDINHA, 2009)

Em relação a tipologia textual em trabalho, tratamos de manuais técnicos. Segundo Costa, o manual pode ser:

- uma obra, espécie de compêndio, de formato pequeno que contém noções ou diretrizes relativas a uma disciplina, a uma ciência, programa escolar, etc.;
- livro que orienta a execução ou o aperfeiçoamento de determinada tarefa ou técnica. Muitas vezes, trata-se de um guia (v.) prático, como por exemplo um manual de digitação, de corte e costura;
- livreto descritivo e explicativo que acompanha determinados produtos, orientando acerca do uso, do funcionamento, da conservação, instalação, etc., como os manuais que acompanham eletrodomésticos ou eletroeletrônicos, etc. (COSTA, 2008)

Para nosso trabalho utilizamos a última definição. Ainda para o autor, o manual caracteriza-se pelas orientações dadas usando o imperativo e o infinitivo, numa conversa direta com o leitor, com o predomínio do discurso instrucional e didático.

Outro conceito operador fundamental na condução do trabalho é a definição dos termos do ESP. Segundo Rosângela Munhoz (2000), o ESP (English for Especific Propouses) é: “[...] uma metodologia de ensino de língua inglesa na qual todas as decisões com relação ao conteúdo ser ministrado e suas estratégias estão baseadas nas necessidades do Educando.” De modo a se pensar o inglês instrumental enquanto uma forma do ensino da língua inglesa que compreende o idioma como uma língua técnica e científica e de caráter geral, tendo por objetivo o uso de estratégias específicas. Além do mais, tende a desenvolver a compreensão de textos de diversas áreas do conhecimento escritos em inglês, usando para isso estratégias intuitivamente aplicadas na língua nativa e propondo outras (GUANDALINI, 2002).



III CONEDU

CONGRESSO NACIONAL DE
E D U C A Ç Ã O

Metodologicamente organizamos o trabalho a partir do estudo sistemático das abordagens e teorias da Linguística de *Corpus* a fim de que se possa compreender como se dão os trabalhos de análise linguística. Em seguida, com o objetivo de compreender o cenário da pesquisa, fez-se levantamento do histórico do ensino e parâmetros da abordagem ESP, no ensino de Língua Inglesa. Definido os conceitos que embasam a pesquisa e o cenário, foi feita a escolha do conjunto de textos que foi analisado, manuais técnicos. Com base nestes textos, que, com auxílio de ferramentas específicas (ferramentas computacionais como *Manyeyes* e *Word Smith Tools*), os léxicos foram contabilizados para análise.

RESULTADOS E DISCUSSÕES

A análise de *corpora* contou com o uso de cinco manuais técnicos, a saber:

- PowerFlex 40 Quick Starter;
- Thermoelectirc Air Conditioner installation and operation manual;
- Water Pump and Motor Assembly Installation Manual;
- High-Capacity Cold Carbonator Assemblies;
- Eletronic Controllors for refrigerating units.

A partir do uso do Software online WordCounter e do Word Smith Tools obtivemos o trabalho com: *corpora* composta com 17.778 palavras, totalizando um montante de 125.880 caracteres. De acordo com Sardinha (1999) *apud* Fromm (2003) o *corpus* pode ser classificado como pequeno por possuir menos de 80.000 palavras.

O conjunto de texto apresenta o nível de leitura College Graduate. De acordo com o Blog WordCounter.net (2015), a ferramenta Reading Level, do software wordcounter.net, tem por função indicar o nível de escolaridade que uma pessoa deve ter para o desenvolvimento da leitura competente do texto, ou conjunto de textos, em análise. A análise é feita a partir do Dale-Chall readability formula que consiste na comparação do texto produzido com uma lista de 3.000 palavras comuns da língua inglesa. Caso o texto produzido possua a maioria das palavras dentro desta lista de palavras, é considerado um texto de nível básico, que pode ser entendido por uma turma de quarta série. Quanto maior o número de palavras fora desta lista, maior é o nível de dificuldade.

O tempo de leitura estimado para *corpora* reunida é de aproximadamente 65 minutos – número baseado em uma velocidade de leitura de 275 palavras por minuto – sendo o tempo médio



III CONEDU

CONGRESSO NACIONAL DE
E D U C A Ç Ã O

de leitura de cada manual de 9 a 21 minutos, tomando como base o número de palavras.

A lista das 50 palavras-chaves mais frequentes são:

- | | |
|------------------------|--------------------------|
| 1. 262 (3%) water | 23. 47 (0%) voltage |
| 2. 148 (2%) unit | 24. 46 (0%) equipment |
| 3. 107 (1%) co2 | 25. 45 (0%) tank |
| 4. 107 (1%) line | 26. 42 (0%) control |
| 5. 88 (1%) pump | 27. 41 (0%) assembly |
| 6. 88 (1%) carbonator | 28. 41 (0%) installation |
| 7. 88 (1%) drive | 29. 41 (0%) output |
| 8. 85 (1%) power | 30. 39 (0%) system |
| 9. 74 (1%) check | 31. 39 (0%) air |
| 10. 67 (1%) valve | 32. 38 (0%) service |
| 11. 65 (1%) input | 33. 37 (0%) operation |
| 12. 64 (1%) inlet | 34. 37 (0%) start |
| 13. 61 (1%) supply | 35. 37 (0%) type |
| 14. 60 (1%) time | 36. 36 (0%) failure |
| 15. 56 (1%) motor | 37. 36 (0%) plain |
| 16. 56 (1%) carbonated | 38. 36 (0%) fault |
| 17. 53 (1%) manual | 39. 35 (0%) injury |
| 18. 52 (1%) damage | 40. 35 (0%) terminal |
| 19. 52 (1%) enclosure | 41. 34 (0%) phase |
| 20. 51 (1%) safety | 42. 34 (0%) side |
| 21. 50 (1%) set | 43. 33 (0%) digital |
| 22. 50 (1%) parameter | 44. 32 (0%) result |

Dessas 50 palavras, 6 foram retiradas por se tratarem de unidades de medidas.

CONCLUSÃO

Embora a *corpora* aqui analisada seja classificada como pequena, a reunião do conjunto de textos foi trabalhosa. Demanda tempo e principalmente a organização dos dados para o trabalho com os softwares linguísticos.

A definição do nível de leitura influencia diretamente no público-alvo quanto à elaboração de materiais. A atenção nesta parte do desenho de um curso é imprescindível para a qualidade e garantia de sucesso no processo de ensino- aprendizagem. Por exemplo, no caso específico, os cursos Técnicos Subsequentes de nosso Instituto, teríamos problemas/dificuldades em pensar a construção de curso a partir deste *corpus*. Além de extenso, está escrito em língua estrangeira e apresenta grau de leitura acadêmico.

Já a análise do tempo de leitura pode ser útil para definir a duração de cada aula. E mais, qual o tempo que será gasto pelos estudantes na execução de atividades.



III CONEDU

CONGRESSO NACIONAL DE
E D U C A Ç Ã O

Outras inferências também podem ser retiradas dos dados coletados como, por exemplo, a não presença significativa de grupos pronominais entre os termos-chave mais frequentes. Fato esse, que nos ajuda direcionar o desenho / planejamento das disciplinas de Inglês Aplicado na instituição.

AGRADECIMENTO

Agradecemos ao apoio institucional do Instituto Federal de Ciência, Educação e Tecnologia de Mato Grosso, através de disponibilidade de auxílio financeiro ao pesquisador e de bolsas de Iniciação Científica da Pró – Reitoria de Pesquisa e Inovação do IFMT (PROPES/IFMT).

REFERÊNCIAS BIBLIOGRÁFICAS

- BAKER, M. Corpus in Translation Studies: an overview and some suggestions for future research. In: *Target 7:2*. Amsterdam: John Benjamins, 1995.
- BERBER SARDINHA, T. B. *Pesquisa em Linguística de corpus com WordSmith Tools*. 1ª ed. Campinas, São Paulo: Mercado de Letras, 2009.
- BERBER SARDINHA, T. *Linguística de Corpus*. Barueri-SP: Manole, 2004.
- BERBER SARDINHA, Tony. Linguística de Corpus: uma entrevista com Tony Berber Sardinha. *Revista Virtual de Estudos da Linguagem - ReVEL*. Vol. 2, n. 3, agosto de 2004.
- BIDERMANN, M.T.C. *Teoria Linguística*. 2. ed. São Paulo: Martins Fontes, 2001.
- COSTA, S. R. *Dicionário de Gêneros Textuais*. Belo Horizonte, Minas Gerais: Autêntica Editora. 2008.
- COX, M. I. P. & ASSIS-PETERSON, A. A. Ser/Estar professor de inglês no cenário da escola pública: em busca de um contexto eficaz de ensino/aprendizagem. In: *Polifonia – Revista de Linguagens: Universidade Federal de Mato Grosso*. Ano 5. No. 05. Cuiabá: EdUFMT, 2002.
- Estudo do PageGroup. *Mapeamento da proficiência do idioma inglês na América Latina*. Disponível em: < http://www.michaelpage.com.br/productsApp_br/PDF_MP/EnglishStudy.pdf > acesso em 07/03/2015.
- FROMM, G. O Uso de Corpora na Análise Linguística. *Revista Factus*, São Paulo, v. 1, n. 1, p. 69-76, 2003.
- FROMM, G. *Proposta para um modelo de glossário de informática para tradutores*. Dissertação de Mestrado. São Paulo: FFLCH/USP, 2002.
- GUANDALINI, E. O. *Técnicas de Leitura em Inglês: ESP – English For Specific Purposes: estágio I*. São Paulo: Textonovo, 2002.
- MUNHOZ, R. *Inglês Instrumental: estratégias de leitura: módulo I*. São Paulo: Textonovo, 2000.