

PESQUISAS E  
PERSPECTIVAS  
EM LINGUÍSTICA  
DE CORPUS

## **Série Espaços da Linguística de Corpus**

### *Editor*

Tony Berber Sardinha, Pontifícia Universidade Católica de São Paulo, Brasil

### *Corpo Editorial*

Ana Frankenberg-Garcia, ISLA, Portugal

Anise D Orange Ferreira, Universidade Estadual Paulista, Araraquara, Brasil

Deise Prina Dutra, Universidade Federal de Minas Gerais, Brasil

Diva Cardoso de Camargo, Universidade Estadual Paulista, São José do Rio Preto, Brasil

Eckhard Bick, Universidade do Sul da Dinamarca

Elisa Duarte Teixeira, Projeto Comet, Universidade de São Paulo, Brasil

Gladis Barcellos Almeida, Universidade Federal de São Carlos, Brasil

Guillermo Rojo, Universidade de Santiago de Compostela, Espanha

Heliana Mello, Universidade Federal de Minas Gerais, Brasil

Helmara Moraes, Consulado dos Estados Unidos da América, São Paulo, Brasil

Marcia Veirano Pinto, GELC, Pontifícia Universidade Católica de S.Paulo, Brasil

Maria Cecília Lopes, GELC, Pontifícia Universidade Católica de S.Paulo, Brasil

Maria José Bocorny Finatto, Universidade Federal do Rio Grande do Sul, Brasil

Mark Davies, Universidade Brigham Young, Estados Unidos da América

Mike Scott, Aston University, Reino Unido

Oto Vale, Universidade Federal de São Carlos, Brasil

Patricia Bertoli Dutra, GELC, Pontifícia Universidade Católica de S.Paulo, Brasil

Simone Sarmento, Universidade Federal do Rio Grande do Sul, Brasil

Stella Tagnin, Universidade de São Paulo, Brasil

Tania Shepherd, Universidade do Estado do Rio de Janeiro, Brasil

ANA MARIA T. IBAÑOS  
LÍVIA PRETTO MOTTIN  
SIMONE SARMENTO  
TONY BERBER SARDINHA  
(ORGANIZADORES)

PESQUISAS E  
PERSPECTIVAS  
EM LINGUÍSTICA  
DE CORPUS

---

**Dados Internacionais de Catalogação na Publicação (CIP)**  
**(Câmara Brasileira do Livro, SP, Brasil)**

Pesquisa e perspectivas em linguística de corpus /  
organizadores Ana Maria T. Ibaños...[et al.] . – Campinas, SP  
: Mercado de Letras, 2014. – (Série Espaços da Linguística de  
Corpus)

Outros organizadores: Lívia Pretto Mottin, Simone Sarmento,  
Tony Berber Sardinha.  
ISBN 978-85-7591-341-3

1. Análise linguística 2. Linguagem e línguas - Ensino auxiliado  
por computador 3. Linguagem e línguas – Estudo e ensino 4.  
Linguística – Metodologia 5. Linguística – Processamento  
de dados I. Ibaños, Ana Maria T.. II. Mottin, Lívia Pretto. III.  
Sarmiento, Simone. IV. Sardinha, Tony Berber. V. Série.

14-13450

CDD-410.285

**Índices para catálogo sistemático:**

1. Linguística de corpus : Análise linguística via computador :  
Linguística aplicada 410.285

capa egerência editorial: Vande Rotta Gomide  
foto de capa: Marina Meirelles Gomide  
preparação dos originais: Editora Mercado de Letras  
editoração: DPG Editora

1ª edição  
**FEVEREIRO/2015**  
IMPRESSÃO DIGITAL  
*IMPRESSO NO BRASIL*

---

Esta obra está protegida pela Lei 9610/98.  
É proibida sua reprodução parcial ou total  
sem a autorização prévia do Editor. O infrator  
estará sujeito às penalidades previstas na Lei.

---

## AGRADECIMENTOS

*Primeiramente, gostaríamos de agradecer à CAPES, Fapergs, Macmillan, STB, SBS, Cambridge University Press, DISAL, Consulado dos Estados Unidos (São Paulo), British Council e John Benjamins, pelo apoio e auxílio à realização do evento que deu origem a este livro. Somos também gratos às equipes da Faculdade de Letras e ao Programa de Pós-Graduação em Letras da Pontifícia Universidade Católica do Rio Grande do Sul por não terem medido esforços para sediar e garanti o sucesso do evento.*

*Agradecemos a todos que de alguma forma participaram do evento: organizadores, comissão científica, monitores, conferencistas e pesquisadores que contribuíram com a apresentação de seus trabalhos. De forma especial, agradecemos a todos os autores que nos enviaram seus trabalhos para publicação neste livro. Aos pareceristas anônimos pelas leituras cuidadosas, correções e sugestões, as quais contribuíram para a qualidade do material publicado.*

*Finalmente, agradecemos à Sylviane Granger e à John Benjamins Publishing Company por permitirem a publicação da versão traduzida do artigo The contribution of learner corpora to second language acquisition and foreign language teaching: A critical evaluation, originalmente publicado na obra Corpora and Language Teaching (Aijmer 2008), publicada pela mesma editora.*

*Simone, Tony, Livia e Ana*



## SUMÁRIO

|  |     |
|--|-----|
| PREFÁCIO .....   | 11  |
| APRESENTAÇÃO .....   | 13  |
| CORPORA DE APRENDIZES  |     |
| A CONTRIBUIÇÃO DE CORPORA DE APRENDIZES<br>ÀS ÁREAS DE AQUISIÇÃO DE SEGUNDA LÍNGUA<br>E ENSINO DE LÍNGUA ESTRANGEIRA:<br>UMA AVALIAÇÃO CRÍTICA ..... | 23  |
| Sylviane Granger   |     |
| PACOTES LEXICAIS EM CORPORA DE APRENDIZES .....  | 57  |
| Deise Prina Dutra, Tony Berber Sardinha  |     |
| BELC: BRAZILIAN ENGLISH LEARNER CORPUS .....   | 81  |
| Aline Pacheco  |     |
| CONSTRUÇÃO E CODIFICAÇÃO DE CORPUS   |     |
| BLOGS, AMAZÔNIA E A FLORESTA SINTÁ(C)TICA:<br>UM CORPUS DE UM NOVO GÊNERO? .....   | 123 |
| Cláudia Freitas, Diana Santos  |     |

TRATAMENTO DA AMBIGUIDADE DOS SEGMENTOS  
INTRODUZIDOS POR PREPOSIÇÃO –  
UMA ABORDAGEM LEXICAL . . . . . 151  
*Magali Sanches Duran, Sandra Maria Alúísio*

VARRA: UM SERVIÇO PARA A VALIDAÇÃO, AVALIAÇÃO  
E REVISÃO DE RELAÇÕES SEMÂNTICAS NO AC/DC . . . . . 199  
*Cláudia Freitas, Diana Santos, Hugo Gonçalo Oliveira,  
Violeta Quental*

AELIUS: UMA FERRAMENTA PARA ANOTAÇÃO  
AUTOMÁTICA DE CORPORA USANDO O NLTK. . . . . 233  
*Leonel Figueiredo de Alencar*

MINERANDO TWEETS . . . . . 283  
*Larissa Astrogildo de Freitas, Ulisses Brisolara Corrêa,  
Angélica Alves Fernandes*

#### QUESTÕES DE LINGUAGEM E LINGUÍSTICA APLICADA

PADRÕES LÉXICO-GRAMÁTICAIS NA ESPECIFICAÇÃO  
DE PROPÓSITO E RESULTADO EM ABSTRACTS DE  
ARTIGOS CIENTÍFICOS: APLICAÇÕES NO ENSINO  
DE EAP E NA CONSTRUÇÃO DE FERRAMENTAS DE  
SUPORTE À ESCRITA CIENTÍFICA. . . . . 303  
*Carmen Dayrell, Arnaldo Candido Jr., Mariana Curi,  
Stella Tagnin, Sandra Alúísio*

MEDIDAS DE COMPLEXIDADE TEXTUAL ENTRE  
TRADUÇÕES BRASILEIRAS E ORIGINAIS DE LITERATURA  
INGLESA: UM ESTUDO-PILOTO BASEADO EM CORPUS . . . 347  
*Bianca Pasqualini, Maria José Bocorny Finatto, Aline Evers*

A UTILIZAÇÃO DE UM CORPUS DE OPERAÇÕES  
AERONÁUTICAS (COPAER) PARA A DESCRIÇÃO DA  
LINGUAGEM DE ESPECIALIDADE DA AVIAÇÃO:  
SUBSÍDIOS PARA O ENSINO DE ESP . . . . . 373  
*Ana Eliza Pereira Bocorny*

|  |     |
|--|-----|
| CORPORA E OPERAÇÕES ENUNCIATIVAS:<br>UM ESTUDO SOBRE AS ADVERSATIVAS DO<br>PORTUGUÊS BRASILEIRO. . . . .   | 407 |
| <i>Marion Celli</i>  |     |
| ESTILO DE TRADUTORES: ESTUDO BASEADO<br>NO CORPUS HEART OF DARKNESS/(NO) CORAÇÃO<br>DAS TREVAS . . . . .   | 441 |
| <i>Célia Maria Magalhães, Carolina Pereira Barcellos</i>   |     |
| HIGH FREQUENCY ITEMS IN A CORPUS OF SITCOM<br>DISCOURSE: SOME DIFFERENCES BETWEEN PSEUDO<br>AND REAL CONVERSATION. . . . .                             | 469 |
| <i>Bárbara Malveira Orfanò</i>   |     |
| NOTAS METODOLÓGICAS PARA A ELABORAÇÃO<br>DE CORPORA DIGITAIS DE EXCERTOS DE PROSA<br>GREGA ANTIGA BASEADOS EM KEYWORDS PARA<br>FINS DIDÁTICOS. . . . . | 493 |
| <i>Anise A. G. D'Orange Ferreira</i>   |     |
| USO DE CORPORA NO ENSINO DE LÍNGUAS<br>ESTRANGEIRAS PARA PROFISSIONAIS DA ÁREA<br>DE PUBLICIDADE . . . . .   | 525 |
| <i>Cristina Mayer Acunzo</i>   |     |
| SOBRE OS AUTORES. . . . .  | 549 |



## PREFÁCIO

A Linguística de Corpus vem ganhando espaço no Brasil nas duas últimas décadas, cobrindo todos os campos de estudos da linguagem, desde a descrição das línguas naturais, aspectos sociais do uso da linguagem, estudos de tradução, além do ensino/aprendizagem de línguas materna e estrangeira.

Sua história ganhou um marco especial em 1999, com os esforços de alguns visionários ao organizarem a primeira edição do Encontro Brasileiro de Linguística de Corpus. Naquele ambiente generoso de discussões acadêmicas a respeito da interface computador e dados textuais de natureza empírica, começou a se fazer no país a difusão de pesquisas, intercâmbio de recursos e formação de parcerias entre pesquisadores e instituições.

Devido à necessidade de dar visibilidade à crescente produção em Linguística de Corpus no Brasil, o Encontro passou a ter periodicidade anual. O XIX Encontro, realizado na Universidade Católica em Porto Alegre em 2010 foi resultado desses esforços para a difusão dos estudos com e sobre corpora. Os melhores trabalhos dentre aqueles selecionados para apresentação no XIX Encontro estão reunidos no presente volume que, por sua vez, compõe o número dois da Coleção *Espaços da Linguística de Corpus*. Estes estudos oferecem um testemunho do retorno

sem volta dos dados empíricos ao cerne da Linguística feita no Brasil. Além disso, mostram o desenvolvimento e criatividade das pesquisas, a constante exploração da área como teoria, como método e como prática, e uma revisita às suas múltiplas possibilidades e às suas igualmente numerosas áreas de desafios.

As inúmeras possibilidades oferecidas pela Linguística de Corpus, focadas de forma tímida nas primeiras versões de nossas reuniões, se fazem cada vez mais numerosas, presentes e consistentes nos programas dos Encontros: Linguística de Corpus aparece junto à Tradução, à Literatura, à Linguística Aplicada, à Lexicografia, Terminologia, ao Ensino de Línguas Estrangeiras e Materna, Análise do Discurso e Estilística.

Nos Encontros e nos dois volumes da série publicados até a data de hoje, há uma aceitação tácita de que “a linguística de corpus é mais do que uma (...) metodologia para se estudar a língua, mas também um novo tipo de pesquisa, além de ser de fato, uma abordagem filosófica ao assunto” (Leech 1992, p. 106, grifos meus). Portanto, ao lado de trabalhos com conteúdo neo-Firthiano, aqui têm lugar igualmente proeminente as pesquisas baseadas em corpus. Não há lugar para divisões porque as pesquisas na área não podem esperar; precisam ocupar seus espaços acadêmicos no país (daí o nome da série).

O objetivo deste segundo livro é, pois, registrar os avanços feitos nas formas de compilar, etiquetar e processar textos digitais e as múltiplas possibilidades de enfoques linguísticos com o auxílio do computador. Através das reflexões sobre as questões-chave da Linguística de Corpus feitas neste volume, pode-se ver mais um esforço que certamente ajudará a alavancar a área no país.

*Tania M. G. Shepherd*

## APRESENTAÇÃO

*Livia Pretto Mottin*  
*Simone Sarmiento*

A partir dos anos 60, uma nova área da linguística, chamada Linguística de Corpus (LdC), conferiu um novo sentido à palavra *corpus*. Na LdC, um corpus é uma coleção de textos produzidos naturalmente na língua (em contraposição a textos induzidos e à língua da máquina), armazenados em formato eletrônico e com o intuito de serem alvo de investigações linguísticas. Através da utilização de coleções de textos naturais, a LdC cresceu consideravelmente nos últimos anos e vem impactando diversas áreas de pesquisa em linguística. Seu crescimento se deve não apenas ao seu caráter essencialmente empírico, mas à sua capacidade de gerar evidências inéditas sobre a língua, tais como frequência de palavras e palavras que tendem a co-ocorrer umas com as outras.

A LdC se ocupa da coleta criteriosa de porções de linguagem armazenadas em formato eletrônico com o propósito de servirem para investigações linguísticas. Assim, é uma forma empírica de estudo da língua. Por depender do uso de computadores, o surgimento do primeiro corpus eletrônico aconteceu

em um contexto histórico pouco favorável aos seus avanços. Os entraves tecnológicos existentes eram muitos e as ferramentas computacionais limitadas. Como na linguística moderna, o termo corpus é quase sinônimo do termo corpus em formato eletrônico (Mcenery e Wilson 2004 [1996]), pode-se imaginar as dificuldades enfrentadas na época para digitalizar os corpora e acessá-los através de computadores.

O primeiro corpus linguístico eletrônico, o *Brown University Standard Corpus of Present-day American English*, lançado nos anos 60, foi o marco do início dos trabalhos com corpora. Pouco menos de dez anos antes do lançamento do corpus *Brown*, Chomsky havia lançado seu livro *Syntactic Structures*, no qual divulgava o gerativismo e defendia uma visão racionalista da linguagem em oposição à abordagem empírica da LdC. Para Chomsky, o que interessava era o estudo da competência (as normas internalizadas que o falante sabe sobre a língua) e, segundo ele, os dados necessários para tal análise provinham da intuição do linguista que os buscava em sua mente por meio da introspecção (Mcenery e Wilson 2004[1996]; Berber Sardinha 2000). Dados empíricos seriam úteis apenas para a investigação do desempenho (o uso que os falantes fazem da língua) dos usuários da língua. A compilação desse corpus, um tanto quanto desafiadora para a época, e a mudança de paradigmas linguísticos ocorrida na época (o racionalismo predominando em relação ao empirismo) foram fatores determinantes que vieram a tornar o *Brown* uma referência na LdC. A partir de então, o desenvolvimento e aprimoramento de computadores e ferramentas utilizadas para a análise de corpora vêm permitindo e possibilitando progressos na área.

Os progressos alcançados na área nos últimos anos se dão devido a fatores como a alta capacidade que os computadores atuais têm de armazenar dados e o desenvolvimento de ferramentas capazes de manipular corpora com acurácia. A Internet

foi também um fator importante na história da LdC, à medida que através dela os textos não precisam mais ser digitalizados. Se textos disponíveis *online* atenderem às necessidades do linguista, podem ser facilmente retirados do ambiente virtual e armazenados em computadores para a compilação de corpora. Além de facilitar a coleta dos dados para a compilação de um corpus, a Internet proporciona ao linguista de corpus o acesso a uma variada gama de textos das mais diferentes fontes que abrange desde livros, jornais, revistas e periódicos de áreas específicas do conhecimento a bate-papos informais, por exemplo. Os progressos da LdC trazem consigo um aumento de interesse na área, o qual pode ser sentido nos eventos anuais ELC (Encontro de Linguística de Corpus) e EBRALC (Escola Brasileira de Linguística Computacional). Apesar de incipientes, esses eventos já contam com um número considerável de participantes todos os anos e de apresentações de trabalhos pertencentes a áreas como descrição de linguagens especializadas, tradução, aquisição de línguas, sociolinguística, entre outras.

A presente obra foi organizada para registrar trabalhos apresentados durante o IX Encontro de Linguística de Corpus e a IV Escola Brasileira de Linguística Computacional. O IX Encontro de Linguística de Corpus aconteceu imediatamente após a IV Escola Brasileira de Linguística Computacional, ambos na Pontifícia Universidade Católica do Rio Grande do Sul, em Porto Alegre. O sumário desta obra evidencia a multiplicidade de olhares e aplicações da LdC no contexto brasileiro e internacional. Os temas abordados compreendem pesquisas relacionadas à aquisição de segunda língua, linguística textual, tradução, linguagens especializadas e gramática. Os trabalhos contribuem para o entendimento de línguas como o português, o inglês e até o grego. É importante ressaltar que os trabalhos aqui apresentados passaram por rigorosa seleção por pares e posteriormente, por uma série de revisões, de forma a incorporar as sugestões

dos pareceristas. Este livro é um passo importante na direção de consolidar a história da Linguística de Corpus no Brasil e de partilhar com todos os interessados na área o trabalho de colegas que têm se dedicado à pesquisa em corpora.

O primeiro bloco de textos é intitulado *Corpora de Aprendizizes*. Até o início dos anos 90, nenhum esforço havia sido feito na tentativa de compilar um corpus de linguagem autêntica de aprendizes de inglês,<sup>1</sup> revelando uma lacuna no conhecimento sobre a produção desses aprendizes, dada a quantidade de aprendizes de inglês no mundo todo (Granger 1998, 2003). Em meados dos anos 90, acadêmicos passaram a reconhecer o valor dos corpora de aprendizes e das evidências que eles poderiam gerar para a descrição e o melhor entendimento da linguagem de aprendizes de línguas. Projetos foram então lançados com o intuito de preencher tal lacuna. Um dos destaques foi o processo de compilação do *International Corpus of Learner English* (ICLE), lançado por Sylviane Granger. Este campo da LdC coloca-se como uma nova perspectiva na abordagem de questões referentes à aquisição e aprendizagem de línguas, exercendo, através de suas descrições, impacto em áreas subjacentes como o ensino de LEs e a produção de material didático.

Dentre os três trabalhos publicados sob o escopo deste tema, o primeiro, *A contribuição de corpora de aprendizes às áreas de Aquisição de Segunda Língua e Ensino de Língua Estrangeira: uma avaliação crítica*, previamente publicado em inglês, na obra *Corpora and Language Teaching* (Aijmer 2008), é de autoria de Sylviane Granger, um dos nomes mais reconhe-

---

1. Granger (1998) faz menção à compilação de um corpus de aprendizes de inglês especificamente, pois, segundo ela, a língua inglesa foi a língua mais estudada sob a perspectiva da LdC e o primeiro corpus linguístico eletrônico, o corpus *Brown*, é um corpus de inglês. Portanto, se a inexistência de corpora de aprendizes já representava uma lacuna na LdC, a falta de um corpus de aprendizes de inglês representava uma lacuna ainda maior.

cidos mundialmente nas pesquisas sobre corpora de aprendizes. Sua influência pode ser atestada nos dois capítulos subsequentes: *Pacotes lexicais em corpora de aprendizes*, de Deise Prina Dutra e Tony Berber Sardinha; e *BELC: Brazilian English Learner Corpus*, por Aline Pacheco.

A seção seguinte inclui cinco trabalhos e trata da *Construção e Codificação de Corpora*. Conforme Santos (1998), os pesquisadores da área de linguística de corpus podem ser divididos em dois grandes grupos: os usuários e os compiladores de corpora. Os usuários visam a extrair informações a partir dos corpora já coletados. Os compiladores, por sua vez, preocupam-se principalmente com a criação, estruturação e anotação de corpora. Há também um terceiro grupo, que busca desenvolver ferramentas para que os usuários e compiladores de corpora, já citados, possam desenvolver seu trabalho. Os trabalhos apresentados nessa seção são os seguintes: *Blogs, Amazônia e a Floresta Sintá(c)tica: um corpus de um novo gênero?*, de Cláudia Freitas e Diana Santos; *Tratamento da ambiguidade dos segmentos introduzidos por preposição – uma abordagem lexical*, de Magali Sanches Duran e Sandra Maria Aluísio; *VARRA: um serviço para a Validação, Avaliação e Revisão de Relações semânticas no AC/DC*, de Cláudia Freitas, Diana Santos, Hugo Gonçalo Oliveira e Violeta Quental; *Aelius: uma ferramenta para anotação automática de corpora usando o NLTK*, de Leonel Figueiredo de Alencar; e *Minerando Tweets*, de Larissa Astrogildo de Freitas, Ulisses Brisolara Corrêa e Angélica Alves Fernandes.

Os capítulos finais desta obra tratam de *Questões de Linguagem e Linguística Aplicada*. Parece haver um crescente interesse na aplicação de pesquisas baseadas em corpus no ensino de línguas. Essa aplicação pode ser de duas formas: o uso direto de corpora com os aprendizes e o uso indireto. No uso direto de corpora em aula, os alunos “tornam-se pesquisadores e não meros ‘receptores da língua’” (Berber Sardinha 2010, p. 297).

Ainda segundo Berber Sardinha, o uso de corpora por aprendizes possibilita que eles assumam o controle de seu próprio aprendizado, definindo metas, levantando e observando dados, criando hipóteses e tirando suas próprias conclusões. Johns (1997) denominou esse tipo de metodologia de ensino de *Data Driven Learning* (DDL). Corpora podem também ser usados de forma indireta, através da elaboração de materiais baseados em linhas de concordância.

Além desses usos mais diretos, os corpora vêm sendo cada vez mais usados na elaboração de materiais didáticos. A LdC pode oferecer informações relacionadas a vocabulário, gramática, formalidade e informalidade, diferenças entre a linguagem escrita e falada, como as pessoas começam e terminam uma conversa, entre outros aspectos. Desta forma, estudos baseados em corpora podem sugerir os itens linguísticos e processos que serão mais provavelmente encontrados por usuários de uma língua e que, portanto, merecem mais investimento em termos de tempo. Para a utilização de um corpus na elaboração de material didático (livros, polígrafos ou exercícios) é necessário, primeiramente, decidir (no caso da língua inglesa) quanto ao tipo e variedade de inglês que servirá como base para a elaboração do material, uma vez que corpora diferentes apresentarão palavras diferentes e, frequentemente, diferentes usos e funções das palavras a serem ensinadas. A palavra *nice*, por exemplo, é uma das quinze palavras mais frequentes no inglês falado (McCarten 2007). Entretanto, ela é bastante rara no inglês acadêmico escrito, ocorrendo sempre em citações de literatura ou em entrevistas. Portanto, a escolha (ou a compilação) de um corpus pode afetar as palavras a serem incluídas nos materiais didáticos, assim como seus sentidos e usos.

Além dessas áreas, McEnery *et al.* (2006) mencionam aproximações da LdC com os estudos linguísticos diacrônicos, a pragmática, a semântica, a sociolinguística, a análise do dis-

curso crítica, a estilística e os estudos literários. Como visto, a LdC pode auxiliar na maioria das áreas da linguística. Nesta obra, questões dessa natureza são abordadas em oito capítulos: *Padrões Léxico-Gramaticais na especificação de propósito e resultado em abstracts de artigos científicos: aplicações no ensino de EAP e na construção de ferramentas de suporte à escrita científica*, de Carmen Dayrell, Arnaldo Candido Jr., Mariana Curi, Stella Tagnin e Sandra Aluísio; *Medidas de complexidade textual entre traduções brasileiras e originais de literatura inglesa: um estudo-piloto baseado em corpus*, de Bianca Pasqualini, Maria José Bocorny Finatto e Aline Evers; *A utilização de um corpus de operações aeronáuticas (COPAER) para a descrição da linguagem de especialidade da aviação: subsídios para o ensino de ESP*, de Ana Eliza Pereira Bocorny; *Corpora e operações enunciativas: um estudo sobre as adversativas do português brasileiro*, de Marion Celli; *Estilo de tradutores: estudo baseado no corpus Heart of Darkness/(No) Coração das Trevas*, de Célia Maria Magalhães e Carolina Pereira Barcellos; *High frequency items in a corpus of sitcom discourse: some differences between pseudo and real conversation*, de Bárbara Malveira Orfanò; *Notas metodológicas para a elaboração de corpora digitais de excertos de prosa grega antiga baseados em keywords para fins didáticos*, de Anise A. G. D'Orange Ferreira; e *Uso de corpora no ensino de línguas estrangeiras para profissionais da área de publicidade*, de Cristina Mayer Acunzo.

Acreditamos que todos os trabalhos aqui apresentados, respeitando suas diferentes vertentes teórico-metodológicas, são extremamente relevantes e contribuem para o nosso aperfeiçoamento e formação constantes.

Desejamos a todos uma ótima leitura!

## Referências

- AIJMER, K. (ed.) (2008). *Corpora and Language Teaching*. Benjamins.
- BERBER SARDINHA, Tony (2000). “Linguística de Corpus: histórico e problemática.” *DELTA*, São Paulo, vol. 16, n.º 2, pp. 323-367.
- GRANGER, Sylviane (2003). “The International Corpus of Learner English: A New Resource for Foreign Language Learning and Teaching and Second Language Acquisition Research.” *TESOL Quarterly*, vol. 37, n.º 3, pp. 538-546, autumn.
- GRANGER, Sylviane (1998). “The computer learner corpus: a versatile new source of data for SLA research”, in: GRANGER, Sylviane (ed.) *Learner English on Computer*. New York: Longman, pp. 3-18.
- JOHNS, T. (1997).” Contexts: the background, development and trialling of a concordance-based CALL program”, in: WILCHMANN, Anne; FLIGELSTONE, Steven; MCENERY, Tony and KNOWLES, Gerry (eds.) *Teaching and Language Corpora*. London: Longman, pp. 100-115.
- MCCARTEN, Jeanne (2007). *Teaching Vocabulary: Lessons from the corpus, lessons for the classroom*. New York: Cambridge University Press.
- MCENERY, Tony; XIAO, Richard and TONO, Yukio (2006/2007) *Corpus-Based Language Studies – An advanced Resource Book*. Oxon: Routledge.
- MCENERY, Tony and WILSON, Andrew (1996[2004]) *Corpus Linguistics: An introduction*. Edinburgh: Edinburgh University Press.
- SANTOS, Diana (1998/1999). “Disponibilização de corpora de texto através da WWW.” *Actas Workshop da APL Sobre Linguística Computacional*, 1, Lisboa. Lisboa: Colibri, pp. 323-346.