

PRÁTICAS ORAIS EM INGLÊS E LC: UTILIZANDO O VOICE

ORAL PRACTICES IN ENGLISH AND LC: USING VOICE

Edison Gomes Junior¹

Recebimento do Texto: 25/10/2022

Data de Aceite: 20/11/2022

RESUMO: esse artigo discute a Linguística de Corpus (LC) na área do ensino do Inglês Acadêmico (IA), a partir da proposição de uma sequência pedagógica baseada na ideia de gêneros orais acadêmicos, e inglês como língua franca (ILF), e partindo de uma referência linguística um corpus multimodal (oral e escrito) de não-nativos em contexto acadêmico, encontrado no VOICE (*Vienna-Oxford International Corpus of English*). Em nossa abordagem, algumas características operacionais desse corpus serão explicadas, e aplicações pedagógicas, sugeridas.

PALAVRAS-CHAVE: Linguística de Corpus. Ensino de Inglês. Prática Oral em Inglês. VOICE.

ABSTRACT: this article discusses Corpus Linguistics (CL) in the area of Academic English (AE) teaching, from the proposition of a pedagogical sequence based on the idea of academic oral genres and English as a lingua franca (ILF), and which departs from the linguistic reference of a multimodal corpus (oral and written) of non-natives in an academic context found in VOICE (*Vienna-Oxford International Corpus of English*). In our approach, some operational characteristics of this corpus will be explained, and pedagogical applications will be suggested.

KEYWORDS: Corpus Linguistics. English Teaching. Oral Practice in English. VOICE.

¹ Professor Doutor da Universidade do Estado do Mato Grosso (UNEMAT). e-mail: edigomes2000@uol.com.br

Introdução

Segundo Viana (2011, p.27), um corpus deve ser compreendido como um conjunto de textos, escritos e orais, selecionados com base em princípios claros e bem definidos, que são uma reprodução da produção linguística de toda a população que se quer investigar, ou uma amostra representativa dessa população, consistindo em uma forma de representar empiricamente o uso que se faz de uma língua em seu sentido geral ou específico. Tais conjuntos existem em forma eletrônica e podem ser utilizados para a pesquisa e o ensino. Coletando os corpora orais dos mais variados gêneros, a Linguística de Corpus (LC) atualmente trabalha com, e oferece, uma gama de textos falados que podem ser visualizados e estudados, propondo uma nova perspectiva do estudo da linguagem que é basicamente multimodal e multicultural, pois possibilita a visualização e a audição de textos de diferentes tipos de falantes da língua alvo, nativos e não-nativos.

Feak (2013) acredita que o cenário do campo dos estudos de Inglês para Necessidades Especiais (ou *English for Specific Purposes - ESP*), dentro do qual o Inglês Acadêmico (IA) se encontra, tem sido fortemente alterado devido à facilidade com que corpora falado pode ser criado. Mais e mais, os estudos de IA têm se debruçado sobre os gêneros orais produzidos em inglês no contexto acadêmico, pois é sabido que, muito além da leitura, da escrita de artigos e da compreensão de palestras, o estudante acadêmico precisa estar apto a produzir gêneros textuais orais nesse contexto, tais como os relacionados a apresentações, seminários e reuniões^{xiii}, ou mesmo em contextos sociais indiretamente ligadas à vida acadêmica, ou seja, tudo que se produz linguisticamente em outra língua, além dos tópicos acadêmicos.

Outro aspecto importante do ensino do IA relacionado à oralidade, é o fato de os encontros acadêmicos serem, em muitos casos, multilinguísticos, pois uma grande maioria dos falantes de inglês que atende tais eventos seja não-nativa, e de modo que a língua alvo seja utilizada a partir de “diferentes processos de internalização e desempenho, ou externalização”, que apontam para a característica internacional e “franca” do idioma alvo. Posto isso, propõe-se, nesse artigo, sugerir como a LC, a partir de um corpus oral, pode ser utilizada para prover um conjunto de referenciais discursivas orais que ocorre em um contexto acadêmico. Para a

nossa atividade didática, voltada para A com pouca experiência no idioma, e que poderíamos qualificar de A1, escolhemos como tema discursivo e introdutório “a apresentação pessoal”, que pode ser mais ou menos breve, e existe em contextos formais e informais.

Segundo O’Keffe, McCarthy e Carter (2007, p. 216), o discurso acadêmico oral, que é comparado ao discurso para negócios, possui características institucionais, tais como domínios irrealis de hipóteses e especulações, é dirigido por objetivos, presidido ou professor, discussão conduzida, etc. Esse tipo de discurso é também derivado da conversa do dia-a-dia, e possui características com a conversa banal da vida cotidiana, exibindo a orientação humana primária para a cortesia, convergência e relacionamentos bons e não ameaçadores, e ocorrendo em face de papéis institucionais sancionados hierarquicamente. Assim, enquanto muitos usuários da linguagem acadêmica falada usam o inglês como língua franca em contextos não-nativos, os intercâmbios e relacionamentos acadêmicos bem-sucedidos repousam, em última análise, na construção e manutenção de um bom relacionamento. Tais características, relacionadas a um letramento social acadêmico discursivo, devem ser discutidas, exploradas e praticadas, pois englobam importantes questões discursivas, como noções de formalidade, educação e empatia.

A partir dos ganhos pedagógicos da LC para o ensino de línguas, e da necessidade de se explorar o inglês acadêmico como língua internacional e franca, propomos uma sequência didática baseada em um corpus oral de falantes não-nativos, que será um corpus de referência para a sequência didática, e por meio do qual o aprendiz (A) vai poder ouvir, visualizar, discutir e praticar funções e vocabulários na língua alvo. Dentre os corpora pesquisados na Internet, o VOICE se mostrou o mais adequado para a nossa proposta pedagógica^{xiv}, pois oferece uma série de recursos para a pesquisa de contextos e gêneros, disponibiliza com facilidade as gravações de algumas interações, e fornece um corpus oral autêntico de não nativos de inglês, em diferentes contextos, “que utilizam a língua com sucesso em suas vidas pessoais, profissionais ou acadêmicas, podendo ser vistos menos como aprendizes do que como usuários do idioma” (VOICE, 2013)^{xv}.

A partir dessa abordagem pedagógica para A brasileiros, que une DDL (*data-driven learning*) com ILF, pretende-se, também, de um lado, demonstrar a

língua menos como modelo ideal produzido por um nativo ideal, do que língua “glocal”, ou seja, compartilhada e utilizada global e localmente (“glocalmente”), e internalizada de maneiras diferentes e por diferentes falantes; de outro, ao demonstrar diferentes “performances” da língua, reduzir o filtro afetivo do A, conscientizando-o da normalidade do sotaque, que não deve ser relacionado ao erro. Nesse caso, procura-se trabalhar também com a autoestima dos A, um fator preponderante na aquisição de uma segunda língua (KRASHEN, 1981, p.23). Para realizar a sequência pedagógica, composta de funções utilizadas em apresentações pessoais, partiu-se da ideia básica da função *my name is*, que foi pesquisada no corpus VOICE. O corpus utilizado para a análise e prática linguística é propositalmente pequeno devido ao nível dos alunos. Nesse sentido, seguimos Brown (2005), que acredita que “um corpus pequeno pode ser mais relevante pedagogicamente” (apud Boulton and Tyne, 2011).

Da mesma maneira, O’Keffe, McCarthy e Cartes (2007, p. 198) acreditam que pequenos corpora especializados podem levar a percepções que não podem ser obtidas facilmente a partir de grandes corpora gerais, ajudando a entender funcionamentos e distinções de linguagem e, diríamos, discursos. Segundo esses autores, os corpora especializados, como o discutido em nosso artigo, oferecem as vantagens de serem direcionados e possuírem dados que provavelmente representarão o domínio de destino mais fielmente; seus léxicos e estruturas especializadas tendem a ocorrer com padronização e distribuição mais regulares, mesmo com quantidades relativamente pequenas de dados; e seu uso e aplicação são mais fáceis de definir e delimitar em termos de objetivos pedagógicos.

O corpus VOICE dentro da sequência didática proposta

O VOICE compreende transcrições de interações face a face sem *script* que ocorrem naturalmente em inglês como língua franca (ILF). Atualmente o VOICE compreende 1 milhão de palavras de interações faladas, o que equivale a aproximadamente 120 horas de fala transcrita. Além disso, 23 gravações de eventos de fala transcritos também podem ser ouvidas. Os falantes gravados em VOICE são experientes de ILA, e pertencentes de uma ampla gama de origens de primeiro idioma. Até o momento, o corpus inclui aproximadamente 1250 falantes

de ELF com aproximadamente 50 primeiras línguas diferentes (desconsiderando as variedades das respectivas línguas). As interações gravadas cobrem uma gama de eventos de fala diferentes em termos de domínio (profissional, educacional, lazer), função (troca de informações, atuação de relações sociais) e papéis e relacionamentos dos participantes (conhecido x não conhecido, simétrico x assimétrico). Os tipos de eventos de fala são: entrevistas; conferências de imprensa; encontros de serviço; discussões de seminário; discussões do grupo de trabalho; discussões de *workshops*; encontros; painéis; sessões de perguntas e respostas e conversas.

Em uma busca, o corpus fornece o seguinte campo:

Figura 1: a página de busca e os domínios sócio-discursivos

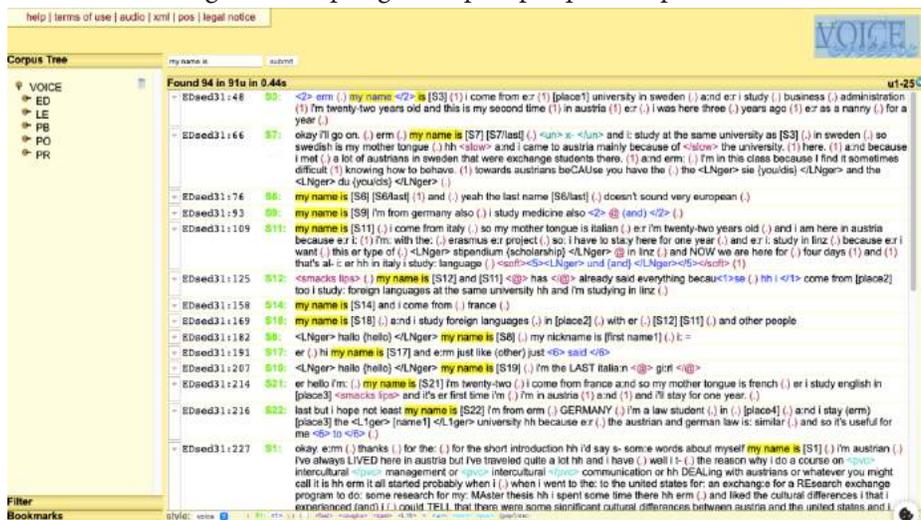


Fonte: *VOICE online*^{xvi}

As siglas à direita do campo informam cinco domínios, que abrangem diferentes contextos sociodiscursivos: O domínio educacional (ED), que abrange situações relacionadas as pessoas ou instituições envolvidas em ensino, treinamento ou estudo; o domínio do lazer (LE), que abrange situações relacionadas a escolhas de momentos e atividades desconectadas com trabalho e estudo; a dimensão dos negócios profissionais (PB), que abrange situações relativas à manufatura, compra, venda ou fornecimento de mercadorias ou serviços em troca de dinheiro; a dimensão da organização profissional (PO), que abrange situações relativas a atividades de redes ou organizações que não estejam fazendo pesquisa ou negócios e compra; e o domínio da pesquisa/ciência profissional (PR), que abrange situações relativas ao estudo criterioso de um assunto, especialmente com o intuito de descobrir novos fatos sobre ele.

Ao se fazer uma busca linguística no VOICE (no nosso caso, *my name is*) a ferramenta exhibe em que arquivos a palavra ou frase procurada ocorre:

Figura 2: corpus gerado pela pesquisa de palavras



No estilo VOICE de visualização dos dados, que pode ser modificado no rodapé da página, o campo principal oferece informações em três colunas. A coluna à esquerda localiza o resultado da pesquisa no corpus, listando a sua identificação (ID) e o número do evento de fala. Essas duas informações são separadas por dois pontos, como em EDsed25: 36). A coluna do meio fornece a ID do locutor no evento de fala transcrito, seguido por dois pontos: S1: ou S2: A terceira coluna, mais ampla, à direita, contém o enunciado real produzido pelo falante, com uma série de etiquetas em diferentes cores.

Esses marcadores apontam para diversas ocorrências linguísticas e prosódicas no texto oral de cada falante, como pausas, risadas, velocidades, alturas, pausas, entonação, etc., que podemos entender como marcadores discursivos linguísticos e paralinguísticos. O VOICE oferece a seguinte legenda:

Figura 3: legenda de marcadores discursivos

S1:, S2:, S3:, ...	Identified speakers
SS:	Group of speakers
SX:, SX-f., SX-m., SX-1:, ...	Speakers not identified
text?, text.	Intonation
TEXT	Emphasis
(.), (1), (3),...	Pauses
<1> </1>, <2> </2>, ...	Overlaps
=	Other-continuation
te:xt	Lengthening
tex-	Word fragments
@	Laughter
<@> </@>	Laughingly spoken
(text)	Uncertain transcription
<pvc> </pvc>	Pronunciation variations and coinages
<ono> </ono>	Onomatopoeic noises
<L1scc> </L1scc>, <LNfre> </LNfre>, ...	Non-English speech
<spel> </spel>	Spelling out
<fast> </fast>, <whispering> </whispering>, <imitating> </imitating>, ...	Speaking modes
hh, hhh	Breath
<coughs>, <applauds>, <clears throat>, ...	Speaker noises
[S1], [org1], [place1], [last name1], ...	Anonymization
{parallel conversation between S1 and S2 starts}, {telephone rings}, {S1 leaves the room}, ...	Contextual events
<un> xxxx </un>	Unintelligible speech
(gap 00:02:23), (nrec 00:50:00), ...	Transcription borders, untranscribed portions

Na busca da frase *my name is*, o VOICE mostra, na moldura amarelada do campo, à esquerda, os arquivos relacionados à busca, e aponta para uma concentração dessas estruturas na sigla ED, que marca o domínio educacional. O acesso a esse domínio mostra que alguns arquivos possuem um ícone de som, indicando que eles estão gravados. Assim, entre os arquivos EDsed 31 e os EDsed 251, preferiu-se o segundo:

No documento HEADER o tipo de encontro é descrito da seguinte maneira:

Figura 6: descrição do evento^{xvii}

Event Description
 Words: 15445
 This is a seminar discussion among European international students (with a predominance of law students) and people working in academic mobility contexts. The interaction is clearly guided by S1, who also delivers an initial presentation on the history of academic mobility in Europe and argues that Europe has always been integrated. The discussion centers around the usefulness and limitations of academic mobility and internationalism for the labor market. The speakers discuss different approaches to knowledge and education, the debates around the Bologna process and the introduction of bachelor's degrees in continental European academic programs. They weigh advantages and disadvantages of academic and professional mobility.

Finalmente, outras duas interessantes maneiras de visualização dos corpora, e que podem ser adotadas para o ensino são, além do estilo VOICE, os estilos PLAIN (simples) e o estilo KWIC (*Keyword in Context*, ou Palavra Chave em Contexto). O primeiro trata-se de uma versão reduzida da transcrição fornecida, que se torna mais legível, pois possui menos marcadores discursivos^{xviii}; o segundo mostra os resultados da pesquisa no centro da visualização, e seu contexto linguístico, ou cotexto:

Figura 7: o estilo PLAIN (simples) e o estilo KWIC



A partir dessas características, pode-se perceber o alcance pedagógico desse corpus, que combina o registro oral do inglês utilizado como Língua Franca por não-nativos, e as várias possibilidades de transcrição desse texto para a análise da língua e do discurso oral, incluindo fenômenos cotextuais e paralinguísticos. A partir daí, propomos uma sequência didática analítica e comunicativa.

Sequência didática

- AULA comunicativa para alunos de IA;
- Nível: A1(mas pode ser adaptado para A2);
- Tema: Apresentação pessoal em contexto acadêmico;
- Habilidades linguísticas mais trabalhadas: escuta e oralidade;
- Aspectos linguísticos que podem ser trabalhados: vocabulário (países e matérias acadêmicas); classes de palavras como verbos e tempos verbais (presente simples, presente progressivo e passado);
- Material: computador, papel, caneta, corpus de papel e/ou eletrônico
- Corpus: VOICE (domínio da educação – seminário)
-

Passo 1: Aquecimento (*warm up*)

Comentário: Como se trata de uma gravação de um seminário com estudantes de várias áreas de ensino que serão mencionadas, o professor (P) pode fazer o aquecimento da aula a partir de um exercício escrito ou oral de revisão de nomes de diferentes matérias acadêmicas, incluindo as que estarão no corpus. Em um nível mais básico, países também podem ser elicitados.

Passo 2: Contextualização e audição do material

Comentário: O P pode, e tratando o corpus escrito como um exercício de compreensão oral, pode trabalhar com as noções de pré, durante e pós-escuta. Como exercício de pré-escuta, pode-se discutir com a classe experiências educacionais fora do país, maneiras de se apresentar em uma sala no Brasil e em outro país, e a partir da descrição do evento disponibilizada pelo VOICE (figura 6),

contextualizar o texto oral a ser trabalhado. A partir daí, como exercício de escuta, os A ouvirão o material autêntico e anotarão os dados pessoais apresentados pelos falantes, tais como país de origem e formação acadêmica. Para o exercício oral, utilizamos apenas a parte da gravação que contem as apresentações dos falantes, cortando toda a introdução e explanação inicial do P, que pode ser utilizada em outro momento, já que é uma aula introdutória. Após a discussão e a correção do exercício, e como atividade de pós-escuta, o P pode propor um exercício aural de sensibilização no qual os A tentam ligar pronúncias a países, trabalhando com memórias sonoras de outras línguas, e a descrição dos alunos no VOICE (Figura 5), que abre espaço para um exercício comunicativo de leitura, do tipo quem é quem. Ainda no que tange à pós-escuta, o P pode, ainda, trabalhar com o estilo VOICE de visualização (figura 2) e discutir e sensibilizar os A para os aspectos paralinguísticos do texto, tais como hesitações, entoações, pausas, risos etc., ou propor comparações culturais e comportamentais. Segundo Mauranen (1996, p. 92), as gravações de diálogos falados mantêm um status de ‘observação não participante’, que é relevante para os A e revela-se como uma importante estratégia para aprender ou lidar com grupos de falantes da língua alvo. As gravações de interação falada que não foram originalmente produzidas para fins de didática da linguagem (e, portanto, eram autênticas ou genuínas para os falantes originais) têm uma forte reivindicação de autenticidade no sentido de “similaridade” para alunos em sala de aula, sejam elas pesquisadas com métodos de corpus ou outros meios.

Passo 3: Visualizando o corpus e reconhecendo estruturas linguísticas

Comentário: O P, após o exercício de escuta, pode agora mostrar o corpus ao aluno. Para essa etapa, uma vez que o corpus é pequeno, sugerimos um corpus de papel copiado do estilo *plain*, para facilitar a leitura. Outra importante característica do corpus de papel, é que ele pode ser editado pelo P, que pode arrumar, manter ou retirar os marcadores do texto. Assim, pode-se obter o seguinte material de trabalho^{xix}:

1. My name is X and I'm now studying medicine here in Vienna, that's all.
2. Well, my name is X I'm trying to finish my studies in law at the moment, University of X and work for the X parliament.
3. My name is X I studied civil engineering.
4. Well, my name is X I'm studying international business administration here at the u- university of X for economics business administration and I did an Erasmus term in X.
5. My name is X I'm from the University of X.
6. Okay, it's my turn, hello, my name is X. I'm from France and I'm doing an exchange in X and this year I'm writing my final year's er thesis mobility problems in Europe. Well I'll say more later. Thank you.
7. Hello, my name is X. I study educational science in Germany.
8. My name is X. I'm from Slovakia and I study law at the university of X.

A partir da visualização simplificada do corpus, o P pode explorar com mais profundidade, e indutivamente, o texto verbal oral, tais como os marcadores de início de turno, como *well*. Com o corpus de papel, o A pode marcar verbos e preposições, e discutir o uso dos tempos verbais, observando suas particularidades formais e discursivas^{xx}. Outra possibilidade de análise linguística, é partir da visualização do estilo KWIC, que nesse caso vai além do material trabalhado pelo A, e fornece um cotexto geral do uso da expressão, que pode ser utilizado para observação e discussão linguística:

Figura 8: o estilo PLAIN do texto oral estudado

erm	my name is	[S3] i come from er [place1] university in s
okay i'll go on erm	my name is	[S7] [S7/last] x- and i study at the same un
	my name is	[S6] [S6/last] and yeah the last name [S6/
	my name is	[S9] i'm from germany also i study medicine a
	my name is	[S11] i come from italy so my mother tongue i
	my name is	[S12] and [S11] has already said everything
	my name is	[S14] and i come from france
	my name is	[S18] and i study foreign languages in [plac
hallo	my name is	[S8] my nickname is [first name1] i
er hi	my name is	[S17] and erm just like other just said
hallo	my name is	[S19] i'm the last italian girl
er hello i'm	my name is	[S21] i'm twenty-two i come from france and s
last but i hope not least	my name is	[S22] i'm from erm germany i'm a law student
roduction hh i'd say s- some words about myself	my name is	[S1] i'm austrian i've always lived here in
e er of the lectures in in [place1] hh er so	my name is	[S1] [S1/last] i i studied history and law i
okay	my name is	[S3] [S3/last] erm i studied political scienc
	my name is	[S4] [S4/last] and i'm now studying medicin
well	my name is	[S5] [S5/last] i'm trying to finish my stud
	my name is	[S6] [S6/last] er i studied civil engineering
well	my name is	[S7] [S7/last] i'm studying international bus
hh erm	my name is	[S2] [S2/last] i'm from the university of [p
okay it's my turn hello	my name is	[S8] [S8/last] hh i'm from france and i'm doing
hhh er	my name is	[S9] [S9/last] ich komme aus der ukraine oop
hh hello	my name is	[S10] [S10/last] hh i study educational scien
	my name is	[S11] [S11/last] i'm from slovakia and i stud

Passo 4: Prática oral comunicativa 1(mais dirigida)

Comentário: A partir das discussões e práticas com o corpus do VOICE, exercícios orais podem ser propostos, nos quais os A receberão identidades novas, com nomes, nacionalidades e áreas ou projetos de estudo diferentes, que podem ser retirados do corpus analisado, e poderão simular uma apresentação nos moldes do corpus. Exemplo de instrução oral: Name: Yuri / Country: Germany / Study area: computer science

Passo 4a: Prática oral comunicativa 2 (menos dirigida).

Comentário: Pretende-se fazer uma transferência para o mundo do A, que preparará a realizará uma apresentação oral nos moldes apresentados, na qual deverão utilizar seus nomes, nacionalidades e áreas de estudo verdadeiros. O tipo de informação que o A apresenta sobre si mesmo pode ser aumentado tanto a partir do corpus estudado, pois os falantes também mencionam outros assuntos ao se apresentarem (a idade, por exemplo), como também a partir de outros corpora: o BNC falado (*spoken*), por exemplo, apresenta uma série de apresentações pessoais que vão além do escopo acadêmico.

Conclusão

Esperamos ter demonstrado como a área de LC, baseada no estudo da linguagem mediado por computador, oferece uma série de possibilidades de práticas pedagógicas da língua inglesa, que não são necessariamente relacionadas ao inglês avançado e nativo, ou à gramática e tradução, e são de ordem multimodal. Partindo da ideia de um gênero oral acadêmico básico e do VOICE, foi possível delimitar um pequeno corpus oral que pode ser utilizado no ensino de IA, adequado à ideia de ILF e Internacional, por suas várias características: ser autêntico, básico, prático, multimodal, falado por não-nativos, e que disponibiliza marcadores paralinguísticos e contextuais que podem ser utilizados para aprofundamento e discussão da modalidade oral da linguagem e outros aspectos. Além disso, a estratégia pedagógica pretendeu reduzir o filtro afetivo do A, apresentando diferentes versões da língua inglesa, internalizada por não-falantes. De um lado, o corpus dispensa os diálogos mais artificiais dos livros didáticos, sendo uma fonte rica de informações linguísticas que podem ser exploradas em todos os níveis da análise linguística, do fonético ao discursivo, abrindo espaço para o inglês a partir do ponto da diversidade de falantes; de outro lado, ele pode ser utilizado conjuntamente com outros corpora e concordanciadores, sendo um material dinâmico que aponta para a LC como possibilidade de inclusão em metodologias ativas de ensino e aprendizagem.

Referências

BOLTON, A.; TYNE H. Corpus linguistics and data-driven learning: a critical overview. Henry Tyne (ed.). **Vals-alsa**. Bulletin Suisse de linguistique appliquée. Número 97. 2013. p.97-118.

FEAK, Christine B. ESP and Speaking. In: **The Handbook of English for Specific Purposes** (ed) Brian Paltridge, Sue Starfield. 2013. p. 35-53.

MAURANEN, Anna. Spoken corpus for an ordinary learner. In: **How to use corpora in language teaching** (Edited by John Sinclair). Amsterdam / Philadelphia: John Benjamins Publishing Company, 1996. p. 89-105.

KRASHEN, Stephen. **Second Language Acquisition and Second Language Learning**. Oxford: Pergamon Press Inc, 1981.

O'KEFFE Anne; MACARTHY, Michael; CARTES Ronald. **From Corpus to Classroom: language use and language teaching**. Cambridge University Press; 2007.

VIANA, Vander. Linguística de corpus: Conceitos, técnicas & análises. In: **Corpora no ensino de línguas estrangeiras**. HUB Editorial, São Paulo, 2011. pp. 25-95.

VOICE. 2013. **The Vienna-Oxford International Corpus of English** (versão 2.0 XML). Director: Barbara Seidlhofer; Researchers: Angelika Breiteneder, Theresa Klimpfinger, Stefan Majewski, Ruth Osimk-Teasdale, Marie-Luise Pitzl, Michael Radeka. Disponível em <https://www.univie.ac.at/voice/>.