

**PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO
PUC-SP**

Cristina Mayer Acunzo

**USO DE *CORPORA* PARA O ENSINO DE LÍNGUA INGLESA PARA
PROFISSIONAIS DE PUBLICIDADE**

MESTRADO EM LINGUÍSTICA APLICADA E ESTUDOS DA LINGUAGEM

**SÃO PAULO
2012**

PUC-SP

Cristina Mayer Acunzo

**USO DE *CORPORA* PARA O ENSINO DE LÍNGUA INGLESA PARA
PROFISSIONAIS DE PUBLICIDADE**

MESTRADO EM LINGUÍSTICA APLICADA E ESTUDOS DA LINGUAGEM

Dissertação apresentada à Banca Examinadora da Pontifícia Universidade Católica de São Paulo, como exigência parcial para obtenção do título de Mestre em Linguística Aplicada e Estudos da Linguagem, sob a orientação do Prof. Dr. Antonio Paulo Berber Sardinha.

**SÃO PAULO
2012**

BANCA EXAMINADORA

Aos meus pais
Ao Danilo

AGRADECIMENTOS

Ao Prof. Dr. Tony Berber Sardinha, pela dedicação ao ensino, pela paciência e pelo incentivo. Pelo exemplo excepcional de pesquisador. Por ter me dado oportunidades de participar ativamente do mundo acadêmico.

À Capes, pela concessão da bolsa.

Aos professores do LAEL, em especial à Prof^a Dr^a Rosinda Guerra Ramos, por me ajudar a me aprofundar na área de Ensino de Línguas, pelo carinho e confiança. À Prof^a Dr^a Maximina Freire, por me apresentar ao mundo da Complexidade.

Às dedicadas colaboradoras do LAEL Maria Lúcia Reis e Márcia Martins pelo auxílio e direcionamento quanto às questões burocráticas e pela ajuda nas pesquisas no CEPRIL.

Ao GELC e grupo de orientandos do Prof. Dr. Tony Berber Sardinha, pelo apoio e pelas contribuições que muito acrescentaram a este trabalho, dentre os quais, os de maior convivência: Carlos H. Kauffmann, Denise Delegá Lúcio, Eduardo Cassimiro, Patrícia Bértoli Dutra, Rosana de Barros Silva e Teixeira. À Telma de Lurdes São Bento Ferreira, pela leitura e revisão. À Maria Cecília Lopes, pelas ricas contribuições na qualificação e pelo apoio constante dentro e fora do LAEL.

À minha tutora e amiga Renata Condi de Souza, por me mostrar o caminho do LAEL, pelo direcionamento, pela leitura e contribuições valiosas na qualificação e pelos momentos de descontração.

À querida Marcia Veirano Pinto, por plantar a semente da mudança de paradigmas e do caminho da pesquisa. Pelas contribuições valorosas e confiança constante, pelas conversas virtuais, pela amizade.

A todos os meus companheiros de trabalho, coordenadores e professores, pela troca, e que ao longo desses quinze anos me ajudaram a entender a difícil tarefa de ser professor de língua estrangeira. Aos meus alunos, pela confiança e disposição a experimentar o novo.

À minha tão querida amiga Patricia Mozelli, pelas ideias irresistíveis e a quem eu tenho seguido na busca pelo crescimento profissional.

A todos os meus familiares, unidos pelo sangue ou não, que sempre torcem por mim.

Aos meus pais, com quem sempre pude, posso e poderei contar, pelo amor incondicional e incessante incentivo aos estudos e ao meu crescimento pessoal.

Ao meu marido Danilo, pela compreensão e apoio incondicionais, mesmo quando me encontrava no quarto do fundo, atrás de uma pilha de livros e papéis digitando. Por ouvir pacientemente minhas ideias sobre *corpus*. Pelo amor constante e por sempre acreditar no meu potencial.

A Deus, pela benção da vida e por propiciar caminhos para minha evolução.

RESUMO

Esta pesquisa teve como objetivo principal a extração dos padrões lexicogramaticais de um *corpus* composto por textos da área de Publicidade para a preparação de atividades de ensino de Inglês como Língua Estrangeira com *corpora* para estudantes do idioma que atuam nessa área.

Para tanto, encontramos suporte teórico principal na Linguística de *Corpus*, uma área que proporciona a pesquisa, o estudo e a exploração da língua em uso (BERBER SARDINHA, 2004; 2009; 2011); em especial, na sua vertente que trata da interface *corpus* e ensino de Língua Estrangeira. Assim, servem de embasamento teórico para a pesquisa as abordagens desenvolvidas por Sinclair (1991; 2004a; 2004b), Johns (1991), Willis (1990), Berber Sardinha (2009; 2011) e Willis e Willis (2007), que dizem respeito à exploração de *corpora* no ensino por meio da identificação e da análise de padrões lexicogramaticais da língua e ao ensino por meio da abordagem lexical.

Os resultados da análise do *corpus* sugerem que existe uma grande variedade de padrões de linguagem relativos a Publicidade a serem estudados. Encontramos diversos padrões e colocados que, se bem explorados por professores e alunos, podem auxiliar na comunicação dos profissionais dessa área. Propomos, assim três atividades de ensino de Língua Inglesa com esses padrões.

A pesquisa pretendeu ter feito uma contribuição original para o desenvolvimento de atividades de ensino baseadas em *corpora*, bem como para estudantes que atuam na área de Publicidade, para a qual existe pouco material didático no mercado.

Palavras-chave: Linguística de *Corpus*, Ensino de Língua Inglesa como Língua Estrangeira

ABSTRACT

This study aimed at identifying the lexicogrammatical patterns of a *corpus* of texts about Advertising for the development of activities for the teaching of English as a foreign language with *corpora* to professionals who work in this area.

In order to achieve this aim, the main theoretical framework for the research is provided by Corpus Linguistics, which is an area that supports the research and the study of the language in use (BERBER SARDINHA, 2004; 2009, 2011); this study is mainly concerned with the interface between *corpus* and the teaching of foreign languages. Therefore, it is also based on the approaches developed by Sinclair (1991; 2004a; 2004b), Johns (1991), Willis (1990), Berber Sardinha (2009; 2011) and Willis e Willis (2007), which refer to the exploration of *corpora* in language teaching through the identification and analysis of lexicogrammatical patterns and the teaching based on a lexical approach.

The results of the analysis of the *corpus* showed that there is a great variety of language patterns related to Advertising to be studied. We have found a range of patterns and collocates which, if adequately explored by both teachers and students, can help the professionals in this field communicate effectively. We proposed, then, three activities for the teaching of English with those patterns.

The study presented here intended to have made an original contribution to the development of classroom activities based on *corpora*, as well as to professionals in Advertising, for whom there is few teaching materials available.

Keywords: Corpus Linguistics, Teaching of English as a Foreign Language

SUMÁRIO

INTRODUÇÃO	13
CAPÍTULO 1: FUNDAMENTAÇÃO TEÓRICA	20
1.1 Linguística de <i>Corpus</i>	20
1.1.1 Histórico e aplicação	20
1.1.2 <i>Corpora</i>	26
1.1.2.1 Definição e exemplos	26
1.1.2.2 Tipologia.....	27
1.1.2.3 Construção, extensão e representatividade.....	30
1.1.3 Conceitos na Linguística de <i>Corpus</i>	32
1.2 Ensino de Língua Inglesa.....	37
1.2.1 Linguística de <i>Corpus</i> e Ensino de Língua Inglesa	43
1.2.1.1 Propostas para uso de corpora no ensino de Língua Inglesa e no desenvolvimento de material didático	49
1.2.1.2 Concordâncias na sala de aula	59
1.2.1.3 Uma proposta para o ensino de ILE <i>in company</i>	64
CAPÍTULO 2: METODOLOGIA DE COLETA E ANÁLISE DOS DADOS	70
2.1 Coleta dos dados.....	70
2.1.1 <i>Corpus</i> de estudo	70
2.1.2 <i>Corpus</i> de referência.....	72
2.2 Procedimentos de análise dos dados.....	73
2.3 Análise dos dados	80
2.3.1 Análise de <i>ad</i>	83
2.3.1.1 Análise de <i>ad spending</i>	85
2.3.1.2 Análise de <i>ad network(s)</i>	90

2.3.1.3	Análise de <i>ad revenue</i>	96
2.3.1.4	Análise de <i>ad campaign(s)</i>	98
2.3.1.5	Análise dos verbos colocados de <i>ad</i>	101
CAPÍTULO 3: DESENVOLVIMENTO DAS ATIVIDADES		103
3.1	Contexto	105
3.2	Atividade 1	106
3.3	Atividade 2	116
3.4	Atividade 3	120
CONSIDERAÇÕES FINAIS		127
REFERÊNCIAS BIBLIOGRÁFICAS		131
ANEXOS		139

LISTA DE FIGURAS

Figura 1:	Trecho de <i>A Concordance to Shakespeare</i> com a palavra <i>welcome</i>	22
Figura 2:	Tela do <i>WordList</i> do <i>Wordsmith Tools</i> com resultado estatístico do <i>corpus</i> de estudo	74
Figura 3:	Tela do <i>WordList</i> do <i>Wordsmith Tools</i> com resultado parcial de lista de palavras do <i>corpus</i> de estudo em ordem alfabética	75
Figura 4:	Tela do <i>WordList</i> do <i>Wordsmith Tools</i> com resultado parcial de lista de palavras do <i>corpus</i> de estudo por frequência.....	76
Figura 5:	Tela do <i>KeyWords</i> do <i>Wordsmith Tools</i> com resultado parcial de chavidade (<i>keyness</i>) da comparação entre a lista de palavras do <i>corpus</i> de estudo e da lista de palavras do <i>corpus</i> de referência.....	77
Figura 6:	Tela do <i>Concord</i> , aba <i>collocates</i> , do <i>Wordsmith Tools</i> com os colocados de <i>ad</i>	78
Figura 7:	Tela do <i>Concord</i> , aba <i>patterns</i> , do <i>Wordsmith Tools</i> com os padrões de <i>ad</i>	79
Figura 8:	<i>Slide</i> 1 da atividade 1	107
Figura 9:	<i>Slide</i> 2 da atividade 1	107
Figura 10:	<i>Slide</i> 3 da atividade 1	108
Figura 11:	<i>Slide</i> 4 da atividade 1	108
Figura 12:	<i>Slide</i> 5 da atividade 1	109
Figura 13:	<i>Slide</i> 6 da atividade 1 com concordâncias com os nódulos <i>ad</i> , <i>ads</i> e <i>advertising</i> do <i>ad corpus</i>	110
Figura 14:	<i>Slide</i> 7 da atividade 1 – resultado da pesquisa feita no COCA (<i>Corpus of Contemporary American English</i>) sobre a frequência de ocorrência dos verbos <i>run</i> , <i>place</i> e <i>put</i> com os nódulos <i>ad</i> , <i>ads</i> e <i>advertising</i>	110
Figura 15:	<i>Slide</i> 8 da atividade 1	111
Figura 16:	<i>Slide</i> 9 da atividade 1 – resultado da pesquisa feita no COCA (<i>Corpus of Contemporary American English</i>) sobre as ocorrências em diferentes tipos de textos dos verbos <i>run + ad</i> , <i>run+ ads</i> , <i>place + ad</i> e <i>place + ads</i>	111
Figura 17:	<i>Slide</i> 10 da atividade 1	112
Figura 18:	<i>Slide</i> 11 da atividade 1	112

Figura 19: Tela da atividade 1 do slide 11 – tela do <i>Compleat Lexical Tutor</i> com as palavras frequentes do texto usado na atividade 1	113
Figura 20: Tela da atividade 2 do slide 11 – tela do COCA (<i>Corpus of Contemporary American English</i>) com resultados da pesquisa sobre <i>at the same time</i>	113
Figura 21: Slide 12 da atividade 1	114
Figura 22: Slide 13 da atividade 1	115
Figura 23: Slide 1 da atividade 2	116
Figura 24: Slide 2 da atividade 2	116
Figura 25: Slide 3 da atividade 2	117
Figura 26: Slide 4 da atividade 2	118
Figura 27: Slide 5 da atividade 2 - Exemplo de como encontrar um texto que contenha as palavras ou padrões para desenvolver a atividade usando a ferramenta <i>Concordance Plot</i> do programa <i>AntConc</i>	118
Figura 28: Slide 6 da atividade 2	119
Figura 29: Slide 1 da atividade 3	120
Figura 30: Slide 2 da atividade 3	121
Figura 31: Slide 3 da atividade 3	121
Figura 32: Slide 4 da atividade 3	122
Figura 33: Slide 5 da atividade 3	123
Figura 34: Slide 6 da atividade 3	123
Figura 35: Slide 7 da atividade 3	124
Figura 36: Slide 8 da atividade 3	124
Figura 37: Slide 9 da atividade 3	125

LISTA DE TABELAS

Tabela 1:	Resultados da extração de palavras-chave – 25 primeiras em ordem de chavicidade (<i>keyness</i>)	81
Tabela 2:	Colocados principais de <i>ad</i> em ordem de frequência	83
Tabela 3:	Colocados principais de <i>ad</i> em ordem de IM	84
Tabela 4:	Colocados de <i>ad</i> com frequência na posição R1	84
Tabela 5:	Colocados de <i>ad spending</i>	85
Tabela 6:	Colocados de <i>ad network</i>	91
Tabela 7:	Colocados de <i>ad networks</i>	91
Tabela 8:	Colocados de <i>ad revenue</i>	96
Tabela 9:	Colocados de <i>ad campaign</i>	99
Tabela 10:	Colocados de <i>ad campaigns</i>	99
Tabela 11:	Verbos colocados de <i>ad</i>	102

LISTA DE QUADROS

Quadro 1:	Linhas de concordância de <i>ad</i> , <i>ads</i> e <i>advertising</i> do <i>corpus</i> de estudo ...	61
------------------	--	----

LISTA DE ESQUEMAS

Esquema 1:	Procedimentos e ferramentas usados na pesquisa	103
Esquema 2:	Primeiro ciclo da Tarefa 1	125
Esquema 3:	Segundo ciclo da Tarefa 1.....	126
Esquema 4:	Terceiro ciclo da Tarefa 1	126

INTRODUÇÃO

O ensino de Línguas Estrangeiras para profissionais de áreas específicas é um desafio, pois existe pouco material disponível no mercado que atenda às necessidades dos alunos de comunicar-se em seu meio profissional. A minha experiência profissional, como professora de Inglês como Língua Estrangeira, em aulas particulares e em empresas (*in company*), especialmente em agências de Publicidade, levou-me a observar que esse contexto de ensino possui algumas especificidades e características diferentes das de um curso regular em uma escola de idiomas; conseqüentemente, o conteúdo e o material usado nessas aulas também precisam ser adequados.

Aulas *in company* são comumente custeadas pela empresa, a qual espera que o funcionário esteja capacitado a usar o idioma em algum momento em seu meio profissional. Isso requer que o material usado em aula esteja adequado às necessidades do aluno no que concerne a essa comunicação. Em nossa experiência, pudemos verificar que os alunos normalmente conhecem os termos usados em sua área profissional, porém apresentam dificuldades no uso desse vocabulário específico. É por meio de materiais autênticos¹, como textos e vídeos, que os alunos podem ter contato com essa língua em uso; observamos que os materiais didáticos disponíveis no mercado proporcionam muito pouco contato com esse tipo de material.

Diante desse panorama e de nosso contexto de ensino de Inglês como Língua Estrangeira, doravante ILE, em agências de Publicidade, buscamos realizar esta pesquisa a fim de estudar a linguagem usada no meio profissional em que os alunos estavam inseridos para ajudá-los a comunicar-se nesse meio. Assim, este trabalho advém da minha prática como professora de ILE e da minha reflexão como aluna do curso de extensão ‘Ensinando Inglês de Verdade por meio de Filmes e Programas de TV em DVD’, coordenado pelo Prof. Dr. Tony Berber Sardinha, na COGEAE². Conhecendo a Linguística de *Corpus* e usando suas ferramentas em nossas aulas, percebemos que se trata de uma área que auxilia grandemente nesse processo, visto que ela ocupa-se do estudo da língua em uso e permite a análise dessa linguagem em textos de áreas específicas.

¹ Consideramos material autêntico o que não foi produzido ou criado para fins pedagógicos. Detalharemos essa questão na seção 1.2.1.3.

² COGEAE – PUC-SP – Coordenadoria Geral de Especialização e Aperfeiçoamento e Extensão.

O primeiro objetivo deste trabalho foi o de extrair e analisar o uso de alguns dos padrões lexicogramaticais mais frequentes no meio Publicitário; o segundo objetivo foi o de utilizar esses padrões na preparação de atividades³ de ensino de ILE para profissionais dessa área. Entendemos por padrão lexicogramatical a combinação de duas ou mais palavras que ocorre com frequência com um significado associado a ela (BERBER SARDINHA, 2004).

Para atingir o objetivo proposto, desenvolvemos a seguinte metodologia de pesquisa: (1) coleta de um *corpus*⁴ de mais de 1 milhão de palavras composto por artigos escritos e transcrição de vídeos com textos lidos e entrevistas de uma revista digital de Publicidade, que chamamos de *ad corpus*; (2) análise do *corpus* e sua comparação com um *corpus* de referência, o BNC (*British National Corpus*) para a identificação dos padrões lexicogramaticais distintivos do *corpus* de estudo, por meio do programa *Wordsmith Tools* (SCOTT, 2008); (3) transposição dos achados para as atividades de ensino, ilustrando como preparar aulas e desenvolver tais atividades.

Encontramos suporte teórico e metodológico principais para a pesquisa na Linguística de *Corpus*, doravante LC, que, dedicando-se à coleta criteriosa e exploração de conjuntos de dados linguísticos textuais, denominados *corpora*, proporciona a pesquisa, o estudo e a exploração da língua em uso, de forma empírica e por meio do uso de computador (BERBER SARDINHA, 2004). No Brasil, sabemos que a LC ainda é jovem, como apontado em um levantamento realizado por Berber Sardinha e Almeida (2008). Pretendemos, com nosso trabalho, ajudar a consolidar a área no país, na sua vertente que trata da interface *corpus* e sala de aula de Língua Estrangeira.

Dessa forma, o trabalho aqui proposto fundamenta-se na área de pesquisa baseada em *corpus* que se preocupa com o ensino de Línguas Estrangeiras. Uma das questões principais da LC nesta área é como incorporar o estudo e a análise da língua por meio de *corpora* no ensino-aprendizagem de línguas. Os trabalhos desta área com relevância para a pesquisa são de Sinclair (1991; 2004a; 2004b), Johns (1991), Willis (1990), Berber Sardinha (2009; 2011) e Willis e Willis (2007).

Sinclair (1991; 2004a; 2004b) e Johns (1991) apresentam formas de se trabalhar os conceitos e as ferramentas da LC no ensino por meio de concordâncias, que

³ Entendemos atividade como “uma sequência de trabalhos feitos pelos alunos na sala de aula, individualmente ou não, veiculados em um material de ensino” (BERBER SARDINHA, 2011, p. 307).

⁴ *Corpus* refere-se à coletânea de textos em formato eletrônico coletados de forma criteriosa a fim de servir para a pesquisa de uma língua ou variedade linguística (BERBER SARDINHA, 2004). O plural de *corpus* é *corpora*.

são “listagens de uma palavra específica (o ‘nódulo’, *node word* ou *search word*) juntamente com parte do texto onde ocorreu” (BERBER SARDINHA, 2009, p. 9). Johns é o criador do *Data Driven Learning* (DDL), abordagem na qual o aluno é incentivado a estudar os padrões lexicogramaticais da língua por meio de concordâncias, desenvolvendo, assim a habilidade de um pesquisador. Cabe ao professor guiar o aluno, facilitando esse trabalho de formular hipóteses sobre a língua, buscar respostas para suas questões, identificar padrões recorrentes e fazer generalizações a respeito dos achados.

O *Lexical Syllabus*, desenvolvido por Dave Willis (1990), é uma das propostas que aplica diretamente os conceitos da LC no ensino de Língua Estrangeira. O Currículo Lexical, assim traduzido por Berber Sardinha (2004), utilizou a descrição de um *corpus* da Língua Inglesa, a princípio coletado para a produção do dicionário *Cobuild*, para o desenvolvimento de uma série de livros didáticos. Os conceitos dessa proposta que usamos nesta pesquisa são o da exploração dos itens lexicogramaticais mais comuns da língua e o do desenvolvimento de atividades baseadas no léxico.

Berber Sardinha (2009; 2011) também apresenta maneiras de se usar *corpora* em sala de aula por meio de concordâncias; o autor propõe três tipos de atividades para o ensino de ILE, quais sejam: centradas na concordância, centradas no texto e multimídia/multigênero. As atividades centradas na concordância focam na identificação dos padrões lexicogramaticais e são sugeridos exercícios acompanhando as concordâncias, o que propomos em nossas atividades. O segundo tipo de atividade parte de um texto e trabalha as questões que o envolvem, como tema, gênero, ideias principais, além de trazer concordâncias com palavras do texto. Da mesma maneira, propomos atividades em que os alunos usam o texto, não só concordâncias. As atividades multimídia/multigênero envolvem diferentes mídias e gêneros e incentivam o aluno a entender as diferenças entre eles e seus padrões frequentes ou típicos. Usamos a ideia de atividade multimídia em nossa proposta.

Para a elaboração de uma das atividades, baseamo-nos ainda na abordagem desenvolvida por Dave Willis e Jane Willis (2007) intitulada Ensino baseado em Tarefas (*Task-based Teaching*). Segundo os autores, “a forma mais eficaz para ensinar uma língua é envolvendo alunos com o uso real da língua em sala de aula”⁵ (WILLIS E

⁵ “the most effective way to teach a language is by engaging learners in real language use in the classroom” (WILLIS e WILLIS, 2007, p. 1).

WILLIS, 2007, p. 1)⁶. A maneira que eles propõem para que isso aconteça é por meio de tarefas⁷ que levem os alunos a usar a língua em discussões, resolução de problemas, jogos, entre outros. O foco principal das tarefas não está em o aluno usar a língua somente de forma correta, mas sim com significado. O segundo foco da tarefa é na linguagem que o aluno irá usar para comunicar suas ideias de acordo com a proposta da tarefa. Por último, há o foco na forma, que pode ser feito de acordo com a linguagem usada pelos alunos, ou seja, o professor pode trabalhar suas dificuldades ou direcionar sua atenção para alguma forma relacionada à linguagem usada no contexto da tarefa.

No Brasil, as pesquisas na área da aplicação de *corpus* no ensino vêm crescendo, como podemos observar em dissertações, teses, artigos e congressos⁸ em diversas universidades brasileiras nas quais a LC está presente. Nossa pesquisa insere-se no grupo de pesquisa GELC/CNPq⁹, que tem como objetivos principais consolidar e divulgar a LC no país, promover o intercâmbio com grupos de pesquisa nacionais e internacionais, integrar a LC com outras áreas do saber, criar e disponibilizar *corpora* à comunidade nacional e internacional e tornar a LC mais acessível a todos os interessados.

Na universidade em que a pesquisa foi desenvolvida, a LC está dentro do Programa de Pós-Graduação em Linguística Aplicada. Acreditamos que nossa pesquisa corresponda à Linguística Aplicada, visto que entendemos essa área estudar a língua em uso nas práticas sociais (ALMEIDA FILHO, 2008; MOITA LOPES, 2008) e nossa pesquisa surgiu de nossa experiência em sala de aula, onde pudemos observar que o que ensinávamos não atendia às necessidades de comunicação de nossos alunos. Essa concepção nos leva à relação entre teoria e prática, de forma que não podemos mais formular teorias sem que os envolvidos nas práticas tenham voz (MOITA LOPES, 2008). Outra questão é a da articulação de diversas áreas e disciplinas, de forma que não conseguiríamos realizar essa pesquisa estando dentro de uma só área, da Linguística, do Ensino de Línguas, da Publicidade ou da Pedagogia. Além disso, acreditamos que os

⁶ Tradução minha. Doravante todas as traduções são feitas pela autora.

⁷ Tarefa, para esses autores, é uma atividade na qual o idioma é usado pelo aluno para um propósito comunicativo, com o objetivo de atingir um resultado; o significado é primordial, em detrimento da forma, e existe uma relação com situações do mundo real (WILLIS E WILLIS, 2007).

⁸ Artigos de revistas como *Specialist*, *Delta*, *Newsletter* e *Intercâmbio* (LAEL – PUC-SP), eventos como InPLA (Intercâmbio de Pesquisas em Linguística Aplicada) e Encontro de Linguística de *Corpus* (vide Berber Sardinha e Almeida, 2008), projetos como o da USP-SP: COMET (*Corpus* Multilíngue para Ensino e Tradução). Disponíveis em: <http://www.pucsp.br/pos/lael/contents.php> e <http://www.fflch.usp.br/dlm/comet/>. Acesso em março de 2010.

⁹ Grupo de Estudos de Linguística de *Corpus*, do LAEL – PUC-sp. Uma das esferas de atuação do GELC/CNPq é a produção de material didático advindo da pesquisa com *corpora*.

conhecimentos que podemos construir servirão para diversas áreas e disciplinas, de forma que eles não sejam estáticos e pouco abrangentes.

Destacamos a seguir as pesquisas do Programa de Estudos Pós-Graduados em Linguística Aplicada e Estudos da Linguagem (LAEL) da PUC-SP no âmbito da LC. Jacobi (2001) analisa padrões na língua espanhola utilizando-os na preparação e aplicação de material didático para o ensino de espanhol a brasileiros. Bértoli-Dutra (2002) usa a LC na produção de atividades didáticas baseadas em letras de música. Ferrari (2004) mostra como elaborar atividades para o ensino de inglês para alunos de Comércio Exterior baseando-se na abordagem sociointeracionista. Barbosa (2004) propõe uma metodologia de elaboração de material didático *online* baseado em *corpus* para o ensino-aprendizagem de inglês instrumental apoiando-se em princípios teóricos da Autonomia na Aprendizagem de Línguas, na Aprendizagem Experiencial e na Aprendizagem Baseada em Tarefas. Condi de Souza (2005) descreve a coleta, a análise e a utilização de *corpora* na elaboração de uma tarefa para o aprendizado de padrões lexicogramaticais em Língua Inglesa. Ferreira (2006) mostra como a LC auxilia no ensino da pronúncia do inglês. Vicentini (2006) sugere a elaboração de atividades para ensino de padrões selecionados de um *corpus* de *scripts* do seriado *Friends* com base na teoria sociointeracionista de Vygotsky. Alonso (2006) objetivou a preparação de atividades para o ensino de falsos cognatos em espanhol com o uso de concordâncias. Moreira Filho (2007) desenvolve um *software* para a preparação semiautomática de aulas de leitura em inglês com *corpora*. Pinto (2008) descreve o uso de padrões em *corpora* de falantes nativos e de alunos de ILE para preparar atividades de ensino baseando-se na abordagem Pós-Método. Bissaco (2010) analisa a interação entre alunos e professor enquanto realizam atividades com concordâncias no ensino de língua espanhola.

Embora haja uma série de publicações, observamos que dentre esses trabalhos somente um foca o ensino de ILE para profissionais de uma área específica, que é o objetivo de nossa pesquisa. Ademais, a maioria deles baseou-se em uma abordagem de ensino específica, ao passo que propomos dois tipos de atividades, sendo uma baseada em uma abordagem e outra aberta, com propriedades que englobam diversas abordagens, ambas voltadas especialmente para o ensino de aulas em empresas.

Buscamos preencher duas lacunas principais na literatura com nosso trabalho: (1) a escassez de pesquisa em desenvolvimento de material de ensino com *corpus* e (2)

a falta de pesquisa em LC na preparação de material de ensino para a área de Publicidade. Os materiais para o ensino de inglês para publicitários são geralmente preparados pelo professor e baseados em textos da *Internet*, de *sites* das áreas que fornecem informações sobre a atuação das agências no mundo publicitário. As atividades preparadas pelo professor, assim como a escolha dos itens a serem explorados nestes textos são feitas sem acesso a recursos da Linguística de *Corpus* e os alunos acabam por estudar somente itens gramaticais retirados dos textos ou então extraem o vocabulário sem analisar o seu uso.

A fim de atingir os objetivos mencionados acima, formulamos as seguintes questões de pesquisa:

- 1) Quais os padrões lexicogramaticais mais característicos do *ad corpus*?
- 2) Como os padrões lexicogramaticais distintivos do *ad corpus* podem ser usados no desenvolvimento de atividades para o ensino de ILE para o público-alvo?

Explicamos brevemente a seguir a organização dessa dissertação. O capítulo 1 expõe e detalha a fundamentação teórica que embasa a pesquisa. A primeira seção corresponde à Linguística de *Corpus* e engloba seu histórico e aplicação, sua definição e exemplos, sua tipologia, sua construção, extensão e representatividade e seus conceitos. A seção 2 volta-se para o Ensino de Língua Estrangeira e aborda os pressupostos das abordagens de ensino de línguas relevantes para a pesquisa. A seção inclui a relação entre a LC e o ensino de línguas, apresentando as propostas para uso de *corpora* no ensino de língua estrangeira e no desenvolvimento de material didático, o uso de concordâncias na sala de aula e, por fim, a nossa proposta para o ensino de ILE com *corpora* em empresas.

O capítulo 2 é dedicado à metodologia de coleta e análise dos dados e nele apresentamos a coleta dos dados, o *corpus* de estudo e o *corpus* de referência, bem como os procedimentos de análise dos dados, especificamente a análise de padrões com a palavra *ad* e seus colocados.

No capítulo 3 partimos para a exposição sobre o desenvolvimento das atividades, explicando como os achados da análise foram nelas usados.

Em seguida, encontram-se as seções que abordam as considerações finais, as referências bibliográficas e os anexos.

Pretendemos, com esta pesquisa, contribuir para que o ensino de ILE possa modernizar-se e que tecnologias sejam usadas para o desenvolvimento de aulas mais significativas e customizadas às necessidades dos alunos em seu meio profissional.

CAPÍTULO 1: FUNDAMENTAÇÃO TEÓRICA

Apresentaremos aqui o arcabouço teórico da pesquisa. O capítulo está dividido em duas seções. A primeira seção aborda a LC, embasamento principal da pesquisa. Nessa seção, explicamos seu histórico e aplicação, sua definição e exemplos, sua tipologia, sua construção, extensão e representatividade e seus conceitos. A segunda seção trata do Ensino de Língua Estrangeira e nela apresentamos as abordagens de ensino de ILE, as propostas para o ensino com *corpora* e com concordâncias e, por fim, como nossa proposta concilia os pressupostos dessas abordagens e propostas.

1.1 Linguística de *Corpus*

Este trabalho tem como fundamentação teórica principal a Linguística de *Corpus* que, segundo Berber Sardinha (2004, p. 3),

ocupa-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. Como tal, dedica-se à exploração da linguagem por meio de evidências empíricas, extraídas por computador.

A escolha por essa área deve-se ao fato de que a coleta de textos é feita de forma científica, sistemática e estruturada, com um objetivo específico, e esses dados podem servir de base empírica para o reconhecimento, a exploração e a análise dos elementos da língua em uso.

1.1.1 Histórico e aplicação

Nos dias de hoje, em que o computador é uma ferramenta indispensável, é difícil imaginar a coleta e análise manual de *corpora*. Considerando *corpus* no seu sentido original, conjunto de textos ou documentos, Berber Sardinha (2004) assinala que na Grécia Antiga já havia o *Corpus* Helenístico, definido por Alexandre, o Grande, em 320 a.C. Evidências de um trabalho efetivo de coleta e análise de palavras e frases em grandes quantidades de textos datam do século XIII, quando estudiosos da Bíblia e suas equipes de seguidores e protegidos indexavam as palavras de citações bíblicas

manualmente e as organizavam alfabeticamente, juntamente com anotações de onde a passagem ocorria no texto (McCARTHY e O'KEEFE, 2010). Curiosamente, já era o início do trabalho com concordâncias¹⁰. O primeiro trabalho com concordâncias da Bíblia é o *Concordantiae Morales*, anônimo, mas associado a Antônio de Pádua, um doutor da Igreja¹¹ que viveu nos séculos XII e XIII. Esse trabalho foi sucessivamente replicado com outras versões da Bíblia¹².

Já no século XVIII, o trabalho com *corpora* e com concordâncias foi apropriado por estudiosos literários como Andrew Becket, que produziu, em 1787, *A Concordance to Shakespeare*¹³, obra na qual “passagens destacadas e paralelas das peças do merecidamente admirado escritor são metodicamente organizadas”¹⁴. O texto constitui-se de palavras organizadas em ordem alfabética seguidas de passagens das peças de Shakespeare nas quais elas ocorrem. Veja a Figura 1 a seguir com um trecho da obra citada.

¹⁰ Atualmente, o termo concordâncias refere-se a “recortes do conteúdo de vários textos” (BERBER SARDINHA, 2009). O conceito e o uso serão explicados na seção 1.2.2.2.

¹¹ Doutores da Igreja (do latim *Doctores Ecclesiae*) são homens e mulheres que enaltecem o Cristianismo da Igreja Católica por meio de pensamentos, pregações ou escritos. Ver: CHAPMAN, J. Doctors of the Church. In *The Catholic Encyclopedia*. New York: Robert Appleton Company, 1909. Disponível em: <http://www.newadvent.org/cathen/05075a.htm>. Acesso em fevereiro de 2012.

¹² Ver <http://www.1911encyclopedia.org/Concordance>. Acesso em agosto de 2011.

¹³ Título completo: *A Concordance to Shakespeare: suited to all the editions, in which the distinguished and parallel passages in the play of that justly admired writer are methodically arranged. To which are added, three hundred notes and illustrations, entirely new.* (BECKET, 1969)

¹⁴ “*distinguished and parallel passages in the play of that justly admired writer are methodically arranged.*” (BECKET, 1969, título)

W E L C O M E.

Sir, you are very welcome to our house :
It must appear in other ways than words,
Therefore I scant this breathing courtesy.
Merchant of Venice, A. 5, S. 1.

A hundred thousand welcomes : I could weep,
And I could laugh; I am light, and heavy. Wel-
come.

A curse begin at very root of's heart,
That is not glad to see thee !
Coriolanus, A. 2, S. 1.

— Therefore as a stranger give it welcome.
There are more things in heaven and earth, Horatio,
Than are dreamt of in your philosophy.
Hamlet, A. 1, S. 5.

————— Hermione,
How thou lov'st us, shew in our brother's welcome ;
Next to thyself, and my young rover, he's
Apparent to my heart. *Winter's Tale, A. 1, S. 2.*

————— Pray you, bid
These unknown friends to us welcome ; for it is
A way to make us better friends, more known.
Winter's Tale, A. 4, S. 3.

Figura 1: trecho de *A Concordance to Shakespeare* com a palavra *welcome*.¹⁵

É interessante observar que o trabalho com concordâncias realizado hoje por linguistas de *corpus* ainda é considerado metódico e exaustivo, embora essas propriedades estejam associadas à coleta e à análise e não ao processamento, que hoje é feito por programas computacionais em segundos. No mesmo século, o exaustivo e longo trabalho de coleta e análise manuais de dados feito pelo lexicógrafo Samuel Johnson, conhecido por Dr. Johnson, deu origem ao *Dictionary of the English Language*, cujo foco era nos significados do uso das palavras e que foi considerado um modelo, sendo o dicionário mais usado na época.

Durante o século XX, vários pesquisadores ainda construíam *corpora* não eletrônicos. Dentre eles, citamos o *Survey of English Usage* (SEU), compilado em 1959 por Randolph Quirk e sua equipe em Londres; esse *corpus*, bem como as categorias gramaticais resultantes de sua análise, serviram de referência para outros *corpora* e forneceram a base para a famosa gramática *A Comprehensive Grammar of the English Language*, de Quirk, Greenbaum, Leech e Svartvik (QUIRK et al., 1985). Foi na década de 50 que linguistas como Firth, Sinclair, Halliday, Harris e Fries intensificaram

¹⁵ Disponível em http://reader.digitale-sammlungen.de/de/fs1/object/display/bsb10749956_00001.html?leftTab=toc. Acesso em fevereiro de 2012.

as pesquisas com coletas de dados linguísticos reais (LEECH, 1992; SINCLAIR, 2004a).

Firth (1957) é considerado um dos estudiosos mais importantes para a fundamentação da Linguística de *Corpus*, especialmente no que diz respeito ao estudo de uma língua em função do seu contexto de uso (McENERY e WILSON, 2001). Os estudos de Firth nos anos 50 contribuíram para o surgimento de uma nova tradição de linguistas, chamados neo-firthianos, como John Sinclair e Michael Halliday. Halliday (1994) desenvolve a Gramática Sistemico-funcional sob uma perspectiva sociosemiótica, partindo do interesse pelos significados da língua em uso e os contextos em que ela é realizada. Sinclair (2004a) propõe questionamentos sobre a centralidade do léxico e sobre padrões colocacionais¹⁶ na linguagem, o que demandava análise de grandes quantidades de textos que só podia ser feita manualmente, pela falta da tecnologia computacional da época.

Entretanto, no final dos anos 50, o empirismo, e conseqüentemente as pesquisas com *corpora*, perderam força com a chegada da teoria racionalista de Chomsky, que rejeitava a pesquisa com dados autênticos e criticava a confiabilidade dos *corpora* coletados manualmente. Enquanto Chomsky considerava a linguagem como um fenômeno cognitivo e que suas estruturas gramaticais – separadamente do significado – deveriam ser estudadas por meio de metodologias que se apoiassem na intuição do pesquisador, Sinclair (2004a, p. 2) argumentava que “a linguagem deveria ser estudada em contextos de uso reais, deveria trabalhar com exemplos abrangentes – quando possível textos completos – e deveria ter a análise de significado como ponto central”¹⁷. A palavra chave de Sinclair era ‘evidência’, o que demonstra uma posição fundamentalmente empírica.

As pesquisas com *corpora* reapareceram e consolidaram-se nos anos 80, com o advento de uma tecnologia que possibilitou a análise de grandes quantidades de dados linguísticos por meio do computador. Segundo McEnery et al. (2006), o termo ‘Linguística de *Corpus*’ só começou a ser usado no início dos anos 80. Foi quando teve início o projeto *COBUILD*, uma parceria entre a Universidade de Birmingham e a editora *Collins*, cujo editor-chefe foi Sinclair; atualmente o projeto é considerado um

¹⁶ A seção 1.1.3 explica esse conceito.

¹⁷ “language should be studied in naturally occurring contexts of use, should work with extended examples – where possible complete texts – and should have at its centre the analysis of meaning” (SINCLAIR, 2004, p. 2)

dos trabalhos mais marcantes para a LC. Juntos publicaram dicionários, gramáticas e livros didáticos baseados na pesquisa com *corpora* para o ensino de inglês, deixando uma contribuição importante para pesquisas subsequentes. Segundo Berber Sardinha (2007, p. 127), Sinclair criou uma escola de LC que talvez seja “a mais praticada e a mais dinâmica do mundo”.

Na década de 90, suítes de *software* como o *Wordsmith Tools*, de Mike Scott, facilitaram ainda mais o trabalho na área e atualmente existem diversos programas e ferramentas à disposição de linguistas de *corpus*, que atuam em universidades em diversos países. Segundo Berber Sardinha (2004), um dos centros de pesquisa mais desenvolvidos é o da Grã-Bretanha, com universidades em cidades como Londres, Birmingham, Lancaster, Liverpool e Brighton. Nos Estados Unidos, a Linguística Gerativo-transformacional é mais dominante, o que pode ser um dos motivos pelos quais a LC tem uma presença menos marcante. No entanto, é lá que atua Douglas Biber, um dos maiores linguistas de *corpus* da atualidade; além disso, o país apresenta um grande desenvolvimento em pesquisas em Processamento de Linguagem Natural (PLN), um campo multidisciplinar que envolve Linguística, Informática e Inteligência Artificial.

Conforme McCarthy e O’Keefe (2010), nunca houve um período tão fértil para a LC como o atual. Segundo os autores, a LC é, para muitos, um fim em si mesma; no entanto, as pesquisas atuais mostram que ela também é um meio para um fim. Como um fim em si mesma, a LC fornece meios para análises empíricas da linguagem, contribuindo para a sua descrição e definição, o que verificamos em dicionários e gramáticas. Como um meio para um fim, a LC vai além dos domínios do léxico e da gramática, aplicando suas técnicas a outras questões de diferentes áreas como Ensino, Análise de Discurso, Pragmática, Linguística Forense, entre outras.

No Brasil, considera-se que as pesquisas em LC estão em fase inicial, sendo que o primeiro livro da área foi publicado em 2004 por Berber Sardinha. Em contrapartida, podemos dizer que a LC já vem ganhando força, haja vista publicações em revistas, jornais, dissertações, e encontros entre pesquisadores da área¹⁸ como o

¹⁸ Artigos de revistas como *Especialist*, *Delta*, *Newsletter* e *Intercâmbio* (LAEL - Programa de Pós-Graduação em Linguística Aplicada ao Ensino de Línguas – PUC-SP), eventos como InPLA (Intercâmbio de Pesquisas em Linguística Aplicada) e Encontro de Linguística de *Corpus* (Ver BERBER SARDINHA e ALMEIDA, 2008), projetos como o da USP-SP, COMET (*Corpus* Multilíngue para Ensino e Tradução). Disponíveis em: <http://www.pucsp.br/pos/lael/contents.php> e <http://www.flch.usp.br/dlm/comet/>. Teses e dissertações em Linguística Aplicada pelo LAEL, PUC-SP, disponíveis em: http://www.pucsp.br/pos/lael/lael-inf/def_teses.html e do

Encontro de Linguística de Corpus, que está em sua décima edição e promove discussões entre pesquisadores de diversos países e universidades brasileiras. Apesar desse crescimento, observamos que ainda é bastante restrito o número de pesquisadores na área. Um levantamento realizado por Berber Sardinha e Almeida (2008) mostra que nas agências de fomento CNPq e CAPES¹⁹, 132 pesquisadores mencionaram a expressão ‘Linguística de *Corpus*’ em suas pesquisas, o que não significa que são pesquisadores diretos do campo. Já a menção à ‘Linguística Aplicada’ traz 880 nomes. Os grupos de pesquisa em LC, nesse mesmo levantamento, totalizam 12 em contraposição a 74 em Linguística Aplicada. Podemos verificar em nossa prática e na de colegas que atuam no ensino superior que nas universidades, especialmente em cursos de Letras, a LC ainda não se consolidou como disciplina, embora diversos pesquisadores na área que atuam como professores procurem levá-la às salas de aula.

A aplicação da LC deu-se inicialmente na Lexicografia, o que evidentemente ainda ocorre. Outras áreas, no entanto, estão cada vez mais utilizando a LC, como Ensino e Aprendizagem de Línguas²⁰, Análise de Discurso, Literatura e Estilística Literária, Tradução, Linguística Forense, Pragmática, Sociolinguística, entre outras.

No âmbito da evolução das tecnologias, Tognini Bonelli e Sinclair (2006, p. 208) resumem os estudos em LC em três estágios. Os primeiros vinte anos, de 1960 a 1980, constituem o primeiro estágio, quando *corpora* de até 1 milhão de palavras tinham que ser transformados em formato digital por meio de digitação. O segundo estágio é dividido em duas décadas: a primeira, de 80, quando foram construídos *corpora* de 20 milhões de palavras, com o auxílio de aparelhos de digitalização de textos (*scanners*) e a segunda, dos anos 90, é considerada a Primeira Serendipidade²¹ (*First Serendipidity*), com uma nova magnitude quanto ao tamanho dos *corpora*, haja vista a disponibilidade maior de textos em formato digital. A Segunda Serendipidade (*Second Serendipidity*) e o novo milênio possuem uma quantidade ilimitada de textos já criados em formato digital, disponíveis na *Internet*.

Departamento de Letras Modernas pela USP-SP, disponíveis em http://www.fflch.usp.br/dlm/ingles/pos/html/tese_traducao.html. Acesso em março de 2010.

¹⁹ O Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) é uma agência do Ministério da Ciência e Tecnologia (MCT) destinada ao fomento da pesquisa científica e tecnológica e à formação de recursos humanos para a pesquisa no país. A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) atua para a institucionalização da pós-graduação e seu reconhecimento público.

²⁰ Essa área será enfatizada na seção 1.2.1.

²¹ Neologismo que se refere a uma descoberta feita ao acaso.

Podemos verificar que a LC trouxe uma nova visão sobre a linguagem e não se limita mais somente às questões iniciais formuladas por Sinclair (1991), apesar de ainda haver muito a discutir sobre elas. As diversas áreas que a utilizam, os diversos *corpora* sendo coletados e as novas tecnologias sendo utilizadas trazem novas e múltiplas questões que fazem da LC uma área do conhecimento e uma disciplina emergente que está conquistando seu espaço.

1.1.2 *Corpora*

Nesta seção abordamos as questões que envolvem *corpora* no que se refere à sua definição, tipologia, construção, extensão e representatividade.

1.1.2.1 Definição e exemplos

Como dissemos, o termo *corpus* antes do século XX referia-se a uma coletânea de textos escritos (McCARTHY e O'KEEFE, 2010). No dicionário *Oxford English Dictionary* (OED) de 1956, a definição do termo aproxima-se da que utilizamos hoje, sendo o corpo de material escrito ou falado no qual uma análise linguística se baseia. Berber Sardinha (2004, p. 18) expõe pontos importantes que envolvem a elaboração de um *corpus*, quais sejam: ter dados autênticos, ser um objeto de estudo linguístico, ser criteriosamente selecionado, ser legível por computador e ser vasto para que seja representativo; assim, o autor coloca a seguinte definição como sendo a mais completa:

Um conjunto de dados linguísticos (pertencentes ao uso oral ou escrito da língua, ou a ambos), sistematizados segundo determinados critérios, 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 a descrição e análise.

Dentre os principais *corpora* eletrônicos de Língua Inglesa atualmente disponíveis para pesquisa, citamos o pioneiro *Brown* (*Brown Corpus*), compilado nos anos 60 e composto por inglês americano escrito, com 1 milhão de palavras. Já o *British National Corpus* (BNC) foi o primeiro *corpus* a conter 100 milhões de palavras, compilado nos anos 90 e composto por inglês britânico escrito e falado. Atualmente, o

maior *corpus* de Língua Inglesa americana disponível para pesquisa *online* é o *Corpus of Contemporary American English* (COCA), composto por mais de 425 milhões de palavras e está em crescimento desde seu início, em 1990. Existem também disponíveis *corpora* eletrônicos de diversas outras línguas, incluindo da língua portuguesa, dos quais citamos o *Corpus Brasileiro*²², composto por português escrito e falado, com 1 bilhão de palavras.

1.1.2.2 Tipologia

Um *corpus* é sempre coletado para servir a um propósito específico e sua tipologia está atrelada a esse propósito. Berber Sardinha (2004) agrupa os principais tipos de *corpora* em categorias, quais sejam: modo, tempo, seleção, conteúdo, autoria, disposição interna e finalidade.

A tipologia mais geral é em relação ao modo e ao tempo. Quanto ao modo, um *corpus* pode ser escrito, composto por textos escritos, ou falado, composto por transcrições de fala. É maior a quantidade de *corpora* escritos pelas dificuldades que a coleta de um *corpus* de língua falada traz como transcrições, por exemplo, que consomem muito esforço e tempo. Uma das mudanças que estão para ocorrer, segundo Reppen (2010) é uma maior disponibilidade de *corpora* falados resultantes do desenvolvimento da tecnologia, especialmente no reconhecimento de fala; esse desenvolvimento também pode contribuir para a construção de *corpora* multimodais, que envolvem linguagem gestual, por exemplo. Dentro do *corpus* falado, Tognini Bonelli (2010) destaca o *corpus* de fala (*speech corpus*), que tem como objetivo o estudo dos sons e das características fonéticas da língua. Com relação ao tempo, um *corpus* sincrônico compreende um período de tempo, enquanto um diacrônico compreende vários, ou seja, são ‘fotos instantâneas em intervalos de tempo’ (*snapshots at intervals of time*), segundo Tognini Bonelli (*idem*). Um *corpus* histórico representa um período de tempo passado e um contemporâneo, por sua vez, representa o período de tempo corrente.

²² O *Corpus Brasileiro* foi desenvolvido e está sediado no Centro de Pesquisas, Recursos e Informação de Linguagem (CEPRIL), Programa de Estudos Pós-Graduados em Linguística Aplicada e Estudos da Linguagem (LAEL) da PUC-SP, com apoio da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP). Disponível em <http://corpusbrasileiro.pucsp.br/cb/Inicial.html>. Acesso em agosto de 2011.

Com relação à seleção, existem os *corpora* de amostragem (*sample corpora*), que são amostras finitas de uma língua em um tempo específico, portanto, estáticos; Tognini Bonelli (2010) refere-se a eles como ‘fotos instantâneas’ (*snapshots in time*) e acrescenta que são coletados com o objetivo de mostrar o estado de uma língua com o tempo paralisado. Em contrapartida, os *corpora* monitores são reciclados e refletem o estado atual da língua, ou seja, são dinâmicos ou orgânicos, pois podem ser modificados. Um *corpus* é equilibrado (*balanced*) quando seus componentes são distribuídos em quantidades semelhantes.

O *corpus* especializado abarca textos de um tipo específico, como um gênero, um registro, um tempo, ou um tópico e são usados para investigar um tipo particular de linguagem. Conforme Tognini Bonelli (2010, p. 22), eles são “escolhidos não por serem ordinários (comuns), mas por serem extraordinários”²³ e são coletados pelo fato de que as escolhas linguísticas neles são únicas (ímpares). Já o *corpus* geral (HUNSTON, 2002) inclui a maior variedade de textos possíveis com o objetivo de serem representativos²⁴ de uma língua.

Ainda quanto ao conteúdo, há os *corpora* dialetais ou regionais, compostos por variedades sociolinguísticas específicas. Há também os *corpora* multilíngues e bilíngues, que incluem idiomas diferentes. Inicialmente, o desenvolvimento de *corpora* bilíngues e multilíngues estava associado às necessidades de tradução mecânica (TOGNINI BONELLI, 2010). Hoje, os trabalhos em Estudos de Tradução se utilizam dos *corpora* paralelos, compostos por textos comparáveis, usados para estudos comparativos envolvendo o texto original e sua tradução, por exemplo, e dos *corpora* alinhados, nos quais as traduções aparecem abaixo de cada linha do original. Ambos são categorizados por Berber Sardinha (2004) relacionados à ‘disposição interna’.

Os textos podem ter como autores falantes nativos da língua e formam, assim, um *corpus* de língua nativa. Os *corpora* podem ser formados por aprendizes de uma língua e são, portanto, considerados *corpora* de aprendiz.

De forma geral, os *corpora* tiveram como objetivo o registro do uso dos falantes de uma comunidade e esses falantes deveriam ser nativos da língua em questão. Atualmente, existe uma questão que envolve quem é o falante ‘ideal’ de uma língua e qual seria seu modelo padrão (*standard*), haja vista sua variedade. Além disso,

²³ “*chosen not for their ordinariness but for their extraordinariness*” (TOGNINI BONELLI, 2010, p. 22)

²⁴ A questão da representatividade será abordada na seção 1.1.2.3.

recentemente tem havido um interesse crescente na linguagem usada por não nativos e uma discussão sobre a importância de considerá-la nos estudos de *corpora*. O Inglês como Língua Franca (*English as a Lingua Franca*), doravante ELF, surgiu para referir-se à comunicação em inglês entre falantes de diferentes línguas. De cada quatro falantes de Língua Inglesa no mundo, somente um é nativo (CRYSTAL, 2003). Dessa forma, a Língua Inglesa está sendo influenciada e modificada tanto por falantes nativos quanto não nativos e, apesar de ainda haver muita controvérsia, existe uma necessidade de estudos em ELF com relação à questão do nativo ainda ser considerado o ‘guardião’ do padrão e do que é correto, bem como suas implicações no ensino da língua (SEIDLHOFER, 2005). Conforme aponta Mauranen (2009), as pesquisas em ELF demoraram em decolar, mas têm crescido rapidamente e é hoje uma área de pesquisa bastante ativa. Mauranen é líder do projeto *ELFA Corpus*²⁵, que compilou um *corpus* de 1 milhão de palavras de linguagem acadêmica falada. Seidlhofer²⁶, outra pesquisadora da área, questiona os estudos com *corpora* que consideram um *corpus* representativo da língua em geral, sugerindo que é necessário explicitar que ele é composto por um ‘tipo de inglês’, e não geral, pois não engloba falantes não nativos. A autora, que coordena um projeto de coleta de um *corpus* de ELF²⁷, ainda aponta para a necessidade de se discutir essas questões na LC, que ainda não incorporou tais conceitos. Cabe destacar, portanto, que surge uma nova autoria de *corpora*, a de ELF e/ou não nativos de uma língua, cujos estudos estão voltados para o uso da linguagem e não para os erros, como ocorre com os *corpora* de aprendiz.

Por fim, os *corpora* podem ter diversas finalidades. Denomina-se *corpus* de estudo o *corpus* a ser descrito; um *corpus* de referência é usado para servir de contraste com o *corpus* de estudo. Há ainda o *corpus* de treinamento ou teste, construído para o desenvolvimento de ferramentas computacionais. Tognini Bonelli (2010) também identifica os *corpora* para comparação. A princípio, qualquer *corpus* pode ser usado para comparação; no entanto, alguns *corpora* são construídos especificamente com essa finalidade e possuem, assim, características diferenciadas. Os *corpora* comparáveis

²⁵ O *ELFA Corpus* foi compilado por pesquisadores das seguintes universidades: *University of Tampere, the University of Helsinki, Tampere University of Technology e Helsinki University of Technology*. Ele é composto por gravações e transcrições de palestras, apresentações, seminários, defesas de teses e discussões em conferências. Ver <http://www.helsinki.fi/englanti/elfa/elfacorporus>. Acesso em janeiro de 2012.

²⁶ Ideias expostas em conferência no IX Congresso de Linguística Aplicada, no Rio de Janeiro em julho de 2011.

²⁷ Barbara Seidlhofer trabalha em um projeto na Universidade de Viena na coleta do VOICE (*Vienna Oxford International Corpus of English*). Ver <http://www.univie.ac.at/voice/page/index.php>. Acesso em julho de 2011.

(*comparable*) podem ser coletados para fins de comparação de diferentes línguas ou variedades linguísticas com relação a diversos fatores como geográficos, históricos, de tópico e contrastivos. O *corpus* contrastivo tem como objetivo o estudo de variedades e é constituído por componentes principais selecionados para facilitar tal estudo. Uma última finalidade que acreditamos importante mencionar é o *corpus* pedagógico, constituído pela linguagem a qual um aprendiz foi exposto (HUNSTON, 2002). Diversos autores de materiais didáticos o compilam para auxiliar na confecção de materiais extras, testes ou até para informar ao aluno sobre o conteúdo ou a linguagem que ele já estudou.

1.1.2.3 Construção, extensão e representatividade

A construção e a extensão de um *corpus* estão diretamente ligadas ao seu propósito e objetivos. Reppen (2010) explica que ter uma pergunta de pesquisa bem elaborada ou um objetivo de estudo claro é um dos passos essenciais para a construção de um *corpus*. Sinclair (2004a) afirma que “para um *corpus* ser confiável, os critérios estruturais devem ser escolhidos com cuidado porque as questões de equilíbrio e representatividade dependem dessas escolhas”²⁸. Os critérios mais comuns apontados por Sinclair são relativos ao modo (se originário de língua falada, escrita ou, atualmente, eletrônica), tipo de texto (se for de língua escrita, por exemplo, livro, revista e carta), domínio (acadêmico ou popular, por exemplo), língua(s) ou variação linguística, localização dos textos (por exemplo, inglês dos Estados Unidos ou Austrália) e data.

A esses critérios, acrescentamos alguns dos pré-requisitos que Berber Sardinha (2004) lista para a formação de um *corpus* computadorizado, quais sejam: (1) deve ser formado por textos autênticos, ou seja, não pode ter sido produzido para servir de objeto de pesquisa; (2) deve ter um conteúdo selecionado criteriosamente, de acordo com as características e objetivos estabelecidos pelo criador e (3) deve ser um conjunto representativo de uma variedade linguística ou de uma língua.

Berber Sardinha (2004, p. 23) explica que “a representatividade está ligada à questão da probabilidade”, ou seja, como alguns traços linguísticos ocorrem com maior

²⁸ “for a corpus to be trusted, the structural criteria must be chosen with care, because the concerns of balance and representativeness depend on these choices.” (Sinclair) Disponível em <http://ota.ahds.ac.uk/documents/creating/dlc/chapter1.htm>. Acesso em julho de 2011.

frequência do que outros, quanto maior um *corpus*, maior a probabilidade de ocorrência dos traços mais raros ou de baixa frequência. Entretanto, muitas vezes é difícil aproximar-se da população da qual o *corpus* é coletado, ou essa população é desconhecida. Assim, é importante verificar a população de onde provém a amostragem para especificar ‘do que’ o *corpus* é representativo, além de adequá-lo ao objetivo de estudo. Sinclair (1990) exemplifica essa relação com o precursor do *Bank of English*²⁹, que continha uma grande quantidade de textos de ficção para que houvesse exemplos de ‘boa escrita’. Entretanto, observou-se que o *corpus* tinha como objetivo servir de suporte para o ensino de Língua Inglesa e algumas características dessa linguagem iam de encontro a esse propósito; houve, portanto, um conflito entre o conteúdo do *corpus* e seu objetivo.

No âmbito da extensão, o *corpus* dever estar de acordo com a representatividade e a praticidade, que se refere a questões de tempo e disponibilidade (Reppen, 2010). A ideia inicial de ‘quanto maior, melhor’ pode ser atribuída a dois fatores, segundo Evison (2010): primeiro, pela possibilidade (e facilidade) de coleta de grandes quantidades de dados por meios eletrônicos; segundo, porque *corpora* extensos eram necessários para o trabalho de lexicógrafos e gramáticos. No entanto, conforme a pesquisa com *corpora* foi atendendo a novos interesses, essa questão de extensão tornou-se mais aberta e atualmente *corpora* pequenos são considerados como de igual importância a *corpora* grandes, especialmente nas áreas de Inglês para Fins Específicos e Inglês para Fins Acadêmicos (FLOWERDEW 2002, 2004; TRIBBLE, 2002). De qualquer forma, mesmo *corpora* grandes são limitados, conforme aponta Leech (2007), que menciona o *Brown Corpus* para retratar a limitação de qualquer *corpus* quanto à população total e quanto à variedade linguística.

Berber Sardinha (2004, p. 26) relaciona a extensão de um *corpus* a três dimensões: número de palavras, número de textos e número de gêneros, registros e tipos textuais. Quanto ao tamanho, o autor sugere a seguinte classificação:

- Menos de 80 mil palavras: pequeno
- De 80 a 250 mil palavras: pequeno-médio
- De 250 mil a 1 milhão de palavras: médio

²⁹O *Bank of English* é um *corpus* monitor, em expansão, composto principalmente por inglês britânico, com mais de 450 milhões de palavras, sendo parte de um projeto conduzido pela equipe do *COBUILD* juntamente com a *HarperCollins Publishers*, na Universidade de Birmingham, Reino Unido.

- De 1 milhão a 10 milhões de palavras: médio-grande
- 10 milhões de palavras ou mais: grande

Observamos na literatura da LC que a maioria dos autores aponta a adequação do *corpus* aos seus objetivos como um dos pontos principais para sua construção, pois isso é o que determina sua extensão e representatividade.

1.1.3 Conceitos na Linguística de *Corpus*

Apresentamos aqui alguns dos conceitos importantes que envolvem a LC pertinentes em nossa pesquisa, como lexicogramática e visão probabilística da língua.

Dentre as observações de Sinclair (1987; 1990) que emergiram dos estudos do projeto *Cobuild* e subsequentes, apontamos as principais, que formaram a base para alguns dos conceitos da LC (SINCLAIR, 2004a, p. 17-19):

- 1) Nem todas as formas de um lema³⁰ compartilham o mesmo significado; portanto, não é possível simplesmente considerar todas as palavras derivadas de um lema como tendo significados similares. A palavra *moving*, por exemplo, tem um significado associado à afeição emocional, o qual não está associado a outros derivados como *moved* ou ao próprio lema *move*.
- 2) Se uma palavra pode ser usada em mais de uma classe lexical, seu significado pode estar associado a ela, como vemos em *combat* que, como substantivo concerne à questão física e como verbo relaciona-se à questão social; já *locked in combat*, apesar de ser um substantivo, está associado ao social. Essa exceção mostra que a relação entre o significado e a classe gramatical é quebrada quando se trata de um termo técnico ou expressão idiomática.
- 3) Palavras e expressões carregam uma ‘aura de significado’ (*aura of meaning*) que só podemos entender observando várias ocorrências.
- 4) Alguns grupos de palavras de mesma classe como substantivos, por exemplo, ocorrem com maior frequência em uma função, como orações preposicionais; uma palavra que pertence a um desses grupos é *lap*.

³⁰ Lema é como uma entrada em um dicionário, ou seja, a forma de uma palavra sem flexão.

- 5) As categorias tradicionais não são tão sólidas como parecem. Em estudos, verificou-se que *of*, por exemplo, não tem ocorrência mais frequente como preposição, como em *remind of*; esse item geralmente possibilita um grupo nominal a estender sua estrutura pré-nuclear ou oferece um segundo núcleo. Os padrões frequentes com *of* são, por exemplo, *a lot of work* e *a glass of milk*.
- 6) As palavras não são escolhidas independentemente, ou seja, “a escolha de uma palavra condiciona a escolha da próxima, e novamente da seguinte”³¹.

A LC não faz a separação rígida entre léxico e sintaxe, ao contrário, aponta a existência de um nível do sistema linguístico conhecido por lexicogramática, que engloba o vocabulário e a gramática (BERBER SARDINHA, 2004). Um dos conceitos que justificam o estudo de padrões lexicogramaticais é o de linguagem enquanto sistema probabilístico. Segundo essa visão, a linguagem é padronizada, isto é, muitos traços linguísticos teoricamente possíveis não ocorrem na língua em uso com a mesma frequência (BERBER SARDINHA, 2004). Halliday (2008) aponta para o fato de que a questão da probabilidade é bem aceita quando se trata do léxico, ou seja, as pessoas têm consciência das palavras que mais usam, portanto assimilam a ideia da probabilidade. Por outro lado, há uma resistência na aceitação da probabilidade de uso da gramática, pois dizer que mais frases afirmativas do que negativas são usadas, por exemplo, é visto como uma ameaça à liberdade, as pessoas dizem ter poder para escolher entre as duas livremente. O autor explica que isso pode ocorrer porque é mais fácil entender o significado de escolhas lexicais do que das gramaticais, que são mais inconscientes.

Os trabalhos de Sinclair (1991) e Halliday (1994), na questão da visão da linguagem enquanto probabilidade, opõem-se à visão racionalista de Chomsky, para o qual a linguagem é um fenômeno cognitivo e deve ser estudada baseada em metodologias que se apoiam na intuição do pesquisador. Enquanto a tradição chomskyana apoia-se em “evidências que advêm do conhecimento sobre a língua de um único linguista geralmente, por sua vez, baseadas em exemplos que ele ou ela inventou para fins de exemplificações de teorias”³² (SINCLAIR, 2004a, p. 4), a LC apresenta uma posição fundamentalmente empírica, ou seja, o pesquisador baseia-se em

³¹ “the choice of one word conditions the choice of the next, and of the next again” (SINCLAIR, 2004a, p. 19)

³² “the evidence provided by a single linguist’s knowledge of the language often in turn based on examples which he or she has invented for the purposes of theoretical exemplification” (SINCLAIR, 2004a, p. 4)

evidências; no caso do linguista, em dados linguísticos, a partir de um *corpus* (BERBER SARDINHA, 2004).

A linguística chomskyana gerativista enfatiza a determinação de quais agrupamentos sintáticos são *possíveis (permissíveis) dado o conhecimento que um falante nativo possui de sua língua*. Já a linguística hallidayana descreve a *probabilidade dos sistemas linguísticos, dados os contextos em que os falantes os empregam*. (BERBER SARDINHA, 2004, p. 30, grifo meu)

Quanto à intuição, Sinclair (1997, p. 32) não descarta totalmente o seu uso, mas sim explica com relação a que ela deve ser associada:

Conheça sua intuição. Sua intuição sobre a linguagem (...) é seu bem mais importante (...). (...) ela fornece a você informações instantâneas, especialmente sobre duas coisas (a) o significado das palavras isoladas e (b) a formação adequada de sentenças isoladas. Ela também fornece uma variedade de informações sobre estilos e registros e sobre palavras, fraseologias e sentenças consideradas isoladamente. Ela não é, entretanto, confiável com relação à forma que as palavras e sentenças são combinadas na comunicação real.³³

Na visão chomskyana, a língua deve ser ensinada por meio de estruturas gramaticais, pois funciona como um esquema de *slot and filler*, ou seja, o aluno aprende as estruturas e as preenche com o léxico, de forma livre e intuitiva. Já as pesquisas em LC, especialmente na aplicação no ensino-aprendizagem de línguas, descrevem a linguagem do ponto de vista lexical, ou seja, a escolha do léxico e da sintaxe é feita conjuntamente (SINCLAIR, 1987).

Segundo Berber Sardinha (2009, p. 12), “um padrão léxico-gramatical é uma sequência recorrente de (pelo menos duas) palavras, dentro de um espaço delimitado (normalmente equivalente a até quatro palavras de distância), que possui um sentido específico”. Em nossa língua nativa, usamos esses padrões de forma automática, ou seja, não selecionamos palavras isoladas para a construção dos discursos e é esse uso de unidades já padronizadas que nos permite usar a língua com fluência (SINCLAIR,

³³“Know your intuition. Your intuition about language (...) is your most important asset. (...) it gives you instant information, mainly about two things (a) the meanings of words in isolation (b) the well-formedness of sentences in isolation. It also gives a variety of information about styles and registers, and of words, phrases and sentences considered in isolation. It is not, however, reliable about the way words and sentences are combined in actual communication.” (SINCLAIR, 1997, p. 32)

1991). No item 2 das observações de Sinclair que mencionamos anteriormente, o autor aponta para esse sentido específico que as expressões idiomáticas assumem.

Halliday (2008) explica que a lexicogramática não é uma nova estratégia de representação; é somente uma elaboração do que já existe, assim como a linguagem, que não criou uma nova maneira de realizar significados, mas encontrou uma variedade de maneiras de combinar o léxico e a gramática. Dessa forma, com a lexicogramática, a linguagem liberta-se da bipolaridade, permitindo que, com um número finito de unidades, possamos realizar um número infinito de significados (EGGINS, 2004).

Segundo Berber Sardinha (2004) os padrões lexicogramaticais podem ser formalizados em três conceitos principais, que são: colocação, coligação e prosódia semântica, que explicaremos na sequência.

Firth é o autor de uma das frases mais famosas na LC: “*you shall know a word by the company it keeps*” [você conhece uma palavra pela sua companhia] (FIRTH, 1957, p. 11). O autor aponta para a importância das relações e dos padrões que envolvem um item lexical, gerando, assim, o conceito de colocação (*collocation*).

A colocação pode ser considerada como “a tendência de duas palavras de co-ocorrerem, ou como a tendência de uma palavra atrair outra”³⁴ (HUNSTON, 2002, p. 68), ou seja, uma associação de itens lexicais. Segundo Sinclair (1991), a colocação ilustra o princípio idiomático (*idiom principle*) em contraposição ao princípio de escolha aberta (*open-choice*), que mencionamos anteriormente como *slot and filler*, pois “palavras não ocorrem de forma aleatória em um texto”³⁵ (SINCLAIR, 1991, p. 110). Halliday (2008) denomina esse tipo de estrutura de *bricks & mortar* (tijolos e argamassa), de forma que as palavras seriam tijolos, unidas entre si por meio da argamassa, que é a gramática. Um exemplo de colocação é *shed tears*; alguns outros colocados de *shed* são: *light, jobs* e *blood* (HUNSTON, 2002).

A coligação é, segundo Firth, “a inter-relação das categorias sintáticas”³⁶ (PALMER, 1968, p. 23), ou seja, está no nível gramatical e concerne à associação entre itens lexicais e gramaticais. O exemplo bastante claro que mencionamos é o selecionado por Berber Sardinha (2004, p. 40): “*start* é mais comum com sintagmas nominais e orações *-ing*, enquanto *begin* é mais usado com um complemento *to*”. Segundo Tognini

³⁴ “the tendency of two words to co-occur, or as the tendency of one word to attract another” (HUNSTON, 2002, p. 68)

³⁵ “words do not occur at random in a text” (SINCLAIR, 1991, p. 110)

³⁶ “the interrelations of the syntactical categories” (PALMER, 1968, p. 23)

Bonelli (2001), o conceito de coligação está sendo estendido para incluir as preferências sintáticas que uma palavra apresenta dentro de um contexto; ‘palavra’, nesse caso, é vista como um item lexical e não como membro de uma classe gramatical. Esse conceito relaciona-se com as ideias dos itens 4 e 5 mencionados anteriormente que apresentamos sobre as ideias de Sinclair, ao considerar que as preferências sintáticas de uma palavra estão relacionadas ao significado que elas possuem no contexto.

A prosódia semântica é a associação entre itens lexicais e conotação (negativa ou positiva) ou instância avaliativa (BERBER SARDINHA, 2004). Um dos exemplos de prosódia semântica mais citados é o dado por Stubbs (1996): a palavra *cause* tem uma prosódia semântica negativa por coocorrer com palavras de conotação negativa como *problem(s)*, *illness*, *disaster*, *damage* e *concern*.

Conforme Sinclair (2004a, p. 34),

a prosódia semântica tem um papel primordial na integração de um item com seu arredor. Ela expressa algo próxima à função do item – ela mostra como o restante do item pode ser interpretado funcionalmente. Sem ela, a cadeia de palavras apenas ‘significa’ – não é colocada em uso em uma comunicação viável.³⁷

Segundo Hunston (2002), a prosódia semântica deve ser analisada com a observação de uma grande quantidade de ocorrências de um item porque está associada à tipicidade de uso deste item. Hunston ainda afirma que esse conceito não é geralmente ensinado porque não pode ser explicado conscientemente pelo falante nativo da língua. À semelhança, Tognini Bonelli (2001) explica que a prosódia semântica é menos reconhecida de maneira imediata do que a colocação, a menos que o texto esteja com etiquetas morfossintáticas³⁸.

As principais áreas de interesse da Linguística de *Corpus* de acordo com Kennedy (1998) são, de forma sucinta, a compilação e a construção de *corpora*, o desenvolvimento de ferramentas para análise de *corpora*, a descrição da linguagem e o ensino-aprendizagem de línguas. O foco neste trabalho é no uso de *corpora* no ensino-aprendizagem de língua estrangeira, o que detalhamos na seção seguinte.

³⁷ “the semantic prosody has a leading role to play in the integration of an item with its surroundings. It expresses something close to the ‘function’ of the item – it shows how the rest of the item is to be interpreted functionally. Without it, the string of words just ‘means’ – it is not put to use in a viable communication.” (SINCLAIR, 2004, p. 34)

³⁸ Etiquetagem morfossintática consiste na inserção no texto de códigos que indicam a classe gramatical e a estrutura sintática de cada palavra e, respectivamente, frase.

1.2 Ensino de Língua Inglesa

Nesta seção, iremos abordar o ensino de Língua Inglesa, doravante LI, em especial algumas das abordagens em voga. A seguir, iremos detalhar o ensino de LI com *corpora* para finalmente indicar como essas abordagens influenciam a nossa proposta para o ensino de ILE nesta pesquisa.

Atualmente é difícil afirmar que exista uma única abordagem ou método de ensino de LI que seja o mais completo, utilizado de maneira exclusiva ou aceito por professores e pesquisadores. Uma das abordagens mais seguidas em escolas de línguas é a Abordagem Comunicativa, que possui diversas interpretações e aplicações.

Considerando o histórico de ensino de LI, observamos que os profissionais da área acabam por mesclar técnicas e procedimentos utilizados nos diferentes métodos que foram desenvolvidos ao longo do tempo. Os anos 40 e 50 foram marcados pelo ensino de estruturas linguísticas por meio do condicionamento, baseado no behaviorismo. Já a década de 60 trouxe a gramática gerativa de Chomsky e a questão girava em torno de como encaixá-la no ensino. Nos anos 70, houve um foco nos aspectos afetivos, emocionais e socioculturais. As décadas de 80 e 90 presenciaram um desenvolvimento de abordagens fundamentadas em propriedades comunicativas e as salas de aulas passaram a ser caracterizadas por tarefas significativas, simulações de mundo real e autenticidade (BROWN, 2007).

Conforme mencionamos, a Abordagem Comunicativa (*Communicative Language Teaching*), doravante CLT, é a mais difundida hoje em escolas de idiomas no Brasil, nas quais o CLT pode não ser usado em sua essência, mas a maior parte das práticas de ensino está baseada em seus pressupostos. Apontamos a seguir as características fundamentais do CLT, conforme Brown (2007):

- 1) Objetivo no todo: foco em todos os componentes da competência comunicativa, sendo eles o gramatical, o discursivo, o funcional, o sociolinguístico e o estratégico.
- 2) Relação entre forma e significado: as técnicas de ensino devem envolver o aluno no uso da língua pragmático, autêntico e funcional para propósitos variados.
- 3) Fluência e precisão: ambos são importantes e focados em diferentes situações do ensino e aprendizagem.

- 4) Foco em contextos do mundo real: os alunos devem ser estimulados a desenvolver estratégias para usar a língua em contextos diversos fora da sala de aula.
- 5) Autonomia e envolvimento estratégico: os alunos desenvolvem a consciência de sua própria aprendizagem, entendendo suas estratégias e estilos para produção e compreensão da língua.
- 6) Papel do professor: facilitador e guia; os alunos constroem significados por meio da interação com outros alunos e com o professor.
- 7) Papel do aluno: participantes ativos em seu processo de aprendizagem, que é focada no aluno.

Apesar de algumas escolas e professores atestarem que utilizam essa abordagem, observamos que muitas aulas e materiais não condizem com algumas de suas características, como por exemplo, a prática da repetição (*drills*), o uso de textos, de diálogos e de frases inventados e o foco na forma e no ensino da gramática descontextualizada. Portanto, baseados em nossa prática, reiteramos que, atualmente, o ensino de ILE em algumas escolas que se dizem comunicativas está pautado em uma mistura de itens emprestados de diversas abordagens.

O *Task-based Teaching*, doravante TBT, proposto por Willis e Willis (2007), é uma abordagem que orienta os alunos a desempenharem tarefas (*tasks*) com foco no significado e não na forma e está, assim, segundo Ellis (2003), no coração do CLT.

O conceito de tarefa já foi amplamente discutido por pesquisadores, professores e profissionais engajados no TBT. Destacamos aqui a definição considerada mais completa por Willis e Willis (2007):

Uma tarefa é uma atividade na qual

1. o significado é primordial;
2. existe alguma questão comunicativa para resolver;
3. existe algum tipo de relação com atividades comparáveis do mundo real;
4. a conclusão da tarefa tem certa prioridade;
5. a avaliação da tarefa ocorre em termos de produção.

Tarefas:

1. não fornecem ao aluno a linguagem de outras pessoas para regurgitar;
2. não estão interessadas em exposição da linguagem;
3. não são orientadas pela conformidade;
4. não são orientadas pela prática;
5. não embutem a linguagem em materiais para que estruturas específicas sejam focadas.³⁹ (SKEHAN, 1998, p. 95)

³⁹A task is an activity in which 1. meaning is primary; 2. there is some communication problem to resolve; 3. there is some sort of relationship to comparable real-world activities; 4. task completion has some priority; 5. the assessment

Vários autores desenvolveram diversos modelos, com nomes distintos, para o uso de tarefas no ensino de línguas estrangeiras: *Task-based Language Learning* (TBLL), *Task-based Learning* (TBL), *Task-based language teaching* (TBLT) e *Task-based instruction* (TBI).

Segundo Willis (2004), o TBT foi introduzido por Prabhu (1987), que desenvolveu e aplicou o projeto *Communicational Teaching Project* em uma escola secundária, na Índia, utilizando tarefas sem o ensino explícito de gramática para que os alunos tivessem experiência do inglês em uso. Prabhu (1987) baseou seu projeto no pressuposto de que seria possível aprender a língua por meio de uma tarefa que não focasse um problema linguístico da mesma forma que se aprende focando questões linguísticas. Assim, o TBT evoluiu como um ramo da Abordagem Comunicativa (*Communication Language Teaching – CLT*), baseado nos resultados de pesquisas em *Second Language Acquisition – SLA* (ELLIS, 2003).

Um exemplo apontado por Ellis (2003) é o da prática conhecida como *Presentation-Practice-Production* (PPP), amplamente usada no ensino de ILE na Abordagem Comunicativa, cujo foco é na forma, ou seja, o aluno é exposto a uma apresentação do professor sobre algum item ou estrutura gramatical, tem oportunidade de praticá-lo e na fase de produção (*production*), tenta usar o item ou estrutura em uma atividade comunicativa. No entanto, foi observado que o aluno tem a capacidade de fazer uso do que aprendeu no momento da prática (*practice*) ou produção (*production*) quando monitorado, mas não há garantia de que ele conseguirá lembrar ou aplicar essa linguagem em uma situação comunicativa cujo foco é o significado, como em uma situação real, por exemplo (WILLIS, 2004).

Em uma abordagem baseada no significado, o objetivo é encorajar os alunos a usar a língua em situações de comunicação real preparadas pelo professor, mesmo que ela não esteja correta. Assim, a forma está subordinada ao significado, ou seja, deverá vir depois dele. Para Willis e Willis (2007), essa ordem é justificada por duas razões. A primeira é que se o aluno estiver preocupado em usar as estruturas gramaticais a que foi exposto, os benefícios do foco no significado no momento da atividade comunicativa serão perdidos, pois o esforço do aluno será somente em reproduzir as formas. Assim, a

of the task is in terms of outcome. Tasks: 1. do not give learners other people's language to regurgitate; 2. are not concerned with language display; 3. are not conformity orientated; 4. are not practice orientated; 5. do not embed language into materials so that specific structures can be focused upon. (SKEHAN, 1998, p. 95)

probabilidade de o aluno melhorar sua fluência e confiança será menor e ele estará pouco preparado para usar a língua fora da sala de aula. A segunda razão é a extrema dificuldade do aluno para conseguir incorporar a estrutura a que foi exposto em sua produção espontânea dentro de um espaço de tempo limitado.

Visto que o significado deve preceder à forma, uma aula baseada em TBT, segundo Willis e Willis (2007), deve conter uma sequência de tarefas. O planejamento de uma sequência de tarefas inicia-se com as tarefas alvo (*target tasks*), seguidas pelas formuladas pelo professor e que irão preceder as tarefas principais, chamadas de facilitadoras (*facilitating tasks*). Na primeira parte do ciclo, chamada de *task*, os alunos desenvolvem uma tarefa cujo foco é o significado, que pode ser, por exemplo, uma discussão sobre algum tópico. Nessa fase, ao introduzir o tópico, o professor pode inserir algum vocabulário relevante, em questões ou em frases para a discussão, com o objetivo de preparar (*prime*) o aluno para a tarefa. No entanto, o professor não deve corrigir os alunos ou fazer qualquer tipo de apresentação com relação à forma. Na segunda parte (*planning*), os alunos irão preparar-se para reportar suas ideias ou o que foi discutido em grupos, por exemplo, para a classe. Nesse momento, os alunos, individualmente, podem organizar suas ideias e planejar como irão expressá-las aos seus respectivos grupos, bem como planejar com seus grupos como irão reportar o que foi discutido para os demais. Os alunos podem usar livros de referência e o professor pode ajudá-los, corrigindo ou reformulando sua linguagem. Essa tarefa pode também ser suplementada por uma redação, a ser feita em sala de aula ou em casa. Assim, há um foco na linguagem, ou seja, “temos alunos explorando a linguagem em resposta a uma necessidade de expressar significados necessários”⁴⁰ (WILLIS e WILLIS, 2007, p. 24).

Na fase seguinte (*report*), os alunos apresentam suas descobertas e opiniões individuais e/ou do grupo. O foco é tanto no significado, espera-se que o aluno comunique suas ideias, como na obtenção de um nível de precisão, que a própria circunstância de comunicação requer (WILLIS, 1990).

Ao final da sequência de tarefas, há, então, o foco na forma. Nessa fase, o professor especifica e isola itens para estudo e trabalha com eles fora do contexto da atividade comunicativa, por meio de algum exercício. A linguagem selecionada pelo professor pode estar ligada às próprias tarefas (itens selecionados de um texto ou

⁴⁰ “we have learners exploring the language in response to a need to express required meanings” (WILLIS e WILLIS, 2007, p. 24)

diálogo), bem como à linguagem usada pelos alunos ao realizarem as tarefas. Por outro lado, os próprios alunos podem identificar e selecionar itens a serem estudados.

O foco na forma deve vir ao final da tarefa por três motivos principais, conforme Willis e Willis (2007). O primeiro é que os alunos poderão explorar a linguagem, baseados em sua própria experiência nas tarefas: ouvindo o professor, preparando o relato e apresentando-o (*report*). Dessa forma, “como esse foco na forma vem depois de os alunos terem experienciado a língua em uso, eles têm um contexto que irá ajudá-los a trazer sentido para a nova linguagem”⁴¹ (WILLIS e WILLIS, 2007, p. 25). O segundo motivo é que o foco na forma fará com que os alunos possam perceber a linguagem usada nas tarefas em futuras exposições à língua. O terceiro é o da motivação, pois os alunos querem saber o que aprenderam. É improvável que os alunos consigam usar a nova linguagem com consistência em uma única aula; entretanto, é importante mostrar as oportunidades de aprendizado que eles têm nas tarefas. Na fase do foco na forma, os alunos podem estar mais receptivos ao estudo da linguagem que pode ser usada para expressar os significados, a qual eles esforçaram-se por encontrar.

Segundo Willis e Willis (2007), uma das vantagens do TBT em aula é a do contato com uma comunicação real, envolvendo os alunos em diferentes tipos de discurso e promovendo uma interação espontânea. É evidente que o contexto de sala de aula ainda é restrito; fora dela, o aluno tem contato com uma infinidade de situações sociais para as quais precisa adaptar sua linguagem. No entanto, ele terá muito mais dificuldade ao enfrentar situações do mundo exterior se o contexto de sala de aula não refletir tais situações, mesmo que artificialmente, já que “tarefas artificiais podem não refletir precisamente o mundo real, mas elas com certeza obrigam os alunos a engajarem em atos de significados do mundo real e de discursos do mundo real”.⁴² (WILLIS e WILLIS, 2007, p. 142).

Nas tarefas, a linguagem usada pelos alunos para resolver um problema ou uma questão reflete a linguagem que os alunos iriam precisar em tal situação no mundo real, já que eles usam essa linguagem para trocar informações. Isso pode ser mais bem observado em aulas de Inglês para Fins Específicos (*English for Specific Purposes* –

⁴¹ “since this focus on form comes after learners have experienced the language in use, they have a context which will help them to make sense of the new language” (WILLIS e WILLIS, 2007, p. 25)

⁴² “artificial tasks may not offer a precise reflection of the real world, but they do oblige learners to engage in real-world meanings and real-world discourse acts” (WILLIS e WILLIS, 2007, p. 142)

ESP) (HUTCHINSON e WATERS, 1987), nas quais os alunos têm oportunidade de usar a linguagem que irão precisar em situações profissionais. Ainda segundo Willis e Willis (2007, p. 136),

é impossível reproduzir todos estes relacionamentos sociais e profissionais na sala de aula de língua, mas isso não nos absolve da responsabilidade de preparar os alunos para lidar com a grande variedade de encontros linguísticos que eles podem vir a experienciar no mundo real⁴³

Em sala de aula, para que as tarefas sejam desenvolvidas, o papel do professor não pode ser aquele tradicional de transmissor do conhecimento, mas em poucas palavras, o professor terá o papel de conduzir e organizar discussões e supervisionar o trabalho dos alunos nas tarefas. Citamos abaixo os papéis do professor no TBT, segundo Willis e Willis (2007):

- 1) Condutor e organizador de discussões, iniciando-as ou organizando os alunos em grupos ou pares.
- 2) Administrador dos grupos e/ou pares, monitorando-os e reagrupando os alunos quando necessário ou para que tenham oportunidade de reformular suas ideias.
- 3) Facilitador, promovendo tarefas motivadoras e certificando-se de que os alunos tenham capacidade de executá-las.
- 4) Motivador, sendo positivo ao dar retorno (*feedback*) e destacando resultados positivos.
- 5) ‘Conhecedor’ da língua e conselheiro, participando das discussões como alguém que tem maior conhecimento e experiência com a língua do que os outros e auxiliando, assim, os alunos com significados, não com a forma. O professor pode destacar as contribuições dos alunos, repetindo-as ou reformulando-as, tendo o cuidado de não fazer correções.
- 6) Professor da língua, em alguns estágios, explicando, demonstrando e levantando formas da língua.

Apresentamos aqui duas propostas mais usadas atualmente no ensino de LI e, em seguida, iremos focar o ensino de LI com *corpora*.

⁴³“it is impossible to reproduce all these social and working relationships in the language classroom, but this does not absolve us from the responsibility of preparing learners to deal with the wide range of language encounters they may experience in the real world” (WILLIS e WILLIS, 2007, p. 136)

1.2.1 Linguística de *Corpus* e Ensino de Língua Inglesa

Nesta seção abordamos uma das áreas de interesse da LC na qual esta pesquisa está inserida, o ensino-aprendizagem de Língua Inglesa.

O interesse de profissionais da área de ensino e aprendizagem de línguas no trabalho com *corpora* advém do reconhecimento da necessidade de priorizar o ensino de padrões lexicais e de que esses padrões não poderiam ser selecionados apenas pela introspecção (SINCLAIR, 2004b). Esse interesse e o uso de *corpora* por esses profissionais ocorrem há mais de 20 anos, como podemos ver pelas pesquisas de Sinclair (1987, 1991, 2004a), e tem sido crescente, especialmente pela disponibilidade de programas e ferramentas de fácil utilização e da maior facilidade de acesso a *corpora*, que estão disponíveis *online* para serem baixados e/ou comprados (RÖMER, 2006, 2009; BERNARDINI, 2004).

Berber Sardinha (2004) destaca os marcos importantes relativos à influência da LC no ensino de línguas, dos quais mencionamos: (a) a obra de Tim Johns (1991), trabalho pioneiro que defende o uso de computadores no ensino, com o objetivo de tornar o aluno um pesquisador; (b) a gramática baseada em *corpus* do projeto *COBUILD*, liderado por Sinclair (1990); (c) o *ELR Journal*⁴⁴, dedicado à concordância em sala de aula; (d) o início da conferência *TALC (Teaching and Language Corpora)*, em 1994, dedicada a aplicações de LC no ensino de línguas; a nona edição foi realizada na República Tcheca, em 2010; (e) a chegada da aplicação de corpus no ensino ao professor de línguas estrangeiras no *TESOL*⁴⁵, em 1996; e (f) a coletânea editada por Granger (1998) sobre *corpora* de aprendiz, *Learner English on Computer*.

Segundo Berber Sardinha (2004), existem quatro principais áreas de concentração do acesso e exploração de *corpora* no ensino. A primeira é a da descrição da linguagem nativa, escrita e falada, em pesquisas conduzidas no meio acadêmico. A introdução da LC no ambiente pedagógico ocorre por meio dos materiais didáticos que incorporam os resultados dessas pesquisas.

A segunda é a descrição da linguagem usada por estudantes de língua estrangeira, compilada em *corpora* de aprendiz (*learner corpora*). Essa é uma atividade

⁴⁴ Disponível em <http://ejournals.org.uk/ELR>. Acesso em março de 2010. Inicialmente lançado como *English Language Research*, e relançado como *Empirical Language Research*.

⁴⁵ O *TESOL - Teachers of English to Speakers of Other Languages, Inc.* é uma associação global para profissionais de ensino de língua inglesa e está sediada na cidade de Alexandria, em Virginia, nos Estados Unidos.

que tem tomado impulso recentemente e que e traz, conforme Granger (2003), perspectivas pedagógicas bastante interessantes para o desenvolvimento de materiais, de currículos, de avaliações e de metodologias.

A terceira área é a da transposição de metodologias de pesquisa acadêmica para a sala de aula, por meio, geralmente, de concordâncias e listas de palavras. Segundo o autor, a concordância é o instrumento mais empregado e serve de base para materiais de ensino com diversos propósitos, como para a solução de dúvidas sobre o emprego de palavras específicas ou até para fins de conscientização. Com relação a essa área, Berber Sardinha (2004) explica que a inserção da LC em sala de aula é raramente feita pelo professor em virtude de haver um distanciamento entre o profissional de ensino e a academia.

A quarta área é a do desenvolvimento de metodologias e abordagens de ensino baseado em pesquisa com *corpora* ou em conceitos da Linguística de *Corpus*, que abordaremos na seção 1.2.1.1.

No que diz respeito ao material didático, mais especificamente às gramáticas, destacamos as obras de Biber et al. (1999), Conrad e Biber (2009) e a de Carter e McCarthy (2006). A *Longman Grammar of Spoken and Written English (LGSWE)*, de Biber et al. (1999), revela dados estatísticos sobre o uso da Língua Inglesa escrita e falada, em suas estruturas, tempos verbais e léxico em diferentes registros⁴⁶. Dez anos mais tarde, Conrad e Biber (2009) publicam a *Real Grammar*, que foca nos padrões gramaticais mais comuns da Língua Inglesa e suas diferenças na escrita e na fala dentro de contextos discursivos. Trata-se de um material inovador, no qual os autores apresentam conteúdos que os alunos aprendem nos livros de gramática e o que o *corpus* informa sobre eles, seguido de exercícios contextualizados que englobam análise de discurso e prática de escrita, de forma simples e que promovem a reflexão sobre o uso da língua. Carter e McCarthy (2006) oferecem explicações claras sobre o uso do inglês escrito e falado baseadas no uso real da língua na *Cambridge Grammar of English*.

Com relação a materiais didáticos para cursos, podemos destacar a série *Touchstone* (McCarthy et al., 2005), um material ‘informado pelo *corpus*’ (*corpus-informed*⁴⁷), no qual os alunos encontram as palavras, expressões e gramática mais

⁴⁶ Registro é o termo usado por Biber para referir-se às variedades linguísticas associadas a diferentes situações e propósitos.

⁴⁷ Termo usado por McCarthy et. al. (2005) e Carter e McCarthy (2006) para descrever a pesquisa baseada em *corpus* (*corpus-based*).

usadas, visto que é baseado em pesquisas no *corpus* de inglês americano do *Cambridge International Corpus*⁴⁸. O material traz informações ao aluno sobre a frequência de uso de palavras e expressões como, por exemplo, ‘*Really?*’, uma das 50 palavras mais frequentes no *corpus*, para expressar interesse ou surpresa e ‘*yeah*’ ser mais frequente do que ‘*yes*’. Outro exemplo é a série *Face2Face* (REDSTON e CUNNINGHAM, 2009), baseada no *Cambridge English Corpus*⁴⁹ (antigo *Cambridge International Corpus*), que traz exercícios para a aprendizagem e prática de colocações.

Com exceção desses materiais, nos quais a aplicação da LC ocorre de forma indireta, observamos que, apesar dos vinte anos de pesquisa e uso de *corpora* no ensino de línguas estrangeiras, esse tipo de material ainda não está sendo produzido em larga escala. Além disso, Cheng (2010) sugere que ainda há necessidade de mais materiais como o *Reading Concordances*, de Sinclair (2003), visto que ele mostra como examinar e interrogar um *corpus* de maneira efetiva e eficaz por meio de sete passos para a análise de linhas de concordância. Foi o trabalho pioneiro de Tribble e Jones (1990), com o *Concordances in the Classroom*, que o uso de concordâncias no ensino tomou rumo. Os autores descrevem maneiras práticas de transpor os resultados da análise de *corpora* em exercícios ou tarefas para o ensino de línguas.

Leech (1997) aponta três formas para o uso direto de *corpora* no ensino. A primeira é ensinar sobre *corpora*. O estudo da LC como disciplina deve cobrir as principais áreas da Linguística, quais sejam: histórico, objeto de estudo, métodos de investigação e modelos ou teorias empregados. A segunda forma é a de ensinar os alunos a explorar a linguagem por meio de *corpora*; nesse caso, o professor atua como um facilitador, auxiliando o aluno a desenvolver habilidades de pesquisa e de análise da língua. A terceira forma é a da exploração para o ensino, ou seja, o uso seletivo de *corpora* no ensino quando não há o uso do computador em sala de aula.

Existem algumas divergências com relação ao uso de *corpora* no ensino, especialmente no que diz respeito a dois aspectos. O primeiro refere-se ao lugar do

⁴⁸ O *Cambridge International Corpus* é um *corpus* de “conversas e textos do dia a dia que mostram como as pessoas realmente usam o inglês” (“*a large database of everyday conversations and texts that show how people actually use English*”) (McCARTHY et al., 2005 p. iv).

⁴⁹ O *Cambridge English Corpus* (CEC) tem mais de 1 bilhão de palavras e foi construído pela *Cambridge University Press* para auxiliar no desenvolvimento de materiais para aprendizes de ILE. É composto por textos de jornais, livros de ficção e de diversos assuntos, sites, revistas, correspondências comerciais (*junk mail*), programas de rádio e TV, gravações de conversas do dia a dia, entre outros. Mais informações em http://www.cambridge.org/br/elt/catalogue/subject/custom/item3637700/Cambridge-English-Corpus-Cambridge-English-Corpus/?site_locale=pt_BR. Acesso em janeiro de 2012.

corpus e dos resultados da descrição do mesmo. Por um lado, Sinclair (2004a) e os pesquisadores do projeto *COBUILD* defendem fervorosamente o uso direto de *corpora* no ensino, ou seja, a descrição feita pelos linguistas de *corpus* deve ser incorporada e os professores devem usar somente linguagem autêntica, nunca inventada. Por outro lado, Widdowson (1991) defende a separação entre descrição da língua e ensino, afirmando que as análises feitas por linguistas fornecem fatos, mas não garantem relevância pedagógica.

O segundo aspecto de divergência é com relação à mediação entre o *corpus* e o aluno. Para Sinclair (2004a) e Tim Johns (1991), em sua proposta de Aprendizado Movido por Dados (*Data Driven Learning*)⁵⁰, doravante DDL, o aluno deve ter acesso direto a um *corpus* e explorar a linguagem respondendo às suas próprias questões e hipóteses, desenvolvendo a habilidade de um pesquisador. Já Aston (1997) afirma que o professor pode simplificar os resultados das buscas no *corpus*, selecionando e filtrando os textos de acordo com o nível de compreensão do aluno, especialmente no caso de iniciantes. Segundo o autor, “talvez o uso mais óbvio de *corpora* seja tratá-los como fontes de materiais que o professor pode selecionar e adaptar de acordo com o que for necessário”.⁵¹ (ASTON, 1997, p. 52)

O professor pode analisar a linguagem e focar o que achar mais relevante, não se baseando somente no que já foi descrito pelos pesquisadores. Concordamos com Berber Sardinha (2004) ao assinalar que a descrição feita pelo professor não necessariamente conflita com a descrição feita pelos pesquisadores; ela tanto pode acrescentar, no caso de algum elemento que ainda não tenha sido analisado, como pode ser comparada à descrição já existente, levando-se em consideração que são descrições feitas em contextos diferentes. Assim, é possível haver um meio termo entre as várias posições em relação ao uso de *corpora* no ensino.

Dentro da LC, a área que pode ter maior influência para o ensino é a da descrição da linguagem nativa (BERBER SARDINHA, 2004). De acordo com Sinclair (1988), existe uma mitologia acerca do inglês que professores aceitam como verdade, mas que as evidências do *corpus* desafiam. Essa mitologia é proveniente do fato de que “as metodologias de ensino distanciam-se dos desenvolvimentos em Linguística, e isso

⁵⁰ Aprendizado movido a dados (*Data Driven Learning* – DDL) é uma abordagem cuja ênfase é desenvolver no estudante da língua a habilidade de pesquisa e de descoberta, colocando à sua disposição dados de um *corpus*.

⁵¹ “perhaps the most obvious pedagogic use of corpora is to treat them as sources of classroom materials which the teacher can select from and adapt according to requirements.” (ASTON, 1997, p. 52)

é o que ocorre atualmente, e tem ocorrido há alguns anos”⁵² (SINCLAIR, 1997, p. 30), considerando que não há mais nada a ser descrito sobre ela. Berber Sardinha (2004) aponta quatro crenças que envolvem essa mitologia. A primeira é a da separação entre léxico e sintaxe (ou vocabulário e gramática). A segunda é de que o vocabulário serve para preencher as lacunas sintáticas, ou seja, deve-se aprender primeiro a gramática e depois o léxico. Em terceiro lugar, não é possível analisar a fluência nativa, pois se trata de algo subjetivo que está na mente dos falantes nativos. Por fim, o mais importante na linguagem é seu caráter criativo; portanto, não é necessário aprender os modos típicos de expressão da linguagem, os padrões e traços linguísticos.

Embora já existam materiais baseados na descrição da linguagem, é necessário que o professor mude sua visão com relação a esses mitos para que essa descrição possa realmente ser útil no ensino da língua. A principal crença que precisa ser mudada é a do ensino da sintaxe (estruturas gramaticais) em primeiro lugar e do vocabulário em segundo, este como um ‘material’ que irá ser encaixado nas estruturas. A LC trabalha com a lexicogramática, como já mencionamos ao abordar os seus conceitos; assim, são apresentados ao aluno as estruturas e seu léxico correspondente, ao invés de o aluno ter que juntar os dois de acordo com sua criatividade e intuição, o que pode trazer resultados artificiais, diferentes da linguagem usada por um nativo.

Com relação ao uso de *corpora* em sala de aula de língua estrangeira, Aijmer (2009) afirma que ainda há certa resistência ou falta de habilidade por parte de professores, mas prevê um futuro promissor. Römer (2009) sustenta que os linguistas de *corpus* têm um trabalho árduo para enfrentar o desafio de professores e alunos que estão acostumados a métodos mais tradicionais e sugere que uma tarefa importante é divulgar a LC e informar os professores sobre o que existe disponível para ajudá-los em sua prática pedagógica, como *corpora online*, materiais baseados em *corpora*, materiais de referência sobre *corpora*, entre outros. A autora lembra que é igualmente importante conhecer as necessidades dos professores que possuem uma rotina difícil e estressante. A esse respeito, O’Keefe et al. (2007) apontam que, talvez por causa dessa rotina, professores se assustam com a palavra *corpus*, por a relacionarem a uma tecnologia complexa; entretanto, quando eles conhecem materiais baseados em *corpora*, tornam-se

⁵² “the methodologies of language teaching distances itself from developments in linguistics, and this is the case at present, and has been for some years” (SINCLAIR, 1997, p. 30)

interessados e adeptos à LC. Os autores apontam para uma questão importante, a da necessidade de professores informarem linguistas de *corpus*.

Professores sabem o que os alunos precisarão aprender sete dias da semana e eles sabem isso pela prática, não pela teoria. Seu conhecimento tácito precisa ser valorizado mais explicitamente em relação às descobertas sobre *corpus* e suas aplicações práticas. Professores de línguas devem continuamente reafirmar seu papel de mediadores entre descobertas sobre *corpus* e prática. Suas questões de pesquisa sobre a língua que surgem da prática precisam ser consideradas e incorporadas nas pesquisas. Isso seguramente acontecerá conforme mais professores souberem como construir e usar um *corpus* e interpretar criticamente seus achados para o ensino de línguas. (O'KEEFE et al., 2007, p. 246⁵³)

Existem várias críticas com relação ao uso da descrição da linguagem nativa no ensino. Berber Sardinha (2004, p. 264) resume a posição dos críticos em duas instâncias:

Posição moderada: modelagem (*modelling*) do ensino a partir da descrição do *corpus*, cujo expoente seria Ronald Carter. Posição extremista: espelhamento ou atrelamento do ensino na Linguística de *Corpus*, sustentada por John Sinclair, Dave Willis e Michael Lewis.

Cook (1998), um dos maiores críticos da posição extremista, faz considerações sobre o atrelamento dos materiais didáticos ao *corpus* e a limitação do mesmo. De acordo com suas observações, os materiais de ensino podem tornar-se escravos da descrição do *corpus*. Owen (1993, p. 184) é outro crítico que afirma que na gramática pedagógica do *COBUILD* houve “a proibição da intuição e uma dependência insensata na análise computacional do *corpus*”⁵⁴.

Quanto à limitação, existem várias características que tornam um *corpus* limitado para o uso no ensino (COOK, 1998). A primeira é a de que *corpora* são unilaterais, ou seja, leva-se em consideração somente a produção e não a recepção, haja vista serem constituídos por textos de alta circulação e não englobarem textos pouco lidos, mas muito produzidos como, por exemplo, bilhetes. Os *corpora* também são incompletos, ou seja, um adolescente, por exemplo, tem contato com um montante de

⁵³ “Teachers know that learners will need to learn all seven days of the week, and they know this from practice, not from theory. Their tacit knowledge needs to be brought to bear more explicitly in relation to corpus findings and their practical applications. Language teachers must continually assert their role as mediators between corpus findings and practice. Their research questions about language that arise out of practice need to be pursued and incorporated into the research agenda. This will surely be realised as more language teachers are made aware of how to build and use corpora and critically interpret corpus findings for language teaching.” (O'KEEFE et al., 2007, p. 246)

⁵⁴ “prohibition on intuition and a simple-minded reliance on computation analysis” (OWEN, 1993, p. 184)

linguagem (300 milhões de palavras) maior do que o de um *corpus*. Outra característica é a de serem parciais, isto é, o aluno tem mais contato com um tipo de linguagem diferente da linguagem do nativo. O último ponto é o da descontextualização dos *corpora*; visto que no ensino, são apresentados padrões desprovidos do contexto em que ocorreram ou apenas em um pequeno contexto.

Segundo Berber Sardinha (2004), as questões sobre a limitação do *corpus* podem ser explicadas ao aluno, que poderá analisar os padrões selecionados, bem como ter contato direto com os textos que formam o *corpus*, entendendo sua limitação e, ainda assim, podendo explorá-los.

1.2.1.1 Propostas para uso de corpora no ensino de LI e no desenvolvimento de material didático

Nesta seção, apresentamos algumas das propostas para utilização de *corpora* no ensino de LI que constam na literatura no que concerne a preparação de aulas e material didático.

Berber Sardinha (2004) relaciona três propostas: o Currículo Lexical (*Lexical Syllabus*), desenvolvido por Dave Willis (1990); a Abordagem Lexical (*Lexical Approach*), de Michael Lewis (1993); e o Aprendizado Movido a Dados (*Data Driven Learning – DDL*), formulado por Tim Johns (1991), que abordaremos a seguir.

O Currículo Lexical (*Lexical Syllabus*), proposto por Sinclair e Renouf (1988) e desenvolvido por Willis (1990), aplica diretamente os conceitos da LC no ensino de língua estrangeira. Nesta proposta, Willis (1990) utilizou a descrição de um *corpus* da Língua Inglesa, a princípio coletado para a produção do dicionário *COBUILD*, no desenvolvimento de uma série de livros didáticos.

O autor argumenta que a abordagem comunicativa e os métodos que nela se baseiam separam a linguagem em blocos e os apresentam aos alunos, que devem então assimilá-los. O problema desse tipo de abordagem, evidenciado pela sua prática de ensino e por pesquisas (ELLIS, 1984) é que “*‘input’ does not equal ‘intake’*”, ou seja, nem tudo o que é ensinado é efetivamente aprendido; assim, o autor conclui que o professor tem pouco controle sobre o que está sendo aprendido pelo aluno. A constante repetição do professor ao ensinar um determinado item mostra que o aluno leva um tempo para assimilar e usar o que aprende em uma língua. Por isso, Willis questiona se,

ao invés da apresentação controlada pelo professor, não seria a constante exposição à língua o que realmente proporciona a aprendizagem. Assim, dever-se-ia utilizar metodologias que proporcionem a exposição significativa do aluno não só à língua, mas também ao seu uso.

Para o desenvolvimento de tal abordagem, é necessário levar em consideração a tendência natural do aluno de usar sua lógica para entender a língua. Dessa forma, Willis (1990) apresenta duas ações para que essa abordagem tenha sucesso. A primeira é que a metodologia deve ser feita de forma a desenvolver a habilidade do aluno de aprender a partir da observação e da decodificação da língua a fim de que ele possa usá-la em sua comunicação. A segunda é que é necessário verificar a que tipo de linguagem os alunos estão sendo expostos. O autor sugere a organização dessa exposição em três formas:

- 1) A linguagem a que o aluno estará exposto deve ser estruturada de forma que ele não enfrente dificuldades e desmotive-se ao entrar em contato com uma atividade muito complexa para seu entendimento.
- 2) Deve haver uma seleção da linguagem a ser exposta ao aluno de forma a mostrar os padrões e significados mais comuns da língua.
- 3) É importante organizar o currículo de forma que os alunos sejam expostos à linguagem e que sejam destacadas as questões levantadas pela sua experiência com a língua.

As questões mais críticas com relação a esses aspectos foram sobre quais seriam esses padrões e significados da linguagem mais comuns e de como organizar um currículo que os incluíssem. Willis então partiu para o desenvolvimento do *Collins COBUILD English Course* utilizando os materiais usados no projeto *COBUILD* proposto e desenvolvido por Sinclair et al. (1987) na compilação e análise de um *corpus* da Língua Inglesa para a criação do dicionário *COBUILD*. O material didático foi criado com base na descrição da linguagem feita pelos pesquisadores do projeto a partir de um *corpus* composto por 7.3 milhões de palavras. Foram selecionados os pontos principais da descrição e os textos que os exemplificassem foram reunidos em outro *corpus*, intitulado *learner's corpus*⁵⁵. Assim, a linguagem dos materiais é autêntica e centrada no léxico.

⁵⁵ O *learner's corpus* é diferente do *learner corpus* (*corpus* de aprendiz), o qual é um *corpus* composto por textos de aprendizes, ou seja, falantes não nativos.

Segundo Willis (1990, p. vii),

o currículo lexical, portanto, proporciona ao aluno uma oportunidade de aprendizado coerente. Ele não dita o que vai ser aprendido e em qual ordem. Ele oferece ao aluno a experiência com um pequeno, mas balanceado, *corpus* de linguagem natural do qual é possível fazer generalizações sobre a língua como um todo.⁵⁶

O Currículo Lexical, como o próprio nome já evidencia, é baseado em uma visão lexical da língua, ou seja, a linguagem não é apresentada e analisada por meio de estruturas gramaticais. Um exemplo dado por Willis (1990) é o do uso do *would*. A maioria dos materiais didáticos apresenta as três estruturas condicionais isoladas: (1) *If it rains we'll get wet.*, (2) *If it rained we would get wet.* e (3) *If it had rained we would have got wet.*, sendo que existe um grande número de possibilidades de uso das formas condicionais, como vemos no exemplo: *You can always explore the neighbourhood if you have half an hour to spare. Even if I had the time I feel too tired. If it got out it might kill someone. If it's all right by you we could start now.* Segundo o autor, os alunos geralmente aprendem que *might* e *could* são usados para expressar possibilidade; assim, os alunos poderiam criar frases como os dois últimos exemplos sozinhos, sem a necessidade de aprender a regra ou a estrutura. A ideia, portanto, é de que se o *would* for ensinado lexicalmente com a ideia de hipótese, os alunos podem produzir frases como o segundo exemplo (*If it rained we would get wet.*) sem a necessidade de saber regras gramaticais. Com esse exemplo, vemos que “a estratégia de destacar significados de palavras é muito mais produtiva do que a estratégia de ensinar padrões estruturais”.⁵⁷ (WILLIS, 1990, p. 19).

Para Willis (1990), uma metodologia que proporcione uma comunicação real deve conter atividades que tenham seu foco no significado e não na forma, ou seja, é esperado do aluno que ele demonstre a habilidade de usar a língua e não de usar alguma estrutura específica que ele tenha aprendido. Para que isso ocorra em sala de aula, cabe ao professor criar um ambiente no qual o aluno esteja exposto a diversas situações em que ele tenha que adequar seus conhecimentos da língua para comunicar-se. A forma que Willis encontrou para que haja esse contexto comunicativo em sala de aula é por

⁵⁶ “the lexical syllabus, therefore, affords the learner a coherent learning opportunity. It does not dictate what will be learned and in what order. It offers the learner experience of a tiny but balanced *corpus* of natural language from which it is possible to make generalizations about the language as a whole” (WILLIS, 1990, p. vii)

⁵⁷ “the strategy of highlighting word meanings is a much more productive one than the strategy of teaching structural patterns” (WILLIS, 1990, p. 19)

meio da metodologia baseada em tarefas (*task-based methodology*), que explicamos na seção anterior, visto que esse tipo de abordagem apresenta uma visão sobre a língua como um sistema de significados. Dessa forma, o aluno desempenha uma tarefa (*task*) cujo foco é produção, ou seja, no resultado, e não na linguagem usada para chegar ao resultado. (WILLIS, 1990).

Segundo Berber Sardinha (2004), as críticas ao Currículo Lexical são que as atividades apresentadas nos livros não são baseadas em tarefas, como prometido, mas sim na forma e muitos exemplos não parecem autênticos.

A segunda proposta de ensino baseada na LC apontada por Berber Sardinha (2004) é a de Michael Lewis (1993; 1997), a Abordagem Lexical (*Lexical Approach*). Assim como o Currículo Lexical, trata-se de uma abordagem que possui como papel central o léxico; a diferença entre as duas abordagens está no fato de que a Abordagem Lexical enfatiza o caráter composto do léxico, que é descrito por meio de porções (*chunks*) lexicogramaticais. Os alunos têm contato com essas porções por meio de textos, registrando-os em cadernos lexicais. Dessa forma, as palavras gramaticais, como *would*, por exemplo, são tratadas como itens lexicais e o ensino de colocações é priorizado. Diferentemente do Currículo Lexical, as atividades utilizadas nessa abordagem não são centradas em tarefas (*task-based*), mas sim na forma (*form-based*); ademais, as atividades envolvem identificação de porções lexicais em textos, comparação de itens, preenchimento de lacunas (*gap filling*), categorização de itens, de acordo com características variadas, sequenciamento de itens lexicais e retirada de itens de uma lista (BERBER SARDINHA, 2004, P. 288).

Quanto às críticas a essa abordagem, Berber Sardinha (2004) menciona seu papel ambíguo em relação à LC, pois está de acordo com ela quanto à lexicogramática e aos padrões, entretanto não trata claramente a questão da descrição da linguagem, sobre como foi feita a pesquisa com os *corpora*, que ferramentas foram usadas na listagem dos itens lexicais ou que critérios foram usados para a identificação das porções e colocações. Outrossim, o conceito de porção é vago e é colocado em prática por meio de atividades baseadas na repetição (*drills*).

A última proposta é uma das mais sólidas para a utilização de *corpora* no ensino. Intitulada por Tim Johns (1991) de *DDL – Data Driven Learning* (Aprendizado movido a dados), trata-se de uma abordagem indutiva cuja ênfase é desenvolver no estudante da língua a habilidade de pesquisa e de descoberta, colocando à sua

disposição dados de um *corpus*. Assim, o aluno é estimulado ao trabalho individual e independente na busca de respostas por meio de concordâncias diretamente no computador ou impressas, preparadas pelo professor no caso da impossibilidade de uso do computador em sala de aula.

McEnery e Wilson (1997) situam o primeiro uso de concordâncias no ensino de línguas em 1969, feito por Peter Roe, na Universidade de Aston, em Birmingham. No entanto, foi somente no final dos anos 80 que essa prática foi apresentada e discutida na comunidade acadêmica, por meio dos trabalhos de Tribble e Jones (1990) e Tim Johns (1986). Johns (1986; 1991) introduziu o DDL no ensino de Inglês para Fins Específicos para alunos não nativos de Língua Inglesa com o uso do concordanciador⁵⁸ *MicroConcord*. Nesta abordagem, os alunos trabalham em sala de aula em computadores acessando um *corpus* por meio do concordanciador para identificar padrões ou aspectos recorrentes, interpretar esses padrões, classificando-os de acordo com critérios próprios e, por fim, fazer generalizações a respeito da ocorrência (BERBER SARDINHA, 2004). Johns (1986) considera o aluno um ‘Sherlock Holmes’, visto que ele é um investigador e descobridor da língua.

As grandes vantagens dessa abordagem são (BERNARDINI, 2004; BERBER SARDINHA, 2004; O’SULLIVAN, 2007; CHAMBERS, 2010; GILQUIN e GRANGER, 2010):

- 1) Uso e exploração de linguagem autêntica na sala de aula e grandes quantidades de exemplos, não para a simples repetição (*drills*), mas sim para o que Ellis (2002, p. 177) chama de repetição consciente (*mindful repetition*), ou seja, a leitura das mesmas instâncias visando o entendimento dos padrões e itens linguísticos. Esse acesso à linguagem autêntica e a uma quantidade de exemplos maior do que pode ser encontrado em livros de referência usados por alunos e professores, como dicionários, gramáticas e livros didáticos, leva à expansão do vocabulário e à percepção e conscientização de padrões lexicogramaticais.
- 2) A modificação do papel do professor de ‘fonte de conhecimento’ para facilitador, auxiliando os alunos a interpretar os dados e a fazer suas buscas

⁵⁸ Um concordanciador é um programa de computador que fornece automaticamente uma concordância.

e o acesso a exemplos de língua em uso para dar suporte ao ensinamento da língua.

- 3) O fato de o aluno desenvolver maior autonomia e independência, além de tomar responsabilidade sobre sua aprendizagem, aprendendo a ser um pesquisador e a refletir sobre as questões da língua; além disso, o elemento da descoberta, que faz a aprendizagem ser mais motivadora e divertida. Isso caracteriza a mudança da aprendizagem dedutiva, na qual a regra é dada ao aluno (*rule-driven*) para a indutiva, na qual o aluno descobre a regra (*rule-discovery*).
- 4) A possibilidade de o aluno corrigir-se comparando seu uso da língua ao uso de nativos ou ótimos usuários (*expert*).
- 5) Desenvolvimento de habilidades de aprendizagem, que envolvem fazer previsões e observações, pensar, notar, analisar, raciocinar, interpretar, deduzir, comparar, entre outras.

A partir dos anos 90, concordanciadores mais sofisticados, como o *Concord* do *WordSmith Tools* (SCOTT, 2008) e o *MonoConc* (BARLOW, 2000) tornaram-se disponíveis a preços mais acessíveis, além de outros, como o *AntConc*⁵⁹, gratuitamente disponíveis. Outrossim, cresceu a possibilidade de consulta a grandes *corpora online*, como o *British National Corpus*⁶⁰. Isso encorajou pesquisadores e professores a experimentar o DDL em sala de aula, o que levou a algumas mudanças e adaptações da abordagem. A impossibilidade de uso de computadores individuais em sala de aula, por exemplo, levou ao uso de impressos com linhas de concordância selecionadas pelo professor. Esse preparo prévio também ‘soluciona’ uma das questões discutidas com relação ao professor que se sente inseguro quando os alunos pesquisam diretamente no *corpus*, visto que eles podem se deparar com itens não conhecidos pelo professor.

A principal crítica feita a essa abordagem refere-se ao fato de que seu sucesso ocorreu em um contexto específico, universitário, no qual os alunos estariam mais acostumados ao uso do computador e, sendo estrangeiros em um país diferente do seu, estariam abertos a ideias e metodologias inovadoras, conforme aponta Berber Sardinha (2004). Por outro lado, alguns pesquisadores descobriram que, mesmo alunos que mostraram uma reação positiva à abordagem, relataram que a atividade de análise de

⁵⁹ Disponível em <http://www.antlab.sci.waseda.ac.jp/software.html>. Acesso em março de 2011.

⁶⁰ Disponível em <http://www.natcorp.ox.ac.uk/>. Acesso em março de 2011.

corpus consome tempo, é trabalhosa e tediosa (CHENG et al., 2003; CHAMBERS, 2005). Uma solução apresentada por Chambers (2010) é mesclar o estudo do *corpus* ao estudo de textos; para tanto, os alunos fazem pesquisas em *corpora* que permitem o acesso ao texto completo ou usam *corpora* coletados pelo professor e/ou pelos alunos, como um *corpus* pedagógico, por exemplo.

O uso de *corpora* coletados pelo professor e/ou alunos ou cujos textos estão disponíveis foi testada e é defendida por pesquisadores que utilizaram esse recurso, como Stevens (1991), Cobb (2004) e O'Sullivan e Chambers (2006) e também resolve outra questão levantada, a do contexto, visto que o aluno pode acessar o texto completo e não só a linha de concordância. Esses *corpora*, que explicamos anteriormente como *corpora* especializados, podem ser menores e focar as necessidades dos alunos. Com relação a essa prática, Gilquin e Granger (2010) destacam o fato de ser trabalhoso e demandar tempo (*time consuming*). Ademais, o professor precisa ter um bom conhecimento e habilidade para essa atividade.

Bernardini (2004, p. 22), por sua vez, seguindo as premissas do DDL, propõe o DL (*Discovery Learning*), uma abordagem na qual os alunos “são guiados a observar grandes coletâneas de textos variados de uma maneira aberta e exploratória”⁶¹. A autora opta pelo acesso direto dos alunos a *corpora*; entretanto, diferentemente de Johns (1991), considera que os alunos devam assumir o papel de descobridores, e não de pesquisadores.

Enquanto a abordagem de aprendizagem como pesquisa favorecida por Johns (1991) e Gavioli (2000) presumem (implicitamente) que os alunos dividam os mesmos interesses, competências e capacidades de professores e linguistas (adultos), a visão da aprendizagem como descoberta não faz tal alegação.⁶² (BERNARDINI, 2004, p. 22-23)

Ao contrário, para a autora, a Aprendizagem como Descoberta (*Discovery Learning*) incentiva os alunos a seguir seus interesses e desenvolver capacidades de forma que “suas buscas tornem-se mais focadas, suas interpretações mais precisas e seu

⁶¹ “are guided to browse large and varied text collections in open-ended, exploratory ways” (BERNARDINI, 2004, p. 22)

⁶² “Whereas the learning as research approach favoured by Johns (1991) and Gavioli (2000) implicitly assumes that learners share the same interests, competences and capacities as (adult) teachers or linguists, the learning as discovery view makes no such claim.” (BERNARDINI, 2004, p. 22-23)

entendimento mais afiado sobre o uso de *corpus* e sobre a língua”⁶³ (BERNARDINI, 2004, p. 23). Essa abordagem foi usada com alunos de Língua Inglesa de nível avançado de graduação na Universidade de Bolonha e a autora revela que, apesar de alguns alunos apresentarem uma tecnofobia (*technophobia*), eles demonstram interesse e, aos poucos, ficam mais experientes na exploração de *corpora*.

Conforme aponta Chambers (2010), ninguém defende a ideia do DDL como o principal ou único componente no ensino de línguas, mas sim como um recurso para encorajar a observação da língua em uso, a descoberta, a formulação de hipóteses e, conseqüentemente, a autonomia, não fugindo às principais premissas da abordagem, quais sejam: o princípio idiomático da língua e a descoberta de padrões lexicogramaticais, bem como o uso de concordâncias como instrumento para a observação desses padrões. A maneira como o professor irá usar o *corpus* e as concordâncias irá depender do contexto de aprendizagem, do nível dos alunos, da disponibilidade de materiais, entre outros fatores. Segundo Gilquin e Granger (2010), por exemplo, o uso do DDL ‘puro’ (*hard*) pode apresentar melhores resultados em um contexto de aulas individuais do que em grupos. Além disso, as autoras concluem que os pesquisadores da área também não concordam no que diz respeito ao uso do DDL somente com alunos que possuem conhecimento avançado na língua ou também com iniciantes. Cabe destacar que, apesar das dificuldades, Gilquin e Granger (*idem*) consideram que “existem sinais encorajadores de que o DDL merece um espaço na sala de aula”⁶⁴.

Berber Sardinha (2011) assevera que, ao longo desses 20 anos, a aplicação da LC na sala de aula foi e é fundamentada principalmente na filosofia do *classroom concordancing*, com o uso de concordâncias, como vimos nas propostas apresentadas. O autor indica que existem outras alternativas e propõe três tipos de atividades de ensino com *corpora*, explicando os passos a serem seguidos para sua preparação. As atividades são: centradas na concordância, centradas no texto e multimídia/multigênero.

Para as atividades centradas na concordância, Berber Sardinha (2011) lista o que deve ser levado em conta ao prepará-las, sendo: (a) escolha da palavra de busca, que é guiada pelos objetivos da atividade, da aula, da unidade e do curso; (b) modo de

⁶³ “*their searches become better focused, their interpretation of results more precise, their understanding of corpus use and their language awareness sharper*” (BERNARDINI, 2004, p. 23)

⁶⁴ “*there are encouraging signs that DDL may deserve a place in the language classroom*” (GILQUIN e GRANGER, 2010)

apresentação, ou seja, se a aula será em um laboratório ou na sala de aula, o que implica na necessidade de o professor trazer a concordância pronta impressa ou para projetar; (c) tipo de concordância: geral (com todas as ocorrências), selecionada, editada (sem símbolos ou palavras cortadas, por exemplo), preenchida (com a palavra de busca) ou lacunada (com lacunas no lugar da palavra de busca); (d) quantidade de concordâncias e (e) exercícios, como perguntas, completar frases, ligar colunas, etc. O autor então explica os ingredientes que o professor precisa para preparar a atividade, o *corpus* e um *software* para produção de concordâncias e indica que o professor deve ensinar ao aluno como ler uma concordância e como observar os padrões lexicogramaticais. O professor pode selecionar as concordâncias previamente, atentando a questões como número, seleção e simplificação das linhas; outra opção é o aluno produzir as concordâncias, atividade que requer que o professor ensine o aluno a usar o *software*, formular buscas, selecionar linhas, etc.

Algumas das desvantagens com relação ao uso de concordâncias são: (a) o tempo de preparação das atividades, que é alto, (b) o fato dos alunos ficarem entediados e desmotivados quando expostos continuamente a esse tipo de atividade e (c) o fato de sentirem falta de ler textos inteiros, ao invés de partes. Além disso, as concordâncias podem acabar sendo usadas somente para se encontrar regras gramaticais e não padrões.

Nas atividades centradas no texto, a concordância está presente, mas assume um papel secundário, visto que o foco é no texto. Nesse tipo de atividade, não há a desvantagem do aluno sentir falta do contexto ou de uma situação concreta de uso da língua. Berber Sardinha (2011) apresenta uma atividade desenvolvida por Moreira Filho (2007), que desenvolveu um *software* chamado *Reading Class Builder*, o qual cria automaticamente as atividades, o que reduz o tempo de preparo.

A atividade é baseada em um texto escolhido pelo usuário e é feita uma análise preliminar com informações como tamanho do texto, dificuldade presumida e riqueza vocabular. O usuário pode escolher as palavras de busca para o exercício de concordância, a língua dos enunciados, o *corpus*, entre outras opções. A atividade traz, em primeiro lugar, as palavras mais frequentes e questões para que o aluno faça previsões sobre o assunto do texto; em seguida, o aluno deve passar os olhos no texto para encontrar itens que o ajude a entendê-lo. Outra lista de palavras retiradas do texto é mostrada para o aluno reconhecer se são cognatas ou não e se são familiares a ele; a partir dessas palavras, o aluno é questionado sobre o tipo de texto, seu objetivo e

características. Após a leitura do texto, são feitas perguntas sobre a ideia central e os alunos podem rever suas previsões feitas anteriormente. O aluno então trabalha com algumas palavras do texto, identificando sufixos, prefixos e classe gramatical. Por fim, é feito um exercício com linhas de concordância e o aluno opina sobre o conteúdo do texto.

Dentre as vantagens desse tipo de atividade, citamos a contextualização dos padrões enfocados e o tempo de preparação que, se feito com o uso do *software*, é curto. No entanto, segundo Berber Sardinha (2011), o ensino de padrões fica em segundo plano, o foco das atividades é somente na habilidade de leitura e o *software* apresenta alguns problemas em versões recentes do *Windows*. A atividade pode ser desenvolvida sem o uso do *software*; dessa forma essa última desvantagem pode ser desconsiderada.

A última atividade proposta por Berber Sardinha (idem, p. 329) é a mais inovadora e atual, pois compreende mídias e características centrais das diversas teorias de gênero⁶⁵, e visa “capacitar os alunos a lidar com (produzir e/ou receber) a linguagem envolvida em determinada atividade social”. O desenvolvimento do material proposto seguiu quatro critérios, quais sejam: (a) modularidade, o que significa que o material é dividido em partes que podem ser preparadas independentes umas das outras; (b) replicabilidade, de forma que os enunciados gerais podem ser aplicados a outros materiais, reduzindo o tempo de preparação; (c) curso, de inglês geral, nível avançado e (d) conhecimento prévio, visto que é esperado que o aluno já conheça conceitos como gênero, atividade social, padrão, frequência e significado e que tenha familiaridade com instrumentos da LC, como concordância, palavras-chave, padrões, entre outros.

O material envolve quatro partes:

- 1) Foco na atividade social: a atividade selecionada pelo autor é a de escolher um *videogame* e aprender a jogá-lo melhor. Essas duas atividades sociais abarcam outras, como ler textos ou assistir a vídeos com comentários sobre o jogo (*game reviews*), comprar o jogo, entender como o jogo funciona e como atingir os objetivos, que o usuário encontra na *Internet* em textos ou vídeos com instruções passo a passo (*walkthroughs* e *let's plays*), conhecer melhor a história do jogo e seus criadores em textos enciclopédicos interativos como *wikis*, por exemplo.

⁶⁵ Para saber mais sobre a visão de gênero que envolve essa atividade, ver Berber Sardinha, 2011, p. 324-325.

- 2) Foco no gênero: como vimos no item anterior, o aluno terá acesso a diversos gêneros e poderá explorar seu contexto de uso.
- 3) Foco no texto: estudo da lexicogramática de exemplares de textos que ilustram os gêneros. Os alunos podem gerar as concordâncias diretamente no computador ou o professor as traz para a aula.
- 4) Foco no *corpus*: explorar o *corpus* de um dos gêneros estudados.

As grandes vantagens desse tipo de atividade são o estudo da língua por meio de tarefas do mundo real, o que, conseqüentemente, faz com que o aluno explore diversos gêneros e sua linguagem característica, por meio de padrões. As questões e enunciados são padronizados e podem ser usados na preparação de novas atividades, diminuindo o tempo de preparo. Como contras, Berber Sardinha (2011) aponta a complexidade da preparação, devido ao trabalho com diferentes gêneros e aos recursos utilizados, além da necessidade de o professor conhecer LC e as teorias de gênero; ademais, pode ser vista como uma abordagem para ensinar gêneros ou simplesmente uma atividade comunicativa a partir de um tópico.

A seguir, iremos detalhar o uso de concordâncias no ensino de LI.

1.2.1.2 Concordâncias na sala de aula

Como vimos anteriormente, uma das formas de se utilizar textos autênticos e *corpora* em sala de aula é por meio das concordâncias. Segundo Berber Sardinha (2004, p. 273), “a concordância é empregada para exemplificar o uso (*usage*) de traços linguísticos e as situações nas quais ocorre”. Usando o programa *Wordsmith Tools*, de Mike Scott (2008), por exemplo, podemos obter uma listagem das ocorrências de um item selecionado no conteúdo de vários textos ao mesmo tempo. Assim, a palavra estudada fica centralizada na tela e é possível fazer uma leitura hipertextual das palavras que ocorreram junto a ela no *corpus*. Dessa forma, a palavra está em seu contexto original e, se for necessário, o programa traz uma porção de texto maior da ocorrência a ser estudada.

Segundo Hunston (2002), existem quatro principais tipos de observação que podem ser feitas a partir de linhas de concordância. A primeira observação é com relação à centralidade e à tipicidade. Ao contrário das descrições linguísticas

tradicionais, que focam no que pode ou não ser usado na língua, a LC usa os *corpora* para observar o que é frequente ou o que raramente ocorre no uso da língua.

Corpora não podem ser usados para se determinar o que é impossível em uma língua, sendo que eles não oferecem evidências negativas, e não podem nem ser usados para determinar o que é possível, já que um *corpus* pode muito bem conter expressões que qualquer falante de uma língua rejeitaria como ‘incorreto’.⁶⁶ (HUNSTON, 2002, p. 42)

O exemplo dado por Hunston é de *more healthier*, ao invés de *healthier*, que é usado e considerado correto pelos falantes da língua.

Dessa forma, ao invés de observar o possível, as linhas de concordância fornecem evidências do que é típico, ou seja, dos significados, das colocações e dos padrões de um item. Hunston (2002) afirma que a intuição dos falantes da língua nem sempre está de acordo com o que é apresentado na análise de frequências de palavras. A centralidade está relacionada a categorias e não a palavras individuais. Uma estrutura gramatical, por exemplo, pode indicar vários tempos, mas a frequência do uso em um tempo específico indica a centralidade neste tempo.

A segunda possibilidade de observação de linhas de concordância é das distinções de significado, especialmente com relação a palavras que possuem significados similares. O exemplo dado por Hunston (2002) advém dos estudos de Partington (1998), com relação a palavras que ele considera semigramaticais, como *sheer, pure, complete, utter* e *absolute*. Embora muitos dicionários apresentem algumas como sinônimas, seus colocados refutam essa aparente similaridade.

A terceira forma de observação está relacionada à segunda, pois diz respeito à observação de significado e de padrão. Como vimos, em sua maioria, os significados das palavras são diferenciados pelos padrões em que eles frequentemente ocorrem. Assim, “distinguir entre os significados é uma questão de distinguir entre padrões de uso”⁶⁷ (HUNSTON, 2002, p. 47). Para que isso seja feito, Hunston (2002) divide as linhas de concordância em grupos diferentes, de acordo com os significados dos itens que estão sendo analisados para mostrar que cada significado está claramente associado a um padrão específico.

⁶⁶ “*Corpora cannot be used to determine what is impossible in a language, as they do not offer negative evidence, and they cannot really even be used to determine what is possible, as a corpus may well contain utterances which any speaker of a language would reject as ‘incorrect’.*” (HUNSTON, 2002, p. 42)

⁶⁷ “*distinguishing between the meanings is a matter of distinguishing between patterns of usage*” (HUNSTON, 2002, p. 47)

A quarta e última observação é com relação aos detalhes. Hunston (2002) aponta que as concordâncias nos três tipos de observação descritos acima fornecem ideias mais gerais sobre como as palavras são usadas, sobre seus significados e como eles podem estar associados a padrões. No entanto, a análise por meio de linhas de concordâncias pode ser muito mais detalhada e observações mais específicas podem ser feitas sobre o comportamento de palavras individuais.

O Quadro 1 a seguir mostra linhas de concordância das palavras *ad*, *ads* e *advertising* no *corpus* de estudo. As palavras a serem analisadas aparecem no centro da listagem e podem ser chamadas de nódulos. Fizemos a seleção das linhas que proporcionassem o estudo dos verbos colocados das palavras em questão.

1	omain names has purchased one 30-second ad in the Super Bowl, said Bob Parsons,
2	s-delivery service will run a 30-second ad in the second quarter touting the fac
3	Buy: The website will air one 60-second ad in the second quarter. Creative: The
4	all come together. When people see a TV ad , do they go to the store, go online,
5	is optimized. SEM involves buying the ads -- or sponsored results -- that run
6	ites and then grow a network by placing ads on related sites in their niche. Mea
7	are making money from apps: by selling ads or sponsorships on them. The Style a
8	e views were on YouTube, which can sell advertising against only a very small fr
9	tars and characters from the movie. The ad will include a "modern interpretation
10	sers evaluate in looking to place their ads ; they want to be in contextually rel
11	er. Creative: GoDaddy is producing two ads , both of which will feature race-car
12	idn't have. For example, you can put an ad through Google's TV program and test
13	ther than a medium on which to purchase advertising . The videos were produced
14	th search behavior and say you ran this ad for Product X on Wednesday at 10 and

Quadro 1: linhas de concordância de *ad*, *ads* e *advertising* do *corpus* de estudo.

Com o estudo dessas linhas de concordância podemos verificar que *advertising* e sua abreviação e respectiva flexão ocorre com palavras como *purchase* (1 e 13), *run* (2 e 14), *air* (3), *see* (4), *buy* (5), *place* (6 e 10), *sell* (7 e 8), *produce* (11), *put* (12), *include* (9) e *feature* (11). Observamos também a repetição de *XX-second ad*. Não podemos afirmar se esses são colocados sem analisar uma quantidade maior de linhas e ocorrências. Esse quadro é apenas ilustrativo do que pode ser observado em linhas de concordância.

Berber Sardinha (2004, p. 273) indica que a atividade de estudar concordâncias em sala de aula, além de fornecer exemplos do uso de traços linguísticos, “também tem validade psicológica, pois está relacionada com processos mentais fundamentais envolvidos na aprendizagem”. O autor exemplifica o emprego das concordâncias no ensino das seguintes formas:

- 1) Investigação do aspecto idiomático da linguagem. Pode ser feita a seleção de uma ou mais concordâncias que mostrem a repetição de determinados padrões em contextos similares.
- 2) Levantamento (*Brainstorming*) de tópicos. Uma seleção de concordâncias com exemplos de realização de um determinado assunto em contextos diferentes pode servir de ponto de partida para discussões e composições.
- 3) Meio de investigação de questões retóricas. O exemplo dado pelo autor é o da produção de linhas de concordância com o ponto de interrogação como palavra de busca para o estudo da formação de perguntas, já que os aspectos centrais das interrogativas do inglês ficam em evidência.
- 4) Encorajamento de posição crítica a respeito de dogmas da gramática prescritiva (FLIGELSTONE, 1993, p. 102 *apud* BERBER SARDINHA, 2004, p. 274). As concordâncias podem trazer exemplos contrários à intuição e a regras difundidas, como por exemplo, o uso de algum item na linguagem oral e na linguagem escrita, que encontramos em gramáticas e livros didáticos.

Entretanto, existem duas questões a serem superadas no uso de concordâncias no ensino, que são a necessidade de treinamento e de conscientização (BERBER SARDINHA, 2004). Em primeiro lugar, o professor precisa saber lidar com um programa e estar treinado para o uso e para a análise das linhas de concordâncias para que possa ensinar aos alunos. Como vimos, a leitura das linhas de concordância não é feita de forma linear, mas pode ser feita em qualquer ordem, sendo geralmente guiada pela palavra-chave e pelos itens ao seu redor. A questão mais complicada pode ser com relação à análise. Professores e alunos precisam desenvolver estratégias de análise da linguagem para que possam observar as regularidades e coocorrências e chegar aos padrões. Se o professor não tiver o conhecimento e a experiência para encontrar esses padrões, o aproveitamento do uso das concordâncias pode ser reduzido e os alunos podem desmotivar-se. É importante apontar que esse conhecimento é com relação à

tarefa em si, ou seja, o conhecimento da análise e não da língua, pois se deve ter em mente que “ninguém, nem mesmo as obras de referência mais importantes, possui conhecimento total de uma língua, e há muitas exceções, empregos divergentes e inovações que não constam dessas obras” (BERBER SARDINHA, *idem*, p. 279).

Por isso, a conscientização é muito importante; torna-se necessário que professores e alunos entendam seus papéis: o papel mais ativo do aluno, de explorador e pesquisador; o papel do professor de facilitador e mediador, não de detentor do conhecimento. Berber Sardinha (2004) ressalta que muitos professores e instituições tradicionais ainda não estão preparados para uma metodologia que utilize concordâncias, pois o professor estará delegando a tarefa de encontrar respostas aos alunos, que podem acreditar não ser esse seu papel.

Ainda segundo Berber Sardinha (2004), em contextos nos quais pode haver o uso de concordâncias, é necessária a conscientização dos alunos com relação a dois aspectos: (a) dos benefícios da concordância como instrumento e (b) da utilização da concordância como meio de explorar a linguagem. O aluno deve entender que a concordância permite a compreensão do funcionamento da linguagem, por meio da observação de uma grande quantidade de palavras e estruturas. Os benefícios de seu uso incluem: obtenção de respostas que nem sempre são encontradas em obras de referência, desenvolvimento do espírito pesquisador e de uma atitude mais ativa, maior independência em relação ao professor e ao curso ou ao material didático e centralização no aluno e individualização do aprendizado.

No âmbito dos concordanciadores, podemos observar que estão se tornando cada dia mais acessíveis e simples de serem usados. Mike Scott (2011) relata a trajetória de Tim Johns na idealização do DDL e na criação do primeiro concordanciador que permitiria o esforço do aluno em descobrir a língua ao invés de ser informado sobre ela. Scott (2011), o criador de um dos concordanciadores mais usados e completos, o *WordSmith Tools*, conta que Tim Johns o aconselhou a comprar um dos primeiros microcomputadores disponíveis no mercado, mas, como não encontrava programas (*software*) compatíveis, decidiu aventurar-se a programar. De posse do programa desenvolvido por Tim Johns, o *Micro-concord*, Scott juntou-se a Johns e daí nasceu o *MicroConcord* (SCOTT e JOHNS, 1993).

Conforme mencionamos anteriormente, hoje contamos também com o *AntConc*, um concordanciador gratuitamente disponível, desenvolvido por Laurence

Anthony, da Universidade de Waseda, no Japão, e que tem sido considerado de uso mais fácil (*user-friendly*) do que o *WordSmith*, por ter uma interface que permite o uso de diversas ferramentas em uma só janela.

Quanto aos concordanciadores *online* que permitem o acesso e exploração a *corpora* grandes, de 1 bilhão de palavras, citamos os disponibilizados compilados por Mark Davies, da Universidade Brigham Young⁶⁸. Compreendem cinco *corpora* de Língua Inglesa, um de língua portuguesa e um de espanhol e são usados, de acordo com Davies, por mais de 80 mil pessoas por mês. Dentre os *corpora*, encontramos o BNC (*British National Corpus*) e o COCA (*Corpus of Contemporary American English*), aos quais fizemos menção anteriormente. A interface é bem estruturada e completa e permite não somente observar linhas de concordância, mas também obter informações sobre o uso do item pesquisado ao longo do tempo e nos diferentes tipos de textos que compõem os *corpora*. Outro concordanciador disponível gratuitamente é o *Compleat Lexical Tutor*, que proporciona o acesso a diversos *corpora* de línguas inglesa, francesa, espanhola e alemã; foi desenvolvido por Tom Cobb, da *Université du Québec à Montréal* (UQAM).

1.2.1.3 Uma proposta para o ensino de ILE *in company*

Esta pesquisa envolve o ensino de ILE em um contexto específico, de aulas para um público de uma determinada área profissional e no local de trabalho do aluno. Explicaremos nesta seção o que envolve esse contexto e os pressupostos nos quais as atividades que desenvolvemos foram baseadas.

A modalidade de aulas em empresas, comumente chamada '*in company*' surge para atender às necessidades de alunos de aprendizagem de uma Língua Estrangeira e de empresas, que precisam de funcionários com conhecimento do idioma para se comunicar com clientes, com fornecedores, com a matriz, com filiais ou dentro da própria empresa, como no caso de multinacionais. Algumas empresas oferecem cursos de línguas a seus funcionários apenas como um bônus, por não haver uma exigência imediata de conhecimento e uso do idioma; entretanto, mesmo nesse último caso, espera-se que o funcionário que recebe tal benefício tenha a capacidade de utilizar a

⁶⁸ Disponível em <http://corpus.byu.edu/>. Acesso em março de 2011.

língua em alguma ocasião em que esse uso seja necessário, como em um congresso, para fazer um curso no exterior, receber um visitante, etc.

Abordamos a seguir algumas das questões que envolvem o ensino de inglês *in company*, relacionando-as às nossas decisões quanto ao desenvolvimento das atividades, haja vista que essa ação depende inteiramente dessas questões.

A princípio, citamos a necessidade do aluno de comunicar-se em seu meio profissional, usando vocabulário específico. Verificamos em nossa prática que o ensino do vocabulário da área, de maneira contextualizada ou não, muitas vezes não é uma necessidade, visto que os alunos comumente já sabem esse vocabulário. Em contrapartida, os alunos mostram dificuldades em seu uso, nas colocações e padrões que o envolvem. No livro didático voltado para o ensino de inglês para profissionais de Marketing e Publicidade que encontramos no mercado, observamos que o vocabulário apresentado está presente nos textos, mas não é trabalhado de forma a ajudar o aluno a usá-lo; os exercícios consistem em palavras cruzadas, exercícios de ligar (*match*) e completar, nos quais o aluno descobre a palavra que está sendo trabalhada por meio de definições, e não pelo uso.

Nossa proposta, ao contrário, visa a desenvolver no aluno a capacidade de observação desse vocabulário na língua em uso a fim de detectar os padrões que o envolvem para poder usá-los em sua comunicação. Essa proposta condiz com as abordagens apresentadas anteriormente no que concerne a exposição do aluno a linhas de concordância com questões formuladas pelo professor que irão guiá-lo a encontrar os padrões mais frequentes (JOHNS, 1991; SINCLAIR, 1991; HUNSTON, 2002; BERNARDINI, 2004; BERBER SARDINHA, 2004, 2009, 2011). Concordamos com a ideia de Bernardini (2004) ao tratar o aluno como um descobridor e não como pesquisador, como sugere Johns (1991); entretanto, usamos como base a análise da linguagem por meio de concordâncias, premissa do DDL.

Para que o aluno tenha contato com a língua em uso, é necessária a inserção de materiais autênticos em sala de aula. Conforme aponta Mishan (2005), esta presença do texto autêntico em sala de aula suscitou um longo debate sobre a natureza da autenticidade, relacionada tanto ao texto autêntico em si, quanto à percepção e manipulação do texto pelo leitor (MISHAN, 2005). Nesta segunda visão, a autenticidade está na tarefa e na interação do aluno com o texto. Neste debate, também vieram à tona diversas visões sobre o texto autêntico. Dentre elas, destacamos a

definição de Morrow (1977, p.13, apud MISHAN, 2005): “um texto autêntico é uma porção de língua real, produzida por um falante ou escritor real para um público real e desenhado para transmitir uma mensagem real”⁶⁹. Segundo Mishan (2005), a ‘explosão’ da presença da autenticidade em sala de aula é consequência da relação de dois movimentos. O primeiro é sociológico e diz respeito à revolução das tecnologias da informação e comunicação. Tais tecnologias como canais de TV internacionais, *softwares*, *Internet* e *Intranet* abriram o acesso à língua em uso por meio de todo tipo de texto em todas as línguas. O segundo é pedagógico e baseia-se em uma mudança no ensino-aprendizagem, na transferência da responsabilidade pela aprendizagem para o aluno, ou seja, o ensino não estar mais centrado no professor.

Considerando esta questão da autenticidade em materiais baseados em *corpora*, esta pesquisa considera dois aspectos centrais colocados por Mishan (2005), que trabalha com a autenticidade em materiais didáticos de Língua Estrangeira. O primeiro aspecto é o da motivação; a autora propõe que a motivação é um dos elementos mais importantes para a aprendizagem e concentra-se nos dois tipos de motivação de Gardner e Lambert (1972) que são ‘motivação integrativa’ e a ‘motivação instrumental’. A primeira diz respeito ao desejo do indivíduo de integrar-se à comunidade que usa a língua. A segunda refere-se ao desejo de alcançar algum objetivo prático. A autora argumenta que o uso de textos de língua autêntica está relacionado à motivação integrativa, de forma que estes textos são recursos eficientes para levar o aluno ao entendimento e integração com a cultura em questão. Acreditamos que *corpora* trazem grande motivação ao aluno, especialmente por seus elementos apontados ainda por Mishan (2005), que são: (a) cultura, por incorporarem e representarem a cultura dos falantes nativos da língua estudada, (b) atualidade, por trazerem tópicos atuais e língua em uso relevantes para os alunos, e (c) desafio, por proporcionarem ao aluno o desafio de compreender a língua usada por falantes nativos.

O segundo aspecto é da questão do desafio. Ainda é amplamente discutida a ideia de que o aluno em níveis iniciais no estudo da língua ainda não tem condições para compreender o texto autêntico, o que tem levado professores e autores de materiais didáticos a simplificar os textos e a linguagem usada nestes materiais, acreditando que esta simplificação facilita o aprendizado da língua. No entanto, mais uma vez

⁶⁹ “an authentic text is a stretch of real language, produced by a real speaker or writer for a real audience and designed to convey a real message of some sort” (Morrow, 1977: 13, apud Mishan, 2005).

verificamos tanto a partir de nossa prática, quanto em estudos, que a simplificação não desperta no aluno o desafio de compreender a língua em uso e não promove o enriquecimento de sua linguagem (BERBER SARDINHA, 2009; MOREIRA FILHO, 2007; CONDI DE SOUZA, 2005; BARBOSA, 2004). Dessa forma, acreditamos que os materiais autênticos podem ser utilizados em todos os níveis porque o que deve estar adequado ao nível de conhecimento do aluno é a atividade preparada e não o material em si. Esta linha de argumentação de que a dificuldade do texto é independente da dificuldade da tarefa (CELANI et al., 2005) é utilizada por Berber Sardinha (2009) em uma proposta de atividade com o uso de textos autênticos e *corpora* no ensino de línguas.

As aulas *in company* visam preparar o aluno para as atividades do mundo real que ele poderá vir a enfrentar em seu ambiente profissional. Assim, o material autêntico torna-se uma ferramenta importante. Nosso trabalho propõe a compilação de um *corpus* composto por textos relacionados à área profissional do aluno, de forma a expor a linguagem nela usada e a incentivar a discussão de tópicos relevantes ao aluno, por meio dos textos que compõem o *corpus*. Trata-se, portanto, de uma das propriedades típicas do CLT, trazer a autenticidade para a sala de aula. *Corpora* podem ser representativos ou não de uma língua ou realidade, porém, o importante é que eles proporcionam condições para que os alunos envolvam-se com um discurso real, autenticando-o (GAVIOLI e ASTON, 2001).

Uma vantagem de compilar um *corpus* para o desenvolvimento do material é a do aluno ter acesso não só a trechos, mas também ao texto. Dessa forma, desenvolvemos atividades que focam no uso de padrões lexicogramaticais que o aluno pode observar tanto nas concordâncias como no texto. Ademais, nossa proposta objetiva a discussão de temas relacionados à realidade profissional do aluno e os textos servem de ponto inicial para essas discussões; esse tipo de atividade é geralmente solicitado por alunos nesse contexto de aula. Dessa forma, procuramos tornar a aula mais dinâmica e variada, visto que pesquisadores que aplicaram o DDL revelaram que os alunos sentem-se entediados se forem expostos somente a essa tarefa (CHENG et al., 2003; CHAMBERS, 2005).

As atividades com concordâncias podem ser apresentadas de maneiras diversas de forma a não expor o aluno sempre ao mesmo tipo de atividade e também com o propósito de apresentar o contexto mais explicitamente. Gilquin e Granger (2010)

mostram que, apesar de pesquisadores relatarem que o formato KWIC⁷⁰ torna os padrões mais visíveis, essa variação é interessante para diversificar a forma de apresentar as concordâncias. Pensando nessas questões, sugerimos em uma das atividades concordâncias em frases ao invés do formato KWIC.

As aulas *in company* são normalmente dadas em grupos de alunos menores do que em uma escola de idiomas ou regular. Os grupos variam de 2 a 6 alunos e são ministradas em salas de reunião, nas quais os alunos sentam-se em torno de uma mesa, e não em carteiras. Isso proporciona uma atuação maior do aluno, que tem mais oportunidades de colocar questões que ele tenha sobre a língua e as respostas a essas questões podem ser buscadas em aula, se o professor tiver um computador disponível, o que comumente acontece. Assim, o professor ou o próprio aluno pode buscar as respostas no *corpus* quando a dúvida surgir; o *corpus* e o concordanciador, nesse caso, servem como recursos ‘dormentes’ (*sleeping resources*), como sugere Johns (1988). O autor afirma que assim como se tornou um hábito a presença de dicionários na sala de aula para os alunos consultarem, ele gostaria de ver, algum dia, todas as salas de aula equipadas com um computador para que os alunos fizessem buscas em *corpora*.

O aluno que estuda na empresa geralmente não tem tempo de fazer um curso em uma escola, por isso opta por essa modalidade. É bastante comum o aluno não comparecer à aula devido a alguma atividade profissional que esteja desempenhando no horário determinado. Verificamos, em nossa prática, que quando usamos um material didático regular, ou seja, livros didáticos para o ensino da língua em cursos, os alunos acabavam se sentindo desmotivados por perderem algum conteúdo do livro, especialmente os relacionados à gramática.

Em contrapartida, quando organizamos as aulas baseadas a partir do léxico, utilizando textos e *corpora* para exploração de padrões lexicogramaticais, os alunos mostraram-se mais interessados e motivados, pois, apesar de terem perdido alguma aula, sabiam que isso não iria atrapalhar a continuidade do curso ou implicar em dificuldades futuras. Além disso, os alunos mostravam-se cansados de estudar a gramática de forma tradicional nos livros didáticos repetidamente. Em nossa proposta, os alunos entram em contato com os diversos itens gramaticais apresentados nos livros

⁷⁰ *Key-word-in-context*: formato de apresentação de linhas de concordância no qual a palavra sendo pesquisada fica centralizada e os trechos não consistem em frases completas, mas sim trechos em que algumas palavras podem até aparecer cortadas.

didáticos de maneira mais livre e significativa, em ocasiões diversas e sem um sequenciamento rígido. Em nossa experiência, os alunos inclusive apresentam suas dúvidas que surgem quando estão usando a língua ou lendo o texto e usamos os *corpora* para exemplificar seu uso. Conforme Willis, “considerar o léxico como ponto inicial permite a identificação dos significados e padrões mais comuns e oferece aos alunos uma imagem do que é típico na maneira como o inglês é usado”⁷¹ (WILLIS, 1990, p. 129), o que, a nosso ver, desperta o interesse dos alunos.

Por fim, acreditamos importante mencionar o fato do uso de tecnologia em sala de aula, incorporando o uso de programas para análise de *corpora* compilados pelo professor e/ou pelos alunos, como o *WordSmith Tools* e o *AntConc*, bem como de concordanciadores *online*, para análise em *corpora* de língua geral. Ademais, optamos por incluir em nosso *corpus* transcrições de vídeos para que também pudessem ser usados nas atividades, com o objetivo de desenvolver no aluno a habilidade auditiva e não somente a de leitura. Esse uso de ferramentas, de vídeos e da *Internet* em sala de aula condiz com a proposta de Berber Sardinha (2011) de atividades multimídia.

Finalizamos aqui este capítulo de fundamentação teórica, no qual apresentamos os pressupostos teóricos relevantes para nossa pesquisa e procuramos apontar como eles estão relacionados à proposta feita neste trabalho para o ensino de ILE para publicitários. A seguir, detalharemos a metodologia da pesquisa e do desenvolvimento das atividades propostas.

⁷¹ “taking lexis as a starting point enables us to identify the commonest meanings and patterns and to offer students a picture which is typical of the way English is used” (WILLIS, 1990, p. 129)

CAPÍTULO 2: METODOLOGIA DE COLETA E ANÁLISE DOS DADOS

Há dois aspectos que iremos abordar em nossa metodologia. O primeiro é a coleta dos dados, que compreende considerações sobre os dois *corpora*, o de estudo e o de referência. O segundo refere-se à análise dos dados.

2.1 Coleta dos dados

Explicamos a seguir a composição e os procedimentos de coleta dos dois *corpora* usados na pesquisa, o *corpus* de estudo e o *corpus* de referência.

2.1.1 *Corpus* de estudo

A compilação do *corpus* de estudo, denominado *ad corpus*, foi feita segundo os seguintes critérios levantados por Berber Sardinha (2004): (a) composição, por textos autênticos utilizando linguagem natural, (b) autenticidade, textos produzidos por falantes nativos e em situações reais e (c) seleção criteriosa de acordo com os objetivos da pesquisa.

O *ad corpus* é composto por 1.118 textos em formato eletrônico e tem 1.004.733 palavras (*tokens*). Trata-se de um *corpus* sincrônico de porte médio-grande e consideramos ser especializado, por ser composto por textos de uma área específica para a exploração da linguagem nela usada. Os textos foram coletados de fonte autêntica, da *Internet*, de uma revista digital que publica artigos sobre a atuação das agências de Publicidade na área. O *site* selecionado é o *AdAge – Advertising Agency*⁷², que fornece informações, notícias e análise de dados sobre *marketing* e mídia, possibilitando a comunicação para as comunidades de Publicidade e *Marketing*. A revista teve início em 1930, em formato de jornal e atualmente, em seu formato impresso, é semanal e distribuída em diversos países do mundo, incluindo o Brasil. A revista digital traz publicações em forma de artigos, colunas, vídeos, *blogs*, *podcasts*, entre outros.

⁷² Disponível em <http://adage.com>. Acesso de julho de 2009 a fevereiro de 2012.

A escolha por essa fonte para o *corpus* deve-se a diversos fatores. O primeiro é a questão da disponibilidade e variedade de tipos de textos, de autores e de informações. Em segundo lugar, verificamos que a maior parte dos profissionais da área a consideram uma importante fonte de informações e a utilizam diariamente. Por último, em nossa experiência com aulas nesse contexto, observamos que diversos professores de ILE que trabalham com esses profissionais utilizam essa revista em suas aulas e essa prática é aprovada e até mesmo solicitada pelos próprios alunos, que geralmente desejam ter aulas com temas e linguagem relacionados à área em que atuam. Anteriormente a essa pesquisa, utilizamos textos dessa revista e de diversas outras fontes em nossas aulas; de acordo com os alunos, em conversas informais em sala de aula, a *Adage* é a mais completa e eles sentiram que o uso dela em sala de aula foi bastante proveitoso para sua aprendizagem do idioma.

Fizemos a seleção dos textos no período de setembro de 2009 a fevereiro de março de 2011 e os textos selecionados datam de maio de 2009 a março de 2011. Selecionamos textos gratuitamente disponíveis e salvos em arquivos do processador de texto *Microsoft Word* em formato tipo texto sem formatação (.txt). Como dissemos anteriormente, a revista traz uma diversidade de tipos e categorias em sua publicação. Selecionamos, dentre elas as seguintes:

- 1) Artigos informativos, denominados *stories*, que são publicados semanalmente tanto no formato digital, como impresso e contêm notícias sobre a atuação das agências de Publicidade ao redor do mundo, novas campanhas, eventos de premiação, dentre outros temas;
- 2) Artigos informativos e opinativos, chamados de *blogs*, publicados diariamente. Esses textos apresentam algum tema relacionado a Publicidade juntamente com o comentário e/ou opinião do autor;
- 3) Transcrições de vídeos chamados *3 minute Adage*. Trata-se de vídeos de 3 minutos que apresentam algum tema ou campanha realizada por alguma agência de Publicidade. Eles podem ser assistidos no próprio *site* ou pode ser feito o *download* (baixar) gratuito por meio do *iTunes*⁷³. Vários vídeos também estão disponíveis em outros sites de vídeos, como o *Youtube*. Os vídeos são similares a um jornal de TV, ou seja, é como uma reportagem,

⁷³ *iTunes* é um reprodutor de áudio e vídeo desenvolvido pela [Apple](#), para reproduzir e organizar música digital, arquivos de vídeo e para a compra de arquivos de mídia digital.

um apresentador lê um texto escrito e alguns incluem entrevistas com pessoas que trabalham na área ou nas agências que criaram a campanha sendo apresentada.

Os 1.118 textos que compõem o *corpus* estão divididos em 920 *stories*, 178 *blogs* e 20 vídeos. Sabemos que existem diferenças no uso da língua entre os textos das *stories* e *blogs* e os textos transcritos. No entanto, em uma comparação, verificamos que não havia uma grande diferença com relação aos padrões distintivos da área nos quais focamos esta pesquisa. Assim, optamos por manter essa quantidade menor de textos transcritos, devido ao tempo limitado para as transcrições e para que pudéssemos também desenvolver uma atividade com o uso de um vídeo.

A coleta dos textos foi feita da seguinte forma: com relação aos textos escritos, utilizamos o processo de selecionar, copiar e colar cada texto em um arquivo do *Microsoft Word*, tanto em formato de documento do *Word* como em formato texto sem formatação (.txt). Salvamos nos dois formatos para ter a possibilidade de acessar os textos com as figuras e outros itens que são eliminados no formato .txt com mais facilidade, quando fôssemos usar os textos nas atividades. Os vídeos foram assistidos e transcritos por nós e salvos da mesma maneira que os textos escritos. Nomeamos os arquivos da seguinte forma:

- 1) Artigos (*stories*): *st_* + numeração. Exemplo: *st_0001*, *st_0002*, etc.
- 2) Artigos opinativos (*blogs*): *bl_* + numeração. Exemplo: *bl_0001*, *bl_0002*, etc.
- 3) Vídeos: *vi_* + numeração. Exemplo: *vi_0001*, *vi_0002*, etc.

2.1.2 *Corpus* de referência

O *corpus* de referência serve de contraponto ao *corpus* de estudo, para fins de comparação e de extração dos padrões lexicogramaticais mais típicos do *corpus* de estudo, que emergem a partir do contraste dos dois *corpora*. Segundo Berber Sardinha (2009, p. 226-227), um *corpus* de referência, também denominado *corpus* de controle, normalmente possui as seguintes características:

- 1) É pelo menos 5 vezes maior do que o *corpus* de estudo, já que tem a função de representar as frequências da língua de um modo geral; o tamanho maior

também garante que o maior número de palavras do *corpus* de estudo esteja representada no *corpus* de referência.

- 2) Não deve conter o *corpus* de referência para que não seja feita uma comparação do *corpus* de estudo com ele mesmo.
- 3) Deve ser selecionado com um propósito específico para que não sejam comparadas listas de palavras que não tenham relação alguma uma com a outra; “um *corpus* de referência geral, que inclua vários gêneros, é tida como a escolha não-marcada para estudos de palavras-chave”.

O *corpus* de referência selecionado nesta pesquisa foi o *British National Corpus*, doravante *BNC*, composto por inglês escrito e falado, com cerca de 100 milhões de palavras (*tokens*). A parte escrita inclui textos de jornais, livros acadêmicos e de ficção, cartas, redações, entre outros. A parte falada consiste em transcrições de conversas informais, de idade, regiões e classes sociais diferentes, reuniões governamentais, programas de rádio, entre outros.

Essa escolha deveu-se ao fato desse *corpus* ser referência da Língua Inglesa em uso. Além disso, no momento da pesquisa, era um dos únicos *corpora* disponíveis gratuitamente para uso em programas computacionais. A lista de palavras apresenta 99.465.296 ocorrências (*tokens*) e 512.588 itens (*types*)⁷⁴.

2.2 Procedimentos de análise dos dados

Nesta seção expomos os procedimentos de análise dos dados provenientes do *corpus* de estudo e de sua comparação com o *corpus* de referência.

Um *corpus* não traz informações novas sobre a linguagem, é apenas uma coletânea de exemplos, um banco de dados da língua em uso. O ‘novo’ vem dos programas computacionais que nos oferecem possibilidades diversas de análise da língua, ou seja, uma nova perspectiva do que já conhecemos (HUNSTON, 2002). Os *corpora* foram, portanto, analisados por meio da ferramenta computacional *WordSmith Tools*, versão 5.0 (SCOTT, 2008), que oferece meios de analisar *corpora* em busca da padronização lexicogramatical.

⁷⁴ Em “Paulo viajou com o seu pai e com o seu irmão”, por exemplo, temos 11 ocorrências (*tokens*) e 8 itens (*types*).

O primeiro passo para a análise foi o uso da ferramenta *WordList*, que faz uma lista de palavras de um *corpus*.

Ao fazer uma lista de palavras, o programa produz três abas principais:

- 1) dados estatísticos do *corpus* (*statistics*) (Ver Figura 2);
- 2) uma lista de palavras por ordem alfabética (*alphabetical*) (Ver Figura 3);
- 3) uma lista de palavras por frequência (*frequency*) (Ver Figura 4).

Na lista de estatísticas (*statistics*) podemos verificar o tamanho total do *corpus* de estudo. Como vemos na Figura 2, o *corpus* de estudo é constituído de 1.004.733 *tokens* ou ocorrências, e 32.166 *types*, que chamamos de itens. A proporção item/ocorrência (*Type/Token Ratio*) é 3,27. Essa proporção é feita dividindo-se o número de itens pelo número de ocorrências; o resultado é multiplicado por 100 para transformar o valor em porcentagem. Isso significa que quanto maior o valor, mais palavras diferentes e quanto menor o valor, maior o número de repetições; esse valor pode ser interpretado para verificar a riqueza lexical de um texto ou *corpus*. (BERBER SARDINHA, 2009)⁷⁵.

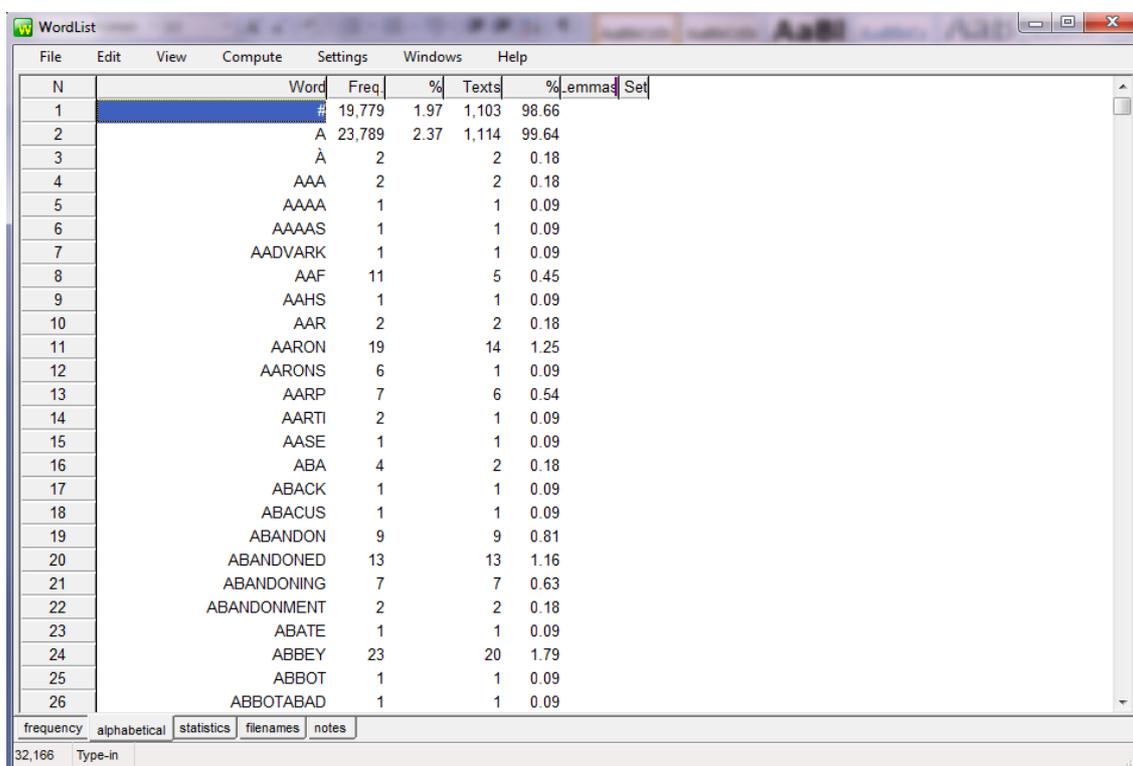
	N	Overall	1	2	3	4	5	6	7
text file		Overall	bl_0001.txt	bl_0002.txt	bl_0003.txt	bl_0004.txt	bl_0005.txt	bl_0006.txt	bl_0007.txt
file size	6,003,697	1,580	2,562	1,022	907	552	1,493	1,137	
tokens (running words) in text	1,004,733	272	463	184	160	90	254	206	
tokens used for word list	984,954	266	452	171	155	85	249	199	
sum of entries									
types (distinct words)	32,166	186	258	109	108	75	165	136	
type/token ratio (TTR)	3.27	69.92	57.08	63.74	69.68	88.24	66.27	68.34	
standardised TTR	45.63								
standardised TTR std.dev.	54.61								
standardised TTR basis	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	
mean word length (in characters)	4.75	4.58	4.35	4.34	4.42	4.89	4.66	4.24	
word length std.dev.	2.54	2.48	2.26	2.47	2.38	2.85	2.38	2.50	
sentences	51,878	19	25	7	12	3	13	10	
mean (in words)	18.99	14.00	18.08	24.43	12.92	28.33	19.15	19.90	
std.dev.	12.91	16.71	13.64	12.34	14.83	8.50	14.38	12.68	
paragraphs	1,118	1	1	1	1	1	1	1	
mean (in words)	881.00	266.00	452.00	171.00	155.00	85.00	249.00	199.00	
std.dev.	430.32								
headings									
mean (in words)									
std dev.									

Figura 2: tela do *WordList* do *Wordsmith Tools* com resultado estatístico do *corpus* de estudo.

⁷⁵ As demais informações apresentadas na janela de estatística não serão aqui abordadas; para explicações detalhadas sobre elas, vide Berber Sardinha (2009).

Nas listas em ordem alfabética e de frequência (Ver Figuras 3 e 4) encontramos as seguintes informações:

- 1) *Word*: palavra ou item;
- 2) *Freq.*: frequência das palavras no *corpus* de estudo;
- 3) %: porcentagem da frequência das palavras considerando a porcentagem do total de ocorrências das palavras;
- 4) *Texts*: quantidade de textos nos quais a palavra ocorre;
- 5) %: porcentagem da frequência de ocorrência das palavras nos textos.



The screenshot shows the WordList application window with a menu bar (File, Edit, View, Compute, Settings, Windows, Help) and a table of word frequency data. The table has columns for N, Word, Freq, %, Texts, %_lemmas, and Set. The data is sorted alphabetically by word.

N	Word	Freq	%	Texts	%_lemmas	Set
1	#	19,779	1.97	1,103	98.66	
2	A	23,789	2.37	1,114	99.64	
3	À	2		2	0.18	
4	AAA	2		2	0.18	
5	AAAA	1		1	0.09	
6	AAAAA	1		1	0.09	
7	AADVARK	1		1	0.09	
8	AAF	11		5	0.45	
9	AAHS	1		1	0.09	
10	AAR	2		2	0.18	
11	AARON	19		14	1.25	
12	AARONS	6		1	0.09	
13	AARP	7		6	0.54	
14	AARTI	2		1	0.09	
15	AASE	1		1	0.09	
16	ABA	4		2	0.18	
17	ABACK	1		1	0.09	
18	ABACUS	1		1	0.09	
19	ABANDON	9		9	0.81	
20	ABANDONED	13		13	1.16	
21	ABANDONING	7		7	0.63	
22	ABANDONMENT	2		2	0.18	
23	ABATE	1		1	0.09	
24	ABBAY	23		20	1.79	
25	ABBOT	1		1	0.09	
26	ABBOTABAD	1		1	0.09	

At the bottom of the window, there are tabs for 'frequency', 'alphabetical', 'statistics', 'filenames', and 'notes'. The 'frequency' tab is selected. The status bar at the bottom left shows '32,166 Type-in'.

Figura 3: tela do *WordList* do *Wordsmith Tools* com resultado parcial de lista de palavras do *corpus* de estudo em ordem alfabética.

N	Word	Freq	%	Texts	%_emmas	Sel
1	THE	48,322	4.81	1,118	100.00	
2	TO	28,187	2.81	1,116	99.82	
3	AND	26,029	2.59	1,113	99.55	
4	A	23,789	2.37	1,114	99.64	
5	OF	23,670	2.36	1,117	99.91	
6	#	19,779	1.97	1,103	98.66	
7	IN	17,757	1.77	1,115	99.73	
8	THAT	11,428	1.14	1,105	98.84	
9	FOR	10,955	1.09	1,107	99.02	
10	IS	10,042	1.00	1,089	97.41	
11	ON	8,178	0.81	1,111	99.37	
12	WITH	7,556	0.75	1,086	97.14	
13	IT	7,395	0.74	1,080	96.60	
14	AS	6,986	0.70	1,063	95.08	
15	ARE	5,718	0.57	1,037	92.75	
16	SAID	5,417	0.54	818	73.17	
17	BUT	5,243	0.52	1,073	95.97	
18	AT	5,188	0.52	1,052	94.10	
19	BE	4,824	0.48	1,030	92.13	
20	HAVE	4,521	0.45	1,016	90.88	
21	FROM	4,328	0.43	1,024	91.59	
22	HAS	4,179	0.42	980	87.66	
23	ITS	4,127	0.41	903	80.77	
24	MORE	4,115	0.41	969	86.67	
25	BY	4,014	0.40	1,052	94.10	
26	AN	3,907	0.39	1,012	90.52	
27	THIS	3,706	0.37	999	89.36	
28	WAS	3,680	0.37	886	79.25	
29	OR	3,633	0.36	944	84.44	
30	YOU	3,599	0.36	850	76.03	
31	NEW	3,517	0.35	931	83.27	
32	NOT	3,372	0.34	980	87.66	
33	MEDIA	3,350	0.33	764	68.34	

Figura 4: tela do *WordList* do *WordSmith Tools* com resultado parcial de lista de palavras do *corpus* de estudo por frequência.

Podemos observar que na lista das palavras mais frequentes do *corpus* de estudo encontramos artigos, preposições, pronomes e verbos, sendo que o primeiro item relevante para estudo dentro da área de Publicidade, *media*, está na 33ª posição. Para a observação e seleção das palavras mais representativas no *corpus* de estudo, fizemos a comparação entre a lista de palavras dele e do *corpus* de referência especificado anteriormente.

Dessa forma, o segundo passo para a análise foi o uso da ferramenta *Keywords*, “que permite a seleção de itens de uma lista de palavras (ou mais) por meio da comparação de suas frequências com uma lista de referências” (BERBER SARDINHA, 2004, p. 96). A Figura 5 mostra uma janela com uma parte dos resultados.

N	Key word	Freq.	%	RC. Freq.	RC. %	Keyness	P	Lemmas
1	AD	2,987	0.30	3,078	19,175.07	0.0000000000		
2	MEDIA	3,350	0.33	7,862	17,348.12	0.0000000000		
3	MARKETING	2,821	0.28	5,198	15,692.83	0.0000000000		
4	BRAND	2,283	0.23	2,032	15,105.70	0.0000000000		
5	MARKETERS	1,390	0.14	61	12,299.38	0.0000000000		
6	FACEBOOK	1,139	0.11	0	10,491.78	0.0000000000		
7	COM	1,212	0.12	281	9,725.86	0.0000000000		
8	BRANDS	1,355	0.13	770	9,714.46	0.0000000000		
9	ADVERTISING	1,872	0.19	4,400	9,687.03	0.0000000000		
10	CONSUMERS	1,525	0.15	2,290	8,959.66	0.0000000000		
11	ADS	1,220	0.12	599	8,944.69	0.0000000000		
12	ONLINE	1,185	0.12	597	8,654.87	0.0000000000		
13	AGENCY	1,804	0.18	5,645	8,484.58	0.0000000000		
14	GOOGLE	905	0.09	0	8,336.11	0.0000000000		
15	CEO	886	0.09	93	7,548.24	0.0000000000		
16	DIGITAL	1,279	0.13	1,915	7,519.67	0.0000000000		
17	TV	1,610	0.16	6,294	6,966.77	0.0000000000		
18	TWITTER	727	0.07	11	6,582.24	0.0000000000		
19	ADAGE	692	0.07	115	5,715.39	0.0000000000		
20	U	1,210	0.12	4,344	5,410.42	0.0000000000		
21	ADVERTISERS	644	0.06	310	4,734.98	0.0000000000		

Figura 5: tela do *KeyWords* do *WordSmith Tools* com resultado parcial de chavidade (*keyness*) da comparação entre a lista de palavras do *corpus* de estudo e da lista de palavras do *corpus* de referência.

As informações principais da lista de palavras-chave (Ver Figura 5) são:

- *Keyword*: palavras-chave do *corpus* de estudo;
- *Freq.*: a frequência das palavras-chave no *corpus* de estudo;
- *%*: a porcentagem frequência das palavras-chave no *corpus* de estudo;
- *RC. Freq.*: a frequência da palavra-chave no *corpus* de referência;
- *Keyness*: chavidade, ou seja, o resultado da estatística de comparação (*log-likelihood*);
- *P*: valor da estatística *p*, que indica o grau de significância da estatística de comparação.

As palavras-chave aparecem em ordem de chavidade, de forma que no topo da lista estão as positivas, ou seja, as que possuem frequência significativamente mais alta no *corpus* de estudo do que no de referência. Quando o contrário acontece, as palavras-chave são denominadas negativas.

A partir da lista de palavras-chave, selecionamos nesta pesquisa as mais frequentes, tendo em mente a transposição dos achados para atividades didáticas. A análise foi feita com a ferramenta *Concord* do *WordSmith Tools*. A Figura 6 mostra a aba *Collocates*, na qual podemos verificar as 5 palavras à direita e à esquerda de *ad* que

ocorrem com maior frequência. Na aba *Patterns*, também podemos verificar essas colocações, ou seja, as palavras que mais ocorrem juntamente com *ad*, como vemos na Figura 7.

N	Word	Ratio	Texts	Total	Left	Right	L5	L4	L3	L2	L1	Centre	R1	R2	R3	R4	R5
1	AD	0.000	746	3,063	44	44	11	13	11	8	1	2,975	1	8	11	13	11
2	THE	0.000	498	1,253	779	474	132	106	125	127	289	0	5	83	151	131	104
3	TO	0.000	359	657	369	288	63	60	95	73	78	0	15	56	62	59	96
4	A	0.000	334	602	354	236	51	79	98	126	0	12	7	52	76	48	53
5	AND	0.000	337	579	278	301	46	64	56	61	51	0	34	112	39	54	62
6	AGE	0.000	243	553	11	542	3	2	2	2	2	0	540	0	0	1	1
7	OF	0.000	325	539	360	179	52	53	67	102	86	0	4	14	42	49	70
8	IN	0.000	306	522	291	231	36	56	49	94	56	0	31	80	39	42	39
9	FOR	0.000	225	328	158	170	28	39	23	38	30	0	43	67	18	21	21
10	AN	0.000	216	322	270	52	11	11	22	17	209	0	2	9	16	17	8
11	THAT	0.000	195	273	123	150	17	17	22	36	31	0	28	51	26	25	20
12	ON	0.000	182	239	95	144	13	23	21	19	19	0	23	37	28	20	36
13	WITH	0.000	172	229	129	100	20	23	15	31	40	0	5	40	14	24	17
14	IS	0.000	155	207	69	138	18	21	17	7	6	0	29	59	23	13	14
15	AS	0.000	115	155	93	62	19	33	19	13	9	0	4	17	17	12	12
16	FROM	0.000	135	152	60	92	12	5	7	19	17	0	13	34	13	15	17
17	SPENDING	0.000	73	140	7	133	2	2	2	0	1	0	130	0	1	1	1
18	AGENCY	0.000	92	131	14	117	2	7	3	2	0	0	79	9	11	10	8
19	ARE	0.000	111	131	45	86	12	13	12	7	1	0	1	37	22	20	6
20	IT	0.000	102	128	62	66	22	19	11	9	1	0	3	12	26	12	13
21	BY	0.000	104	116	45	71	4	8	6	5	22	0	4	22	17	9	19
22	YOU	0.000	58	113	29	84	10	11	5	1	2	0	1	19	26	22	16
23	HAS	0.000	98	110	37	73	12	10	8	4	3	0	4	35	18	9	7
24	NETWORKS	0.000	58	106	17	89	4	6	5	1	1	0	76	1	3	2	7
25	ITS	0.000	97	106	72	34	8	9	14	12	29	0	0	7	10	11	6

Figura 6: tela do *Concord*, aba *Collocates*, do *Wordsmith Tools* com os colocados de *ad*.

N	L5	L4	L3	L2	L1	Centre	R1	R2	R3	R4	R5	
1	THE	THE	THE	THE	THE	AD	AGE	AND	THE	THE	THE	
2	TO	A	A	A	AN	A	SPENDING	THE	A	TO	TO	
3	OF	AND	TO	OF	OF		AGE'S	IN	TO	AND	OF	
4	A	OF	OF	IN	TO		AGENCY	FOR	IN	OF	AND	
5	AND	TO	AND	TO	IN		NETWORKS	IS	OF	A	A	
6	IN	IN	IN	AND	AND		REVENUE	TO	AND	IN	IN	
7	FOR	FOR	FOR	FOR	TOLD		DOLLARS	A	ON	THAT	ON	
8	IS	AS	ON	SUPER	ONLINE		CAMPAIGN	THAT	IT	ON	THAT	
9	IT	ON	THAT	THAT	WITH		AGENCIES	WITH	IS	FOR	WITH	
10	WITH	IT	AS	ACCORDING	BOWL		NETWORK	ON	THAT	WITH	FOR	
11	AS	WITH	AN	WITH	THAT		INDUSTRY	BUT	YOU	YOU	YEAR	
12	THAT	IS	IS	U	TV		SALES	ARE	FOR	ARE	BY	
13	ON	AD	ITS	ON	MOBILE		FOR	HAS	ARE	AD	IS	
14	WAS	THAT	WITH	FROM	THIS		AND	FROM	THIS	IS	FROM	
15	BE	ARE	YEAR	30	FOR		SUPPORTED	WHAT	THEY	AN	IT	
16	HAS	WILL	ARE	AN	ITS		IN	BY	AS	IT	AS	
17	FROM	HAVE	AD	IT	THEIR		CAMPAIGNS	AT	CENTER	HAS	FROM	YOU
18	IT'S	AT	FIRST	AS	NEW		BUSINESS	HOW	AN	ITS	WHO	
19	ARE	NOT	BE	ITS	BY		THAT	AT	WITH	WAS	AD	
20	YOU	MORE	ACCORDING	AD	S		IS	YOU	BY	HAS	AT	
21	AN	AN	IT	THEIR	DISPLAY		INVENTORY	IT	WAS	MORE	BUT	
22	AD	WAS	AT	NEW	MORE		BREAKS	SO	HAVE	MEDIA	ARE	
23	MR	YOU	THIS	AT	SECOND		BUYERS	WILL	SAID	AS	THIS	
24	HAVE	IT'S	MORE	BILLION	ON		MARKET	LAST	FROM	BY	WHICH	
25	MEDIA	BUT	HAVE	WHEN	VIDEO		PAGES	HAVE	PAPER	BE	WILL	

Figura 7: tela do *Concord*, aba *Patterns*, do *Wordsmith Tools* com os padrões de *ad*.

Ainda dentro da ferramenta *Concord*, analisamos os padrões lexicogramaticais típicos com as palavras selecionadas por meio de cálculos estatísticos para medir o grau de associação entre os itens, ou seja, o grau de coocorrência entre um nóculo e seus colocados.

Dentre as medidas de associação disponíveis, utilizamos a Informação Mútua (*Mutual Information*) e o Escore T (*T Score*). Ambas as medidas consideram a razão O/E, ou seja, Observado/Esperado, levando em conta, segundo Berber Sardinha (2004, p. 201), “quantas vezes de fato duas palavras ocorreram juntas dentro do horizonte delimitado (o valor observado) e quantas vezes seria esperado que elas ocorressem (o valor esperado), dado o tamanho do *corpus* e a frequência de cada palavra”.

O cálculo de Informação Mútua, doravante IM, identifica as associações que possuem ocorrências não-aleatórias, ou seja, constantes, mas não tão frequentes na língua (BERBER SARDINHA, 2004). Um exemplo retirado do nosso *corpus* de estudo é *ad network*. A fórmula para o cálculo da IM é:

$I = \log_2 O/E$	O = valor observado E = valor esperado
------------------	---

O valor de O (observado) é calculado com a seguinte fórmula:

$O = f(n,c)/N$	$f(n,c)$ = frequência de ocorrência mútua do nóculo (n) e do colocado (c) N = tamanho do <i>corpus</i>
----------------	---

O valor de E (esperado) é calculado com a seguinte fórmula:

$E = f(n)/N * f(c)/ N$	$f(n)$ = frequência do nóculo no <i>corpus</i> $f(c)$ = frequência do colocado no <i>corpus</i> N = tamanho do <i>corpus</i>
------------------------	--

O Escore T, por sua vez, identifica as associações não tão constantes, porém mais frequentes na língua (BERBER SARDINHA, 2004). Um exemplo do corpus de estudo desta pesquisa é *ad to*. A fórmula para o cálculo do Escore T é:

$T = (O - E) / (\text{raiz quadrada de } f(n,c)/N)$	$f(n)$ = frequência do nóculo no <i>corpus</i> $f(c)$ = frequência do colocado no <i>corpus</i> N = tamanho do <i>corpus</i>
---	--

Segundo Berber Sardinha (2004), podemos afirmar, com base nos cálculos de IM, que uma associação entre duas palavras é, de fato, uma colocação, se o resultado for superior a 3. Já no caso do Escore T, uma palavra é considerada colocada de um nóculo quando o resultado for superior a 2. Levando em conta esses cálculos, selecionamos em nossa análise os colocados que tivessem resultado de IM (Informação Mútua) acima de 3 e de Escore T acima de 2.

2.3 Análise dos dados

Neste capítulo apresentaremos a análise dos dados pertinentes para o desenvolvimento das atividades.

Observamos na Tabela 1 a seguir as 20 primeiras palavras-chave do *corpus* de estudo em ordem de chavicidade.

	Palavra	Frequência no <i>corpus</i> de estudo	% estatística da frequência no <i>corpus</i> de estudo	Frequência no <i>corpus</i> de referência	% estatística da frequência no <i>corpus</i> de referência	Chavicidade (<i>log-likelihood</i>)
1	ad	2.987	0,30	3.078	0,0030945	19.175,07
2	media	3.350	0,33	7.868	0,0079102	17.348,12
3	marketing	2.821	0,28	5.198	0,0052259	15.692,83
4	brand	2.283	0,23	2.032	0,0020392	15.105,70
5	marketers	1.390	0,14	61	0,0000613	12,299.38
6	facebook	1.139	0,11	0	0	10.491,78
7	com	1.212	0,12	281	0,0002825	9.725,86
8	brands	1.355	0,13	770	0,0007741	9.714,46
9	advertising	1.872	0,19	4.400	0,0044236	9.687,03
10	consumers	1.525	0,15	2.290	0,0023023	8.959,66
11	ads	1.220	0,12	599	0,0006022	8.944,69
12	online	1.185	0,12	597	0,0006002	8.654,87
13	agency	1.804	0,18	5.645	0,0056753	8.484,58
14	google	905	0,09	0	0	8.336.11
15	ceo	886	0,09	93	0,0000934	7.548,24
16	digital	1.279	0,13	1.915	0,0019252	7.519,67
17	tv	1.610	0,16	6.294	0,0063278	6.966,77
18	twitter	727	0,07	11	0,000011	6.582,24
19	adage	692	0,07	115	0,0001156	5.715,39
20	u	1.210	0,12	4.344	0,0043673	5.410,42

Tabela 1: resultado da extração de palavras-chave - 25 primeiras em ordem de chavicidade (*keyness*).

O valor p usado em testes estatísticos varia de 0 a 1. O valor .01, por exemplo, indica que existe 1% de possibilidade de erro em se afirmar a existência de uma relação, ou seja, pode ter ocorrido ‘por acaso’. Todas as palavras da Tabela 1 possuem valor menor do que 0.000001, o que mostra que elas possuem grande significância, visto que a probabilidade de a diferença ter ocorrido por acaso é muito baixa.

A seleção dos itens foi feita de acordo com o objetivo da pesquisa, que foi o de proporcionar o estudo do uso do vocabulário da área de Publicidade. Assim, optamos por eliminar as palavras que nomeiam *sites* da *Internet* e redes de relacionamentos, como *Facebook*, *Google* e *Twitter*. Apesar de relevantes para o ensino da língua, preferimos desenvolver, neste trabalho, atividades com palavras mais específicas da área. A palavra *com* no *corpus* de estudo refere-se, em geral, ao domínio que compõe um endereço da *Internet*; *adage* refere-se ao nome da revista e ‘*u*’ refere-se a ‘*U.S.*’ – Estados Unidos (*United States*) ou ‘*U.K.*’ – Reino Unido (*United Kingdom*).

A análise dos itens foi feita por meio da observação dos resultados dos cálculos de Informação Mútua e Escore T, como explicamos na seção 2.2, e das linhas de concordância.

2.3.1 Análise de *ad*

A primeira palavra selecionada para análise foi *ad*, por estar na primeira posição na lista de palavras-chave quando da comparação do *ad corpus* com o *corpus* de referência. Para tanto, usamos a ferramenta *Concord*, do *Wordsmith Tools*, como especificado na seção 2.2. Dentre elas, observamos que as que mais ocorrem na posição L1, ou seja, logo à esquerda de *ad* são palavras gramaticais, isto é, determinantes e preposições, como *the*, *to*, *an* e *of*. À direita, encontramos *age* com maior frequência, o que pode ser explicado pelo nome da revista (*Adage*). A seguir, vêm as palavras *spending*, *agency* e *networks*. A tabela a seguir apresenta os colocados de *ad* em ordem de frequência, considerando-se MI acima de 3 a Escore T acima de 2, como dito anteriormente:

	COLOCADO	FREQUÊNCIA	IM	ESCORE T
1	<i>to</i>	657	3,030	25,863
2	<i>a</i>	602	3,157	25,304
3	<i>and</i>	579	2,834	24,276
4	<i>age</i>	553	9,700	23,488
5	<i>in</i>	522	3,116	22,101
6	<i>for</i>	328	3,589	18,990
7	<i>an</i>	322	4,941	17,360
8	<i>that</i>	273	3,068	17,165
9	<i>on</i>	239	3,401	16,789
10	<i>with</i>	229	3,487	13,783
11	<i>from</i>	152	3,257	11,260
12	<i>spending</i>	140	9,422	11,815
13	<i>agency</i>	131	9,551	11,430
14	<i>are</i>	131	3,207	10,206
15	<i>has</i>	110	3,814	9,742
16	<i>its</i>	106	4,415	9,813
17	<i>networks</i>	106	10,819	10,290
18	<i>revenue</i>	100	9,656	9,988
19	<i>age's</i>	98	18,135	9,899
20	<i>new</i>	91	4,563	9,136
21	<i>more</i>	86	3,728	8,574
22	<i>year</i>	81	5,251	8,764
23	<i>will</i>	80	3,363	8,075
24	<i>campaign</i>	67	8,032	8,684
25	<i>TV</i>	75	8,589	8,638

Tabela 2: colocados principais de *ad* em ordem de frequência.

Nesta pesquisa, nossa proposta foi analisar os colocados que apresentam forte associação; dessa forma, optamos por analisar os que tivessem valores de IM altos. Esses colocados não são tão frequentes na língua, mas possuem ocorrência constante; consideramos também que eles tivessem Escore T maior do que 2, o que indica que não deixam de ser frequentes. A Tabela a seguir mostra os 10 colocados de *ad* de acordo com esses critérios.

	COLOCADO	FREQUÊNCIA NO <i>CORPUS</i>	IM	ESCORE T
1	<i>age's</i>	98	18,135	9,899
5	<i>networks</i>	106	10,819	10,290
1	<i>age</i>	553	9,700	23,448
6	<i>revenue</i>	100	9,656	9,988
4	<i>agency</i>	131	9,551	11,430
2	<i>spending</i>	140	9,422	11,815
	<i>TV</i>	75	8,589	8,638
7	<i>campaign</i>	76	8,032	8,684
8	<i>year</i>	81	5,251	8,764
9	<i>an</i>	322	4,941	17,360
10	<i>new</i>	91	4,563	9,136

Tabela 3: colocados principais de *ad* em ordem de IM.

Dentre esses colocados, encontramos somente uma palavra gramatical, que é *an*; entendemos que esse resultado corrobora nossa ideia inicial de estudar as palavras de conteúdo, por terem uma relação mais direta com o nóculo. Observando as linhas de concordância, vimos que os colocados selecionados possuem frequência alta na posição R1, à direita do nóculo. Fizemos, portanto, uma busca dos colocados nessa posição e encontramos, dentre as 10 primeiras, em ordem de frequência:

	COLOCADO	FREQUÊNCIA NO <i>CORPUS</i>	FREQUÊNCIA NA POSIÇÃO R1	IM	ESCORE T
1	<i>age</i>	553	540	9,700	23,448
2	<i>spending</i>	140	130	9,422	11,815
3	<i>Age's</i>	98	95	18,135	9,899
4	<i>agency</i>	131	79	9,551	11,430
5	<i>networks</i>	106	76	10,819	10,290
6	<i>revenue</i>	100	74	9,656	9,988
7	<i>campaign</i>	76	63	8,032	8,684
8	<i>dollars</i>	67	63	10,374	8,179
9	<i>agencies</i>	67	55	9,325	8,472
10	<i>industry</i>	60	49	6,655	7,699

Tabela 4: colocados principais de *ad* com frequência na posição R1.

Encontramos também *campaigns*, no plural, em 16º lugar e *network*, no singular, na 11ª posição. Observando novamente as linhas de concordância, pudemos concluir que esses padrões apresentam um sentido específico sendo, portanto, relevantes para estudo e, conseqüentemente, para o ensino. O próximo passo foi, então, selecioná-los para uma observação de seu uso. Os padrões *ad age*, *ad age's* e *ad agency* estavam, em sua grande maioria, relacionados ao nome da revista, *AdAge*. Dessa forma, selecionamos os seguintes para a análise: *ad spending*, *ad network(s)*, *ad revenue* e *ad campaign*.

Para analisar o uso desses padrões, utilizamos a ferramenta *Concordance*, que traz as linhas de concordância e os colocados do item selecionado. Vejamos a seguir a análise dos padrões e seus colocados; para essa análise, optamos por considerar a ordem de frequência, pois a intenção neste momento é de entender como os padrões são usados, quais os itens que frequentemente aparecem junto a eles e que relação eles possuem com os padrões. Para cada um dos padrões, analisamos os 5 itens que ocorrem com eles com mais frequência e depois discutiremos os que podem ser relevantes para serem ensinados com os padrões.

2.3.1.1 Análise de *ad spending*

Encontramos 130 linhas de concordância com esse padrão (Ver Anexo 1) e podemos observar na Tabela 5 a seguir seus colocados.

	COLOCADO	FREQUÊNCIA	INFORMAÇÃO MÚTUA	ESCORE T
1	<i>in</i>	54	9,430	7,338
2	<i>the</i>	42	7,430	6,443
3	<i>to</i>	34	8,345	5,813
4	<i>of</i>	23	7,551	4,770
5	<i>and</i>	21	7,636	4,560
6	<i>for</i>	16	8,819	3,991
7	<i>million</i>	15	13,899	3,873
8	<i>measured</i>	14	16,692	3,742
9	<i>will</i>	12	10,214	3,461
10	<i>with</i>	12	8,821	3,456

Tabela 5: dez colocados principais de *ad spending* em ordem de frequência.

Podemos observar que dentre os colocados de *ad spending* somente 2 itens são palavras de conteúdo, sendo as demais palavras gramaticais, preposições, artigo, conjunção e verbo modal. Apresentamos a seguir as relações que pudemos observar entre o padrão *ad spending* e seus colocados.

1) Linhas de concordância com o nódulo *ad spending* e seu colocado *in*:

1 Disney XD grossed \$82 million in measured **ad spending in** 2009, down 1.1% from Toon
 2 e hours, and saw a 9.2% bump in measured **ad spending in** 2010, according to Kantar. Di
 3 .4 million). Fox News' overall measured **ad spending in** 2009 was down 0.7% from 2008,
 4 ser and business development in Shanghai. **Ad spending in** China's online video category
 5 akers' marketing dollars. Chrysler cut its TV **ad spending in** 2009 by half -- with the othe
 6 Associates projects that overall new-vehicle **ad spending in** 2010 will rise to \$19.2 billi
 7 companies concentrate the bulk of their own **ad spending in** the U.S. Eli Lilly & Co., Mer
 8 Either way, such factors will lead to record **ad-spending in** a number of states. What to w
 9 st cerebral in a way." Ford actually cut U.S. **ad spending in** 2009 as it consolidated its f
 10 planning more modest single-digit growth for **ad spending in** line with sales growth this y
 11 2009, and it grossed \$415 million in measured **ad spending in** 2009, according to Kantar,
 12 than the \$300 million-plus worth of issue-**ad spending in** 2009. The 2009 issue-ad spend
 13 sumer brands: only \$11 million on measured **ad spending in** 2009 and \$29 million in the f
 14 its rates and dwelling as a distant third in **ad spending in** the category. When rival Alls
 15 layers have also maintained or even increased **ad spending in** the last year, unlike others
 16 n issue that generates millions of dollars in **ad spending in** these competitive federal rac
 17 ssue that generates millions of dollars **in ad spending** in these competitive federal races. T
 18 bout two-thirds of the retailer's increase **in ad spending** for the fiscal third quarter and a su
 19 f Mexico, it unleashed \$100 million **in ad spending**, largely on network TV, to stem the d
 20 its rates and dwelling as a distant third **in ad spending** in the category. When rival Allstate
 21 ts for 2011 suggest moderate growth **in ad spending**, reflecting the economy's slow recove
 22 t. Fueling that growth -- and a jump **in ad spending** of more than 20% in the first 10 mont
 23 King over a proposed 25% increase **in ad spending**, to be funded through soft-drink reba
 24 ts Asia to show at least 5.9% growth **in ad spending** this year and in Latin America by 9.3
 25 ly, it is unlikely that we'll see any lulls **in ad spending**. Further, the nation is layered with

É possível observar que *in* com maior ocorrência após o nóculo (linhas 1 a 16) está comumente relacionado às palavras subsequentes, como ano e local. Já à esquerda (linhas 17 a 25), verificamos que se trata da preposição usada com o padrão para mencionar algo que ocorre com *ad spending*, como aumento (*growth* e *jump*), crescimento (*increase*) e período de calmaria (*lull*).

2) Linhas de concordância com o nóculo *ad spending* e o colocado *the*:

1 its rates and dwelling as a distant third in **ad spending** in **the** category. When rival Alls
 2 y, it is unlikely that we'll see any lulls in **ad spending**. Further, **the** nation is layered
 3 orecasts for 2011 suggest moderate growth in **ad spending**, reflecting **the** economy's slow
 4 ly without incumbents running for reelection. **Ad spending** on **the** state level will be drive
 5 ding. Overall, China represents 3.4% of total **ad spending** for **the** Global 100, slightly bel
 6 e clarity to one of the most opaque areas of **ad spending**, and **the** lifeblood of many Ameri
 7 t. And while P&G added \$1 billion back to its **ad spending** for **the** fiscal year ended June 3
 8 son. It also banked \$18.4 million in measured-**ad spending** during **the** first 11 months of
 9 bout two-thirds of the retailer's increase in **ad spending** for **the** fiscal third quarter and
 10 layers have also maintained or even increased **ad spending** in **the** last year, unlike others
 11 l companies concentrate the bulk of their own **ad spending** in **the** U.S. Eli Lilly & Co., Me

A maior ocorrência de *the* com o padrão está à direita do nóculo (R2) e vemos que não se relaciona diretamente a *ad spending*, mas é usado como um artigo determinante relacionado às palavras que aparecem após ele.

3) Linhas de concordância com o nóculo *ad spending* e o colocado *to*:

1 recent report that Colgate will need **to** boost **ad spending** by \$150 million to \$200 million
 2 Either way, such factors will lead **to** record **ad-spending** in a number of states. What to w
 3 nd personal-care companies he covers **to** hike **ad spending** as a share of sales by more than
 4 ral experts expect the 2010 census **to** boost **ad spending** toward minority groups as awaren

A maior ocorrência de *to* é na posição L2, ou seja, duas palavras à esquerda do nóculo, como vemos nas concordâncias 1 a 4. Observamos que não é um item diretamente relacionado a *ad spending*, mas sim aos verbos, expressando propósito.

4) Linhas de concordância com o nóculo *ad spending* e o colocado *of*:

1 o. 5 national advertiser in the U.S., with **ad spending of** \$2.3 billion, up 14.1% from 2002,
 2 o. 5 national advertiser in the U.S., with **ad spending of** \$2.3 billion, up 14.1% from 2002,
 3 ount for only 1% of 2010's total political **ad spending of** \$4.2 billion. It's not uncommon fo
 4 st. Fueling that growth -- and a jump in **ad spending of** more than 20% in the first 10 mont

O colocado *of* aparece tipicamente à direita do nóculo e notamos que é usado para referir-se a valores relacionados a *ad spending*.

5) Linhas de concordância com o nóculo *ad spending* e o colocado *and*:

1 a market in flux, shrinking revenue, weak **ad spending and** little upside offered by a tepid
 2 ix Could Look Like A Peek at the **Ad Spending and** Brand Power the Combination Would
 3 ity to one of the most opaque areas of **ad spending, and** the lifeblood of many American b
 4 ger, yes, but after so much sputtering of **ad spending and** cost cutting at media companies,
 5 ny profiles, agencies, brands, executives, **ad spending and** sales for the nation's 100 top ad

Conforme observamos, *and* apresenta função de conector e não podemos dizer que forma um padrão com *ad spending*, já que os itens que aparecem depois dele não se repetem.

6) Linhas de concordância com o nóculo *ad spending* e o colocado *for*:

1 ding. Overall, China represents 3.4% of total **ad spending for** the Global 100, slightly b
 2 s in marketing spending. Online **ad spending for** small and medium-sized U.S. businesses
 3 ll slice of its budget. McDonald's total U.S. **ad spending for** 2009 was \$873.3 million, c
 4 bout two-thirds of the retailer's increase in **ad spending for** the fiscal third quarter and
 5 Borrell estimates the total online political **ad spending for** 2010 will reach \$44.5 millio
 6 t. And while P&G added \$1 billion back to its **ad spending for** the fiscal year ended Ju
 7 ckelodeon grossed \$833.87 million in measured **ad spending for** full-year 2009, down
 8 move eliminated some \$300 million in regional **ad spending for** uniform messages, tho

É possível observar que *for*, com maior ocorrência à direita do nóculo (R1), relaciona-se às palavras que o seguem, referindo-se a empresa/negócios (linhas 1 e 2), ano/período (linhas 3, 4, 5, 6 e 7) e propósito (linha 8).

7) Linhas de concordância com o nóculo *ad spending* e os colocados *million* e *measured*:

1 consumer brands: only \$11 **million** on **measured ad spending** in 2009 and \$29 million in t
 2 ckelodeon grossed \$833.87 **million** in **measured ad spending** for full-year 2009, down 10.
 3 ing to Ad Age's analysis. First-half **measured ad spending** rose 5.7%, with gains in every
 4 ABMiller.) For the Global 100, total **measured ad spending** tumbled 8.7% to \$107.2 billi
 5 tor to see a drop in first-half 2010 **measured ad spending**, according to WPP's Kantar Me
 6 son. It also banked \$18.4 **million** in **measured-ad spending** during the first 11 months of 2
 7 take: Cartoon banked \$230 **million** in **measured ad spending** during 2009, down 1% from
 8 U.S. For the Global 100 group, total **measured ad spending** rose 3.1% to \$117.9 billion in
 9 2009, and it grossed \$415 **million** in **measured ad spending** in 2009, according to Kantar,
 10 but Disney XD grossed \$82 **million** in **measured ad spending** in 2009, down 1.1% from T
 11 time hours, and saw a 9.2% bump in **measured ad spending** in 2010, according to Kantar.
 12 8, \$6.4 million). Fox News' overall **measured ad spending** in 2009 was down 0.7% from 2

Podemos observar pelas concordâncias e pelos textos que se trata de um padrão referente a investimentos, gastos e dinheiro relativos ao gasto/custo de anúncios padrão. Encontramos na observação dessas linhas os verbos usados com o padrão *gross* e *bank* à esquerda, usado para referir-se à quantidade investida ou gasta e *rise* e *tumble* à direita, usados quando se refere à porcentagem de aumento e queda dos anúncios.

8) Linhas de concordância com o nóculo *ad spending* e o colocado *will*:

1 hat, contrary to the pro-digital pundits, TV **ad spending will** not be shifting dramaticall
 2 y 7.8%. The crucial question is when global **ad spending will** return to the 5% or 6% grow
 3 from Ad Age DataCenter reports Worldwide **ad spending will** grow 5.3% in the new year,

O colocado *will* aparece tipicamente à direita do nóculo, que nesse caso assume a função de sujeito; como observamos nas concordâncias, esse uso indica as mudanças relativas a *ad spending* no futuro.

9) Linhas de concordância com o nóculo *ad spending* e o colocado *with*:

1 NCHANG, China (AdAge.com) -- **With ad spending** growth still in the double digits, ad
 2 o. 5 national advertiser in the U.S., **with ad spending** of \$2.3 billion, up 14.1% from 2002,
 3 n't provide much incentive to stick **with ad-spending** plans. But big marketers like Procter

Novamente encontramos um colocado que não está diretamente relacionado a *ad spending*; a preposição *with* em sua posição mais frequente, à esquerda do nóculo, diz respeito a causa (linhas 1 e 2) e junto a (linha 3).

Com essa análise, podemos concluir que, dentre os dez colocados mais frequentes, os que têm relação direta com *ad spending* e podem ser, portanto, relevantes para serem ensinados são: *in ad spending*, *ad spending of*, *million in ad spending* e *measured ad spending*. Esses padrões provam ser bastante úteis em discussões sobre aspectos financeiros de uma campanha publicitária.

2.3.1.2 Análise de *ad network(s)*

Seguimos os mesmos procedimentos mencionados na análise anterior para observar o uso de *ad network(s)* e podemos verificar nas Tabelas 6 e 7 os colocados encontrados.

	COLOCADO	FREQUÊNCIA	IM	ESCORE T
1	<i>the</i>	24	6,623	4,849
2	<i>and</i>	19	7,492	4,335
3	<i>an</i>	11	9,657	3,313
4	<i>of</i>	10	6,349	3,123
5	<i>on</i>	9	8,257	2,990
6	<i>to</i>	8	6,428	2,965
7	<i>mobile</i>	6	16,860	2,828
8	<i>a</i>	6	6,096	2,414
9	<i>its</i>	5	9,860	2,447
10	<i>is</i>	4	6,996	2,219

Tabela 6: colocados principais de *ad network*.

	COLOCADO	FREQUÊNCIA	IM	ESCORE T
1	<i>and</i>	40	8,556	6,308
2	<i>the</i>	24	6,623	4,849
3	<i>to</i>	22	7,714	4,668
4	<i>of</i>	16	7,028	3,969
5	<i>on</i>	12	8,673	3,456
6	<i>are</i>	12	9,347	3,459
7	<i>a</i>	9	6,681	2,971
8	<i>that</i>	9	7,733	2,986
9	<i>publishers</i>	8	16,998	2,828
10	<i>from</i>	8	8,867	2,822

Tabela 7: colocados principais de *ad networks*.

Selecionamos para análise os itens em comum entre a forma singular e plural, que são: *the*, *and*, *to*, *of*, *on*, *a* e o verbo *be*, sendo *is* e *are*. Dentre as palavras de conteúdo, encontramos *mobile* e *publishers*, as quais adicionamos à análise. Desconsideramos *a* por não estar relacionado diretamente ao nóculo *ad network(s)*.

- 1) Linhas de concordância com o nóculo *ad network(s)* e o colocado *the*:

1 ion in stock on the leader in **the** mobile **ad-network** space, AdMob, and got its claws in ret
 2 bile, 4INFO could be in play. **The SMS ad network** has the reach, technology platform, us
 3 " said Alec Andronikov, head of Movox, **the SMS ad network** behind the campaign.
 4 one myself; my blog is on **the** BlogHer **ad network**), none of these networks is ideal unle
 5 enry Copeland, founder of **the** blog **ad network** Blogads.com. (Claim to fame? They've g
 6 ion in stock on the leader in **the** mobile **ad-network** space, AdMob, and got its claws in ret

Assim como no padrão analisado anteriormente, *the* aparece como determinante de *ad network*. Encontramos, no entanto, em todas as concordâncias, itens entre *the* e o nóculo que identificam tipo ou nome, como *mobile*, *SMS*, *BlogHer* e *blog*. Analisando *ad networks*, a posição mais típica foi três itens à direita (R3) e nesses casos *the* não apresenta relação direta com o nóculo, como podemos ver a seguir:

1 some of their peers, both publisher and **ad networks**, who feared **the** buy -- probably the b
 2 sold directly or optimize any number of **ad networks** vying for the impression. That's a lo
 3 e advertising prices, but Mr. Morgan said **ad networks** unfairly took **the** fall: The real prob
 4 because it is. It's the same formula that **ad networks** promised on **the** web -- take low-cost
 5 n for Google (and has given birth to 400 **ad networks**), and represents **the** best thinking of

2) Linhas de concordância com o nóculo *ad network(s)* e o colocado *and*:

1 the site. Evernote recently dumped its **ad network and** started curating its own ads, maki
 2 ly scant reference was made to AOL's **ad network and** technologies -- the onetime Platfo
 3 ERE's desktop site, mobile apps and **ad network, and** PayPal) looks to be moving toward
 4 the situation has brought for both the **ad network and** the global conservation network, o
 5 ose publishers' core properties, but from **ad network and** long-tail content partners as well
 6 CBS and ABC to the portals and video **ad networks and** is planning to spend significantl
 7 S Today, advertisers use a myriad of **ad networks and** exchanges to buy demographics or
 8 y channels: many publishers, many **ad networks and** now several ad exchanges. To make
 9 uarter, monetizing almost entirely via **ad networks and** Google AdSense. Are the kittes t
 10 ker or CBS, to stop doing business with **ad networks and** to focus on brand dollars. It's a
 11 igital-media-buying platforms that usurp **ad networks and** portals. It's a big pool and ever

12 rush-to-the-bottom pricing driven **by ad networks and** exchanges, online display has had
 13 for Google (and has given birth to 400 **ad networks**), **and** represents the best thinking of
 14 encies and vendors, but a survey of **ad networks and** publishers reveals unanimity: CPG
 15 e (QPS) being purchased by DSPs, **ad networks and** agency trading desks on a RTB bas
 16 ising model is under pressure from **ad networks and** Facebook. Ad networks have allowe
 17 ales force. Our ads are sold through **ad networks and** rep firms, MTV Networks, Gorilla
 18 y is coming mainly from two sources: **ad networks and** Google's AdSense. But in the thir
 19 e forums subsist on revenues from **ad networks and** Google AdSense, yielding CPMs oft
 20 at complicates "impressions" is that **ad networks and** demand side platforms (DSPs) that

O colocado *and* funciona como um conector e, na sua posição mais frequente, à direita do nóculo, o que podemos observar é a relação entre ad network(s) com redes, agentes, organizações e sistemas, como *technologies, PayPal, global conservation network, exchanges, publishers, agency trading desks, Facebook, rep firms, Google's AdSense e demand side platforms*.

3) Linhas de concordância com o nóculo *ad network(s)* e os colocados *to e of*:

1 s that while one party is in control **of an ad network**, an exchange operates on behalf of all
 2 roduct marketing manager **of mobile ad network** AdMob. "Layering device targeting on t
 3 sold directly or optimize any number **of ad networks** vying for the impression. That's a lo
 4 ES Today, advertisers use a myriad **of ad networks** and exchanges to buy demographics o
 5 agencies and vendors, but a survey **of ad networks** and publishers reveals unanimity: CPG
 6 e last decade, with literally hundreds **of ad networks**, data companies, yield managers, ad
 7 -channel mix The traditional rules **of ad networks** -- aggregate massive amounts of impre
 8 sands of websites (and even hundreds **of ad networks**) -- do you want to spend all day on t
 9 hanges were largely invented **to enable ad networks** to more efficiently and cheaply trade
 10 mass will enable CrowdGather **to bypass ad networks** altogether. Fair enough, but perhaps
 11 for Google (and has given birth **to 400 ad networks**), and represents the best thinking of
 12 n iPhone users, according **to mobile ad network** AdMob, and 10% of the ad requests on A
 13 nly scant reference was made **to AOL's ad network** and technologies -- the onetime Platfo
 14 ressions were passed back forward **to an ad network**, then passed back if unsold. So while
 15 ness from premium inventory **to its ad network**, exchange and newspaper consortium. FA

Nas concordâncias, observamos que as preposições *to* e *of* não estão relacionadas diretamente a *ad network(s)*, mas sim às palavras ao redor do nóculo. No caso de *of*, encontramos, por exemplo, *in control of*, *number of*, *myriad of* e *rules of*. Nesses casos, vemos que os substantivos referem-se ao gerenciamento ou controle das redes. Com *to*, temos *give birth to*, *according to*, *was made to*, e *passed to*. Podemos ainda observar os verbos usados com *ad network(s)* nas linhas com *to* como colocado, como *enable*, *bypass*, *be born (give birth)*.

4) Linhas de concordância com o nóculo *ad network(s)* e o colocado *on*:

1 ly to keyword and display ads **on** Google's **ad network**, a download button on YouTube an
 2 to remnant display-ad inventory sold **on** an **ad network**. Think news video, casual gaming
 3 e off 10% of our TV budget and run **on ad networks!**" said Chris Allen, VP-video innovat

Podemos dizer que *on* é a preposição frequente usada antes de *ad network(s)*, como podemos ver nas concordâncias. Esse colocado, juntamente com a análise dos usos anteriores, reforça a ideia da *ad network* como de uma rede, como se fosse um local, na qual as empresas negociam e apresentam seus anúncios e campanhas publicitárias, como vemos com os verbos *display*, *sold* e *run*.

5) Linhas de concordância com o nóculo *ad network(s)* e os colocados *is* e *are*:

1 mmodate blogs of all niches and sizes. Until then, **ad networks are** the best option.
 2 pretty impressive." "It's a reflection that **ad networks are** beginning to rapidly shift online
 3 . Many of these vendors, publishers and **ad networks are** working together on initiatives t
 4 ed in the space -- and more-established **ad networks are** starting to offer dynamic creativ
 5 some niche retailers simply don't have. Ad **networks are** built on the promise of scalabili

O verbo *are* liga *ad network(s)* a um complemento ou funciona como um auxiliar; não há recorrência entre os verbos que os sucedem, mas podemos observar o uso desses verbos com o padrão, determinando as ações realizadas pelas *ad networks*, como *offer*, *shift* e *work together*. Quanto a *is*, não encontramos relação alguma com o nóculo, visto que esse item ocorre em posições afastadas dele.

6) Linhas de concordância com o nóculo *ad network* e o colocado *mobile*:

1 evelopers: More than 120,000 **Mobile Ad Network**: Quattro Wireless Price tag: \$275 mill
 2 g to a Google spokeswoman. **Mobile Ad Network**: AdMob Price tag: \$750 million in stoc
 3 ertising business by acquiring **mobile-ad network** Quattro Wireless. Though the deal was
 4 on in stock on the leader in the **mobile ad-network** space, AdMob, and got its claws in ret
 5 roduct marketing manager of **mobile ad network** AdMob. "Layering device targeting on t
 6 n in stock on the leader in the **mobile ad-network** space, AdMob, and got its claws in ret
 7 than iPhone users, according to **mobile ad network** AdMob, and 10% of the ad requests

Podemos observar a frequência de *mobile* à esquerda do nóculo, uma associação interessante, pois indica o tipo de *ad network*, relacionado a celulares. Vemos que *mobile* está relacionado tanto a *ad network*, quanto à palavra que o segue, como nas linhas 4, 5, 6 e 7, nas quais *mobile* está relacionado a *space* e *AdMob*, uma empresa de propaganda focada em anúncios em celulares. Assim, forma-se os padrões *mobile ad network + space* e *mobile ad network + nome da empresa*.

7) Linhas de concordância com o nóculo *ad networks* e o colocado *publishers*:

1 ny of these vendors, **publishers** and **ad networks** are working together on initiatives t
 2 y channels: many **publishers**, many **ad networks** and now several ad exchanges. To make
 3 ve negotiations between **publishers** and **ad networks**. Networks' sales forces were still du

Essa concordância nos leva a observar a relação de trabalho e/ou negócios entre *publishers* e *ad networks*, mas não há um padrão propriamente dito com esse colocado e o nóculo.

Considerando a análise de *ad network(s)*, concluímos que podemos selecionar os seguintes padrões relevantes para o ensino: *the + substantivo + ad network*, *on + ad network*, *mobile + ad network*. Observamos também por meio dos colocados *and* e *are* outros itens relacionados a *ad network* que poderiam ser explorados no ensino.

2.3.1.3 Análise de *ad revenue*

Com relação a *ad revenue*, não encontramos palavras de conteúdo como colocados. Vejamos na Tabela a seguir os itens que tipicamente ocorrem com *ad revenue* e em seguida a análise:

	COLOCADO	FREQUÊNCIA	IM	ESCORE T
1	<i>the</i>	31	6,992	5,524
2	<i>of</i>	17	7,115	4,093
3	<i>to</i>	17	7,345	4,098
4	<i>in</i>	17	7,763	4,104
5	<i>and</i>	15	7,151	3,846
6	<i>from</i>	11	9,327	3,311
7	<i>is</i>	9	7,844	2,987
8	<i>that</i>	7	7,370	2,630
9	<i>a</i>	7	6,318	2,613
10	<i>but</i>	7	8,606	2,639

Tabela 8: colocados principais de *ad revenue*.

1) Linhas de concordância com o nóculo *ad revenue* e o colocado *the*:

1 onated the block of time and agreed **the ad revenue** would be donated to charity. All told,
 2 f the ad dollars spent online. Not that **the ad revenue** is anything to sneeze at. Nielsen just
 3 , gain more viewer attention -- and **the ad revenue** that comes with it. Fox News captured
 4 n. What gives? Mr. Armstrong said **the ad-revenue** drop was largely "self-inflicted," mea

Verificamos que *the* ocorre com maior frequência à esquerda do nóculo e como um determinante, marcando a existência do rendimento proveniente da campanha; na linha 4, observamos que pode ser usado na função de adjetivo, com o hífen, referente ao substantivo *drop*.

2) Linhas de concordância com o nóculo *ad revenue* e o colocado *of*:

1 h would put 12% of page views and 9% **of ad revenue** at risk, but Mr. Brill said he doesn't
 2 xecutives believe TV is the sole driver **of ad revenue**, that could leave rivals eager to sell

3 s susceptible to the cyclical swings **of ad revenue**. Consider: News Corp. said revenue for
4 applications; soak up new categories **of ad revenue**, including obituaries; and pursue new

As concordâncias nos mostram que *of* relaciona *ad revenue* a propriedades como oscilações (*swings*) e categorias (*categories*), nas linhas 3 e 4.

3) Linhas de concordância com o nóculo *ad revenue* e o colocado *from*:

1 r to boost their own followings and **ad revenue from** YouTube. YouTube pro Jill Hanner,
2 d similar companies get most of their **ad revenue from** commercials that appear alongside
3 nd are anxious to keep \$60.5 billion in **ad revenue from** going the way of Philco. A host o
4 r regular videos, which means more **ad revenue from** YouTube, which is placing ads acr

As concordâncias revelam que *from* é uma preposição tipicamente usada com *ad revenue* indicando origem do rendimento referente a Publicidade.

4) Linhas de concordância com o nóculo *ad revenue* e os colocados *is* e *that*:

1 f the ad dollars spent online. Not that the **ad revenue is** anything to sneeze at. Nielsen just
2 esses. Third-quarter domestic-display **ad revenue is** down 8% year over year, but up from
3 half what the daily required. Online **ad revenue is** down some 20%, which the Monitor at
4 2010. Last year's take: Disney Channel **ad revenue is** not tracked by Kantar Media, but Di
5 d. BAD NEWS U.S. newspapers' digital **ad revenue is** not keeping pace with digital adver
6 the bill. Media that relies primarily on **ad revenue is** suffering far more than those compa
7 es believe TV is the sole driver of **ad revenue, that** could leave rivals eager to sell
8 re viewer attention -- and the **ad revenue that** comes with it. Fox News captured

As ocorrências do nóculo com *is* e *that* nos permitem expandir as ideias relacionadas aos rendimentos de uma campanha, ligando-o a preposições *down* e a verbos como *sneeze at*, *track*, *keep pace with*, *suffer* e *come with*.

5) Linhas de concordância com o nóculo *ad revenue* e o colocado *in*:

1 All told, the program generated \$6 million **in ad revenue** with the biggest chunk -- some \$
 2 r ended June 2010, with a modest 3% rise **in ad revenue** but a whopping 19% jump in fees
 3 t. The company earned close to \$1 billion **in ad revenue** in 2008, compared to an estimated
 4 he network has taken in about \$3.1 billion **in ad revenue** during the first seven months of
 5 Bravo. Add that to the \$585.7 million **in ad revenue** logged by Comcast Entertainment's
 6 Blip.tv's top earners pull down \$500,000 **in ad revenue** a year based on 50/50 revenue sha
 7 dia to finish 2010 with an overall decline **in ad revenue**. "I don't know that the positive
 8 That sector has outperformed network TV **in ad-revenue** growth. Cable networks also get t
 9 lity and are anxious to keep \$60.5 billion **in ad revenue** from going the way of Philco. A h
 10 entureBeat, was making \$25,000 a month **in ad revenue**. While there was no media spendi

O colocado ‘*in*’ ocorre na maioria das vezes na posição L1, ou seja, diretamente à esquerda de ‘*ad revenue*’. Observamos nas concordâncias que, como ‘*ad revenue*’ refere-se a rendimentos ou receita, as palavras ao redor estão relacionadas a valores em dinheiro, como vemos em *million*, *billion* e *\$*. Assim, podemos encontrar ‘*rise + in + ad revenue*’, ‘*earn / take in / pull down / decline / keep / make + amount of money + ad revenue*’.

Os colocados *to*, *and*, *a* e *but* não ocorreram próximos, nem relacionados ao nóculo. Dessa forma, os padrões que pudemos verificar com *ad revenue* são: *of ad revenue*, *in ad revenue* e *ad revenue from*.

2.3.1.4 Análise de *ad campaign(s)*

Com o padrão *ad campaign(s)*, encontramos uma coocorrência de preposições, como se pode ver nas Tabelas 9 e 10.

	COLOCADO	FREQUÊNCIA	IM	ESCORE T
1	<i>the</i>	27	6,793	5,149
2	<i>an</i>	22	10,657	4,688
3	<i>a</i>	20	7,833	4,453
4	<i>of</i>	10	6,349	3,123
5	<i>new</i>	10	10,965	3,161
6	<i>with</i>	10	8,558	3,154
7	<i>to</i>	10	6,580	3,129
8	<i>from</i>	9	9,037	2,994
9	<i>and</i>	8	6,244	2,791
10	<i>that</i>	8	7,563	2,813
11	<i>in</i>	7	6,483	2,616
12	<i>for</i>	6	7,404	2,435
13	<i>by</i>	5	7,920	2,227

Tabela 9: colocados principais de *ad campaign*.

	COLOCADO	FREQUÊNCIA	IM	ESCORE T
1	<i>and</i>	17	7,332	4,098
2	<i>of</i>	12	6,612	3,429
3	<i>the</i>	11	5,497	3,243
4	<i>for</i>	7	7,627	2,632
5	<i>a</i>	5	5,833	2,197

Tabela 10: colocados principais de *ad campaigns*.

- 1) Linhas de concordância com o nóculo *ad campaign(s)* e os colocados *to, for, from e by*:

1 a sweet deal in which LBL will create a new **ad campaign to** stimulate sales at 50% off th
2 nized the need for an organized, strategic **ad campaign to** drive the ongoing demobilizat
3 ng, pricing strategies, retail promotions and **ad campaigns to** reach consumers effectively
4 the recession, earning itself a new national **ad campaign from** Publicis' Saatchi & Saatchi
5 f course, to the software giant's ongoing **ad campaign from** agency Crispin Porter & Bog
6 mployees. Chrysler breaks the "New Day" **ad campaign from** BBDO. The spots use third
7 Fox somewhat, but he largely credits an **ad campaign from** Publicis Worldwide, New Y
8 y a just-launched umbrella marketing and **ad campaign from** WPP's Y&R that includes s
9 mployees. Chrysler breaks the "New Day" **ad campaign from** BBDO. The spots use third

10 rm details about the upcoming national **advertising campaign for** Via, other than the offi
 11 t's not to say there aren't marketing and **advertising campaigns for** the Beatles' selected p
 12 n the course of 2010, Kraft has launched new **ad campaigns for** 15 of its "power brands,"
 13 new standard of the world," and forthcoming **ad campaigns for** the brand will center on a
 14 is year." The cross-platform **ad campaign by** the Atlanta shop Lawler Ballard Va
 15 better known as FARC. An **ad campaign by** Bogota-based Lowe SSP3 is helping
 16 is being promoted with an **ad campaign by** WPP's Y&R with a "Heaven on Earth
 17 00 by the end of the year. An **ad campaign by** Sancho BBDO highlights that the s

Novamente verificamos palavras gramaticais como colocados. Analisamos as linhas de concordância e vimos que o uso de *ad campaign(s)* com as preposições, *to*, *for* e *from* à direita mostrou-se interessante. O primeiro bloco de linhas de concordância nos mostra o uso de adjetivo + *ad campaign* + *to* + verbo. Esse uso refere-se ao propósito da campanha publicitária. Em segundo lugar, verificamos o padrão adjetivo/nome da campanha + *ad campaign* + *from/by* + nome da agência; vemos aqui que esse padrão é usado em geral para referir-se a uma certa campanha e que agência a criou ou desenvolveu. Por último, encontramos *ad campaign* + *for* + marca /empresa, relacionando a marca ou empresa para a qual a campanha foi criada ou desenvolvida.

2) Linhas de concordância com o nódulo *ad campaign(s)* e o colocado *new*:

1 t's because they're the stars of a **new ad campaign** that broke during Fox's NFL broadcast
 2 thed" after Mt. Lutz criticized a **new ad campaign** that she defended, said Chris Hamer,
 3 et deal in which LBL will create a **new ad campaign** to stimulate sales at 50% off their n
 4 2010, Kraft has launched **new ad campaigns** for 15 of its "power brands," includ

Como podemos observar nas concordâncias, o adjetivo *new* é um colocado típico à esquerda de *ad campaign(s)*, provavelmente pela natureza dos negócios feitos nas agências, que é a criação de campanhas.

3) Linhas de concordância com o nódulo *ad campaign(s)* e o colocado *that*:

1 cathed" after Mt. Lutz criticized a new **ad campaign that** she defended, said Chris Hamer,
 2 lasphemy that takes it too far. And an **ad campaign that** uses rapper Eminem to personify
 3 t's because they're the stars of a new **ad campaign that** broke during Fox's NFL broadcast
 4 10, it's "Write the Future" for Nike, an **ad campaign that** leveraged Nike's biggest stars i
 5 lasphemy that takes it too far. And an **ad campaign that** uses rapper Eminem to personify

Essas concordâncias nos permitem observar como expandir as ideias relacionadas às campanhas e quais os verbos comumente usados para isso, como *use*, *break* e *leverage*.

4) Linhas de concordância com o nóculo *ad campaign(s)* e o colocado *in*:

1 ordered for possible use **in** a national **ad campaign**. Winners received \$500. That every ma
 2 talk about what it's like to have a role **in** an **ad campaign** for the first time. While it's not a

Conforme podemos observar nas concordâncias, *in* é uma preposição que coocorre com *ad campaign*.

Os padrões que pudemos encontrar com o nóculo estudado são: *ad campaign* + *to* / *for* / *from* / *by*, *new* + *ad campaign*, *in* + *a/an* + (adjetivo) + *ad campaign* e *ad campaign* + *that*.

2.3.2 Análise dos verbos colocados de *ad*

Por fim, achamos relevante analisar os verbos colocados de *ad*. Eles não apareceram entre os 25 mais frequentes (Ver Tabela 2), entretanto entendemos que seria uma análise pertinente por *ad* ser um substantivo, além do fato de que, em nossa experiência, os alunos têm dificuldades no uso desse padrão. Na Tabela 11 podemos verificar os verbos relacionados diretamente ao nóculo. Os verbos *did* e *put* apresentaram valor de Informação Mútua abaixo de 3; porém, decidimos mantê-los na análise por estarem próximos a 3 e pelo valor de Escore T ser superior a 2.

	COLOCADO	FREQUÊNCIA	IM	ESCORE T
1	<i>do</i>	40	2,868	5,458
2	<i>run</i>	24	5,167	4,763
3	<i>launched</i>	20	7,179	4,441
4	<i>did</i>	17	2,577	3,432
5	<i>ran</i>	16	5,976	3,936
6	<i>running</i>	15	5,156	3,764
7	<i>put</i>	14	2,973	3,265
8	<i>placed</i>	5	4,274	2,120

Tabela 11: principais verbos colocados de *ad*.

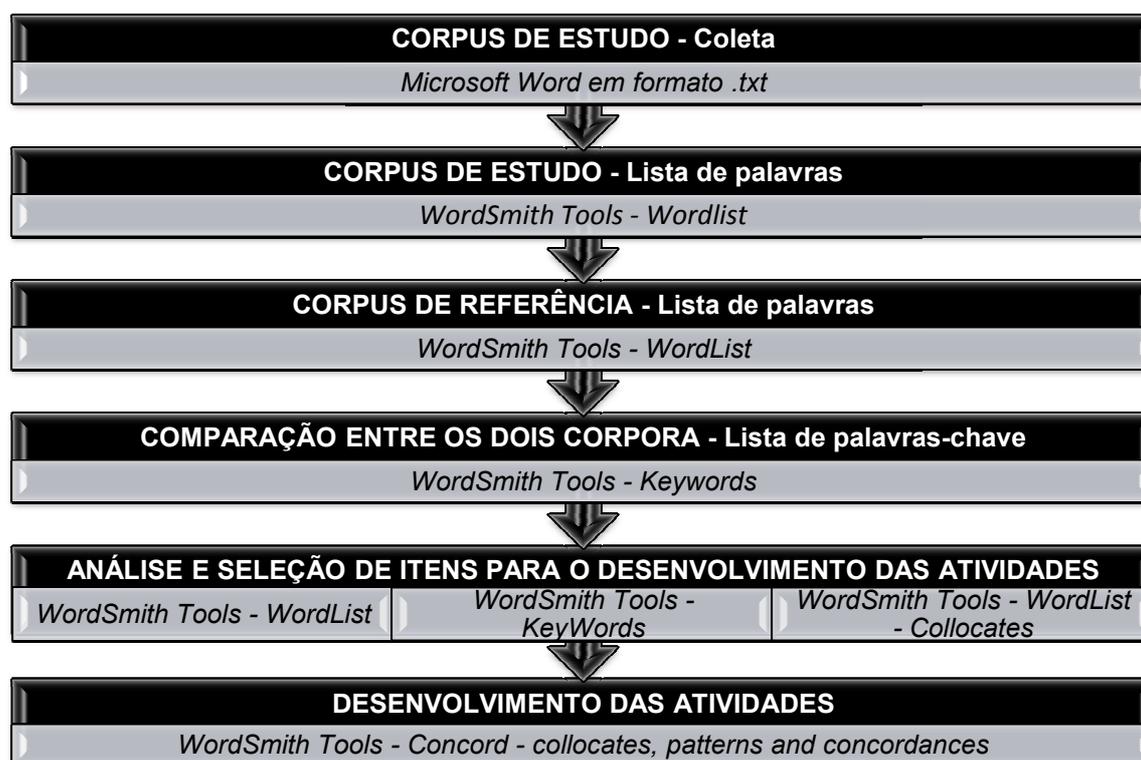
Vejamos a seguir algumas linhas de concordância com os verbos selecionados.

1 Sullivan noted: "When you sit down to **do** an **ad**, you're competing with every brand out th
2 d the agency and warned them not to **run** the **ad**, on the grounds of racism, imperialism an
3 gan "Future prime minister." It's also **run** an **ad** on the massively popular music service Sp
4 l reign supreme, the models for how to **run** an **ad**, and drive emotion, and get a lift in the
5 e company had no guarantee ABC would **run** the **ad** the way it did, despite multiple entrea
6 ive, Groupon has since **launched** a replacement **ad** that's running in prime time. It's also w
7 rizon. The carrier recently **launched** a teaser **ad** attacking the iPhone for all the things i
8 In 2001, for instance, the two **launched** a TV **ad** starring musical luminaries ranging from
9 ver accept such a 'flashing' image. We **did** an **ad** at the time for After Six Formals with a
10 you can believe BBDO, Istanbul, just **did** an **ad** so much like Y&R, Detroit, made in 2002.
11 t of "Gossip Girl" this season, the CW **ran** an **ad** that talked about "evil" characters not o
12 City Department of Health **running** a major **ad** push asking consumers if they're "pouring
13 ay special. JCPenney began **running** a holiday **ad** promoting its relationship with the Salva

Finalizamos este capítulo em que fizemos a análise dos padrões formados por *ad* e seus colocados. Neste trabalho não apresentamos a análise das outras palavras-chave encontradas por questões de espaço, apesar de terem sido feitas e nossos estudos. Acreditamos que essa análise tenha mostrado nossa trajetória para a seleção de alguns itens que usamos no desenvolvimento de nossas atividades, assunto do qual trataremos no capítulo seguinte.

CAPÍTULO 3: DESENVOLVIMENTO DAS ATIVIDADES

Antes de abordarmos a questão das atividades, apresentamos no esquema a seguir os procedimentos que envolveram a pesquisa.



Esquema 1: procedimentos e ferramentas usados na pesquisa.

Neste Capítulo, explicaremos os procedimentos de transposição dos achados da pesquisa para as atividades. Para tanto, baseamo-nos nas propostas apresentadas na seção 1.2.2.1, e especial nos seguintes autores: Willis (1990), Johns (1991), Berber Sardinha (2004; 2009; 2011), Bernardini (2004) e Willis e Willis (2007). Esses autores discutem critérios para selecionar padrões de *corpora* e aplicá-los na preparação de material didático; tais critérios envolvem a pertinência dos padrões para o contexto de ensino, em geral, e para as atividades em questão, em particular.

Baseado nas propostas dos autores mencionados, elencamos os seguintes pressupostos para o desenvolvimento das tarefas:

- 1) Uso de textos autênticos, não criados com um propósito pedagógico, mas sim para uma comunicação real.
- 2) Foco no significado e na forma.

- 3) Criação de oportunidades para os alunos usarem a língua para comunicação real.
- 4) Uso de linhas de concordância para a exploração da língua em uso.
- 5) Foco na lexicogramática, nos padrões lexicogramaticais e nas colocações;
- 6) Desenvolvimento das habilidades de exploração e descoberta dos alunos.
- 7) Papel do professor de observador, mediador e facilitador, guiando os alunos para as descobertas.

Sugerimos três atividades para o uso de textos autênticos e de resultados da análise de *corpora* em sala de aula. Procuramos desenvolver atividades em que os alunos pudessem lidar com o texto e com o *corpus*, pois acreditamos que “texto e *corpus* são duas unidades de língua que se completam naturalmente, até porque a segunda é uma coletânea da primeira” (BERBER SARDINHA, 2009). Conforme explica o autor, um *corpus* pode ser interessante, porém não pode ser lido de forma linear como um texto, mas sim convida a uma leitura hipertextual; um *corpus* explicita a intertextualidade, enquanto o texto a supõe. A leitura e discussão dos textos da área são importantes para que os alunos possam comunicar-se sobre assuntos de sua realidade, o que, em nossa experiência, mostrou-se ser um fator de motivação. Verificamos também que a observação e o questionamento sobre o uso da língua parte dos próprios alunos por meio dos textos. Em diversas situações os alunos solicitaram o estudo de algo que leram no texto.

Acreditamos que com a observação de linhas de concordância os alunos podem

ver muitos exemplos de uma palavra ou expressão, em seu contexto. Tem-se uma ideia melhor do uso de uma palavra observando muitos exemplos dela e é vendo ou ouvindo novas palavras em contexto muitas vezes que você pode compreender o significado da maioria das palavras em sua língua nativa. Um dicionário pode dar o significado, mas não é muito bom em mostrar como usar uma palavra⁷⁶ (SCOTT, 2008⁷⁷)

Procuramos desenvolver atividades com ferramentas e recursos variados, de forma que o professor pode preparar os exemplos previamente, se não tiver os recursos em sala de aula, assim como usá-los juntamente com os alunos, se houver a

⁷⁶ “see lots of examples of a word or phrase, in their contexts. You get a much better idea of the use of a word by seeing lots of examples of it, and it's by seeing or hearing new words in context lots of times that you come to grasp the meaning of most of the words in your native language. It's by seeing the contexts that you get a better idea about how to use the new word yourself. A dictionary can tell you the meanings but it's not much good at showing you how to use the word.” (SCOTT, 2008)

⁷⁷ Explicação sobre os propósitos de uma concordância do manual do programa, versão 5.0.

possibilidade. O tempo das atividades varia de acordo a disponibilidade de um computador em sala de aula ou não, com o número de alunos e seu rendimento, entre outros fatores. A princípio, no entanto, calculamos que cada uma possa ser realizada em 2 aulas de 1 hora e trinta minutos cada.

Cabe ressaltar que todas as atividades podem ser usadas em aulas individuais ou em grupos. No caso de aulas individuais, os alunos podem refletir sozinhos e reportar suas ideias ao professor ou então discutir as questões juntamente com o professor, que terá um papel de participante da discussão, e depois reportar a discussão por escrito, em forma de redação, por exemplo. Além disso, sugerimos em algumas atividades que o aluno busque mais exemplos em linhas de concordância do item que está estudando. Isso pode ser feito pelo professor antecipadamente, de forma impressa, ou, se houver a possibilidade, com o uso de um ou mais computadores em sala de aula.

3.1 Contexto

Como já mencionamos, essa pesquisa teve como ponto de partida nossa experiência como professora de inglês como língua estrangeira para profissionais que atuam na área de Publicidade. As aulas são geralmente ministradas no local de trabalho do aluno, em salas de reunião, e os grupos variam de um a quatro alunos. É possível utilizar computadores ou não, dependendo da necessidade e conforme a aula preparada pelo professor. Apresentamos nossas atividades neste trabalho em *slides* no programa *PowerPoint*, da suíte *Microsoft Office*, visto que é o meio que utilizamos com os alunos.

O foco principal dos alunos é a comunicação, especialmente no âmbito profissional, e em geral não se utiliza materiais didáticos. Quando há o uso desses materiais, atividades extras são ministradas para complementá-los com o objetivo de proporcionar maior contato com a linguagem da área. Portanto, as atividades foram elaboradas para alunos de níveis intermediário e/ou superior que utilizam a língua em suas atividades profissionais e que possuem certo conhecimento tanto de alguns termos, como do mercado de trabalho na área de Publicidade.

3.2 Atividade 1

A primeira atividade teve como ponto de partida um vídeo que pode ser selecionado pelo professor ou pelos alunos. Essa atividade está dividida em três partes, embora essa divisão não seja rígida; trata-se, portanto de uma atividade aberta, que tem seus objetivos, porém pode ser modificada ao longo de sua execução, de acordo com as sugestões e necessidades dos alunos. O objetivo é a discussão do tema proposto e a exploração da linguagem do texto, que pode ser escolhida pelo professor ou sugerida pelos alunos. Neste caso, o tema é *Green Marketing*, conhecido como Marketing ambiental ou Marketing verde, que é a estratégia de marketing voltada ao processo de venda de produtos e serviços baseados em seus benefícios ao meio ambiente.

Na primeira parte, a princípio, são sugeridas questões sobre o tema do vídeo para que os alunos possam familiarizar-se com os tópicos; chamamos esta parte de *Discussing the topic* (Ver Figuras 8 e 9). É dado um trecho com a explicação sobre *Green Marketing* retirado de uma enciclopédia da *Internet*, cujo objetivo é informá-los sobre o assunto e instigá-los a observar sua origem a partir da linguagem utilizada e do layout do texto. Os alunos também discutem outro tópico apresentado no vídeo: *media vehicles*. Nesse momento, é interessante que o professor verifique o quanto os alunos sabem sobre o assunto e quanto à linguagem relacionada a ele. Em seguida, na parte que intitulamos *Getting to know about it* há uma atividade para os alunos preverem o que verão no vídeo; os alunos então assistem ao vídeo e discutem as questões que ele apresenta; essas questões podem partir dos alunos ou podem ser dadas sugestões, como em nossa atividade (Ver Figuras 10 e 11).

GREEN MARKETING

DISCUSSING THE TOPIC: GREEN MARKETING

Take a look at the excerpt below and discuss:

- *Where is it from? How do you know that?*
- *Discuss the topic Green Marketing and what you know about agencies/companies that use it.*

According to the [American Marketing Association](#), **green marketing** is the marketing of products that are presumed to be environmentally safe.^[1] Thus green marketing incorporates a broad range of activities, including product modification, changes to the production process, packaging changes, as well as modifying advertising. Yet defining green marketing is not a simple task where several meanings intersect and contradict each other; an example of this will be the existence of varying social, environmental and retail definitions attached to this term.^[1] Other similar terms used are **Environmental Marketing** and **Ecological Marketing**.

The legal implications of marketing claims call for caution. Misleading or overstated claims can lead to regulatory or civil challenges. In the USA, the Federal Trade Commission provides some guidance on environmental marketing claims.^[2]

Contents [\[hide\]](#)

- 1 History
- 2 Greenhouse gas reduction market
- 3 Popularity and effectiveness
 - 3.1 Ongoing debate
 - 3.2 Confusion
 - 3.3 Statistics

Figura 8: slide 1 da atividade 1.

GREEN MARKETING

DISCUSSING THE TOPIC: *MEDIA VEHICLES*

- *What are the most common media vehicles used in Brazil? Why?*
- *Which are the most effective ones in your opinion? Explain.*

Media vehicle:

- *the specific program, publication, or promotional piece used to carry an advertising message*
(<http://en.mimi.hu/marketingweb/vehicle.html>)
- *specific print or electronic medium employed in an advertising campaign*
(<http://www.businessdictionary.com/definition/media-vehicle.html>)

Figura 9: slide 2 da atividade 1.

GREEN MARKETING

GETTING TO KNOW ABOUT IT

"Green marketing has found your clothes closet and is turning it into a new kind of advertising venue."

- What do you think they're talking about?
- Watch the video to find out.



Figura 10: slide 3 da atividade 1.

GREEN MARKETING

GETTING TO KNOW ABOUT IT

Discuss:

- Have you heard about this idea before? If so, where and what else do you know about it?
- What do you think of this kind of advertising?
- Could it work in Brazil/your city? Why (not)?

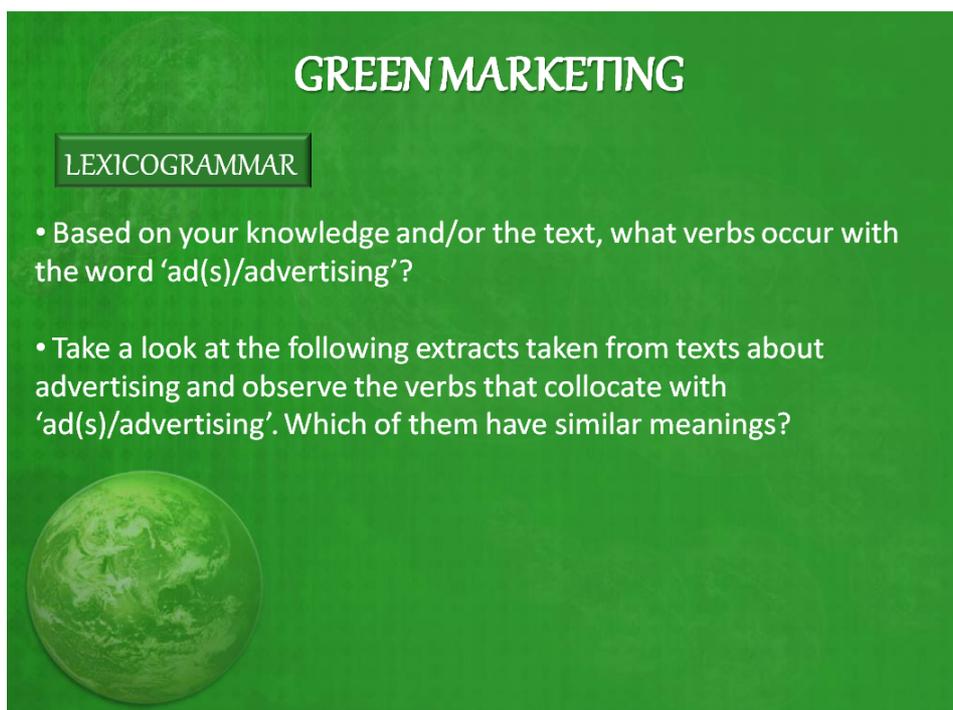


Figura 11: slide 4 da atividade 1.

A segunda parte da atividade, chamada de *Lexicogrammar*, consiste na exploração da linguagem; o que será explorado pode ter sido sugerido pelos alunos, ao

assistirem ao vídeo, por exemplo, pode ser algo que o professor considere interessante explorar baseado no uso da língua dos alunos ao longo da primeira parte da atividade ou em aulas, ou pode ter sido escolhido pelo professor ao analisar o texto do vídeo.

Propomos aqui o estudo dos verbos que coocorrem com *ad* por meio de linhas de concordância retiradas do *ad corpus* e por meio do *COCA (Corpus of Contemporary American English)*⁷⁸ (Ver Figuras 12, 13, 14, 15, 16 e 17). Essa escolha deveu-se ao fato de que verificamos que os alunos mostravam dificuldade no uso desses colocados. Nessa atividade, os alunos têm a oportunidade de comparar o uso dos verbos colocados de *ad* em *corpora* diferentes, de língua geral e da área em que atuam. É sugerida também a exploração do padrão *at the same time*, recorrente no texto, como podemos verificar nos resultados de sua análise por meio de uma ferramenta do *site Compleat Lexical Tutor*⁷⁹, que faz uma lista automática das palavras no texto e suas frequências (Ver Figuras 19 e 20). Outra sugestão é o estudo de *advertising venue* (Ver Figura 21).



GREEN MARKETING

LEXICOGRAMMAR

- Based on your knowledge and/or the text, what verbs occur with the word 'ad(s)/advertising'?
- Take a look at the following extracts taken from texts about advertising and observe the verbs that collocate with 'ad(s)/advertising'. Which of them have similar meanings?

Figura 12: slide 5 da atividade 1.

⁷⁸ <http://corpus.byu.edu/coca/>

⁷⁹ <http://www.lextutor.ca/>

GREEN MARKETING

LEXICOGRAMMAR

give online ads a bad name. Would the Times put an ad with a festering sore on the front page of the paper could never accept such a 'flashing' image. We did an ad at the time for After Six Formals with a model unfurl the first brands to take a leap of faith and place advertising in the uncharted "cyberspace" territory. But several roll ads to its video; the others have been selling advertising for a while now. "Brands need content – they need an attack, traditional administration allies are running ads meant to hold Democrats' feet to the fire. The United A CEO Bob Lord: "If you want a company to do banner ads and microsites, there are agencies that can do that and results? Is there something wrong with Google selling ads next to their search results? Why isn't it good for consumers more ad revenue from YouTube, which is placing ads across the bottom third of her videos. "I have noticed that men's offerings 20% in the past year and ramped up ads in men's fashion magazines and online to appeal to guys. EA Sports, publisher of the Tiger games, ramped up ads and PR for about six weeks before the game launched ads next to tweets, like Google. Twitter could run the ads based on the content of tweets combined with what broadcast vehicle of choice for Starbucks, which runs TV ads rarely. The chain ran its first TV ad in nearly a year just ads next to tweets, like Google. Twitter could run the ads based on the content of tweets combined with what broadcast vehicle of choice for Starbucks, which runs TV ads rarely. The chain ran its first TV ad in nearly a year just Part of the solution we developed was not just doing ads, but creating properties and platforms that acted as content stuck to his story and said if Gawker wouldn't run his ads, he'd take them elsewhere. Mr. Caruso said the scene change in 2010. Unilever, for example, places video ads across the gamut from Hulu and network sites like CBS

Figura 13: slide 6 da atividade 1 com concordâncias com os nós *ad*, *ads* e *advertising* do *ad corpus*.

GREEN MARKETING

LEXICOGRAMMAR

- Take a look at the results of search on the verbs that occur with *ad(s)/advertising* taken from *Corpus of Contemporary American English*.
- Do you think there are differences in the way these verbs are used?

 <div style="display: flex; align-items: center; justify-content: center;"> <div style="border: 1px solid black; border-radius: 50%; width: 40px; height: 40px; display: flex; align-items: center; justify-content: center; margin-right: 10px;">run</div> <div style="border: 1px solid black; padding: 10px 20px; font-size: 24px;">625</div> </div>	<div style="display: flex; align-items: center; justify-content: center; margin-right: 10px;"> <div style="border: 1px solid black; border-radius: 50%; width: 40px; height: 40px; display: flex; align-items: center; justify-content: center; margin-right: 10px;">place</div> <div style="border: 1px solid black; padding: 10px 20px; font-size: 24px;">228</div> </div>	<div style="display: flex; align-items: center; justify-content: center;"> <div style="border: 1px solid black; border-radius: 50%; width: 40px; height: 40px; display: flex; align-items: center; justify-content: center; margin-right: 10px;">put</div> <div style="border: 1px solid black; padding: 10px 20px; font-size: 24px;">127</div> </div>
--	--	--

Figura 14: slide 7 da atividade 1 – resultado da pesquisa feita no COCA (*Corpus of Contemporary American English*) sobre a frequência de ocorrência dos verbos *run*, *place* e *put* com os nós *ad*, *ads* e *advertising*.

GREEN MARKETING

LEXICOGRAMMAR

- Let's compare differences in use between *run* and *place* (by looking at data from the *Corpus of Contemporary American English*). By looking at the following charts, what can you find out about the uses of both *run* and *place* with ad(s)?



Figura 15: slide 8 da atividade 1.



Figura 16: slide 9 da atividade 1 – resultado da pesquisa feita no COCA (*Corpus of Contemporary American English*) sobre as ocorrências em diferentes tipos de textos dos verbos *run + ad*, *run + ads*, *place + ad* e *place + ads*.

GREEN MARKETING

LEXICOGRAMMAR

- Now go to <http://corpus.byu.edu/coca/> (*Corpus of Contemporary American English*) and read some examples of those verbs. Does 'ad' always have the same meaning? Is it related to Advertising in all the examples?
- What can you say about the differences between a corpus of General English and a specialized corpus of Advertising? What can you learn from both?



Figura 17: *slide 10* da atividade 1.

GREEN MARKETING

LEXICOGRAMMAR

- Watch the video again, but this time have a look at its script, and take notes of any words related or not to advertising that you find useful.

Example 1:

- 1) Using the tool *Concordance* in *Compleat Lexical Tutor*, we can find the pattern 'at the same time' (<http://www.lextutor.ca>).

SCREEN

- 2) Check how it's used in *COCA* (<http://corpus.byu.edu/coca/>).

SCREEN



Figura 18: *slide 11* da atividade 1.

TEXT CONCORDANCE OUTPUT - Windows Internet Explorer
 http://www.letutor.ca/concordancers/text_concord/u_conc.pl

Home > Concordancers > Text Concordance input > Concordance output

Tune-Up July 09 ... Concordance index for text: **TITLE** With secondary sort right of keyword. [Sort left] [View lines as text] [Check all] [none] [extract]

236 types (different words):
 Click to locate in concordance
 A 20
 ADAGE 2
 AD 1
 ADVERTISERS 1
 ADVERTISING 3
 AGE 1
 ALL 1
 ALTERNATIVE 1
 AM 2
 AMERICAN 1
 AN 7
 ANNUAL 12
 ANTHONY 2
 ANY 1
 ARE 2
 AS 2
 ASKED 1
 AT 9
 ATTORNEYS 1
 AWAY 1
 A LOT 1 2
 BAD 1
 BASED 0
 BE 2
 BEGAN 1
 BENEFIT 1
 BENEFITS 0
 BEST 2
 BIG 1
 BILLION 1
 BOB 1
 BUSINESS 1
 BUY 3
 CAMPAIGNS 1

322. leaner network, but we've also expanded our distribution at RETAIL and at hospitality. So as an example, we are in over
 323. developed new ways to connect hanger advertising to larger RETAIL chain marketing campaigns. "Part of the key to our gr
 324. r, as an example, we can_l be within a mile radius of every RETAIL store that they distribute to in the entire country o
 325. their business. So if a marketer has distribution at a key RETAILER, as an example, we can_l be within a mile radius of
 326. itality. So as an example, we are in over . . . Penney stores RIGHT now, we are in Hershey' hotels, theme parks, and virtu
 327. product that was a great media marketing vehicle and at the SAME time had significant environmental benefits." In , the
 328. ted in , it has reported annual triple digit growth. At the SAME time, it has developed new ways to connect hanger adver
 329. ted more than million eco hangers to , dry cleaners. At the SAME time, it offered a growing list of big advertisers a ne
 330. rtising venue. The Hanger Network and In home Media Company SAW new opportunities in an eco friendly clothes hanger made
 331. tonio Neves and this is minute Adage, your daily video news SHOW. Green marketing has found your clothes closet and is t
 332. e as a great media marketing vehicle and at the same time had SIGNIFICANT environmental benefits." In , the Hanger Network
 333. e to in the entire country on a national or a local basis." SINCE the Network started in , it has reported annual triple
 334. lso expanded our distribution at retail and at hospitality. SO as an example, we are in over . . . Penney stores right now
 335. e distribution footprint that best supports their business. SO if a marketer has distribution at a key retailer, as an e
 336. munications." And that' it for this minute AdAge. Thanks SO much for being with us. I am Antonio Neves at Advertising
 337. duced to the person who had invented the eco hanger through SOME attorneys that I worked with and they asked me to take
 338. ked me to take a look at an opportunity that on the surface SOUNDED a little crazy, but then a_lot_1 of the best market
 339. little crazy, but then a_lot_1 of the best marketing ideas SOUNDED crazy the first time a_lot_1 of us heard them, and t
 340. ing list of big advertisers a new way to effectively target SPECIFIC demographics. "Because dry cleaning is such a dense
 341. country on a national or a local basis." Since the Network STARTED in , it has reported annual triple digit growth. At
 342. n example, we can_l be within a mile radius of every retail STORE that they distribute to in the entire country on a nat
 343. at hospitality. So as an example, we are in over . . . Penney STORES right now, we are in Hershey' hotels, theme parks, an
 344. hannel and the product which was replacing wire hangers had SUCH a bad impact on the environment and there was a chance
 345. vely target specific demographics. "Because dry cleaning is SUCH a dense distribution wearing all DMAs in the country, b
 346. keter, we can_l follow the distribution footprint that best SUPPORTS their business. So if a marketer has distribution a
 347. they asked me to take a look at an opportunity that on the SURFACE sounded a little crazy, but then a_lot_1 of the best
 348. cycled paper. It began putting ads on the hanger' cardboard SURFACES and giving away truckloads of them to dry cleaners

Figura 19: tela da atividade 1 do slide 11 – tela do *Compleat Lexical Tutor* com as palavras frequentes do texto usado na atividade 1

Corpus of Contemporary American English (COCA) - Windows Internet Explorer
 http://corpus.byu.edu/coca/

CORPUS OF CONTEMPORARY AMERICAN ENGLISH
 410 MILLION WORDS, 1990-2010
 COMPARE RESULTS: COCA COHA TIME BNC

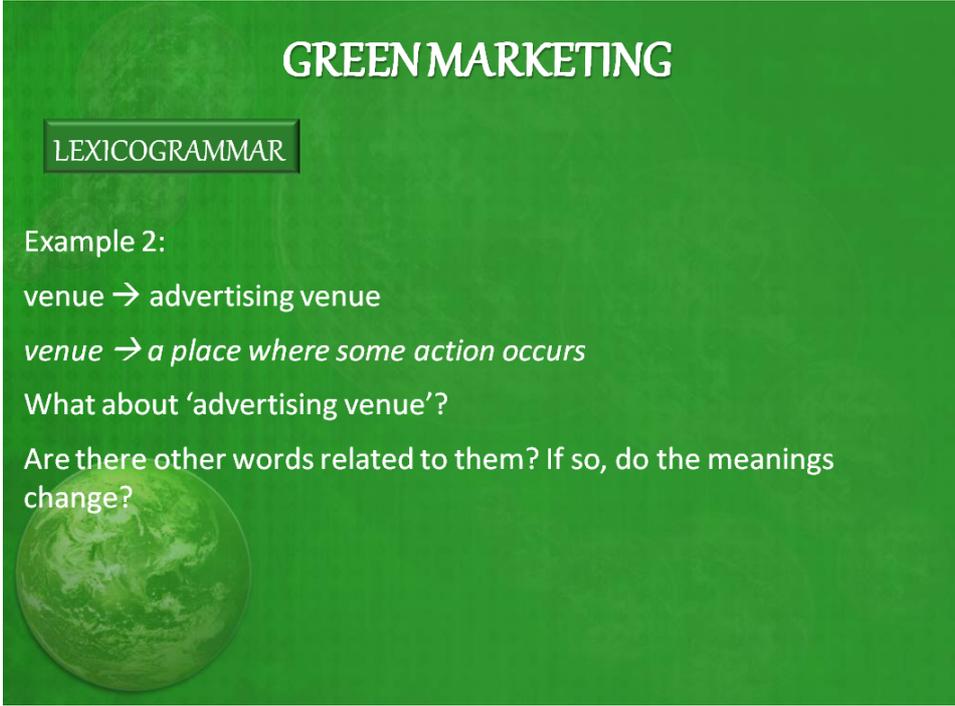
SEARCH STRING: "at the same time"
 WORD(S):
 COLLOCATES:
 POS LIST:
 SECTIONS: SHOW

CONTEXT	TOT
1 . AT THE SAME TIME	7194
2 BUT AT THE SAME TIME	2488
3 , AT THE SAME TIME	1926
4 AND AT THE SAME TIME	1926
5 AT THE SAME TIME	1816
6 WHILE AT THE SAME TIME	1047
7 " AT THE SAME TIME	500
8 ALL AT THE SAME TIME	352

KEYWORD IN CONTEXT DISPLAY
 SECTION: NO LIMITS
 PAGE: 1 / 11
 SAMPLE: 100 200 500 1000

CLICK FOR MORE CONTEXT	CONTEXT
1 2010 SPOK NPR_TalkNat A B C	not actually seem yet to be a sustainable model, whereas - while at the same time providing some j
2 2010 SPOK PBS_NewsHour A B C	How do you negotiate your role as journalist, as writer, while at the same time almost being, I do nt
3 2010 FIC Bk:ThiefShadows A B C	a skillet for placing my hand where she had given no warrant while at the same time I was attempti
4 2010 FIC Bk:WarStars A B C	them to swing a thick armor shield towards any attacks from Nergal while at the same time keepi
5 2010 FIC Bk:GhostsDoingOrange A B C	and move and change expression - a sudden, exaggerated grimace, while at the same time the mai
6 2010 FIC Bk:GhostsDoingOrange A B C	By its light I could see the battle in the corn, while at the same time we were reinforced quite unexp
7 2010 FIC Bk:HowSeducerSinner A B C	her mind racing to formulate a response that would firmly discourage him while at the same time sp
8 2010 MAG Esquire A B C	most powerful politician in the world and responded to the pain politically while at the same time rer
9 2010 NEWS Denver A B C	, the focus is on large sums of TV contract money, while at the same time the struggles that student

Figura 20: tela da atividade 2 do slide 11 – tela do *COCA (Corpus of Contemporary American English)* com resultados da pesquisa sobre *at the same time*.



GREEN MARKETING

LEXICOGRAMMAR

Example 2:
venue → advertising venue
venue → a place where some action occurs

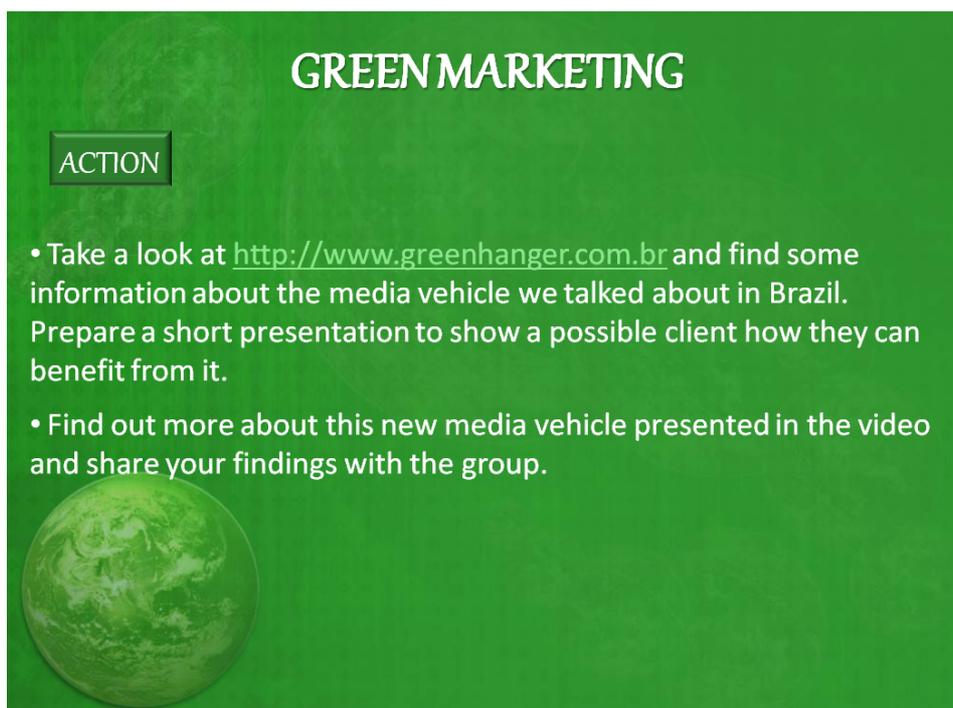
What about 'advertising venue'?

Are there other words related to them? If so, do the meanings change?



Figura 21: *slide 12 da atividade 1.*

Por fim, propomos que os alunos façam uma pesquisa sobre o assunto discutido em aula ou sobre a empresa mencionada no vídeo para apresentar em aula (Ver Figura 22). Essa parte, que chamamos de '*Action*', instiga os alunos a usar os padrões e a linguagem em uma atividade real, ou seja, que pode ocorrer em sua atuação profissional.



GREEN MARKETING

ACTION

- Take a look at <http://www.greenhanger.com.br> and find some information about the media vehicle we talked about in Brazil. Prepare a short presentation to show a possible client how they can benefit from it.
- Find out more about this new media vehicle presented in the video and share your findings with the group.

Figura 22: slide 13 da atividade 1.

3.3 Atividade 2

Essa atividade parte de um padrão a ser explorado, que é o *ad campaign*. À semelhança da Atividade 1, é aberta, não segue um formato rígido. Desenvolvemos a atividade utilizando o *corpus* de estudo e uma das ferramentas do programa *AntConc*⁸⁰, por ser de mais fácil uso, de acordo com pareceres de professores que já utilizaram o *Wordsmith Tools* e o *AntConc*. Sugerimos também que o professor explore o uso dos padrões usando o COCA (*Corpus of Contemporary American English*), mencionado na Atividade 1. Da mesma maneira, é possível preparar as linhas de concordância e as atividades previamente e utilizá-las em sala de aula em formato impresso. Calculamos 2 aulas de 1 hora e 30 minutos para essa atividade, lembrando que esse tempo depende do contexto de ensino, do rendimento dos alunos, dos recursos disponíveis pelo professor em sala de aula, entre outros fatores.

A princípio, sugerimos que os alunos reflitam e discutam sobre uma campanha publicitária, já que o padrão a ser explorado é *ad campaign*. A primeira parte da atividade traz linhas de concordância para o estudo do padrão e seus colocados. Os

⁸⁰ <http://www.antlab.sci.waseda.ac.jp/software.html>

alunos observam as linhas do *corpus* de estudo e associam os usos aos propósitos (Ver Figuras 23, 24 e 25).



Figura 23: slide 1 da atividade 2.

2) Look at the right of 'ad campaign' and mark:

a → for the ones that are about the agency (that created the campaign)
b → for the ones that are about the product / brand
c → for the ones that are about the purpose of the campaign

() athan Fox somewhat, but he largely credits an **ad campaign** from Publicis Worldwide, New York, th
() 75 employees. Chrysler breaks the "New Day" **ad campaign** from BBDO. The spots use third-party
() or a sweet deal in which LBL will create a new **ad campaign** to stimulate sales at 50% off their nor
() sen, CMO of the clinic, which started a national **ad campaign** earlier this year from Adworks, Washi
() cognized the need for an organized, strategic **ad campaign** to drive the ongoing demobilization eff
() new standard of the world," and forthcoming **ad campaigns** for the brand will center on a theme
() h new storytelling techniques and high-impact **ad campaigns** like the great ads for Apple. But ev
() rs to attend "the unveiling of its first global **advertising campaign**" from TBWA/Chiat/Day tomo
() nfirm details about the upcoming national **advertising campaign** for Via, other than the offi
() at's not to say there aren't marketing and **advertising campaigns** for the Beatles' selected p

2) Underline the words that common to each one. What pattern can you see?

ad campaign ← from
for
to

Figura 24: slide 2 da atividade 2.

3) Read the excerpts below and identify:

- ◆ which / what kind of ad campaign they are about;
 - ◆ in which ones you know about the company or agency involved in the campaign.
- a) Several years ago, the Syracuse University football team did an **ad campaign** in New York City...
 - b) "The Beatles: Rock Band," for instance, launched last week with an estimated \$20 million **ad campaign**...
 - c) Sergio Jaramillo, Colombia's vice minister for defense, recognized the need for an organized, strategic **ad campaign** to drive the ongoing demobilization effort...
 - d) ...LBL will create a new **ad campaign** to stimulate sales at 50% off their normal sticker price for creative services.
 - e) Chrysler breaks the "New Day" **ad campaign** from BBDO.
 - f) It was a nod, of course, to the software giant's ongoing **ad campaign** from agency Crispin Porter & Boguysky.
 - g) The company wouldn't confirm details about the upcoming national **advertising campaign** for Via...
 - h) New York-based journalists have gotten invites from the French tire maker and awardee of restaurant stars to attend "the unveiling of its first global **advertising campaign**" from TBWA/Chiat/Day tomorrow...

Figura 25: slide 3 da atividade 2.

Em seguida, sugerimos que os alunos pesquisem no COCA as palavras que coocorrem com o padrão, comparando esse uso com o uso encontrado no *corpus* de estudo (Ver Figura 26). Para realizar uma atividade com um texto com mais exemplos do padrão, selecionamos um texto do *corpus* de estudo por meio da ferramenta *Concordance Plot* do programa *AntConc* (Ver Figura 27). Selecionado o texto, desenvolvemos uma atividade em que retiramos algumas palavras que coocorrem com *ad campaign* para que os alunos as reposicionem no texto, como uma atividade de preencher lacunas (Ver Figura 27).

3) Do some research on the corpus collected or on COCA (*Corpus of Contemporary American English*) and find more instances of:

- what collocates with ad campaign
- verbs that collocate with ad campaign(s)
- adjectives that collocate with ad campaign(s)

COCA – * + ad campaign*	
	CONTEXT
1	AN AD CAMPAIGN
2	NEW AD CAMPAIGN
3	THE AD CAMPAIGN
4	MILLION AD CAMPAIGN
5	" AD CAMPAIGN
6	'S AD CAMPAIGN
7	TELEVISION AD CAMPAIGN

COCA – verb + an ad campaign	
	CONTEXT
1	LAUNCHED AN AD CAMPAIGN
2	DO AN AD CAMPAIGN
3	DEVELOPING AN AD CAMPAIGN
4	DESIGNING AN AD CAMPAIGN
5	DESIGNED AN AD CAMPAIGN
6	LAUNCHING AN AD CAMPAIGN
7	BEGAN AN AD CAMPAIGN

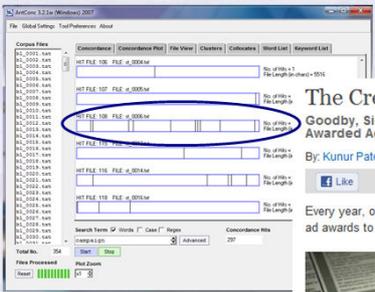
<http://corpus.byu.edu/coca/>

4) Are these collocations similar or different from the ones found in the Advertising corpus?

Figura 26: slide 4 da atividade 2.

FURTHER PRACTICE WITH A TEXT

Finding a text for further practice using the corpus collected and *AntConc* (*Concordance Plot*)



The Creativity Awards Report
 Goodby, Silverstein & Partners Tops Agency Winners, While BBDO Remains the Most-Awarded Ad Network
 By: Kunur Patel | Published: September 28, 2009

Every year, once the gold dust has settled, Ad Age's *Creativity combs* through the results of the leading global ad awards to determine which people, companies, and campaigns claimed the most hardware.

This year, top agencies earned distinction largely due to interactive work. But it's not just interactive know-how that sets an agency apart to awards juries. The most-awarded agency of 2009 is Goodby & Silverstein & Partners, a shop that simultaneously was named Cannes' digital agency of the year and, according to our count, had the most-awarded TV spot this season with "Rabbit" for Comcast.

What's more, the San Francisco-based Omnicom agency didn't earn this distinction for hitting one home run that won everything. Last year's top winners could chalk success up to standout campaigns like HBO's "Voyeur" and "Halo 3." A broad range of award-winning work and strength across media proved fruitful for Goodby, which won for work such as the YouTube-quaking "Wanoland Shake It" for Nintendo, the split-screen "There Can Only Be One" spots for the NBA that inspired spoofs on the cover of Time magazine and on SNL, the "Save the Honey Bees" campaign for Honey Bunches of Cereal, and the virtual haunted "Hotel 626" for Doritos. Goodby also won for the first augmented-reality campaign for a major advertiser, the GE Plug into the Smart Grid website, as well as the data-crunching Sprint Now widget.

See all the winning work | Last year, Goodby was the eighth agency on our awards list. Agency co-

Figura 27: slide 5 da atividade 2 - Exemplo de como encontrar um texto que contenha as palavras ou padrões para desenvolver a atividade usando a ferramenta *Concordance Plot* do programa *AntConc*.

FURTHER PRACTICE WITH A TEXT

5) Read the first part of the text below and decide in which of the black spaces the following words fit:

standout
for
"Save the honey bees"
most-awarded
for

6) Talk about:

- ♦ the campaigns mentioned in the text;
- ♦ ongoing / forthcoming campaigns.

The Creativity Awards Report

Goodby, Silverstein & Partners Tops Agency Winners, While BBDO Remains the Most-Awarded Ad Network

By: Kunur Patel Published: September 28, 2009

Like Tweet 0 + 0 [+ Share | Email | Reprints | Print |]

Every year, once the gold dust has settled, Ad Age's Creativity **combs through the results** of the leading global ad awards to determine which people, companies and campaigns claimed the most hardware.

This year, top agencies earned distinction largely due to interactive work. But it's not just interactive know-how that sets an agency apart to awards juries. The most-awarded agency of 2009 is Goodby, Silverstein & Partners, a shop that simultaneously was named Cannes' digital agency of the year and, according to our count, had the most-awarded TV spot this season with "Rabbit" for Comcast.

What's more, the San Francisco-based Omnicom agency didn't earn this distinction for hitting one home run that won everything -- last year's top winners could chalk success up to **campaigns** like HBO's "Voyeur" and "Halo 3." A broad range of award-winning work and strength across media proved fruitful for Goodby, which won for work such as the YouTube-making "Wanland Shake It" for Nintendo, the edit.

This **campaign** for Tourism Queensland was the **campaign** of 2009.

<http://adage.com/article/agency-news/creativity-awards-report-2009/139270/>

Figura 28: slide 6 da atividade 2.

3.4 Atividade 3

Esta atividade teve como ponto de partida um tema dentro da área de Publicidade. Escolhemos um texto que fosse relevante para a discussão do tema *digital* e que trouxesse possibilidades de exploração de alguns padrões relacionados à área, bem como da língua em geral. Nesta atividade, seguimos os passos apresentados por Willis e Willis (2007); portanto, segue uma ordem mais rígida do que as duas atividades anteriores. Como os autores usa o termo 'tarefa', como explicamos na seção 1.2.2.1, faremos o mesmo nesta seção. Segundo esses autores, "uma aula baseada em tarefa provavelmente envolveria não uma única tarefa, mas uma sequência de tarefas. Essas tarefas relacionam-se umas com as outras"⁸¹ (WILLIS e WILLIS, 2007, p. 21).

O primeiro ciclo da tarefa (Ver Figuras 29, 30 e 31), intitulado *Talk about it*, tem o objetivo de introduzir o tópico do texto, envolvendo os alunos em uma discussão (*priming*). Nesse momento, os alunos podem compartilhar suas opiniões e experiências relacionadas ao tópico e o professor verifica seu conhecimento com relação ao tópico em si e ao vocabulário que o envolve. Para que haja uma preparação (*preparation*) dos alunos, são sugeridas algumas questões associadas a trechos retirados do texto para

⁸¹ "a task based lesson would probably involve not a single task, but a sequence of tasks. These tasks relate to one another" (WILLIS e WILLIS, 2007, p. 21)

discussão. Após a preparação, os alunos iniciam a tarefa principal do ciclo (*target task*), que é a previsão de algumas informações do texto. Essa tarefa é feita individualmente para que depois seja discutida. Se a discussão for feita em pares ou grupos, os alunos podem resumir as previsões do par ou do grupo para a sala. Se for feita individualmente, pode-se pedir que o aluno escreva um parágrafo ou tópicos em sala de aula ou antecipadamente (como tarefa para casa) sobre o que foi discutido com o professor ou refletido. Esse é, portanto, o momento em que os alunos focam na linguagem que irão usar para reportar suas ideias (*planning*). O professor irá observar e mediar quando necessário. A seguir, os alunos irão reportar (*report*) suas ideias para a sala ou para o professor.



TASK I – DIGITAL ADVERTISING

1) TALKABOUT IT

1a) Discussion:

- What are some of the predictions about the growth in digital advertising?
- How do you feel about this growth both nationally and globally?
- Comment on these elements of digital advertising regarding their uses, effect and relevance:

banner ad – mobile – SMS – blog – RSS – spam

Figura 29: slide 1 da atividade 3.

TASKI – DIGITAL ADVERTISING

1b) These are some predictions about digital advertising for 2008 taken from an article published on the site AdAge (www.adage.com). Read them and discuss the questions that follow.

ONLINE ADVERTISING WILL GROW 29% IN 2008
How much did online advertising grow in 2008?

ADVERTISERS WILL LEARN TO LOVE FACEBOOK AND YOUTUBE
How are advertisers taking advantage of *Facebook* and *Youtube*?

A BIG DAILY PAPER WILL GO ALL-DIGITAL
How much percent of the staff did the paper that went digital fire?

2008 WILL BE THE YEAR OF MOBILE ADVERTISING
How long have we been hearing this prediction?

BROADCAST TV DIES, WHILE WEB VIDEO BREAKS THROUGH
How much in millions do you think online video grew in advertising?



Figura 30: *slide 2 da atividade 3.*

TASKI – DIGITAL ADVERTISING

1c) Individually, read the following extracts and match them to the predictions above. (There are only four extracts and five predictions.) Then guess the numbers that are missing.

Well, online advertising did grow in 2008, but just ★ in the first six months...

The paper fired ★ of its staff (142) and will focus on digital.

You've been hearing that prediction since ★...

...online video took a pittance in advertising -- perhaps ★ million if you believe eMarketer.

1d) Share your ideas in pairs or groups and decide on the most probable predictions to be presented to the class.

1e) Report the selected predictions to the class. In order to check your guesses, scan the text to find your answers.

Figura 31: *slide 3 da atividade 3.*

O segundo ciclo da tarefa (Ver Figura 32) é chamado de *Know about it*, já que os alunos irão ter contato com o texto autêntico, comparando suas previsões com as do artigo, relacionando-o ao seu conhecimento prévio. A primeira parte do ciclo (*target*

task) pode ser feita em casa e os alunos trazem para a sala de aula os pontos relevantes para serem discutidos. Assim, os alunos têm oportunidade de mais bem preparar a apresentação dos pontos selecionados para discussão (*planning*). Em sala de aula, os alunos podem compartilhar os pontos relevantes (*report*) e participar de uma discussão sobre um tópico de seu interesse profissional.



TASKI – DIGITAL ADVERTISING

2) KNOW ABOUT IT

2a) Read the text and choose some of the points that called your attention to be discussed in class. After you read it, think about what the heading of the article is. The teacher will tell you after you share your guesses.

2b) Discuss the points you selected and the possibilities for the heading with the class.

Figura 32: slide 4 da atividade 3.

A terceira parte da tarefa (Ver Figuras 33, 34, 35 e 36), *Explore it*, tem o foco no estudo do léxico. A princípio, há o desenvolvimento de uma tarefa (*priming + preparation*) que tem como objetivo despertar o aluno para a observação de colocações e padrões. As tarefas seguintes (*task + planning e target task*) focam nas palavras-chave selecionadas, levando os alunos à sua exploração no texto e por meio de linhas de concordância, encontrando colocações e padrões. O foco dessas tarefas é na forma, mas os alunos estão engajados em discussões e reflexões, passando pelas fases de planejamento e relato (*planning + report*) da mesma forma que fizeram nas outras partes.

TASKI – DIGITAL ADVERTISING

3) EXPLORE IT

3a) Read these extracts from the text and try to fill in with the missing words. If necessary, look at some options at the end. Think about the meanings and uses of these expressions.

Even if the economy takes a dive, online ads won't. In _____¹, online will benefit from a recession, as marketers move spending from TV to cheaper, more measurable media.

ADVERTISERS WILL LEARN TO LOVE FACEBOOK AND YOUTUBE – Neither Facebook's _____² YouTube's importance to marketers matches its scale on the web.

DIGG WILL BE ACQUIRED – It's no _____³ it's a perennial prediction -- every year Digg does seem to do a little acquisition dance with someone.

You've been hearing that prediction since 1996, and every year people trot it out again. Sure, there were _____⁴ of advertising milestones this past year...

Back in January, J.P. Morgan's Imran Khan predicted a 34% earnings-per-share rise for major internet stocks. Needless to _____⁵, it didn't happen.

1) fact / actually 2) nor / or 3) fact / wonder 4) great / lots 5) speak / say

Figura 33: slide 5 da atividade 3.

TASKI – DIGITAL ADVERTISING

STUDY OF ONLINE AND DIGITAL

3b) Take a look at the text and find the uses of the word *online*, focusing on what comes before and after it and its meanings. Take notes of your findings to discuss with the class.

3c) Now study these two blocks of lines from other texts about digital advertising:

comes down to is relatively simple: The **online** ad and media industries must come rence between success or failure for an **online** campaign." The ability to target tive. The screens are too small, some **online** creatives say. And offline create

What other things can be online? Take a look at what comes after online. Think of some more examples or take a look at some more concordance lines to find them.

Figura 34: slide 6 da atividade 3.

TASKI – DIGITAL ADVERTISING

your identity with you wherever you go **online**. Facebook Connect, Google Friend Connect, and MySpace emphasize on community. "Women go **online** more for communication," said a woman who saw a TV ad, do they go to the store, go **online**, research more? When they search for a man's world out there -- at least it is **online**. Women's-content websites, such as MySpace, sell personal stories. Clips can be seen **online**, and fans can vote on the stories. People who order **online** are also more likely to jump on to buy from each individual store.

What are the verbs used with online? What other verbs can be used? Think of them or take a look at more concordance lines to find them.

3d) Share your ideas with the class.

Figura 35: *slide 7 da atividade 3.*

TASKI – DIGITAL ADVERTISING

3e) Now study the uses of the word *digital*, the same way you did with *online*. Here are some other lines, but mixed up:

digital markets in the world when it comes to **digital**. About 20% of media spending is in the United States. Nike continues to be the envy of many **digital** agencies that too often have a "creative and rigorous" look at how to attack **digital**. "And certainly in the branding of National Advertisers added a **digital** category this year to its Multichannel Advertiser (AdAge.com) -- Preschoolers are going **digital** -- some as young as 2 years old. The agency's staff are now dedicated to **digital**, including the additions that came with banner ads (25%). Unfortunately, **digital** marketing is still wired for the dollars move from traditional media to **digital**. MySpace has a much bigger **digital** showed that when it comes to building **digital** platforms for marketers, it's on a **digital** with no standard metric yet in place for **digital** out of home, the ad model is con-

3f) In what ways are the uses of *online* and *digital* similar?

Figura 36: *slide 8 da atividade 3.*

Por fim, sugerimos o *Follow-up* (Ver Figura 37), cujo objetivo é dar oportunidade para que o aluno retome suas ideias, usando o vocabulário usado nas aulas de forma mais elaborada e planejada. Em um dos itens o aluno tem a possibilidade de

escrever com um objetivo real, postando seu comentário no *site* e participando de uma discussão genuína.

TASK1 – DIGITAL ADVERTISING

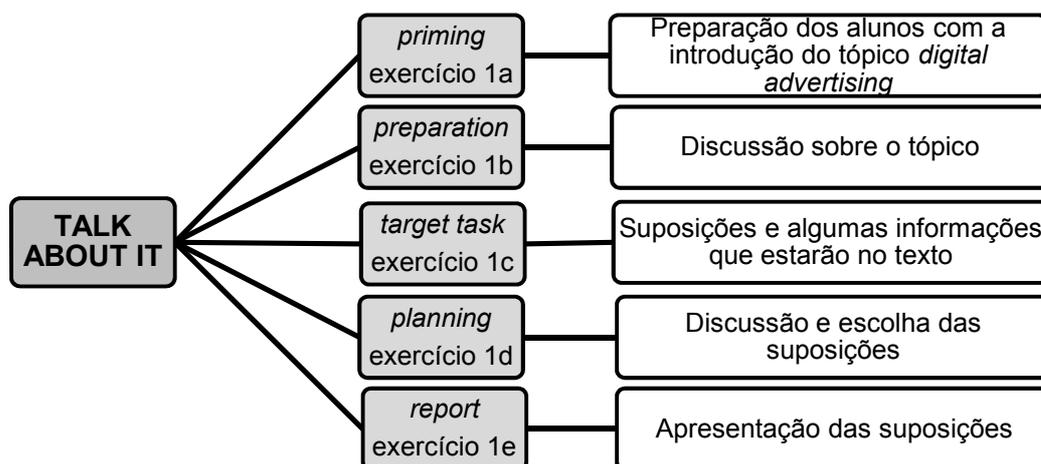
FOLLOW-UP

Choose one or more topics:

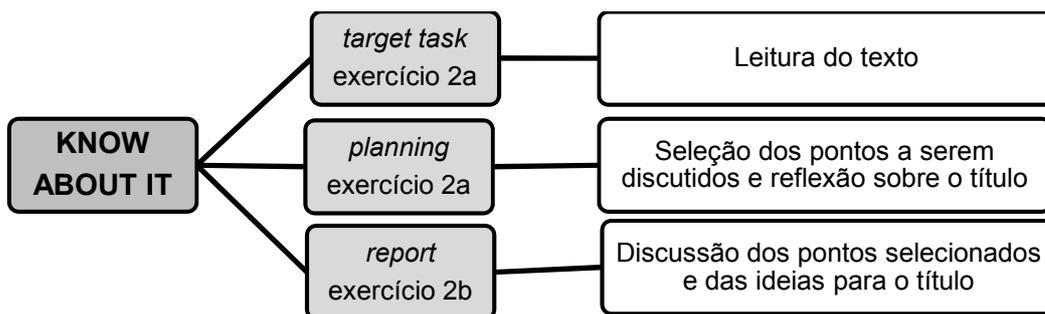
- Write a paragraph using the ideas discussed in class stating your own opinions on the predictions.
- Write a paragraph with your comments on the text to be posted on the site.
- Write some predictions for 2009 related to *digital*.
- Read the [eMarketer's](http://www.emarketer.com/Article.aspx?R=1006813) predictions for 2009 on <http://www.emarketer.com/Article.aspx?R=1006813> and write a paragraph commenting on them.

Figura 37: *slide 9 da atividade 3.*

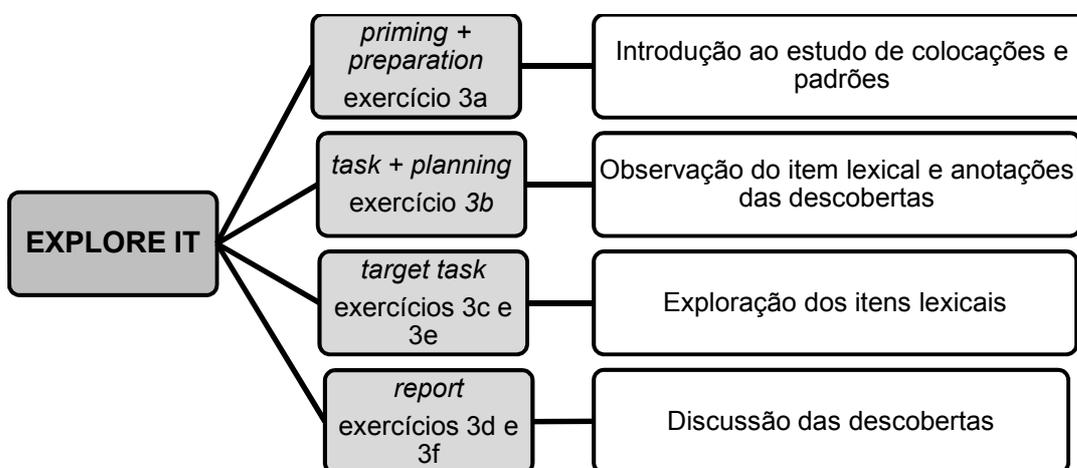
Nos esquemas 2, 3 e 4 a seguir podemos verificar os três ciclos da Atividade 3.



Esquema 2: primeiro ciclo da Atividade 3



Esquema 3: segundo ciclo da Atividade 3



Esquema 4: terceiro ciclo da Atividade 3

Finalizamos aqui o Capítulo 3, no qual explicamos o desenvolvimento das atividades de ensino de ILE para profissionais de Publicidade a que nos propomos. Na seção a seguir, faremos as considerações finais sobre a pesquisa, incluindo uma discussão sobre as atividades propostas.

CONSIDERAÇÕES FINAIS

Retomamos aqui os pontos principais do percurso de nossa pesquisa, apontando limitações e sugerindo ações futuras com relação a pesquisas, desenvolvimento de material didático e aplicações dos resultados obtidos no ensino.

Nossa prática de ensino nos permitiu verificar as necessidades dos alunos que atuam na área de Publicidade na aprendizagem de ILE. Observamos que os alunos possuíam conhecimento do vocabulário do meio, mas apresentavam dificuldades em utilizá-lo. Devido à falta de materiais didáticos voltados para esse público que possibilitam aos alunos o contato com a linguagem da área, professores optam por criar suas próprias atividades com textos retirados da *Internet*; entretanto, essas atividades comumente focam primordialmente questões gramaticais e de vocabulário separadamente, o que não auxilia o aluno a entender o funcionamento da língua em uso. Com a LC, que vê a padronização da língua e considera a lexicogramática, professores e alunos podem enxergar as relações entre os itens da língua, ou seja, as palavras frequentemente usadas com o vocabulário da área, formando padrões que auxiliam os alunos a comunicar-se com mais fluência (SINCLAIR, 1991).

Com nossa pesquisa, procuramos mostrar o desenvolvimento de atividades para o ensino de ILE com base no estudo dos padrões lexicogramaticais distintivos do *ad corpus*, um *corpus* compilado com o objetivo de servir para a pesquisa dos padrões distintivos da área de Publicidade. Encontramos diversos padrões com as palavras-chave resultantes da comparação entre o *ad corpus* e o BNC, que usamos como *corpus* de referência. Pudemos analisar os diversos padrões e colocados em nossa trajetória; por motivos de tempo e espaço, apresentamos a análise dos padrões com o nóculo *ad* e seu uso, quais sejam: *ad spending*, *ad network(s)*, *ad revenue* e *ad campaign(s)*. Analisamos ainda alguns dos verbos colocados de *ad*, por ser uma questão de difícil uso pelos alunos conforme nossa experiência em aulas em agências de Publicidade.

Selecionamos alguns dos padrões encontrados para mostrar o desenvolvimento das atividades propostas. As duas primeiras atividades, que consideramos ‘abertas’ não foram desenvolvidas seguindo somente uma abordagem de ensino. Elas envolveram diversos pressupostos de diferentes abordagens de ensino em geral e de ensino com *corpora*, como o CLT (BROWN, 2007), o DDL (JOHNS, 1991), Lexical Syllabus (WILLIS, 1990) e as propostas de Berber Sardinha (2009; 2011) de atividades centradas

na concordância, no texto e multimídia. A terceira atividade foi baseada na proposta de Willis e Willis (2007) do TBT e, por seguir alguns passos delineados pelos autores, consideramos mais ‘fechadas’. Ao contrário do que imaginávamos, a última atividade demandou mais tempo de preparo do que as primeiras, talvez pela necessidade de se seguir os passos e encaixar as ideias dentro de uma estrutura já delineada. As duas primeiras permitiram maior criatividade e liberdade, visto que não seguimos uma estrutura pronta, mas sim pressupostos de diversas abordagens que dialogam entre si. Mostramos, assim, que o uso de *corpora* na sala de aula pode ser feito em concordância com diversas abordagens de ensino estruturas (*frameworks*), contanto que seus pressupostos sejam seguidos.

Com o uso da LC em sala de aula, objetivamos levar o aluno à reflexão sobre a linguagem e ao uso da língua com eficiência, dando-lhe voz, de forma que o ensino possa partir de seus questionamentos sobre esse uso da língua e o professor possa agir como um guia e facilitador, ajudando-o a encontrar suas respostas e customizando o ensino às suas necessidades. Acreditamos, assim, que essa pesquisa poderá servir de incentivo para que professores possam mudar seus paradigmas, questionar e refletir sobre suas práticas, livrando-se da resistência ao novo e do conformismo (CELANI, 1998), e engajarem na pesquisa científica, no mundo acadêmico (MOITA LOPES, 2002), para que nossas práticas de ensino não sejam baseadas em teorias ‘exportadas’, mas que possam estar efetivamente adequadas ao contexto de nossos alunos e às suas necessidades sociais com relação ao uso do idioma.

Segundo Moita Lopes (2002, p. 86, grifo meu), muitas pesquisas em Linguística Aplicada estão “orientadas *para* a sala de aula” e não “*na* sala de aula”, ou seja, o pesquisador e os envolvidos no processo, professores e alunos, estão distantes entre si. No entanto, o autor aponta para uma tendência de pesquisa, que estamos procurando fazer, relacionada ao movimento do professor-pesquisador, a pesquisa-ação, na qual o pesquisador está envolvido diretamente com a investigação crítica de sua prática.

Essa é uma das questões mais discutidas hoje com relação ao ensino de LC com *corpora*. Conforme O’Keefe et al. (2007, p. 246), para que a LC possa ‘informar’ para o ensino, professores precisam ‘informar’ a LC, visto que os professores são os melhores mediadores entre as descobertas com *corpora* e a prática. Os autores fazem algumas sugestões para ajudar professores a sentirem-se confortáveis em usar *corpora* no ensino, sendo elas maior disponibilidade de *corpora* e ferramentas *online* e uma maior

diversidade de *corpora*. Mostramos em nossa pesquisa algumas ferramentas e *corpora* disponíveis e como o professor pode compilar um *corpus* e analisar o uso da língua, transpondo seus achados para atividades de ensino.

Uma das limitações desse trabalho corresponde à compilação do *corpus*, que pode ser trabalhosa e demandar tempo. A análise e seleção dos itens lexicogramaticais requer certa prática e a mudança da visão da língua como estruturas a serem preenchidas com o léxico para a lexicogramática. O desenvolvimento das atividades também demanda tempo e criatividade para que as atividades sejam variadas e compreendam assuntos variados e interessantes ao aluno. Esses são alguns dos ‘contras’ assinalados por Berber Sardinha (2011) e por professores e pesquisadores que usam a LC no ensino em conversas informais. Acreditamos, no entanto, que quando o professor adquire certa prática, a preparação de aulas e de atividades com *corpora* torna-se menos trabalhosa e é feita em menos tempo, visto que o professor pode reutilizar algumas estruturas de atividades, somente modificando o texto. Em nossa opinião, esse trabalho compensa, pois não é mais possível conceber o ensino da língua sem considerar seu uso e os padrões lexicogramaticais que a envolvem; as descobertas que diversos pesquisadores já fizeram e fazem e que nós estamos começando a fazer são inestimáveis. Acreditamos que ensinar o idioma dessa maneira traz mudanças importantes para os alunos, tanto com relação à visão da linguagem no que diz respeito à padronização, quanto ao desenvolvimento de habilidades e capacidades de uso de tecnologias, de manipulação de dados, de descoberta, reflexão e conclusões.

Entendemos hoje que não precisamos nos rebelar contra as abordagens e métodos de ensino que foram desenvolvidos e utilizados ao longo do tempo. Ao contrário, precisamos, sim, ter um olhar crítico com relação a eles, tirando proveito do que eles nos oferecem. Um exemplo disso é proposta de Willis (1990), do Currículo Lexical (*Lexical Syllabus*), da qual usamos alguns pressupostos em nossa pesquisa. Ela é exemplificada por Moita Lopes (2002) como um equívoco de ignorar que um fato linguístico não tem relação direta com o ato de ensinar e aprender línguas. Como explica o autor, os itens lexicais mais comuns da língua podem não ser os que mais bem facilitam o processo da aprendizagem, ou seja, essa é uma proposta que priorizou aspectos linguísticos e não práticos, de ensino-aprendizagem. Por outro lado, em nossa concepção, dois dos maiores equívocos na criação de materiais didáticos, que Willis (1990) leva em consideração, são (a) o de não serem baseados em ou trazerem exemplos

de língua autêntica e (b) o da separação entre o léxico e a sintaxe, em uma abordagem estruturalista, na qual os alunos utilizam o léxico para preencher lacunas na sintaxe (BERBER SARDINHA, 2004). Assim, podemos manter um diálogo entre as propostas e pesquisas já realizadas com nosso contexto, refletindo e questionando para ir adiante.

Infelizmente, em nosso país ainda são poucas e limitadas as propostas de LC para o ensino de ILE (BERBER SARDINHA, 2004). Se por um lado podemos tirar proveito do que já foi feito, por outro essa situação abre-nos a oportunidade de ir à busca da construção de novos conhecimentos. Por esses motivos e pela realidade com a qual nos deparamos em sala de aula é que nos propomos a desenvolver essa pesquisa, que estabelece uma comunicação entre diversas áreas, sendo: (a) a área de Publicidade, por meio de nossos próprios alunos, que nos trazem suas necessidades de comunicação e suas experiências profissionais, bem como materiais autênticos (textos de revistas, da *Internet*, e-mails, entre outros) que podemos utilizar nas aulas; (b) a LC, que nos possibilita explorar materiais autênticos, focando na lexicogramática, seu significado e uso, em contraposição ao ensino focado em estruturas e formas e (c) a área de ILE, por meio de suas abordagens, métodos e experiências de interação em sala de aula. Essas são as áreas com as quais dialogamos em nossa pesquisa, mas com certeza outras podem ser acrescentadas para futuros desenvolvimentos. É esse caráter dinâmico da LA que nos possibilita o aprender e desaprender, uma constante movimentação, construção e desconstrução de conhecimentos pelos quais passamos como pesquisadores e professores.

REFERÊNCIAS BIBLIOGRÁFICAS

- AIJMER, K. Introduction. In: AIJMER, K. (ed.) **Corpora and language teaching**. Amsterdam and Philadelphia: John Benjamins, 2009.
- ALMEIDA FILHO, J. C. A Linguística Aplicada na grande área da linguagem. In: SILVA, K. A. e ALVAREZ, M. L. **Perspectivas de Investigação em Linguística Aplicada**. Campinas: Pontes Editores, 2008.
- ALONSO, M. C. **Corpus linguístico e a aquisição de falsos cognatos em espanhol como língua estrangeira**. Dissertação de Mestrado em Linguística Aplicada. LAEL, PUC, São Paulo, 2006.
- ASTON, G. Enriching the learning environment: corpora in ELT. In: WICHMANN, A. S. et al. **Teaching and language corpora**. London/New York: Longman, 1997.
- BARBOSA, M. **Material didático para o ensino de inglês instrumental online: uma abordagem experiencial baseada em corpus, gênero e tarefa**. Dissertação de Mestrado em Linguística Aplicada. LAEL, PUC, São Paulo, 2004.
- BARLOW, M. **MonoConc Pro**. Houston, TX: Athelstan, 2000.
- BECKET, A. **A Concordance to Shakespeare: suited to all the editions, in which the distinguished and parallel passages in the play of that justly admired writer are methodically arranged. To which are added, three hundred notes and illustrations, entirely new**. Palgrave Macmillan, 1969.
- BERBER SARDINHA, T. **Linguística de Corpus**. São Paulo: Manole, 2004.
- _____. Preparação de material didático para Aprendizagem Baseada em Tarefas com Wordsmith Tools e corpora. **Calidoscópio**, v. 4, n. 3, p. 148-155, 2006. Disponível em: www.unisinos.br/publicacoes_cientificas/images/stories/pdfs_calidoscopio/vol4n3/art03_berber.pdf. Acesso em maio de 2008.
- _____. Corpus linguistics: readings in a widening discipline. **DELTA** [online], v. 23, n. 1, p. 127-137, 2007. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-44502007000100006. Acesso em julho de 2009.
- _____. **Pesquisa em Linguística de Corpus com Wordsmith Tools**. Campinas: Mercado de Letras, 2009.
- _____. Como usar a Linguística de Corpus no ensino de língua estrangeira. Por uma Linguística de Corpus educacional brasileira. In: VIANA, V e TAGNIN, S. **Corpora no ensino de línguas estrangeiras**. São Paulo: Hub Editorial, 2011.

_____ e ALMEIDA, G. M. B. A Linguística de Corpus no Brasil. In: Tagnin, S. E. O., Vale, O. A. (Orgs.). **Avanços da Linguística de Corpus no Brasil**. São Paulo: Humanitas, 2008.

BERNARDINI, S. Corpora in the classroom: an overview and some reflections on future developments. In: SINCLAIR, J. (ed.) **How to use corpora in language teaching**. Amsterdam and Philadelphia: Routledge, 2004.

BÉRTOLI-DUTRA, P. **Explorando a Lingüística de Corpus e letras de música na produção de atividades pedagógicas**. Dissertação de Mestrado em Linguística Aplicada. LAEL, PUC, São Paulo, 2002.

BIBER, D., JOHANSSON, S. LEECH, G. CONRAD, S. e FINNEGAN, E. **Longman Grammar of Spoken and Written English**. London: Longman, 1999.

BISSACO, C. **Ensinando com corpora: Mediação e interação em aulas de espanhol como língua estrangeira**. Tese de Doutorado em Linguística Aplicada. LAEL, PUC, São Paulo, 2010.

BOULTON, A. DDL: Reaching the parts other teaching can't reach? In: FRANKENBURG-GARCIA, A. (ed.) **Proceedings of the 8th Teaching and Language Corpora Conference**. Lisbon, Portugal: Associação de Estudos e de Investigação Científica do ISLA-Lisboa, p. 38-44, 2008. Disponível em: http://hal.archives-ouvertes.fr/docs/00/32/67/06/PDF/2008_boulton_TaLC_reaching.pdf. Acesso em fevereiro de 2009.

BROWN, H. D. **Teaching by principles: an interactive approach to language pedagogy**. NY: Pearson, 2007.

CARTER, R. e McCARTHY, M. **Cambridge Grammar of English: a comprehensive guide. Spoken and written English grammar usage**. Cambridge: Cambridge University Press, 2006.

CELANI, M. A. Transdisciplinaridade na Linguística Aplicada no Brasil. In: SIGNORINI, I. e CAVALCANTI, M. (orgs.). **Linguística Aplicada e Transdisciplinaridade**. Campinas: Mercado de Letras, 1998.

_____, DEYES, F. e HOLMES, J. L. **ESP in Brazil - 25 Years of Evolution and Reflection**. São Paulo / Campinas: EDUC, Mercado de Letras, 2005.

CHAMBERS, A. Integrating Corpus Consultation in Language Studies. **Language Learning and Technology**, v. 9, n. 2, p. 111–125, 2005.

_____. What is data-driven learning? In: MCARTHY, M. e O'KEEFE, A. (eds.) **The Routledge handbook of Corpus Linguistics**. New York: Routledge, 2010.

CHENG, W., Warren, M. and Xu, X. The Language Learner as Language Researcher: Corpus Linguistics on the Timetable, **System**, v. 31, n. 2, p. 173–186, 2003.

CHENG, W. What can a corpus tell us about language teaching? In: MCARTHY, M. e O'KEEFE, A. (eds.) **The Routledge handbook of Corpus Linguistics**. London e New York: Routledge, 2010.

CONDI DE SOUZA, R. **Dois corpora, uma tarefa: O percurso de coleta, análise e utilização de corpora eletrônicos na elaboração de uma tarefa para ensino de inglês como Língua Estrangeira**. Dissertação de Mestrado em Linguística Aplicada. LAEL, PUC, São Paulo, 2005.

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

COOK, G. The uses of reality: a reply to Ronald Carter. **ELT Journal**, Oxford: Oxford University Press, v. 52, n. 1, p. 57-63, 1998.

CRYSTAL, D. **English as a Global Language** (2 ed.). Cambridge: Cambridge University Press, 2003.

EGGINS, S. **An Introduction to Systemic Functional Linguistics** (2 ed.) London: Continuum International Publishing Group, 2004.

ELLIS, R. **Classroom and Second Language Development**. Oxford: Oxford University Press, 1984.

ELLIS, N. Frequency Effects in Language Processing. A Review with Implications for Theories of Implicit and Explicit Language Acquisition. **Studies in Second Language Acquisition**, v. 24, p.143–188, 2002.

_____. **Task-based language teaching and learning**. Oxford: Oxford University Press, 2003.

EVISON, J. What are the basics of analysing a corpus? In: MCARTHY, M. e O'KEEFE, A. (eds.) **The Routledge handbook of Corpus Linguistics**. London e New York: Routledge, 2010.

FERRARI, J. **ESP, Linguística de Corpus e Sócio-Interacionismo na elaboração de uma unidade de um material didático para comércio exterior**. Dissertação de Mestrado em Linguística Aplicada. LAEL, PUC, São Paulo, 2004.

FERREIRA, E. **Palavra freqüente, pronúncia diferente: a lingüística de corpus auxiliando o ensino da pronúncia do inglês como língua estrangeira**. Dissertação de Mestrado em Linguística Aplicada. LAEL, PUC, São Paulo, 2006.

FIRTH, J. R. **Papers in linguistics – 1934-1951**. Oxford: Oxford University Press, 1957.

FLOWERDEW, L. Corpus-based Analyses in EAP. In FLOWERDEW, J. (ed.) **Academic Discourse**. London: Pearson, 2002.

GAVIOLI, L. e ASTON, G. Enriching Reality: Language Corpora in Language Pedagogy. **ELT Journal**, v.55, n. 3, p. 238–246, 2001.

GARDNER, R. e W. LAMBERT. **Attitudes and Motivation in Second Language Learning**. Rowley, Massachusetts: Newbury House, 1972.

GILQUIN, G e GRANGER, S. How can data-driven learning be used in language teaching? In MCARTHY, M. e O'KEEFE, A. (eds.) **The Routledge handbook of Corpus Linguistics**. London e New York: Routledge, 2010.

GRANGER, S. The International Corpus of Learner English: A New Resource for Foreign Language Learning and Teaching and Second language Acquisition Research, **TESOL Quaterly**, v. 37, n. 3, p. 538-546, 2003.

HALLIDAY, M.A.K. **An Introduction to Functional Grammar** (2 ed.). London: Edward Arnold, 1994.

HALLIDAY, M.A.K. **Complementarities in language**. Beijing: Commercial Press, 2008.

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

HUTCHINSON T. e WATERS. A. **English for Specific Purposes**. Cambridge University Press, 1987.

JACOBI, C. **Linguística de Corpus e ensino de espanhol a brasileiros: descrição de padrões e preparação de atividades didáticas (decir/hablar ; mismo; mientras /en cuanto/ aunque)**. Dissertação de Mestrado em Linguística Aplicada. LAEL, PUC, São Paulo, 2001.

JOHNS, T. Micro-Concord: a language learner's research tool. **System**, v. 14, n. 2, p. 151-162, 1986.

_____. Whence and Whither Classroom Concordancing? In: BONGAERTS, T. et al. (eds.) **Computer Applications in Language Learning**. Dordrecht: Foris, 1988.

_____. Should you be persuaded: two examples of data-driven learning. **ELT Journal**. Birmingham: University of Birmingham v. 4, p. 1-16, 1991.

KENNEDY, G. D. **An introduction to corpus linguistics**. New York: Longman, 1998.

LEECH, G. Corpora and theories of linguistic performance. In: Svartvik, J. (ed.) **Directions in Corpus Linguistics: Proceedings of 283 Nobel Symposium 82, Stockholm, 4-8 August 1991**. Berlin: Mouton de Gruyter, 1992.

LEECH, G. Teaching and language corpora: a convergence. In: WICHMANN, A. S. et al. **Teaching and language corpora**. London/New York: Longman, 1997.

_____. New resources, or just better old ones? The Holy Grail of representativeness. In: HUNDT, M., N. NESSELHAUF, N. e BIEWER, C. (eds.) **Corpus Linguistics and the Web**. Amsterdam: Rodopi, p. 133-149, 2007.

LEWIS, M. **The Lexical Approach. The State of ELT and a Way Forward**. London: LTP, 1993.

_____. **Implementing the lexical approach: putting theory into practice**. Hove, Language Teaching Publications, 1997.

_____. Corpora. In: TEUBERT, W. e KRISHNAMURTHY, R. **Corpus Linguistics: critical concepts in Linguistics**. New York: Routledge, 2007.

MAURANEN, A. Introdução. In: MAURANEN, A. e RANTA, E. (eds.) **English as a Lingua Franca: Studies and findings**. Cambridge Scholars Publishing, 2009.

McCARTHY, M., McCARTEN, J. e SANDIFORD, H. **Touchstone 1 Teacher's Book**. Cambridge: Cambridge University Press, 2005.

McARTHY, M. e O'KEEFE, A. Historical perspective: What are corpora and how have they evolved? In: MCARTHY, M. e O'KEEFE, A. (eds.) **The Routledge handbook of Corpus Linguistics**. London e New York: Routledge, 2010.

McENERY, T., XIAO, R. e TONO, Y. **Corpus-based language studies: an advanced resource book**. London: Routledge, 2006.

McENERY, T. e WILSON, A. Teaching and language corpora. **ReCall**, v. 9, n. 1, p. 5-14, 1997.

_____. **Corpus Linguistics**. Edinburgh: Edinburgh University Press, 2001.

MISHAN, F. **Designing Authenticity into Language Learning Materials**. Bristol: Intellect Books, 2005.

MOITA LOPES, L. P. **Oficina de Lingüística aplicada – A natureza social e educacional dos processos de ensino-aprendizagem de línguas**. Campinas: Mercado de Letras, 2002.

_____. (org.) **Por uma Lingüística Aplicada Indisciplinar** (2 ed.). São Paulo: Parábola Editorial, 2008.

MOREIRA FILHO, J. **Desenvolvimento de um software para preparação de aulas de inglês com corpora**. Dissertação de Mestrado em Linguística Aplicada. LAEL, PUC, São Paulo, 2007.

O'KEEFE, A., McCARTHY, M. e CARTER, R. **From corpus to classroom**. Cambridge: Cambridge University Press, 2007.

O'SULLIVAN, Í. Enhancing a Process-Oriented Approach to Literacy and Language Learning: The Role of Corpus Consultation Literacy. **ReCALL**, v. 19, n. 3, p. 269–86, 2007.

_____. e CHAMBERS, A. Learners' Writing Skills in French: Corpus Consultation and Learner Evaluation. **Journal of Second Language Writing**, v. 15, p. 49–68, 2006.

OWEN, C. Corpus-based grammar and the Heineken effect: lexico-grammatical description for language learners. **Applied Linguistics**. Oxford: Oxford University Press, v. 50, n. 3, p. 219-224, 1993.

PALMER, F. R. (Ed.) **Selected Papers of J.R. Firth 1952-1959**. Longman, 1968.

PARTINGTON, A. **Patterns and meanings: using corpora for English language research and teaching**. Amsterdam: John Benjamins, 1998.

PINTO, M. **O uso de things, thing, anything, something e everything em corpora de aprendiz**. Dissertação de Mestrado em Linguística Aplicada. LAEL, PUC, São Paulo, 2008.

PRABHU, N. S. **Second Language pedagogy**. Oxford: Oxford University Press, 1987.

QUIRK, R. et al. **A Comprehensive Grammar of the English Language**. New York: Longman, 1985.

REDSTON, C. e CUNNINGHAM, G. **Face2Face Elementary Student's Book**. Cambridge: Cambridge University Press, 2005.

REPPEN, R. Building a corpus: what are the key considerations? In: MCARTHY, M. e O'KEEFE, A. (eds.) **The Routledge handbook of Corpus Linguistics**. New York: Routledge, 2010.

RÖMER, U. Pedagogical applications of corpora: Some reflections on the current scope and a wish list for future developments. **Zeitschrift für Anglistik und Amerikanistik**, v. 54, n. 2, p. 121-134, 2006.

_____. Corpus research and practice: what help do teachers need and what can we offer? In AIJMER, K. (ed.) **Corpora and language teaching**. Amsterdam and Philadelphia: John Benjamins, 2009.

_____. Pedagogical Applications of Corpora: Some Reflections on the Current Scope and a Wish List for Future Developments. **Zeitschrift für Anglistik und Amerikanistik**, v. 54, n. 2, p. 121-134, 2006.

SCOTT, M. e JOHNS, T. **MicroConcord**. Oxford: Oxford University Press, 1993.

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

_____. Aprendizagem direcionada por dados: uma homenagem a Tim Johns (1993 – 2009). In: VIANA, V e TAGNIN, S. **Corpora no ensino de línguas estrangeiras**. São Paulo: Hub Editorial, 2011.

SEIDLHOFER, B. Key concepts in ELT: English as a Lingua Franca. **ELT Journal**. Oxford University Press, v. 59, p. 339-341, 2005.

SINCLAIR, J. et al. **Cobuild English Dictionary**. London and Birmingham: Collins COBLUID, 1987.

SINCLAIR, J. (ed.) **Collins Cobuild English grammar**. London: Collins, 1990.

_____. **Corpus, Concordance, Collocation**. Oxford: Oxford University Press, 1991.

_____. Corpus evidence in language description. In: WICHMANN, A. S. et al. **Teaching and language corpora**. London/New York: Longman, 1997.

_____. **Reading Concordances**. London: Pearson, 2003.

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

_____. Introduction. In: SINCLAIR, J. (ed.) **How to use corpora in language teaching**. Amsterdam and Philadelphia: Routledge, 2004b.

_____ e RENOUF, A. A lexical syllabus for language learning. In CARTER, R. e McCARTHY (Eds.) **Vocabulary and Language Teaching**. London: Longman, 1988.

STEVENS, V. Concordance-based Vocabulary Exercises: A Viable Alternative to Gap-fillers. In JOHNS, T. e KING, P. (eds.) **Classroom Concordancing: English Language Research Journal 4**. Birmingham: Centre for English Language Studies, University of Birmingham, p. 47–63, 1991.

STUBBS, M. **Text and corpus analysis: computer-assisted studies of language and culture**. Oxford: Blackwell, 1996.

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

_____ e SINCLAIR, J. Corpora. In: BROWN, K. (ed.) **Encyclopedia of Language and Linguistics** (2 ed.). Amsterdam: Elsevier, p. 206–19, 2006.

TRIBBLE, C. Corpora and Corpus Analysis: New Windows on Academic Writing. In: FLOWERDEW, J. (ed.) **Academic Discourse**. London: Longman, 2002.

_____ e JONES, G. **Concordances in the classroom: a resource book for teachers**. Harlow: Longman, 1990.

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

_____. Theoretical overview of the evolution of Corpus Linguistics. In: MCARTHY, M. e O'KEEFE, A. (eds.) **The Routledge handbook of Corpus Linguistics**. New York: Routledge, 2010.

VICENTINI, G. **A linguística de corpus e o seriado Friends como base para o ensino de chunks em sala de aula de língua inglesa**. Dissertação de Mestrado em Linguística Aplicada. LAEL, PUC, São Paulo, 2006.

WIDDOWSON, H.G. The description and prescription of language. In: ALATIS, J. (org.): **Georgetown University Round Table on Languages and Linguistics**. Washington, D.C.: Georgetown University Press, 1991.

_____. On the limitations of linguistics applied. **Applied Linguistics**, v. 21, n. 1, p. 3-25, 2000.

WILLIS, D. **The Lexical Syllabus**. London: Collins ELT, 1990.

_____ e WILLIS, J. **Doing Task-based Teaching**. Oxford: Oxford University Press, 2007.

WILLIS, J. Perspectives on task-based instruction: understanding our practices, acknowledging different practitioners. In: LEAVER, B. L. e WILLIS, J. **Task-based instruction in foreign language education: practices and programs**. Washington, DC: Georgetown University Press, 2004.

ANEXOS

ANEXO 1: Linhas de concordância de *ad spending*

1 expects Asia to show at least 5.9% growth in ad spending this year and in Latin America b
 2 better than advertising. U.S. measured-media ad spending last year plunged 12.3%, accordi
 3 government-sanctioned bankruptcy. GM slashed ad spending last year, but it remained the a
 4 increase their share of voice as they step up ad spending this year, Sanford C. Bernstein
 5 ,600 a year ago. Meanwhile, growth in overall ad spending among wireless operators has dec
 6 the Middle East market, and its double-digit ad spending growth was forecast to continue
 7 nch has them thinking about how to parcel out ad spending among various audiences -- a not
 8 s recent ruling that will allow for increased ad spending by unions, special interests and
 9 move eliminated some \$300 million in regional ad spending for uniform messages, though mos
 10 f choppiness. PricewaterhouseCoopers projects ad spending on total U.S. TV will grow to \$8
 11 of issue-ad spending in 2009. The 2009 issue-ad spending related to federal issues, such
 12 son to a year when the company slashed global ad spending 13% to \$7.6 billion. Too much sk
 13 timedia in October cut its forecast for world ad spending growth to an anemic 0.5% from a
 14 tor to see a drop in first-half 2010 measured ad spending, according to WPP's Kantar Media
 15 n issue that generates millions of dollars in ad spending in these competitive federal rac
 16 its rates and dwelling as a distant third in ad spending in the category. When rival Alls
 17 y, it is unlikely that we'll see any lulls in ad spending. Further, the nation is layered
 18 Forecasts for 2011 suggest moderate growth in ad spending, reflecting the economy's slow r
 19 ly without incumbents running for reelection. Ad spending on the state level will be drive
 20 ding. Overall, China represents 3.4% of total ad spending for the Global 100, slightly bel
 21 me clarity to one of the most opaque areas of ad spending, and the lifeblood of many Ameri
 22 t. And while P&G added \$1 billion back to its ad spending for the fiscal year ended June 3
 23 son. It also banked \$18.4 million in measured-ad spending during the first 11 months of 20
 24 bout two-thirds of the retailer's increase in ad spending for the fiscal third quarter and
 25 layers have also maintained or even increased ad spending in the last year, unlike others
 26 l companies concentrate the bulk of their own ad spending in the U.S. Eli Lilly & Co., Mer
 27 players such as Kellogg Co., which increased ad spending but still saw sales fall 5%, lar
 28 Business NANCHANG, China (AdAge.com) -- With ad spending growth still in the double dig
 29 ars in lump sums work in digital media, where ad spending is still ruled by quarter-to-qua
 30 ories Supreme Court Ruling Will Put Political Ad Spending 'on Steroids' *Corporate Speech*,
 31 ctor's billions in marketing spending. Online ad spending for small and medium-sized U.S.
 32 But they're not panicking and, in fact, some ad spending forecasts show the markets havin
 33 year of the recession. It's no surprise that ad spending is severely depressed: U.S. meas
 34 sibling Automotive News. Auto sales and auto ad spending have seen gains in 2010, though
 35 Kantar Media's Campaign Media Analysis Group, ad spending this season will top \$3 billion.

36 mpany profiles, agencies, brands, executives, ad spending and sales for the nation's 100 t
 37 by 7.8%. The crucial question is when global ad spending will return to the 5% or 6% grow
 38 2015, up from \$34 million in 2009. Including ad spending with promotions, events and rese
 39 To Catch Up, Colgate May Ratchet Up Its Ad Spending Analysts Predict Up to \$200M Boos
 40 the estimated U.S. share of overall worldwide ad spending. ZenithOptimedia, part of Public
 41 been flocking to the medium with enthusiasm. Ad spending on overall TV increased 10.5% in
 42 e than double when it comes to auto insurance ad spending. Progressive outspends it by nea
 43 d by radio in 2011. Marketers can't hold back ad spending forever, or they run the risk of
 44 Gulf of Mexico, it unleashed \$100 million in ad spending, largely on network TV, to stem
 45 suality of all the cost-cutting and mergers is ad spending -- especially on TV -- while the
 46 Ad Spending Is on the Rise, but Growth Rate M
 47 arketers and retailers as the cause. Local-TV ad spending was off 2.8% in 2008, per TNS. O
 48 uly's forecast when ZenithOptimedia projected ad-spending growth of 1.6% next year. And it
 49 that, contrary to the pro-digital pundits, TV ad spending will not be shifting dramaticall
 50 inting unions to lower costs; more government ad spending with newspapers, both in print a
 51 ce a 26%, or \$108 million, increase in global ad spending to more than \$500 million. U.S.-
 52 orecast. Fueling that growth -- and a jump in ad spending of more than 20% in the first 10
 53 tural experts expect the 2010 census to boost ad spending toward minority groups as awaren
 54 as tumbled 42% since its 1981 peak. Cigarette ad spending -- \$51 million in 2006, accordin
 55 week. Unilever last week said it had boosted ad spending \$500 million during the first ha
 56 Allstate offer. SHELLING OUT: Auto insurers' ad spending vs. market share. The campaign -
 57 ream probably offers some cause for optimism. Ad spending on local English-language TV ove
 58 media market in flux, shrinking revenue, weak ad spending and little upside offered by a t
 59 planning more modest single-digit growth for ad spending in line with sales growth this y
 60 ets, too. Zenith Optimedia boosted its global ad-spending forecast last week for the secon
 61 was overtaken by retail as the second-biggest ad spending category last year, according to
 62 100 roster, 12 firms do more than half their ad spending abroad, largely tracking with wh
 63 relative bright spots last year. Personal-care ad spending slipped just 1.6%, displacing au
 64 forecasts a staggering 9.9% drop in worldwide ad spending, said Jonathan Barnard, ZenithOp
 65 es right now is 'bang for the buck.' " As for ad spending, it is the third and fourth quar
 66 mes a report from ZenithOptimedia that global ad spending should increase 0.5% in 2010. Me
 67 and spending is up 119% in 2009, a year when ad spending overall has suffered double-digi
 68 to hold a party, there's a growing sense that ad spending is haltingly stabilizing -- or a
 69 - just 5% of total worldwide advertising. But ad spending is growing fast. Publicis Groupe
 70 ings from Ad Age DataCenter reports Worldwide ad spending will grow 5.3% in the new year,
 71 ckelodeon grossed \$833.87 million in measured ad spending for full-year 2009, down 10.7% f
 72 accounted for 47% of second-quarter internet ad spending, up from 44% a year ago, accordi
 73 advertising and a shrinking market. Cigarette ad spending has fallen 95% from its 1985 pea
 74 in Fazal As brands continue to increase their ad spending on Facebook, there is an ongoing

75 g on here? Can advertisers pare back their TV ad-spending allocation even as TV wins more
 76 the network a mid-tier player in terms of net ad spending, it's enough to make the network
 77 nds Amid Greek Economic Crisis With World Cup Ad Spending Way Down, Region Regroups
 78 anufacturers both intend to increase in-store ad spending by double digits, despite the re
 79 in operating margin in the form of increased ad spending). He didn't specify the total. B
 80 Meager, yes, but after so much sputtering of ad spending and cost cutting at media compan
 81 rom 1997 to 2007, the COI quickly doubled its ad spending. However, Conservative leader Da
 82 ional sales that they do more than half their ad spending abroad. Coca-Cola Co. allocates
 83 spent in 2000, according to industry-reported ad spending figures cited by J.D. Power & As
 84 t, user and business development in Shanghai. Ad spending in China's online video category
 85 e as the main determinant of consumer choice. Ad spending trends certainly support this co
 86 100 LNA including top spenders by medium and ad spending by category. Links to Ad Age's L
 87 ntiate its product from competitors'. Delta's ad spending fell by nearly 50% to \$32.9 mill
 88 doesn't provide much incentive to stick with ad-spending plans. But big marketers like Pr
 89 aft/Cadbury Mix Could Look Like A Peek at the Ad Spending and Brand Power the Combinatio
 90 s heightened spending there, too. P&G boosted ad spending \$1 billion for the year ended Ju
 91 g Burger King over a proposed 25% increase in ad spending, to be funded through soft-drink
 92 ed. A year ago, forecasters figured 2010 U.S. ad spending would be flat or down a bit. For
 93 Spending Across Categories Back on the Rise, Ad-Spending Forecasts Are Also Up That's the
 94 Spending Across Categories Back on the Rise, Ad-Spending Forecasts Are Also Up BATAVIA,
 95 teresting and expensive primaries fueling the ad spending. Adding an even greater signific
 96 eer sales -- and the overwhelming majority of ad spending -- and all five suffered through
 97 at China accounted for 4.1% of 2008 worldwide ad spending. Ad Age's 23rd annual Global Mar
 98 gration with "X Factor," while PepsiCo boosts ad spending 30% across beverage brands. If P
 99 and personal-care companies he covers to hike ad spending as a share of sales by more than
 100 Either way, such factors will lead to record ad-spending in a number of states. What to w
 101 to Kantar Media. Meanwhile, Nick at Nite saw ad spending rise a whopping 20% during the s
 102 Age.com) -- Marketers in 2011 will boost U.S. ad spending 2.8%, down slightly from 2010's
 103 ABMiller.) For the Global 100, total measured ad spending tumbled 8.7% to \$107.2 billion i
 104 Three Reasons Not to Panic When Online Ad Spending Drops 5.4% Why You Should be Op
 105 ing to Ad Age's analysis. First-half measured ad spending rose 5.7%, with gains in every s
 106 ng is severely depressed: U.S. measured-media ad spending fell 4.1% in 2008 and then plung
 107 account for only 1% of 2010's total political ad spending of \$4.2 billion. It's not uncomm
 108 U.S. For the Global 100 group, total measured ad spending rose 3.1% to \$117.9 billion in 2
 109 e-time hours, and saw a 9.2% bump in measured ad spending in 2010, according to Kantar. Di
 110 Associates projects that overall new-vehicle ad spending in 2010 will rise to \$19.2 billi
 111 Borrell estimates the total online political ad spending for 2010 will reach \$44.5 millio
 112 0 Could Be Record-Breaking Year for Political Ad Spending The 2010 elections will be viewe
 113 ll slice of its budget. McDonald's total U.S. ad spending for 2009 was \$873.3 million, com

114 her than the \$300 million-plus worth of issue-ad spending in 2009. The 2009 issue-ad spend
115 take: Cartoon banked \$230 million in measured ad spending during 2009, down 1% from 2008,
116 2009, and it grossed \$415 million in measured ad spending in 2009, according to Kantar, a
117 but Disney XD grossed \$82 million in measured ad spending in 2009, down 1.1% from Toon Dis
118 consumer brands: only \$11 million on measured ad spending in 2009 and \$29 million in the f
119 akers' marketing dollars. Chrysler cut its TV ad spending in 2009 by half -- with the othe
120 st cerebral in a way." Ford actually cut U.S. ad spending in 2009 as it consolidated its f
121 8, \$6.4 million). Fox News' overall measured ad spending in 2009 was down 0.7% from 2008,
122 e. The Cleveland Clinic more than doubled its ad spending from 2008 to 2009, according to
123 m is so steadfast that it cut its traditional ad spending from 2008 and 2009 more than 50%
124 e No. 5 national advertiser in the U.S., with ad spending of \$2.3 billion, up 14.1% from 2
125 e No. 5 national advertiser in the U.S., with ad spending of \$2.3 billion, up 14.1% from 2
126 recent report that Colgate will need to boost ad spending by \$150 million to \$200 million
127 nths of 2010, Apple boosted its measured U.S. ad spending by 13.6%, according to Kantar Me
128 elcome change from the recent past: 2009 U.S. ad spending tumbled 11.9% (average of three
129 is pretty fully embraced by now: U.S. search-ad spending reached \$10.7 billion in 2008, a

ANEXO 2: Linhas de concordância de *ad network*

1 billion Developers: More than 120,000 Mobile Ad Network: Quattro Wireless Price tag: \$275
 2 Times Co.'s digital ad revenue grew 14.6% and ad network ValueClick saw a slight increase
 3 ilder has what it takes to build a proprietary ad network on par with Amazon. At the moment
 4 the advertising business by acquiring mobile-ad network Quattro Wireless. Though the deal
 5 it gets divided like this: The agency (\$.75), ad network (\$2), data provider (\$0.75), ad e
 6 engine, WHERE's desktop site, mobile apps and ad network, and PayPal) looks to be moving t
 7 lay-type problem is relatively easy. Ask your ad network -- or the sites you're purchasing
 8 ssment the situation has brought for both the ad network and the global conservation netwo
 9 l as cable operators like Charter to build an ad network for TV. Actually, it's an audienc
 10 terms, Amazon is using brand data to build an ad network that delivers relevant advertisin
 11 m those publishers' core properties, but from ad network and long-tail content partners as
 12 e internet," said Richard Jalichandra, CEO of ad network Technorati Media. "And with all t
 13 iciency drive the day, it's not clear various ad-network maneuvers will be enough. "What's
 14 re and JumpTV, while brands that come through ad network serving include Citibank, Ford an
 15 re and JumpTV, while brands that come through ad network serving include Citibank, Ford an
 16 h Mr. Brien now running the company's largest ad network, many say it only makes sense tha
 17 using the site. Evernote recently dumped its ad network and started curating its own ads,
 18 a situation where the ad-serving entity -- an ad network, a DSP, or any of the other compa
 19 O of video ad server FreeWheel. So far, video ad network Brightroll is the biggest recipie
 20 outbid the computer maker on its first-choice ad network, AdMob, there was talk that Apple
 21 and whistles; and paying a fledgling Twitter ad network to spread the word among its army
 22 as Campaign Cornerstone 'Hashtags,' Fledgling Ad Network Help Get the Word out Across Micr
 23 Winners, While BBDO Remains the Most-Awarded Ad Network Posted by Kunur Patel on 09.2
 24 Winners, While BBDO Remains the Most-Awarded Ad Network Posted by Kunur Patel on 09.2
 25 e is that while one party is in control of an ad network, an exchange operates on behalf o
 26 ke Fyall, product marketing manager of mobile ad network AdMob. "Layering device targeting
 27 argely to keyword and display ads on Google's ad network, a download button on YouTube and
 28 le to remnant display-ad inventory sold on an ad network. Think news video, casual gaming
 29 for lots of cheap impressions is an order an ad network like 24/7 Real Media would be lik
 30 h them? People who block third-party cookies (ad-network cookies) are still reached, you j
 31 deployed across Google's third-party display ad network. The effort is similar to Microso
 32 and network," said Mike Cassidy, president of ad network Undertone Networks. "It's a reall
 33 buying aQuantive and with it ad server Atlas, ad network Drive PM and Avenue A/Razorfish.
 34 s," according to a Google spokeswoman. Mobile Ad Network: AdMob Price tag: \$750 million in
 35 ictim of an attack and initially suspected an ad network was responsible. It turned out th
 36 million in stock on the leader in the mobile ad-network space, AdMob, and got its claws i
 37 ns in mobile, 4INFO could be in play. The SMS ad network has the reach, technology platfor

38 said Alec Andronikov, head of Movox, the SMS ad network behind the campaign.
39 se I am one myself; my blog is on the BlogHer ad network), none of these networks is ideal
40 er," said Henry Copeland, founder of the blog ad network Blogads.com. (Claim to fame? They
41 million in stock on the leader in the mobile ad-network space, AdMob, and got its claws i
42 n-win-win scenario. If you take this retailer/ad network model and combine manufacturer da
43 re impressions were passed back forward to an ad network, then passed back if unsold. So w
44 of the business from premium inventory to its ad network, exchange and newspaper consortiu
45 e apps than iPhone users, according to mobile ad network AdMob, and 10% of the ad requests
46 adget. Only scant reference was made to AOL's ad network and technologies -- the onetime P
47 well, with players like MTV's Tribes, Travel Ad Network, Complex Media and others try to
48 : March 30, 2009 What's the point of using an ad network? Aren't they just about cheap ads
49 ne," said Marcus Startzel, senior VP sales at ad network Millennial Media. Sixty percent o

ANEXO 3: Linhas de concordância de *ad networks*

1 a boon for Google (and has given birth to 400 ad networks), and represents the best thinki
 2 and Google, data providers like BlueKai, and ad networks like Audience Science currently
 3 reative and go wide with video on portals and ad networks. "I definitely think that CPGs h
 4 among some of their peers, both publisher and ad networks, who feared the buy -- probably
 5 to becoming exchanges in their own right. And ad networks -- at least those of the 400 or
 6 there. Many of these vendors, publishers and ad networks are working together on initiati
 7 plicative negotiations between publishers and ad networks. Networks' sales forces were sti
 8 0s he founded one of the first ad servers and ad networks at Real Media; later he founded
 9 se of ad exchanges, additional parties can be ad networks, which commonly use them to buy
 10 uyers to compare and gauge real scale between ad networks, DSPs and others. This is a prev
 11 op-quality journalists and one of the biggest ad networks in the world to pay for everythi
 12 my overall QPS volume share from RTB buyers (ad networks, DSPs, agency trading desks)? Ho
 13 nits and rush-to-the-bottom pricing driven by ad networks and exchanges, online display ha
 14 itical mass will enable CrowdGather to bypass ad networks altogether. Fair enough, but per
 15 at because of the scope of its data. Consumer ad networks or vertical-related networks are
 16 ut CEO The latter two deals are about display ad networks attempting to add video capabili
 17 irst phase of growth ("Network 1.0"), display ad networks focused almost exclusively on ta
 18 display volume (QPS) being purchased by DSPs, ad networks and agency trading desks on a RT
 19 Ad exchanges were largely invented to enable ad networks to more efficiently and cheaply
 20 entioned in the space -- and more-established ad networks are starting to offer dynamic cr
 21 under pressure from ad networks and Facebook. Ad networks have allowed garbage inventory f
 22 e money out of a still-weak market. Watch for ad networks eying publisher optimizers, whic
 23 in a down advertising market. Initially, four ad networks -- all venture-funded startups -
 24 s of now, the forums subsist on revenues from ad networks and Google AdSense, yielding CPM
 25 You can buy audience-targeting inventory from ad networks, exchanges and digital-service p
 26 the advertising model is under pressure from ad networks and Facebook. Ad networks have a
 27 that plague most multi-billion-dollar global ad networks in an era of rapid change, but a
 28 Digital Marketing Guide: Ad Networks Why Targeting, Pricing and the Li
 29 which some niche retailers simply don't have. Ad networks are built on the promise of scal
 30 ost agencies are still buying from individual ad networks that aggregate inventory in the
 31 en the case for several years, even as larger ad networks continue to make significant rev
 32 next. Large publishers are looking more like ad networks. Publisher-side brokers are movi
 33 ompete better on those fronts. There are many ad networks that include newspapers and offe
 34 e many supply channels: many publishers, many ad networks and now several ad exchanges. To
 35 Cunningham, CEO of Appsavvy. "A lot of mobile ad networks have completely dropped the ball
 36 Display Publishers That Resemble Ad Networks, Ad Networks Focused on Brand Safety and Br
 37 r the last decade, with literally hundreds of ad networks, data companies, yield managers,

38 ." Sales-channel mix The traditional rules of ad networks -- aggregate massive amounts of
39 EXCHANGES Today, advertisers use a myriad of ad networks and exchanges to buy demogra
40 to be sold directly or optimize any number of ad networks vying for the impression. That's
41 f thousands of websites (and even hundreds of ad networks) -- do you want to spend all day
42 n their agencies and vendors, but a survey of ad networks and publishers reveals unanimity
43 ying publisher optimizers, which have cut off ad networks from publishers. Perhaps no one
44 advance what sites the buy will be placed on. Ad networks often use audience targeting to
45 t's shave off 10% of our TV budget and run on ad networks!" said Chris Allen, VP-video in
46 istribute malware. "They have been focused on ad networks for a while and they've tightene
47 luable to marketers. Some people blame online ad networks for dragging down online adverti
48 e-mail domain. Gawker hasn't used third-party ad networks for several years, which is why
49 brand. This rule also applies to third-party ad networks, which are paid by advertisers a
50 ble coupons, barcodes and augmented reality. AD NETWORKS & EXCHANGES Today, adver
51 n of Digital Display Publishers That Resemble Ad Networks, Ad Networks Focused on Brand Sa
52 nline advertising prices, but Mr. Morgan said ad networks unfairly took the fall: The real
53 ting, but without fighting another 30 similar ad networks with newspaper inventory in the
54 ry puts content owners and probably even some ad networks out of business. Looking up For
55 that money is coming mainly from two sources: ad networks and Google's AdSense. But in the
56 wouldn't discover the malicious code. Target: ad networks The attack reflects a growing so
57 t's because it is. It's the same formula that ad networks promised on the web -- take low-
58 pages. What complicates "impressions" is that ad networks and demand side platforms (DSPs)
59 "pretty impressive." "It's a reflection that ad networks are beginning to rapidly shift o
60 te blogs of all niches and sizes. Until then, ad networks are the best option.
61 direct sales force. Our ads are sold through ad networks and rep firms, MTV Networks, Gor
62 e and runs a portion of its inventory through ad networks. Sabah Karimi, a contributor to
63 mechanics of digital media. At the same time, ad networks persisted as a sales channel for
64 ignedificant portion of direct remnant sales to ad networks. With the leadership of the agen
65 y true when it comes to brands (as opposed to ad networks, DSPs, and data companies). Whil
66 cant money into mobile media support with two ad networks -- Quattro and Jump-Tap -- as we
67 to digital-media-buying platforms that usurp ad networks and portals. It's a big pool and
68 ll. Meanwhile, the trend toward more vertical ad networks, many of which create their own
69 first quarter, monetizing almost entirely via ad networks and Google AdSense. Are the kitt
70 demanding such a low CPM rate that even video ad networks took a pass, let alone publisher
71 ted, but also the intent of the viewer. Video ad networks, paid distribution services and
72 tes like CBS and ABC to the portals and video ad networks and is planning to spend signifi
73 to have every intention of growing it. Video ad networks Tremor Media, Yume, BBE and Brig
74 ne as far as discontinuing relationships with ad networks. But Walker Jacobs, senior VP-di
75 r, Gawker or CBS, to stop doing business with ad networks and to focus on brand dollars. I
76 without having to maintain relationships with ad networks." Web Browser: Safari Browser Ma

ANEXO 4: Linhas de concordância de *ad revenue*

1 spring, it's more a production tactic than an ad-revenue strategy. "We had a drop-off in a
2 sed the car to boost their own followings and ad revenue from YouTube. YouTube pro Jill Ha
3 ata To wit: in the most recent quarter, AOL's ad revenue declined 27% from the same period
4 e better than CBS Corp. CEO Les Moonves. Auto ad revenue in the tank? Les will tell you it
5 's Attempt to Spruce Up Its Sites Washes Away Ad Revenue Ad Impressions Drop, but So Do Pa
6 -month period, and Comcast's collective cable ad revenue could eclipse the broadcast piece
7 ter of 2010. Last year's take: Disney Channel ad revenue is not tracked by Kantar Media, b
8 rs instead. BAD NEWS U.S. newspapers' digital ad revenue is not keeping pace with digital
9 revenue 17%, The New York Times Co.'s digital ad revenue grew 14.6% and ad network ValueC
10 ant. Slow growth Newspapers' share of digital ad revenue has fallen from 16.2% in 2005 to
11 g Web Woes Pummel Newspapers Slipping Digital Ad Revenue, Emerging Content Farms Prese
12 o ComScore. Mr. Millstein declined to discuss ad revenue but said he expects the site to t
13 her. And they're all dipping into the display-ad revenue stream. In a not-atypical scenari
14 al businesses. Third-quarter domestic-display ad revenue is down 8% year over year, but up
15 hey also do still have comparatively enormous ad revenue even in their depleted states. Th
16 than half, according to ITZBelden estimates; ad revenue would hold up in that case, becau
17 this point, between 60% and 70% of Facebook's ad revenue comes from small businesses. Such
18 a show, with the idea that the proceeds from ad revenue would go to charity. Originally,
19 luding advancing startup funds against future ad revenue and even offering \$1,000 coupons
20 y-Backed Venture Will Use YouTube to Generate Ad Revenue by Michael Learmonth Published
21 g accounted for 14% of Time Inc.'s first-half ad revenue. In 2007, before the recession, o
22 d contrast the two. AD REVENUE American Idol: Ad revenue for the 2009 season slipped to \$8
23 and Bravo. Add that to the \$585.7 million in ad revenue logged by Comcast Entertainment's
24 All told, the program generated \$6 million in ad revenue with the biggest chunk -- some \$2
25 by VentureBeat, was making \$25,000 a month in ad revenue. While there was no media spendin
26 nt. The company earned close to \$1 billion in ad revenue in 2008, compared to an estimated
27 dia to finish 2010 with an overall decline in ad revenue. "I don't know that the positive
28 lity and are anxious to keep \$60.5 billion in ad revenue from going the way of Philco. A h
29 8. That sector has outperformed network TV in ad-revenue growth. Cable networks also get t
30 ear ended June 2010, with a modest 3% rise in ad revenue but a whopping 19% jump in fees f
31 he network has taken in about \$3.1 billion in ad revenue during the first seven months of
32 . Blip.tv's top earners pull down \$500,000 in ad revenue a year based on 50/50 revenue sha
33 dvertising accounted for only 7% of Time Inc. ad revenue. The newspaper industry continues
34 the bulk of the traditional radio industry's ad revenue. "Digital can be a challenge to l
35 publisher Time Inc., said first-half magazine ad revenue rose 5%, driven by a \$26 million
36 to displace SMS and display as the top mobile ad revenue driver, according to the Kelsey G
37 driver, according to the Kelsey Group. Mobile ad revenue will more than quadruple from nex

38 to displace SMS and display as the top mobile ad revenue driver, according to the Kelsey G
39 a revenue difference," he said of his mobile ad revenue. Now, generally ad-free isn't the
40 driver, according to the Kelsey Group. Mobile ad revenue will more than quadruple from nex
41 ter than her regular videos, which means more ad revenue from YouTube, which is placing ad
42 rm to drive a better user experience and more ad revenue?" Google, as a rule, is about bui
43 king for a show that could send some of NBC's ad revenue over to Time Warner's TBS. Mr. Bu
44 . Mr. Rittenberg contends the network lost no ad revenue as a result of clients pulling th
45 n executives believe TV is the sole driver of ad revenue, that could leave rivals eager to
46 pproach would put 12% of page views and 9% of ad revenue at risk, but Mr. Brill said he do
47 obile applications; soak up new categories of ad revenue, including obituaries; and pursue
48 is less susceptible to the cyclical swings of ad revenue. Consider: News Corp. said revenu
49 4, according to SNL Kagan estimates. Based on ad revenue, E! is on par with NBCU's Bravo,
50 that they need to reduce their dependence on ad revenue, and -- as Bloomberg well knows -
51 foot the bill. Media that relies primarily on ad revenue is suffering far more than those
52 Denver news. Newspapers also have real online ad revenue to protect, hauling in \$653.1 mil
53 ass than half what the daily required. Online ad revenue is down some 20%, which the Monit
54 ution is more complicated. They risk not only ad revenue but affiliate revenue from cable
55 all the other broadcast networks, saw overall ad revenue drop in recession-plagued 2009 by
56 the big mismatch between actual and potential ad revenue. According to Comscore, social ne
57 s in Videos and News Stories Can Help Predict Ad Revenue, Optimize Web Content NEW YOR
58 ting to decay. But it could also help predict ad revenue for online videos or pieces of co
59 onitor attributes to the recession, but print ad revenue has actually held steady, the Mon
60 that in its fourth quarter of 2009, it reaped ad revenue of \$114.9 million, up 32% from \$8
61 luding mobile) passed newspapers in 2010 U.S. ad revenue, making the internet the second-l
62 n ad revenue. Comcast on top Time Warner said ad revenue fell 15% in the first half of 200
63 ery, Scripps, Cablevision and Time Warner saw ad revenue increase 13% to 21%, while those
64 ly features Bleacher writers. The site splits ad revenue. An insider says Bleacher will be
65 revenues, for example, are faring better than ad revenue. Comcast on top Time Warner said
66 of the ad dollars spent online. Not that the ad revenue is anything to sneeze at. Nielsen
67 ESPN donated the block of time and agreed the ad revenue would be donated to charity. All
68 8 million. What gives? Mr. Armstrong said the ad-revenue drop was largely "self-inflicted,
69 MSNBC, gain more viewer attention -- and the ad revenue that comes with it. Fox News capt
70 works and similar companies get most of their ad revenue from commercials that appear alon
71 satellite and telecommunications-provider TV-ad revenue. Entities ranging from Cablevisio
72 Age decided to compare and contrast the two. AD REVENUE American Idol: Ad revenue for th
73 E!'s cachet is rising among advertisers, with ad revenue expected to surpass the \$200 mill
74 ou risk 12% of your page views and 9% of your ad revenue, but Journalism Online Co-founder

ANEXO 5: Linhas de concordância de *ad campaign*

1 is to reach 4,500 by the end of the year. An ad campaign by Sancho BBDO highlights that t
 2 l limited rollout, whipping up frenzy with an ad campaign featuring fans in loving embrace
 3 athan Fox somewhat, but he largely credits an ad campaign from Publicis Worldwide, New Yor
 4 Forces of Colombia, better known as FARC. An ad campaign by Bogota-based Lowe SSP3 is h
 5 Forces of Colombia, better known as FARC. An ad campaign by Bogota-based Lowe SSP3 is h
 6 is a blasphemy that takes it too far. And an ad campaign that uses rapper Eminem to perso
 7 alk about what it's like to have a role in an ad campaign for the first time. While it's n
 8 is a blasphemy that takes it too far. And an ad campaign that uses rapper Eminem to perso
 9 top 10, it's "Write the Future" for Nike, an ad campaign that leveraged Nike's biggest st
 10 ally in January and is being promoted with an ad campaign by WPP's Y&R with a "Heaven on E
 11 seeking comment. So yes, "What is 230?" is an ad campaign. It's for a GM product. But that
 12 ad of the Cash for Clunkers' program, with an ad campaign saying it was already offering t
 13 ear, hiring independent agency Anomaly for an ad campaign called "Grab Some Buds," which s
 14 etting a religious figure down on paper in an ad campaign, substitute Jesus or, better yet
 15 the Syracuse University football team did an ad campaign in New York City in an attempt t
 16 act and does allow calls while navigating. An ad campaign, created in-house, includes TV a
 17 ked by a just-launched umbrella marketing and ad campaign from WPP's Y&R that includes sev
 18 e directors named Little Caesars as the "best ad campaign of 1994." Maybe you remember the
 19 te its hard-hitting "Save money. Live better" ad campaign and recent ads claiming Walmart
 20 te its hard-hitting "Save money. Live better" ad campaign and recent ads claiming Walmart
 21 . Bulgarelli said. Said Mr. Fay of the coming ad campaign: "We need to make an emotional c
 22 ake that Pepsi Max has launched a comparative ad campaign. After all, its rival in Atlanta
 23 mission needed a redo. An expensive consumer ad campaign around the launch mocked the cel
 24 675 employees. Chrysler breaks the "New Day" ad campaign from BBDO. The spots use third-p
 25 675 employees. Chrysler breaks the "New Day" ad campaign from BBDO. The spots use third-p
 26 timately hagiographic "Imported From Detroit" ad campaign. "The company has invested great
 27 timately hagiographic "Imported From Detroit" ad campaign. "The company has invested great
 28 tal difference is attributable to the display-ad campaign. The effort itself is meant to r
 29 ony. "When we launched our 'Panel of Experts' ad campaign this year with Justin Timberlake
 30 gn bubble," the idea of working on one finite ad campaign after another. Further to hiring
 31 gns. But as Bill Bernbach once said, "A great ad campaign will make a bad product fail fas
 32 will make a bad product fail faster." A great ad campaign will also make a bad marketing s
 33 nfortunate it would be if the Axe fell on his ad campaign at point of contact. Not only wo
 34 d be featured as the new stars of Pizza Hut's ad campaign, with Martin Agency's Andy Azula
 35 ons. The launch day of her Viva Glam lipstick ad campaign alone generated nearly 20 millio
 36 nched last week with an estimated \$20 million ad campaign, including a TV commercial with
 37 Wahl is stoking the trend with a \$10 million ad campaign launched earlier this year. THE

38 MOBILE AD CAMPAIGN Earlier this year, the New York K
39 Motorola's rise and fall 2004: "Hello, Moto" ad campaign launches. JANUARY 2004: Former S
40 its "Get Every Dollar You Deserve" multimedia ad campaign. Like H&R Block, it employs tax
41 s they ordered for possible use in a national ad campaign. Winners received \$500. That eve
42 the recession, earning itself a new national ad campaign from Publicis' Saatchi & Saatchi
43 , CMO of the clinic, which started a national ad campaign earlier this year from Adworks,
44 r a sweet deal in which LBL will create a new ad campaign to stimulate sales at 50% off th
45 ed unscathed" after Mt. Lutz criticized a new ad campaign that she defended, said Chris Ha
46 ut. That's because they're the stars of a new ad campaign that broke during Fox's NFL broa
47 d, of course, to the software giant's ongoing ad campaign from agency Crispin Porter & Bog
48 t in getting out that information through our ad campaign, through our public relations ou
49 e second time the Bureau has developed a paid ad campaign, and the last one was successful
50 cation behind its ubiquitous "I am a Phoenix" ad campaign. Initiative is Phoenix's local m
51 plant peanuts this year." The cross-platform ad campaign by the Atlanta shop Lawler Balla
52 han two dozen other unions have begun a print-ad campaign saying they will oppose any legi
53 the need for less noise was clear. Samsung's ad campaign, breaking Sept. 30, touts its fr
54 cognized the need for an organized, strategic ad campaign to drive the ongoing demobilizat
55 cognized the need for an organized, strategic ad campaign to drive the ongoing demobilizat
56 s important to continue the effort beyond the ad campaign. Fidelity's Mr. Speros advises c
57 for the just released product PRO MERIS, the ad campaign pitches the anti flea treatment
58 ting local villages with the Afghan army. The ad campaign has three themes, each backed by
59 nattended or taken down after the life of the ad campaign. Still, it offered a lot of work
60 r. Kelly attributed some of the growth to the ad campaign. "We experienced record traffic
61 Although Smartbox has yet to do a traditional ad campaign, it has pushed out an aggressive
62 -home software option.* A general national TV ad campaign serves as "air cover" and comple
63 jifilm, has also been first with a widespread ad campaign, buying up 3-D movie screens thi

ANEXO 6: Linhas de concordância de *ad campaigns*

1 ttnacht, raved about the success of his radio ad campaigns. We've been using radio to real
 2 ot to: Hotels, motels and resorts are running ad campaigns with either drastically discoun
 3 one another for both search and display based ad campaigns Regardless of how we got here,
 4 comprehensive program of education (including ad campaigns targeting businesses and consum
 5 regionally and we used to run 25 different big ad campaigns. By being inclusive with the re
 6 text messages from my college-age kids about ad campaigns -- just recently they texted ab
 7 In the course of 2010, Kraft has launched new ad campaigns for 15 of its "power brands," i
 8 en suggested for public relations and digital ad campaigns); fee plus bonus (where retaine
 9 e new standard of the world," and forthcoming ad campaigns for the brand will center on a
 10 11.16.09 NEW YORK (AdAge.com) -- Not a lot of ad campaigns target armed guerrillas who liv
 11 the past tumultuous year with some successful ad campaigns. In fact, the Fiat brand won Go
 12 rom Travelocity, whose best-performing online ad campaigns came after the No. 2 online tra
 13 ustomizable Unlike traditional online display ad campaigns bound by a beginning and end da
 14 s Sony, Samsung and Vizio, which have started ad campaigns around LED sets for the season
 15 . Both banks have featured mobile services in ad campaigns. Bank of America's Doug Brown,
 16 ng, pricing strategies, retail promotions and ad campaigns to reach consumers effectively
 17 bill, which appeared headed for passage, with ad campaigns on several fronts. Notably, it
 18 ation -- have launched or are about to launch ad campaigns, the biggest of which is an eff
 19 mber 2009 when it launched one of the boldest ad campaigns the restaurant industry has see
 20 ng. Ad Age: What about traditional marketing? Ad campaigns and direct mail and all of that
 21 The result was efficiently deployed terrible ad campaigns the world over. That's what you
 22 immediate freeze on spending, declaring that ad campaigns are not priorities, and that it
 23 ual community around new product launches and ad campaigns. They're very good at starting
 24 ne are the days of hiding marketing ideas and ad campaigns from employees for fear of info
 25 for advice and ideas on marketing, previewing ad campaigns with them before the ads go pub
 26 errillas NEW YORK (AdAge.com) -- Not a lot of ad campaigns target armed guerrillas who liv
 27 odby Silverstein & Partners demand that their ad campaigns contain not just a PR component
 28 Where's Betty?" ? Ad Age: You've done several ad campaigns in recent years, including 1-80
 29 ed products start to hit shelves and roll out ad campaigns. "With these new restrictions i
 30 Gamers so enjoy value-exchange ads that most ad campaigns are reposted on Facebook by gam
 31 11:16 AM Not content to let its international ad campaigns make fun of Mexicans or degrade
 32 h new storytelling techniques and high-impact ad campaigns like the great ads for Apple. B