

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

BRUNA RODRIGUES DA SILVA

**Ensino de vocabulário no Ensino de Jovens e Adultos (EJA) –
atividades a partir do *corpus* do jornal Diário Gaúcho**

Porto Alegre

2011

BRUNA RODRIGUES DA SILVA

**Ensino de vocabulário no Ensino de Jovens e Adultos (EJA) –
atividades a partir do *corpus* do jornal Diário Gaúcho**

Trabalho de Conclusão de Curso – TCC –
apresentado ao Curso de Letras da
Universidade Federal do Rio Grande do Sul
– UFRGS, como requisito parcial para a
obtenção do título de Licenciado em Letras –
Português e Literatura Portuguesa.

Orientação: Prof. Dr. Valdir do Nascimento
Flores
Co-orientação: Prof^ª. Dr^ª. Maria José
Bocorny Finatto

Porto Alegre

2011

BRUNA RODRIGUES DA SILVA

**Ensino de vocabulário no Ensino de Jovens e Adultos (EJA) –
atividades a partir do *corpus* do jornal Diário Gaúcho**

Trabalho de Conclusão apresentado à Universidade Federal do Rio Grande do Sul – UFRGS como requisito parcial para obtenção do título de Licenciado em Letras – Português e Literatura Portuguesa.

Aprovado em _____

BANCA EXAMINADORA

Professor Orientador: Valdir do Nascimento Flores

Professora: Jane da Costa Naujorks

Doutorando UFRGS: Leonardo Zilio

AGRADECIMENTOS

Em primeiro lugar, agradeço aos meus pais – Alcides e Tania – pelo amor incondicional. Agradeço a eles e também à minha irmã Bianca por apoiarem e incentivarem meus estudos, confiando em minhas escolhas e torcendo sempre pela minha felicidade.

À minha família, especialmente minha tia e minhas avós, agradeço pelo carinho, pela força e pela torcida em todos os momentos de minha trajetória.

Aos meus padrinhos – Gilberto, Luisa e Camila – agradeço pela confiança depositada desde sempre.

Aos meus queridos Greice e Rafael, agradeço por instigar o gosto pelos livros, pela leitura, pelo ensino, pelos alunos e, sobretudo, pela educação, despertando meu interesse pela licenciatura, me mostrando que o mundo precisa de bons professores e, principalmente, me fazendo acreditar que isso é possível.

Aos meus amigos, agradeço pelo companheirismo e pela amizade, sem os quais não seria possível a conclusão desta etapa.

Aos professores que tive ao longo de minha caminhada, agradeço não só pelo conhecimento, mas também por me ensinarem o que é ser um bom professor – ou, em alguns casos, o que não é.

Ao professor Valdir, agradeço pela disponibilidade de orientação deste trabalho.

Por fim e, com certeza, muito importante, agradeço à professora e amiga Maria José, minha orientadora de monitoria e de pesquisa, meu exemplo de sabedoria, de pesquisadora, de professora, de pessoa. Agradeço pelos ensinamentos, pelo incentivo, pela dedicação e pelas oportunidades dentro e fora da Universidade, sem as quais minha formação – pessoal e profissional – não seria a mesma.

Enfim, agradeço a todos que, de uma forma ou de outra, acreditaram que este sonho fosse possível, torcendo pelo meu sucesso.

Muito Obrigada!

“Ensinar é um exercício de imortalidade. De alguma forma continuamos a viver naqueles cujos olhos aprenderam a ver o mundo pela magia da nossa palavra. O professor, assim, não morre jamais.”

(Rubens Alves)

RESUMO

Este trabalho integra a pesquisa PADRÕES DO PORTUGUÊS POPULAR ESCRITO: O VOCABULÁRIO DO JORNAL DIÁRIO GAÚCHO - FASE 1 e pretende demonstrar como, partindo de um *corpus*, podem ser geradas atividades de ensino de vocabulário para alunos do Ensino de Jovens e Adultos (EJA), mesmo que estes nem sempre tenham acesso a computadores nas escolas. Visa-se apresentar uma proposta de ensino de léxico pautada na observação do uso real das palavras, oriunda da Lingüística de Corpus (LC), partindo de um tipo de texto jornalístico que, embora popular entre os estudantes, é pouco explorado como material didático: o jornal popular porto-alegrense Diário Gaúcho. Primeiramente, caracterizam-se a pesquisa, o jornal e o *corpus* reunido, em parte disponível gratuitamente na internet. Depois, apresentam-se atividades já realizadas com alunos da graduação em Letras da UFRGS, buscando adaptação para o público do EJA. Na sua versão adaptada, vê-se que o primeiro elemento para que o professor do EJA consiga aproveitar o material é que conheça noções sobre LC, *corpora* e diferentes padrões de frequência de palavras. Em seguida, explica-se como é possível desenhar uma unidade de ensino de vocabulário aproveitando o jornal popular e a visão de *corpus*. Por fim, avalia-se como o professor poderá, tratando de vocabulário e produzindo atividades de sala de aula associadas à manipulação do jornal impresso, estimular os alunos a terem contato com a internet, realizando atividades *on-line* com o *corpus* DG, por meio das ferramentas de exploração deste disponíveis no ambiente virtual de aprendizagem PorPopular.

Palavras-chave: Ensino de Jovens e Adultos (EJA). Vocabulário. Jornal Popular. Lingüística de Corpus (LC). *Corpus/Corpora*.

ABSTRACT

This paper integrates the research THE PATTERNS OF POPULAR WRITTEN PORTUGUESE LANGUAGE: THE VOCABULARY OF DIÁRIO GAÚCHO NEWSPAPER - Phase 1 and intends to demonstrate how we can create vocabulary activities to students of the Youngs and Adults Teaching system starting from a *corpus* even though they are not used to access computers at school. It aims to show a learning proposal of the lexicon based on the observation of the real usage of the words, which comes from the Corpus Linguistics, starting from a kind of journalistic text that is unexplored as a didactic material, but very common among the students, the popular newspaper: Diário Gaúcho (DG). Firstly, the research presents the newspaper and the organized *corpus*, which is available in part for free on the internet. After that, it shows the activities that have already been performed with the students from UFRGS Language graduation course, seeking for an adaptation to Youngs and Adults Teaching system students. In its adapted version, we can see that the first element to Youngs and Adults Teaching system teachers take advantage of this material is to know some notions about Corpus Linguistics, *corpora* and different patterns of the frequency of words. Then it explains how it is possible to draw a unit of vocabulary learning taking advantage of the popular newspaper and the vision of *corpus*. At the end it estimates how teachers, dealing with vocabulary and producing activities associated with the manipulation of the printed newspaper in the classroom, can stimulate students to keep in touch with internet executing online exercises with *corpus* DG, through its exploration tools available in the virtual environment of PorPopular learning.

KEY WORDS: Youngs and Adults Teaching system. Vocabulary. Popular Newspaper. Corpus Linguistics. *Corpus/Corpora*.

LISTA DE ILUSTRAÇÕES

Figura 01 – Lista de palavras da página 14 do dia 03 de março de 2008 do jornal Diário Gaúcho.....	15
Figura 02 - Contraste entre as 10 palavras mais freqüentes: DG (março/2008) - Banco do Português (recorte jornalístico).....	16
Figura 03 - Contraste entre as 20 palavras mais freqüentes: DG (março/2008) - ZH (março/2008).....	16
Figura 04 - Glossário esportivo da turma de Léxico e Dicionários – UFRGS 2009.....	22
Figura 05 - 50 palavras mais freqüentes no <i>corpus</i> de 6 meses do DG.....	46
Figura 06 - Tela inicial do <i>site</i> PorPopular.....	54
Figura 07 - Tela da ferramenta geradora de listas de palavras do <i>site</i> PorPopular.....	55
Figura 08 - Tela da ferramenta geradora de contextos do <i>site</i> PorPopular.....	56
Figura 09 - Tela da ferramenta geradora de <i>n-gramas</i> do <i>site</i> do PorPopular.....	57

SUMÁRIO

1 INTRODUÇÃO	10
2 ANTECEDENTES	14
2.1 HISTÓRICO DA PESQUISA POR POPULAR.....	14
2.1.1 Sobre jornalismo popular e sobre o jornal Diário Gaúcho	18
2.1.2 Sobre o <i>corpus</i> organizado	19
2.1.3 Experiência com estudantes de Letras da UFRGS	21
2.2 ENCAMINHAMENTO DO TRABALHO.....	23
3 NOÇÕES FUNDAMENTAIS E REVISÃO DE BIBLIOGRAFIA	24
3.1 PALAVRA.....	24
3.2 LÉXICO – VOCABULÁRIO.....	27
3.3 LINGÜÍSTICA DE CORPUS.....	29
3.3.1 Princípios da Lingüística de Corpus	29
3.3.2 <i>Corpus/Corpora</i>	31
3.3.3 Estatística Lexical	32
3.3.4 Padrões de frequência de palavras	34
3.4 PRÓXIMOS PASSOS.....	36
4 CENÁRIOS DE ENSINO/APRENDIZAGEM ENVOLVIDOS	37
4.1 EJA - ENSINO PARA JOVENS E ADULTOS.....	37
4.1.1 Alunos e Professores em perfil	39
4.2 O ENSINO DE VOCABULÁRIO.....	40
4.2.1 O espaço do ensino de vocabulário nas aulas de Língua Portuguesa	41
4.2.2 Enfoques do ensino de léxico no EJA – opções e princípios envolvidos	43
5 EXPLORAÇÃO DO <i>CORPUS</i> E PROPOSTAS DE ENSINO	45
5.1 PADRÕES DO LÉXICO NO DG: DESCRIÇÃO.....	45

5.2 IDÉIAS PARA O ENSINO DE VOCABULÁRIO NO EJA.....	47
5.2.1 Unidade de ensino de vocabulário.....	48
5.2.2 Idéias de atividades para sala de aula.....	49
5.2.3 Opções para exploração do <i>corpus</i> DG em ambientes digitais.....	53
6 CONSIDERAÇÕES FINAIS.....	59
REFERÊNCIAS.....	62
ANEXO A – Reportagem <i>Efetivo Reduzido Investiga Crime</i> , do jornal <i>Zero Hora</i> (09/01/2009).....	66
ANEXO B – Reportagem <i>Menos policiais no Caso Becker</i> , do jornal <i>Diário Gaúcho</i> (09/01/2009).....	67
ANEXO C – Capa do jornal <i>Diarinho – Diário do Litoral</i> (19/01/2009).....	68
ANEXO D - Manchete da capa do jornal <i>Meia Hora</i> (30/10/2008).....	69
ANEXO E - Exemplos de exercícios de ensino de vocabulário.....	70
ANEXO F - Exemplos de exercícios de ensino de vocabulário(2).....	71
ANEXO G - Exemplos de exercícios de ensino de vocabulário(3).....	72
ANEXO H – Reportagem <i>Morte em micro-ônibus</i> , do jornal <i>Diário Gaúcho</i> (16/06/2010).....	73
ANEXO I - Perguntas de interpretação.....	74
ANEXO J – Reportagem <i>Dois homens executam jovem dentro de ônibus e fogem a pé na Região Metropolitana</i> , do jornal <i>Zero Hora</i> (16/06/2010).....	75
ANEXO L - <i>Wordlist</i> texto <i>Morte em micro-ônibus</i> (DG 16/06/2010).....	76
ANEXO M - Palavras mais freqüentes na Língua Portuguesa segundo o Banco do Português.....	77
ANEXO N - Parte da <i>wordlist</i> do dia 14/01/2008 do <i>site</i> PorPopular.....	78
ANEXO O - Parte dos contextos da palavra ‘ <i>de</i> ’ do dia 14/01/08 do <i>site</i> PorPopular.....	79
ANEXO P - Parte dos contextos da palavra ‘ <i>é</i> ’ do dia 14/01/08 do <i>site</i> PorPopular.....	80
ANEXO Q – Parte dos <i>tri-gramas</i> do dia 14/01/2008 do <i>site</i> PorPopular.....	81

1 INTRODUÇÃO

O ensino de vocabulário desempenha papel importante no ensino de línguas. No caso da língua estrangeira, o método de ensino de vocabulário mais comum é a apresentação de imensas listas de palavras com a respectiva tradução em determinada língua ao lado. Ao questionar seu professor sobre o significado de alguma palavra na leitura de textos ou na resolução de exercícios, o aluno é instigado a procurá-lo em glossários e dicionários bilíngües, atribuindo normalmente uma definição específica para cada vocábulo, geralmente associada apenas a uma tradução literal e descontextualizada, como se o sentido de uma palavra fosse algo único.

No caso da língua materna, o mais comum nas escolas é o ensino de vocabulário por associação, em que o aluno é incentivado a associar o significado de palavras desconhecidas - que aparecem, na maioria das vezes, em frases soltas - com imagens ou com palavras semelhantes na sua forma (raiz, prefixo, sufixo, etc.). Os exercícios de vocabulário mais freqüentes nos livros didáticos são os que trazem um modelo como, por exemplo:

Desagradável é aquele que não é agradável.

Bondoso é aquele cheio de bondade.

Após a leitura das frases, o aluno é solicitado a completar as lacunas seguindo o mesmo raciocínio:

Desonesto é _____.

Maldoso é _____.

Assim, conforme o modelo acima, o exercício seria realizado sucessivamente. Além desse tipo de atividade, há o ensino de vocabulário por associação de palavras com outras dessemelhantes (os antônimos), e a presença de glossários ao final dos textos ou dos livros didáticos, assim como no caso da língua estrangeira, também é recorrente.

Esse tipo de prática mostra que o ensino de vocabulário – seja em língua materna, seja em língua estrangeira – tem de ser problematizado, já que o aluno, exposto apenas a atividades desse tipo (ensino de vocabulário por associação entre palavras, geralmente soltas e descontextualizadas), fica privado de relacionar o(s) significado(s) das palavras com o seu contexto de uso e de perceber, por exemplo, os diversos usos, combinatórias, e sentidos possíveis de uma mesma palavra.

É imprescindível ressaltar que não avaliamos tais práticas como equivocadas ou absurdas, pelo contrário, sabemos da importância do uso de dicionários, por exemplo, e inclusive propomos atividades pensando nisso. O que questionamos é o fato de, na maioria das vezes, o ensino de vocabulário ficar centrado e restrito somente nisto: na definição única de uma palavra, sem presença de contexto; e no uso de glossários que já trazem pronto o significado das palavras, sem que o aluno tenha a chance de descobri-lo de outra maneira e, inclusive, de pesquisá-lo em diferentes fontes.

Segundo Paiva (2004, [s.p.]), “As listas de palavras descontextualizadas sempre foram alvo de críticas dos especialistas em ensino de línguas”. Em função disso, neste trabalho de conclusão de curso, pretendemos tratar da temática do ensino de vocabulário à luz dos conceitos da Linguística de Corpus, que acredita que a língua é um sistema no qual uma palavra sempre será definida em função das outras que a acompanham em diferentes situações, observando-se, dessa maneira, um uso real das palavras.

O *corpus* textual utilizado advém de pesquisa de Iniciação Científica, associada ao Projeto PorPopular – Padrões do Português Popular Escrito (http://www6.ufrgs.br/textecc/index_porpopular.php) – que pretendeu reconhecer padrões de vocabulário em textos escritos e, para tanto, buscou explorar o material do jornal popular porto-alegrense Diário Gaúcho (DG). Esse jornal foi lançado no ano 2000 em Porto Alegre e Região Metropolitana e foi escolhido como material de análise da pesquisa PorPopular, entre outros, em função da aceitação por parte do seu público-alvo, pois, segundo Amaral (2006, p.80), “parte dos consumidores do DG não eram leitores de jornal, e após seu lançamento, em 2000, a região metropolitana de Porto Alegre passou a ser a primeira em índice de leitura de jornais no Brasil.”

O público-alvo de nossa proposta de ensino de vocabulário são os alunos do Ensino de Jovens e Adultos (EJA) das escolas públicas. Acreditamos que o estudo

do léxico ainda é pertinente nessa fase, já que “os falantes nativos continuam a expandir seu vocabulário mesmo na fase adulta” (MACHADO, 2003, p.41 *apud* RICHARDS, 1976). Além disso, nossa pesquisa pretende contribuir para a diminuição da lacuna de estudos existentes sobre a educação de jovens e adultos e, principalmente, acerca do desenvolvimento de recursos didáticos para esses alunos.

O foco principal do trabalho, então, está em apontar opções de como tratar da temática do ensino de vocabulário com o aluno do EJA, aproveitando o jornal popular e a visão de *corpus* e de significado de palavras advindas da Linguística de Corpus. O objetivo é contribuir para o alargamento do repertório lexical que os alunos já têm, bem como “para a conscientização do seu uso [do vocabulário] adequado e para a sua correção quanto a aspectos morfológicos, sintáticos e [...] semânticos” (BARBEIRO, [entre 2003 e 2007], p.01).

A idéia é sistematizar dados descritivos do *corpus* do DG já organizado e oferecer subsídios para que os professores aproveitem as ferramentas de exploração (quantidade de palavras, freqüência, contextos, combinações mais recorrentes) dos textos reunidos, disponíveis gratuitamente na internet com seus alunos, além de apresentar atividades possíveis para discutir o tema.

A proposta parte de atividades de ensino do léxico já desenvolvidas com alunos da graduação em Letras da UFRGS, durante nossa experiência no programa de Monitoria em Educação a Distância da UFRGS e que, agora, serão adaptadas para o Ensino de Jovens e Adultos. Tais atividades aconteceram na disciplina Léxico e Dicionários, nos anos de 2009 e 2010, e consistiram na observação de padrões de vocabulário do DG e na análise de freqüências e de usos das palavras, para, a partir disso, organizar-se a base de um pequeno glossário esportivo da turma, idealizando como público-alvo um estudante de português como língua estrangeira. Visto que foram realizadas por alunos da graduação em ambiente de laboratório de Informática do Instituto de Letras da UFRGS, as atividades contaram com o auxílio de ferramentas computacionais e do uso da internet.

O presente estudo apresenta a descrição do *corpus* que será utilizado nas atividades de ensino sugeridas, com o intuito de aproximar o professor do material utilizado, para que ele consiga se integrar às propostas. Além disso, a idéia é adaptar as atividades já realizadas com os alunos da graduação e criar uma unidade de ensino de vocabulário para alunos do EJA, levando em conta uma ainda inegável dificuldade de recursos de infra-estrutura e de informatização das escolas públicas,

especialmente no turno da noite. Ficam, então, como desafio da nossa proposta, duas perguntas: a) Até que ponto é possível planejar certas atividades de sala de aula sem o auxílio de tecnologia? B) Atividades baseadas em *corpus* são possíveis sem o uso direto, pelo aluno, de ferramentas computacionais?

Buscando responder tais questionamentos, no primeiro capítulo fazemos uma apresentação da pesquisa, destacando o jornal analisado, os dados do *corpus* organizado e quais atividades baseadas nesse *corpus* já foram realizadas. No segundo capítulo fazemos uma revisão da bibliografia da área, apresentando os principais conceitos e perspectivas teórico-metodológicas envolvidos, para, no terceiro capítulo, focarmos o Ensino de Jovens e Adultos, o perfil de alunos e de professores e a questão do espaço do ensino de vocabulário nas aulas de Língua Portuguesa. Por fim, no quarto capítulo apontamos atividades para o ensino de vocabulário baseado em *corpora*, tanto na versão impressa - tradicional de sala de aula -, quanto na versão digital – por meio do ambiente virtual de aprendizagem PorPopular.

2 ANTECEDENTES

Conforme já mencionado, a idéia deste trabalho de conclusão de curso surgiu ao longo da minha atuação como bolsista de Iniciação Científica. Como, para situar a proposta deste trabalho, é importante ter conhecimento dessa experiência de IC, a seguir estão sintetizados os seus principais pontos.

2.1 HISTÓRICO DA PESQUISA PORPOPULAR

A pesquisa PorPopular – cujo título oficial, junto ao CNPq, é PADRÕES DO PORTUGUÊS POPULAR ESCRITO: O VOCABULÁRIO DO JORNAL *DIÁRIO GAÚCHO* - FASE 1 – teve início em novembro de 2009 e deve ser concluída até novembro de 2011. Seu objetivo é obter uma descrição sobre padrões do vocabulário e da linguagem como um todo exibidos por textos de jornais populares de grande circulação e tiragem. Esses textos são feitos, em tese, de modo mais simplificado para serem compreendidos com facilidade por públicos de menor poder aquisitivo e que tenham níveis de escolaridade não muito elevados. A descrição e os estudos do Projeto são feitos à luz de referenciais teórico-metodológicos da Linguística de Corpus (BERBER SARDINHA, 2004a), havendo uma tendência bastante forte para o tratamento estatístico do vocabulário dos textos.

O *corpus* de estudo inicial foi uma amostra seriada de edições diárias dos 12 meses do ano de 2008 do jornal popular porto-alegrense Diário Gaúcho. Mais adiante, esse *corpus* acabou sendo complementado por um semestre de amostra do ano de 2009. O material textual do jornal foi cedido à coordenadora da pesquisa, com a devida autorização, emitida pelo editor-chefe do jornal, para armazenamento, compartilhamento e publicação *on-line* na versão somente texto.

A primeira etapa do trabalho consistiu na produção e na organização do *corpus* do Diário Gaúcho em formato somente texto, a partir dos arquivos originais em formato PDF. Isso foi necessário porque nesse formato de arquivo é possível utilizar sistemas computacionais especialmente desenvolvidos para a realização de

diversas estatísticas lexicais e para a obtenção de listagens de palavras de diferentes perfis.

Com o auxílio desses sistemas e das ferramentas disponíveis no *site* do grupo TEXTQUIM (<http://www6.ufrgs.br/textquim/index.php>) foram observadas as palavras mais freqüentes, mês a mês, em amostras que incluíram, a cada mês, dez diferentes dias da semana. Quando pronta, a lista de freqüências (*wordlist*) trouxe informações sobre a quantidade de palavras (*tokens*) e a quantidade de palavras diferentes (*types*) presentes nos textos. Isso é o que se pode observar na Figura a seguir.

Lista de palavras		4	faça	2	capricórnio	2	pressagius
Tokens: 690		4	suas	2	virgem	2	intermédio
Types: 380		3	se	2	câncer	2	dieta
21	de	3	proteção	2	leão	2	professor
19	que	3	conseguir	2	sagitário	2	www
15	e	3	seu	2	sonho	2	sucesso
15	o	3	ano	2	situações	1	precisa
13	trabalho	3	neste	2	significa	1	firme
13	a	3	ao	2	aquário	1	decidido
12	para	3	vontade	2	áries	1	touro
12	saúde	3	santa	2	ficar	1	ser
12	amor	3	já	2	nasci	1	contra
12	com	3	resposta	2	gostaria	1	caminho
11	bom	3	fazer	2	favorável	1	só
10	uma	2	emprego	2	terá	1	salmão
10	do	2	cada	2	pacote	1	percorrido
7	os	2	vela	2	saber	1	fique
7	um	2	todas	2	nos	1	tarefas
7	as	2	este	2	gastos	1	atento
7	em	2	h	2	combinações	1	mal-estar
7	nathanael	2	das	2	procure	1	peixes
7	dia	2	agradável	2	números	1	roxa
5	boa	2	pessoas	2	azul	1	gordurosos
5	por	2	velas	2	vida	1	invista
5	muito	2	planos	2	graça	1	quem
5	mais	2	imagem	2	idéias	1	atividades
5	da	2	amigos	2	força	1	exerce
4	anjo	2	não	2	grande	1	no
4	br	2	gêmeos	2	têm	1	setor

Figura 01: Lista de palavras da página 14 do dia 03 de março de 2008 do jornal Diário Gaúcho.

A partir da *wordlist*, foi feita a comparação do *corpus* DG com dados de padrões de vocabulário colhidos do Banco do Português, um *corpus* de linguagem geral do Brasil (BERBER SARDINHA, 2004a, p.150), que possui mais de 120 milhões de palavras (tomando como base o ano 2000). Esse *corpus* é mantido na PUC-SP e é aberto, ou seja, seu conteúdo está em constante renovação. No contraste feito, percebeu-se que as palavras mais freqüentes (de, a, o, e, que) se repetem em ambos os *corpora*, variando apenas a posição, como mostra a Figura a seguir:

	Diário Gaúcho (mar/2008) Total de palavras = 160869 Total de palavras diferentes = 18127		Banco do Português (recorte jornal) Total de palavras = 223.371.280 Total de palavras diferentes = 582.372	
1	7135	de	6022939	de
2	5136	a	4289463	a
3	5081	o	4135372	o
4	4018	e	2906593	e
5	2859	que	2763756	que
6	2671	do	2433919	do
7	2337	da	2169947	da
8	1998	com	1760984	em
9	1951	para	1403295	para
10	1942	em	123337	no

Figura 02: Contraste entre as 10 palavras mais freqüentes: DG (março/2008) - Banco do Português (recorte jornalístico).

Além disso, foram feitas também observações e contrastes com padrões de vocabulário do jornal *Zero Hora* (ZH), publicado pela mesma empresa do DG e dirigido a um público de maior poder aquisitivo. Esse contraste foi realizado em três etapas: primeiro, a partir das listas de palavras de cada *corpus*, em seguida, com as combinações de palavras mais freqüentes (*n-gramas*) em cada um e, por fim, com textos sobre um mesmo assunto publicados nos dois jornais.

O paralelo entre as *wordlists* mostrou que, assim como na comparação com o Banco do Português, não houve diferença significativa entre as palavras mais freqüentes nos dois *corpora* (de, a, o, e, que), conforme exemplifica a Figura abaixo:

	Diário Gaúcho (mar/2008) Total de palavras = 160869 Total de palavras diferentes = 18127		Zero Hora (mar/2008) Total de palavras = 557258 Total de palavras diferentes = 32881	
1	7135	de	27087	de
2	5136	a	18506	a
3	5081	o	16747	o
4	4018	e	12678	e
5	2859	que	11145	do
6	2671	do	9970	que
7	2337	da	8945	da
8	1998	com	8080	em
9	1951	para	6455	com
10	1942	em	6417	para

Figura 03: Contraste entre as 20 palavras mais freqüentes: DG (março/2008) - ZH (março/2008).

Já o contraste dos *n-gramas* revelou forma de escrita e interesses diferenciados dos jornais, uma vez que as combinações mais freqüentes eram

diferentes nos *corpora* analisados: data de nascimento, local de nascimento e ser taxista é, no DG; e de acordo com, fim de semana e Caxias do Sul, no ZH. As combinações que se repetiram foram apenas as que independem do tipo de texto/leitor, como por exemplo: Porto Alegre, Rio Grande do Sul, das h min, o que demonstrou diferença não só de formato dos jornais, mas também de assuntos abordados e, conseqüentemente, de público-alvo.

Ao comparar as notícias publicadas simultaneamente nos dois jornais também não houve grandes diferenças, nem nas palavras e nem na quantidade delas. O que se pode observar foi que o texto de ZH esclareceu mais o fato em foco, trazendo alguma informação a mais, afinal, seus leitores são, teoricamente, diferenciados em relação aos leitores do DG, o que é perceptível inclusive em parte do título dos textos em questão: efetivo reduzido (ZH – ANEXO A) x menos policiais (DG - ANEXO B).

O trabalho de contraste com o jornal Zero Hora foi feito por meio da parceria entre a equipe PorPopular, da qual sempre fiz parte, e os investigadores de Processamento da Linguagem Natural (PLN) do Núcleo Interinstitucional de Lingüística Computacional da USP (NILC-USP). O NILC-USP (<http://www.nilc.icmc.usp.br/nilc/index.html>) desenvolve pesquisas sobre padrões de texto e de vocabulário em diferentes frentes de estudo, cabendo salientar o projeto PorSimples¹, que estuda a simplificação de textos com vistas a atender portadores de dificuldades de leitura (MARGARIDO *et al*, 2009 *apud* FINATTO, 2009).

Tendo feito essa breve síntese da pesquisa PorPopular, cabe ressaltar que uma parte do *corpus* DG, parcialmente organizado, já está oferecida *on-line* no site do projeto (<http://www6.ufrgs.br/textecc/experimente.php>.) em formato somente texto, devidamente identificado, para ser utilizado por pesquisadores interessados. A pesquisa de IC, naturalmente, foi um recorte de uma pesquisa mais ampla, que segue sendo desenvolvida.

¹ Disponível em <<http://caravelas.icmc.usp.br/wiki/index.php/Principal>>

2.1.1 Sobre o jornalismo popular e sobre o jornal Diário Gaúcho

O Jornalismo produz conhecimentos e textos sobre o cotidiano. Segundo Amaral (2006), a partir do ano 2000 surge um novo conceito de jornal popular, que não se resume mais às matérias policiais e sensacionalistas, mas que busca proximidade e empatia com o público-alvo, geralmente associado com as camadas de menor poder aquisitivo da população. Jornais populares como o Diário Gaúcho, atualmente, procuram dar voz ao seu leitor (preferencialmente das classes B, C e D), que é sua principal fonte de informação. Esse segmento da imprensa tem como diferencial um 'tom' de texto e de escrita mais popular, justamente porque objetiva aproximação com seu público (SILVA; FINATTO, 2009).

O texto do jornal popular é construído por jornalistas, pessoas com nível cultural, econômico e social privilegiados em relação à grande massa da população brasileira. Esses jornalistas precisam, então, interagir com um tipo de leitor cujas condições econômicas e sociais e a bagagem cultural são, em geral, diferentes das suas. Assim, em tese, o cenário da produção do texto do jornal popular é o de um redator de formação universitária que precisa adaptar a feição de sua escrita para que haja a interação desejada entre o veículo, o jornalista e seu público-alvo. Na pesquisa PorPopular, parte-se do pressuposto que o texto desse tipo de jornal integra um uso específico da Língua Portuguesa, denominado de Português Popular Escrito.

A grande maioria das pesquisas em *corpora* sobre vocabulário, sobre neologismos ou sobre outros elementos mórficos ou gramaticais da Língua Portuguesa, feitas no Brasil até hoje, não utilizam materiais desse gênero popular (FINATTO, 2009). Os pesquisadores têm utilizado principalmente materiais oriundos do jornal *Folha de São Paulo* (KAUFMANN, 2008 *apud* FINATTO, 2009), que tem um público de alto poder aquisitivo. Em função dessa lacuna de dados lingüísticos sobre o vocabulário presente na escrita do jornalismo popular, optou-se por estudar esse material, em tese, diferenciado do material jornalístico que usualmente se tem explorado.

Na primeira etapa da pesquisa (fase1), foi explorado apenas um jornal. Entre várias publicações do gênero disponíveis no Brasil, optou-se pelo jornal Diário Gaúcho. Esse jornal foi escolhido em função da aceitação por parte do seu público-

alvo na capital gaúcha e em todo um entorno de municípios vizinhos da região metropolitana. O jornal tem grande tiragem: dados do IVC (Instituto Verificador de Circulação) indicam que cada 01 exemplar tende a ser lido por 05 pessoas em média, fato que redimensiona a relação entre tiragem/venda/número de leitores².

Outro fator digno de nota é que, conforme Amaral (2006, p.80), “parte dos consumidores do DG não eram leitores de jornal, e após seu lançamento, em 2000, a região metropolitana de Porto Alegre passou a ser a primeira em índice de leitura de jornais no Brasil”.

Vale ressaltar, ainda, que o DG foi escolhido também porque, na comparação com veículos similares, pode-se dizer, a princípio, que esse jornal traz uma linguagem séria, simples, direta e acessível. Jornais semelhantes, tais como o *Diarinho* (ANEXO C), de Santa Catarina ou o *Meia Hora* (ANEXO D), do Rio de Janeiro, exibem um texto, via de regra, marcado pelo humor bastante escrachado em todas ou na maioria das suas seções.

2.1.2 Sobre o *corpus* organizado

Um *corpus*, na acepção da Lingüística de Corpus, é uma reunião de textos em formato digital que serve para estudo de uma língua. Usamos, para o estudo do vocabulário do jornal Diário Gaúcho, arquivos de edições diárias completas do ano de 2008, gentilmente cedidos pelo jornal à coordenadora da pesquisa PorPopular.

No que se refere à coleta e seleção de amostra das edições diárias do DG, foi utilizada a mesma sistemática adotada por Ieda Maria Alves no Projeto TermNEO, acessível em <http://www.fflch.usp.br/dlcv/neo>, que corresponde à iniciativa Projeto ‘Observatório de Neologismos Científicos e Técnicos do Português Contemporâneo’ e no seu Projeto ‘Base de Neologismos do Português Brasileiro Contemporâneo’, que utilizam, desde 1993, textos de jornais diários e de revistas semanais como *corpus* (MARONEZE, 2009 *apud* FINATTO, 2009)

² Em março, a circulação paga do DG foi de 163.825 exemplares, e o número de leitores é de 1,5 milhão de pessoas (dados creditados ao IBOPE e IVC, conforme informações publicadas no *site* da empresa RBS em maio de 2011: <http://gruporbs.clicrbs.com.br/blog/2011/05/17/circulacao-de-zh- cresce-pelo-14%C2%BA-mes-em-relacao-ao-mesmo-mes-do-ano-anterior/>)

Resumidamente, a seleção foi procedida da seguinte forma: foram tomados todos os jornais do ano de 2008 e, a cada mês, foram selecionadas as edições completas de 10 a 12 dias, distribuídos nas 4 ou 5 semanas de 1 mês. A intenção foi obter uma amostra composta pelo todo do jornal de cada dia (excluídos apenas informes publicitários, classificados, indicações de expediente e datação) por diferentes dias não consecutivos de cada semana. Como o DG circula de segunda a sábado, em janeiro de 2008, na primeira semana, foram selecionados as edições de segunda, quarta e sexta-feira. Na segunda semana de janeiro, as edições de terça, quinta e sábado e, assim, sucessivamente. Esse mesmo procedimento foi aplicado aos 12 meses do ano de 2008.

Quanto à metodologia de trabalho com esse *corpus*, ocorreram diferentes etapas até que ele ficasse pronto para o uso, em formato somente texto, idêntico ao formato PDF original, que corresponde à sua versão impressa em papel:

- 1) conversão de cada página PDF do DG da edição de determinado dia para o formato TXT (somente texto);
- 2) exclusão de textos publicitários, datações, textos de expediente e outros cuja presença seja fixa todos os dias;
- 3) conferência manual dos arquivos para eliminação de caracteres e de junção de palavras indesejáveis, gerados pela conversão;
- 4) conferência manual dos arquivos para colocar em ordem (conforme o PDF) frases e textos;
- 5) organização dos arquivos com nomenclatura específica;
- 6) junção em arquivo único dos arquivos de cada dia do jornal;
- 7) junção em arquivo único dos arquivos de cada mês do jornal.

Após todas essas etapas, o *corpus* organizado até o momento compreende amostra com edições completas do jornal publicadas ao longo dos primeiros 06 meses do ano de 2008. Cada mês inclui 10 dias de edição, selecionados, conforme já mencionado, de modo a alternar dias da semana. Esse *corpus* tem 974.672 *tokens* (palavras) e 44.456 *types* (palavras diferentes).

Desse material é que foram identificadas as palavras e as construções mais frequentes, como já citado, com apoio das ferramentas geradoras de *wordlist* (lista de palavras) e de *n-gramas* (agrupamentos de palavras) oferecidas no *site* do Projeto TEXTQUIM.

2.1.3 Experiência com estudantes de Letras da UFRGS

A disciplina de Léxico e Dicionários, do curso de graduação em Letras da UFRGS - que atende tanto futuros professores, quanto futuros tradutores -, trata, entre outros assuntos, da exploração do vocabulário em *corpora* textuais para produção de dicionários. Em função disso, a vivência do Projeto PorPopular foi levada para a sala de aula de Léxico e Dicionários, com o intuito de mostrar a importância e a presença do texto jornalístico como *corpus* na produção de dicionários e nos estudos sobre vocabulário de escrita.

A primeira parte da atividade de ensino, associada à minha experiência de monitoria nessa disciplina, envolveu apresentar aos alunos os passos de obtenção dos dados sobre o vocabulário presente nos textos. Isso aconteceu por meio da manipulação de arquivos em formato PDF e TXT, pelo exame das listagens de palavras e pelo reconhecimento dos diferentes padrões de frequências, feitos pelos próprios alunos. Eles geraram as suas listas de palavras, com o apoio das ferramentas do *síte* TEXTQUIM, e observaram quais palavras apareciam e quais se repetiam.

A segunda parte da atividade de ensino envolveu a apresentação da natureza de um dicionário de português como língua estrangeira, por meio da identificação de referenciais teóricos de Lexicografia destinada a aprendizes de línguas, com destaque aos dicionários monolíngües para estrangeiros. A partir disso, foram explorados com os alunos os diferentes momentos de planejamento da seleção do vocabulário que figura nesse tipo de dicionário. Em seguida, o texto jornalístico foi apresentado como aquele que geralmente está associado ao reconhecimento de padrões de vocabulário.

Os alunos foram, então, incentivados a produzir um pequeno glossário de português 'popular' ou 'simplificado' para usuários estrangeiros, tomando como base os dados obtidos do jornal DG. A idéia é que tivessem em mente, como usuários, os estudantes do Programa de Português para Estrangeiros da UFRGS. A partir das listas de palavras que cada aluno tinha produzido, eles escolheram duas que achavam que se relacionavam à temática de esportes e, assim, a turma montou o

seu glossário esportivo, destinado a um estudante de português como língua estrangeira. Como era um glossário pensado para um estrangeiro, as explicações tinham de ser simples e diretas, a fim de que o usuário, mesmo com uma proficiência limitada em português, conseguisse atribuir sentido à definição dada para as palavras que encabeçavam os verbetes.

Essas atividades aconteceram na plataforma *Moodle*, de educação a distância da UFRGS, e o glossário foi organizado em ordem alfabética de verbetes, que continham a seguinte estrutura: a classificação gramatical, o significado, uma observação em relação à algum dicionário *on-line* consultado e o exemplo de uso (a frase do DG da qual cada aluno escolheu a sua palavra). Abaixo uma Figura com o *layout* do glossário dos alunos na plataforma *Moodle* da UFRGS:

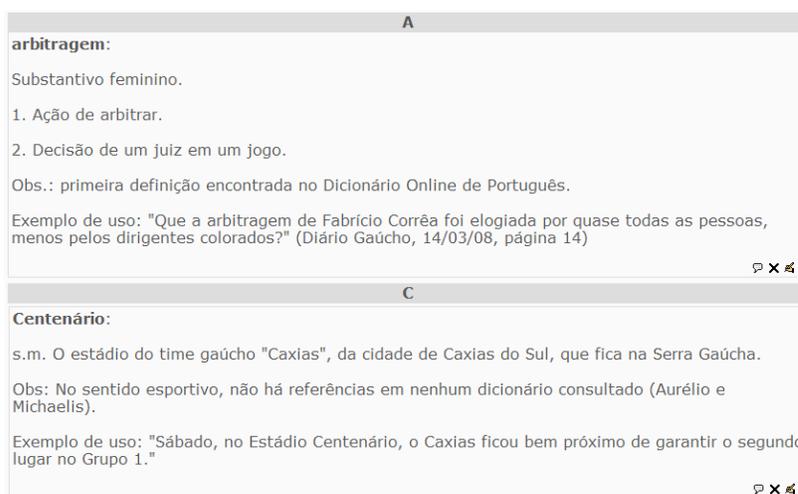


Figura 04: Glossário esportivo da turma de Léxico e Dicionários – UFRGS 2009.

Salientou-se aos alunos que é importante, tanto para o Bacharelado, quanto para o Licenciando em Letras, ter contato com a metodologia básica das pesquisas sobre vocabulário que partem de *corpora* jornalísticos. Além disso, explorar a condição da Língua Portuguesa como língua estrangeira e a produção de materiais como dicionários é ainda uma atividade pouco presente nos currículos dos cursos de Letras. Podemos afirmar que os alunos de Letras ficaram empolgados com a atividade, especialmente no momento em que questionamos o quanto cada verbebo criado seria mais ou menos compreendido pelo usuário.

Com as atividades feitas e baseadas no *corpus* DG, os alunos perceberam diferenças de tratamento e de apresentação entre as palavras de um *corpus* de estudo (no caso, as páginas do jornal DG) e suas definições presentes (ou, muitas

vezes, ausentes) em um dicionário *on-line* consultado, concluindo que, enquanto um *corpus* representa o uso, um dicionário traz descrições – que podem, muitas vezes, ser limitadas ou até opostas ao que se verifica num *corpus* ou em diferentes contextos de uso.

2.2 ENCAMINHAMENTO DO TRABALHO

Neste trabalho de conclusão de curso, tentaremos mostrar como as atividades de ensino do léxico desenvolvidas com alunos da graduação em Letras poderiam ser adaptadas e recontextualizadas para o EJA. O objetivo é que a temática do vocabulário seja apresentada aos educandos de maneira que represente o uso, o cotidiano deles, e que os alunos percebam as diversas possibilidades de significado e de uso das palavras.

A idéia é desenhar uma unidade de ensino de vocabulário para os alunos da 8ª série do EJA, tomando como base o *corpus* do jornal Diário Gaúcho. Para tanto, utilizamos como principal referencial teórico os conceitos da Lingüística de Corpus, expostos no capítulo seguinte. Além disso, no próximo capítulo também são apresentadas as noções teóricas fundamentais que norteiam nossa proposta de ensino de vocabulário, que aproveita o *corpus* do DG como material didático para os alunos do EJA.

3 NOÇÕES FUNDAMENTAIS E REVISÃO DE BIBLIOGRAFIA

Este capítulo traz as noções teóricas fundamentais que norteiam nossa proposta de ensino para os alunos do EJA. Em função disso, trazemos aqui uma revisão da bibliografia da área da Lexicologia e dos estudos em *corpora*. É oferecido um quadro de conceitos relacionados à temática do ensino de vocabulário, tais como palavra - vocábulo, léxico - vocabulário, *corpus/corpora*, estatística lexical, frequência de palavras e riqueza lexical.

3.1 PALAVRA

Palavras, palavras, nada mais
que palavras.
(SHAKESPEARE)

Segundo Rosa (2000, p.73) “de acordo com o uso comum do termo, que tem por base nosso conhecimento da escrita, parece trivial definir o que seja uma palavra”. Isso porque “[...] em todo falante existe uma consciência intuitiva de uma unidade léxica, seja qual for a sua língua materna” (BIDERMAN, 1978, p.72), ou seja, qualquer pessoa sabe identificar uma palavra estando ela numa frase, num texto ou solta. Entretanto, explicar exatamente o que é uma palavra, “quando se pretende uma definição de validade universal” (BIDERMAN, 1978, p.73), não é algo claro e nem mesmo unânime até para os mais renomados lingüistas. Existem várias concepções sobre o que seja uma palavra, e cada ponto de vista parece indicar um objeto diferente (FINATTO, 2009).

O dicionário Houaiss apresenta palavra como “unidade da língua escrita, situada entre dois espaços em branco, ou entre espaço em branco e sinal de pontuação”. Nesse sentido, Mattoso Câmara afirma que:

[...] a apresentação do vocábulo na escrita se faz pelo critério formal. Deixa-se entre eles, obrigatoriamente, um espaço em branco, porque, mesmo quando sem pausa entre si num único grupo de força, cada um é considerado uma unidade mórfica de per si. (CÂMARA, 1982, p.69)

O critério para o estabelecimento do vocábulo formal foi basicamente instituído por Bloomfield (1933), tendo em vista o seu funcionamento na frase. Conforme o lingüista, as unidades formais da língua são de duas espécies:

1) formas livres, quando constituem uma seqüência que pode funcionar isoladamente como comunicação suficiente [...]; 2) formas presas, que só funcionam ligadas a outras [...]. O vocábulo formal é a unidade que se chega, quando não é possível uma nova divisão em duas ou mais formas livres. (CÂMARA, 1982, p.69 *apud* BLOOMFIELD, 1933)

Um exemplo de forma livre indivisível seria luz, de duas ou mais formas presas seria in+pre+vis+ível e de uma forma livre e uma ou mais presas seria in+feliz.

Para abarcar as partículas proclíticas e enclíticas em português, Câmara (1967, p.88) acrescentou uma terceira espécie de unidade formal da língua: “as formas dependentes”. Uma forma dependente seria aquela que não é livre, porque não funciona como comunicação suficiente e também não é presa porque admite intercalação de novas formas e variação posicional na frase. Ou seja, seriam as partículas proclíticas átonas: artigos, preposições, etc.

A existência da forma dependente atesta a falta de correspondência rigorosa entre vocábulo formal e vocábulo fonológico, já que uma forma dependente como as não é um vocábulo fonológico, mas sim parte dele ao se juntar com uma forma livre como amigas, em as amigas, por exemplo. É dessa falta de correspondência que advém a imprecisão do termo, de como conceituá-lo, pois, conforme Mattoso:

O grande problema (...) é que por vocábulo se entendem duas unidades diferentes. De um lado, há o vocábulo fonológico, que corresponde a uma divisão espontânea na cadeia da emissão vocal. De outro lado, há o vocábulo formal ou mórfico, quando um segmento fônico se individualiza em função de um significado específico que lhe é atribuído na língua. (CÂMARA, 1982, p.34)

Houaiss acrescenta que palavra é uma “unidade pertencente a uma das grandes classes gramaticais”, além de ser sinônimo de vocábulo. Quanto à última acepção do dicionário, Mattoso Câmara diz:

Há os dois termos, grosso modo equivalentes, “vocábulo” e “palavra”, cuja distribuição complementar de uso não está bem fixada. O melhor critério para essa distribuição parece ser o de atribuir a “vocábulo” uma significação geral e considerar “palavra” um tipo especial de vocábulo, de aplicação restrita aos nomes e verbos [...]. (CÂMARA, 1982, p.34)

Para ele, toda palavra é um vocábulo, mas nem todo vocábulo é uma palavra. Nesse sentido, o vocábulo é mais abrangente, já que um sufixo, por exemplo, seria um vocábulo, mas não necessariamente uma palavra.

Conforme Biderman (1978), entre os critérios utilizados para definir o termo palavra, três são os mais relevantes: o fonológico, o gramatical e o semântico. Do ponto de vista fonológico, “A palavra pode ser imperfeitamente caracterizada como uma seqüência fonológica que recorre sempre com o mesmo significado” (BIDERMAN, 1978, p.104). Nesse sentido, uma palavra seria uma seqüência fônica completa. Do ponto de vista gramatical, Biderman analisa-o sob duas perspectivas (morfo-sintática) que, segunda ela, atuam de maneira simultânea: “em função de marcadores morfossintáticos que ela [a palavra] apresenta” e em relação à “função exercida pela palavra na sentença” (BIDERMAN, 1978, p.109). Do ponto de vista semântico, Biderman segue Ullmann, que define palavra como “a unidade semântica mínima do discurso” (BIDERMAN, 1978, p.116 *apud* ULLMANN, 1952), mas ressalta que, “Se a informação semântica não congregar todos os outros elementos no topo da hierarquia, os resultados da análise lingüística serão distorcidos” (BIDERMAN, 1978, p.119). Ou seja, ela acredita que para determinar uma unidade léxica é necessário levar em conta todos os três critérios, simultaneamente.

Na perspectiva da Lingüística de Corpus temos algumas percepções bastante ‘operacionais’ de palavra. A mais usual é a consideração da palavra gráfica, sem ignorarmos os estudos que trabalham como formas lematizadas.

Como se percebe, não há coincidência de entendimento para o termo palavra. Cada ponto de vista parece demandar uma compreensão diferente do seu conceito. Neste estudo, cujo encaminhamento é essencialmente aplicado e não teórico, além de restrito à palavra escrita, acreditamos que a melhor opção para definir o termo seja acompanhar Bisognin (2009), em que palavra é “(...) toda unidade lingüística mínima que pode constituir significado, delimitada na escrita por dois espaços em branco e/ou sinal de pontuação” (BISOGNIN, 2009, p. 25).

No sentido de um reconhecimento gráfico da noção de palavra, Basilio (2005) menciona que é preciso considerar seqüências gráficas possíveis e impossíveis em uma língua. Assim, uma seqüência como ‘uthw’, por exemplo, não será considerada como palavra em português.

Além disso, aderindo à LC, neste trabalho cada item gráfico será tomado como um item-palavra, independentemente de sua repetição, variação ou flexões. Isso é que se denomina, em Lingüística de Corpus, de *tokens* (itens, ocorrências) e que corresponde, grosso modo, ao número de palavras gráficas que há num texto

(FINATTO, 2009). Também não distinguiremos palavra de vocábulo, usando ambos como sinônimos, conforme o dicionário.

Sabemos que tais escolhas são um tanto redutoras, todavia, o desenho de uma unidade de ensino requer a adoção de posicionamentos que sejam operacionais.

3.2 LÉXICO – VOCABULÁRIO

Quem fala uma língua sabe
muito mais do que aprendeu.
(CHOMSKY)

Segundo o dicionário Houaiss, léxico é o “repertório de palavras existente numa língua”. E vocabulário é o “conjunto de vocábulos de uma língua”, ou seja, o léxico.

Em outras palavras, podemos afirmar que o léxico é o conjunto de palavras que um indivíduo tem disponível para se expressar numa determinada língua, seja de maneira oral ou escrita. Conforme Bisognin (2009, p.29) “(...) é ao léxico [...] que as pessoas recorrem para entender um povo. O léxico é considerado um espelho fiel da vida, refletindo sua organização”.

Neste sentido, uma característica do léxico é a constante mutação, visto que ele está sempre se modificando. Assim como algumas palavras deixam de ser usadas, se tornando arcaísmos na língua, outras novas palavras se incorporam ao vocabulário, evidenciando essa mutabilidade do léxico.

Para se comunicar, os seres humanos fazem uso do léxico que têm conhecimento, adquirindo, nessa interação comunicativa, novas palavras e formando novas unidades quando necessário. Sobre isso, Basilio (2004) afirma que:

O léxico, portanto, não é apenas um conjunto de palavras. Como sistema dinâmico, apresenta estruturas a serem utilizadas em sua expansão. Essas estruturas [...] permitem a formação de novas unidades no léxico como um todo e também a aquisição de palavras novas por parte de cada falante. (BASILIO, 2004, p. 09)

Basilio defende que ‘um sistema fechado de palavras’ não é suficiente para conceituar léxico, pois o ser humano está constantemente conhecendo e criando

coisas novas, de maneira que precisa expressá-los de alguma forma. Para isso serve o léxico, um "(...) sistema dinâmico, capaz de se expandir à medida que se manifesta a necessidade de novas unidades de designação e construção de enunciados" (BASILIO, 2004, p.09).

Em função dessa mutabilidade, Lepschy (1984, p.161) afirma que o léxico "é o reino da irregularidade", pois estamos sempre criando e recriando palavras. Entretanto, Basilio (2004, p.07) contrapõe esse pensamento, assegurando o alto teor de regularidade do léxico: "Os diferentes processos derivacionais de mudança e extensão de classe servem a funções pré-determinadas, traduzidas em estruturas morfológicas do português". Na verdade, o que acontece, mesmo quando novas palavras são incorporadas à língua, é que estas seguem um padrão, sendo normalmente baseadas em vocábulos existentes. Por exemplo, quando o ministro Antonio Magri³ disse que 'O salário do trabalhador é imexível', muito foi questionado, porque a palavra não existia na língua. Contudo, ele não estava fazendo nada mais que seguir um processo já existente em outras construções da língua, como imbatível, inalterável, etc., acrescentando o prefixo de negação. Tal vocábulo, que na época gerou bastante controvérsia, hoje está dicionarizado e pertence ao léxico da língua portuguesa. A mesma coisa acontece com os estrangeirismos. O verbo *to delete*, do inglês, hoje está dicionarizado como deletar, seguindo o mesmo princípio de outros verbos como cantar, dançar, etc.

Em face dessas considerações acerca do assunto, em nosso estudo, não distinguiremos léxico de vocabulário, usando os dois termos como sinônimos. Além disso, inspirados em Basilio (2004), consideraremos o vocabulário como uma espécie de banco de dados da língua, um depósito de elementos de designação, que fornece unidades básicas e que é capaz de se expandir, originando novas unidades. Nesse sentido, não diferenciaremos elementos lexicais e elementos gramaticais. Sabemos que há concepções que consideram o léxico apenas como o conjunto das palavras que possuem um sentido autônomo na língua, como os verbos e substantivos, desconsiderando elementos estritamente gramaticais, como

³ Foi ministro do Trabalho durante o governo Fernando Collor de Mello, de 1990 até 1992. É associado a um neologismo da década de 90, quando respondeu a um repórter que questionara se o salário também seria reduzido, dizendo: "O *salário do trabalhador é imexível*" (fonte: http://pt.wikipedia.org/wiki/Ant%C3%B4nio_Rog%C3%A9rio_Magri).

os artigos e as preposições (SILVA; FINATTO, 2009). Entretanto, não faremos essa distinção, de maneira que os diferentes itens do vocabulário da língua (tanto ‘mulher’ ou ‘viajar’, quanto ‘de’ ou ‘mas’), embora tenham características distintas, fazem parte do léxico aqui considerado.

3.3 LINGÜÍSTICA DE CORPUS

A observação da linguagem real é a única maneira de se descrever uma língua.
(FIRTH, HALLIDAY e SINCLAIR)

A referência que guia nosso trabalho é a Lingüística de Corpus (BERBER SARDINHA, 2004a), que é o estudo do uso da língua em grandes conjuntos de textos autênticos, sendo que tal estudo é feito com apoio de ferramentas informatizadas e da observação estatística.

A Lingüística de Corpus ocupa-se da coleta e da exploração de corpora, ou conjunto de dados lingüísticos textuais coletados criteriosamente, com o propósito de servirem para a pesquisa de uma língua ou variedade lingüística. (BERBER SARDINHA, 2004a, p.3)

3.3.1 Princípios da Lingüística de Corpus

A Lingüística de Corpus entende a língua como um sistema probabilístico, pois, conforme Sardinha (2000, p.350) “embora muitos traços lingüísticos sejam possíveis teoricamente, eles não ocorrem com a mesma freqüência”. O exemplo que Berber Sardinha traz é a freqüência dos substantivos no nível morfossintático, que, no português, é muito maior que qualquer outra categoria, uma vez que 25% das palavras são substantivos (KENNEDY, 1998 *apud* BERBER SARDINHA, 2000). Com isso, a probabilidade de uma palavra qualquer de um texto ser um substantivo é muito maior do que ela ser de outra classe, já que existem muito mais substantivos na língua.

As bases teórico-metodológicas da Lingüística de Corpus vêm do britânico Firth, que acreditava que o significado de uma palavra se configura pelo seu contexto de uso. O que ele afirmava é que as palavras preferem determinadas associações e ainda rejeitam outras (FINATTO, 2009). Nesse sentido, se fala em combinações possíveis e prováveis. As possíveis são as combinações que se constituem na língua, sendo realizações possíveis de acontecer, e as prováveis são as usadas com maior freqüência, pois têm maior probabilidade de ocorrer. Sobre isso, Maciel declara que:

A recorrência, freqüência e estabilidade dessas combinações mostram que elas não são feitas ao acaso, mas são semântica e pragmaticamente restritas. De fato, embora outras combinações léxico-gramaticais sejam teoricamente possíveis, algumas fraseologias e colocações se repetem, enquanto outras são raras ou inexistentes. (MACIEL, 2002, p.5)

Sobre essa estabilidade das combinações, Berber Sardinha (2000, p.351) afirma que “A linguagem forma padrões que apresentam regularidade [...] e variação sistemática”, por isso, para descobrir quais são esses padrões, a LC se ocupa de análise empírica, extensiva e contrastiva de um *corpus*. A análise acontece de maneira experimental em grandes conjuntos de textos autênticos, os *corpora*, pois, conforme Maciel (2002, p.02) “O objeto de estudo da Lingüística de Corpus é a língua em contexto de uso”. Tal material de estudo tem de ser em linguagem natural, visto que, dessa maneira, registra “a linguagem realmente utilizada por falantes e escritores da língua em situações reais” (BERBER SARDINHA, 2000, p.352). Além disso, o *corpus* deve ser extenso para que assim represente uma variedade lingüística. E essa observação é feita de maneira contrastiva, uma vez que há padrões que têm “freqüência comparável em corpora distintos” (BERBER SARDINHA, 2000, p. 351).

Para que tais análises dos *corpora* obtenham resultados mais precisos, são utilizadas ferramentas eletrônicas. Berber Sardinha (2004a) destaca que já havia *corpus* e a própria LC antes do surgimento do computador, porém, como, nessa época, as análises eram feitas manualmente, havia risco de falhas e equívocos, justamente porque os *corpora* eram analisados a olho humano. Nesse sentido, o computador se torna imprescindível para tais pesquisas, já que, mesmo com uma grande equipe disponível, a análise estaria sujeita ao cansaço e à imperfeição humanos, diferentemente do computador, já que a máquina é incansável (MACIEL, 2002). Maciel ainda ressalta que aparecem erros e formas não aceitáveis nos

corpora, porque se prioriza a autenticidade e a grande quantidade de contextos possíveis e não a gramaticalidade. Em função dessa prioridade é que as formas desviantes aparecem.

3.3.2 Corpus/Corpora

Segundo Berber Sardinha (2004a, p.03), “Havia corpus antes do computador”. *Corpus*, na etimologia da palavra, vem de corpo, representando, no caso, um ‘corpo’ de documentos, um conjunto de dados. Das várias definições de *corpus* propostas por diversos autores, Berber Sardinha (2004a, p.18) considera a mais completa:

Um conjunto de dados lingüísticos (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 lingüístico ou de algum de seus âmbitos, dispostos de tal modo que possam ser processados por computador, com a finalidade de propiciar resultados vários e úteis para descrição e análise. (SANCHEZ, 1996, p.8-9)

Vale ressaltar características importantes, mencionadas na definição de Sanchez:

➤ A origem: o *corpus* deve ser composto por textos autênticos, ou seja, em linguagem natural. Esses textos não devem ter sido produzidos com o fim específico da pesquisa lingüística, se tornando, dessa maneira, artificiais. A importância da autenticidade do texto está no fato de que ela “registra a linguagem natural realmente utilizada por falantes e escritores da língua em situações reais” (BERBER SARDINHA, 2000, p. 352) e não por potenciais falantes idealizados. Além disso, o *corpus* deve ser formado por textos de falantes nativos. Quando isso não acontece, deve ser identificado como *corpora* de aprendizes.

➤ O propósito: o *corpus* deve ter como intenção ser objeto de estudo de pesquisa lingüística.

➤ A composição: o conteúdo do *corpus* tem de ser escolhido de maneira criteriosa, sobretudo, valorizando a naturalidade e a autenticidade. Além disso, deve obedecer a critérios específicos de acordo com o que se pretende pesquisar. Por exemplo, como pretendemos observar o vocabulário do jornal Diário Gaúcho, nossa

seleção de textos foi feita com a intenção de obter uma amostra composta pelo todo do jornal de cada dia, para tanto, selecionamos diferentes dias não consecutivos de cada semana.

➤ A formatação: o *corpus* tem de ser formatado de maneira que os dados sejam legíveis por computador.

➤ A representatividade: o *corpus* deve ser representativo daquilo que se pretende observar, normalmente uma língua ou variedade. Em nosso caso, nosso *corpus* de estudo, o *corpus* do DG é representativo do jornal como um todo. “Para que seja representativo, um *corpus* deve conter o maior número possível de sentidos de cada forma” (BERBER SARDINHA, 2004a, p.24). Ou seja, essa característica está associada à próxima:

➤ A extensão: para que seja representativo, um *corpus* deve ser extenso, o maior possível, segundo Berber Sardinha (2004a). Não há critérios específicos que determinem a extensão, entretanto, dela depende os resultados da análise. Como para LC, a linguagem é um sistema probabilístico, na qual certos traços são mais freqüentes que outros (BERBER SARDINHA, 2000), quanto mais extenso o *corpus*, maior a quantidade de palavras e, portanto, mais precisamente se consegue identificar quais são esses traços mais recorrentes e também quais são os de menor recorrência.

Em nosso trabalho, então, um *corpus* será definido como um conjunto de textos autênticos, em linguagem natural, extenso (de maneira que seja representativo) e legíveis por computador, nossa ferramenta para descrição e análise do vocabulário do jornal Diário Gaúcho.

3.3.3 Estatística Lexical

Segundo Gerber e Vasilévski, além do desenvolvimento acelerado dos recursos computacionais, a matemática também apóia, atualmente, a pesquisa lingüística, “pois se podem pesquisar grandes volumes de dados, contando-se com processamento automático e dados numéricos” (GERBER e VASILÉVSKI, 2007, p.176). Para as autoras a “Lingüística com números é a possibilidade que emerge”

(GERBER e VASILÉVSKI, 2007, p.176), ou seja, é um método de estudo novo, mas promissor.

Conforme Biderman (1998, p.162) “Muitas teorias foram elaboradas para tratar o fenômeno da linguagem. Uma delas, a estatística lingüística, considera a face quantitativa da linguagem”. A autora justifica tal abordagem em função da elevadíssima frequência dos fenômenos lingüísticos que, ao serem analisados, originaram resultados eficazes nas décadas de 60 e 70. A estabilidade dos signos lingüísticos torna possível a sua previsibilidade, visto que, segundo Biderman:

Na verdade a frequência de letras, fonemas, número de sílabas e comprimento da palavra são independentes do estilo individual e constituem um condicionamento lingüístico. Assim é possível prever os fonemas, grafemas, vocábulos e unidades gramaticais que poderão ocorrer nos discursos oral e escrito dos falantes e escritores. (BIDERMAN, 1998, p.162)

Sendo assim, o uso de determinada forma em detrimento da outra pelo falante, acontece porque isso independe deste, se relacionando com as formas mais aceitas e usadas pela comunidade dos falantes como um todo, e não pela vontade de cada indivíduo.

Isso significa que, no uso da linguagem, há um grande número de predeterminações, condicionadas pela língua em questão e alheias à interferência do indivíduo, que não escolhe os elementos lingüísticos de que se vai servir na sua expressão oral ou escrita. (BIDERMAN, 1978, p.05)

Nesse sentido, a autora comenta que, se alguém analisar uma vasta quantidade de dados da linguagem, conseguirá prever quais fonemas, grafemas, vocábulos e unidades gramaticais ocorrerão nos discursos oral ou escrito dos usuários daquele sistema. Isso acontece porque, “em qualquer corpus lingüístico, notam-se [...] certas constantes na distribuição dos signos” (BIDERMAN, 1978, p.09). Mesmo assim, Gerber e Vasilévski (2007) acreditam que, por se tratar de língua, temos que ponderar que ela varia continuamente, permanecendo em estado de renovação constante. Além disso, as autoras afirmam que:

[...] a classificação pode não ser fácil nem unânime, pois a língua é uma população complexa, peculiar e relativamente nova em pesquisas que envolvem Estatística, e sempre pode haver uma ou outra ocorrência que insista em destoar das demais. (GERBER E VASILÉVSKI, 2007, p.180)

No caso específico do Português, depois de realizar pesquisa e análise com *corpora* variados (textos literários, jornalísticos, etc.), Biderman (1998) obteve o léxico mais freqüente da nossa língua. Dessa investigação, a autora concluiu que embora o léxico da língua seja muito vasto, os falantes não usam todas as formas possíveis, muito pelo contrário, “o uso desse tesouro lexical por parte dos usuários

da língua é bem modesto” (BIDERMAN, 1998, p.178). Ela revela que, mesmo na língua escrita, na qual o indivíduo dispõe de um vocabulário mais rico e mais variado, a multiplicidade de formas usadas também não é muito grande. Ainda sobre o vocabulário do Português, Biderman constata que, ao seguir uma ordem decrescente de freqüência das palavras, é possível observar que os vocábulos gramaticais são usados mais frequentemente, estando no topo da lista elementos como ‘de’, ‘e’, ‘a’. A estes seguem palavras de uso menos freqüente, chegando, ao final da lista, aos vocábulos específicos, “cuja freqüência diferirá de texto para texto, até atingirmos os ‘*hapax legomena*’ que caracterizam cada corpus em particular” (BIDERMAN, 1978, p.09). Essas freqüências maiores ou menores que cada *corpus* apresenta serão discutidas no item seguinte.

3.3.4 Padrões de freqüência de palavras

Segundo Berber Sardinha (2004a, p.23) “a linguagem é um sistema probabilístico de combinatórias” e, nesse sentido, alguns traços são mais recorrentes que outros. “No caso do léxico, pode-se diferenciar entre aqueles de maior freqüência e os de menor freqüência, sendo que a diferença entre eles é relativa” (BERBER SARDINHA, 2004a, p.23). Essa ‘relatividade’ é citada pelo autor porque enquanto alguns vocábulos têm alta freqüência, outros têm freqüência muito rara.

A importância de se colocar o foco no léxico é discutida por Nation (2001), que defende que a aprendizagem de vocabulário deve ser baseada em um estudo cuidadoso das freqüências lexicais. Segundo este autor, há que se ter em mente que as palavras de uma língua se dividem em dois grupos: palavras de alta freqüência e palavras de baixa freqüência⁴. (VICENTINI, 2006, p.35)

Segundo Vicentini, Nation esclarece a diferença entre a análise de um *corpus* geral de uma língua e de um *corpus* de determinado texto desta língua, uma amostra em situação específica. O *corpus* geral é composto de grande maioria de palavras de baixa freqüência, pois nele existem diversos assuntos e temáticas

⁴ Em Vicentini (2006) não há explicação sobre o que seria considerado especificamente como ‘palavras de alta freqüência’ e ‘palavras de baixa freqüência’. Segundo ela, os dados são de Nation (2001), mas para efeito ilustrativo, pois o autor não apresenta valores exatos para suas considerações.

envolvidas, apontando para o fato das palavras provavelmente serem mais diversificadas, se repetindo menos. Por outro lado, na análise de determinado texto desta língua, por ser uma amostra da língua em uso em situação específica, os valores se invertem. Neste caso, a grande maioria das palavras tende a ser de alta freqüência, porque a situação analisada é específica e, por isso, aparecem as mesmas palavras (daquele contexto específico) várias vezes. Vicentini (2006, p.37) complementa citando como exemplo palavras do âmbito da LC, como ‘*corpora*’ e ‘*colocação*’, que com muita freqüência aparecerão quando analisado um *corpus* da área, mas não aparecerão quando analisado um *corpus* geral da língua. Neste caso, da análise de um *corpus* geral, as formas normalmente mais encontradas são ‘e’, ‘um’, ‘de’, que “não estão relacionadas a apenas um tipo de assunto e podem ser encontradas em qualquer tipo de texto” (VICENTINI, 2006, p.37).

Ainda sobre as palavras de rara freqüência, Sardinha defende que, para que apareçam num *corpus*, é necessário que ele seja o mais extenso possível, pois “quanto maior a quantidade de palavras, maior a probabilidade de aparecerem palavras de baixa freqüência” (BERBER SARDINHA, 2004a, p.23). Essas formas, conhecidas como *hapax legomena*, segundo ele, são a maioria. Em função disso, pode-se afirmar que essas palavras são as mais comuns, ou seja, a maioria das palavras de uma língua são as que ocorrem poucas vezes. Por ocorrerem poucas vezes e por formarem a maior parte do vocabulário de uma língua, são necessárias amostras amplas para que elas apareçam e possam ser analisadas, daí a importância da extensão de um *corpus* de estudo. (BERBER SARDINHA, 2004a)

Em nosso trabalho, interessou observar os diferentes padrões de freqüência das palavras que aparecem no *corpus* do jornal Diário Gaúcho, para, a partir disso, caracterizá-las de acordo com a sua presença e função nos textos. Isso auxiliou nas atividades de ensino de vocabulário propostas para os alunos do EJA. Além disso, conforme os padrões de freqüência, também são analisados os possíveis sentidos de um vocábulo, o que também é explorado com os alunos, fazendo parte da unidade de ensino sugerida.

3.4 PRÓXIMOS PASSOS

Revisados e apresentados os conceitos teórico-metodológicos que norteiam nossa proposta de ensino sobre vocabulário para os alunos do EJA, no próximo capítulo discutimos sobre o ensino para jovens e adultos, bem como sobre o público-alvo do nosso trabalho: professores e alunos do EJA. Além disso, o capítulo seguinte apresenta como é tratada a temática do vocabulário nas aulas de língua portuguesa atualmente e como pretendemos abordá-la em nossa unidade de ensino para os alunos do EJA.

4 CENÁRIOS DE ENSINO/APRENDIZAGEM ENVOLVIDOS

Neste capítulo caracterizamos o cenário de ensino/aprendizagem no qual nossa proposta de ensino sobre vocabulário está inserida. Assim, importa identificar, como segmento social, os alunos do Ensino de Jovens e Adultos. Em geral, são adultos, também trabalhadores, com idades variadas. Em muitos casos, abandonaram os estudos ainda jovens. Abordamos também o perfil dos professores, mestres que, muitas vezes, são uma 'luz no fim do túnel' para essas pessoas.

Outro ponto importante que levantamos neste capítulo é o ensino de vocabulário. A intenção é ponderar sobre qual seria o seu espaço nas aulas Língua Portuguesa e como poderia ser adequadamente tratado na unidade de ensino que idealizamos para turmas de EJA.

4.1 EJA – ENSINO PARA JOVENS E ADULTOS

A educação é direito de todos
e dever do Estado e da família
[...].
(Constituição Federal de 1988,
Art.205)

Conforme a Constituição Federal Brasileira de 1988, a educação é um direito de todos, entretanto, nem todos conseguem usufruir deste na idade apropriada. Foi para tentar suprir esse 'atraso' que o ensino supletivo foi implantado, em 1971, sendo marco importante na história da educação de jovens e adultos no Brasil (LOPES e SOUZA, 2005).

Durante o período militar, a educação de adultos adquiriu pela primeira vez na sua história um estatuto legal, sendo organizada em capítulo exclusivo da Lei nº 5.692/71, intitulado ensino supletivo. O artigo 24 desta legislação estabelecia com função do supletivo suprir a escolarização regular para adolescentes e adultos que não a tenham conseguido ou concluído na idade própria. (Vieira, 2004, p. 40).

Em 1971, então, foi legalizado o Ensino de Jovens e Adultos – EJA – (ainda intitulado supletivo) buscando levar de volta à escola aqueles que por algum motivo

não concluíram seus estudos. A idéia inicial do EJA, assegurada, logo em seguida, pela Constituição de 1988, era disponibilizar acesso ao ensino fundamental e ao ensino médio àqueles que não completaram seus estudos, a fim de garantir o direito que essas pessoas têm à educação básica.

Essa modalidade de ensino era oferecida apenas pelas escolas públicas, segundo a Constituição, garantindo o direito à educação básica obrigatória e gratuita. Com o passar do tempo, o ensino passou a ser livre à iniciativa privada, desde que autorizado e avaliado qualitativamente pelo Poder Público.

Além disso, ao final dos anos 90, o Ensino de Jovens e Adultos passou a ser oferecido não mais somente àqueles que iniciaram os estudos e, por algum motivo, tiveram de interromper, mas também na modalidade inicial de alfabetização, àqueles que nunca tinham freqüentado uma sala de aula.

O EJA, por tratar de assegurar educação básica àqueles que não tiveram acesso na idade adequada, lida com pessoas de diferentes idades, com vivências e cultura próprias, de acordo com as experiências vividas. O ensino para esses alunos, então, deve levar em conta esse fato, já que, diferentemente dos alunos matriculados regularmente, os jovens e adultos que freqüentam os cursos de EJA, na sua maioria, não estão em formação, tendo identidade e opinião formadas.

Sabe-se que educar é muito mais que reunir pessoas numa sala de aula e transmitir-lhes um conteúdo pronto. É papel do professor, especialmente do professor que atua na EJA, compreender melhor o aluno e sua realidade diária. Enfim, é acreditar nas possibilidades do ser humano, buscando seu crescimento pessoal e profissional. (LOPES e SOUZA, 2005, p. 02).

Segundo Lopes e Souza (2005), o ensino para jovens e adultos precisa utilizar os conhecimentos trazidos pelo aluno, aproveitando seus saberes no ensino das disciplinas do currículo. Essa seria uma forma de cativar o estudante que, vendo em aula fatos do seu cotidiano, tende a ficar motivado com as atividades e com escola em si. Além disso, com a utilização do conhecimento do educando, ele provavelmente perceberá que a sua vida até aquele momento também foi um aprendizado e que, agora, suas experiências de vida podem servir como base para novos aprendizados.

4.1.1 Alunos e Professores em perfil

Segundo Loch (2009, p.21), “O trabalho de planejamento e avaliação do ensino na EJA [...] é uma atividade pedagógica complexa dada a evidência da grande heterogeneidade presente neste grupo.” A autora afirma, com isso, que não só a faixa etária dos alunos é bastante diversa - dos 15 até quase os 90 anos - nas turmas de EJA, mas também a bagagem cultural, que é justamente o que os caracteriza, “a diversidade cultural, étnica, de gênero, de ofício, etc.” (LOCH, 2009, p.19), de maneira que ambas precisam ser pensadas nos planejamentos de aula.

As turmas de EJA são compostas, na sua maioria, por jovens que procuram a conclusão dos estudos em busca de novas oportunidades e por adultos que acreditam na formação básica como mola propulsora para melhorias nas condições de vida.

Quando visitamos uma sala de aula de EJA e vemos que, nessa sala, convivem um jovem morador de rua, um jovem deficiente, um jovem-adulto trabalhador, avós de 70, 80 anos, negros, brancos, pardos, homossexuais, todos e todas com um desejo ardente de aprender mais, que têm histórias para contar, não apenas pessoais, mas históricas, nos damos conta da riqueza desta diversidade para o planejamento [...]. (LOCH, 2009, p.19)

Nessa afirmação de Loch percebemos a diversidade presente nas salas de aula do EJA, diversidade com a qual o professor tem de saber lidar. Os alunos do EJA sofrem com discriminação, preconceito, vergonha, etc. e, cabe ao professor identificar essas questões, utilizando-as a favor da educação, mostrando aos alunos que “a EJA é uma educação possível e capaz de mudar significativamente a vida de uma pessoa, permitindo-lhe reescrever sua história de vida” (LOPES e SOUZA, 2005, p. 02).

Lopes e Souza afirmam que o professor do EJA deve ser “um profissional especial, capaz de identificar o potencial de cada aluno” (LOPES e SOUZA, 2005, p.02) e nessa perspectiva criar estratégias de ensino que incentivem os alunos a frequentarem e a participarem das aulas, bem como a realizarem as tarefas e atividades sugeridas. Caso contrário, as autoras afirmam que o ensino ficará limitado à imposição de um padrão, tornando as atividades mecânicas e com fim em si mesmas, não objetivando que os alunos reconheçam a importância dos conhecimentos adquiridos e a presença deles nas suas vidas. Sobre isso, Oliveira afirma que

De outra parte, no que se refere a seleção dos conteúdos, cabe ressaltar a necessidade de uma lógica que os compreenda não como uma finalidade em si, mas como meio para uma interação mais plena e satisfatória do aluno com o mundo físico e social à sua volta, oportunizando a essas populações a valorização dos saberes tecidos nas suas práticas sociais em articulação com os saberes formais que possam ser incorporados a esses fazeres/saberes cotidianos, potencializando-os técnica e politicamente. Na seleção dos conteúdos, a prioridade seria, então, a abordagem de conhecimentos relacionados à vida social e à compreensão dos elementos que intervêm na vida cotidiana. (OLIVEIRA, 2010, p.107)

Nesse sentido, Lopes e Souza destacam a proposta de Paulo Freire, educador ligado à educação popular e, sobretudo, à alfabetização. Freire propõe que as aulas do EJA sejam baseadas na realidade do educando, levando em conta suas experiências e opiniões. O papel do educador é de organizar esses dados de maneira que o conteúdo, a metodologia e o material utilizados nas aulas sejam compatíveis com a realidade dos alunos, pois, assim, o aluno-adulto-trabalhador compreenderá o que está sendo ensinado e saberá aplicar em sua vida os conhecimentos adquiridos na escola (LOPES e SOUZA, 2005).

O aluno ativo no processo de ensino-aprendizagem, sobretudo o adulto, segundo Lopes e Souza, vai se sentir mais interessado pela escola e mais responsável por aquilo que está aprendendo. Essa é uma forma de recuperar a auto-estima desses adultos, normalmente esquecida. Com isso, o educando perceberá que está modificando a sociedade e a sua realidade, tudo isso por meio da educação.

4.2 O ENSINO DE VOCABULÁRIO

O ensino do vocabulário é a base para o ensino de qualquer língua, por isso, tal temática é recorrente nas aulas de língua, seja ela materna ou estrangeira.

No caso da língua estrangeira, a aquisição do vocabulário acontece a fim de que o indivíduo se comunique e entenda aquele idioma, de maneira oral ou escrita. Inicialmente, o que geralmente é ensinado ao estudante de uma determinada língua estrangeira é o léxico básico daquela língua, a fim de que o aprendiz consiga entender e se fazer entender com palavras e expressões simples e cotidianas, tais como 'bom dia', 'boa noite', 'olá', etc. Em seguida, são ensinadas algumas frases e, com isso, a estrutura sintática básica do idioma em questão, para que o aprendiz

consiga se comunicar minimamente dizendo qual seu nome, sua idade, perguntando onde fica algum lugar, etc.

A partir daí, o repertório vocabular do aprendiz vai aumentando e, assim, ele vai conhecendo novas palavras, expressões, estrutura sintática e aprendendo o funcionamento daquela língua, uma vez que “a utilização do vocabulário não é feita de forma isolada” (BARBEIRO, [entre 2003 e 2007], p.02).

O ensino de vocabulário de língua estrangeira se diferencia do ensino de vocabulário de língua materna, pois, neste caso, nada é completamente estranho ao indivíduo, visto que ele já esteve e está constantemente em contato com a língua em foco desde pequeno, aumentando seu repertório vocabular com tempo e conforme as experiências que tem/teve. Sobre esse fato, da criança já entrar na escola com bom conhecimento das palavras da língua, o Projeto Diversidade Lingüística na Escola Portuguesa⁵ afirma que:

O papel da escola, da aprendizagem formal, pode então orientar-se para o seu reforço e alargamento, para a conscientização do seu uso adequado e para a sua correção quanto a aspectos morfológicos, sintáticos e, também, quanto a aspectos de rigor semântico. (BARBEIRO, [entre 2003 e 2007], p.01)

O que o Projeto defende é que a escola precisa atentar para esse repertório já trazido pelo aluno, buscando seu alargamento, bem como a discussão do uso (se é adequado ou não e em quais situações) e da forma correta. Nesse sentido, o vocabulário transita por vários dos conteúdos específicos das aulas de língua materna. Sendo assim, imagina-se que tal temática seja permanentemente tratada e estudada, o que analisaremos no item seguinte.

4.2.1 O espaço do ensino de vocabulário nas aulas de Língua Portuguesa

Segundo Dargel (2005), o elemento central das aulas de Língua Portuguesa é o texto. Para que haja recepção e compreensão deste, os alunos precisam,

⁵ Este projeto estuda a diversidade lingüística presente nas escolas de ensino básico da região metropolitana de Lisboa, em consequência do crescente número de imigrantes provenientes dos mais diversos países. Seu objetivo é a valorização da presença do multilingüismo nas escolas e da língua materna dos alunos, além da criação de materiais que melhorem a competência em língua portuguesa dos alunos de outras línguas maternas.

primeiramente, compreender as palavras que o compõem. Com isso, é possível perceber a importância do ensino de vocabulário e a onipresença dele nas aulas de português.

A autora afirma que os vocábulos não devem ser trabalhados de maneira isolada, fora de contexto, mas sim de maneira que auxilie os alunos na compreensão do sentido global do texto.

Acredita-se que a partir da compreensão do sentido de uma palavra no texto, o aluno, seguindo orientações do professor, poderá perceber outros sentidos, ou oposições, que essa palavra pode ter em outros contextos diferentes daquele que foi estudado no texto. Assim, esse aluno estará ampliando seu conhecimento lexical. (DARGEL, 2005, p.01)

Ela acredita que o trabalho com o vocabulário específico de um texto fará o aluno perceber que as palavras têm diferentes sentidos e usos, de acordo com cada situação. Em relação a isso, Dargel (2005) defende que o educando que lida com unidades desconhecidas, entendendo-as e incorporando-as, tende a expandir seu repertório vocabular e também sua visão de mundo.

O recurso didático utilizado com maior frequência nas aulas de Língua Portuguesa para o auxílio da compreensão do vocabulário são os dicionários. Conforme Rodrigues:

[...] o dicionário é um dos recursos mais utilizados e adequados à consulta de novos vocábulos de língua estrangeira ou materna e, conseqüentemente, ao enriquecimento lexical do usuário de dada língua. (RODRIGUES, 2009, p.01)

Os dicionários de língua são usados freqüentemente na compreensão do significado das palavras, bem como no auxílio às dúvidas de ortografia e de acentuação, sendo muito presentes nas aulas de português. Dargel (2005) defende que esse uso deve ser assessorado pelo professor, já que muitas obras apresentam diversas acepções para o mesmo verbete, além de fazer referências a outros verbetes, encaminhando o usuário a uma nova consulta, o que, muitas vezes, confunde o aprendiz.

Outro espaço do ensino de vocabulário nas aulas de português é trazido pelos livros didáticos: a associação entre palavras e/ou entre palavras e imagens (ANEXOS E e F). A associação entre palavras ocorre de diversas formas, como, por exemplo, entre palavras primitivas e palavras derivadas (cujas bases são iguais e, por isso, o aluno é instigado a desvendar o sentido da palavra derivada a partir da palavra primitiva) ou entre palavras de mesmo prefixo ou sufixo (em que o aluno descobre o significado desmembrando a palavra nos seus constituintes).

Nos livros didáticos também é recorrente a prática de ensino de vocabulário por sinonímia ou por antonímia. Nesses casos, o aluno descobre o significado de um vocábulo desconhecido ao relacioná-lo com seu sinônimo ou antônimo, geralmente palavra mais utilizada na língua, que é, portanto, conhecida pelo aprendiz (ANEXO G).

O grande equívoco desses livros, segundo Silva (2005), é trazer exercícios desse tipo com frases soltas, fora de contexto, privando o aluno de relacionar o(s) significado(s) das palavras com o seu contexto e também de perceber os diversos usos e sentidos possíveis de uma palavra.

Por fim, ainda sobre os livros didáticos, quando o vocabulário é explorado dentro de textos, a presença de glossários ao final destes ou ao final dos livros é recorrente. Essa prática é questionável, visto que não estimula os alunos a buscarem o significado das palavras nem nos dicionários, nem no contexto, pois eles já recebem o sentido de cada uma delas pesquisado.

O ensino de vocabulário, conforme apresentado, transita por vários dos conteúdos específicos das aulas de língua materna. Contudo, na maioria das vezes, ainda não é uma prática cotidiana das aulas de Língua Portuguesa. Ao longo dos textos e das atividades desenvolvidas com os alunos, imagina-se que o ensino do léxico se faça presente, visto que, conforme Barbeiro ([entre 2003 e 2007], p.02) “a utilização do vocabulário não é feita de forma isolada”. Entretanto, infelizmente, isso nem sempre acontece.

4.2.2 Enfoques do ensino de léxico no EJA – opções e princípios envolvidos

Segundo Machado (2003), Richards (1976 *apud* Carter e McCarthy, 1998) faz considerações sobre o aprendizado do vocabulário de uma língua: “[...] os falantes nativos continuam a expandir o seu vocabulário na fase adulta, embora pouco se saiba sobre a média de palavras que um adulto adquire” (MACHADO, 2003, p.41 *apud* RICHARDS, 1976).

Ou seja, mesmo para os alunos do EJA, com idades diversas, sendo muitos adultos e até idosos, é importante abordar a temática do vocabulário, já que o ser

humano continua sempre expandindo seu repertório lexical, independentemente da sua idade.

Outro ponto abordado por Richards (1976) *apud* Machado (2003) é sobre conhecer uma palavra:

[...] conhecer uma palavra é saber as probabilidades que outras palavras têm em se associar a esta. [...] É conhecer as limitações que esta tem em relação à função e à situação em que ela se encontra. [...] Além disso, significa conhecer sua posição dentro da língua, associando-se com outras palavras dessa língua; mais ainda, significa conhecer o seu valor semântico. E por último, conhecer uma palavra, [...] significa conhecer os seus diferentes significados. (RICHARDS, 1976 *apud* MACHADO, 2003, p.41)

Nesse trecho, o autor aborda diferentes possibilidades de ensino sobre o léxico de uma língua, mostrando como tal temática é rica, abrindo um leque de possibilidades de ensino. É dessa forma diversificada, levando em conta o que está presente na vida do aluno do EJA (para que ele reconheça a importância e consiga atribuir valor ao que está aprendendo), que pretendemos abordar a temática do vocabulário e sugerir atividades para o ensino deste nas aulas de Língua Portuguesa do EJA, pois, segundo Silva (2005, p.69) “constatamos a necessidade de explorar o sentido das palavras em situações reais de uso para que os alunos as compreendam de forma plena”.

Nossa idéia é utilizar textos do jornal Diário Gaúcho, tendo em vista a sua provável familiarização⁶ com os alunos do EJA, como fonte para discussões e atividades em torno do léxico de nossa língua. As atividades englobam, como enfatiza Machado (Richards 1976 *apud* MACHADO, 2003), a palavra e seu(s) comportamento(s), além de seus contextos e sua associação com outros elementos da língua. Buscamos que o aluno compreenda o funcionamento do léxico da língua portuguesa, a sua língua materna, de acordo com seu uso real, conforme a Lingüística de Corpus, e “a partir de elementos que são parte da experiência do aluno” (SILVA, 2005, p.65).

⁶ O grau de instrução dos seus leitores está assim distribuído: 60% com Ensino Fundamental, 34% com Ensino Médio e 6% com Ensino Superior. A renda média do leitor está na faixa de até 5 salários mínimos. (AMARAL, 2006, p.38-39 e 80-85).

5 EXPLORAÇÃO DO *CORPUS* E PROPOSTAS DE ENSINO

Este capítulo traz a descrição dos padrões lexicais mais presentes no *corpus* de seis meses do jornal Diário Gaúcho do ano de 2008. O desenho desses padrões e sua percepção por parte do professor e de seus alunos contribuíram para as atividades de ensino, principalmente, para as discussões sobre vocabulário propostas para os alunos do Ensino de Jovens e Adultos.

Além disso, neste capítulo também são expostas idéias práticas de atividades de sala de aula para o ensino de vocabulário no EJA. A maior parte dos recursos provém do DG, o qual entendemos ser material de fácil acesso aos alunos e professores da rede pública.

5.1 PADRÕES DO LÉXICO NO DG: DESCRIÇÃO

Como já foi mencionado, o *corpus* de estudo da pesquisa PorPopular reúne os seis primeiros meses do DG do ano de 2008. Essa amostra de textos, colhida de acordo com a metodologia exposta no capítulo 1, contém 974.672 palavras (*tokens*) sendo que, destas, 44.456 são palavras diferentes (*types*).

Conforme exposto no capítulo 2, de acordo com a Lingüística de Corpus, “a linguagem é um sistema probabilístico” (BERBER SARDINHA, 2004a, p.23). Nesse sentido, alguns traços são mais recorrentes que outros, por isso, na análise de um *corpus*, é possível diferenciar o léxico mais freqüente do léxico menos freqüente (BERBER SARDINHA, 2004a).

As palavras que mais aparecem no *corpus* DG são os artigos (*o, a, um, uma*) e as preposições (*de, do, da, para, com, em, no, na*), ou seja, as palavras chamadas gramaticais da língua. Essas palavras exprimem relações entre as idéias e não têm sentido próprio específico, uma vez que seu sentido está mais ligado à combinatória em que se encontram, como, por exemplo, o uso de ‘*de*’, em escova *de* dentes (em que *de* implica ‘que é usada nos/para os dentes’) e em escada *de* ferro (com o sentido do *de* associado a ‘que é feita de ferro’). Além destas, entre as unidades mais freqüentes também estão algumas conjunções e alguns pronomes. Segundo

Santos, essas palavras constituem um sistema fechado da língua, existindo em número reduzido e limitado (SANTOS, 2000).

Por outro lado, as palavras conhecidas como palavras lexicais, de conteúdo nocional mais específico, constituem um sistema aberto da língua, estando em renovação constante, já que por meio de adaptações gráficas, morfológicas ou semânticas nessas palavras, surgem novos vocábulos que são incorporados à língua (SANTOS, 2000). Entre as palavras lexicais, a que mais aparece no *corpus* é o verbo *ser*, conjugado nas suas formas de terceira pessoa, no presente *é* e no passado *foi*.

O advérbio de negação *não* aparece bastantes vezes, sendo acompanhado por *mais*. E os substantivos encontrados com mais freqüência são *anos*, *día* e *vagas*.

Na Figura abaixo, é possível visualizar a quantidade das 50 palavras mais freqüentes nos seis meses analisados e a representatividade delas no *corpus*.

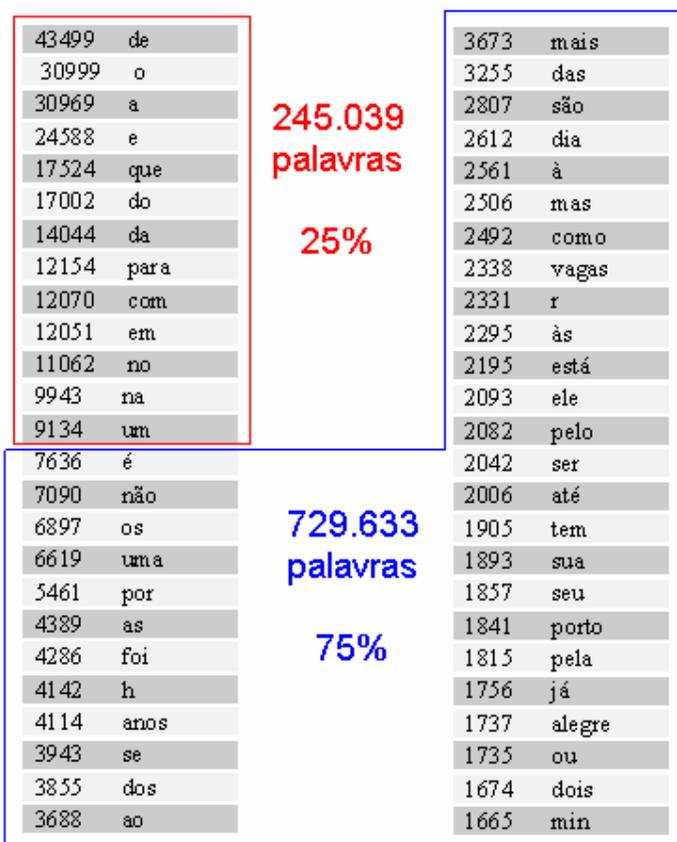


Figura 05: 50 palavras mais freqüentes no *corpus* de 6 meses do DG.

Por meio dessa Figura é possível perceber que as palavras gramaticais aparecem sem interrupções nas 13 primeiras posições, além de continuar aparecendo na seqüência desse *corpus* analisado. Somente essas 13 primeiras palavras que aparecem representam 25% do total de palavras desse *corpus*, o que já é uma quantidade alta. Entre os 75% restantes também existem palavras gramaticais, apontando para o fato de que elas são, muito provavelmente, a maioria das palavras que aparecem nesse *corpus*.

Isso mostra que, no DG, assim como em outros *corpora*, as palavras gramaticais aparecem em grande quantidade. Segundo Bueno, Tavares e Paz (2009, [s.p.]), “as palavras gramaticais são pouco numerosas, mas de altíssima freqüência”. Isso acontece porque estas são palavras funcionais na língua, utilizadas para indicar, determinar, identificar, apontar ou substituir termos já utilizados, para relacionar frases e/ou elementos, etc.

O restante das palavras, chamadas lexicais, geralmente identificam o(s) assunto(s) principal (is) e o que é importante para aquele texto, caracterizando-o. No caso do *corpus* DG, a recorrência da palavra *vagas* é a que mais nos chama atenção, pois o jornal realmente disponibiliza muitas oportunidades de emprego em suas páginas, o que vai de encontro à sua proposta de assistencialismo, exposta no manual de redação do jornal.

Esses padrões identificados no *corpus* DG, sobretudo os diferentes padrões de freqüência e as palavras mais utilizadas, forneceram dados para as atividades e, principalmente, contribuiram para a discussão sobre vocabulário proposta aos alunos do EJA no item seguinte.

5.2 IDÉIAS PARA O ENSINO DE VOCABULÁRIO NO EJA

A idéia é aprimorar os conhecimentos sobre o léxico da língua portuguesa dos alunos EJA, visto que eles já têm um repertório vocabular próprio, já que se trata de um público jovem e adulto. Utilizamos, para tanto, como material de estudo, textos do DG, uma vez que o jornal provavelmente faz parte do universo dos alunos e/ou

de suas famílias, pois nosso público-alvo principal são os educandos das escolas públicas⁷.

As atividades são guiadas pelos princípios teórico-metodológicos da Lingüística de Corpus, conforme apresentado neste trabalho, entretanto, os termos e a metodologia utilizados são adaptados ao contexto escolar de uma aula de Língua Portuguesa, sendo acessíveis aos professores e alunos do Ensino de Jovens e Adultos.

Idealizamos como público-alvo das atividades os alunos da 8ª série do ensino fundamental do EJA de escolas públicas, pois, nesse momento em que eles estão concluindo o ensino fundamental, é importante um trabalho cuidadoso com a leitura e a escrita e, a partir delas, com léxico, buscando o alargamento do repertório vocabular do aluno.

5.2.1 Unidade de ensino de vocabulário

A temática do vocabulário é a base para o ensino de línguas, conforme exposto no capítulo anterior, por isso, se faz quase onipresente nas aulas de língua portuguesa. Entretanto, as abordagens mais comuns são aquelas que propõem trabalhar com o vocabulário isolado (fora de um texto), com a utilização de dicionários sem maior exploração do que estes trazem a cada definição, com a descoberta do significado das palavras por associação, etc. Tais práticas, de certa forma, privam o aluno de relacionar o(s) significado(s) das palavras com seus usos em contexto e, da mesma maneira, não apresentam ao educando os diversos usos e sentidos possíveis de uma palavra, aliás, nem sequer expõem que uma palavra pode ter mais de um uso e/ou significado.

Em função disso e porque “constatamos a necessidade de explorar o sentido das palavras em situações reais de uso para que os alunos a compreendam de forma plena” (SILVA, 2005, p.69), utilizamos os princípios da Lingüística de Corpus,

⁷ O DG, quando lançado, foi definido como uma publicação destinada a um casal com filhos de 8 a 12 anos, o pai vigilante, a mãe faxineira, com renda mensal de R\$ 1.200,00. O grau de instrução dos seus leitores está assim distribuído: 60% com Ensino Fundamental, 34% com Ensino Médio e 6% com Ensino Superior. [AMARAL, 2006, p.38-39 e 80-85].

que acredita que o significado de uma palavra se configura pelo seu contexto de uso, observando, dessa maneira, o uso real das palavras.

Nesse sentido, idealizamos uma unidade de ensino que permita que o aluno perceba os diferentes modos de funcionamento do léxico da sua língua materna, a Língua Portuguesa, de acordo com seu uso em um tipo de texto que trata de elementos presentes no seu cotidiano, na sua vida, contribuindo, dessa forma, para o reforço e alargamento do repertório vocabular do educando. Tal unidade foi pensada para as turmas de 8ª série do ensino fundamental do EJA, contudo, nada impede que seja aplicada a outras séries e até mesmo a turmas do ensino regular.

5.2.2 Idéias de atividades para sala de aula

Parte 1: Preparação para a leitura, apresentação do texto do jornal Diário Gaúcho, leitura (individual e em grupo), interpretação, seleção de palavras.

★ Preparação para a leitura: para iniciar, é feita uma sondagem sobre o hábito da leitura e o hábito de leitura de jornal dos alunos. Eles comentam o que preferem/gostam de ler e, no caso dos jornais, qual(is) preferem e o porquê. Sobre o Diário Gaúcho, os alunos são questionados se lêem (quantos lêem) e o que acham dele.

★ Apresentação do texto do jornal Diário Gaúcho (ANEXO H): discussão sobre o título '*Morte em micro-ônibus*' (a partir dele, do que os alunos acham que se trata o texto) e sobre figuras.

★ Leitura: individual (cada aluno lê silenciosamente) e em grupo (cada aluno lê um trecho do texto para a turma, exercitando, assim, a oralidade).

★ Interpretação: perguntas sobre o texto são colocadas no quadro (ANEXO I), os alunos têm um tempo determinado para responder/pensar sobre as questões que, depois, serão discutidas no grande grupo.

★ Seleção e marcação de palavras: após a discussão sobre a reportagem, cada aluno destacará, em seu texto, quais palavras não fazem parte do seu uso cotidiano, do meio onde vivem e das pessoas com as quais convivem, ou seja, as palavras 'diferentes', com as quais não estão acostumados.

OBS: pedir para que os alunos tragam minidicionários de Língua Portuguesa e incentivar os que têm dicionário a carregarem sempre consigo, não só para as aulas de Português.

Parte 2: Retomada do texto e das palavras selecionadas. Uso do dicionário

- ★ Retomada oral do assunto do texto.
- ★ As palavras selecionadas por cada aluno são expostas ao grande grupo e, com elas, os alunos fazem as atividades.
- ★ Atividade 1: Você selecionou palavras do texto que não são do uso comum/freqüente no seu cotidiano. Agora, pense sobre elas: você sabe o que tais palavras significam? É possível compreendê-las por meio do texto/da frase do texto em que ela(s) aparece(m)? Tente explicá-las, escrevendo um significado para essas palavras.
- ★ Atividade 2: Agora observe o que o dicionário traz sobre as palavras que você selecionou, qual a classificação gramatical, quais os sentidos possíveis, etc. Selecione a definição que o dicionário traz que mais se encaixa no contexto em que a palavra aparece.
- ★ Atividade 3: Compare sua explicação sobre as palavras e o que você selecionou sobre elas no dicionário: os significados são semelhantes? Ou diferem? No que? Você acha que é possível entender alguma palavra desconhecida por meio do contexto em que ela aparece? Por quê?
- ★ Os alunos realizam as atividades e depois elas são discutidas no grupo, entre colegas e professor, estabelecendo uma conversa sobre vocabulário, uso, contexto, significado das palavras, uso do dicionário, etc.
- ★ Tema de casa: encontrar novos/diferentes contextos para as palavras analisadas.

OBS: usar dicionários da biblioteca, caso os alunos não tenham/tragam.

Parte 3: Retomada do texto e das palavras trabalhadas. Discussão no grupo dos novos contextos trazidos pelos alunos. Comparação de notícias sobre o mesmo assunto publicadas no DG e no ZH. Proposta de produção.

- ★ Retomada do texto e das palavras trabalhadas de maneira oral.

★ Discussão no grupo dos novos contextos trazidos pelos alunos: cada aluno conta/lê os contextos em que encontrou as palavras trabalhadas em aula. O professor leva alguns contextos, para o caso dos alunos não levarem, já que se trata de um público adulto, que normalmente tem diversas atribuições e pode esquecer os deveres escolares. Por meio dessa atividade, aparecerão diferentes contextos com as mesmas palavras, demonstrando aos alunos os usos e os sentidos possíveis de cada uma delas.

★ Entre o material que o professor leva (diversos contextos para as palavras selecionadas pelos alunos) está uma reportagem sobre o mesmo assunto do texto do DG analisado, publicada por um jornal tradicional: Zero Hora (ANEXO J). A idéia é que os alunos, assessorados pelo professor, leiam os dois textos, observando as semelhanças e diferenças de escrita, de vocabulário, de enfoque, etc. Dessa maneira, os alunos perceberão como se dá a transposição do texto de um jornal popular para um jornal tradicional.

★ Proposta de produção: Faça você mesmo! Agora que você conheceu palavras diferentes e estudou sobre elas, observando seus possíveis sentidos e seus diferentes contextos, crie uma propaganda/um anúncio usando tais palavras. Seja criativo!

Parte 4: Retomada das palavras estudadas, análise e discussão das freqüências das palavras presentes no texto, apresentação indireta da Lingüística de Corpus, exposição do Projeto PorPopular.

★ Retomada das palavras estudadas: os alunos podem ler e expor para os colegas o que produziram com aquelas palavras ‘diferentes’, selecionadas na primeira parte.

★ Atividade 1: Análise das freqüências das palavras presentes naquele texto: os alunos contam as palavras e observam quais se repetem mais e quais quase não aparecem.

★ A partir disso, são discutidas as palavras mais e menos freqüentes naquele texto, suas funções, características e porque umas aparecem mais que as outras. A idéia é que os alunos percebam as características distintas de elementos lexicais e de elementos gramaticais (sem necessariamente denominá-las desta forma) e que compreendam que ambos fazem parte do léxico da língua portuguesa, embora tenham *status* diferenciados. A professora leva a *wordlist* (ANEXO L) para

sistematizar a contagem e os dados levantados pelos alunos, explicando que isso que eles fizeram manualmente pode ser feito por um programa de computador, caso o texto esteja digitalizado.

★ As palavras mais freqüentes na língua portuguesa (fala x escrita), por meio de dados do Banco do Português (ANEXO M), também são comentadas com os alunos.

★ Então, é introduzida, indiretamente, a Lingüística de Corpus. A idéia é explicar para os alunos que uma grande quantidade de textos semelhantes aos que eles estudaram em aula, por exemplo, podem compor um *corpus*. E um *corpus* pode servir para diversos estudos/fins: ensino de línguas, produção de grandes dicionários, gramáticas e livros didáticos - sobretudo de língua inglesa (SARDINHA, 2000, p.329), desenvolvimento de ferramentas computacionais, etiquetadores eletrônicos, informatização de grandes bases de dados, (SARDINHA, 2000, p.329), corretores ortográficos e resumidores automáticos como os do *Word* (SARDINHA, 2004b, p.02), além dos variados tipos de pesquisas e descrições da linguagem. O propósito é que eles que conheçam a ciência que lida com isso (LC), como ela se desenvolve (com auxílio do computador, de programas, de *softwares*), e para o que serve, sem necessariamente utilizar a terminologia da área.

★ Por fim, o Projeto PorPopular é exposto aos alunos: é explicado que o texto do DG trabalhado em aula faz parte do material de estudo da pesquisa e, além disso, é apresentado o objetivo do Projeto, o que já foi estudado, quais os resultados obtidos e, ainda, quais as perspectivas de estudos futuros do Projeto. O *site* do PorPopular é colocado no quadro para que os interessados anotem e algumas possibilidades de exploração da pesquisa e do *corpus on-line* são comentadas com alunos, que poderão visitar o *site* caso tenham interesse e possibilidade de acesso à internet.

Com tais atividades e, sobretudo, com a discussão final, espera-se que os alunos percebam que se pode aprender palavras novas com os jornais e, inclusive, com um jornal popular como o DG.

Além disso, apresentando as palavras com variados contextos, assumindo diferentes sentidos, acredita-se que o aluno compreenda que os vocábulos da Língua Portuguesa não têm um sentido único, mas sim variam conforme a situação de uso. Nesse sentido, objetiva-se alargar o repertório lexical já trazido pelo aluno

que, ao pesquisar sobre as palavras, seus diversos significados e usos, e produzir com elas, mais facilmente as incorporará ao seu vocabulário.

A partir da apresentação dos diferentes padrões de freqüência das palavras, imagina-se que os alunos percebam que há diferenças entre elas, seja de função, de freqüência, de características, etc.

Por fim, ao incentivar a busca por contextos e a produção dos alunos, presume-se que eles tragam elementos do seu cotidiano para sala de aula. Dessa maneira, os alunos tendem a reconhecer a importância dos conhecimentos adquiridos e a presença deles nas suas vidas.

5.2.3 Opções para exploração do *corpus* DG em ambientes digitais

Assim como a temática do ensino de vocabulário foi proposta para sala de aula nos moldes tradicionais, utilizando elementos habituais como quadro, giz, caderno, texto impresso, fotocópias, etc., também pretendemos expor idéias para exploração de vocabulário *on-line*.

Em um segundo momento, então, após as atividades realizadas em sala de aula, a idéia é que, sendo possível na realidade da escola, os alunos tenham a oportunidade de vivenciar atividades diferentes daquelas feitas em aula. Como sabemos que nem todas as escolas públicas contam com sala de informática e computadores disponíveis, sobretudo para os alunos da noite, entendemos que esse tipo de proposta nem sempre é realizável nesses ambientes. Entretanto, apresentamos a idéia, pois entendemos que ambientes virtuais constituem uma fonte riquíssima de aprendizagem, que precisa ser explorada com todo tipo de aluno, incluindo os estudantes do EJA.

A pesquisa PorPopular, apresentada no primeiro capítulo deste trabalho, conta com um *site*: http://www6.ufrgs.br/textecc/index_porpopular.php. Esse *site* oferece dados da pesquisa e ambiente de estudo, no qual está amostra *on-line* do *corpus* DG organizado pelo grupo de pesquisadores até o momento.

A primeira parte do *site* apresenta dados dos projetos, dos parceiros, da equipe que integra a pesquisa, além de objetivos, resultados obtidos, dificuldades

passadas e perspectivas de estudo. A segunda parte, que é o ambiente de estudo, é a que permite explorar os textos do DG.



Figura 06: Tela inicial do site PorPopular

O site oferece, no ambiente de estudo, a sessão *experimente* (destacada na Figura acima), com diversas possibilidades de investigação do *corpus*. As duas primeiras opções são para visualizar o objeto de estudo: o jornal Diário Gaúcho, na sua versão de texto e na sua versão com imagens (assim como vai para as bancas), respectivamente. As três opções seguintes trazem as ferramentas de exploração do *corpus*: lista de palavras, gerador de contextos e gerador de *n-gramas*. Com estas, os alunos podem de fato experimentar o *corpus*, analisando seu vocabulário, as palavras mais presentes, as combinações frequentes e os contextos em que estas aparecem. Nesse sentido, para alunos da 8ª série do EJA de escolas públicas com ambientes informatizados, pensamos nas seguintes atividades, realizadas e debatidas em sala de aula entre professor e alunos:

Atividade 1: Analisando palavras

Assim como foi feito em sala de aula, agora você contará como auxílio de um programa de computador, disponível *on-line* pelo site PorPopular, para analisar as palavras de um texto. No item 3 do *menu* *experimente* do site (http://www6.ufrgs.br/textecc/index_porpopular.php), você abrirá uma janela: 3

Criando uma lista de palavras. Nela, selecione a ordem da lista por frequência, e escolha o *corpus* de 1 dia de jornal: 14/01/2008. Clique em 'entrar' e o sistema apresentará uma lista de palavras por ordem de frequência daquele dia do Diário Gaúcho (ANEXO N).

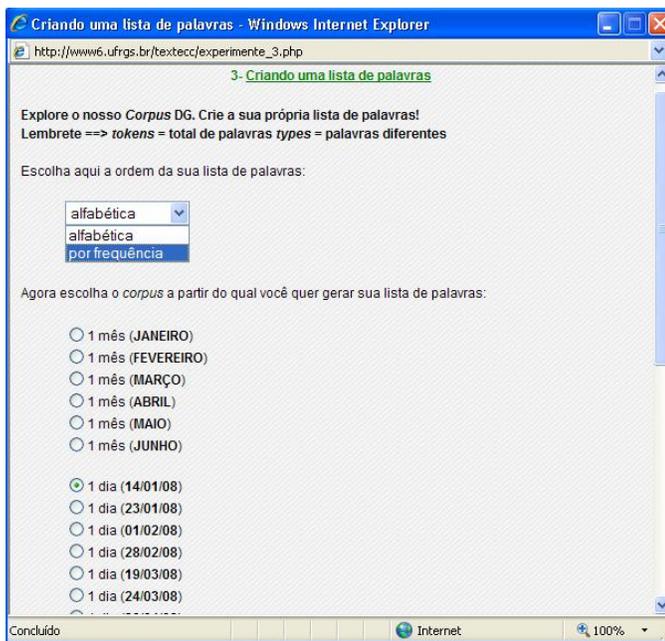


Figura 07: Tela da ferramenta geradora de listas de palavras do *site* PorPopular

Parte 1: Agora, a partir dessa lista, pense sobre as questões:

- Quais palavras são mais frequentes na página do DG analisada?
- De que tipo são essas palavras?
- Porque você acha que isso acontece?
- Você acha que tais palavras são frequentes somente nos jornais ou em outros textos também?
- Tente escrever um pequeno texto usando o mínimo possível dessas palavras. Agora conte as palavras do texto. Quantas são semelhantes às palavras mais presentes no Diário Gaúcho? Elas aparecem mais ou menos que as outras palavras?

Parte 2: Utilizando a mesma lista de palavras criada para o exercício anterior, responda:

- Qual a palavra de conteúdo mais frequente? Em que posição ela aparece?
- Quais as outras palavras de conteúdo que aparecem entre as 40 palavras mais frequentes?

c) O que você sabe sobre essas palavras?

d) Elas diferem daquelas analisadas no exercício anterior? No quê?

Atividade 2: Analisando contextos

No item 4 do *menu* experimente do *site* (http://www6.ufrgs.br/textecc/index_porpopular.php), você abrirá uma janela: 4 Buscando palavras – gerador de contextos. Nela, escreva a palavra ‘de’. Selecione 5 palavras antes e após a sua busca e o dia 14/01/2008.

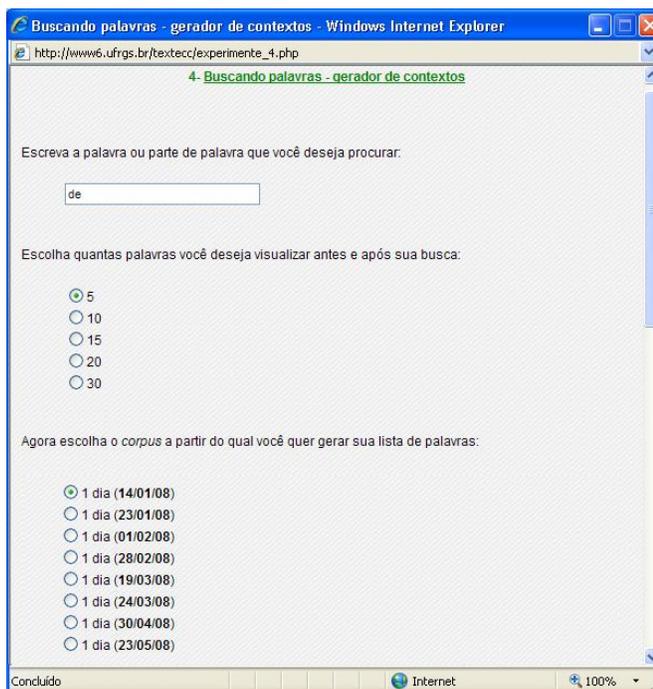


Figura 08: Tela da ferramenta gerador de contextos do *site* PorPopular

Faça o mesmo, trocando apenas a palavra ‘de’ por ‘não’. Com os dois resultados (ANEXOS O e P), reflita:

a) Quantas ocorrências de cada palavra você encontrou?

b) Onde essas duas palavras ocorrem mais: no início, no meio ou no fim das frases?

c) Existe alguma ocorrência que apareça em início da frase? Qual ou quais?

d) Entre estas de início de frase, alguma construção se repete? Você acha que essa pode ser uma característica do jornal?

e) Escreva duas frases, uma com ‘de’ e outra com ‘não’ (diferente das encontradas no jornal), em que elas apareçam no início da frase.

f) Agora escreva duas frases, uma com ‘de’ e outra com ‘não’ (diferente das encontradas no jornal), em que elas apareçam no meio da frase.

g) Qual dupla de frases você teve mais facilidade para criar?

h) A dupla de frases que você criou com mais facilidade é a mesma que aparece bastante no Diário Gaúcho? Por que você acha que isso aconteceu?

Atividade 3: Analisando grupos de palavras

No item 5 do *menu* experimente do *site* (http://www6.ufrgs.br/textecc/index_porpopular.php), você abrirá uma janela: 5 Gerador de *n-gramas*. Escolha o dia 14/01/2008 e a dimensão de tamanho nº3. Da mesma forma que com as outras ferramentas, o sistema apresentará uma lista de combinações de três palavras por ordem de frequência daquele dia do Diário Gaúcho (ANEXO Q).

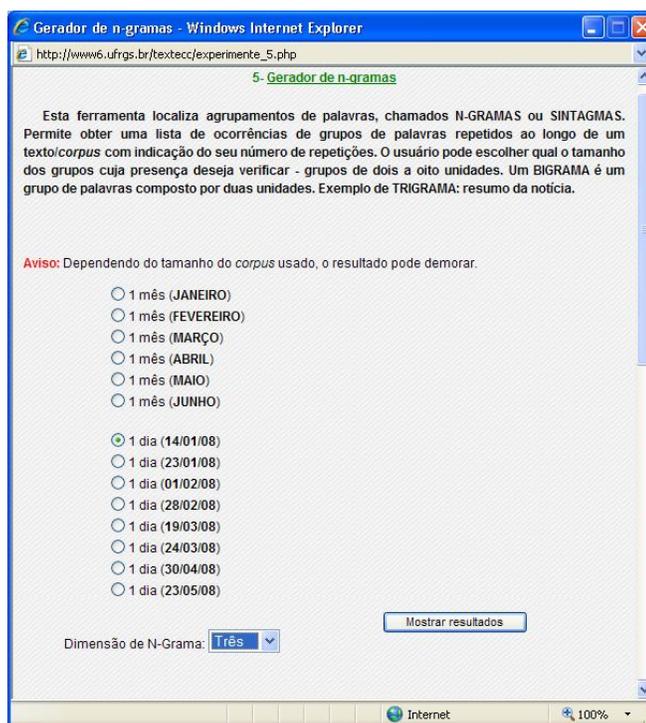


Figura 09: Tela da ferramenta geradora de *n-gramas* do *site* do PorPopular

A partir dela, opine:

- Qual a combinação de três palavras você encontrou como mais freqüente?
- Ela apresenta alguma das palavras mais recorrentes no jornal ('de' ou 'não')?
- Porque você acha que isso acontece?
- Entre as combinações encontradas, qual a que mais se repete contendo 'de'?
- Entre as combinações encontradas, qual a que mais se repete contendo 'não'?

f) Entre as combinações encontradas, qual a que mais se repete contendo as duas palavras: 'de' e 'não'? Em qual posição ela aparece (entre as primeiras, mais frequentes ou não?) Essa combinação se repete muito?

g) Escreva um parágrafo contendo essa combinação de três palavras encontrada na letra f (que mais se repete contendo as duas palavras: 'de' e 'não') dizendo o que você pensa sobre atividades escolares realizadas em ambientes digitais.

Como tais atividades são uma continuação das idéias pensadas para sala de aula, exposta no item anterior, os objetivos são os mesmos, sobretudo a percepção do aluno em relação às diferentes palavras que compõem o nosso vocabulário e, por isso, a maior ou menor utilização destas, além da questão do uso, abordada por acreditarmos que apresentando palavras com variados contextos, o aluno perceberá que os sentidos dos vocábulos variam de acordo com a situação, não tendo sentido único. Ao final, sugerimos colher a opinião do aluno do EJA sobre a questão de atividades escolares em ambientes digitais, porque entendemos que a inclusão digital também deve integrar o ensino para jovens e adultos, então, pensamos em conhecer a opinião desses alunos sobre o assunto.

Por fim, é importante salientar que as atividades idealizadas para ambientes informatizados podem ser realizadas em sala de aula também, uma vez que o professor pode levar fotocopiadas as listas geradas pelas ferramentas do *site* e, da mesma forma, realizar as atividades em sala de aula com seus alunos.

6 CONSIDERAÇÕES FINAIS

O presente trabalho buscou apontar opções de como tratar da temática do ensino de vocabulário com aluno do EJA, aproveitando o jornal popular e a visão de *corpus*.

Conforme pesquisado e exposto neste estudo, o ensino de vocabulário nas escolas, atualmente, está centrado na associação das palavras desconhecidas com outras semelhantes no sentido (sinônimos) ou na forma (raiz, prefixo ou sufixo), e também com vocábulos contrários (antônimos) ou imagens. Essa prática é recorrente em exercícios que apresentam palavras ou frases descontextualizadas. Nos poucos momentos em que o contexto se faz presente, são freqüentes os glossários ao final dos textos ou dos livros didáticos, apresentando ao aluno o significado dos vocábulos desconhecidos já pesquisado.

Tais práticas nos levaram a questionar o espaço do ensino de vocabulário nas aulas de Língua Portuguesa, bem como a apresentar propostas que envolvessem o empenho do aluno na busca por um significado, o uso do dicionário - auxiliado pelo professor, a questão do contexto e dos diversos significados que um vocábulo pode adquirir e, sobretudo, a percepção disto por parte do educando.

Para tanto, buscamos os referenciais teórico-metodológicos da Lingüística de Corpus, que acredita que o significado de uma palavra se configura pelo seu contexto de uso, observando, portanto, o uso real das palavras.

O público-alvo de nossa proposta foram os alunos do Ensino de Jovens e Adultos das escolas públicas, pois nosso estudo pretendeu contribuir para a diminuição da lacuna de estudos existentes sobre a educação de jovens e adultos e, sobretudo, acerca do desenvolvimento de recursos didáticos para esses alunos.

Utilizamos como material didático para as atividades propostas o jornal Diário Gaúcho, por ser de fácil acesso, por apresentar notícias e fatos do cotidiano e por dialogar com nosso público-alvo: os alunos do EJA.

O estudo alcançou seus objetivos, pois as idéias de atividades que contribuíssem para alargamento e reforço do vocabulário dos alunos do EJA, a partir do jornal popular e da visão de *corpus*, foram propostas, tanto na modalidade tradicional de sala de aula, quanto na opção para exploração de vocabulário em ambiente virtual de aprendizagem. Isso mostra que é possível realizar atividades

baseadas em *corpora* em sala de aula, sem necessariamente o auxílio de tecnologias. Claro que, assim como com as primeiras pesquisas baseadas em *corpora*, realizadas manualmente, pode haver equívoco com relação aos dados estatísticos, entretanto, isso não afeta diretamente as atividades de ensino, já que o essencial é que o aluno perceba as palavras, a quantidade delas, suas freqüências, combinações e contextos de uma maneira geral. Isso é o que vai fazê-lo compreender a diferença entre os vocábulos, seus sentidos, suas características, suas funções, etc.

Sobre as atividades propostas para ambientes informatizados, sabemos que, infelizmente, nem todas as escolas públicas contam com sala de informática e computadores disponíveis, sobretudo, aos estudantes da noite. Mesmo assim, propusemos atividades em ambientes virtuais porque entendemos que a inclusão digital também deve integrar os objetivos do EJA, pois não deixa de ser uma tentativa de reinserção desses indivíduos no mundo, já que, hoje em dia, a tecnologia está presente em praticamente tudo. Além disso, tais idéias podem ser realizadas em sala de aula também, uma vez que o professor pode levar fotocopiadas as listas geradas pelas ferramentas do *sítio* e realizar as atividades da mesma forma com os seus alunos.

Então, apesar das atividades de ensino de vocabulário propostas não terem sido efetivamente testadas nas turmas de 8ª série do EJA, é possível inferir que tanto as atividades de sala de aula, quanto as atividades previstas para ambientes informatizados contribuem para o alargamento do repertório lexical do aluno. Ao pesquisar sobre as palavras, seus diversos significados e usos, e, principalmente, ao produzir com elas, mais facilmente o aluno as incorporará ao seu vocabulário.

Conforme Dargel (2005):

Acredita-se que a partir da compreensão do sentido de uma palavra no texto, o aluno, seguindo orientações do professor, poderá perceber outros sentidos, ou oposições, que essa palavra pode ter em outros contextos diferentes daquele que foi estudado no texto. Assim, esse aluno estará ampliando seu conhecimento lexical. (DARGEL, 2005, p.01)

Fica, então, como perspectiva, a aplicação da unidade de ensino proposta. Por meio da efetiva realização das atividades perceberemos se o público imaginado como público-alvo está de acordo com as propostas ou se necessita ser repensado. Além disso, a partir da prática em sala de aula, observaremos quais mudanças ou

adaptações precisam ser feitas a fim de aperfeiçoar nossa unidade de ensino de vocabulário que aproveita o jornal popular e a visão de *corpus*.

Em suma, tratar da temática do ensino de vocabulário à luz dos conceitos da Lingüística de Corpus possibilita a observação do uso real das palavras. Dessa maneira, o aluno compreende o sentido das palavras de forma plena, passando “a realmente se apropriar [...] dos inúmeros conjuntos de palavras que constituem sua língua” (SILVA, 2005, p.69), o que contribui significativamente para o alargamento do conhecimento do mundo de uma pessoa (DARGEL, 2005).

Por fim, é importante ressaltar que a educação de jovens e adultos tem sido pensada com a finalidade de realmente permitir o acesso à educação, independentemente da idade. Nesse sentido, muito já foi feito, mas sempre há mais o que fazer, visto que, segundo Lopes e Souza “Não se pode acomodar com os avanços já conseguidos, é necessário vislumbrar novos horizontes na busca da total erradicação do analfabetismo em nosso país, pois a educação é direito de todos” (LOPES e SOUZA, 2005, p.09-10).

REFERÊNCIAS

AMARAL, M. F. **Jornalismo Popular**. São Paulo: Contexto, 2006.

BARBEIRO, L. F. Diversidade Lingüística na **Escola Portuguesa**. In: **Projecto Diversidade Lingüística na Escola Portuguesa (ILTEC)**. Lisboa: [s.n., entre 2003 e 2007]. Disponível em: <http://www.iltec.pt/divling/_pdfs/cd2_vocabulario_quadro-geral.pdf> Acesso em: 30 de maio de 2011.

BASILIO, M. **Formação e classes de palavras no português do Brasil**. São Paulo: Contexto, 2004.

BASILIO, M. **Teoria Lexical**. São Paulo: Ática, 2005.

BERBER SARDINHA, T. Lingüística de Corpus: histórico e problemática. **D.E.L.T.A.**, [São Paulo], v.16, n°02, p.323-367, 2000.

BERBER SARDINHA, T. **Lingüística de Corpus**. Barueri-SP: Manole, 2004a.

BERBER SARDINHA, T. Lingüística de Corpus: uma entrevista com Tony Berber Sardinha. **Revista Virtual de Estudos da Linguagem – ReVEL**, [S.l.] v.02, n°03, agosto de 2004b. ISSN 1678-8931 Disponível em: <http://www.revel.inf.br/site2007/_pdf/3/entrevistas/revel_3_entrevista_tony_berber_sardinha.pdf> Acesso em: 04 de julho de 2011.

BIDERMAN, M. T. **Teoria Lingüística (lingüística quantitativa e computacional)**. Rio de Janeiro: Livros Técnicos e Científicos, 1978.

BIDERMAN, M. T. A FACE QUANTITATIVA DA LINGUAGEM: UM DICIONÁRIO DE FREQUÊNCIAS DO PORTUGUÊS. **Alfa**, São Paulo, v.42, número especial, p.161-181, 1998.

BIDERMAN, M. T. Conceito lingüístico de palavra. In: BASILIO, M. (org). **Palavra**. Rio de Janeiro: Grypho, 1999. v.1, p.81-97.

BISOGNIN, T. R. **Sem medo do Internetês**. Porto Alegre: AGE, 2009.

BRASIL. Constituição (1988). Constituição da República Federativa do Brasil. Brasília, DF: Senado Federal, 1988. Capítulo III – Sessão I. Disponível em: <<http://www.soleis.com.br/ebooks/1-Constituicoes5.htm#89>> Acesso em: 16 de maio de 2011.

BUENO, E; TAVARES, F; PAZ, J. Estatística da Palavra: Conceito, Palavras Gramaticais e Palavras Lexicais. In: **Acadêmica de letras: Língua Portuguesa – Literatura Brasileira – Literatura Francesa – Prática da Língua e Cultura Francesa – Revisão e Editoração de Textos**. [S.l.: s.n.], 2009. Disponível em: <<http://jps-ltf-eterna-estudante.blogspot.com/2009/01/estilstica-da-palavra-conceito-palavras.html>> Acesso em: 30 de maio de 2011.

CAMARA, J. M. **Estrutura da Língua Portuguesa**. Rio de Janeiro: Vozes, 1982.

CEGALLA, D. P. **Novíssima Gramática a Língua Portuguesa**. Companhia Editoria Nacional: São Paulo-SP, 2001. 34ª edição.

DARGEL, A. P. T. P. A utilização do dicionário no ensino do léxico. **Ave Palavra (UNEMAT)**, n°06, p. 58-68, 2005. Disponível em: <<http://www2.unemat.br/avepalavra/EDICOES/06/artigos/DARGEL.pdf>> Acesso em: 19 de maio de 2011.

DUBOIS, Jean *et al.* **Dicionário de Lingüística**. São Paulo: Cultrix, 1978.

FINATTO, M. J. B. **PADRÕES DO PORTUGUÊS POPULAR ESCRITO: O VOCABULÁRIO DO JORNAL DIÁRIO GAÚCHO. FASE 1**. Projeto de Pesquisa contemplado Edital MCT/CNPq, n°14/2009 – Universal. Novembro de 2009. Disponível em: <<http://www6.ufrgs.br/textecc/arquivos/PorPopularPropostaCNPqaprovada.pdf>> Acesso em: 02 de maio de 2011.

GERBER, R. M.; VASILÉVSKI, V. **Um percurso para pesquisas com base em corpus**. Florianópolis: Editora da UFSC, 2007.

HOUAISS, A. **Dicionário Eletrônico Houaiss**. Rio de Janeiro: Objetiva, 2001. 1CD-ROM.

LEPSCHY, G. Léxico. In: **Enciclopédia EINAUDI. Linguagem – Enunciação**. Lisboa: Imprensa Nacional – Casa da Moeda, 1984, v. 02, p.156-178.

LOCH, J. M. P. Planejamento e avaliação em EJA. In: BINS, K. L. G. *et al* (Org.). **EJA: planejamento, metodologias e avaliação**. Porto Alegre: Mediação, 2009.

LOPES, S. P. , SOUZA; L. S. EJA: UMA EDUCAÇÃO POSSÍVEL OU MERA UTOPIA? **Revista Alfabetização Solidária (Alfasol)**, v.05, 2005. Disponível em: <http://www.cereja.org.br/pdf/revista_v/Revista_SelvaPLopes.pdf> Acesso em: 16 de maio de 2011.

MACHADO, B. S. **O uso do dicionário e as atividades de pré-leitura: prática e análise exploratórias em sala de aula de língua estrangeira**. Rio de Janeiro, 2003. Dissertação (Mestrado em Letras), Departamento de Letras, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2003.

MACIEL, A. M. B. Lingüística de *Corpus*: bases teórico-metodológicas. In: COLÓQUIO NACIONAL DE LETRAS EM DIÁLOGO E EM CONTEXTO: rumos e desafios, 2002, Porto Alegre. **Anais...** Porto Alegre: Programa de Pós-Graduação em Letras, 2003. 1CD-Rom.

MANUAL DE REDAÇÃO DO DIÁRIO GAÚCHO. Agosto de 2005.

OLIVEIRA, I. B. As interfaces educação popular e EJA: exigências de formação para a prática com esses grupos sociais. **Educação. Faculdade de Educação, Programa de Pós-Graduação da PUCRS**. Porto Alegre, v.33, nº02, (mai/ago 2010). Porto Alegre: EDIPUCRS, 2010.

PAIVA, V. L. M. O. A gramática e o vocabulário no ensino de inglês: novas perspectivas. In: DUTRA, D.P & MELLO, H (Org.). **Ensino de vocabulário**. Belo Horizonte: Faculdade de Letras/UFMG, 2004. (Estudos Lingüísticos; 7). Disponível em: <<http://www.veramenezes.com/vocabulario.htm>> Acesso em:07 de abril de 2011.

RODRIGUES, G. ; PARREIRA, M. C. Lexicografia e o ensino de expressões idiomáticas da Língua Portuguesa. In: VII ENGTLEX Encontro Intermediário de Lexicologia, Lexicografia e Terminologia da ANPOLL, 2009, São José do Rio Preto-SP. **RESUMOS EXPANDIDOS**. São Paulo: GTLEX, 2009.
Disponível em:
<http://www.mel.ileel.ufu.br/gtlex/viiengtlex/pdf/resumos/Gislaine_Rodrigues.pdf>
Acesso em: 05 de julho de 2011.

ROSA, M. C. **Introdução à Morfologia**. São Paulo: Contexto, 2000.

SANTOS, V. Empréstimos lingüísticos: tradição e atualidade. In: **Idioma e Soberania - Nossa Língua, Nossa Pátria**. Câmara dos Deputados: Brasília, 2000. Disponível em: <<http://www2.camara.gov.br/a-camara/conheca/historia/historia/cdnos500anos/seminarios/semin2/fala7.html>>. Acesso em: 30 de maio de 2011.

SILVA, B. R. ; FINATTO, M. J. B. . PORTUGUÊS POPULAR ESCRITO: O VOCABULÁRIO DO JORNAL DIÁRIO GAÚCHO. In: X Salão de Iniciação Científica da PUCRS, 2009, Porto Alegre. **Anais do X Salão de Iniciação Científica da PUCRS**. Porto Alegre : EDIPUCRS, 2009. p. 3332-3334.

SILVA, S. Um esboço acerca do léxico em enunciação: concepção e consequências pedagógicas para o ensino fundamental. **Revista Língua e Literatura**. Frederico Westphalen, v.10/11, nº6/7, p.57-71 Frederico Westphalen: URI, 2005.

VICENTINI, G. **A Lingüística de Corpus e o seriado *Friends* como base para o ensino de *chunks* em sala de aula de Língua Inglesa**. São Paulo, 2006. Dissertação (Mestrado em Lingüística Aplicada e Estudos da Linguagem), Pontifícia Universidade Católica de São Paulo, São Paulo, 2006.

VIEIRA, M. C. **Fundamentos históricos, políticos e sociais da educação de jovens e adultos – Volume I: aspectos históricos da educação de jovens e adultos no Brasil**. Universidade de Brasília: Brasília, 2004.

ANEXO A – Reportagem *Efetivo Reduzido Investiga Crime*, do jornal *Zero Hora* (09/01/2009).

Efetivo reduzido investiga crime



A Polícia Civil reduziu o efetivo que investiga o assassinato do médico Marco Antonio Becker, ocorrido em 4 de dezembro. A investigação começou com uma força-tarefa de 27 policiais dedicados prioritariamente a tentar elucidar o crime. Passado mais de um mês da execução, o caso continua como prioridade na Delegacia de Homicídios, mas os agentes não atuam em tempo integral no episódio.

Policiais da 15ª DP (bairro Partenon), convocados para ajudar, agora são chamados esporadicamente para as apurações. Muitos agentes das duas delegacias foram deslocados ao Litoral e outros estão em férias, o que reduz a força-tarefa a um máximo de 15 homens. Dois dos delegados que estavam na proa das investigações, Rinaldo Vieira Junior e Omar Abud, estão de férias. Ontem, Rodrigo Bozzetto era o único delegado no caso, que continua sob sigilo de Justiça. Bozzetto deve entrar em licença paternidade neste final de semana.

ANEXO B – Reportagem *Menos policiais no Caso Becker*, do jornal *Diário Gaúcho* (09/01/2009).

Menos policiais no Caso Becker

A Polícia Civil reduziu o efetivo que investiga o assassinato do médico Marco Antonio Becker, ocorrido em 4 de dezembro. A investigação começou com uma força-tarefa de 27 policiais dedicados prioritariamente a tentar elucidar o crime. Passado mais de um mês da

execução, o caso continua prioridade na Delegacia de Homicídios, mas os agentes não atuam em tempo integral no episódio.

Muitos agentes foram deslocados ao Litoral (para a Operação Verão) e outros estão em férias, o que reduz a força-tarefa a um máximo

de 15 homens. Dois dos delegados que estavam na proa das investigações, Ranolfo Vieira Junior e Omar Abud, estão de férias. Ontem, Rodrigo Bozzetto era o único delegado no caso, que continua sob sigilo de Justiça. Bozzetto deve entrar em licença paternidade neste final de semana.

ANEXO C – Capa do jornal *Diarinho – Diário do Litoral* (19/01/2009)

Bom dia!



Como TEMPO Sol
mín. 22, máx. 23

di@rinho 29 anos
Diário do Litoral

Litoral Catarinense, Segunda - Feira, 19 de Janeiro de 2009

Caminhoneiro *enraba* Sílvio Brito
O cantor sofreu o acidente no mesmo dia em que deu entrevista ao DIARINHO



Fim da linha
Polícia Civil *grampeia* estelionatário que arrepiava a região



Moça morre em porradaço na BR-101
Na Penha, *carango* tromba em poste e fica de pentelhos pro ar



Prêmio da Mega Sena tá acumulado em 30 milhões

UM PONTINHO
Marinheiro começa o Catarinão empatando com o Metropolitano

CONFIRA NO TTS
Dicas pra não ser passado pra trás nas compras pela internet

FINAL DE SEMANA SANGRENTO

Dimenor morre a facadas em Navega e dois são baleados em Itajaí

ANEXO D - Manchete da capa do jornal *Meia Hora* (30/10/2008).



ANEXO E - Exemplos de exercícios de ensino de vocabulário

Faça o relacionamento dos antônimos e diga o que significam:

- (a) homogêneo () neologismo
(b) arcaísmo () policultura
(c) monocultura () microcéfalo
(d) macrocéfalo () heterogêneo

Assinale a idéia comum expressa pelas palavras de cada grupo:

a) telescópio – microscópio – endoscopia – radioscopia – autópsia

() tocar () ler () olhar, examinar () ouvir

b) antropônimo – homônimo – pseudônimo – topônimo – heterônimo

() homem () estudo () lei () nome

c) aristocrata – escravocrata – plurocrata – tecnocrata – autocrata

() domínio () dinheiro () arte () glória

(CEGALLA, D. P. **Novíssima Gramática a Língua Portuguesa**. Companhia Editoria Nacional: São Paulo-SP, 2001. 34ª edição. p.104)

ANEXO F - Exemplos de exercícios de ensino de vocabulário (2)

Festa de Aniversário

Adoro quando faço anos porque a minha mãe faz sempre muita comida boa. Nesse dia,

convido sempre os meus amigos para ir almoçar a minha casa. Começamos por comer um bom

_____  com _____ , _____  e _____ . Para

aqueles que não gostam de _____ , a minha mãe faz também peixe assado com

_____ . Depois do prato principal, comemos _____

_____  ou _____ . Os meus frutos preferidos são a _____  e as

_____ , mas no dia do meu aniversário prefiro comer as deliciosas sobremesas

que a minha mãe faz: _____ , arroz doce, _____ . Há

sempre muitos doces! Depois do almoço, eu e os meus amigos vamos brincar. À hora do lanche

comemos uns _____ , depois cantamos os parabéns e eu apago as velas do meu

_____ .

Pôr a Mesa

Lá em casa todos ajudamos a pôr a _____ . Começamos por pôr a _____ .

Depois pomos os _____ , os _____  e os talheres: o _____

_____  à esquerda do prato; a _____  e a _____  à direita.

Pomos também um guardanapo para cada um. Assim que a mesa está posta, a minha mãe traz o

_____  para a mesa, sentamo-nos nas _____  e

começamos a comer.

ANEXO G - Exemplos de exercícios de ensino de vocabulário (3)

E X E R C Í C I O S

1. No quadro 1 existem palavras que se relacionam com as do quadro 2 por serem sinônimas ou por serem antônimas. Copie-as, formando pares e identifique com **S** os sinônimos e com **A** os antônimos:

quadro 1	façanha	construir	entusiasmo
	debate	degustar	paulatino
	abandonar	definitivo	estupendo

quadro 2	amparar	discussão	formidável
	provisório	demolir	indiferença
	provar	proeza	rápido

(MESQUITA, R. B.; MARTOS, C. R. **Gramática Pedagógica**. Editora Saraiva: São Paulo-SP, 1998. 27ª edição. p.309)

ANEXO H – Reportagem *Morte em micro-ônibus*, do jornal *Diário Gaúcho* (16/06/2010)

EXECUÇÃO EM GUAÍBA

Morte em micro-ônibus

RESUMO DA NOTÍCIA

Dois homens entraram no ônibus que faz a linha Guaíba-Porto Alegre, disseram para os passageiros abaixarem as cabeças e atiraram em um jovem.

RENATO GAVA

renato.gava@claro.com.br

Uma execução ontem dentro de ônibus em Guaíba, levou a Brigada Militar a intensificar a fiscalização nos coletivos que circulam pela Avenida das Indústrias, uma das mais movimentadas do município. Nos últimos dois dias, três jovens foram alvejados dentro de coletivos perto do Bairro São Francisco. Um homem morreu e dois adolescentes ficaram feridos – um deles permanecia

ontem em estado gravíssimo no HPS da Capital.

● Bandidos tinham a vítima escolhida

– Nossa intenção é trabalhar mais forte nessas áreas. De dez dias para cá, começaram a ocorrer mais crimes nessa região. Vamos remanejar policiais e, em alguns momentos, haverá ações específicas nos ônibus – prometeu o comandante do

31º BPM, major Marcelo Mello.

Ontem, às 6h40min, dois homens ingressaram no micro-ônibus da linha Guaíba-Porto Alegre, pediram para os passageiros abaixarem a

cabeça e executaram Everton Dias de Azambuja, 19 anos. A vítima levou pelo menos dois tiros, na cabeça e no pescoço. Morreu antes da chegada da Brigada Militar.

Na segunda-feira, dois meninos de 16 anos haviam sido atingidos na mesma via, a cerca de 50m do local. Eles estavam em um ônibus quando dois homens entraram. Um dos adolescentes foi atingido na cabeça, na barriga e

no pescoço e, ontem, seguia na UTI do HPS, na Capital, em estado gravíssimo. O rapaz responde a quatro procedimentos de adolescente infrator por roubo, porte ilegal de arma e envolvimento



Everton

com drogas. Na ocorrência, outro jovem foi baleado de raspão na mão e liberado após ser atendido no hospital.

– Provavelmente, os crimes de segunda-feira e de hoje (ontem) estão relacionados, e o autor é o mesmo. Já identificamos o suspeito. É questão de tempo até que tenha sua prisão solicitada – informou o chefe de investigações da DP de Guaíba, Leonardo Gardel.

Lei do silêncio

Como nada foi roubado nos dois casos, a Polícia Civil não tem dúvidas de que se trata de uma execução. O motivo seriam rixas de gangues ligadas ao tráfico, que atuam na região dos Bairros Nova Cohab (nas proximidades da Cohab) e São

Francisco. Passageiros que estavam no ônibus começaram a ser ouvidos ontem.

– Está difícil obter informações, o pessoal tem muito medo de represálias – lamentou o chefe de investigações da DP.

COMO ACONTECEU

1. Dois homens ingressam no ônibus em um ponto de ônibus na Avenida das Indústrias, Bairro São Francisco.
2. Em seguida, cada um saca uma arma. Eles pedem para que os passageiros abaixem a cabeça, vão até a parte traseira do veículo e um deles dispara pelo menos três tiros na vítima.
3. Os dois saem do coletivo calmamente, caminhando. O jovem atingido, de 19 anos, morre antes da chegada da polícia.

ANEXO I - Perguntas de interpretação

- Sobre o que fala o texto?
- É possível prever sobre o que será lido a partir do título?
- Houve vítimas no caso descrito?
- Qual a atitude da polícia em relação ao acontecimento?
- A polícia faz alguma ligação deste com algum outro acontecimento? Qual? Por quê?
- Como aconteceu o crime que matou Everton Azambuja?

ANEXO J – Reportagem *Dois homens executam jovem dentro de ônibus e fogem a pé na Região Metropolitana*, do jornal Zero Hora (16/06/2010)

Dois homens executam jovem dentro de ônibus e fogem a pé na Região Metropolitana

Vítima é Everton Dias Azambuja, de 19 anos, que foi alvejado com disparos na cabeça e no pescoço.

Dois homens armados executaram com quatro tiros Everton Dias Azambuja, de 19 anos, por volta das 7h de hoje dentro de um ônibus que faz a integração Guaíba-Porto Alegre. O veículo circulava no sentido bairro-Centro, na Avenida das Indústrias, quando a cerca de 500 metros da BR-116, a vítima foi alvejada com disparos na cabeça e no pescoço.

Os criminosos fugiram a pé. O local do assassinato foi isolado e a Brigada Militar (BM) está fazendo buscas na região.

Na tarde de ontem, um crime parecido também ocorreu em Guaíba. um jovem de 16 anos estava dentro de um ônibus que faz linha interna, quando um homem o atingiu na cabeça, no abdômen e no pescoço. Ele foi encaminhado em estado gravíssimo ao HPS de Porto Alegre e está internado na UTI.

Outro adolescente, Henrique Martins Gonçalves, de 16 anos, levou um tiro de raspão, foi atendido e liberado.

ANEXO L - Wordlist texto Morte em micro-ônibus (DG 16/06/2010)

Tokens 396 - Types 226

19	de	3	foi	1	arma	1	gardel	1	outro
12	a	3	guaíba	1	às	1	h	1	ouvidos
12	e	3	homens	1	as	1	haverá	1	pediram
10	na	3	mais	1	até	1	havam	1	pela
9	o	3	os	1	atendido	1	hoje	1	pelo
8	da	3	para	1	atingido	1	homem	1	permanecia
8	dois	3	passageiros	1	atingidos	1	hospital	1	perto
8	em	3	uma	1	atiraram	1	identi	1	pessoal
7	no	2	abaixarem	1	atuam	1	ilegal	1	polícia
6	do	2	adolescentes	1	autor	1	indústrias	1	policiais
6	ontem	2	alegre	1	avenida	1	informações	1	por
6	que	2	anos	1	azambuja	1	informou	1	porte
5	ônibus	2	brigada	1	bairro	1	infrator	1	prisão
5	um	2	capital	1	bairros	1	ingressaram	1	procedimentos
4	nos	2	chefe	1	baleado	1	intenção	1	prometeu
3	cabeça	2	cohab	1	bandidos	1	intensi	1	provavelmente
3	dias	2	coletivos	1	barriga	1	já	1	proximidades
3	é	2	começaram	1	bpm	1	jovens	1	quando
3	execução	2	crimes	1	cá	1	lamentou	1	quatro
3	foi	2	das	1	cabeças	1	lei	1	questão
3	guaíba	2	dentro	1	camos	1	leonardo	1	rapaz
3	homens	2	dos	1	car	1	liberado	1	raspão
3	mais	2	dp	1	caram	1	ligadas	1	relacionados
3	os	2	entraram	1	cas	1	local	1	remanejar
3	para	2	estado	1	casos	1	m	1	represálias
3	passageiros	2	estavam	1	cerca	1	major	1	responde
3	uma	2	francisco	1	chegada	1	mão	1	resumo
2	abaixarem	2	gravissimo	1	circulam	1	marcelo	1	rixas
2	adolescentes	2	guaíba-porto	1	civil	1	medo	1	roubado
2	alegre	2	hps	1	co	1	mello	1	roubo
2	anos	2	investigações	1	com	1	meninos	1	scalização
2	brigada	2	jovem	1	comandante	1	menos	1	se
2	capital	2	levou	1	como	1	mesma	1	seguia
19	de	2	linha	1	deles	1	mesmo	1	seriam
12	a	2	micro-ônibus	1	dez	1	min	1	sido
12	e	2	militar	1	difícil	1	momentos	1	silêncio
10	na	2	morreu	1	disseram	1	morte	1	solicitada
9	o	2	pescoço	1	drogas	1	motivo	1	sua
8	da	2	região	1	dúvidas	1	movimentadas	1	suspeito
8	dois	2	são	1	eles	1	muito	1	tempo
8	em	2	segunda-feira	1	envolvimento	1	município	1	tenha
7	no	2	ser	1	escolhida	1	nada	1	tinham
6	do	2	tem	1	especi	1	não	1	tiros
6	ontem	2	vítima	1	está	1	nas	1	trá
6	que	1	ações	1	estão	1	nessa	1	trabalhar
5	ônibus	1	adolescente	1	everton	1	nessas	1	trata
5	um	1	alguns	1	executaram	1	nossa	1	três
4	nos	1	alvejados	1	faz	1	notícia	1	últimos
3	cabeça	1	antes	1	feridos	1	nova	1	uti
3	dias	1	ao	1	foram	1	obter	1	vamos
3	é	1	após	1	forte	1	ocorrência	1	via
3	execução	1	áreas	1	ganges	1	ocorrer		

ANEXO M - Palavras mais freqüentes na Língua Portuguesa segundo o Banco do Português.

	FALA	ESCRITA
Posição	Frequência do Banco do Português Falado	Frequência do Banco do Português Escrito
1	E (113.061)	DE (1.537.460)
2	QUE (108.883)	A (1.082.233)
3	A (77.882)	O (1.026.380)
4	É (75.609)	E (726.548)
5	O (71.329)	QUE (667.850)

(BISOGNIN, T. R. **Sem medo do Internetês**. Porto Alegre: AGE, 2009. p.164)

ANEXO N – Parte da *wordlist* do dia 14/01/2008 do site PorPopular

Lista de palavras		
Tokens: 14799		
Types: 4292		
Ordem	Ocorrências	Type
1	615	de
2	477	o
3	476	a
4	358	e
5	281	que
6	238	do
7	227	da
8	198	em
9	194	com
10	193	para
11	193	no
12	161	um
13	141	na
14	128	não
15	126	é
16	115	uma
17	97	os
18	96	foi
19	75	h
20	75	por
21	74	se
22	62	as
23	59	anos
24	53	das
25	52	dos
26	50	mas
27	45	ao
28	43	às
29	40	r
30	40	ele
31	39	à
32	36	pelo
33	36	como
34	36	mais
35	35	até
36	34	tem
37	34	dia
38	33	ou
39	32	tempo
40	31	sua

ANEXO O – Parte dos contextos da palavra ‘de’ do dia 14/01/08 do *site* PorPopular

Concordanciador
Total de concordâncias: 593
Donders O QUE HÁ DE NOVO INTERNAÇÃO O
vice-presidente da República, José Alencar, de 76 anos, voltou a ser
para o próximo sorteio é de R\$ 15 milhões.
pessoas morreram queimadas na explosão de um caminhão-tanque no Sul da
Chaves afirmou que a tentativa de golpe de Estado de 2002
que a tentativa de golpe de Estado de 2002 contra seu
tentativa de golpe de Estado de 2002 contra seu governo tinha
VÔO Mais de 5 mil pessoas aguardam viagem
aguardam viagem no aeroporto argentino de Ezeiza, em Buenos Aires. Passageiros
depredaram guichês e provocaram fuga de funcionários no terceiro dia da
FRIO Uma onda de frio matou 52 pessoas na
na última semana na província de Herat, no Afeganistão. O frio
pronunciamento à nação em cadeia de rádio e tevê ontem à
afirmar que não há risco de epidemia de febre amarela no
não há risco de epidemia de febre amarela no aís. O
morarem, ou visitarem, áreas de risco (florestas e
Não existe risco de epidemia garantiu o ministro.
o ministro. Apesar da existência de 24 notificações pelas secretarias de
de 24 notificações pelas secretarias de Saúde, sendo dois casos confirmados (um
o Brasil não tem casos de febre amarela desde 1942. Os
desde 1942. Os casos registrados de lá para cá foram de
de lá para cá foram de febre amarela silvestre, ou seja,
febre amarela silvestre, ou seja, de pessoas que contraíram a doença
deve procurar o posto de saúde se morar, ou for
for visitar, as chamadas áreas de risco (florestas e matas).
ou se foi vacinado antes de 1999. Tomando
após dez dias. Em janeiro de 2008, foram enviados 3,23 milhões
2008, foram enviados 3,23 milhões de doses da vacina contra febre
A vacina oferece dez anos de proteção e o seu reforço
segundo Estado com maior número de universidades federais. A posição será
assumida hoje, com a oficialização de duas instituições.
transformação da Fundação Faculdade Federal de Ciências Médicas em Universidade Federal
Ciências Médicas em Universidade Federal de Ciências da Saúde de Porto
Federal de Ciências da Saúde de Porto Alegre. ÁRVORE
levou um susto na noite de sábado. Uma árvore caiu em
muro e derrubou um poste de luz após uma forte ventania.
17 mil em um posto de gasolina de Marília (SP) devolveu
em um posto de gasolina de Marília (SP) devolveu o dinheiro
Samba. Mais uma vez, milhares de crianças vão tomar conta da
Samba, desenvolvido pela Secretaria Municipal de Esportes, Recreação e Lazer (SME),
todo ano monta uma escola de samba mirim. Projeto
dos ensaios Aos sábados, de forma itinerante, eles comandam os
comandam os ensaios da escola de samba mirim. Optamos sempre
Optamos sempre por uma quadra de escola próxima das comunidades que

ANEXO P – Parte dos contextos da palavra ‘é’ do dia 14/01/08 do site PorPopular

Concordanciador
Total de concordâncias: 100
Chapéu do Sol. Hoje é dia do enfermo O
prêmio para o próximo sorteio é de R\$ 15 milhões.
reforço durante este período "não é necessário, nem recomendado". UNIVERSIDADES
e pela cultura. Nosso propósito é dar uma integração maior com
Bosco Vaz. Sábado é dia dos ensaios Aos
brutal, inexplicável. O que apavora é que a maldade deixou de
do poço se aproxima, se é que já não chegamos lá.
O valor das inscrições é de R\$ 36,45 para os
para os ônibus cujo embarque é feito pela porta da esquerda
o número de viagens diárias é reduzido em 10,3%. Caiu
em alguns bairros. A previsão é de que o abastecimento seja
devolvido ao mar. A espécie é originária do Sul da Argentina.
do povo". O programa preferido é à beira-mar, mas já é
é à beira-mar, mas já é comum o visitante "tirar o
banho de lagoa. Da praia é só seguir 2,5km para se
Quem já foi diz que é uma grande festa! Como
começo, sem dúvida. O trabalho é longo e tomara que, neste
lembrando que o Flamengo é um time amador e também
clube. Sua saída do Zaragoza é tida como "inviável", pelos R\$
argentino deixar o time espanhol é quase nula. SOUZA
Grêmio (onde seu salário aumentaria) é o trunfo do assessor Paulo
uma melhor avaliação. No entanto, é notório o desentrosamento entre os
há uma nova formação e é o primeiro jogo da temporada.
Porto Alegre. Meta é impedir o tri do tradicional
informação de que Santa Maria é um caldeirão em janeiro, surpreendeu.
Braga apostou em larley. Sorte é que, até a estréia no
Pato foi natural e isto é suficiente para empolgar. O garoto,
fora de forma, mostrou que é o Fenômeno. Jogou muito, fez
fez gol ele que é apontado como sucessor natural de
deles (companheiros). O verdadeiro Fenômeno é o Ronaldo disse. Esta
Gargano, que comete falta e é expulso GUERRA TOTAL
garotada do time júnior. Tanto é verdade que o próprio treinador
time de qualidade. O problema é que como alguns negócios não
em Renato, D'Alessandro e Souza é pensar alto. É disso que
e pela Seleção de Togo, é um dos reforços do Veranópolis.
bem ao Gauchão. Nossa meta é ficar na Primeira Divisão
preferem ainda não dizer qual é o problema de saúde.
para banco de sangue. Sangue é vida. Se todas as pessoas
tive oportunidade. Mas acho que é muito importante fazer isso. Sueli
querer afirmar que o mundo é problemático e não o indivíduo
cérebro. A inteligência que temos é mecânica, herdada da imitação, portanto,
impressão de felicidade. Assim também é a vida. José da Cunha

ANEXO Q – Parte dos *tri-gramas* do dia 14/01/2008 do *site* PorPopular

N-gramas de tamanho 3	
8	às h
8	por volta das
7	volta das h
7	resumo da notícia
6	de porto alegre
6	a partir das
6	partir das h
6	dia às
5	madrugada de sábado
5	das h de
5	até as h
5	às h min
4	ingressos a r
4	h min de
4	de acordo com
4	norte do estado
4	até de janeiro
4	no litoral norte
4	no segundo tempo
4	das h min
4	em porto alegre
4	que o time
4	pelo site www
4	no dia de
3	de febre amarela
3	passar o tempo
3	h ao meio-dia
3	acordo com a
3	escola de samba
3	o tempo passa
3	na madrugada de
3	internado no hospital
3	na praia de
3	no próximo dia
3	era uma pessoa
3	h às h
3	dia de fevereiro
3	porto alegre
3	show de música