

**Vanderlei dos Santos**

**VOCABULÁRIO E LEITURA:**

**A ELABORAÇÃO DE UMA LISTA DE PALAVRAS DE USO ACADÊMICO EM  
PORTUGUÊS DO BRASIL**

**MESTRADO**

**UNIVERSIDADE ESTADUAL DE CAMPINAS**

**Instituto de Estudos da Linguagem**

**2006**

Este exemplar e a redação final da tese  
defendida por Vanderli de

Santos

é aprovada pela Comissão Julgadora em

29/09/2006

Mário R. Guimarães

**Vanderlei dos Santos**

**VOCABULÁRIO E LEITURA:**

**A ELABORAÇÃO DE UMA LISTA DE PALAVRAS DE USO ACADÊMICO EM  
PORTUGUÊS DO BRASIL**

Dissertação apresentada ao Departamento de Lingüística Aplicada do Instituto de Estudos da Linguagem da Universidade Estadual de Campinas, como requisito parcial para a obtenção do grau de Mestre em Lingüística Aplicada na área de Língua Estrangeira.

Orientadora:  
Prof<sup>a</sup>. Dr<sup>a</sup>. Matilde Virgínia Ricardi Scaramucci

**UNIVERSIDADE ESTADUAL DE CAMPINAS**

**Instituto de Estudos da Linguagem**

**2006**

**Ficha catalográfica elaborada pela Biblioteca do IEL - Unicamp**

Sa59L Santos, Vanderlei dos.  
Leitura e vocabulário: a elaboração de uma lista de palavras de uso acadêmico em português do Brasil / Vanderlei dos Santos. -- Campinas, SP : [s.n.], 2006.

Orientador : Matilde Virgínia Ricardi Scaramucci.  
Dissertação (mestrado) - Universidade Estadual de Campinas, Instituto de Estudos da Linguagem.

1. Vocabulário - Estudo e ensino. 2. Língua estrangeira - Leitura. 3. Língua portuguesa - Estudo e ensino - Falantes estrangeiros. I. Scaramucci, Matilde Virgínia Ricardi. II. Universidade Estadual de Campinas. Instituto de Estudos da Linguagem. III. Título.

## BANCA EXAMINADORA

---

Prof.<sup>a</sup>. Dr.<sup>a</sup>. Matilde Virgínia Ricardi Scaramucci  
(Orientadora)

---

Prof.<sup>a</sup>. Dr.<sup>a</sup>. Sandra Regina Buttros Gattolin

---

Prof. Dr. Paulo Sampaio Xavier de Oliveira

Dedico este trabalho a meus pais, que sempre me apoiaram em todos os momentos; e a Evelin Moreno, cujo entusiasmo tem tornado mais significativas minhas realizações.

## **AGRADECIMENTOS**

À Profª. Drª. Matilde Virgínia Ricardi Scaramucci, pela dedicação como professora, o que contribuiu definitivamente para minha formação acadêmica, e pelo apoio e orientação durante a realização deste trabalho.

À Prof. Drª. Sandra Regina Buttros Gattolin e ao prof. Dr. Paulo Sampaio Xavier de Oliveira, membros da banca, pelas valiosas sugestões e pelo profissionalismo e respeito durante o exame de qualificação.

Aos professores do IEL, pelo conhecimento compartilhado.

Aos meus pais, pelo exemplo de vida que sempre me proporcionaram.

À Evelin Moreno, pelo entusiasmo e pelo incentivo.

## RESUMO

Listas de palavras baseadas em corpora têm sido importantes no ensino e aprendizagem de vocabulário em língua estrangeira. Este trabalho visa compilar uma lista de palavras do português do Brasil, freqüentes e de grande alcance no contexto acadêmico brasileiro, e discutir sua utilidade no ensino de vocabulário, na preparação de textos modificados e de testes de vocabulário para aprendizes de Português como língua estrangeira. Discutimos também a questão do ensino direto e indireto de vocabulário. Para a compilação da lista, foi criado um corpus composto de teses e dissertações de 3 universidades públicas brasileiras, perfazendo um total de pouco mais de 7 milhões de palavras (aproximadamente 15 mil páginas de texto). Foram usados os critérios de freqüência e alcance para a seleção das palavras a serem incluídas na lista. As palavras são apresentadas com a freqüência e, nos casos de polissemia, com os diferentes significados com que ocorrem no corpus.

Palavras-chave: ensino e aprendizagem de vocabulário, leitura em língua estrangeira, corpora e listas de palavras, português como língua estrangeira.

## ABSTRACT

Corpora based word lists have been important for the teaching of vocabulary in foreign language. This paper aims to put together a list of high frequency and range Brazilian Portuguese academic words, and discuss direct and indirect vocabulary acquisition, the usefulness of modified texts for language learners and the use of word lists for the elaboration of vocabulary tests. A corpus was assembled from thesis and dissertations from 3 Brazilian public universities, totaling a little over 7 million words (approximately 15.000 pages of text). Frequency and range were the criteria used to select the words to be included in the list. Information about the number of times the words occur in the corpus and the different meanings they have, in the case of polysemy, is provided.

Key-words: vocabulary teaching and learning, reading in a foreign language, corpora and word lists, Portuguese as a foreign language.

## SUMÁRIO

RESUMO .....	v
ABSTRACT .....	vi
<b>1 INTRODUÇÃO .....</b>	<b>1</b>
1.1 O contexto de pesquisa .....	1
1.2 O problema de pesquisa e a justificativa da escolha do tópico .....	2
1.3 O objetivo e as perguntas de pesquisa .....	5
<b>2 LISTAS DE PALAVRAS E O ENSINO DE LÍNGUAS .....</b>	<b>7</b>
2.1 Introdução .....	7
2.2 Corpora pré-eletrônicos .....	8
2.3 Corpora de primeira geração .....	16
2.4 Corpora de segunda geração .....	18
2.5 Corpora de língua portuguesa .....	20
2.6 Listas de palavras acadêmicas .....	21
2.7 Conhecimento de vocabulário e leitura .....	24
2.8 O contexto acadêmico .....	26
2.9 O vocabulário e os métodos de ensino de línguas .....	29
2.10 Aquisição direta e indireta .....	33
2.11 Textos modificados .....	41
2.12 Avaliação de vocabulário .....	48
2.13 Escolha de palavras .....	50
<b>3 METODOLOGIA .....</b>	<b>56</b>
3.1 Compilação do corpus .....	56
3.2 Elaboração da lista .....	59
<b>4 CONSIDERAÇÕES FINAIS .....</b>	<b>78</b>
4.1 Respostas às perguntas de pesquisa .....	78
4.2 Limitações .....	80
4.3 Contribuições .....	80
<b>Lista de palavras acadêmicas do português do Brasil .....</b>	<b>81</b>
<b>REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>103</b>

## LISTA DE TABELAS

Tabela 1 - Exemplo de um verbete da GSL .....	16
Tabela 2 - Composição do CIC .....	20
Tabela 3 - Famílias .....	24
Tabela 4 - Composição do corpus da LPA-BR .....	57
Tabela 5 - Composição do corpus (PUC) .....	60
Tabels 6 - Coberturas da LPA-BR .....	79

## CAPÍTULO I

### INTRODUÇÃO

#### 1.1 O contexto de pesquisa

A leitura de textos acadêmicos é de importância vital para aqueles que se dedicam aos estudos de mestrado e doutorado, sem mencionar os estudantes de graduação. Da mesma forma que o inglês se mostra essencial para os estudos em países de língua inglesa e para a leitura de textos nessa língua, o português é igualmente importante para os estrangeiros que pretendem realizar seus estudos em universidades brasileiras.

A respeito do interesse de estrangeiros pela aprendizagem da língua portuguesa, Almeida Filho (1998) observa:

Uma das línguas que vem crescendo em importância, motivada pela globalização econômica, é o Português. Os acordos político-econômicos como o Mercosul, trazem a necessidade de falar e escrever ambos Português e o Espanhol, já que o tradicional "Portunhol" utilizado não é mais suficiente para o grande volume de negócios a serem efetuados. Assim, nos países de fala hispânica começam a ser implantados cursos de língua portuguesa em faculdades de direito, economia, publicidade e propaganda entre outras.

Um outro fator para o crescimento do ensino de Português como língua estrangeira, é o grande número de alunos estrangeiros nas universidades brasileiras, dentre os quais muitos são falantes de Espanhol. Essas instituições buscam então fornecer o curso de Português para estrangeiros de maneira que os alunos possam minimizar suas dificuldades de comunicação e compreensão na língua portuguesa.

Esse interesse fica ainda mais claro quando se considera o número de candidatos ao exame Celpe-Bras, o Certificado de Proficiência em Língua Portuguesa para Estrangeiros, em 2005.

De acordo com artigo do site Universia Brasil (2005): "O número de candidatos inscritos para

receber o Certificado de Proficiência em Língua Portuguesa para Estrangeiros (CELPE-BRAS) cresce a cada ano. Em 2003, foram 1.710 inscritos; em 2004, 1.807; e, em 2005, 1.856". Acreditamos que pesquisas sobre o ensino de português para estrangeiros são relevantes e úteis para os professores que se dedicam a essa área.

## **1.2 O problema de pesquisa e a justificativa da escolha do tópico**

Durante minha atividade como professor de inglês em escolas de línguas, e como professor particular de português para estrangeiros, tive oportunidade de verificar a dificuldade enfrentada pelos aprendizes para adquirir um vocabulário extenso o suficiente que lhes permita utilizar a língua para suas necessidades. Scaramucci (1995) verificou, em seu trabalho com três aprendizes de inglês de níveis diferentes, que todos consideravam o conhecimento de vocabulário seu maior problema; a esse respeito Gattolin (2005) afirma:

Considerando a defasagem lexical dos alunos ao ingressarem na universidade, a construção do vocabulário limiar, ao que parece, deve ser a meta do ensino, pois é ele o responsável pela possibilidade de o leitor interagir com a linguagem em uma situação de uso real como essa promovida pela leitura.

A aquisição de vocabulário tem sido tradicionalmente deixada a cargo da aprendizagem indireta, feita pelo aprendiz ao ler ou, mais raramente em contexto de LE<sup>1</sup>, ouvir na língua alvo com o objetivo específico de construir o sentido e não adquirir vocabulário. Diversas metodologias de ensino de línguas adotaram essa atitude, como mostraremos na seção 2.9. Por outro lado, vários estudos demonstram que a aprendizagem indireta, usada exclusivamente, pode não ser a melhor

---

<sup>1</sup> Neste trabalho será feita a distinção entre segunda língua (L2), ou seja, a língua aprendida em país falante da língua-alvo, e língua estrangeira (LE), a língua aprendida em país onde não se fala a língua-alvo.

opção para a aquisição de vocabulário pelos aprendizes, por razões que abordaremos no decorrer deste trabalho. De forma geral, os aprendizes têm por volta de 3 horas de aulas semanais nas escolas de ensino de idiomas. Esse tempo normalmente é dedicado à tentativa de desenvolver as habilidades de leitura, escrita, produção e compreensão orais. A exposição do aprendiz à língua é feita através de textos curtos, diálogos gravados por nativos e a comunicação oral com o professor e com os outros aprendizes, com os problemas que a possível falta de proficiência dos mesmos podem apresentar como fonte de insumo. Vários autores (HIRSCH; NATION, 1992; LAUFER, YANO, 2001; LAUFER; NATION, 1995; BENGELEIL; PARIBAKHT, 2004; WARING, 1995; SCARAMUCCI, 1995) consideram a leitura como a melhor fonte de aquisição de vocabulário, não só para aprendizes de L2 e de LE como também para o de L1. Para isso, entretanto, são necessários muitos encontros com a mesma palavra, o que exige um volume de leitura considerável.

Raros aprendizes têm tempo, motivação e interesse para ler a quantidade necessária para uma aquisição indireta de vocabulário bem sucedida em quantidade significativa, como veremos mais adiante. Uma solução apresentada por alguns pesquisadores é o ensino e aprendizagem diretos de vocabulário (MEARA, 1995; WARING, 1995), assim como o uso de textos modificados para facilitar a aquisição através da leitura (BARCROFT, 2003; RIDDER, 2002). Acreditamos que a melhor alternativa seja uma combinação de aprendizagem direta e indireta e que, no caso da aprendizagem indireta, o uso de textos modificados para aprendizes, que permitem um maior número de encontros com as palavras mais úteis, tornaria a aprendizagem de vocabulário através da leitura mais eficiente (WARING, 2003; BARCROFT, 2003; RIDDER, 2002; URANO, 2000). Na seção 2.11 falaremos sobre os textos modificados.

Entretanto, para a elaboração de atividades de aprendizagem direta e textos modificados para aprendizes precisamos saber quais palavras devemos usar, quais palavras serão mais úteis<sup>2</sup> e, portanto, deverão receber maior atenção, e é isso o que buscam mostrar as listas de palavras.

Uma lista de uso geral inclui palavras que são úteis na maioria dos usos da língua, ou seja, é um vocabulário comum essencial, utilizado na comunicação do dia a dia, textos não-técnicos, etc. Palavras de uso geral ocorrem freqüentemente em uma grande diversidade de textos. Sem informações sobre quais palavras serão mais úteis, acaba-se dedicando tempo de estudo valioso a palavras infreqüentes, de pouco alcance<sup>3</sup>, e pouco úteis para o aprendiz iniciante, quando se poderia estar dedicando tempo ao estudo de palavras que são usadas em um maior número de situações.

Se as listas gerais têm grande importância na escolha do vocabulário a ser trabalhado quando se está começando o estudo de uma língua, as listas de palavras acadêmicas têm a mesma importância para aqueles que pretendem atuar em áreas acadêmicas. Em artigo sobre o uso de corpus na preparação dos estudantes para as exigências lingüísticas de aulas regulares no ambiente universitário, Donley e Reppen (2001) afirmam: "A falta de relevância imediata do vocabulário acadêmico torna-o mais passível de ser negligenciado, apesar do papel importante que ele eventualmente terá na vida acadêmica dos estudantes". O papel de uma lista de palavras acadêmicas é, portanto, fornecer uma ponte confiável entre duas zonas relativamente acessíveis de aquisição lexical: entre as palavras que são freqüentes na língua em geral e as palavras que são freqüentes em uma área específica de estudo.

---

<sup>2</sup> Discutimos o que consideramos palavras úteis na seção 2.13.

<sup>3</sup> Alcance (*range*) diz respeito ao número de diferentes tipos de textos no qual uma palavra ocorre. Considerando-se um total de 50 textos diferentes, uma palavra é considerada como tendo um grande alcance quando ela ocorre nos 50 textos e como tendo um alcance pequeno quando ela ocorre em um texto em particular, mesmo que muitas vezes, e nenhuma vez em todos os outros 49 textos.

Podemos encontrar listas de palavras gerais (WEST, 1953), acadêmicas (XUE; NATION, 1984; COXHEAD, 2000) e outras em inglês, por exemplo. Essas listas são frequentemente utilizadas para orientar o ensino direto de vocabulário, a elaboração de testes de vocabulário e de textos modificados para aprendizes. O problema é que não encontramos listas semelhantes no português do Brasil.

A partir do início de minhas leituras acadêmicas na área da Linguística Aplicada, meu interesse foi despertado para o vocabulário utilizado nesses textos e a diferença deste e do vocabulário utilizado nos textos jornalísticos e de ficção, por exemplo. Esse interesse, e a falta de listas de palavras em português, motivou-me a compilar uma lista de palavras acadêmicas do português do Brasil (LPA-BR) que pudesse orientar os professores de português para estrangeiros sobre o vocabulário que será encontrado por seus alunos que vierem a atuar em contexto acadêmico.

### **1.3 O objetivo e as perguntas de pesquisa**

Para alcançar nosso objetivo, que é compilar uma lista com as palavras acadêmicas mais frequentes e de maior alcance do português do Brasil, consideradas as mais úteis para os aprendizes que, já tendo um conhecimento do vocabulário geral do português, tenham a necessidade de ler ou escrever textos acadêmicos, estabelecemos algumas perguntas de pesquisa:

1. Quais itens lexicais ocorrem frequentemente em uma variedade de textos acadêmicos?

Como a LPA-BR tem por objetivo se constituir em uma lista acadêmica geral, ela deve apresentar palavras que sejam frequentes em várias áreas acadêmicas. Daí a segunda pergunta de pesquisa:

2. Qual a frequência com que ocorrem os itens lexicais da LPA-BR nas diferentes áreas de conhecimento?

Para que a LPA-BR se caracterize como uma lista acadêmica, ela deve apresentar palavras que sejam frequentes em textos acadêmicos mas não em textos gerais, pois isso a caracterizará como uma lista específica de palavras acadêmicas e não uma lista geral. Daí a última pergunta de pesquisa:

3. Qual a cobertura da LPA-BR sobre textos acadêmicos e não-acadêmicos?

No capítulo II abordaremos a evolução dos corpora e as listas de palavras, a relação entre vocabulário e leitura, a aquisição direta e indireta de vocabulário, o uso de listas de palavras para a elaboração de textos modificados e testes de avaliação de conhecimento de vocabulário para aprendizes de línguas e a razão de considerarmos as listas de palavras úteis na escolha do vocabulário a ser focado; no capítulo III falaremos sobre os critérios para a compilação do corpus e elaboração da LPA-BR e, finalmente, no capítulo IV responderemos às perguntas de pesquisa e apresentaremos a LPA-BR.

## CAPÍTULO II

### LISTAS DE PALAVRAS E O ENSINO DE LÍNGUAS

#### 2.1 Introdução

A importância de desenvolver o léxico dos aprendizes de LE iniciantes para que adquiram um vocabulário que englobe as palavras mais frequentes tão cedo quanto possível tem sido repetidamente enfatizado pelos pesquisadores (CARTER; MCCARTHY, 1988; COBB; HORST, 2001; KYONGHO; NATION, 1989; MEARA, 1995; NATION, 2001). Foi estimado que as 1.000 famílias de palavras mais frequentes na língua inglesa, segundo a GSL<sup>4</sup>, oferecem uma cobertura de 82.3% dos textos de ficção e 71,4% dos textos acadêmicos segundo estudo de Hu e Nation (2000) e aproximadamente 75% de um agrupamento de textos diversos, incluindo textos de jornais, revistas, ficção e também textos acadêmicos e técnicos (COBB; HORST, 2001). Esses números mostram que o retorno da aprendizagem de um vocabulário formado pelas palavras mais frequentes é significativo.

O estudo de Cobb e Horst (2001), por exemplo, apresenta evidências positivas sobre o uso de listas de palavras para acelerar a aquisição de vocabulário. Nesse estudo foram recrutados 58 estudantes universitários de inglês como L2 para estudarem uma lista de palavras equivalente à GSL, com a ajuda de um programa de concordância, onde as palavras eram apresentadas em contexto. Os resultados mostraram que os aprendizes foram capazes de adquirir 180 palavras em 2 meses.

---

<sup>4</sup> *General Service List*, lista geral elaborada por Michael West. Ver seção 2.2.

Uma outra área de estudo sobre a aprendizagem de vocabulário através de listas de palavras concentra-se em como levar aprendizes que possuem um vocabulário geral básico a níveis mais altos de conhecimento lexical. A *University Word List* (XUE; NATION, 1984) e a *Academic Word List* (COXHEAD, 2000), listas de palavras que ocorrem em contexto acadêmico, são freqüentemente empregadas com aprendizes que possuem um vocabulário básico e aspiram a uma carreira acadêmica. Um estudo de Worthington e Nation (1996) concluiu que o número de palavras desconhecidas de uso acadêmico encontradas por estudantes no primeiro semestre do curso superior é maior do que o recomendável para permitir a inferência bem sucedida. Os autores recomendam o estudo direto de vocabulário acadêmico para minimizar essas dificuldades.

Listas de palavras são importantes na área de ensino e aprendizagem de LE principalmente em 3 áreas: aquisição de vocabulário, elaboração de textos modificados e avaliação de vocabulário.

Neste capítulo, apresentaremos uma breve visão da evolução dos corpora e da elaboração das primeiras listas de palavras. Veremos também como têm sido tratados o ensino e a aprendizagem de vocabulário em diversas metodologias de ensino de línguas, a relação entre vocabulário e leitura, o ensino e aprendizagem diretos e indiretos de vocabulário, a avaliação de vocabulário, os critérios para a escolha das palavras a serem enfocadas em uma abordagem de aquisição direta de vocabulário e os textos modificados.

## 2.2 Corpora pré-eletrônicos

Primeiramente, alguns termos freqüentemente utilizados no estudo dos corpora precisam ser conhecidos: **Ocorrências** são todas as palavras que aparecem em um texto, seja escrito ou falado. São contadas todas as vezes em que uma palavra aparece. Dessa forma a sentença: “Eu não sei se

eu vou” conteria seis palavras ou ocorrências. **Tipos** são ocorrências não repetidas. Essa mesma sentença teria 5 tipos, a palavra “eu” é contada como um tipo. **Lema** consiste de uma forma base e algumas de suas formas flexionadas e reduzidas (n't em inglês, por exemplo). Então *go*, *go*, *went*, *gone* contam como 4 ocorrências, 3 tipos e 1 lema (VERMEER, 2000). **Famílias** incluem a forma base, suas formas declinadas e suas formas derivadas próximas, o que inclui afixos como: *-ly*, *-ness*, e *un-*<sup>5</sup>.

O termo corpus, palavra do Latim para corpo ou conjunto, foi usado no século VI para descrever uma coleção de textos legais, *Corpus Juris Civilis* (FRANCIS; KUCERA, 1979), e tem mantido esse significado, o de um corpo de texto; mas para a lingüística de corpus essa definição não é suficiente. Numa das cinco definições do *Oxford English Dictionary* um corpus é "O corpo de material escrito ou falado sobre o qual uma análise lingüística é baseada". O dicionário *Collins COBUILD* define um corpus como "uma grande coleção de textos escritos ou falados que é usada para a pesquisa lingüística. O dicionário Aurélio define corpus como "conjunto finito de materiais significantes constituído com vistas à análise semiológica". Portanto, o corpus não pode ser visto como apenas uma coleção de textos, mas sim uma coleção de textos considerados representativos de uma dada língua, dialeto ou outro sub-sistema de uma língua, para ser usada para análise lingüística.

Freqüentemente, supõe-se que a pesquisa baseada em corpora<sup>6</sup> tenha começado no início dos anos 60, com a disponibilidade de corpora eletrônicos. Entretanto, antes disso já havia uma considerável tradição de análise lingüística baseada em corpora.

---

<sup>5</sup> Ver seção 2.6 para mais informações sobre o agrupamento em famílias.

<sup>6</sup> Corpora corresponde ao plural de corpus.

Alexander Cruden, em meados da década de 1720, decidiu compilar a mais detalhada concordância<sup>7</sup> da Bíblia (*King James Version of the Bible* - KJVB). A primeira edição da Concordância da Bíblia de Cruden foi publicada em 1737. Esse trabalho, que sempre foi realizado por um grande grupo de pessoas e, mais tarde, pelos computadores, foi feito por Cruden em sua residência, sozinho e à mão. A KJVB tem 777.746 palavras. Cruden ainda escreveu textos explicativos para muitas das palavras; a palavra Sinagoga, por exemplo, remetia a um texto de 4.000 palavras (LARSEN, 2005).

O *Dictionary of the English Language*, publicado em 1775 por Samuel Johnson não foi, como muitos acreditam, o primeiro dicionário da língua inglesa, mas foi o primeiro a apresentar contextualização (150.000 citações ilustrativas), e foi o dicionário padrão de inglês por mais de 150 anos. O dicionário de Johnson permaneceu o mais confiável dicionário da língua inglesa até o aparecimento do dicionário *A New English Dictionary on Historical Principles*, publicado em 10 fascículos entre os anos de 1884 e 1928 e que foi reimpresso em 1933, quando recebeu o nome de *Oxford English Dictionary*; 1.700 das definições de Johnson ainda se encontram nesse dicionário.

Em 1694, porém, a França já tinha publicado seu enorme dicionário, *Le dictionnaire de l'Academie francaise*, resultado de meio século de trabalho de 40 acadêmicos. Johnson acreditava que poderia fazer o trabalho em três anos mas, ao final, levou quase uma década, trabalhando com caneta e papel no sótão de sua casa em Londres. Johnson compilava longas listas de palavras, as quais eram cortadas em tiras e ordenadas alfabeticamente. Ele garimpou a literatura de muitos séculos em busca de citações ilustrativas para suas palavras (LYNCH, 2003).

A idéia para o *Oxford English Dictionary*, um dicionário extremamente complexo, surgiu em 1857. O plano era criar uma coleção vasta e abrangente de palavras inglesas, um léxico da língua mais completo do que jamais tinha sido tentado.

---

<sup>7</sup> Uma listagem, total ou parcial, das palavras de um texto ou corpus, mostrando o contexto de cada ocorrência de cada palavra.

O plano foi formulado pela Sociedade Filológica Inglesa (*English Philologic Society*), um grupo que investigava a estrutura e a história das línguas. Reconhecendo as lacunas que existiam em outros dicionários, a sociedade acreditava que um novo projeto deveria ser executado para examinar o vocabulário da língua inglesa, inclusive aquelas palavras que tinham sido rejeitadas ou tinham passado despercebidas por outros lexicógrafos.

Em 1879, um acordo foi alcançado com os editores para começar o trabalho. O editor do dicionário, James Murray, pediu aos leitores dos países de língua inglesa para participarem. Centenas de voluntários trabalharam como “detetives de palavras”, vasculhando textos históricos e contemporâneos para coletar tantas palavras quanto possível, analisando as diferentes maneiras em que cada uma tinha sido usada. Procuraram na literatura (popular e clássica), jornais, periódicos científicos e tecnológicos, letras de músicas, roteiros de teatro, livros de receitas, testamentos e documentos políticos, coletando uma infinidade de palavras e significados. Como resultado, Murray recebeu milhões de citações, as quais sua equipe passou a avaliar, ordenar e arquivar.

O dicionário viria a incluir palavras perdidas e fora de moda, assim como as mais recentes e também termos técnicos; ele traçava a etimologia de cada palavra, mostrando seu uso mais antigo, mapeando como ela tinha mudado em uso ou significado através do tempo; mostrava também as famílias das palavras, pronúncia e múltiplos significados. Quase 50 anos se passariam antes que a o último fascículo do dicionário fosse publicado, em 1928.

Nascido em Wet Hartford, Connecticut em 1758, Noah Webster acreditava ardentemente no desenvolvimento da independência cultural dos Estados Unidos, e que uma linguagem americana distinta, com suas próprias características, pronúncia e estilo, era essencial.

Em 1806 Webster publicou *A Compendious Dictionary of the English Language*, o primeiro dicionário verdadeiramente americano. Imediatamente, começou a trabalhar na sua obra máxima, o *American Dictionary of the English Language*, para o que ele aprendeu 26 línguas,

incluindo Anglo-Saxão e Sânscrito, a fim de pesquisar as origens da língua do seu país. Esse livro, publicado em 1828, incorporava um novo padrão da lexicografia; era um dicionário com 70.000 registros e considerado por muitos como tendo superado o dicionário de Johnson, não somente em escopo como também em confiabilidade.

Uma faceta da personalidade de Webster era sua disposição para inovar quando ele pensava que a inovação significava evolução. Ele foi o primeiro a documentar palavras distintamente americanas como: *skunk*, *hickory* e *chowder*. Acreditando que muitas convenções da escrita eram artificiais e desnecessariamente confusas, ele encorajou a alteração de várias palavras: *musik* para *music*, *centre* para *center* e *plough* para *plow*, por exemplo (NOAH, 2004).

Edward Lee Thorndike nasceu em Williamsburg, Massachusetts, E.U.A., em 1874. Esse psicólogo educacional americano ofereceu muitas contribuições para a pesquisa sobre educação e aprendizagem. Ele foi um dos primeiros a desenvolver testes para medir a aprendizagem e aptidões. O interesse de Thorndike por contagem de frequência de palavras iniciou-se quando ele veio a saber que os professores de línguas na Alemanha e na Rússia estavam usando contagens de palavras. Eles acreditavam que quanto mais frequentemente uma palavra fosse usada mais útil ela seria para os aprendizes de línguas.

Em 1911, Thorndike começou a contar a frequência das palavras em textos em inglês com a intenção de coletar palavras que seriam úteis para crianças americanas aprendendo a ler nessa língua. Em 1921 ele publicou *The teacher's word book*, o qual listava 10.000 palavras pela frequência de uso, e foi baseado em um corpus de 4.500.000 palavras. Aproximadamente 625.000 palavras de textos da literatura infantil, 3.000.000 de palavras da Bíblia e clássicos ingleses, 300.000 palavras de livros-texto do nível elementar, 50.000 palavras de livros sobre culinária, agricultura, costura, comércio, etc., 90.000 palavras de jornais e 500.000 palavras de correspondências foram usadas para compor esse corpus (MALLIKARJUN, 2002).

Em 1932 ele deu seguimento ao trabalho com a publicação de *A teacher's word book of 20,000 words* (DUBAY, 2004). Em 1944, com a ajuda de Irving Lorge, Thorndike publicou *The teacher's word book of 30,000 words*. Acrescentando a contagem de palavras de revistas que havia sido feita por Lorge, a contagem de Thorndike de 120 livros juvenis e a contagem semântica de Thorndike e Lorge à contagem geral de Thorndike, o corpus usado para a elaboração dessa lista chegou a 18.000.000 de ocorrências. O trabalho de Thorndike foi a base para as primeiras fórmulas de facilidade de leitura<sup>8</sup>. Educadores, editores de livros-texto e professores usam listas de frequência de palavras, assim como fórmulas de facilidade de leitura para casar materiais de leitura com leitores de diferentes níveis.

Educado em uma escola pública em Christ Church, Oxford, Inglaterra, Michael Philip West foi diretamente da graduação, em 1912, para um posto no Serviço de Educação da Índia. Seus anos mais frutíferos e influentes foram passados em Bengala onde ele veio a tornar-se Inspetor de Diretores de Escolas em Chittagong e Calcutá, Diretor da Faculdade de Treinamento de Professores em Dacca e Leitor Honorário em Educação na Universidade de Dacca (HOWATT, 1984).

A mais importante sucessão de eventos na carreira de West, entretanto, aconteceu 10 anos depois de sua nomeação. Em 1923, a Conferência de Educação Imperial tinha pedido uma investigação dos fatos do bilingüismo com referência ao desenvolvimento intelectual, emocional e moral da criança e a importância das questões práticas de educação advindas da investigação de tais fatos. A resposta de West a esse pedido foi um projeto experimental, escrito no relatório sobre bilingüismo. Nele, West desafiava a política educacional, que distribuíra graduações e empregos no governo para um pequeno número de estudantes muito capazes, mas que resultava na desistência da maioria em diversos estágios, sem que qualquer benefício fosse auferido desse processo de educação

---

<sup>8</sup> De acordo com Richards, Platt e Platt (1992), facilidade de leitura (*readability*) se refere à facilidade com que textos podem ser entendidos, e depende de vários fatores, incluindo a extensão das sentenças, o número de palavras novas e a complexidade da linguagem utilizada no texto.

incompleto (TICKOO, 1988). O desperdício era maior devido ao longo período dedicado ao desenvolvimento da habilidade de conversação. A habilidade de leitura, por outro lado, podia ser adquirida mais rapidamente e mais agradavelmente, acreditava West, e não dependia da presença constante de um professor.

A abordagem de West permitia dar acesso a todos os bengaleses, e não só a alguns poucos privilegiados, a materiais de leitura sobre assuntos práticos que não estavam disponíveis na língua mãe. Isto significava o desenvolvimento de materiais melhorados, baseados em princípios de controle estrito do vocabulário. Havia duas maneiras, segundo West, pelas quais os textos existentes podiam ser melhorados. O primeiro era simplificar o vocabulário removendo palavras raras ou antiquadas e substituindo-as por equivalentes mais modernas e comuns. O segundo princípio aplicado por West consistia em distribuir as palavras desconhecidas, o vocabulário alvo, de maneira tal que ocorressem com uma freqüência que permitisse ao leitor absorvê-las e praticá-las a fundo. Nos novos textos que West adaptou ou escreveu, o número de palavras novas caiu, enquanto o número total de palavras aumentou acentuadamente. Em conseqüência, o ritmo pelo qual novas palavras eram apresentadas foi diminuído de 1 palavra desconhecida para 7,4 conhecidas para não mais do que 1 desconhecida para 44,7 conhecidas (HOWATT, 1984).

As idéias de West estavam de acordo com as do movimento de controle de vocabulário. A idéia por trás desse movimento era a de que as palavras mais importantes de uma língua deveriam ser ensinadas aos aprendizes primeiro, e que o número dessas palavras deveria ser limitado.

Em meados dos anos 30, West envolveu-se, com outras figuras do movimento para o controle de vocabulário, na Conferência de Carnegie, a qual aconteceu pela iniciativa de West e sob os auspícios da Corporação Carnegie, primeiro em Nova Iorque, em 1934, e depois em Londres em 1935. O propósito principal da conferência era examinar o papel das listas de palavras no ensino de inglês como língua estrangeira.

A escolha dos itens da *General Service List* foi confiada a West e Palmer. Essa lista viria a ser o maior componente do relatório da conferência, e seria publicada em 1953 separadamente, como a mais confiável e certamente a mais conhecida de todas as listas de palavras. Durante os anos 30, West se tornou um dos mais conhecidos e mais prolíficos autores no campo do inglês como língua estrangeira.

Sua carreira, porém, foi marcada pela controvérsia. Embora suas propostas para uma abordagem centrada no leitor fossem consideradas excelentes, elas não eram vistas com bons olhos nos círculos oficiais da Índia (TICKOO, 1988). Nos anos 50 e 60 ele permaneceu ativo, contribuindo regularmente para o periódico *English Language Teaching* entre outros. Escreveu também um pequeno manual, *Teaching English in Difficult Circumstances*, o qual incluía um apêndice com o que West considerava o vocabulário mínimo para permitir a comunicação. A GSL (West, 1953) é uma lista de aproximadamente 2.000 famílias de palavras escolhidas, em grande parte, com base em sua frequência de uso. Os dados de frequência usados na GSL vieram das contagens feitas por Thorndike e Lorge no início do século 20. Existem algumas significativas diferenças entre a GSL e listas anteriores. A primeira é que as palavras base e suas derivadas são apresentadas com exemplos ilustrativos. Exemplos com a palavra *extend* podem ser vistos na tabela 2. Além disso a lista é enriquecida pela inclusão de expressões idiomáticas e verbos frasais tais como: *break to pieces*, *break one's heart*, *break down*, *break up*.

Tabela 1 - Exemplo de um verbete da GSL

<b>(3) EXTEND</b>	
<b>extend</b> , v.	(1) ( <i>stretch out, be stretched out</i> ) The garden extends as far as the river (2) ( <i>continue, enlarge, protract, lengthen</i> ) Extend one's visit Extend a business
<b>extent</b> , n.	To the full extent of the garden
<b>extension</b> , n.	An extension of the hospital
<b>extensive</b> , adj.	Extensive repairs, enquiries
<b>extensively</b> , adv. (GSL)	

Fonte: West (1953)

Nation e Hwang (1995) mostraram que as 2.000 palavras mais freqüentes continuam sendo o melhor ponto de partida para qualquer estudante que esteja começando os estudos da língua inglesa. A GSL tem sido criticada devido à sua idade, mas as palavras mais usadas, identificadas pela análise por computador de grandes corpora de textos escritos em inglês, têm se mostrado muito similares às da GSL (READ, 2004).

### 2.3 Corpora de primeira geração

O *Brown Corpus of Standard American English* foi o primeiro corpus computadorizado. Ele foi compilado por W. Nelson Francis e Henry Kucera, da Universidade Brown, Providence, Rhode Island. O corpus consiste de 1 milhão de palavras extraídas de textos do inglês americano. Hoje, esse corpus é considerado pequeno e ligeiramente antiquado, porém ainda é usado. Muito da sua utilidade vem do fato de que sua estrutura tem sido copiada por outros compiladores de corpora. O Lancaster-Oslo/Bergen (LOB) Corpus (inglês britânico) e o Kolhapur Corpus (inglês indiano) são

dois exemplos de corpora feitos para serem compatíveis com o Brown Corpus. A disponibilidade de corpora tão similar em estrutura é um recurso valioso para pesquisadores interessados em comparar diferentes variedades da língua (FRANCIS; KUCERA, 1979).

O Corpus Lancastre-Oslo/Bergen (LOB) foi compilado por pesquisadores de Lancaster, Oslo e Bergen, em 1961. Ele consiste de 1 milhão de palavras do inglês britânico extraídas de textos de 15 categorias diferentes como: jornais, textos de ficção adulta e infantil, biografias, revistas, etc. Cada texto tem pouco mais de 2.000 palavras (textos mais longos são cortados na primeira sentença após as 2.000 palavras) e o número de textos varia em cada categoria. O corpus é a contrapartida britânica ao Corpus Brown, o qual contém textos do mesmo ano, de forma que uma comparação entre as duas variedades pudesse ser feita (LOB, 2002).

Outros corpora gerais de primeira geração:

- *The Kolhapur Corpus of Indian English* - 1978, inglês indiano, segue a estrutura do Corpus Brown.
- *The Wellington Corpus of Written New Zealand English* - 1986, inglês da Nova Zelândia, segue a estrutura dos corpora Brown e LOB com algumas modificações.
- *The Australian Corpus of English* - 1986, mesma estrutura e tamanho do Brown/LOB, com algumas modificações.
- *The Corpus of English-Canadian Writing* - 1984, tem a mesma estrutura do Brown/LOB com a adição de textos das categorias feminismo e informática, sendo três vezes maior do que o Brown.
- *The London-Lund Corpus (LLC)* - 1987, parte oral do *Survey of English Usage Corpus*, da *Lund University*. Ele contém 500.000 palavras do inglês britânico falado coletadas entre 1953 e 1987 (TONO, 2003).

## 2.4 Corpora de segunda geração

Em 1980, na Universidade de Birmingham, o professor John Sinclair estabeleceu o projeto COBUILD para realizar uma análise por computador de um corpus totalizando 18 milhões de palavras. Hoje esse corpus cresceu para mais de 400 milhões de palavras, havendo uma previsão de que mais 100 milhões sejam acrescentadas no futuro próximo.

COBUILD é uma sigla de "CO" de Collins, a editora, e "BUILD" de *Birmingham University International Language Database*.

Quase todo o material do corpus foi coletado a partir de 1990, e todo ano ele é revisado, visando substituir dados antigos por mais atuais. Os textos vêm de várias fontes, embora os jornais, sendo uma das fontes mais práticas atualmente, respondam por uma grande parte do corpus. Além dos jornais, o corpus inclui uma extensa coleção de textos de revistas gerais e especializadas e também publicações de negócios como o *Wall Street Journal* e *The Economist*. O inglês falado também é representado, contando com aproximadamente 20 milhões de palavras de transmissões de rádio americanas da *National Public Broadcasting* em Washington, 20 milhões de palavras da *BBC World Service* e rádios locais, além de entre 10 e 20 milhões de palavras de conversas informais, entrevistas e reuniões. Atualmente a equipe do COBUILD está trabalhando para dar ao corpus um caráter mais internacional, com a inclusão de palavras de corpora australianos, canadenses e indianos (THE DICTIONARY, 2002).

O *Longman Corpus Network* compõe-se de um grupo de cinco bancos de dados: o *Longman Learners' Corpus*, composto pela produção escrita de aprendizes de inglês; o *Longman Written American Corpus*, composto por 100 milhões de palavras de livros e jornais americanos; o *Longman Spoken American Corpus*, composto por 5 milhões de palavras do inglês oral americano,

extraídas de situações do cotidiano; o *Spoken British Corpus* (parte do *British National Corpus*), composto por palavras do inglês oral britânico e o *Longman/Lancaster Corpus* com mais de 30 milhões de palavras cobrindo um âmbito extenso de textos escritos que inclui desde a literatura até tabelas de horários de ônibus.

O *British National Corpus* contém mais de 100 milhões de palavras do inglês moderno, falado e escrito. Esse projeto foi executado e é mantido por um consórcio acadêmico/industrial liderado pela *Oxford University Press*. As palavras representativas da linguagem escrita (90%) foram extraídas de jornais regionais e nacionais, periódicos especializados, livros acadêmicos e ficção popular, cartas e memorandos publicados ou não, ensaios escolares e universitários, entre muitos outros tipos de textos. A proporção referente à linguagem oral (10%) provém de uma grande quantidade de conversação informal não ensaiada, gravada por voluntários selecionados de diferentes idades, regiões e classes sociais de uma forma demograficamente equilibrada, além de excertos orais coletados em diversos tipos de contextos, indo de reuniões formais de negócios ou governamentais a programas de rádio e ligações telefônicas. O trabalho de construção do corpus foi completado em 1994 (WHAT, 2004).

O *International Corpus of English (ICE)* teve início em 1990 com o objetivo de coletar material para estudos comparativos de inglês mundialmente. Quinze equipes de pesquisa ao redor do mundo estão preparando corpora eletrônicos de suas próprias variedades nacionais ou regionais de inglês. Cada corpus do ICE consiste de 1 milhão de palavras do inglês falado ou escrito produzido depois de 1989. Para muitos países participantes, o projeto ICE está estimulando a primeira investigação sistemática da variedade nacional. Para garantir a compatibilidade entre os corpora, todas as equipes seguem a mesma estrutura. Cada corpus contém 500 textos de aproximadamente 2.000 palavras, totalizando aproximadamente 1 milhão de palavras (CORPUS, 2003).

O *Cambridge International Corpus (CIC)* foi compilado pela *Cambridge University Press* nos últimos 10 anos. Os textos do corpus vêm de jornais, livros de ficção e não-ficção sobre

vários tópicos, sites da internet, revistas, correspondência, programas de rádio e televisão e conversas cotidianas.

Tabela 2 - Composição do CIC

<b>Inglês britânico</b>	
<b>Nº de palavras</b>	<b>Corpus</b>
450 milhões	Inglês britânico falado
17 milhões	Inglês britânico falado
20 milhões	Inglês acadêmico britânico escrito
30 milhões	Inglês comercial britânico escrito
1 milhão	Inglês comercial britânico falado
<b>Inglês americano</b>	
<b>Nº de palavras</b>	<b>Corpus</b>
200 milhões	Inglês americano escrito
22 milhões	Inglês americano falado
7 milhões	Inglês acadêmico americano escrito
30 milhões	Inglês comercial americano escrito
<b>Inglês de aprendizes</b>	
<b>Nº de palavras</b>	<b>Corpus</b>
19 milhões	Inglês escrito por aprendizes (IEA)
8 milhões	IEA com erros codificados

Fonte: Cambridge (2005)

## 2.5 Corpora de língua portuguesa

O Núcleo Interinstitucional de Lingüística Computacional (NILC) foi criado em 1993, na Universidade de São Paulo (USP) em São Carlos, para estimular a pesquisa e o desenvolvimento

de projetos em Linguística Computacional e Processamento da Linguagem Natural. Por essa ocasião, deu-se início a construção de diversas ferramentas computacionais, visando a dar suporte à tarefa de revisão textual; dentre esse material de suporte, foi elaborado um banco de textos o qual, mais tarde, se tornaria o Corpus de base para diversos aplicativos desenvolvidos pelo NILC, o Corpus NILC (CN) (PINHEIRO; ALUÍSIO, 2003). A construção do CN foi um empreendimento decorrente da ferramenta de revisão desenvolvida pelo NILC em parceria com a ITAUTEC/PHILCO a partir de 1993 – o revisor ReGra. O NILC, em setembro de 2002, contava com 34.092.630 palavras. O CN é composto de textos didáticos (1.147.325 ocorrências), jurídicos (761.852 ocorrências), literários (2.184.620 ocorrências), técnico-científicos (1.767.565 ocorrências), jornalísticos (27.203.360 ocorrências) e universitários (1.027.908 ocorrências).

O corpus de Araraquara teve sua montagem iniciada no começo dos anos 90, por Francisco S. Borba, lingüista especializado em dicionários e autor de diversas obras sobre a língua portuguesa. É um dos mais importantes bancos de dados de língua escrita no país, com aproximadamente 200 milhões de ocorrências de palavras em textos de literatura românica, dramática, técnica, oratória e, predominantemente, jornalística (FRANÇA, 2002).

## **2.6 Listas de palavras acadêmicas**

A *University Word List* (UWL) é uma compilação, feita por Xue e Nation (1984) de quatro listas separadas, as de Campino e Ellery, de 1971, e a de Praninskas, 1972; baseadas em corpora e compostas por palavras que ocorriam em textos gerais; e as de Lynn, de 1973, e a de Ghadessy, de 1979, compiladas a partir das anotações de alunos em livros didáticos. A UWL consiste de 836 famílias, um total de 3.685 palavras, não encontradas na GSL mas frequentes em textos

acadêmicos (NATION; HWANG, 1995). Ela exclui as primeiras 2.000 palavras da GSL e termos técnicos. Assim fazendo, a UWL pretende ser uma lista de palavras acadêmicas gerais<sup>9</sup> que ocorrem através de uma grande extensão de áreas como artes, ciências e direito. A lista contém 11 níveis, com o nível 1 contendo as palavras mais freqüentes e o nível 11 as menos freqüentes (COXHEAD, 2000).

Por considerar que a UWL era um amálgama de outras listas, sem critérios consistentes de seleção, baseada em corpora pequenos e não contendo uma extensão de tópicos ampla e equilibrada, Coxhead (2000) compilou a *Academic Word List* (AWL), constituída de 570 famílias de palavras. O corpus para esse estudo foi composto por textos de 28 diferentes cursos, de 4 áreas: artes, comércio, direito e ciências, que juntos somaram 3,5 milhões de palavras. Para a seleção a autora seguiu três critérios: 1) ocorrência especializada, as famílias de palavras incluídas não podiam estar presentes na GSL; 2) alcance, uma palavra de cada família tinha de aparecer pelo menos 10 vezes em cada uma das quatro áreas e em pelo menos 15 dos 28 cursos que o corpus abrangia; e 3) freqüência, membros de uma família tinham que aparecer pelo menos 100 vezes no corpus. A lista foi dividida em famílias de acordo com o nível 6 da escala de Bauer e Nation.

Essa escala foi elaborada por Bauer e Nation (1993) para o agrupamento de palavras em famílias, na língua inglesa, com o objetivo de orientar os professores de inglês no ensino dos afixos. Bauer e Nation afirmam que “do ponto de vista da leitura, uma família de palavras consiste de uma palavra base e todas as suas formas derivadas e flexionadas que possam ser entendidas por um aprendiz sem que tenha que aprender cada forma separadamente”. Partindo dessa premissa, *watch*, *watches*, *watched* e *watching* podem ser membros da mesma família para um aprendiz que possua o conhecimento dos sufixos flexionais em inglês. Conforme o conhecimento do uso de afixos do aprendiz se desenvolve, o tamanho das famílias de palavras aumenta. Os autores deixam claro que

---

<sup>9</sup> O vocabulário acadêmico, ou sub-técnico, é o vocabulário comum a várias áreas acadêmicas, ao contrário do vocabulário técnico, definido como o vocabulário reconhecidamente específico de um tópico, campo ou disciplina.

o significado da palavra base na forma derivada deve estar claramente relacionado ao significado da palavra base quando encontrada em outras formas derivadas. Por exemplo, *hard* e *hardly* não seriam membros da mesma família.

A classificação do agrupamento em famílias é apresentada pelos autores em sete níveis, cada nível inclui os níveis anteriores e acrescenta novos afixos.

- Nível 1 - Cada palavra é considerada uma família
- Nível 2 - As flexões da palavra base consideradas pertencentes à mesma família neste nível são: plural, terceira pessoa do singular no presente, passado, particípio passado, *-ing*, comparativo, superlativo e possessivo.
- Nível 3 - As famílias incluem as palavras formadas pela adição dos seguintes afixos: *-able, -er, -ish, -less, -ly, -ness, -th, -y, non-, un-*.
- Nível 4 - Inclui os afixos: *-al, -ation, -ess, -ful, -ism, -ist, -ity, -ize, -ment, -ous, in-*.
- Nível 5 - Acrescenta os afixos: *-age, -al, -ally, -an, -ance, -ant, -ary, -atory, -dom, -eer, -en, -en, -ence, -ent, -ery, -ese, -esque, -ette, -hood, -i, -ian, -ite, -let, -ling, -ly, -most, -ory, -ship, -ward, -ways, -wise, anti-, ante-, arch-, bi-, circum-, counter-, en-, ex-, fore-, hyper-, inter-, mid-, mis-, neo-, post-, pro-, semi-, sub-, un-*.
- Nível 6 - Afixos: *-able, -ee, -ic, -ify, -ion, -ist, -ition, -ive, -th, -y, pre-, re-*.
- Nível 7: Afixos: *ab-, ad-, com-, de-, dis-, ex-, sub-*.

Então, usando como exemplo a palavra *develop* teríamos os seguintes agrupamentos em cada nível:

Tabela 3 - Famílias

Níve 1	Agrupamento
2	develop develops developed developing
3	developable undevelopable developer(s) undeveloped
4	development(s) developmental developmentally
5	developmentwise semideveloped antidevelopment
6	redevelop predevelopment

Fonte: Bauer e Nation (1993)

Para maiores informações sobre os critérios para a elaboração dos níveis, limitações do estudo, usos sugeridos e sugestões para pesquisa adicional, vide Bauer e Nation (1993).

Não encontrei estudos semelhantes sobre a língua portuguesa e acredito que tais estudos trariam grandes contribuições à área de ensino de português como língua estrangeira.

## 2.7 Conhecimento de vocabulário e leitura

A relação entre conhecimento de vocabulário e compreensão em leitura tem sido fortemente demonstrada tanto em contextos de L1 (ANDERSON; FREEBODY, 1981; BECK; PERFETTI; MCKEOWN, 1982; MEZYNSKI, 1983) quanto em contextos de LE (SCARAMUCCI,

1995; HULSTIJN, 1997) e L2 (NATION, 2002). Há trabalhos avaliando a influência da extensão<sup>10</sup> de conhecimento de vocabulário (LIU; NATION, 1985; SCARAMUCCI, 1995) como também da profundidade de conhecimento de vocabulário (DE BOT; PARIBAKHT; WESCHE, 1997; QIAN, 1999; SCARAMUCCI, 1995). Na pesquisa sobre leitura em L1 tem havido muitos estudos que demonstram uma forte relação entre vocabulário e leitura. Em um estudo de larga escala Thorndike (1973) pesquisou a leitura em 15 países (com mais de 10.000 estudantes) e relatou correlações através dos países e grupos etários indo de  $r=0,66$  até  $r=0,75$  para leitura e vocabulário<sup>11</sup>. Stanovich (2000) encontrou resultados semelhantes com estudantes da terceira até a sétima série ( $r=0,64$  a  $0,76$ ). Em uma série de estudos, Carver (2003) argumenta que a relação entre compreensão em leitura e conhecimento de vocabulário é tão alta que as pesquisas podem produzir correlações perfeitas. Embora a afirmação seja aparentemente exagerada, Carver apresenta evidências baseadas em dados de avaliação vindos de múltiplas fontes para apoiar sua posição. Levando em conta esses estudos, pode-se afirmar que existe uma forte correlação entre o conhecimento de vocabulário e a compreensão em leitura em L1.

Em contexto de L2, Droop e Verhoeven (2003) demonstraram uma forte relação entre conhecimento de vocabulário e habilidade de leitura com crianças de terceira e quarta séries vindas de grupos lingüísticos minoritários na Holanda ( $r= 0,71$ ). Qian (2002), em estudo com estudantes universitários de vários países, encontrou correlações de  $r=0,68$  a  $0,82$  entre a pontuação obtida no teste de leitura do TOEFL e três medidas de conhecimento de vocabulário. Read (1997) afirma que:

---

<sup>10</sup> De acordo com Qian (1999), quando se fala em conhecimento de vocabulário, extensão de conhecimento é definida como o tamanho do vocabulário, o número de palavras das quais o aprendiz tem pelo menos um conhecimento mínimo do significado; profundidade de conhecimento é definida como o nível de conhecimento que um aprendiz possui de vários aspectos de uma dada palavra, ou quão bem o aprendiz conhece essa palavra.

<sup>11</sup> Um índice “ $r=0$ ” indica que não há nenhuma correlação, um índice “ $r=1$ ” indica uma correlação total.

“Na verdade, a compreensão em leitura está tão fortemente associada com o conhecimento de vocabulário que os elaboradores de testes têm tido dificuldade em distinguir entre os dois”.

A esse respeito Scaramucci (1995, p.27), a partir da análise do cruzamento dos escores obtidos nos testes de conhecimento de vocabulário e os escores do teste de leitura, afirma:

“(...) essa análise mostra de maneira bastante satisfatória a correspondência entre escores de conhecimento de vocabulário e escores de leitura, ou seja, escores baixos em leitura parecem corresponder a escores baixos em conhecimento de vocabulário e vice-versa (...)”.

No mesmo trabalho (p.247) Scaramucci, comparando os padrões de leitura de dois sujeitos, comenta:

Embora os padrões de leitura desses dois leitores não sejam iguais, apresentam semelhanças em muitos aspectos, sendo a mais saliente a dificuldade de inferir palavras e de processar trechos longos de forma automatizada, uma vez que a construção de sentido é sempre truncada por uma palavra desconhecida.

Isto parece mostrar que mesmo o desconhecimento de poucas palavras em um texto pode afetar fortemente a compreensão, particularmente para aprendizes com níveis intermediários ou básicos de proficiência.

## **2.8 O contexto acadêmico**

O vocabulário acadêmico, ou sub-técnico, é o vocabulário comum a várias áreas acadêmicas, ao contrário do vocabulário técnico, definido como o vocabulário específico de um tópico, campo ou disciplina. Frequentemente, os aprendizes conhecem o vocabulário técnico por já estarem atuando na área, o que normalmente lhes falta é o vocabulário acadêmico geral, que muitas vezes serve de ligação entre o vocabulário geral e o técnico.

Sendo o vocabulário acadêmico freqüente em uma grande variedade de textos acadêmicos, mas pouco freqüente em textos não-acadêmicos, caracteriza-se como um vocabulário especializado, útil para os que têm um propósito específico. Segundo Nation (1990), depois de aprender o vocabulário geral mais freqüente, a continuidade da aprendizagem deve levar em conta os objetivos do aprendiz. Se ele pretende ler textos de ficção, revistas e jornais na língua alvo, então deve se preocupar com a aprendizagem do vocabulário de baixa freqüência que ainda desconhece. Isso seria mais eficientemente conseguido, segundo Nation, através da leitura de textos modificados de níveis mais avançados e, assim que possível, a leitura de textos não-modificados. Abordaremos mais profundamente a questão dos textos modificados na seção 2.11. Se o aprendiz pretender ler textos de uma área técnica específica, deve dedicar-se à aprendizagem do vocabulário técnico característico dessa área, embora talvez não possa prescindir do vocabulário acadêmico para uma leitura mais fluente. Se o aprendiz tiver como objetivo atuar na área acadêmica deverá voltar sua atenção para o vocabulário acadêmico (NATION, 1990).

O vocabulário acadêmico é o tipo de vocabulário especializado que é passível de ser ensinado e trabalhado pelo professor, particularmente em aulas de leitura instrumental (LI). Isso não ocorre tão facilmente com o vocabulário técnico devido, em geral, ao desconhecimento da área e vocabulário específicos pelo professor, e pela reunião de aprendizes de diferentes áreas técnicas em uma mesma sala de aula.

Qian (2002) realizou um estudo para determinar a relação entre o conhecimento de vocabulário e desempenho em leitura acadêmica. Os participantes foram 217 estudantes freqüentando um programa intensivo de inglês como L2 na Universidade de Toronto. Esses participantes foram recrutados das classes do nível intermediário em diante. Uma versão da seção de compreensão em leitura do teste de proficiência TOEFL, nunca antes utilizada, foi selecionada como o instrumento de avaliação de compreensão. Essa seção continha cinco passagens de textos, todos versando sobre assuntos acadêmicos gerais de níveis introdutórios, incluindo biologia,

astronomia, geografia, história da arte e biografias. Questões de múltipla escolha foram incluídas após cada passagem para avaliar a compreensão.

O teste de profundidade de vocabulário (TPV), originalmente desenvolvido por Read (1993), foi utilizado para avaliar a profundidade de conhecimento. Para avaliar a extensão de conhecimento foi utilizado o Teste de Níveis de Vocabulário, originalmente desenvolvido por Nation (1990). O estudo concluiu que as duas dimensões de conhecimento de vocabulário, a extensão e a profundidade, são igualmente importantes e estão fortemente associadas uma à outra, assim como a compreensão em leitura, particularmente para aprendizes de nível universitário e em contextos acadêmicos.

Tschirner (2004) realizou um estudo para avaliar a correlação entre tempo de estudo de LE (inglês) e nível de proficiência lingüística em contexto acadêmico. O estudo foi realizado com alunos do primeiro semestre de Língua e Literatura Inglesa da Universidade de Leipzig, Alemanha. Os estudantes tinham entre 5 e 11 anos de estudos de inglês, com uma média de 8 anos. Ao final do estudo, o autor concluiu que o maior obstáculo ao desempenho em contexto acadêmico é a falta de conhecimento de vocabulário, e que mesmo prolongado tempo de estudo de inglês (até oito anos) não dá condições, necessariamente, aos estudantes para alcançar os limiares de vocabulário necessários para propósitos acadêmicos. Particularmente os objetivos produtivos deixam de ser alcançados por uma larga margem. Parece haver pelo menos duas razões para essa situação. Por um lado, parece haver uma grande distância entre o volume de leitura que os aprendizes deveriam efetuar e quanto eles efetivamente efetuem. Por outro lado, muitas escolas parecem não fazer um trabalho suficientemente bom no ensino de vocabulário e de técnicas de aprendizagem de vocabulário, segundo o autor.

## 2.9 O vocabulário e os métodos de ensino de línguas

O principal método de ensino de línguas do início do século 19 foi o da Gramática-Tradução. Uma lição teria, tipicamente, uma ou duas regras gramaticais, uma lista de itens de vocabulário e alguns exemplos práticos para traduzir da L1 para a língua-alvo ou vice-versa. O principal critério para a escolha de vocabulário era freqüentemente sua utilidade para ilustrar uma regra gramatical. Era esperado que os estudantes aprendessem eles próprios o vocabulário necessário através de listas de palavras bilíngües, o que tornou o dicionário bilíngüe uma importante ferramenta de referência. Um dos maiores problemas com a metodologia de Gramática-Tradução era que ela estava focada na habilidade de analisar a língua, e não na habilidade de usar a língua. Além disso, a ênfase na leitura e na escrita pouco faziam para promover uma habilidade de se comunicar oralmente na língua alvo.

No final do século 19, novas idéias baseadas no uso da língua tinham se aglutinado no que se tornou conhecido como o Método Direto. Este procurava imitar a maneira como um nativo aprendia sua língua materna naturalmente, com foco na compreensão oral em primeiro lugar, então a produção oral e, somente mais tarde, a leitura e a escrita. O foco era prioritariamente no uso da língua-alvo, com alguns proponentes mais radicais banindo o uso da L1 na sala de aula. O vocabulário seria adquirido naturalmente através da interação durante as aulas. Evitava-se a tradução, sendo o vocabulário explicado através de figuras ou através de demonstração física (ex.: mímica; andar, pular ou falar para demonstrar esses verbos), com o vocabulário inicial sendo mantido simples e familiar, objetos na sala de aula ou peças do vestuário. Assim, o vocabulário era conectado com a realidade tanto quanto possível. Porém o Método Direto exigia que os professores fossem proficientes na língua alvo, o que nem sempre acontecia. Criticava-se também o fato de que, se fosse seguido fielmente, o Método Direto era contraproducente, exigindo que os professores chegassem

a grandes esforços na tentativa de evitar o uso da língua materna do aprendiz quando uma breve explicação nessa língua seria mais eficiente (RICHARDS; RODGERS, 2000). Ele também imitava a aprendizagem da L1, mas não levava em conta as diferenças entre a aquisição da L1 e da L2 ou da LE. Uma diferença fundamental é que os aprendizes de L1 têm uma exposição abundante à língua, enquanto que aprendizes de L2 têm muito menos, e aprendizes de LE muito menos ainda, usualmente apenas poucas horas por semana, na sala de aula.

Um estudo iniciado em 1923 e publicado em 1929, com o nome de Relatório Coleman, concluiu que o ensino de conversação era impraticável devido ao pouco tempo disponível para o ensino nas escolas e à irrelevância da conversação em língua estrangeira para os alunos universitários americanos. Esse estudo recomendava que a leitura em LE se tornasse o principal objetivo do ensino de línguas nos Estados Unidos (2000). Na mesma época Michael West<sup>12</sup>, ensinando inglês na Índia, concordava que aprender a ler fluentemente era mais importante do que a comunicação oral. Com base nessas idéias surge o Método da Leitura, defendido por educadores americanos e britânicos da época (LEELA MOHD, 1989).

Durante a Segunda Guerra Mundial, as limitações dos métodos acima se tornaram evidentes quando as forças armadas americanas se viram sem pessoal fluente em línguas estrangeiras, que pudessem trabalhar como intérpretes e tradutores. Elas precisavam de uma forma de treinar rapidamente seus soldados nas habilidades comunicativas orais. Foi criado o Programa de Treinamento Especializado do Exército, em 1942, encarregando universidades americanas de desenvolverem um método que pudesse permitir aos estudantes tornarem-se fluentes em conversação em várias línguas estrangeiras. O método desenvolvido derivava sua lógica do behaviorismo, o qual essencialmente dizia que a aprendizagem de línguas era um resultado da formação de hábitos. Então, o método incluía atividades que se acreditava reforçariam bons hábitos de linguagem, tais como atenção à pronúncia, prática oral intensiva (*drills*), foco nos padrões das sentenças e memorização.

---

<sup>12</sup> Mais informações o trabalho de West podem ser encontradas na seção 2.2.

Os aprendizes que passaram por esse "Método do Exército" eram na maioria maduros e altamente motivados e dedicavam um grande número de horas diárias ao estudo da língua, assim seu sucesso foi dramático. Durante muitos anos discutiu-se a possibilidade de que o "Método do Exército" fosse utilizado em cursos regulares de línguas. Porém o sucesso desse método devia-se a características muito específicas dos aprendizes e do ensino. Segundo Richards e Rodgers (2000) o método baseava-se "na intensidade do contato com a língua-alvo ao invés de uma base metodológica bem desenvolvida", mas convenceu diversos lingüistas do valor de uma abordagem intensiva baseada na comunicação oral para a aprendizagem de uma LE.

As origens do Método Audiolingual (AL) podem ser encontradas no "Método do Exército", no trabalho de Fries e Lado no Instituto de Língua Inglesa da Universidade de Michigan, no trabalho de lingüistas comparativos, na tecnologia do laboratório de línguas e no financiamento do governo americano para a pesquisa de ensino de línguas (LEELA MOHD, 1989). A proficiência oral era o principal objetivo desse método, a leitura e a escrita eram introduzidas somente após os estudantes terem praticado as estruturas oralmente. O AL enfatizava fortemente a prática oral intensiva e a repetição e o controle rígido da linguagem do aprendiz para minimizar os erros e maximizar o uso de formas corretas, os quais eram justificados pelas teorias behavioristas. Assumia-se que bons hábitos lingüísticos e exposição à própria língua levariam, eventualmente, a um aumento do vocabulário. Sendo assim, nenhum método específico de ensino de vocabulário seria necessário, já que "a prática inicial deveria se concentrar no domínio de estruturas fonológicas e gramaticais ao invés de no domínio de vocabulário" (RICHARDS; RODGERS, 2000).

No início dos anos 60, o AL era comumente visto como um enorme progresso o qual iria revolucionar o ensino e aprendizagem de línguas. Não mais seriam necessárias as entediadas regras gramaticais, listas de vocabulário e traduções de textos. Entretanto, no final dos anos 60 as limitações do AL estavam começando a se tornar óbvias. Ao contrário da teoria audiolingual, os aprendizes não reagiam de acordo com as expectativas behavioristas, eles queriam traduzir palavras

e textos, exigiam regras gramaticais e consideravam as infundáveis repetições enfadonhas e ineficientes para a aprendizagem. Os ataques de Noam Chomsky às idéias behavioristas, iniciados em 1959 com o artigo *A Review of B. F. Skinner's Verbal Behavior*, e que continuaram pelos anos sessenta, foram decisivos, fazendo com que o AL começasse a perder seguidores. Chomsky argumentava que a língua não era, de acordo com sua visão, a mera formação de hábitos. Chomsky via a língua não como imitação de comportamento, aprendido através da cópia e da repetição, mas sim "envolvendo inovação, formação de novas sentenças e padrões de acordo com regras abstratas e complexas" (CHOMSKY, 1996; citado por RICHARDS; RODGERS, 2000).

No final dos anos 60 e início dos anos 70, a visão estruturalista do ensino de línguas foi substituída por uma ênfase semântica e social (STERN, 1992). O crescimento da psicolinguística e da sociolinguística e um interesse em semântica tiveram implicações importantes para o ensino de línguas, já que destacaram a importância do uso da língua em situações reais. No início dos anos 70 o trabalho do sociólogo americano Hymes resultou no conceito de competência comunicativa. Hymes argumentava que uma sentença não deve ser apenas gramaticalmente correta, mas que ela também deve ser adequada em relação ao contexto no qual é usada. A teoria de Hymes tentava definir o que um indivíduo precisa saber a fim de ser comunicativamente competente (MURPHY, 2000).

Estas mudanças levaram ao surgimento de uma nova abordagem de ensino de línguas, o Ensino de Línguas Comunicativo ou Abordagem Comunicativa (AC). O foco era na comunicação e na fluência ao invés de precisão gramatical. Em qualquer abordagem baseada no significado seria de se esperar que fosse dada especial atenção ao vocabulário. Entretanto, a AC pouco fez além de vê-lo como um suporte para a linguagem funcional (ex. como fazer um pedido, como pedir desculpas). Como em abordagens prévias, esperava-se que o vocabulário da língua-alvo, como o da L1, se resolvesse por si mesmo, através da aprendizagem indireta (SCHMITT, 2000).

O pressuposto, assumido pela maioria das metodologias de ensino de línguas, de que o contato do aprendiz com a língua alvo seria o suficiente para a aquisição de vocabulário, a qual se daria através da inferência do significado das palavras desconhecidas, não pode ser dado como certo devido às dificuldades que a aprendizagem indireta pode apresentar, como veremos a seguir.

### **2.10 Aquisição direta e indireta**

Uma área em que as listas de palavras são muito úteis é a do ensino direto de vocabulário. A aquisição indireta de vocabulário é geralmente definida como a aprendizagem de vocabulário como sub-produto de qualquer atividade não diretamente desempenhada com vistas à aprendizagem de vocabulário, e é contrastada com a aprendizagem intencional, direta, de vocabulário, definida como uma atividade que objetiva gravar informação lexical na memória (RIEDER, 2003). O que verdadeiramente diferencia a aquisição indireta da direta é a intenção. Na aquisição indireta, a atividade não é realizada tendo a aquisição de vocabulário como principal objetivo, enquanto na direta, sim.

Línguas estrangeiras são normalmente aprendidas em escolas de línguas, onde o aluno tem em média 3 ou 4 horas de aulas por semana. A aprendizagem de vocabulário, se indireta, deve se dar tendo como fonte de insumo o discurso do professor, caso em que a proficiência deste influenciará a qualidade desse insumo; dos demais alunos, com um insumo vindo de aprendizes com graus de proficiência variados; e do material didático, textos dos livros e material auditivo. Esse material didático na maioria das vezes é a única fonte de linguagem autêntica, usada por falantes nativos, a que o aprendiz tem acesso. Levando em conta o tempo utilizado com práticas gramaticais e composições escritas, atividades presentes na maioria dos livros para ensino de línguas, o tempo de exposição a insumo que poderia ocasionar a aprendizagem indireta de vocabulário torna-se muito

limitado. Uma alternativa seria a leitura, que o aprendiz poderia fazer fora da sala de aula, aumentando o tempo de contato com a língua e, portanto, suas chances de adquirir vocabulário indiretamente. Alguns autores afirmam ser essa a melhor maneira de abordar a aprendizagem de vocabulário, deixar que o aluno adquira vocabulário através da leitura, indiretamente (KRASHEN; RODRIGO; GRIBBONS, 2004). A aquisição se daria através da inferência das palavras desconhecidas, propiciada pelo contexto. Porém, para que a aquisição de vocabulário através da inferência seja possível, alguns requisitos devem estar presentes, caso contrário a inferência bem sucedida pode até tornar-se impossível.

Segundo os autores citados abaixo, a fim de ser capaz de inferir os significados de palavras desconhecidas em contexto, o aprendiz deve ser capaz de reconhecer automaticamente a maior parte das palavras próximas. Vários pesquisadores têm tentado estimar a natureza e a extensão deste vocabulário. Nation e Hwang (1995) afirmaram que um vocabulário composto das 2.000 famílias de palavras mais frequentes normalmente permite aos aprendizes reconhecer aproximadamente 84% das palavras presentes em uma ampla variedade de textos escritos. Mas 84% não é o suficiente para a compreensão geral de textos<sup>13</sup>, o que, segundo Nation (1990) e Laufer (1997), requer conhecimento prévio de pelo menos 95% das palavras em um texto, as 3.000 famílias mais frequentes, aproximadamente. O conhecimento de 95% das palavras do texto significa que 1 palavra em cada 20, aproximadamente, será desconhecida para o aprendiz e, caso seja relevante para o entendimento do texto, terá de ser inferida. Para que o aprendiz seja capaz de uma leitura que exija apenas uma inferência ocasional de palavras, e para que essa inferência seja bem-sucedida<sup>14</sup>, é

---

<sup>13</sup> A compreensão geral é definida como o entendimento pelo aprendiz da idéia principal de uma passagem através da combinação de itens de informação aparentemente não relacionados. O entendimento específico, por outro lado, exige que o leitor atente para informação factual explicitamente declarada em uma passagem do texto a fim de ser capaz de identificar se afirmações específicas a respeito dessa passagem são verdadeiras ou falsas (Oh, 2001).

<sup>14</sup> Considera-se inferência bem-sucedida aquela que dá à palavra alvo o significado que a maioria dos leitores nativos dariam ao encontrá-la no mesmo contexto.

necessário que conheça 98% das palavras do texto, caso em que apenas 1 palavra em cada 50, aproximadamente, seria desconhecida. Isso exige o conhecimento de aproximadamente 5.000 famílias (HIRSCH; NATION, 1992).

A inferência bem-sucedida encontra ainda outros obstáculos. De acordo com Laufer (LAUFER, 1997), os aprendizes podem não reconhecer palavras desconhecidas como tal. Nesse estudo, foi pedido aos sujeitos que lessem um texto e auto-avaliassem seu conhecimento do vocabulário; a etapa seguinte consistia em um teste no qual era pedido aos sujeitos que traduzissem palavras do texto. A comparação da auto-avaliação subjetiva dos sujeitos com os resultados do teste objetivo mostrou que todos os aprendizes super-avaliaram seu conhecimento do vocabulário, em alguns casos em até 60%. Uma razão para isto pode ser que os aprendizes que entendem a mensagem geral do texto não prestam atenção aos significados precisos de palavras individuais. Uma outra razão é a de que os aprendizes confundem uma palavra desconhecida com outra conhecida. Por exemplo, um aprendiz que encontre a palavra *adapt* pode pensar estar lendo *adopt* ao invés de uma palavra desconhecida, que deveria ser inferida ou procurada no dicionário. Palavras homônimas, falsos cognatos, palavras com uma estrutura enganosa (ex. *shortcomings*, a qual foi inferida como “visitas curtas”; *outline*, a qual foi inferida como “fora da linha”), também induzem à confusão. Em outro estudo de Laufer e Yano (LAUFER; YANO, 2001) todos os aprendizes super-avaliaram seu conhecimento das palavras alvo em um texto, afirmando conhecer palavras que de fato não conheciam. Isso significa que os sujeitos pensaram conhecer a palavra ao encontrá-la ou julgaram ter inferido seu significado corretamente. O estudo demonstrou, ainda, que a eficiência da auto-avaliação de conhecimento lexical é relacionada ao nível do aprendiz. Os aprendizes de níveis básicos e os de nível avançado, de proficiência próxima ao do falante nativo, demonstram um conhecimento mais preciso das palavras que efetivamente conhecem ou não. Os aprendizes que se

situam entre esses níveis declaram, com maior frequência, conhecer palavras as quais na verdade não conhecem.

Para que o aprendiz infira corretamente o significado de uma palavra é preciso que o texto forneça pistas suficientes. Na sentença “Eu vi um X ontem à noite” não há como saber quem ou o que é “X”. As pistas podem até mesmo ser enganosas. Na sentença “As pessoas estavam bebendo, cantando, rindo e brigando”<sup>15</sup>, usada em um estudo de Laufer e Nation (LAUFER; NATION, 1995), a maioria dos sujeitos interpretou “brigando” (*brawling*) como “divertindo-se”.

Em um estudo de Bengelil e Paribakht (BENGELEIL; PARIBAKHT, 2004) sobre inferência a partir do contexto com aprendizes de inglês como LE, os aprendizes intermediários inferiram corretamente 35,5 %, parcialmente 13,7% e erraram 50,8% das palavras. Os avançados estiveram corretos em 46,6%, parcialmente corretos em 19,4% e errados em 34% das inferências. Considerando-se a porcentagem de erros e acertos parciais dos aprendizes intermediários e avançados, pode-se avaliar a dificuldade que aprendizes iniciantes têm para inferir significado.

Se aprendizes intermediários e avançados enfrentam vários obstáculos à aprendizagem indireta através da inferência, como vimos acima, aprendizes iniciantes nem mesmo têm a opção de tentar superar esses obstáculos devido à limitação de seu vocabulário. O que esses aprendizes precisam é de um vocabulário extenso o suficiente para permitir que comecem a usar a língua. Devido às porcentagens de conhecimento de palavras exigido para a inferência, mencionadas acima, Waring (1995) recomenda que se deixe as tentativas de aprendizagem através da inferência para um estágio posterior, quando o aprendiz já possuir um vocabulário mínimo para permitir alguma compreensão, por volta de 1.000 famílias ou até mais. Isso significaria começar com uma ênfase na aprendizagem direta, passando depois à aprendizagem indireta. Scaramucci (1995) registra opinião similar quando, falando das condições necessárias para a aprendizagem de vocabulário de forma indireta, afirma:

---

<sup>15</sup> *People were drinking, singing, laughing, brawling*

(...) essas condições parecem estar presentes mais freqüentemente em contextos de LM, menos em contextos de L2, e muito menos ou quase nunca em contextos de LE. Elas não implicam apenas no engajamento do aprendiz em um grande volume de leitura extensiva, mas, principalmente, no desenvolvimento de uma base lingüística e habilidades de inferência, ou de uso desse léxico através do contexto. A utilização de uma abordagem indireta em estágios iniciais, ou de um ensino não intervencionista de vocabulário, combinado com uma situação em que a qualidade e quantidade de insumo recebido são geralmente bastante limitados, tem sido responsável por grande parte da frustração e incapacidade dos alunos, principalmente frente aos textos de leitura.

Meara (1995) defende a aprendizagem de vocabulário através da memorização de listas de palavras para aprendizes iniciantes. O autor afirma que nesse estágio um número relativamente pequeno de palavras aumenta muito a compreensão dos aprendizes. As 1.000 famílias de palavras mais freqüentes do inglês, por exemplo, permitem uma cobertura de 84,3% da conversação, 82,3% dos textos de ficção e 71,4% dos textos acadêmicos (HU; NATION, 2000). Outra razão mencionada pelo autor é a de que no estágio inicial os aprendizes esperam ter de aprender grande número de palavras por julgarem, muitas vezes, que aprender uma língua significa aprender vocabulário. Faz sentido aproveitar essa expectativa para incentivar a aprendizagem de novas palavras rapidamente e embora, como vimos acima, uma cobertura de aproximadamente 80% não permita ao aprendiz o uso confortável da língua, 1.000 famílias já permitem uma comunicação que, embora imperfeita, coloca as palavras conhecidas em uso e propicia oportunidades para a retenção e aumento do vocabulário.

Outro fator na aprendizagem indireta é o número de encontros necessário para que o aprendiz perceba (*notice*), infira e retenha uma palavra. A freqüência de ocorrências aumenta a possibilidade de que o aluno perceba determinada palavra, infira seu significado e refine suas inferências, chegando a um conhecimento mais profundo das palavras. Não existe acordo, entretanto, sobre quantos e que tipos de encontros são necessários. Nagy, Herman e Anderson (1985), em um estudo em L1, estimaram que a probabilidade de aprender uma palavra indiretamente ao ponto de

ser capaz de fornecer uma definição completa e precisa com apenas um encontro é de 0,11<sup>16</sup>. Um estudo seguinte (HERMAN; ANDERSON; PEARSON; NAGY, 1987), entretanto, chegou a uma probabilidade ainda menor, entre 0,05 e 0,10 para aprendizes menos hábeis em leitura e entre 0,26 e 0,46 para leitores mais hábeis. Uma pesquisa de Nation (1990) desses e de outros estudos concluiu que são necessários entre 5 e 16 encontros para a aquisição. De qualquer forma, existem tantas variáveis envolvidas em aprender uma palavra que é impossível determinar qualquer limiar para o número de encontros. Muito depende de fatores como a saliência da palavra em um dado texto, seu reconhecimento como cognato, sua morfologia, o interesse do aprendiz e a disponibilidade e riqueza de pistas no contexto. Vários outros trabalhos também abordam a dificuldade envolvida na aprendizagem indireta (READ, 2004; ZIMMERMAN; BOYD, 1997; HUCKIN; COADY, 1999; PARIBAKHT; WESCHE, 1999).

Na aprendizagem indireta, os fatores saliência e relevância, que podem alterar o número de encontros necessários para a retenção, não podem ser controlados, dependendo de muitas variáveis. Na aprendizagem direta pode-se tentar interferir nesses fatores, aumentando a relevância e/ou saliência das palavras-alvo.

A força de uma inferência incorreta fica clara na dependência do conhecimento lexical para a construção de sentido mostrada pelos sujeitos da pesquisa de Scaramucci (1995, p.249):

Ao perceber que o significado não se encaixa no contexto, e apesar da dificuldade em se construir um significado com a palavra, o sujeito, entretanto, não consegue abandonar a hipótese. O comportamento é curioso. Ao invés de reformular sua hipótese, o leitor acaba distorcendo o sentido do texto em construção para encaixar a palavra ou apoiar a inferência incorreta.

Acreditamos que a combinação da aprendizagem indireta com a aprendizagem direta proporciona os melhores resultados na aquisição de vocabulário. Durante nossa pesquisa para este trabalho, encontramos poucos estudos sobre a combinação da aprendizagem direta e indireta. Um

---

<sup>16</sup> Uma probabilidade é um número entre 0 e 1. Quanto mais próxima de 1 for a probabilidade de um evento, mais provável que o evento ocorra.

deles, de Paribakht e Wesche (1997), concluiu que o grupo de sujeitos que receberam instrução direta de vocabulário combinada com leitura tiveram maiores ganhos de vocabulário do que o grupo que apenas se dedicou à leitura. Outro estudo, de Mason e Krashen (2004), chegou à conclusão contrária, embora os autores façam a ressalva de que, como os sujeitos do grupo que foi exposto apenas à leitura sabiam quais eram as palavras alvo, ambos os grupos estavam com a atenção voltada para as palavras que seriam testadas posteriormente, e isso descaracteriza o grupo somente leitura como puramente indireto. Acreditamos que essa questão merece estudos mais abrangentes devido a sua importância para a aprendizagem de línguas.

Em seu trabalho sobre ensino direto e indireto de vocabulário Hunt (2005) afirma que o ideal seria a combinação de ambos. Passa, então, a apresentar sugestões para o ensino e aprendizagem diretos e indiretos sem, porém, sugerir qualquer forma de combinar ou agrupar as atividades diretas e indiretas.

Acreditamos que uma alternativa seria o uso de atividades de ensino direto de vocabulário seguidas de atividades que apresentem o vocabulário abordado em contexto, e que favoreçam a aprendizagem indireta.

Uma interessante possibilidade, a meu ver, seria a combinação das atividades utilizadas por Zilles (2001), em seu trabalho sobre aquisição direta e indireta de vocabulário, comparando dois grupos de alunos expostos a 10 palavras alvo. O autor utilizou uma abordagem direta para a apresentação das palavras com um dos grupos e uma abordagem indireta com o outro. Acredito que a combinação das duas abordagens realizadas pelo autor consiste em um bom exemplo de como se pode combinar o ensino direto e o indireto de vocabulário. A seguir apresentamos um exemplo das atividades utilizadas pelo autor.

Na abordagem direta os alunos tiveram 6 encontros com as palavras. No primeiro encontro as 10 palavras alvo foram apresentadas aos alunos e feita a prática de pronúncia. A seguir os alunos, em dois grupos, receberam 10 tiras de papel com as palavras alvo e discutiram sobre o

possível significado das mesmas. Depois receberam 10 tiras com frases contendo as palavras alvo e tentaram inferir o significado pelo contexto. Finalmente receberam 10 tiras com os significados das palavras e foi pedido que combinassem a palavra com o significado e com a frase de exemplo.

No segundo encontro foi realizada uma competição entre os alunos, na forma do jogo da velha com palavras. O professor apresentava a definição ou um exemplo de uso da palavra e os alunos deviam produzi-la.

No terceiro encontro os alunos receberam um cartão com a palavra alvo escrita em um dos lados e sua definição no outro. Os alunos caminhavam pela sala-de-aula e mostravam o lado do cartão com a palavra escrita para um colega de classe, se este soubesse o significado, dizia-o, caso contrário o primeiro aluno dava o significado ao colega.

No quarto encontro o professor afixou 10 tiras de papel contendo perguntas onde apareciam as palavras alvo. Os alunos caminhavam em pares pela sala respondendo as perguntas

No quinto encontro o professor elicitou dos alunos as palavras alvo. Pediu então que cada um escolhesse 4 delas que considerassem as mais difíceis e que escrevessem uma frase com cada uma delas.

No sexto encontro pares de alunos receberam 4 cartões com as palavras alvo e as escreveram em seus cadernos. O professor pediu que escrevessem um diálogo ou parágrafo que contivesse as 4 palavras alvo.

O grupo indireto também teve 6 encontros com as palavras alvo, a partir da leitura de textos apresentados de acordo com os seguintes passos:

- a) uma discussão oral prévia sobre o tópico a ser tratado era realizada, de modo a preparar os alunos para a leitura (objetivo específico do trabalho) e desenvolver habilidades orais (objetivo geral da escola);
- b) todas as discussões eram realizadas em inglês;
- c) era feita a leitura silenciosa do texto em sala de aula;
- d) o professor não forneceu ao aluno o significado de nenhuma palavra dos textos, independentemente de elas pertencerem ou não ao grupo de palavras-alvo;

- e) havia sempre cerca de 4 dicionários à disposição dos alunos, para a verificação opcional de palavras desconhecidas;
- f) as atividades que se seguiam à leitura dos textos eram baseadas no tópico em geral, no desenvolvimento de habilidades produtivas (fala e/ou escrita) ou em pontos gramaticais coerentes com o programa do nível em questão; nenhuma atividade foi baseada na compreensão explícita de nenhuma das palavras-alvo (Zilles, 2001).

Para o conhecimento de outras atividades realizadas e maiores informações sobre as condições e resultados da pesquisa, vide Zilles (2001).

### **2.11 Textos modificados**

Outra área em que as listas são muito úteis é a da preparação de textos modificados para aprendizes. Textos modificados são textos adaptados para facilitar a aprendizagem de línguas. Muitos tipos de modificações são possíveis e se dividem em “simplificação” e “elaboração”. Características típicas da simplificação lingüística incluem o uso de sentenças mais curtas, sintaxe mais simples, vocabulário mais simples, apagamento de elementos da sentença e preferência pela ordem gramatical de palavras (*canonical word order* - sujeito, verbo e objeto). A elaboração envolve, por exemplo, o aumento da redundância, paráfrases, uso de sinônimos e repetições, fala mais lenta e articulação mais clara. As listas de palavras são usadas pelas editoras para orientar o vocabulário a ser utilizado e os níveis dos textos modificados.

Textos simplificados são textos (normalmente livros) com vocabulário e, algumas vezes, gramática simplificados, escritos em vários níveis progressivos de dificuldade, para aprendizes de L2 e de LE. Eles são encontrados em vários gêneros, adaptações de clássicos da literatura, histórias originais, não-ficção como biografias, etc. São escritos para vários públicos, idades e níveis de proficiência. Esses textos permitem que sejam contornados os obstáculos à inferência mencionados acima. Podem ser escritos para que os aprendizes encontrem as palavras alvo com uma frequência

que aumente a chance de aprendizagem, encontrando as palavras mais vezes, os aprendizes têm mais chances de perceber a palavra, de inferi-la e de refinar essa inferência até chegar ao significado correto, caso isso não tenha acontecido na primeira inferência. Como há a probabilidade de se encontrar a palavra várias vezes no mesmo texto, a probabilidade de uma inferência errada sedimentar um significado errado na memória é menor, já que ao encontrar a palavra novamente o aprendiz poderá verificar se o significado anteriormente inferido se adapta ao contexto atual. Podem também fornecer mais pistas para facilitar a inferência das palavras alvo. Os textos modificados, como são escritos com graus progressivos de dificuldade, podem se prestar a uma leitura fluente, quando os textos utilizados estão no nível de proficiência de cada aprendiz, caso em que este encontrará poucas palavras desconhecidas e terá oportunidade de refinar inferências feitas anteriormente, confirmar hipóteses, reencontrar palavras já vistas e inferidas, aumentando a retenção das mesmas. Os textos modificados também podem se prestar a uma leitura mais voltada para a aprendizagem de novo vocabulário. O aprendiz pode utilizar um texto que seja dirigido a aprendizes um nível acima do seu. Nesse caso, terá maior dificuldade para fazer inferência a partir do contexto e isso lhe permitirá praticar a habilidade de inferência, as palavras desconhecidas que não possam ser inferidas podem ser procuradas no dicionário, o que aumentará a percepção sobre a palavra. Além disso, muitos textos modificados vêm com exercícios extras, de vocabulário, interpretação ou gramaticais, que serão úteis ao aprendiz em seu estudo.

Na mesma linha dos estudos sobre o número de encontros necessários para a aprendizagem de uma palavra, um estudo de Waring (2003) sobre a retenção de palavras aprendidas 3 meses após os aprendizes lerem um texto simplificado mostrou que as chances de reconhecer uma palavra encontrada 8 vezes 3 meses depois é de 50 %, número válido para o reconhecimento estimulado, onde são apresentadas alternativas para que o aluno escolha a palavra correta. Para o reconhecimento não-estimulado as chances são de 10 a 15% mesmo depois de terem encontrado a

palavra 18 vezes ou mais. Se a palavra foi encontrada apenas uma vez as chances são próximas de zero. O autor sugere que podem ser necessários 30, 40 ou mais encontros para que uma palavra seja aprendida e retida através da leitura. Esse estudo foi feito com um texto simplificado, com palavras sendo encontradas 18 ou mais vezes. Os resultados seriam ainda mais desencorajadores se fossem usados textos autênticos, já que estes propiciam muito menos encontros quando comparados com textos modificados. Os textos modificados tornam menor o volume de leitura necessário visto que proporcionam um grande número de reencontros das palavras mais úteis para cada nível.

O estudo ainda menciona os textos modificados como úteis para aumentar a profundidade de conhecimento do vocabulário. Para que um aprendiz iniciante aumente seu vocabulário em 10%, podem ser necessárias apenas 50 ou 100 palavras, já um aprendiz avançado pode precisar de centenas ou milhares de palavras para atingir a mesma porcentagem. Por isso a dificuldade em incluir um número significativo de palavras novas com a frequência de encontros exigida aumenta muito nos níveis mais altos. Os textos modificados permitem que, ao invés de se concentrar na aprendizagem de novo vocabulário, os autores apresentem contextos mais ricos para as palavras já conhecidas, novos significados e novas colocações, por exemplo, permitindo aos aprendizes aumentar a profundidade de conhecimento do vocabulário que já possuem.

Os textos modificados também permitem que se destaque as palavras alvo, o que pode, segundo alguns estudos (BARCROFT, 2003; RIDDER, 2002), aumentar a chance do aprendiz perceber a palavra. Algumas editoras de textos modificados já estão fazendo isso em suas séries de livros de níveis iniciais (ex. *Dominoes*, da *Oxford University Press* e *Black Cat*, da *Cideb*). No caso da série *Dominoes*, palavras são destacadas em negrito, e um glossário é incluído em cada página com definições dessas palavras.

Alguns argumentos contra os textos simplificados apontam, na verdade, não para falhas no conceito dos textos simplificados, mas sim para falhas na elaboração dos mesmos, tais como: linguagem pobre, falta de conteúdo (que também ocorrem em textos autênticos, não simplificados); processo de simplificação que deixa o texto incompreensível, etc. Acreditamos que todo e qualquer tipo de material didático pode ser arruinado por falta de cuidado na elaboração, isso não deve anular a validade dos conceitos que o embasam. A esse respeito Bamford e Day (1997) argumentam que:

A artificialidade percebida em textos que foram simplificados ou especificamente escritos para aprendizes de línguas aparece quando os escritores ou editores estão menos preocupados com a comunicação do que com, por exemplo, a necessidade de reduzir uma história complicada a umas poucas páginas de texto. Textos simplificados ruins são escritos, mas existem centenas de livros de excelente qualidade adaptados ou escritos para aprendizes de línguas de todos os níveis de proficiência. A qualidade e variedade de tais textos em inglês e outras línguas é tal que merecem ser chamados de literatura para aprendizes de línguas, exatamente como existe literatura infantil e literatura juvenil.

Uma objeção muito comum aos textos simplificados é a de que eles não são autênticos, que materiais simplificados não representam a língua naturalmente e negam ao aprendiz a oportunidade de encontrar linguagem natural.

A noção de autenticidade tem sido muito debatida no campo da Lingüística Aplicada e pedagogia de línguas e McNeill (1994) afirma que "é difícil para os lingüistas aplicados concordarem sobre o que exatamente autenticidade significa". Lewkowitz (1997) argumenta que "um dos problemas subjacentes à autenticidade é que tem havido poucas tentativas de definir exatamente o que ela significa". O termo autenticidade emergiu no final dos anos 70 quando o ensino de línguas comunicativo começou a ganhar impulso. Ele é, portanto, relacionado ao uso da língua para a comunicação no mundo real. Widdowson (1979) afirma que a autenticidade é uma noção para distinguir entre textos da "vida real" daqueles escritos para fins pedagógicos. Segundo ele, autenticidade é uma característica do relacionamento entre a passagem e o leitor, tem a ver com a

resposta apropriada do leitor ao texto e, portanto, é extrínseca ao texto. Ele enfatiza que a autenticidade não depende da fonte mas da relação do aprendiz com ela (WIDDOWSON, 1978). A expressão "material autêntico" está aberta a um número de interpretações. Um ponto de vista largamente aceito é o de que o insumo é considerado autêntico quando "ele é produzido por e para falantes nativos da língua alvo" (BACON; FINNEMANN, 1990). Little, Devitt e Singleton (1988) descrevem insumo autêntico como "criado para preencher algum propósito social na comunidade lingüística onde ele foi produzido". Nunan (1988) resume perspectivas diferentes sobre materiais autênticos ao dizer que "materiais autênticos são normalmente definidos como aqueles que foram produzidos para propósitos que não sejam o ensino de línguas".

Nos anos 80, muito do debate sobre autenticidade se concentrava nos textos que poderiam ser usados como insumo para os aprendizes (BREEN, 1985; MORRISON, 1989; SWAFFAR, 1985) e um texto era considerado como autêntico se não tivesse sido escrito com propósitos pedagógicos. Carroll (1980) acrescentou que uma aplicação integral do conceito de autenticidade significaria que a linguagem deveria ser o discurso do dia-a-dia, não editado ou modificado no interesse da simplificação mas apresentado com todas as irregularidades; que os contextos das interações deveriam ser realistas, com as interrupções normais, barulhos de fundo e irrelevâncias encontradas nas situações reais. Lewkowicz (1997) aponta que essa interpretação de autenticidade era restrita, já que polarizava a linguagem em autêntica ou simplificada para fins pedagógicos. Ela acredita ser "problemático igualar textos autênticos com textos não-simplificados, visto que textos simples escritos com base em textos genuínos podem ser mais acessíveis a uma dada audiência". Widdowson (1978) também acredita que a simplificação pode ocorrer mantendo a autenticidade do texto no sentido das respostas do aprendiz ao texto. Ele acrescenta que o texto tem de "prender a atenção do aprendiz e parecer a ele como sendo relevante aos seus interesses". Lynch (1982) afirma que "a simplificação contribui para o evento comunicativo atual e para o desenvolvimento da língua a longo-prazo". Lewkowicz (1997) concorda ao dizer que acredita que:

(...) os aprendizes constroem seu conhecimento de uma língua e não se pode exigir que lidem com textos não-simplificados desde o início. Se os aprendizes devem desempenhar no máximo de suas habilidades, eles deveriam ser expostos a textos que sejam apropriados ao seu nível de proficiência lingüística.

O conhecimento prévio dos aprendizes, seu interesse e proficiência são, portanto, assuntos intimamente relacionados quando interpretamos a noção de autenticidade do texto. Se os aprendizes considerarem a linguagem do insumo muito difícil, eles podem não estar dispostos nem mesmo a tentar usá-lo. Como vimos, Widdowson argumenta que autenticidade é o resultado da relação entre o leitor e o texto, não do próprio texto. Ele também aponta que é impossível que um texto não-simplificado seja autêntico em um contexto de LE simplesmente por que os aprendizes estão fora da comunidade para a qual o texto foi dirigido, e que textos autênticos não têm lugar na sala de aula, visto que ele nunca será autêntico e sugere, ao invés, que textos apropriados sejam selecionados para aprendizes (WIDDOWSON, 1998). O que torna o texto escrito por e para falantes nativos autêntico é que eles são ocorrências de comunicação entre escritor e audiência-alvo, texto autêntico é aquele que é lido pela audiência para a qual foi escrito. Então, quando um escritor se comunica com uma audiência-alvo de aprendizes de línguas em um nível específico de proficiência, o texto resultante é autêntico (BAMFORD; DAY, 1997).

Ainda sobre os textos autênticos, cabe a pergunta sobre como os defensores da aprendizagem indireta e do uso exclusivo de textos autênticos esperam que os aprendizes cheguem a um nível de proficiência que lhes permita lê-los. Caso se espere que os aprendizes adquiram vocabulário indiretamente, sem a leitura de textos simplificados, como vão chegar a ter um vocabulário que permita uma cobertura de pelo menos 95% dos textos autênticos, nível necessário para a compreensão geral e inferência a partir do contexto (inferência essa que, como vimos, ainda é laboriosa e passível de erro)? Espera-se que adquiram esse vocabulário a partir do discurso do professor e dos demais alunos? Em 3 horas de aulas por semana? Acreditamos que, caso se descarte

o ensino e aprendizagem diretos e a leitura de textos simplificados, os aprendizes talvez nunca cheguem a ter um vocabulário que lhes permita ler textos não-simplificados.

Quanto à questão de que a leitura dos textos simplificados tornaria os aprendizes incapazes de ler textos não-simplificados, deve-se ter em mente que os textos simplificados são graduais e seu uso deve começar com os de níveis mais básicos e prosseguir com o aumento gradual do nível de dificuldade. Os textos simplificados de níveis avançados diferem muito pouco dos textos não-simplificados, a distância daqueles para estes é muito pequena e essa passagem não apresenta qualquer dificuldade para os aprendizes. Quanto à possibilidade de utilizar textos não-simplificados mais simples para os iniciantes, como por exemplo livros infantis, deve-se levar em conta que os livros infantis também têm vocabulário simplificado, para que a criança não tenha dificuldade na leitura. Porém a simplificação não visa às necessidades do aprendiz de LE, seria mais prodcente uma simplificação que levasse em conta essas necessidades. Além disso, livros infantis freqüentemente usam vocabulário típico do universo infantil, que é muito pouco útil para aprendizes adultos; estes podem, ainda, sentir-se constrangidos e desmotivados por estarem lendo material para crianças, além de que esse tipo de texto dificilmente será interessante para adultos. Vale argumentar ainda que, seguindo-se esse raciocínio, não se deveria permitir que crianças lessem textos infantis pois estas poderiam nunca desenvolver a habilidade de ler textos “adultos”. Os textos simplificados não têm, necessariamente, de ser a única fonte de leitura dos aprendizes, textos simplificados e textos não-simplificados podem ser usados concomitantemente. Zahar, Cobb e Spada (2001) mencionam esse tipo de uso ao sugerirem que :

Visto que os aprendizes aparentemente não aprendem muito bem a partir de contextos ricos<sup>17</sup> (devido ao pouco processamento que esse tipo de texto exige), e não podem aprender nada de contextos totalmente opacos, poderia ser na alternância de contextos mais ou menos informativos que as palavras podem ser aprendidas. Um mecanismo especulativo poderia ser o seguinte: um contexto obscuro ou semi-claro dispara uma necessidade de aprendizagem ou lacuna

---

<sup>17</sup> Contextos ricos são aqueles que oferecem muitas pistas para facilitar a inferência das palavras desconhecidas.

conceitual, a qual é então reativada quando a palavra é eventualmente encontrada em um contexto claro.

Vale lembrar que, pela graduação de níveis de dificuldade que é a própria base dos textos simplificados, eles permitem que se escolha maior ou menor riqueza de pistas contextuais de acordo com o objetivo de ensino.

### **2.12 Avaliação de vocabulário**

Uma terceira área na qual as listas de palavras (gerais, acadêmicas, técnicas) são úteis é a da avaliação de vocabulário. Existem várias razões pelas quais se avalia vocabulário. É importante saber, por exemplo, se os aprendizes estão aprendendo e retendo o vocabulário apresentado pelo material didático utilizado pelo curso para que se possa, caso os resultados não sejam satisfatórios, fazer ajustes na abordagem de ensino e aprendizagem de vocabulário. Nesse caso, naturalmente, o vocabulário a ser testado vai ser aquele apresentado pelo material didático. Testar vocabulário também é importante para se saber a dificuldade da tarefa de aprendizagem de vocabulário que enfrentam estudantes estrangeiros. É útil saber a extensão do vocabulário dos aprendizes para que se possa tomar medidas para ajudá-los a aproximar seu vocabulário daquele dos falantes nativos.

Nation (1993), referindo-se ao fato de que a aprendizagem das 1.000 palavras mais frequentes da língua inglesa oferece um retorno muito grande, ressalta a importância de poder avaliar o conhecimento que os aprendizes têm dessas palavras. Laufer e Nation (1999) argumentam que tem havido um crescimento no interesse sobre testes de vocabulário e que isso pode ser interpretado como significando que existe considerável valor em se ganhar conhecimento sobre partes específicas

do conhecimento lingüístico dos aprendizes, já que esse conhecimento pode ser usado eficientemente para classificação, diagnóstico e desenvolvimento de currículo. Schmitt (2000) argumenta que a atitude dos aprendizes é uma importante razão para se testar vocabulário. Segundo ele, os aprendizes julgam a importância do material ensinado em sala-de-aula em parte por este aparecer ou não nos testes subsequentes, o que é chamado de efeito retroativo. Meara e Fitzpatrick (2000) afirmam que:

(...) na maioria dos contextos práticos, fica claro que a eficiência comunicativa é conquistada com maior sucesso por aprendizes com um vocabulário mais extenso do que por aprendizes com um comando mais detalhado de um vocabulário menor. Não é surpreendente, então, que a mensuração do tamanho do vocabulário tenha se correlacionado positivamente com níveis de proficiência na leitura e escrita e proficiência lingüística geral; daí a importância de se testar vocabulário.

Acreditamos que a avaliação de vocabulário é indispensável para professores e pesquisadores. Os professores podem, através dela, ter uma visão clara do progresso de seus alunos na aquisição de vocabulário. Como em muitos casos o desenvolvimento do vocabulário é deixado a cargo de uma eventual e incerta aprendizagem indireta, a avaliação pode permitir ao professor verificar se os aprendizes estão progredindo nesse aspecto e tomar medidas para corrigir eventuais problemas. Os pesquisadores têm utilizado a avaliação de vocabulário para: avaliar a aquisição e retenção de vocabulário através da leitura (ROTT, 1999), prever o desempenho acadêmico de futuros professores de inglês como L2 (COBB; MORRIS, 2004), medir como a procura de informação sobre palavras influencia a aprendizagem e retenção das mesmas (LAUFER; HILL, 2000), avaliar o papel do vocabulário na compreensão em leitura em um contexto de língua estrangeira (SCARAMUCCI, 1995), medir a extensão do vocabulário receptivo (EYCKMANS, 2004) e avaliar a relação entre conhecimento de vocabulário e desempenho em leitura acadêmica (QIAN, 2002), para citar alguns usos.

Com a exceção dos testes de vocabulário que visam a avaliar a aprendizagem do vocabulário específico de um curso, texto, livro ou material didático, os testes de vocabulário escolhem as palavras a serem testadas a partir de uma lista de palavras, de acordo com o nível que

pretendem testar. A vasta maioria dos testes de vocabulário se utilizam de listas de palavras para orientar os níveis dos mesmos. A UWL e a AWL são freqüentemente utilizadas em testes que visam avaliar o conhecimento de vocabulário acadêmico. Acreditamos que a LPA-BR pode se prestar ao mesmo uso para testes que visem avaliar o conhecimento de vocabulário acadêmico em português.

### **2.13 Escolha de palavras**

Tanto para o ensino direto de vocabulário quanto para a elaboração de textos modificados e de testes de vocabulário faz-se necessário escolher as palavras a serem utilizadas.

Pode-se argumentar que a escolha do vocabulário a ser estudado deve ser feita de acordo com a necessidade dos aprendizes. O vocabulário alvo deve ser aquele que o aprendiz precisará conhecer quando se encontrar em situações reais de comunicação. Mas quão real é o contexto comunicativo de sala de aula? Acreditamos que os aprendizes de LE aprendem uma língua, prioritariamente, para conversar com falantes nativos, ler textos escritos por nativos ou escrever textos que serão lidos por nativos. Isso pode ser menos verdadeiro para a língua inglesa, que se tornou a língua da comunicação mundial de fato, fazendo com que muitos aprendizes da língua inglesa venham a utilizá-la para conversar com falantes não-nativos, ler e escrever textos de e para não-nativos. Mesmo nesse caso, uma grande parcela do uso da língua ainda será com nativos. No caso das outras línguas esse fato é ainda mais verdadeiro. Um aprendiz estrangeiro da língua portuguesa vai, na grande maioria das vezes, utilizá-la com nativos, o mesmo acontecendo com a maioria das outras línguas. O vocabulário que se faz necessário em contexto de sala de aula, em país onde a língua alvo é língua estrangeira, pode não refletir o vocabulário que o aprendiz efetivamente utilizará em suas atividades profissionais, acadêmicas ou sociais com falantes nativos da língua alvo.

Palavras que podem ser muito úteis na cultura e no país de um determinado aprendiz podem não ter qualquer importância no país onde a língua alvo é falada. Palavras que podem ser muito úteis para o aprendiz quando este se encontrar no país da língua alvo ou em comunicação com nativos podem nunca ter sido estudadas. As listas de palavras, baseadas como são em corpora compilados a partir da língua efetivamente utilizada por nativos, são um retrato do vocabulário realmente utilizado por nativos em país onde se fala a língua alvo. Listas bem elaboradas a partir desses corpora são fontes confiáveis de orientação sobre que vocabulário estudar.

Vários critérios podem ser utilizados na escolha dessas palavras: cobertura, facilidade de aprendizagem, semelhança (cognatos), frequência e alcance, entre outros. Acreditamos que frequência e alcance sejam, por si sós, importantes o suficiente para garantir que uma palavra que atenda a esses dois critérios seja incluída em uma lista que visa apresentar as palavras mais úteis. Se uma palavra ocorre frequentemente em vários tipos diferentes de textos, ela fatalmente será encontrada com frequência pelos aprendizes e deve ser considerada para ser ensinada. Uma escolha baseada somente na frequência pode apresentar distorções devido à existência de palavras muito frequentes em apenas um tipo de texto. Coxhead (2000) cita o exemplo do *Collins COBUILD Dictionary* de 1995 que apresenta as palavras *Yemeni* e *Lithuanian* como palavras frequentes, provavelmente devido ao fato de o corpus que serviu de base para o dicionário conter um grande número de jornais do início dos anos 90. A inclusão do critério “alcance” garante que uma palavra só vai ser incluída na lista caso seja frequente e ocorra em um determinado número de textos diferentes. No caso da LPA-BR estabelecemos esse limiar em um mínimo de 4 diferentes áreas do conhecimento.

Vários autores consideram a frequência e o alcance como fatores muito importantes na escolha de determinada palavra. Eyckmans (2004) afirma que: “Parece evidente que as palavras mais frequentes são mais úteis e deveriam ser ensinadas antes (...)” e também que: “Enquanto outros critérios podem ser judiciosamente aplicados, frequência e alcance parecem oferecer o mais objetivo, empiricamente comprovado e menos controverso ponto de partida para avaliar a utilidade básica do vocabulário abordado pelo material didático”. Umamoto (1989) argumenta que: “(...) palavras de alta frequência e grande alcance podem ser vistas como palavras de máxima utilidade e de primeira importância para a aprendizagem de línguas”. Nation e Waring (2002) afirmam que:

A maioria dos estudos sobre frequência têm reconhecido a importância do alcance. Uma palavra não deveria tornar-se parte de uma lista de palavras porque ela ocorre frequentemente. Ela deveria ocorrer frequentemente em uma grande gama de textos. Isso não significa que sua frequência tem de ser aproximadamente a mesma em todos os textos, mas sim que ela deveria ocorrer na maioria dos textos ou grupos de textos.

Bowles (2001) diz: “Há muito tempo tem sido reconhecido que embora a frequência e o alcance de ocorrência textual de uma palavra não sejam os únicos critérios possíveis para sua seleção para o ensino de inglês como L2, eles são cruciais”, e também que: “A evidência de corpus relativa à frequência e alcance de uma palavra fornece dados úteis e objetivos para assegurar que o que estamos ensinando aos aprendizes é altamente útil e característico da língua como um todo”.

Meara (1995) defende uma ênfase na aprendizagem das 1.000 famílias mais frequentes tão rapidamente quanto possível nos estágios iniciais da aprendizagem da língua inglesa; Schmitt (2000) advoga que o vocabulário deveria ser ensinado a aprendizes de LE de acordo com uma perspectiva de custo-benefício, ele menciona que a aprendizagem das 2.000 famílias mais frequentes é o principal objetivo para aprendizes iniciantes e que essas palavras deveriam ser ensinadas diretamente.

Nation (1990), após ressaltar a importância da frequência e do alcance ao afirmar: "Nós podemos obter informação sobre quais palavras serão mais úteis para os aprendizes de inglês olhando as contagens de frequência de vocabulário", e também: "As palavras mais úteis para os aprendizes são palavras de alta frequência com grande alcance", passa a apresentar alguns problemas do uso exclusivo de frequência e alcance como critérios para a elaboração de listas de palavras, os quais passamos a comentar.

Primeiramente, Nation afirma que muitas palavras importantes não ocorrem entre as primeiras 1.000 palavras mais frequentes<sup>18</sup>, isso é verdade mas acreditamos que se existem palavras importantes que não ocorrem entre as primeiras 1.000 isso não muda o fato de que as primeiras 1.000 são muito importantes e devem ser ensinadas ao aprendiz iniciante. O autor também afirma que existem palavras que não são adequadas para aprendizes iniciantes e que se encontram entre as primeiras 1.000, dando como exemplo as palavras *bank* como verbo, *bill*, *company*, *deal*, *issue*, *labor*, *stock* e *supply* e afirmando que essas palavras seriam mais frequentes em cartas comerciais e que não seriam muito frequentes em outros tipos de texto. Acontece que se palavras que ocorrem em apenas um tipo específico de texto foram incluídas nessa lista isso significa que ou o critério alcance não foi levado em conta ou o corpus não era homogêneo, pendendo fortemente para textos comerciais. Um corpus homogêneo, com equilíbrio entre os vários tipos de texto, e que só incluía palavras que tenham um alcance razoável não apresentaria palavras frequentes em apenas um tipo de texto. Outra restrição apresentada pelo autor é a de que listas diferentes apresentam palavras diferentes como mais frequentes. Isso é de se esperar, já que a lista final vai depender do corpus a partir da qual ela foi elaborada. Por isso cada pesquisador deve ter em mente muito claramente que tipo de lista pretende compilar, e formar um corpus de acordo com esse objetivo. Na LPA-BR, por exemplo, deixamos claro que buscamos um vocabulário frequente e de grande alcance usado em

---

<sup>18</sup> Quando fala das primeiras 1.000 palavras mais frequentes Nation está se referindo às palavras da lista de Thorndike e Lorge (Thorndike, 1944).

textos acadêmicos e que os textos que compõem o corpus são oriundos de teses e dissertações de pesquisadores de universidades públicas brasileiras. Nas listas gerais é importante que sejam informados os textos que compuseram o corpus para que se possa utilizar a lista resultante consciente de suas limitações. O autor também afirma que a seqüência em que as palavras são apresentadas não é a melhor seqüência para se ensinar aos aprendizes. Uma lista de palavras serve como orientação sobre que palavras o aprendiz vai encontrar com mais freqüência e em grande diversidade de textos, a abordagem pedagógica a ser utilizada para o ensino desse vocabulário deve ser elaborada pelos professores ou pesquisadores, levando em conta as características dos aprendizes, contexto de ensino, tempo disponível, objetivo, etc. A lista não é feita para ser seguida cegamente como um projeto pedagógico acabado. Além disso, não estamos afirmando que apenas freqüência e alcance devam ser utilizados e que outros critérios devam ser evitados, apenas defendemos que freqüência e alcance, por si só, justificam a inclusão de uma palavra.

Koprowski (2005) realizou um estudo sobre a escolha de frases lexicais a serem apresentadas em materiais didáticos para o ensino de inglês. Foram objeto do estudo três livros para o nível intermediário: *New Headway Upper-Intermediate* (SOARS; SOARS, 1998), *Innovations* (DELLAR; HOCKING, 2000) e *Inside Out Intermediate* (KAY; JONES, 2000). Dos 822 itens lexicais examinados nos três livros, 118 (mais de 14%) não eram encontrados no corpus COBUILD (composto de mais de 330 milhões de palavras) **nem mesmo uma vez**. Isso indica que a probabilidade de serem encontrados pelos aprendizes na vida real é praticamente zero, e levanta dúvidas sobre se deveriam fazer parte de um material dirigido a aprendizes de nível intermediário. A esse respeito o autor comenta: “Visto que existem centenas de milhares de itens lexicais para se escolher, a escolha desses itens ao invés de outros potencialmente mais úteis sugere um processo de escolha sem critérios e descuidado”. Outra descoberta que causa surpresa é o fato de que, entre os 822 itens lexicais presentes nos três livros didáticos, apenas **7** estão presentes em mais de um livro

e **nenhum** item está presente nos três livros, mesmo levando-se em conta que todos são livros de inglês Britânico geral voltados para aprendizes de nível intermediário. O autor afirma que:

Estas descobertas confirmam a suspeita de que os escritores de material didático, de forma geral, começam com um tópico ou tema e então intuem o que eles consideram ser um vocabulário relevante ou útil. Este processo de seleção nos parece ser não-científico e baseado no julgamento pessoal e na intuição dos escritores.

Defendendo uma abordagem mais criteriosa para a seleção de vocabulário em materiais didáticos para ensino de línguas o autor afirma:

Um número de critérios têm sido propostos através dos anos visando oferecer alguma forma de seleção criteriosa de vocabulário para o desenvolvimento de *syllabus*. Entretanto, para determinar a utilidade geral de itens lexicais, frequência e alcance têm sido considerados como os critérios mais adequados.

## **CAPÍTULO III**

### **METODOLOGIA**

#### **3.1 Compilação do corpus**

Discorrendo sobre as características desejáveis de uma lista acadêmica e do corpus ou corpora nos quais ela é baseada, Coxhead (2000) recomenda que o corpus contenha textos cujo tamanho e formato reflitam de maneira precisa o gênero de textos ao qual ele visa representar. A inclusão de uma variedade de textos permite maior variação de vocabulário e a inclusão de textos escritos por vários autores ajuda a neutralizar a distorção que poderia ser causada por idiossincrasias do estilo de um autor, aumentando o número de itens lexicais do corpus.

Um registro como o de textos acadêmicos abrange uma série de sub-registros, os de cada disciplina. Uma lista acadêmica deveria conter uma seleção equilibrada de palavras que ocorram em várias especialidades. Um corpus compilado para o estudo do vocabulário acadêmico deveria ser extenso o suficiente para permitir um número razoável de palavras acadêmicas.

Para compilar uma lista de palavras acadêmicas, é necessário que se forme um corpus significativo. Acreditamos que dissertações e teses de universidades públicas são uma amostra adequada do vocabulário utilizado em textos acadêmicos no Brasil. Além do fato de que podem ser encontradas em formato eletrônico (pdf, doc), o que facilita enormemente o trabalho de alimentar os textos no software de análise, esses textos foram corrigidos pelos orientadores e devem representar bem o vocabulário usado em contexto acadêmico. Como a LPA-BR é voltada para professores que trabalham com alunos estrangeiros que pretendem atuar no ambiente acadêmico brasileiro, acreditamos que teses e dissertações são a melhor amostra real do vocabulário que será exigido desses estudantes tanto na leitura como na produção de textos.

Optamos, por isso, por compor o corpus com esse tipo de textos, utilizando a classificação das áreas de conhecimento definida pelo MEC: Ciências Agrárias, Ciências Biológicas, Ciências Exatas e da Terra, Ciências Humanas, Ciências Sociais Aplicadas, Ciências da Saúde, Engenharias e Linguística, Letras e Arte. O corpus é composto de textos de dissertações e teses de três universidades: UNICAMP, USP e UNESP. Tentamos manter um equilíbrio entre os trabalhos das áreas de conhecimento, mas isso nem sempre foi possível, já que algumas áreas oferecem um número muito maior de trabalhos disponíveis. Tendo em conta essa limitação, foi feito o possível para obter uma amostragem equilibrada. A tabela 4 apresenta a composição do corpus.

Tabela 4 - Composição do corpus

<b>CIÊNCIAS AGRÁRIAS</b>	<b>Nº. de trabalhos</b>	<b>Ocorrências</b>
Agronomia	4	112.391
Veterinária	6	144.330
Zootecnia	4	123.043
<b>Total</b>	<b>14</b>	<b>379.764</b>
<b>CIÊNCIAS BIOLÓGICAS</b>	<b>Nº. de trabalhos</b>	<b>Ocorrências</b>
Biologia	7	146.214
Farmacologia	3	55.805
Fisiologia	2	42.571
Oceanografia biológica	3	147.559
Tecnologia de alimentos	3	103.993
Zoologia	3	50.006
<b>Total</b>	<b>21</b>	<b>579.448</b>
<b>CIÊNCIAS EXATAS E DA TERRA</b>	<b>Nº. de trabalhos</b>	<b>Ocorrências</b>
Física	4	130.904
Geociências	5	294.513
Matemática, estatística e computação científica	5	178.875
<b>Total</b>	<b>14</b>	<b>604.292</b>

<b>CIÊNCIAS HUMANAS</b>	<b>Nº. de trabalhos</b>	<b>Ocorrências</b>
Educação	6	296.055
Filosofia	4	301.611
Geografia	4	232.878
História	4	378.131
Psicologia	4	201.644
Sociologia	4	326.314
<b>Total</b>	<b>26</b>	<b>1.736.633</b>
<b>CIÊNCIAS SOCIAIS APLICADAS</b>	<b>Nº. de trabalhos</b>	<b>Ocorrências</b>
Administração	4	211.697
Comunicação, Jornalismo e Publicidade	2	46.784
Direito	5	267.220
Serviço social	4	143.688
<b>Total</b>	<b>15</b>	<b>457.692</b>
<b>CIÊNCIAS DA SAÚDE</b>	<b>Nº. de trabalhos</b>	<b>Ocorrências</b>
Educação física	5	175.116
Enfermagem	2	86.844
Medicina	6	205.900
Odontologia	5	78.410
Nutrição	3	82.389
Psiquiatria	2	77.938
<b>Total</b>	<b>23</b>	<b>706.597</b>
<b>ENGENHARIAS</b>	<b>Nº. de trabalhos</b>	<b>Ocorrências</b>
Engenharia agrícola	4	120.264
Engenharia civil	2	70.565
Engenharia elétrica e computação	2	89.936
Engenharia mecânica	3	427.527
Engenharia química	3	78.504
<b>Total</b>	<b>14</b>	<b>786.796</b>
<b>LINGÜÍSTICA, LETRAS E ARTES</b>	<b>Nº. de trabalhos</b>	<b>Ocorrências</b>
Artes	7	295.173
Letras	4	192.596
Lingüística	4	203.981
Lingüística aplicada	6	231.688
<b>Total</b>	<b>20</b>	<b>923.438</b>
<b>Total geral</b>	<b>147</b>	<b>6.565.948</b>

A compilação do corpus envolveu a coleta de cada texto em seu formato eletrônico, a conversão para formato que pudesse ser alimentado no programa que faria a análise de frequência e concordância (formato ".txt"), a remoção das bibliografias de todos os textos para evitar que palavras que aparecem com muita frequência nessas bibliografias, como nomes de autores, títulos de periódicos, expressões latinas, etc, viessem a contaminar o corpus, distorcendo o resultado e a alimentação dos textos no programa.

### **3.2 Elaboração da lista**

O software utilizado para gerar a lista de frequência foi o Concordance 3.2 (2002), que faz vários tipos de análises de textos. Foram utilizados os recursos de listagem de frequência e concordância. Os textos foram alimentados no software agrupados por disciplinas (Engenharia química, Engenharia mecânica, etc) para que se pudesse saber a frequência de ocorrência das palavras em cada disciplina. O software analisa o corpus e fornece uma lista das palavras encontradas, com a frequência de cada uma ou seja, o número de vezes em que cada palavra aparece no corpus, e também pode fornecer a porcentagem de cobertura, ou seja, a porcentagem do texto que corresponde àquela palavra. Por exemplo, em um texto de 1.000 palavras uma palavra que apareça 50 vezes tem uma cobertura de 5%. A concordância lista as palavras dentro do contexto em que aparecem. Isso nos permite avaliar se uma palavra é utilizada com sentidos diferentes. Um exemplo que surgiu durante a análise do corpus foi a palavra "ti", que aparecia no corpus 106 vezes. A análise da concordância mostrou que muitos usos se referiam à sigla TI (Tecnologia da Informação), outros se referiam ao pronome "ti", e outros ainda ao símbolo químico do Titânio. Esse exemplo mostra bem como é importante uma análise qualitativa dos dados gerados pelo software para saber se uma

palavra deve ou não ser incluída em uma determinada lista de palavras. Neste caso, nenhum dos significados individuais alcança a frequência mínima para ser incluída na LPA-BR. Sem a análise a palavra seria incluída, já que aparece 106 vezes.

As palavras foram selecionadas para a LPA-BR segundo os seguintes critérios:

**Ocorrência especializada:** as palavras não poderiam estar entre as mais frequentes de uso geral, pois isso a descaracterizaria como uma lista acadêmica, composta de palavras que ocorrem com muita frequência em textos acadêmicos de diversas áreas mas pouco em textos gerais. Para isso, utilizamos a listagem de corpora do Banco de Português, coletado e mantido na Pontifícia Universidade Católica de São Paulo. Quando utilizado, o corpus possuía a seguinte composição:

Tabela 5 - Composição do corpus

<b>Registro</b>	<b>Ocorrências</b>
Imprensa	137471276
Fala*	3178882
Ficção literária	1488195
Escrita Acadêmica	502438
Documentos escritos de negócios	220252
<b>Total</b>	<b>142861043</b>

\*Conversas, entrevistas, aulas, reuniões e conversas telefônicas.

Fonte: Sardinha (2000)

Determinamos que as palavras da LPA-BR não poderiam estar entre as primeiras 3.000 palavras dessa lista. O número de 3.000 palavras foi escolhido com base no fato de que a GSL, com suas 2.000 famílias, contém pouco mais de 3.000 palavras ou ocorrências. Como essa lista não é

lematizada<sup>19</sup> era necessário tomar uma decisão sobre o ponto de corte, consideramos que ao escolher o número de 3.000 palavras estaríamos próximos do número correspondente a um vocabulário geral básico.

Para ser incluída na lista uma palavra deveria atender aos seguintes critérios de alcance e frequência:

**Alcance:** Uma palavra deveria ocorrer em pelo menos 4 áreas para ser incluída na LPA-BR.

**Frequência:** Uma palavra deveria ocorrer pelo menos 90 vezes no corpus para ser incluída.

Esse limiar entre 80 e 100 ocorrências foi estabelecido por Coxhead (2000). Entretanto, tendo em vista o número de diferentes áreas e cursos que compõem o corpus e a exigência de um mínimo de ocorrência em 4 áreas, julgamos mais adequado o limiar de 90 ocorrências. Foram eliminados do corpus siglas (ANATEL, BNH, UV), abreviações (prof., ex.), expressões coloquiais (tá, né), nomes próprios (Evans, Cortez). Muitas vezes só a análise da concordância permite saber se uma palavra está sendo usada como nome próprio. No caso de “machado”, por exemplo, a concordância mostrou se tratar do sobrenome “Machado”.

O próximo passo foi o de comparar a lista gerada a partir do nosso corpus com as primeiras 3.000 palavras do corpus da PUC para eliminar plurais e palavras derivadas de palavras de alta frequência que deveriam estar em uma lista geral (ex.: desenvolvimento, desenvolver faz parte da lista geral; conhecimento, conhecer faz parte da lista). Algumas palavras, embora sendo derivadas de palavras que fazem parte da lista de 3.000 palavras, foram incluídas porque, a nosso ver, têm um significado relativamente distante das palavras base, alguns exemplos são: instituição (a palavra que consta da lista geral é “instituir”), capitalismo e capitalista (a palavra presente na lista geral é capital, acredito que seja freqüente no sentido de cidade capital).

---

<sup>19</sup> Agrupada em famílias. Para maiores informações sobre o agrupamento em famílias ver 2.6.

Após a eliminação das palavras derivadas de palavras freqüentes, o próximo passo foi selecionar as palavras que aparecessem em várias áreas do corpus. Estabelecemos que apenas palavras que aparecessem em pelo menos 4 áreas seriam incluídas. Por exemplo, a expressão “in vitro” era mostrada como aparecendo em 4 áreas; era de se esperar que essa expressão fosse usada nas áreas Ciências Biológicas, Ciências da Saúde e Ciências Agrárias, mas achamos estranho que pudesse aparecer em qualquer das outras áreas. A análise mostrou que a expressão aparecia uma vez em um trabalho da área de Ciências Humanas ocorrendo, portanto, em 4 áreas e sendo incluída na lista.

Optamos por agrupar as palavras em famílias formadas apenas pelas formas singular, plural, masculino e feminino; primeiramente devido ao fato de não termos encontrado uma análise como a feita sobre os agrupamentos em famílias na língua inglesa por Bauer e Nation (1993) e por outro lado por acreditarmos que seria útil para o professor que venha a se utilizar da LPA-BR saber, por exemplo, que a única conjugação da palavra “caracterizar” a aparecer no corpus é “caracteriza”. Com essa informação, o professor pode decidir-se a dar mais ênfase nas flexões verbais que mais ocorrem nos textos acadêmicos. Com o aumento da proficiência do aprendiz, outras flexões virão a fazer parte do vocabulário do aprendiz, aumentando o número dos membros conhecidos da família, como afirma Nation (2001) ao dizer que “o conhecimento de uma palavra envolve o conhecimento dos membros da família dessa palavra e o número de membros conhecidos aumenta de acordo com o desenvolvimento da proficiência”. Dessa forma, quando se encontra a palavra “citado”, por exemplo, deve-se ter em mente que provavelmente as formas “citada”, “citados” e “citadas” também são encontradas no corpus.

A idéia original era a de informar a classe gramatical com que as palavras apareciam no corpus; por exemplo seria informado que “curva” aparece no corpus como substantivo e não como verbo. Entretanto, essa classificação não informa que “curva” aparece como substantivo com o sentido de curva ascendente ou descendente em gráficos, e não como “curva de estrada”, por

exemplo. Nesse sentido a palavra “curva” pode ser usada para se referir a uma reta ascendente em um gráfico, como na expressão “inflação em curva ascendente”.

Uma palavra pode ter diversos significados dentro da mesma classe gramatical, tendo vários significados diferentes como adjetivo, como no caso de “determinado” que aparece com o significado de “resoluto” e também como “definido”, entre outros; ou no caso dos substantivos, “escala” que aparece com o sentido de “série de espaços separados por traços ou pontos, usada para medir distâncias ou quantidades” e também como “lugar de parada de qualquer meio de transporte de viajantes”. Levando esse fato em conta tomei a decisão de apresentar, nos casos de polissemia, todos os significados encontrados no corpus para cada palavra.

Para isso, foi necessária a análise minuciosa da concordância do corpus. Todas as ocorrências das palavras tiveram de ser analisadas para que se pudesse incluir realmente todos os significados presentes no corpus. No caso de uma palavra como “aspecto”, que ocorre 1322 vezes, fica óbvio o tempo necessário para que se verifique o significado de todas as ocorrências. Muitas vezes determinado significado ocorre em número muito pequeno de vezes, como é o caso de “aspecto” com o significado de “aparência”, que ocorre apenas 13 vezes entre 1322 ocorrências. No caso em que um significado ocorra em quantidade muito pequena, menos do que 5% do total das ocorrências de determinada palavra, acrescentamos o sinal (-) após aquele significado.

Os significados foram retirados dos dicionários Michaelis Moderno Dicionário da Língua Portuguesa, formato eletrônico, versão 1.0 e Dicionário Aurélio Eletrônico versão 3.0. As versões eletrônicas foram escolhidas em razão da maior facilidade e rapidez na busca de palavras. O uso de dois dicionários diferentes veio a ser muito útil, já que várias vezes o significado encontrado no corpus não era apresentado em um dos dicionários. Como exemplo podemos citar a palavra “desafiar”. O dicionário Michaelis apresenta os seguintes significados para essa palavra:

- vtd 1 Chamar a desafio; provocar: Desafiar a morte.
- vtd 2 Propor duelo ou combate a: O touro parecia desafiar o seu algoz.
- vtd 3 Incitar, instigar, provocar: Injustiças que desafiam a indignação. Desafiavam-no a descer da cruz. Desafiava a piedade de todos com a paciente resignação na desgraça.
- vtd 4 Afrontar, arrostar: O justo desafia o infortúnio. Desafiava a morte com o seu possante avião.
- vtd 5 Desinquietar, tentar.
- vpr 6 Entrar em desafio: "Cristo para nos dar o exemplo se desafiou com o demônio" (Pe Antônio Vieira).

Porém, o significado mais freqüente no corpus é o de: “Aquilo que representa um desafio, que é difícil, ou perigoso (ex.: Escalar tal montanha é um desafio), só encontrado no Aurélio. O oposto ocorre com a palavra “objetivo”, que aparece no Aurélio com os seguintes significados:

Adj.

1. Relativo ao objeto.
2. Prático, positivo.
3. Filos. Diz-se do que é válido para todos, e não apenas para um indivíduo.
4. Diz-se de fenômeno natural que se determina conforme os critérios científicos vigentes.
- S. m.
5. Mat. No método iterativo, o valor final para o qual convergem progressivamente os resultados das sucessivas interações.
6. Objeto.

O significado ocorrente no corpus só é encontrado no Michaelis: “Meta ou alvo que se quer atingir”.

A decisão de apresentar os significados ocorrentes de cada palavra parece acertada quando se nota que em muitos casos os significados mais freqüentes, aqueles que aparecem em primeiro lugar nos dicionários, não são os que ocorrem no corpus. Como exemplos, podemos citar: “contribuir”, que aparece apenas com o sentido de “cooperar” e não como “concorrer para uma despesa comum” ou “pagar impostos ou taxas”; “envolver”, que não ocorre com os significados de “enrolar(-se), embrulhar(-se), meter(-se) dentro de invólucro” ou “servir de invólucro a” mas sim com o significado de “abranger”.

“Esfera” não ocorre como “superfície esférica” ou “qualquer corpo redondo; bola, globo” mas como “campo, setor, ou ramo dentro do qual se exerce uma atividade; meio”.

O segundo significado mais encontrado da palavra “padrão”, quando esta é utilizada como “desvio padrão” ou seja “a medida mais comum da dispersão estatística” não foi encontrado em nenhum dos dois dicionários mas apenas no dicionário de português da Wikipédia (<http://www.wikipedia.org/>).

Apresentamos a seguir excertos de trabalhos acadêmicos de diferentes áreas, com as palavras presentes na LPA-BR destacadas, para dar uma idéia da cobertura da lista de textos acadêmicos. Com exceção do texto de Linguística Aplicada, escolhido por pertencer à mesma área deste trabalho, os textos foram escolhidos aleatoriamente. Todos os trechos foram escolhidos ao acaso.

Na elaboração da tabela de coberturas da LPA-BR (tabela 6) palavras repetidas foram contadas apenas uma vez, já nos excertos foram contadas as palavras repetidas. Isso aumentou a porcentagem de cobertura e serve para mostrar como o desconhecimento de poucas palavras pode dificultar o entendimento de um texto, devido ao número de vezes em que uma palavra pode ocorrer.

Área: Linguística, Letras e Artes  
Cobertura na área: 4,7%  
Cobertura no excerto: 10,24%

GATTOLIN, S. R. B. **O vocabulário na sala de aula de língua estrangeira**: contribuições para a elaboração de uma teoria de ensino-aprendizagem. Tese (Doutorado em Linguística Aplicada). Instituto de Estudos da Linguagem, UNICAMP, Campinas, 2005. p.350.

O maior objetivo do ensino de línguas, atualmente, é o de propiciar condições para que os aprendizes atinjam a competência comunicativa; por isso, existe uma grande preocupação por parte dos teóricos em desenvolver teorias ou ampliar algumas já existentes no sentido de garantir a aquisição de L2/LE. No entanto, segundo Raptis (1997), todas continuam ignorando o desenvolvimento do vocabulário através da leitura.

Nas décadas de 70 e 80, acreditava-se que, para ler em LE, bastava ao aprendiz ser também um bom leitor em sua língua materna (LM). Assim, ele transferiria as estratégias empregadas na leitura em LM para a leitura em LE.

No entanto, percebendo que nem todos os leitores eram bem sucedidos em suas tarefas, pesquisadores como Hosenfeld (1977) investigaram, por meio de métodos retrospectivos de coleta de dados, as estratégias empregadas pelos bons leitores para, então, ensiná-las aos alunos deficitários. Tais pesquisas ofereceram valiosas contribuições para os interessados no processo de leitura, mas não minimizaram os problemas dos então considerados “maus” leitores.

Alderson (1984) alertou para o fato de que o problema desses leitores talvez não residisse apenas em sua incapacidade de utilizar as estratégias de leitura, mas também em sua falta de conhecimento lingüístico.

Naquele momento, não foram muitos os pesquisadores que deram atenção a esse autor, pois, paradoxalmente, se acreditava que, assim como na LM, era ela - a leitura - o principal meio para o aluno adquirir e expandir o vocabulário na língua-alvo (KRASHEN, 1989).

Essa hipótese de aprendizagem foi denominada, na literatura, *implícita, incidental* ou *indireta*<sup>6</sup> e recebeu grande apoio na área de ensino-aprendizagem de LE ao longo dos anos 80. Fundamentava-se no Modelo Monitor de Krashen (1981), cujo aspecto central é a Hipótese do Insumo que, em termos de leitura, sugere que a mera exposição do aprendiz a textos significativos e compreensíveis é suficiente para uma boa leitura e uma boa aquisição de vocabulário. Para Krashen (1981), a instrução explícita de estratégias conscientes pode ser supérflua, e até mesmo prejudicial para o desenvolvimento natural do leitor.

Dado o prestígio dessa teoria, foi só em meados da década de 90 que estudiosos se conscientizaram da importância do alerta feito por Alderson (1984). A partir de então, começaram a surgir pesquisas investigando tanto o papel do léxico na compreensão em leitura em LE (LAUFER, 1992; 1997; SCARAMUCCI, 1995) quanto o efeito das atividades de vocabulário sobre a aquisição lexical (PARIBAKHT e WESCHE, 1997; WESCHE e PARIBAKHT, 1999; 2000; LAUFER, 2003).

Entre os estudos citados, destaca-se o de Paribakht e Wesche (1997) pelo impacto causado à comunidade científica. Os resultados dessa investigação opõem-se à noção de que o ensino explícito de vocabulário não contribui para o processo de aprendizagem do mesmo, como postula Krashen (1981). Segundo as autoras, a noção da aprendizagem incidental é válida para a expansão do vocabulário em LM, mas, em se tratando de LE, esse processo se torna muito lento, além de não se poder prever *quais* palavras serão aprendidas, *quando* e *em que grau*. Por esse motivo, as autoras decidiram investigar se uma intervenção instrucional poderia vir a tornar o processo de aquisição lexical mais direto e eficiente.

Essa intervenção envolveu um programa de leitura temática e a participação de um grupo de alunos universitários, aprendizes de inglês como LE.

---

<sup>6</sup> Essas três designações opõem-se, respectivamente, aos termos explícito, intencional e direto, empregados na literatura para se referir ao ensino que privilegia a conscientização ou saliência do componente lingüístico.

Tais alunos foram submetidos a dois tratamentos diferenciados para o ensino de vocabulário através da tarefa de compreensão em leitura, cada um enfocando diferentes palavras alvo.

O primeiro procedimento (*Reading Plus* - RP) incluía a leitura de quatro textos selecionados sobre dois temas distintos, perguntas de compreensão e uma série de exercícios de vocabulário visando enfatizar as palavras alvo dos textos principais. O segundo (*Reading Only* - RO) incluía a leitura de quatro textos selecionados sobre outros dois temas, diferentes daqueles da primeira tarefa, e perguntas de compreensão. Os exercícios de vocabulário, propostos no tratamento RP, foram substituídos pela leitura de um texto suplementar que apresentava aos alunos, novamente, as mesmas palavras alvo dos textos principais com o objetivo de expô-los, mais intensamente, a tais palavras através da leitura.

Área: Ciências Agrárias  
Cobertura na área: 5,2%  
Cobertura no excerto: 11%

DEL FAVA, C. **Índices reprodutivos e características de desempenho em bovinos de corte infectados e não-infectados pelo herpesvírus bovino tipo 1 (HPV-1)**. Tese (Doutorado em Clínica Veterinária). Faculdade de Medicina Veterinária e Zootecnia, USP, São Paulo, 2001. p.127.

### CONCLUSÃO

Os efeitos reprodutivos e produtivos da infecção natural pelo Herpesvírus Bovino tipo 1 foram avaliados em um rebanho bovino de corte, composto por matrizes das raças Gir, Guzerá, Nelore e Caracu, localizado na região norte do Estado de São Paulo, Brasil, incluindo uma estação de monta e de nascimento completos.

- A ocorrência de HVB-1 pelo teste ELISA, foi significativamente maior para os touros - 92,5% (37/40) do que para as matrizes - 54,2% (386/712); A taxa de soroconversão das matrizes ao HVB-1 no período de um ano foi igual a 10,3% (58/561), estatisticamente menor que as taxas de animais não reagentes - 37,1% (208/561) e reagentes - 52,6% (295/561);
- Considerando as raças avaliadas, não se verificou diferença estatisticamente significativa entre as taxas de infecção de matrizes reagentes e não reagentes ao HVB-1;
- Com relação aos três grupos genéticos da raça Nelore, o rebanho Controle apresentou maiores taxas de infecção - 85,2% (52/61), quando comparado com os rebanhos Seleção - 45,1% (55/122) e Tradicional - 56,3% (89/158);

As taxas de infecção das **matrizes reagentes** ao HVB-1 aumentaram com o avançar da idade: de 2 a 3 anos – 23,2% (32/138); de 3 a 4 anos – 45,2% (57/126); de 4 a 5 anos – 54,6% (59/108) e = 5 anos – 70,0% (238/340);

- O índice total de prenhez entre **matrizes reagentes** - 80,3% (310/386) e não **reagentes** - 74,5% (243/326) ao HVB-1 não diferiu **estatisticamente**, assim como as raças, soroconversão, grupos **genéticos** Nelore e faixas etárias, ou seja, a infecção pelo HVB-1 não está afetando a fertilidade do rebanho;
- O índice de prenhez de **matrizes reagentes** ao HVB-1 com **escore corporal** na entrada da estação de monta abaixo de 5 - 65,1% (28/43) diferiu do **escore corporal** acima de 5 - 82,2% (282/343), ou seja, o **escore** acima de 5 está favorecendo a fertilidade dos animais;
- O índice de prenhez de **matrizes reagentes** ao HVB-1 com **escore corporal** na saída da estação de monta abaixo de 6 - 36,4% (8/22) diferiu do **escore corporal** acima de 6 - 82,9% (301/363), ou seja, o **escore** acima de 6 na saída da estação de monta está favorecendo a fertilidade;
- O índice de parição de **matrizes reagentes** ao HVB-1 - 97,7% (300/307) foi maior que o de **matrizes** paridas não **reagentes** - 93,8% (225/240), bem como não foi **observado** efeito do HVB-1 na parição das **matrizes reagentes** e não **reagentes** segundo a raça, soroconversão, grupo **genético** Nelore e faixa etária, ou seja, o HVB-1 não está interferindo na parição;

- O coeficiente de natimortalidade de matrizes reagentes ao HVB-1 - 1,3% (4/300) não diferiu do encontrado para matrizes não reagentes ao HVB-1 - 2,2% (5/225), mesmo sendo considerados a raça, soroconversão, grupo genético e faixa etária, ou seja, o HVB-1 não está interferindo na natimortalidade.

Área: Ciências Humanas  
Cobertura na área: 4,9%  
Cobertura no excerto: 16,4%

SOUZA, L. L. d. **Modelos organizadores na resolução de conflitos morais**: Um estudo com adolescentes autores de infração. Dissertação (Mestrado em Psicologia). Faculdade de Ciências e Letras, UNESP, Assis, 2003. p.214.

### 2.1. A questão da moralidade na psicologia.

De acordo com Pérez-Delgado (1991), tem sido grande o interesse pela moralidade humana na comunidade científica a partir da década de 70, o que é expresso pela grande produção (teses, artigos, livros e outros) sobre esse tema.

Um quadro do estudo da moralidade humana na área da psicologia, no Brasil, pode ser visualizado pelo levantamento bibliográfico que realizamos<sup>1</sup> através dos verbetes "Psicologia" e "Desenvolvimento Moral".

Encontramos na última década (1990-2000) um grande número de produções sob a forma de livros, teses e dissertações acadêmicas, artigos, resumos de pesquisas, relacionadas a esses verbetes. Tais trabalhos concentram-se em estudos sobre os níveis de desenvolvimento moral de escolares, sobre a avaliação de ambientes educativos, notadamente a escola e a sala de aula, sobre o desenvolvimento moral de crianças e adolescentes; estudos teóricos sobre a moralidade humana (que discutem fundamentos e metodologias diferenciadas na investigação da moralidade); estudos comparativos e interculturais sobre os níveis e desenvolvimento moral procurando validar teorias; pesquisas que investigam práticas e intervenções educativas na promoção do desenvolvimento moral; estudos que discutem os princípios psicológicos da educação moral.

---

<sup>1</sup> Consultamos as bases de dados eletrônicas via Internet e CD-ROM Dedalus (USP), Athena (Unesp), Acervus (Unicamp), e da UFRGs-SABI.

Nesses trabalhos observamos que as pesquisas psicológicas da **moral**, notadamente as referenciadas em Piaget e Kohlberg, têm servido de marco interpretativo para **diversos** problemas **educacionais** e sociais da atualidade (ver a esse respeito LA TAILLE, 1998b e 1996; MACEDO, 1996; PUIG, 1998; ARAÚJO, 1996).

Tais **referenciais** abordam a **moralidade** através da **perspectiva** cognitiva, descrevendo os processos **conscientes** **envolvidos** no desenvolvimento do **raciocínio moral**. Contudo, existem outras vertentes **teóricas** no campo psicológico que estudam a **moralidade** por outras **perspectivas**. Consideramos **relevante** comentar alguns **referenciais teóricos**, além da via cognitiva, que também são discutidos e utilizados na **produção** das pesquisas **referentes** à **psicologia da moralidade humana**.

Num **artigo** que descreve sumariamente as principais **abordagens teóricas** sobre o estudo psicológico da **moralidade**, Biaggio (1972) também as avalia criticamente na intenção de ampliar o **debate** sobre as pesquisas **decorrentes** desses **enfoques**, **buscando** um diálogo entre as suas principais **contribuições**.

Em sua **análise**, ela primeiramente deixa claro que a **definição** de **moralidade** adotada pelos investigadores "modernos" da **moralidade** (cita Freud e Durkheim como os primeiros) se refere a um "conjunto de regras culturais de ação social que foram internalizadas pelo **indivíduo**" (BIAGGIO, 1972, p. 8).

Área: Engenharias

Cobertura na área: 4,6%

Cobertura no excerto: 15,41%

GEMMA, S. F. B. **Aspectos do trabalho no cultivo orgânico de frutas:** Uma abordagem ergonômica. Dissertação (Mestrado em Engenharia Agrícola). Faculdade de Engenharia Agrícola, UNICAMP, Campinas, 2004. p.180.

Depois de atravessar períodos de avanço e estagnação, na década de 80 a agricultura orgânica começa a se consolidar no campo conceitual e operacional e em 1984 o Departamento de Agricultura dos E.U.A (USDA) reconheceu sua importância formulando a seguinte definição:

"A agricultura orgânica é um sistema de produção que evita ou exclui amplamente o uso de fertilizantes, pesticidas, reguladores de crescimento e aditivos para a alimentação animal compostos sinteticamente. Tanto quanto possível, os sistemas de agricultura orgânica baseiam-se na rotação de culturas, esterco animal, leguminosas, adubação verde, lixo orgânico vindo de fora da fazenda, cultivo mecânico, minerais naturais e aspectos de controle biológico de pragas para a manter a estrutura e produtividade do solo, fornecer nutrientes para as plantas e controlar insetos, ervas daninhas e outras pragas" (citada por EHLERS, 1999, p.55-56).

Vale destacar que a agricultura moderna ou atual, também conhecida como convencional, é caracterizada pela monocultura (cultivo de uma única espécie em grande escala), utilização de insumos químicos, mecanização intensiva, adubação química solúvel; diferente da agricultura orgânica que preconiza a rotação de culturas, a biodiversidade, a não utilização de insumos que tenham como base recursos minerais não renováveis ou compostos sintéticos e no lugar destes a aplicação de adubos verdes, esterco, restos de culturas, palhas e o controle natural de pragas e doenças.

No tocante à agricultura orgânica ORMOND et al. (2000, p.4) relatam que existe uma falta de informações sistematizadas e que ela apresenta-se "como uma retomada do uso de antigas práticas agrícolas, porém adaptando-se às mais modernas tecnologias de produção agropecuária com o

objetivo de aumentar a produtividade e causar o mínimo de interferência nos ecossistemas, além de ser uma das alternativas para viabilizar a pequena propriedade".

A agricultura orgânica vem sendo apontada como uma alternativa para obtenção de alimentos mais saudáveis (sem produtos tóxicos à saúde) através de um tipo de cultivo que por sua vez, não agrida ao meio ambiente, "seja ecologicamente sustentável, economicamente viável e socialmente justo" (IFOAM - *International Federation of Organic Agriculture Movements*).

O apetite por alimentos sem praguicidas e sem aditivos químicos, cultivados em um sistema de produção agrícola, que busca manejar de forma equilibrada o solo e demais recursos naturais (água, plantas, animais, insetos, entre outros), tem crescido muito no Brasil, cerca de 50% ao ano, e no exterior de 20 a 30% (CHAIM, 2002).

Para assegurar ao consumidor que determinados produtos são "orgânicos", ou seja, que estes são cultivados segundo determinado padrão, foram criadas normas de produção e certificadoras que auditam os produtores e lhes conferem um "selo de certificação" 18. Este selo garante aos produtores facilidades na comercialização de seus produtos.

Cerca de 80% dos produtores mantêm estruturas familiares e obedecem a regras rigorosas para obter a certificação de qualidade (CHAIM, 2002).

Existem atualmente no Brasil 19 certificadoras das quais 12 são de origem nacional e 7 estrangeiras (ORMOND et al., 2000). Geralmente, cada certificadora possui um manual que contém as normas de produção e os regulamentos de certificação. O manual da AAO (Associação de Agricultura Orgânica) consta do Anexo 8.2, para que se possa entender um pouco mais a questão da certificação.

Não se tem conhecimento formal dos impactos destas normas de certificação no cotidiano de trabalho dos agricultores, nem da carga de trabalho que elas implicam, seja no aumento do número de tarefas e ou no grau de complexidade das mesmas.

Área: Ciências da Saúde  
Cobertura na área: 4,6%  
Cobertura no excerto: 13,42%

MORANA, H. C. P. **Identificação do ponto de corte para a escala PCL-R (Psychopathy Checklist Revised) em população forense brasileira:** Caracterização de dois subtipos da personalidade; transtorno global e parcial. Tese (Doutorado em Ciências). Faculdade de Medicina, USP, São Paulo, 2003. p.178.

A literatura científica é farta de artigos que buscam correlacionar comportamento anti-social na idade adulta com sinais de agressividade e/ou hiperatividade na infância (JOHNSON, 1999).

Contudo LOEBER (1995) afirma que são poucos os estudos prospectivos que buscam correlacionar o transtorno anti-social na idade adulta com a psicopatologia de crianças e adolescentes em amostras de referencial clínico.

O transtorno do déficit de Atenção/Hiperatividade tem sido observado em cerca de 3 a 7% das crianças, e em torno de 4% dos adultos. Isoladamente ou associado a outros distúrbios, É um fator altamente predisponente para fracasso nos estudos, dificuldades nos relacionamentos, propenso ao uso de drogas, desempenho profissional insatisfatório, maior morbidade, e outros (BOURBON, 1999).

Loeber (1995) em pesquisa realizada com 177 pré-adolescentes, encontrou que o transtorno do déficit de atenção / hiperatividade prediz o início precoce do transtorno de conduta.

A CID-10 e a DSM-IV fazem a distinção o entre o transtorno hipercinético e o transtorno de conduta na infância. A primeira categoria faz referência a comportamento predominantemente irrequieto e turbulento. À segunda categoria são atribuídos o comportamento violento e o desrespeito ao outro.

Bernstein (1996), num estudo longitudinal com adolescentes, verificou que problemas de conduta na infância emergem como um fator consistente e preditivo do diagnóstico de TEP para os três clusters da personalidade.

Rohde e colaboradores (2000) encontraram uma alta taxa de co-morbidade - em torno de 30 a 50% entre o transtorno de déficit de atenção e os transtornos disruptivos do comportamento (transtorno de conduta e transtorno de desafio e oposição).

Juan (1998), agrupou estudos longitudinais, de adolescentes com transtorno de conduta. Entre os trabalhos agrupados, referiu ser o de maior importância, artigo publicado por ROBINS (ROBINS 1996 apud JUAN, 1998).

Neste trabalho, o autor estudou uma amostra de 500 adolescentes e os seguiu durante 30 anos, encontrando uma continuidade psicopatológica entre transtorno de conduta e transtorno anti-social da personalidade em 37% dos casos. Verificou também que todos os casos evoluíram com grande dificuldade de adaptação social, baixo rendimento profissional e atitude de negligência para com os filhos.

Outro estudo, destacado pelo autor, foi o realizado por OFFORD (OFFORD 1992 apud JUAN, 1998), em que se estudou uma população de crianças normais de 04 até 12 anos de idade e outro grupo de crianças de 08 até 16 anos de idade. Seguiu cada amostra por 8 anos. Os resultados revelaram a persistência do diagnóstico de Transtorno de Conduta em 45% dos casos entre o início e o término das avaliações.

## CAPÍTULO IV

### CONSIDERAÇÕES FINAIS

Neste capítulo responderemos às perguntas de pesquisa definidas no Capítulo 1 e abordaremos as limitações e as contribuições deste estudo.

#### 4.1 Respostas às perguntas de pesquisa

1. Que itens lexicais além das primeiras 3.000 palavras mais freqüentes em português ocorrem freqüentemente em textos acadêmicos?

Respondendo a primeira pergunta de pesquisa, comparamos a lista das palavras que atendiam aos critérios de inclusão (ocorrer mais de 90 vezes e ocorrer em pelo menos 4 áreas distintas) com as primeiras 3.000 palavras mais freqüentes listadas pelo corpus do Banco de Português, da Pontifícia Universidade Católica de São Paulo. Do corpus acadêmico, 512 atenderam aos critérios para inclusão na LPA-BR.

2. Com que freqüência as palavras da LPA-BR ocorrem nas diferentes áreas de conhecimento?

A lista oferece uma cobertura ligeiramente maior nas áreas de Ciências Agrárias e Ciências Biológicas, de 5,2 %, sendo ligeiramente inferior nas outras áreas. Quanto ao alcance, 369 palavras da LPA-BR ocorrem em todas as 8 áreas e 61 ocorrem em 7 áreas. Esse fato mostra que a LPA-BR

se caracteriza como uma lista acadêmica geral, já que a grande maioria de suas palavras ocorrem em todas as áreas.

### 3. Qual a cobertura da LPA-BR sobre textos acadêmicos e não-acadêmicos?

Para a análise de cobertura de textos gerais, compilamos um corpus de livros de ficção com 3.113.271 ocorrências. Ao compilar o corpus, íamos verificando a cobertura da LPA-BR e vimos que o aumento da extensão do corpus não afetava a cobertura, devido a isso decidimos não acrescentar mais textos ao corpus.

Como podemos ver na tabela 6, a LPA-BR cobre, em média, 5% dos textos acadêmicos e somente 1,1% dos textos gerais. Isso significa que em um texto acadêmico de 10.000 ocorrências 520 delas estarão na LPA-BR, já em um texto geral, como aquele encontrado em livros de ficção ou jornais, por exemplo, apenas 110 farão parte da LPA-BR. As palavras que compõem a LPA-BR são cinco vezes mais freqüentes em textos acadêmicos do que em textos gerais.

Tabela 6 - Coberturas da LPA-BR

<b>Área</b>	<b>Cobertura</b>
Ciências agrárias	5,2 %
Ciências biológicas	5,2 %
Ciências exatas e da terra	4,6 %
Ciências humanas	4,9 %
Ciências sociais aplicadas	4,6 %
Ciências da saúde	4,6 %
Engenharias	4,6 %
Linguística, letras e artes	4,7 %
Textos gerais	1,1 %

## **4.2 Limitações**

Deve-se levar em conta o tipo de textos utilizados para a elaboração do corpus que serviu de base para a elaboração da LPA-BR, teses e dissertações de universidades públicas. Embora acreditemos que esses textos sejam uma boa amostragem do vocabulário utilizado em contexto acadêmico, existe a possibilidade de que textos oriundos de periódicos e livros acadêmicos possam levar a um resultado diferente.

Também deve-se ter em mente que o objetivo deste trabalho era o de compilar uma lista de palavras que ocorrem em diversas áreas acadêmicas, e que servisse de ponte entre o vocabulário geral e o vocabulário técnico ou específico de cada área acadêmica. Professores interessados exclusivamente em áreas específicas podem se beneficiar de listas compiladas a partir de corpora elaborados com textos de sua área de interesse.

## **4.3 Contribuições**

Da mesma forma que uma lista de palavras de uso geral é utilizada para a escolha de palavras na elaboração de material didático, textos modificados e testes para a avaliação de conhecimento de vocabulário, a LPA-BR pode servir como fonte de informação sobre o vocabulário que será encontrado pelos aprendizes que pretendam se dedicar à atividade acadêmica no Brasil. Professores ou os próprios aprendizes podem se servir da LPA-BR para se orientarem no ensino ou estudo do vocabulário que será encontrado em textos acadêmicos, independentemente da área a que se dediquem já que, como mostram os trechos de trabalhos acadêmicos onde as palavras da LPA-BR foram destacados, apresentados acima, as palavras da LPA-BR são encontradas em um grande número de diferentes textos acadêmicos.

A seguir apresentamos a LPA-BR.

**LISTA DE PALAVRAS ACADÊMICAS DO PORTUGUÊS DO BRASIL**

<b>Palavra</b>	<b>Alcance</b>	<b>Freq.</b>	<b>Definição / Sinônimo</b>
ABERTA	8	102	Que não está fechado; franco, sincero; iniciada; que admite ou aceita novas idéias.
ABORDAGEM	8	584	Tratamento de um assunto, tema
ACADÊMICO	8	289	Relativo a uma escola superior, faculdade ou universidade; membro de uma dessas instituições
ACAMPAMENTO	4	120	Lugar onde se acampa; lugar de permanência provisória
ACERCA	8	164	A respeito de
ACOLHIMENTO	7	185	Ato de acolher, recepção; atenção, consideração
ACOMPANHAMENTO	8	184	Observação da evolução de
ACOMPANHAR	8	92	Observar a evolução de
ACONTECIMENTO	8	196	Episódio, ocorrência
ADEQUADO	8	310	Apropriado, conveniente, oportuno
ADQUIRIR	5	98	Obter; conseguir; alcançar
ADMINISTRATIVO	8	405	Referente à gestão de negócios públicos ou particulares; governo
ADOLESCENTE	7	165	Que está na adolescência
AGENTE	8	566	Autor, causador, propulso; causa, razão, motivo;
AGIR	8	139	Praticar, obrar, operar, atuar
AGRICULTURA	8	251	Cultivo da terra; lavoura; cultura
AJUSTE	8	102	Ajustamento; liquidação dos débitos ou créditos pendentes
ALCANCE	7	114	Penetração do entendimento, inteligência; limite dentro do qual se consegue tocar ou atingir alguma coisa; importância, valor

<b>Palavra</b>	<b>Alcance</b>	<b>Freq.</b>	<b>Definição / Sinônimo</b>
ALIMENTAR	8	153	Alimentício; dar alimento a (-)
ALIMENTO	8	157	Mantimento, sustento, alimentação
ALTAMENTE	8	121	Muito, extremamente
ALTERNATIVA	8	217	Opção entre duas coisas; escolha
AMOSTRA	8	267	Fragmento ou exemplar representativo de alguma coisa
AMPLO	8	108	Extenso; muito grande; considerável
ANÁLISE	8	194	Exame ou estudo da natureza de uma coisa
ANEXO	8	289	Coisa ligada a outra, considerada como principal; incluso, incluído
ANTIGO	8	96	Que existiu outrora; que existe ou é conhecido desde longo tempo; que não está mais em exercício ou atividade
ANUAL	8	127	Que dura um ano; que ocorre todos os anos
APARENTE	8	101	Visível, evidente; imaginário, suposto
APARENTEMENTE	8	121	Ver APARENTE
APONTA	7	132	Direcionar; mencionar; indicar
APRENDIZ	4	101	Aquele que aprende ofício ou arte
APRENDIZAGEM	8	350	Ação de aprender; aprendizado
APROXIMADAMENTE	8	215	Perto de
APUD	8	287	Citado por
AQUILO	6	284	Aquela coisa ou pessoa
AQUISIÇÃO	6	384	Ato ou efeito de adquirir
ARGUMENTO	7	203	Raciocínio de onde se tira uma conclusão; resumo de uma obra; sumário
ARMAZENAMENTO	8	137	Ação de guardar em depósito

<b>Palavra</b>	<b>Alcance</b>	<b>Freq.</b>	<b>Definição / Sinônimo</b>
ARTIGO	8	146	Escrito de jornal, revista, etc; cada uma das divisões, ordinalmente numeradas, de lei, decreto, código, etc
ASPECTO	8	1322	Ponto de vista; relação; aparência (-)
ASSEGURAR	8	120	Garantir
ASSENTADO	4	102	Diz-se de membro de assentamento; firmado, estabelecido com solidez; cravado, engastado (-)
ASSENTAMENTO	5	137	Ato ou efeito de fixar-se, de estabelecer residência em determinado lugar; ato de fixar, firmar
ASSOCIADO	8	501	Agregado, unido; membro
ASSUMIR	8	131	Chamar para si; encarregar-se de; manifestar
ATINGIR	8	150	Chegar a
ATITUDE	8	166	Modo de proceder ou agir, comportamento, procedimento; posição do corpo; porte, jeito, postura (-)
ATIVO	8	200	A totalidade dos bens de uma empresa ou pessoa; que atua; diligente
ATRIBUÍDO	5	96	Considerado como origem ou causa; imputado
ATRIBUTO	8	134	Aquilo que é próprio ou peculiar de alguém ou de alguma coisa; qualidade
ATUAR	8	96	Exercer atividade
AUTOR	8	1262	Escritor de obra literária, científica ou artística
AUTORIDADE	7	269	Agente ou delegado do poder público; direito ou poder de mandar
AUXILIAR	8	94	Ajudar, prestar auxílio a, socorrer; que auxilia, ajudante (-)
AVALIAÇÃO	7	205	Ato de calcular ou determinar o valor, o preço ou o merecimento de; análise
BACIA	6	104	Depressão à superfície do solo, para o centro da qual correm e convergem as águas; vaso redondo, pouco profundo, de bordas largas, usado para vários fins

<b>Palavra</b>	<b>Alcance</b>	<b>Freq.</b>	<b>Definição / Sinônimo</b>
BEM-ESTAR	8	97	Situação agradável do corpo ou do espírito; conforto, tranquilidade
BIBLIOGRAFIA	8	123	Relação das obras consultadas pelo autor, geralmente no fim do trabalho a que serviram de subsídio
BIOLOGIA	8	104	Ciência que estuda os seres vivos e suas relações
BOVINO	5	157	Relativo ou semelhante ao boi
BREVE	7	93	Conciso, resumido; que dura pouco (-)
BUSCANDO	8	187	Ato de buscar
BUSCAR	8	236	Tratar de descobrir ou encontrar; procurar; ir a algum lugar e trazer de lá (-)
CABE	8	186	Ser da responsabilidade de; vir a propósito
CADEIA	8	302	Série ininterrupta de coisas semelhantes; corrente
CAMADA	7	202	Substância situada sobre ou sob outra ou entre duas outras; classe, categoria
CAMPONÊS	6	265	O que mora ou trabalha no campo
CANA	7	339	Cana-de-açúcar; prisão (-)
CANAL	8	101	Intermediário, mediano; cavidade ou tubo que dá passagem a líquidos ou gases
CAPITALISMO	6	273	Sistema econômico e social baseado na propriedade privada dos meios de produção
CAPITALISTA	6	232	Relativo a capital; pessoa que vive no rendimento de um capital
CHARACTERIZA	8	109	Distingue
CARTOGRAFIA	4	175	Arte ou ciência de compor cartas geográficas
CEREBRAL	5	93	Pertencente ou relativo ao cérebro

<b>Palavra</b>	<b>Alcance</b>	<b>Freq.</b>	<b>Definição / Sinônimo</b>
CICLO	8	290	Intervalo de tempo durante o qual se completa uma seqüência de uma sucessão regularmente recorrente de eventos ou fenômenos
CIDADANIA	6	285	Qualidade de cidadão
CITADO	8	197	Mencionado; o que recebeu citação judicial (-)
COBERTURA	8	163	Abrangência, aquilo que cobre ou serve para cobrir
COLETA	8	294	Reunião, arrecadação
COLETIVO	7	262	Pertencente ou relativo a muitas coisas ou pessoas
COLHEITA	6	102	Ato de colher os produtos agrícolas; s produtos agrícolas colhidos (-)
COMPLETO	8	123	A que não falta nada; inteiro
COMPLEXIDADE	8	167	Qualidade do que é complexo
COMPLEXO	8	102	Que abrange ou encerra muitos elementos ou partes; complicado
COMPONENTE	8	532	Parte constituinte
COMPORTAMENTO	8	721	Maneira de se comportar; procedimento
COMPOSTO	8	418	Formado por dois ou mais elementos; mistura de várias espécies de resíduos de matérias orgânicas empregada para a adubação da terra; produzir, inventar
COMPREENDER	8	355	Entender; abranger
COMPRIMENTO	7	251	Extensão medida de um ponto a outro
COMPROMISSO	8	204	Comprometimento, ajuste, obrigação; contrato
CONCEITO	8	761	Idéia; noção; opinião
CONCEITUAL	8	98	Teórico, conceptual
CONCENTRADO	7	111	Que se reúne em um centro, centralizado; rico na substância dissolvida ou misturada

<b>Palavra</b>	<b>Alcance</b>	<b>Freq.</b>	<b>Definição / Sinônimo</b>
CONCLUIR	8	126	Deduzir, inferir; terminar, acabar (-)
CONCLUSÃO	8	208	Ato de concluir, término; ilação, dedução
CONDUTA	7	152	Procedimento moral; comportamento
CONFLITO	8	270	Luta; oposição; pendência;
CONFRONTO	8	95	Ato de defrontar (-se), fazer face; confrontação
CONQUISTA	7	180	Ato de subjugar, submeter pela força das armas; vencer; adquirir à força de trabalho; alcançar
CONSCIENTE	7	131	Que sabe o que faz; cômscio; (psicol.) capaz de perceber a própria experiência
CONSIGO	7	158	Ato de conseguir; em sua companhia; de si para si
CONSISTE	8	193	Compõe-se, é constituído
CONSTITUEM	8	173	Ver CONSTITUIR
CONSTITUI	8	242	Ver CONSTITUIR
CONSTITUIR	8	93	Ser a base ou a parte essencial de; compor, formar
CONSTRUINDO	7	93	Ver CONSTRUIR
CONSTRUIR	8	238	Dar estrutura a; edificar; arquitetar, organizar
CONSULTA	8	104	Pedido de opinião ou parecer; atendimento que médico, advogado ou técnico dá a clientes (-)
CONSUMIDOR	4	103	Aquele que compra para o gasto próprio
CONTRIBUI	8	113	Ver CONTRIBUIR
CONTRIBUIÇÃO	6	175	Ato de contribuir; subsídio moral, social, literário ou científico para algum fim; imposto, tributo (-)
CONTRIBUIR	8	141	Cooperar

<b>Palavra</b>	<b>Alcance</b>	<b>Freq.</b>	<b>Definição / Sinônimo</b>
CONTROLAR	8	98	Exercer a fiscalização ou o domínio sobre alguém ou alguma coisa
COOPERATIVA	5	392	Associação de consumidores ou de produtores; em que há cooperação
CORPORAL	8	335	Relativo ou pertencente ao corpo
CORPO	8	128	A estrutura física do homem ou do animal; corporação, classe; porção limitada de matéria
CORRENTE	8	250	Corrente elétrica; corrente marítima; fluente; aceito por todos
CORRESPONDE	8	226	Está em correlação a, é próprio de
CORRESPONDENTE	8	133	Correlativo
COTIDIANO	7	104	De todos os dias
CRIADO	8	91	Formado, gerado, fundado, instituído
CRISTAL	7	121	Substância sólida cujas partículas constitutivas (átomos, íons ou moléculas) estão arrumadas regularmente no espaço
CUIDADO	8	199	Desvelo, diligência, atenção; ter cautela (-)
CUIDAR	7	109	Tratar; ocupar-se de
CUJO	8	109	De que, de quem, do qual, da qual
CULPA	6	107	Responsabilidade por um ato ou omissão repreensíveis ou criminosos
CULTIVO	6	205	Ato de fazer que nasça e se desenvolva (planta)
CUMPRIMENTO	8	158	Ato de executar pontualmente; ato de tornar efetivo o que foi determinado ou prescrito; obedecer
CURVA	8	309	Traçado, que pode ser uma linha reta, curva ou quebrada, que representa graficamente um elemento variável, conforme é afetado por uma ou mais condições
DADO	8	282	Certo, determinado; que se deu; suposto que

<b>Palavra</b>	<b>Alcance</b>	<b>Freq.</b>	<b>Definição / Sinônimo</b>
DANO	8	275	Prejuízo; estrago
DAQUELE	8	187	Contração da preposição <b>de</b> com o demonstrativo <b>aquele</b>
DAQUILO	7	113	Contração da preposição <b>de</b> com o demonstrativo <b>aquilo</b>
DEBATE	6	98	Troca de idéias em que se alegam razões pró ou contra, com vistas a uma conclusão
DECORRENTE	8	244	Ver DECORRER
DECORRER	8	128	Passar (o tempo); resultar
DECORRÊNCIA	8	111	Decurso, conseqüência
DECRETO	7	242	Determinação escrita, emanada do chefe do Estado, ou de outra autoridade superior; mandado judicial (-)
DEFINE	8	202	Ver DEFINIR
DEFINIÇÃO	8	465	Ato de demarcar, fixar, estabelecer; explicação precisa; significação
DEFINIDO	8	656	Ver DEFINIR
DEFINIR	8	196	Demarcar, fixar, estabelecer;
DEMANDA	8	168	Exigência;
DEMONSTRA	8	156	Ver DEMONSTRAR
DEMONSTRAR	8	126	Provar com um raciocínio convincente; descrever e explicar de maneira ordenada e pormenorizada, com auxílio de exemplos, espécimes ou experimentos; indicar ou mostrar mediante sinais exteriores
DENOMINADO	8	213	Nomeado; designado
DENSIDADE	8	209	Distribuição de uma quantidade por unidade (densidade demográfica); qualidade ou estado do que é denso ou compacto
DENTRE	8	351	Do meio de; no meio de
DEPARTAMENTO	8	252	Repartição pública; divisão, secção de uma empresa

<b>Palavra</b>	<b>Alcance</b>	<b>Freq.</b>	<b>Definição / Sinônimo</b>
DEPOIMENTO	8	206	O que uma ou mais testemunhas, ou as partes em um processo, afirmam verbalmente em juízo; testemunho
DESAFIO	8	127	Aquilo que é difícil, perigoso; Provocação (-)
DESCREVER	8	107	Contar ,expor minuciosamente; percorrer, traçar (-)
DESCRITO	8	147	Ver DESCREVER
DESEMPREGO	7	245	Falta de emprego
DESLOCAMENTO	8	106	Mudança de um lugar para outro; desarticulação de osso (-)
DESTACA	8	149	Ver DESTACA
DESTACADO	8	97	Que sobressai ou se destaca; relevante, eminente
DESTACAR	8	125	Pôr em destaque; fazer sobressair; salientar
DESTINO	8	95	Aquilo que acontecerá a alguém ou algo; lugar a que se dirige ou para onde é expedida alguma pessoa ou coisa; fado, sorte (-)
DESVIO	8	146	Tirar do caminho, rumo ou destino
DETERMINADO	8	370	Específico, certo; definido; estabelecido; resoluto
DETERMINAR	8	186	Indicar com exatidão;
DETRIMENTO	8	91	Dano, prejuízo
DEVER	7	106	Obrigações de fazer ou deixar de fazer alguma coisa, imposta por alguma lei, pela moral, pelos usos e costumes, ou pela própria consciência
DIAGNÓSTICO	7	317	Qualificação dada por um médico a uma enfermidade ou estado fisiológico, com base nos sinais que observa; descrição minuciosa
DIETA	5	260	Alimentação seguida por um indivíduo ou por um grupo de indivíduos; predominância de um alimento na nutrição

<b>Palavra</b>	<b>Alcance</b>	<b>Freq.</b>	<b>Definição / Sinônimo</b>
DIRETAMENTE	8	334	Sem intermediário; sem se desviar
DIRETRIZ	7	307	Conjunto de instruções ou indicações para se tratar e levar a termo um plano, uma ação, um negócio, etc
DISCIPLINA	7	479	Conjunto de conhecimentos em cada cadeira dum estabelecimento de ensino, matéria de ensino; observância de preceitos ou normas
DISPONIBILIDADE	8	137	Qualidade ou estado do que é disponível
DISPOSITIVO	8	219	Regra, preceito; artigo de lei; mecanismo disposto para se obter certo fim
DISSERTAÇÃO	8	466	Trabalho escrito, apresentado a instituição de ensino superior, e defendido, publicamente, por candidato ao grau de mestre
DISTINÇÃO	8	149	Percepção da diferença entre pessoas ou coisas
DISTINTO	8	268	Diverso, diferente
DISTRIBUIÇÃO	6	207	Disposição, ordenamento; classificação; repartição
DISTRITO	7	99	Território que é a sede do governo em uma república federativa (distrito federal); cada uma das partes em que se divide o território do município
DIVERSOS	8	448	Vários; diversos; distintos
DIVERSIDADE	8	189	Variedade
DOCENTE	7	111	Relativo a professores; professor
DOMINANTE	8	129	Que domina, governa, prevalece; predominante
DOUTORADO	8	144	Curso de pós-graduação, mais avançado que o mestrado; o grau de doutor
EDUCACIONAL	7	425	Pertinente a educação
EFETIVAMENTE	8	166	Ver EFETIVO

<b>Palavra</b>	<b>Alcance</b>	<b>Freq.</b>	<b>Definição / Sinônimo</b>
EFETIVO	8	306	Real, verdadeiro; que tem efeito, eficaz
EFICIENTE	8	113	Que produz o efeito desejado
EIXO	8	158	Órgão central de quaisquer movimentos reais ou figurados
ELEMENTO	8	1089	Cada parte de um todo
ELEVADA	8	124	Grande; alçada (-)
EMPREGADO	8	254	Usado, aplicado; aquele que exerce emprego
ENFERMAGEM	6	107	Tratamento dos enfermos; a profissão de enfermeiro
ENFOQUE	7	143	Modo de considerar ou de entender um assunto ou uma questão; ponto de vista
ENUNCIADO	6	227	Exposição sumária de uma verdade expressa, sem desenvolvimento nem explicação, com a finalidade de demonstrá-la
ENVOLVEM	8	100	Abrangem
ENVOLVIDO	8	91	Implicado
ENVOLVIMENTO	8	141	Ação ou efeito de envolver ou envolver-se
ESCALA	4	93	Categoria, graduação; série de espaços separados por traços ou pontos, usada para medir distâncias ou quantidades; lugar de parada de qualquer meio de transporte de viajantes
ESCORE	4	95	Resultado obtido num teste; nota, grau
ESCRITO	7	264	Que se escreveu; composição escrita
ESFERA	7	150	Campo, setor, ou ramo dentro do qual se exerce uma atividade; meio
ESPACIAL	8	175	Concernente ao espaço ou área
ESPÉCIE	7	261	Casta, natureza, qualidade, tipo; conjunto de seres que têm a mesma essência, abstraindo de suas diferenças individuais
ESPECIFICAMENTE	8	192	Exclusivamente, especialmente

<b>Palavra</b>	<b>Alcance</b>	<b>Freq.</b>	<b>Definição / Sinônimo</b>
ESPECIFICIDADE	8	106	Qualidade ou estado de ser específico
ESTABELECE	8	193	Determina; cria; instala
ESTATISTICAMENTE	6	94	De maneira ou de acordo com a estatística
ESTATUTO	6	94	Lei orgânica ou regulamento de um Estado, associação, ou de qualquer corpo coletivo em geral
ESTRATÉGIA	7	407	Arte de aplicar os meios disponíveis com vista à consecução de objetivos específicos
ESTRUTURA	8	342	Organização das partes ou dos elementos que formam um todo; o que é mais fundamental, ou essencial, estável e relevante; esqueleto ou armação de um edifício (-)
ESTUDADO	8	199	Analisado, examinado detidamente; considerado ou preparado cuidadosamente
ESTUPRO	5	162	Coito sem consentimento e efetuado com emprego de força ou ameaça
EVIDENTE	8	142	Que não oferece dúvida; claro, manifesto, patente
EXCESSO	8	115	Aquilo que excede ou ultrapassa o permitido, o legal, o normal; sobra
EXCLUSIVAMENTE	8	121	Unicamente, especialmente, privativamente
EXIGÊNCIA	8	121	Pedido imperioso, repetido
EXISTENTE	8	424	Que existe
EXISTIR	8	168	Ter existência real; ser; haver
EXPANDIR	6	116	Fazer crescer, desenvolver
EXPERIMENTAL	8	362	Fundado na experiência, ou em experiências
EXPERIMENTO	8	286	Ensaio científico para a verificação de relações entre fatos bem definidos
EXPRESSA	8	188	Ver EXPRESSAR; que é enviada rapidamente (-)

<b>Palavra</b>	<b>Alcance</b>	<b>Freq.</b>	<b>Definição / Sinônimo</b>
EXPRESSAR	8	109	Expressar
EXTERNO	8	118	Que é de fora; exterior
EXTREMAMENTE	8	128	Extraordinariamente; imensamente; espantosamente
FACE	8	158	Perante, diante de; em virtude de; superfície; rosto (-)
FATOR	8	451	Qualquer elemento que concorra para um resultado; <i>Mat.</i> cada uma das quantidades que se multiplicam para formar um produto
FERRAMENTA	8	211	Instrumento
FIGURA	8	1782	Imagem, gravura; pessoa ou personalidade representativa; vulto
FILOSOFIA	8	295	Estudo geral sobre a natureza de todas as coisas e suas relações entre si
FINALIDADE	8	247	Fim a que se destina uma coisa; objetivo, alvo; destinação
FINANCIAMENTO	8	139	Ato de prover as despesas de; custear
FIXO	8	189	Imóvel; que não varia;
FLORESTA	6	102	Mata
FLUXO	8	293	Qualquer movimento contínuo ou que se repete no tempo; curso de líquido
FOCO	8	173	Ponto de convergência; centro
FORMADO	8	105	Feito, constituído
FORMAL	8	150	Que se atém às fórmulas estabelecidas; convencional
FORNECER	8	136	Prover do necessário, dar; proporcionar
FRONTEIRA	8	214	Limite ou linha divisória entre dois países, dois Estados, etc;
FRUTO	8	220	Resultado, lucro, proveito; fruta
FUNCIONAL	8	204	Aquilo que é capaz de cumprir com eficiência seus fins utilitários; referente a função

<b>Palavra</b>	<b>Alcance</b>	<b>Freq.</b>	<b>Definição / Sinônimo</b>
FUNCIONAMENTO	8	232	Ato ou efeito de funcionar
FUNDAMENTO	8	220	Motivo, razão; base
GENÉTICO	5	102	Relativo ou pertencente à gênese, à geração; relativo à genética
GERADO	8	96	Criado, produzido, formado
GRAXO	4	577	Gordurento, oleoso
GRIFO	5	122	Sublinhado
HABILIDADE	8	187	Capacidade; aptidão; destreza
HIPÓTESE	8	163	Suposição, conjectura; eventualidade
HUMANO	8	1495	Pertencente ou relativo ao homem
IDEM	5	138	Como antes ou dito antes; da mesma maneira
IDENTIDADE	8	323	O aspecto coletivo de um conjunto de características pelas quais algo é definitivamente reconhecível, ou conhecido
IDEOLOGIA	6	240	Maneira de pensar que caracteriza um indivíduo ou um grupo de pessoas
IMPACTO	8	230	Forte efeito, que impede ou que força mudança; impressão muito forte
IMPLICA	8	198	Traz como consequência
IN-SITU	4	407	No lugar
IN-VITRO	4	133	Em meio artificial
INCONSCIENTE	5	136	<i>Psicol.</i> A parte da nossa vida da qual não temos consciência
INDEPENDENTE	8	331	Que não depende de ninguém ou de nada; autônomo, livre
INDEPENDENTEMENTE	8	101	Ver INDEPENDENTE
INDICA	8	165	Aponta; sugere; demonstra; aconselha
INDICADOR	8	481	Ver INDICA
INDIVIDUAL	8	186	Que é feito ou diz respeito a uma só pessoa; particular; singular

<b>Palavra</b>	<b>Alcance</b>	<b>Freq.</b>	<b>Definição / Sinônimo</b>
INDIVÍDUO	6	914	Pessoa considerada isoladamente em relação a uma coletividade; exemplar de uma espécie qualquer, orgânica ou inorgânica, que constitui uma unidade distinta
INERENTE	8	91	Que por natureza é inseparável de alguém ou alguma coisa
INFERIOR	8	288	Que está abaixo ou por baixo ou mais baixo; que está abaixo de outro em categoria ou dignidade, ou vale menos
INSTITUCIONAL	8	293	Que diz respeito a uma instituição
INSTRUÇÃO	5	131	Ensino, educação intelectual; explicação dada para um determinado fim; <i>Dir.</i> conjunto das formalidades e informações necessárias para elucidar uma causa e pô-la em estado de ser julgada
INSTRUMENTO	8	569	Recurso empregado para se alcançar um objetivo; objeto que serve de agente mecânico na execução de qualquer trabalho; objeto que produz sons musicais (-)
INSUMO	5	113	Elemento que entra no processo de produção ; <i>Fig.</i> tudo aquilo que entra
INTEGRADO	8	103	Juntada, incorporada
INTEGRAL	8	117	Inteiro, total
INTENSIDADE	8	181	Grau, força, gradação
INTERESSANTE	8	219	Que atrai a atenção
INTERNO	8	208	Que está dentro; interior
INTERNACIONAL	8	511	Que diz respeito a diferentes nações
INTERPRETAR	8	98	Explicar ou aclarar o sentido de
INTERVALO	8	246	Espaço entre dois pontos
INTERVENÇÃO	4	93	Ato de intervir; interferência; <i>Med</i> Operação
INTUITO	8	105	Intenção; fim

<b>Palavra</b>	<b>Alcance</b>	<b>Freq.</b>	<b>Definição / Sinônimo</b>
INVESTIMENTO	8	163	Aplicação de dinheiro
JORNADA	8	93	Duração do trabalho diário; viagem, caminho (-)
JUNTAMENTE	8	91	Junto
JUSTAMENTE	8	191	Exatamente
LAZER	8	96	Divertimento, entretenimento
LEITURA	8	774	Ato de ler; aquilo que se lê
LEMBRAR	8	97	Fazer notar; advertir; recordar
LIDA	8	95	Ato de ocupar-se com, tratar de; trabalho; feita a leitura (-)
LIGADO	8	215	Que se ligou, unido; relacionado
LINEAR	8	169	<i>Mat.</i> envolvendo funções ou equações lineares; que diz respeito a linhas; que dá idéia de seguir uma linha reta (-)
LONGO	8	168	Extenso, duradouro
MACHO	6	332	Do sexo masculino
MAPA	8	183	Carta geográfica ou celeste; lista descritiva; catálogo, relação
MARGEM	8	149	Fora de (à margem de); terreno que ladeia um rio ou corrente de água; tolerância
MASSA	8	447	<i>Fis.</i> Quantidade de matéria que forma um corpo; o povo; em grande número; pasta
MATERNA	5	132	Inerente, pertencente ou relativo à mãe
MATRIZ	8	400	Aquilo que é fonte, origem, base; molde para a fundição de qualquer peça; igreja (católica) principal de uma localidade
MECANISMO	8	510	Processo
MEDIANTE	8	228	Através de
MELHORAR	8	160	Tornar melhor ou superior; fazer prosperar

<b>Palavra</b>	<b>Alcance</b>	<b>Freq.</b>	<b>Definição / Sinônimo</b>
MELHORIA	8	156	Transição para melhor; melhoramento
MEMBRANA	6	298	Pele, película
MENTAL	8	135	Relativo à mente; intelectual
MENTE	8	133	Intelecto, pensamento, entendimento; alma, espírito
MESTRADO	8	391	Curso de pós-graduação, mais avançado que a especialização
MESTRE	8	99	Aquele que tem o mestrado; homem superior e de muito saber
METODOLOGIA	8	366	Estudo dos métodos e, especialmente, dos métodos das ciências; conjunto de técnicas e processos utilizados para ultrapassar a subjetividade do autor; teoria dos procedimentos de ensino, geral ou particular para cada disciplina; didática teórica
MISTURA	8	524	Amálgama, mescla; misto
MITO	6	126	Representação de fatos ou personagens reais, exagerada pela imaginação popular, pela tradição; representação de um estado ideal da humanidade; coisa inacreditável, fantasiosa
MODALIDADE	8	113	Forma, aspecto ou característica de uma coisa, ato, pensamento, organização, etc
MODERNIDADE	6	163	Qualidade de moderno; os tempos modernos
MODERNO	8	191	Dos tempos atuais ou mais próximos de nós; recente; atual
MONTAGEM	8	108	Ato ou efeito de montar; conjunto dos preparativos para pôr em cena uma peça teatral
MORAL	6	761	Conjunto de regras de conduta consideradas como válidas, quer de modo absoluto para qualquer tempo ou lugar, quer para grupo ou pessoa determinada
MORALIDADE	5	137	Qualidade do que é moral

<b>Palavra</b>	<b>Alcance</b>	<b>Freq.</b>	<b>Definição / Sinônimo</b>
MORTALIDADE	8	219	Conjunto de mortes ocorridas num espaço de tempo
NAQUELE	8	236	Combinação da preposição em com o pronome demonstrativo aquele
NARRATIVA	6	323	Conto, história
NELE	8	253	Combinação da preposição em com o pronome pessoal ele
NORMA	8	515	Preceito, regra, teor
NUCLEAR	8	123	Diz-se do fenômeno, aparelho, engenho, etc., em que se processam reações de cisão ou fissão nuclear; Que se refere a núcleo
NUTRICIONAL	7	136	Relativo à nutrição ou alimentação
NUTRIENTE	7	131	Que nutre; alimentício
OBESIDADE	4	178	Acumulação excessiva de gordura no corpo
OBJETIVO	8	678	Meta ou alvo que se quer atingir; prático, positivo
OBJETO	8	1034	Assunto, matéria; objetivo, fim, propósito; coisa material
OBSERVA-SE	8	132	Ver OBSERVAR
OBSERVADO	8	764	Ver OBSERVAR
OBSERVAR	8	390	Fazer notar; advertir; estudar, examinar; ver
OBSERVARAM	7	164	Ver OBSERVAR
OBSERVOU-SE	8	101	Ver OBSERVAR
OBSERVOU	8	99	Ver OBSERVAR
OBTENÇÃO	8	205	Ato ou efeito de obter
OBTIDO	8	783	Conseguido, alcançado
OCORREM	8	162	Ver OCORRER
OCORRÊNCIA	7	250	Ato de ocorrer; acontecimento
OCORRER	8	201	Acontecer, dar-se

<b>Palavra</b>	<b>Alcance</b>	<b>Freq.</b>	<b>Definição / Sinônimo</b>
OFERECER	8	117	Apresentar ou propor para que seja aceito
OPORTUNIDADE	8	166	Ocasão favorável; ensejo
ORAL	7	230	Realizado de viva voz; verbal; que diz respeito à boca
ORGÂNICO	7	326	Qualquer composto que encerre carbono em sua molécula; diz-se da lei destinada a desenvolver os princípios exarados numa lei constituinte; diz-se de alimento produzido levando-se em conta os princípios da ecologia, ger. sem o emprego de fertilizantes ou pesticidas sintéticos
ORGANISMO	8	184	Corpo organizado que tem existência autônoma; corporação ou instituição pública; qualquer sistema ou estrutura organizada
ORGANIZAR	8	142	Arranjar, ordenar, preparar
PADRÃO	8	694	Modelo; (desvio padrão) é a medida mais comum da dispersão estatística
PARCIAL	8	112	Que não é total; que se realiza por partes
PARECER	7	115	Opinião fundamentada sobre determinado assunto, emitida por especialista, ter a aparência de
PARTICULAR	8	307	Caso ou circunstância especial; que não é público, que não se destina ao uso público
PARTICULARMENTE	8	179	Ver PARTICULAR
PEDAGOGIA	6	137	Estudo teórico ou prático das questões da educação; conjunto das idéias de um educador prático ou teorista em educação
PEDAGÓGICO	6	600	Relativo ou conforme à pedagogia
PENSAMENTO	8	603	Ato ou faculdade de pensar; o produto intelectual de um determinado indivíduo, grupo, país, ou época; modo de pensar; opinião

<b>Palavra</b>	<b>Alcance</b>	<b>Freq.</b>	<b>Definição / Sinônimo</b>
PERCEBER	8	233	Abranger com a inteligência; compreender, entender; adquirir conhecimento de, por meio dos sentidos
PERCENTUAL	8	189	Relativo a percentagem
PERCURSO	7	107	Trajeto, caminho
PERÍODO	7	169	Qualquer intervalo de tempo, mais ou menos longo, determinado ou indeterminado
PERMANECE	8	118	Continua existindo
PERMANENTE	8	142	Que permanece; contínuo, ininterrupto; constante; definitivo
PERMITEM	8	132	Ver PERMITIR
PERMITINDO	8	142	Ver PERMITIR
PERMITIR	8	141	Tornar possível; dar permissão ou licença para; consentir
PERSONALIDADE	5	242	<i>Psic.</i> Organização constituída por todas as características cognitivas, afetivas, volitivas e físicas de um indivíduo; a maneira habitual de ser de uma pessoa, aquilo que a distingue de outra
PERSPECTIVA	8	812	Modo de ver; ponto de vista; expectativa, esperança, probabilidade
PESQUISADOR	8	257	Que ou aquele que pesquisa
PLANEJAMENTO	8	497	Ato de projetar um trabalho, serviço ou mais complexo empreendimento
PLANTA	8	381	Qualquer vegetal; semear, cultivar
PLATAFORMA	7	142	Superfície plana, horizontal, mais alta que a área adjacente
POBRE	8	273	Desprovido ou mal provido do necessário; pouco fértil; de pouco valor
POEMA	5	125	Obra em verso
POESIA	6	136	Composição poética pouco extensa; pequeno poema; conjunto das obras em verso escritas numa língua

<b>Palavra</b>	<b>Alcance</b>	<b>Freq.</b>	<b>Definição / Sinônimo</b>
POLPA	6	121	Massa, pasta
PONTAL	4	204	Ponta de terra que penetra um pouco no mar ou no rio
PORCENTAGEM	8	127	Parte proporcional calculada sobre uma quantidade de 100 unidades; percentagem
PORTADOR	8	135	Indivíduo que hospeda em seu corpo os organismos específicos de uma doença
POSICIONAMENTO	8	148	Estratégia de marketing que objetiva criar, para o público-alvo, percepção diferenciada de uma marca, empresa, pessoa, produto, serviço, etc., em relação a seus concorrentes; ato ou efeito de posicionar
POSITIVO	8	107	Afirmativo; construtivo; que exprime grandeza superior a zero; real, efetivo, verdadeiro; certo, seguro
POSSIBILITANDO	8	91	Tornando possível
POSSUEM	8	273	Ter; conter; contar com
POSTERIOR	8	136	Que vem ou está depois; ulterior; situado atrás
POSTERIORMENTE	8	236	Ver POSTERIOR
POSTURA	8	174	Ponto de vista; maneira de pensar e agir; atitude; ato de pôr ovos (-)
POTENCIAL	8	342	Com possibilidade ou probabilidade de realização ou aproveitamento; virtual, possível; relativo a potência
PRAZER	7	155	Sensação ou sentimento agradável; satisfação, deleite
PRESSUPOSTO	6	120	Circunstância ou fato considerado como antecedente necessário de outro; suposição, conjetura;
PREVIAMENTE	8	92	Anteriormente, preliminarmente
PRIVADO	7	273	Que não é público, particular; íntimo
PROBABILIDADE	8	217	Motivo ou indício que deixa presumir a verdade ou a possibilidade dum fato; verossimilhança

<b>Palavra</b>	<b>Alcance</b>	<b>Freq.</b>	<b>Definição / Sinônimo</b>
PROCEDIMENTO	8	476	Maneira de agir, de fazer alguma coisa; processo ou ação judicial intentada contra um criminoso
PROCESSAMENTO	8	210	Ato ou efeito de processar
PRODUÇÃO	8	92	Ato ou efeito de produzir, criar, gerar, elaborar, realizar; aquilo que é produzido ou fabricado
PRODUTIVIDADE	8	216	Relação entre a quantidade ou valor produzido e a quantidade ou valor dos insumos aplicados à produção; faculdade de produzir
PRODUTIVO	8	117	Diz-se do trabalho aplicado diretamente no processo de manufaturação; que produz; fértil
PRODUTOR	4	116	Aquele que, com trabalho agropecuário, industrial ou extrativo, cria, aperfeiçoa, industrializa e distribui produtos naturais, alimentícios e transformados; que produz
PRODUZIDO	8	371	Que se produziu
PRODUZIR	8	302	Dar existência; gerar
PROFUNDO	8	119	De grande alcance; sagaz; perspicaz
PROFUNDIDADE	8	135	Ver PROFUNDO
PROMOÇÃO	5	101	Ato ou efeito de PROMOVER; elevação ou acesso a cargo ou categoria superior
PROMOVER	8	207	Dar impulso a; trabalhar a favor de; favorecer o progresso de; fazer avançar; ser a causa de
PROPICIAR	6	91	Proporcionar
PROPOR	7	107	Sugerir; apresentar para consideração, discussão ou solução
PROPOSTO	8	171	Ver PROPOR
PROPRIAMENTE	8	125	No sentido próprio ou direto, não figuradamente; exatamente;

<b>Palavra</b>	<b>Alcance</b>	<b>Freq.</b>	<b>Definição / Sinônimo</b>
PROPRIEDADE	8	713	A coisa possuída, a coisa cuja posse pertence por direito a alguém; qualidade especial; particularidade
PROTAGONISTA	5	142	Pessoa que desempenha ou ocupa o primeiro lugar num acontecimento; principal personagem de uma peça dramática
PROVAVELMENTE	8	159	Que apresenta probabilidades de acontecer
PROVENIENTE	8	102	Que provém; oriundo, procedente
PSICOLOGIA	7	236	A ciência dos fenômenos psíquicos e do comportamento
PUBLICADO	8	93	Impresso para a distribuição ou venda; editado
QUALIDADE	8	110	Atributo; virtude; tipo
RACIOCÍNIO	6	118	Capacidade de raciocinar; juízo, razão; racionalidade; encadeamento, aparentemente lógico, de juízos ou pensamentos
RACIONAL	8	150	Conforme à razão; que raciocina
RACIONALIDADE	5	125	Qualidade de racional
REALIZADO	8	249	Efetado, executado, feito
REAGENTE	5	165	Um dos participantes de uma reação química
RECEPTOR	8	274	Que recebe; recebedor
RECONHECER	8	103	Admitir como certo; aceitar
RECONHECIMENTO	8	248	Ver RECONHECER
RECURSO	8	162	Meio, expediente;
REDOR	8	97	<i>Dir.</i> Ação que compete à pessoa condenada em juízo ou tribunal, para recorrer para outro juízo ou tribunal superior
REDUZIDO	8	194	Que sofreu redução; parco, escasso
REFERE-SE	8	173	Ter relação, dizer respeito; reportar-se, aludir

<b>Palavra</b>	<b>Alcance</b>	<b>Freq.</b>	<b>Definição / Sinônimo</b>
REFERÊNCIA	8	697	(referências bibliográficas) Citação das fontes bibliográficas utilizadas pelo autor; alusão a certa obra ou a um determinado fato;
REFERENCIAL	7	164	Que é utilizado como referência; que constitui ou contém referência
REFERENTE	8	453	Que diz respeito; relativo, concernente, alusivo
REFERIDO	8	118	Exposto por escrito ou oralmente; já mencionado, citado
REFLETIR	8	112	Pensar maduramente; meditar, reflexionar
REGISTRO	8	140	Ato ou efeito de registrar; repartição encarregada de registrar; válvula hidráulica de acionamento manual (-)
RELACIONADO	8	722	Que tem relação com; posto em relação, em rol, em lista
RELACIONAMENTO	7	129	Ato ou efeito de relacionar-se
RELATIVAMENTE	8	149	Ver RELATIVO
RELATIVO	8	199	Referente, concernente; que não é absoluto
RELATO	8	310	Descrição, notícia, informação, relatório
RELEVANTE	8	137	Aquilo que importa; importante
REPRESENTADO	8	198	Ver REPRESENTAR
REPRESENTANTE	8	156	Que se apresenta no lugar de
REPRESENTAR	8	140	Ser a imagem ou a reprodução de; significar, denotar; apresentar-se no lugar de; fazer as vezes de; suprir a falta de
REQUER	8	114	Exige, precisa, necessita;
REQUISITO	8	106	Exigência imprescindível para a consecução de certo fim
RESIDUAL	8	193	Pertencente ou relativo a resíduo
RESOLVER	8	121	Achar a solução de

<b>Palavra</b>	<b>Alcance</b>	<b>Freq.</b>	<b>Definição / Sinônimo</b>
RESPECTIVAMENTE	8	357	Ver RESPECTIVO
RESPECTIVO	8	97	Relativo a cada um em particular ou em separado
RESPONDER	8	120	Dar resposta; responsabilizar-se por
RESPONSABILIDADE	8	346	Qualidade de responder pelos próprios atos ou pelos de outrem
RESSALTA	8	92	Ver RESSALTAR
RESSALTAR	8	137	Fazer sobressair; destacar
RETORNO	8	313	Regresso, volta; lucro advindo de investimento; nas rodovias, desvio próprio para retornar (-)
RUPTURA	8	101	Quebra de relações sociais, ou de compromisso; rompimento; corte; interrupção
RURAL	7	555	Pertencente ou relativo ao campo
SELVA	4	112	Mata, floresta; lugar onde se luta duramente pela sobrevivência
SEMELHANTE	8	213	Análogo, parecido
SENSIBILIDADE	8	186	Limiar muito baixo para certos estímulos; faculdade de sentir; sentimento
SER	8	160	Todo ente vivo e animado
SERVE	8	91	Ser útil, vantajoso; convir
SIGNIFICADO	8	600	Significação, sentido, acepção
SIGNIFICANTE	8	126	<i>Ling.</i> Complexo sonoro audível que encerra o significado do signo lingüístico; Ver SIGNIFICATIVO
SIGNIFICATIVAMENTE	8	121	De forma significativa
SIGNIFICATIVO	8	171	Que contém revelação interessante; expressivo
SIGNO	6	364	Todo objeto, forma ou fenômeno que representa algo distinto de si mesmo; unidade lingüística que tem significante e significado; signo lingüístico

<b>Palavra</b>	<b>Alcance</b>	<b>Freq.</b>	<b>Definição / Sinônimo</b>
SIMILAR	8	116	Que tem a mesma natureza, a mesma função, o mesmo efeito, ou a mesma aparência
SIMULTANEAMENTE	8	106	Que ocorre ou é feito ao mesmo tempo ou quase ao mesmo tempo que outra coisa; concomitantemente
SINTOMA	8	142	Qualquer fenômeno de caráter subjetivo provocado no organismo por uma doença; sinal, indício
SISTEMA	8	676	Conjunto ou combinação de coisas ou partes de modo a formarem um todo complexo ou unitário
SOLIDARIEDADE	8	134	Laço ou vínculo recíproco de pessoas ou coisas independentes; adesão ou apoio a causa, empresa, princípio, etc., de outrem.
SOLO	8	299	Terreno sobre que se constrói ou se anda; chão, pavimento; terra considerada nas suas qualidades produtivas
STATUS	8	117	Posição do indivíduo no grupo; condição, situação
SUBJETIVIDADE	7	122	Qualidade ou caráter de subjetivo
SUBJETIVO	6	90	Diz-se de uma explicação que se baseia em mera concepção; individual, pessoal; particular
SUJEITO	8	1740	Indivíduo que não se nomeia; exposto, passível; assunto, tema
SUPORTE	8	145	Aquilo em que algo se firma ou assenta; pessoa que dá recomendações técnicas para um usuário, explicando o uso de programas ou equipamentos, ou ainda esclarecendo defeitos
SURGE	8	170	Manifestar-se; aparecer
SURGEM	8	96	Ver SURGE
SURGIMENTO	8	153	Ato ou efeito de surgir
SURGIU	8	94	Ver SURGE

<b>Palavra</b>	<b>Alcance</b>	<b>Freq.</b>	<b>Definição / Sinônimo</b>
SUSTENTABILIDADE	7	289	Qualidade de se manter mais ou menos constante, ou estável, por longo período
TAREFA	8	370	Trabalho a ser executado
TECIDO	8	199	Reunião de células com a mesma estrutura, exercendo determinada função; pano, fazenda (-)
TECNOLOGIA	5	174	Conjunto de conhecimentos, especialmente princípios científicos, que se aplicam a um determinado ramo de atividade
TEMPERATURA	4	612	Grau de calor num corpo ou num lugar
TEMPORAL	8	149	Relativo a tempo; temporário
TEOR	8	191	Proporção, em um todo, de uma substância determinada; texto ou conteúdo de uma escrita (-)
TEORIA	8	1448	Princípios básicos e elementares de uma arte ou ciência; doutrina ou sistema fundado nesses princípios; hipótese, suposição
TEÓRICO	8	582	Relativo a teoria
TERAPIA	5	102	Parte da medicina que estuda e põe em prática os meios adequados para aliviar ou curar os doentes, tratamento; terapêutica
TÉRMINO	4	94	Fim, termo
TERRENO	7	113	Espaço de terra mais ou menos extenso; assunto; (ganhar terreno) conquistar vantagens; avançar; difundir-se
TESE	8	388	Monografia para ser defendida em público, em exames de doutorado; argumento, assunto, tema; (em tese) de acordo com o que se supõe; em princípio; em teoria
TODAVIA	8	218	Contudo, porém; entretanto; ainda assim
TOMADA	8	152	Ato ou efeito de tomar; conquista

<b>Palavra</b>	<b>Alcance</b>	<b>Freq.</b>	<b>Definição / Sinônimo</b>
TRANSFORMAR	8	92	Converter, mudar, transfigurar
TRANSTORNO	5	352	Desarranjo, desordem
TRATA-SE	8	178	Tem por assunto, por objeto; versa, consta
TRATAR	8	148	Discutir, debater, questionar; ter por assunto, por objeto; versar, constar; dar o tratamento prescrito ou aconselhado
TRAZER	8	103	Conduzir ou transportar para este lugar; ocasionar, acarretar, causar
TRECHO	7	105	Fragmento de uma obra musical, literária ou artística; excerto; espaço ou intervalo de tempo
TROPICAL	6	115	Relativo aos trópicos ou às regiões da zona tórrida
UMIDADE	7	99	Qualidade ou estado de úmido ou ligeiramente molhado
UNIVERSAL	8	145	Geral, total; comum a todos; que abrange todas as coisas; que se estende a tudo; relativo ou pertencente ao universo (-)
UNIVERSO	7	252	O âmbito em que algo existe ou ocorre; o ambiente preferido; um todo composto de partes, harmonicamente dispostas; conjunto de todos os sóis, planetas, satélites, astros e asteróides do espaço; cosmos
URBANO	7	349	Relativo ou pertencente à cidade
USINA	6	164	Estabelecimento industrial, principalmente para a produção em grande escala, como a de ferro e aço, coque, gás, açúcar, eletricidade etc., ou para o abastecimento de água
UTILIDADE	8	113	Qualidade de útil; serventia
UTILIZAÇÃO	8	490	Ato ou efeito de utilizar
VARIEDADE	8	135	Diversidade, multiplicidade; variante, variação; subespécie
VERIFICAR	8	240	Investigar a verdade de; provar a verdade de; comprovar a exatidão de

<b>Palavra</b>	<b>Alcance</b>	<b>Freq.</b>	<b>Definição / Sinônimo</b>
VERIFICARAM	7	92	Ver VERIFICAR
VERIFICOU	4	93	Ver VERIFICAR
VERSO	4	108	Cada linha de uma composição poética; as composições poéticas em geral
VERTICAL	8	119	Perpendicular ao plano horizontal
VIRTUDE	8	101	(em virtude de) Em consequência de; por causa de
VISANDO	8	238	Ter por fim ou objetivo; ter em vista
VOLTADO	8	98	Apontado em certa direção; com o objetivo de

## REFERÊNCIAS BIBLIOGRÁFICAS

- ALMEIDA FILHO, J. C. P., OEIRAS, J. Y. Y.; ROCHA, H. V. **Português na internet: Questões de planejamento e produção de materiais**. 1998. Disponível em: [http://hera.nied.unicamp.br/teleduc/publicacoes/joeiras\\_ribie98.pdf](http://hera.nied.unicamp.br/teleduc/publicacoes/joeiras_ribie98.pdf). Acesso em: 31 Dezembro 2005.
- ANDERSON, R. C.; FREEBODY, P. Vocabulary knowledge. In: GUTHRIE, J. **Comprehension and teaching: research reviews**. Newark: International Reading Association, 1981. P.77-117.
- BACON, S.; FINNEMANN, M. A study of attitudes, motives, and strategies of university foreign language students and their disposition to authentic oral and written input. **The Modern Language Journal**, v.74, n.4, p.459-473. 1990.
- BAMFORD, J.; DAY, R. R. **Extensive Reading: What Is It? Why Bother?** 1997. Disponível em: <http://www.jalt-publications.org/tlt/files/97/may/extensive.html>. Acesso em: 27 Julho 2005.
- BARCROFT, J. Distinctiveness and bidirectional effects in input enhancement for vocabulary learning. **Applied Language Learning**, v.13, n.2, p.47-73. 2003.
- BAUER, L.; NATION, I. S. P. Word families. **International Journal of Lexicography**, v.6, n.2, p.253-279. 1993.
- BECK, I. L., PERFETTI, C. A.; MCKEOWN, M. G. The effects of long-term vocabulary instruction on lexical access and reading comprehension. **Journal of Educational Psychology**, v.74, p.506-521. 1982.
- BENGELEIL, N. F.; PARIBAKHT, T. S. L2 Reading Proficiency and Lexical Inferencing by University EFL Learners. **The Canadian Modern Language Review**, v.61, n.2, p.225-249. 2004.
- BOWLES, M. What's wrong with Monbusho's prescribed word list? **The Language Teacher**, v.25, n.1, p.102-116. 2001.
- BREEN, M. Authenticity in the language classroom. **Applied Linguistics**, v.6, n.1, p.60-70. 1985.
- CAMBRIDGE International Corpus. Cambridge. 2005. Disponível em: [http://www.cambridge.org/elt/corpus/international\\_corpus2.htm](http://www.cambridge.org/elt/corpus/international_corpus2.htm). Acesso em: 9 Maio 2006.
- CARROLL, B. **Testing Communicative Performance**. Oxford: Pergamon Press. 1980.

CARTER, R.; MCCARTHY, M. **Vocabulary and Language Teaching**. New York: Longman. 1988.

CARVER, R. The highly lawful relationships among pseudoword decoding, word identification, spelling, listening, and reading. **Scientific Studies of Reading**, v.7, p.127-154. 2003.

COBB, T.; HORST, M. Reading academic English: Carrying learners across the lexical threshold. In: FLOWERDEW, J.; PEACOCK, M. **Research perspectives on English for academic purposes**. Cambridge: Cambridge University Press, 2001.

COBB, T.; MORRIS, L. Vocabulary profiles as predictors of the academic performance of Teaching English as a Second Language trainees. **System**, v.32, p.75-87. 2004.

CONCORDANCE, version 3.2. Concordancesoftware, 2002.

CORPUS design. International Corpus of English. 2003. Disponível em: <http://www.ucl.ac.uk/english-usage/ice/>. Acesso em: 2 Maio 2005.

COXHEAD, A. A new academic word list. **Tesol Quarterly**, v.34, n.2, p.213-238. 2000.

DE BOT, K., PARIBAKHT, T. S.; WESCHE, M. Toward a lexical processing model for the study of second language vocabulary acquisition: Evidence from ESL reading. **Studies in Second Language Acquisition**, v.19, p.309-329. 1997.

DELLAR, H.; HOCKING, D. **Innovations: An Intermediate / Upper Intermediate Course**. Hove, UK: Language Teaching Publications. 2000.

DONLEY, K. M.; REPPEN, R. Using corpus tools to highlight academic vocabulary in SCLT. **Tesol Journal**, v.10, n.2/3, p.7-12. 2001.

DROOP, M.; VERHOEVEN, L. Language proficiency and reading ability in first and second language learners. **Reading Research Quarterly**, v.38, p.78-103. 2003.

DUBAY, W. H. **The principles of readability**. 2004. Disponível em: <http://www.nald.ca/fulltext/readab/02.htm>. Acesso em: 28 Abril 2005.

EYCKMANS, J. **Measuring receptive vocabulary size**. Utrecht: LOT. 2004.

FRANÇA, F. **Autor fez pesquisa com textos jornalísticos**. 2002. Disponível em: [http://atica.com.br/imprensa/borba\\_londrina.asp](http://atica.com.br/imprensa/borba_londrina.asp). Acesso em: 4 Maio 2005.

FRANCIS, W. N.; KUCERA, H. **Brown Corpus manual**. 1979. Disponível em: <http://khnt.hit.uib.no/icame/manuals/brown/INDEX.HTM>. Acesso em: 2 Maio 2005.

GATTOLIN, S. R. B. **O vocabulário na sala de aula de língua estrangeira: contribuições para a elaboração de uma teoria de ensino-aprendizagem.** Tese (Doutorado em Linguística Aplicada). Instituto de Estudos da Linguagem, UNICAMP, Campinas, 2005. p.350.

HERMAN, P. A.; ANDERSON, R. C.; PEARSON, P. D.; NAGY, W. E. Incidental acquisition of word meaning from expositions with varied text features. **Reading Research Quarterly**, v.22, n.3, p.263-284. 1987.

HIRSCH, D.; NATION, I. S. P. What vocabulary is needed to read unsimplified texts for pleasure. **Reading in a foreign language**, v.8, p.689-696. 1992.

HOWATT, A. P. R. **A history of English language teaching.** Oxford: Oxford University Press. 1984.

HU, M.; NATION, I. S. P. Unknown vocabulary density and reading comprehension. **Reading in a foreign language**, v.13, n.1, p.403-430. 2000.

HUCKIN, T.; COADY, J. Incidental vocabulary acquisition in second language - A review. **Studies in Second Language Acquisition**, v.21, p.181-193. 1999.

HULSTIJN, J. Mnemonic methods in foreign language vocabulary learning. In: COADY, J.; HUCKIN, T. **Second language vocabulary acquisition.** New York: Cambridge University Press, 1997. p.203-224.

HUNT, A. A framework for developing EFL reading vocabulary. **Reading in a Foreign Language**, v.17, n.1, p.23-59. 2005.

KAY, S.; JONES, V. **Inside Out Student's Book Intermediate.** Oxford: Macmillan Heinemann. 2000.

KOPROWSKI, M. Investigating the usefulness of lexical phrases in contemporary coursebooks. **ELT Journal**, v.59, n.4, p.322-332. 2005.

KRASHEN, S.; RODRIGO, V.; GRIBBONS, B. The effectiveness of two comprehensible-input approaches to foreign language instruction at the intermediate level. **System**, v.32, p.53-60. 2004.

KYONGHO, H.; NATION, P. Reducing the vocabulary load and encouraging vocabulary learning through reading newspaper. **Reading in a Foreign Language**, v.6, n.1, p.323-335. 1989.

LARSEN, T. **The orderly product of a disordered mind.** 2005. Disponível em: <http://www.christianitytoday.com/bc/2005/002/12.16.html>. Acesso em: 9 Maio 2005.

LAUFER, B. The lexical plight in second language reading: Words you don't know, words you think you know and words you can't guess. In: COADY, J.; HUCKIN, T. **Second language vocabulary acquisition**. New York: Cambridge University Press, 1997. p.20-34.

LAUFER, B. Vocabulary acquisition in a second language: do learners really acquire most vocabulary by reading? Some empirical evidence. **The Canadian Modern Language Review**, v.59, n.4, p.567-587. 2003.

LAUFER, B.; HILL, M. What lexical information do l2 learners select in a call dictionary and how does it affect word retention? **Language Learning & Technology**, v.3, n.2, p.58-76. 2000.

LAUFER, B.; NATION, P. A vocabulary-size test of controlled productive ability. **Language Testing**, v.16, n.1, p.33-51. 1999.

LAUFER, B.; NATION, P. Vocabulary Size and Use: Lexical Richness in L2 Written Production. **Applied Linguistics**, v.16, n.3, p.307-322. 1995.

LAUFER, B.; YANO, Y. Understanding unfamiliar words in a text: Do L2 learners understand how much they don't understand? **Reading in a foreign language**, v.13, n.2, p.549-566. 2001.

LEELA MOHD, A. **A Concise Overview of Some ESL Methods**. 1989. Disponível em: <http://www.melta.org.my/ET/1989/main2.html>. Acesso em: 12 Fevereiro 2006.

LEWKOWICZ, J. **Investigating authenticity in language testing**. Thesis (English Language Doctorate). Department of Linguistics and Modern English Language, University of Lancaster, Lancaster, 1997.

LITTLE, D.; DEVITT, S.; SINGLETON, D. **Authentic texts in Foreign Language Teaching: Theory and Practice**. Dublin: Authentik. 1988.

LIU, N.; NATION, P. Factors affecting guessing vocabulary in context. **RELC Journal**, v.16, p.33-42. 1985.

LOB Corpus. LOB. 2002. Disponível em: [http://www.essex.ac.uk/linguistics/clmt/w3c/corpus\\_ling/content/corpora/list/private/LOB/lob.html](http://www.essex.ac.uk/linguistics/clmt/w3c/corpus_ling/content/corpora/list/private/LOB/lob.html). Acesso em: 2 Maio 2005.

LYNCH, A. Authenticity in language teaching: some implications for the design of listening materials. **British Journal of Language Teaching**, v.20, p.9-16. 1982.

LYNCH, J. **Samuel Johnson**. 2003. Disponível em: <http://www.andromeda.rutgers.edu/~jlynch/Johnson>. Acesso em: 27 Abril 2005.

MALLIKARJUN, B. **Vocabulary education**. 2002. Disponível em: <http://www.languageinindia.com/nov2002/vocabulary.html>. Acesso em: 2 Maio 2005.

MASON, B.; KRASHEN, S. Is form-focused vocabulary instruction worthwhile? **RELC Journal**, v.35, n.2, p.179-185. 2004.

MCNEILL, A. What makes authentic materials different? The case of English language materials for education television. In: BIRD, N. **Language and Learning**. Hong Kong: Hong Kong Institute of Education, 1994. p.314-326.

MEARA, P. **The importance of an early emphasis on L2 vocabulary**. 1995. Disponível em: <http://www.jalt-publications.org/ilt/files/95/feb/meara.html>. Acesso em: 16 Maio 2005.

MEARA, P.; FITZPATRICK, T. Lex30: an improved method of assessing productive vocabulary in an L2. **System**, v.28, p.19-30. 2000.

MEZYNSKI, K. Issues concerning the acquisition of knowledge: Effects of vocabulary training on reading comprehension. **Review of Educational Research**, v.53, p.253-279. 1983.

MORRISON, B. Using news broadcasts for authentic listening comprehension. **ELT Journal**, v.43, n.1, p.14-18. 1989.

MURPHY, E. **Strangers in a Strange Land: Teachers' beliefs about teaching and learning French as a second or foreign language in online learning environments**. 2000. Disponível em: <http://www.nlc-bnc.ca/obj/s4/f2/dsk1/tape3/PQDD0020/NQ54025.pdf>. Acesso em: 4 Maio 2006.

NAGY, W. E., HERMAN, P. A.; ANDERSON, R. C. Learning words from context. **Reading Research Quarterly**, v.20, n.2, p.233-253. 1985.

NATION, I. S. P. **Learning vocabulary in another language**. Cambridge: Cambridge University Press. 2001.

NATION, I. S. P. Vocabulary. In: NUNAN, D. **Practical English language**. New York: McGraw-Hill, 2002. p.129-152.

NATION, P. **Measuring readiness for simplified material: A test of the first 1000 words of English**. 1993. Documento do Serviço de Reprodução ERIC Nº: ED 371 585. Acesso em: 20 Outubro 2004.

NATION, P. **Teaching and learning vocabulary**. New York: Newbury House. 1990. 297p.

NATION, P.; HWANG, K. Where would general service vocabulary stop and special purposes vocabulary begin? **System**, v.23, n.1, p.35-41. 1995.

NATION, P.; WARING, R. **Vocabulary size, text coverage and word lists**. 2002. Disponível em: <http://www1.harenet.ne.jp/~waring/papers/cup.html>. Acesso em: 20 Junho 2005.

NOAH Webster and America's first dictionary. Merriam-Webster On Line. 2004. Disponível em: <http://www.m-w.com/info/noah.htm>. Acesso em: 28 Abril 2005.

NUNAN, D. **The learner-centred curriculum: a study in second language teaching**. Cambridge: Cambridge University Press. 1988.

OH, S. Y. Two types of input modification and EFL reading comprehension: Simplification versus elaboration. **TESOL Quarterly**, v.35, n.1, p.69-96. 2001.

Oxford English Dictionary. **History of the Dictionary**. Oxford University Press. 2005. Disponível em: <http://www.oed.com/about/history.html#supp..> Acesso em: 31 Dezembro 2005.

PARIBAKHT, S.; WESCHE, M. Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. In: COADY, J.; HUCKIN, T. **Second Language Vocabulary: a rationale for pedagogy** Cambridge: Cambridge University Press, 1997. p.174-200.

PARIBAKHT, T. S.; WESCHE, M. Reading and "incidental" vocabulary acquisition. **Studies in Second Language Acquisition**, v.21, p.195-224. 1999.

PINHEIRO, G. M.; ALUÍSIO, S. M. **Corpus Nilc: descrição e análise crítica com vistas ao projeto Lacio-Web**. São Carlos, SP: USP. 2003. 60 p. (Série de Relatórios do Núcleo Interinstitucional de Lingüística Computacional)

QIAN, D. D. Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. **Canadian Modern Language Review**, v.56, n.2, p.282-307. 1999.

QIAN, D. D. Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. **Language Learning**, v.52, n.3, p.513-536. 2002.

READ, J. Research in teaching vocabulary. **Annual Review of Applied Linguistics**, v.24, p.146-161. 2004.

READ, J. The development of a new measure of L2 vocabulary knowledge. **Language Testing**, v.10, p.355-371. 1993.

READ, J. Vocabulary and testing. In: SCHMITT, N.; MCCARTHY, M. **Vocabulary: Description, Acquisition and Pedagogy**. Cambridge: Cambridge University Press, 1997. p.303-320.

RICHARDS, J. C., PLATT, J.; PLATT, H. **Longman dictionary of language teaching and applied linguistics**. London: Longman. 1992.

RICHARDS, J. C.; RODGERS, T. S. **Approaches and methods in language teaching**. Cambridge, UK: Cambridge University Press. 2000.

RIDDER, I. Visible or invisible links: Does the highlighting of hyperlinks affect incidental vocabulary learning, text comprehension, and the reading process? **Language Learning and Technology**, v.6, n.1, p.123-146. 2002.

RIEDER, A. **Implicit and explicit learning in incidental vocabulary acquisition**. 2003. Disponível em: [www.univie.ac.at/Anglistik/views/03\\_2/RIE\\_SGLE.PDF](http://www.univie.ac.at/Anglistik/views/03_2/RIE_SGLE.PDF). Acesso em: 9 Agosto 2004.

ROTT, S. The effect of exposure frequency on intermediate language learners' incidental vocabular acquisition and retention through reading. **Studies in Second Language Acquisition**, v.21, p.589-619. 1999.

SARDINHA, T. B. **Os padrões de só – um estudo baseado em corpus**. 2000. Disponível em: [http://lael.pucsp.br/~tony/2000so\\_propor.pdf](http://lael.pucsp.br/~tony/2000so_propor.pdf). Acesso em: 4 Julho 2005.

SCARAMUCCI, M. V. R. **O papel do léxico na compreensão em leitura em língua estrangeira: Foco no produto e no processo**. Tese (Doutorado em Linguística). Instituto de Estudos da Linguagem, UNICAMP, Campinas, 1995.

SCHMITT, N. **Vocabulary in language teaching**. Cambridge: Cambridge University Press. 2000.

SOARS, L.; SOARS, J. **New Headway Upper Intermediate Student's Book**. Oxford: Oxford University Press. 1998.

STANOVICH, K. **Progress in understanding reading: Scientific foundations and new frontiers**. New York: Guilford Press. 2000.

STERN, H. **Issues and options in language teaching**. Disponível em: [http://144.82.31.20/ioe/cms/get.asp?cid=882&882\\_1=830&var1=3&var2](http://144.82.31.20/ioe/cms/get.asp?cid=882&882_1=830&var1=3&var2). Acesso em: 12 Outubro 2004.

SWAFFAR, J. Reading authentic tests in a foreign language: a cognitive model. **The Modern Language Journal**, v.69, n.1, p.15-34. 1985.

THE DICTIONARY revolution. Bangkok Post. 2002. Disponível em: <http://www.bangkokpost.net/education/site2002/cvap0202.htm>. Acesso em: 2 Maio 2005.

THORNDIKE, R. L. **Reading comprehension education in 15 countries**. Stockholm: Almqvist & Wiksell. 1973.

THORNDIKE, E. L.; LORGE, I. **The teacher's word book of 30,000 words**. New York: Teacher's College, Columbia University. 1944.

TICKOO, M. **ESP: State of the art**. Singapore: Singapore University Press / RELC. 1988.

TONO, Y. **Learner corpora and SLA research**. 2003. Disponível em: <http://leo.meikai.ac.jp/~tono/>. Acesso em: 2 Maio 2005.

TSCHIRNER, E. Breadth of vocabulary and advanced English study: An empirical investigation. 2004. Disponível em: <http://e-flt.nus.edu.sg/v1n12004/tschirner.pdf>. Acesso em: 4 Maio 2006.

UMAMOTO, T. **Vocabulary teaching in an EFL context**: Selecting the appropriate words. Dissertation (Master in Education). Faculty of School Education, Hiroshima University, Hiroshima, 1989. p.86.

URANO, K. **Lexical simplification and elaboration**: Sentence comprehension and incidental vocabulary acquisition. Dissertation (Master of Arts in English as a Second Language). University of Hawaii, 2000. p.88.

Universia Brasil. **Cresce o número de candidatos ao Certificado de Língua Portuguesa**. 2005. Disponível em: [http://www.universiabrasil.net/html/materia/materia\\_haii.html](http://www.universiabrasil.net/html/materia/materia_haii.html). Acesso em: 12 Janeiro 2006.

VERMEER, A. Coming to grips with lexical richness in spontaneous speech data. **Language Testing**, v.17, n.1, p.65-83. 2000.

WARING, R. At what rate do learners learn and retain new vocabulary from reading a graded reader? **Reading in a Foreign Language**, v.15, n.2, p.130-163. 2003.

WARING, R. **Second language vocabulary acquisition, linguistic context and vocabulary task design**. 1995. Disponível em: <http://www1.harenet.ne.jp/~waring/papers/BC.html>. Acesso em: 11 Maio 2005.

WEST, M. **A general service list of English words**. London: Longman, Green & Co. 1953.

WHAT is the BNC? British National Corpus. 2004. Disponível em: <http://www.natcorp.ox.ac.uk/what/index.html>. Acesso em: 28 Abril 2005.

WIDDOWSON, H. G. Context, Community, and Authentic Language. **TESOL Quarterly**, v.32, n.4, p.705-716. 1998.

WIDDOWSON, H. G. **Explorations in Applied Linguistics**. Oxford: Oxford University Press. 1979.

WIDDOWSON, H. G. **Teaching Language as Communication**. Oxford: Oxford University Press. 1978.

WORTHINGTON, D.; NATION, P. Using texts to sequence the introduction of new vocabulary in an EAP course. **RELC Journal**, v.27, n.2, p.1-11. 1996.

XUE, G.; NATION, I. S. P. A university word list. **Language Learning and Communication**, v.3, n.215, p.229. 1984.

ZAHAR, R.; COBB, T.; SPADA, N. Acquiring vocabulary through reading: Effects of frequency and contextual richness. **Canadian Modern Language Review**, v.57, n.4, p.541-572. 2001.

ZILLES, M. **O ensino e a aquisição de vocabulário em contexto de instrução de língua estrangeira**. Dissertação (Mestrado em Estudos da Linguagem). Instituto de Letras, UFRGS, Porto Alegre, 2001.

ZIMMERMAN, C.; BOYD, C. Do reading and interactive vocabulary instruction make a difference? An empirical study. **TESOL Quarterly**, v.31, n.1, p.121-140. 1997.