

Alfabetização e Conhecimento Lingüístico: o Projeto e-Labore

Thaís Cristófaró Silva (UFMG)
Raquel Fontes Martins (UFMG)
Leonardo Almeida (UFMG)
Daniela Guimarães (UFMG)

RESUMO

O processo de apropriação da escrita envolve uma série de habilidades diferenciadas, porém co-relacionadas, dentre as quais podemos destacar a organização do conhecimento lingüístico. Entendemos por conhecimento lingüístico um conjunto habilidades físicas, biológicas e sociais que permitem unicamente à espécie humana proceder à comunicação através da linguagem oral (PIERREHUMBERT, 2006, p. 516). Assim, a linguagem é compreendida como um fenômeno social que engloba um conjunto de propriedades cognitivas dos seres humanos.

Este trabalho tem por objetivo apresentar uma proposta de documentação e análise de material lingüístico decorrente de produções textuais de crianças de 6 a 12 anos da cidade de Belo Horizonte: Projeto e-Labore, Laboratório Eletrônico de Oralidade e Escrita. Argumenta-se que a avaliação das produções textuais coletadas oferece informações importantes sobre a organização do conhecimento lingüístico de crianças, subsidiando teorias que sugerem a relevância do uso na organização gramatical da linguagem (KEMMER; BARLOW, 2000; ELLIS, 2002; BYBEE, 2006). Na primeira seção, apresentamos as principais características do projeto e-Labore. Na segunda seção, descrevemos aspectos metodológicos relevantes para o desenvolvimento da pesquisa. A seguir, na terceira seção, tecemos comentários sobre o ensino fundamental no Brasil e

ponderamos sobre os principais métodos utilizados na inserção inicial do aprendiz ao conhecimento da leitura e escrita. Na quarta seção, apresentamos alguns resultados já obtidos pelo projeto e-Labore. Por fim, na quinta seção, apresentamos as contribuições, bem como as próximas etapas do projeto e-Labore.

1 Introdução

O projeto e-Labore tem como objetivo central construir e disponibilizar para a comunidade científica um banco de dados de material escrito por crianças e pré-adolescentes de 6 a 12 anos de idade, falantes do português brasileiro. Este projeto tem caráter inovador, à medida que busca compilar e analisar uma grande quantidade de produções textuais infantis e oferecer tratamento automatizado na análise do material coletado.¹

O projeto e-Labore é também uma iniciativa inovadora por reunir e disponibilizar material para pesquisadores de várias áreas do conhecimento: lingüistas, educadores, fonoaudiólogos, engenheiros da fala etc. Esses pesquisadores se beneficiarão de tal ferramenta para as mais diversas atividades de pesquisa e aplicabilidade tecnológica, tais como, projetos educacionais da língua portuguesa, estudos de patologias da fala infantil, implementação tecnológica de material didático e formulação de teorias sobre a organização da linguagem humana. Na etapa atual de desenvolvimento do projeto, estamos em fase de implementação tecnológica do acervo: digitalização, digitação e arquivamento

¹ Lancaster Corpus Children's Project Writing (<http://bowlandfiles.lancs.ac.uk/lever/index.htm>) e Child Language Data Exchange System (<http://chilides.psy.cmu.edu/>) são dois exemplos de corpora da escrita infantil da língua inglesa com características análogas ao Projeto e-Labore.

das produções textuais. Para informações precisas sobre este processo de documentação, consulte Cristóvão Silva *et al.* (2006) e a página do projeto na Internet (Disponível em: <www.projetoaspa.org/elabore>. Acesso em: abril 2007).

Uma vez que a documentação tenha sido concluída, a equipe do projeto e-Labore pretende empreender, como tarefa inicial, a formulação de uma classificação quantificada dos desvios ortográficos a partir das produções escritas coletadas. Tal taxonomia deverá oferecer informações importantes sobre a superação de desvios ortográficos ao longo da vida escolar da criança e sua relação com a ampliação do léxico infantil. Vale ressaltar que uma taxonomia proposta em Oliveira *et al.* (1989) teve um impacto importante no conhecimento da apropriação da escrita nas séries iniciais. Estes últimos autores contaram com a implementação manual dos dados, uma vez que os recursos tecnológicos disponíveis na época eram mais restritos. Dessa forma, o projeto e-Labore avança na proposta de apresentação de uma taxonomia de desvios ortográficos ao avaliar um grande número de produções textuais e também por avaliar tais desvios automaticamente.

Finalmente, vale ressaltar que a fundamentação teórica de tal projeto segue a proposta de o uso lingüístico refletir a organização gramatical e, portanto, o conhecimento lingüístico dos falantes (KEMMER; BARLOW, 2000; ELLIS, 2002; BYBEE, 2006). Nesta proposta, a aquisição da linguagem é vista como decorrente da aprendizagem gradual de padrões probabilísticos da língua (PIERREHUMBERT, 2003). Uma das hipóteses deste trabalho é a de que, assim como ocorre na aquisição da fala (GIERUT; STORKEL, 2002; MORRISETTE, 2004; YONEYAMA *et al.*, *in press*), a frequência de ocorrência de unidades específicas da língua (como a sílaba, a palavra, a letra) influencia na forma como a criança adquire a escrita. Assim, os resultados a serem disponibilizados pelo projeto e-

Labore podem vir a oferecer elucidações importantes para teorias lingüísticas. A próxima seção apresenta os principais aspectos metodológicos relevantes para o desenvolvimento do projeto e-Labore.

2 Aspectos Metodológicos

O corpus do projeto e-Labore é construído através da coleta e digitação de produções textuais de crianças e pré-adolescentes com idade entre 6 e 12 anos que cursam da 1ª a 6ª série do ensino fundamental. A faixa etária de 6 a 12 anos foi selecionada, por ela compreender o início da aquisição e do aprendizado da escrita, bem como a fixação e a utilização do código escrito (LABOV, 1964; CHAMBERS, 1995; MINISTÉRIO DA EDUCAÇÃO, 2006). Assim, esse é um período importante, pois sendo a fase em que se inicia a aquisição da escrita, observam-se grandes oscilações na grafia das palavras, as quais demonstram a interação da criança com o objeto escrita e suas hipóteses iniciais sobre esse objeto. Há indícios de que a fala de adolescentes apresenta características particulares as quais merecem atenção diferenciada da linguagem infantil (KERSWILL, 1996; ECKERT, 1996).

Há intenção de se realizar três coletas de redações. A coleta piloto, de menor escala, foi realizada e, atualmente, todas as redações dessa coleta já foram digitalizadas, digitadas e arquivadas. Os resultados que serão apresentados posteriormente, neste artigo, consideram dados da coleta piloto. Cristófaró Silva *et al.* (2006) apresentam características detalhadas da metodologia empregada na documentação do projeto e-Labore. Neste artigo,

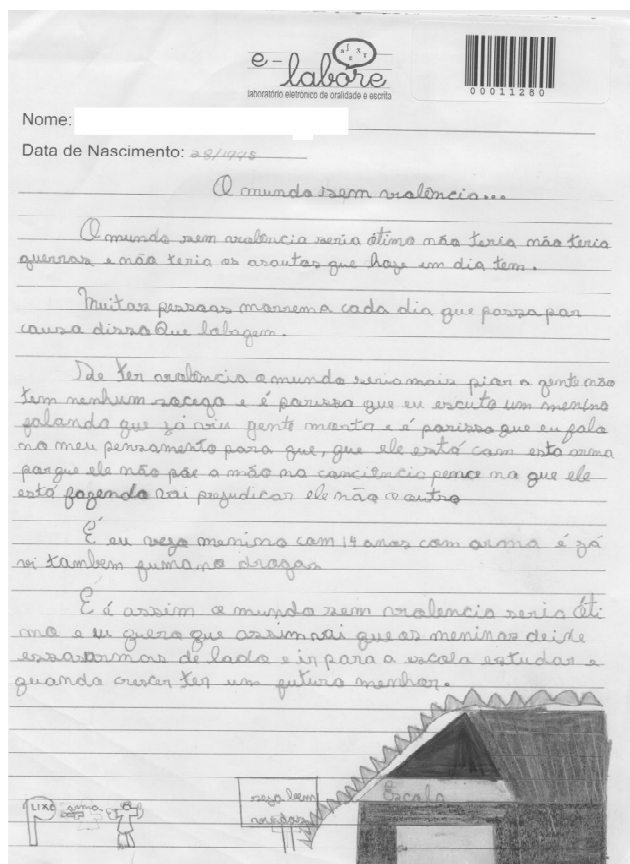
apresentamos, sumariamente, as principais características metodológicas do projeto para oferecer ao leitor instrumentos de avaliação dos resultados discutidos.

As escolas participantes do e-Labore estão situadas na cidade de Belo Horizonte, Minas Gerais. Foram selecionadas escolas da rede pública e da rede particular com o objetivo de avaliar se há diferenças no processo de aquisição da escrita entre os alunos desses dois tipos de escolas. O tipo de escola – se pública ou particular – tem sido considerado fator importante na avaliação de resultados do nível de sucesso de apropriação da escrita (MOLLICA, 2000; OLIVEIRA *et al.*, 1989).

O projeto e-Labore resguarda a privacidade dos alunos, professores e escolas participantes ao alocar um código numérico individual a cada um deles. Sendo assim, será possível consultar produções diversas de um mesmo aluno, considerar o desenvolvimento específico de professores individualmente e avaliar características gerais das escolas individualmente. Vale ressaltar que, também, será possível consultar dados de alunos, professores, ou de escolas de forma conjunta.

A coleta das redações é realizada como uma atividade de produção textual rotineira, pela professora da turma, na expectativa de não modificar as atividades regulares desenvolvidas em sala de aula. O tema da redação em cada turma é definido pelos próprios professores e, sendo assim, há diversificação dos tópicos nas produções textuais coletadas. Essa decisão metodológica busca garantir que o vocabulário utilizado pelas crianças seja o mais próximo possível daquele que elas utilizam habitualmente, em sala de aula. Na Figura 1, é apresentado um exemplo de produção textual da coleta do projeto e-Labore.

Figura 1: Produção textual de estudante da 4a série de uma escola da rede pública de Belo Horizonte.



Na presente etapa do projeto, optamos por documentar, exclusivamente, os desvios ortográficos. Sendo assim, desvios em colocações pronominais, pontuação, regência (verbal ou nominal), concordância (verbal ou nominal) não são indicados na documentação. Contudo, em etapas posteriores, poderá ser empreendida a documentação de parâmetros adicionais como os mencionados acima. As palavras grafadas com desvios ortográficos pelos alunos são marcadas por um par de chaves, e a forma padrão é apontada entre colchetes, como nos exemplos: {patu}[pato] e {anzou}[anzol]. O registro dos dados desta maneira permite que seja possível identificar qual palavra apresentou desvio ortográfico, bem como qual(is) o(s) grafema(s) envolvido(s) no caso em questão. Será possível identificar palavras que não apresentam desvios ortográficos e que, portanto, não

implicam tratamento diferenciado. Por outro lado, serão identificadas as palavras que, freqüentemente, apresentam desvios ortográficos e, neste caso, poderemos avaliar qual é o grau de desvio observado em diferentes séries do ensino fundamental. Na próxima seção, tratamos brevemente de aspectos relacionados ao ensino fundamental e à apropriação da escrita.

3 Ensino Fundamental e Alfabetização

São inúmeros os trabalhos voltados para a apropriação do sistema de escrita no português. Embora com diferentes enfoques, na grande maioria desses trabalhos, pode-se generalizar que: há necessidade de melhorar o processo de apropriação da escrita dos estudantes brasileiros e há necessidade de se obter um conhecimento mais amplo dos problemas enfrentados em diferentes regiões do país.

Iniciativas governamentais e não-governamentais têm sido tomadas com o objetivo de se buscar a melhoria do sistema educacional do país de um modo geral. Uma dessas iniciativas diz respeito à ampliação do ensino fundamental para nove anos em 2006 (BATISTA, 2005). Existe também o documento que rege a organização do ensino: os Parâmetros Curriculares Nacionais – PCN's (Disponível em: <<http://www.mec.gov.br/sef/sef/pcn.shtm>>. Acesso em: abril 2007).

Para o propósito da discussão fomentada neste trabalho, centraremos nos seguintes pontos: caracterização de alfabetização e letramento, relação teoria e métodos na alfabetização e estratégias potenciais de documentação com o objetivo de ter impacto na melhoria do ensino do país.

Consideremos então as noções de alfabetização e letramento. Seguindo Soares (2003), assumimos que a alfabetização diz respeito à apropriação de um sistema de escrita e às habilidades envolvidas em seu uso. Letramento diz respeito ao desenvolvimento de competências para o uso dessa tecnologia em práticas sociais. A alfabetização pode ser compreendida como conhecimento e uso de um sistema convencional que deve ser aprendido de forma sistemática. Ou seja, a alfabetização diz respeito à “aprendizagem da técnica, domínio do código convencional da leitura e da escrita e das relações fonema/grafema, do uso dos instrumentos com os quais se escreve, não é pré-requisito para o letramento” (SOARES, 2003, p. 1). Alfabetização e letramento são processos indissociáveis, mas distintos, dizendo respeito a processos cognitivos e com resultados diferentes e complementares. Aprender o código escrito (alfabetizar-se) é também aprender a fazer uso social da leitura e da escrita (letrar-se).

Contudo, há especificidades da alfabetização. Há uma maneira correta de se grafar uma palavra. Temos, por exemplo, o VOLP² (Vocabulário Ortográfico da Língua Portuguesa) da Academia Brasileira de Letras e inúmeros dicionários que indicam a grafia correta das palavras. De fato, há expectativa em qualquer leitor de encontrar o registro ortográfico oficial das palavras que lê. Por esta razão, os desvios ortográficos causam profundo impacto nos leitores (sejam professores ou não!). A escola, de maneira geral, indica com veemência os desvios ortográficos nas produções textuais dos alunos. Portanto, grafar uma palavra como o esperado faz parte da expectativa do leitor e também de quem a escreveu. O aprendizado dos padrões ortográficos que se inicia na alfabetização vigora por toda a vida do leitor-escritor.

² (Disponível em: <<http://www.academia.org.br/>>, procurar aba “nossa língua” > “vocabulário ortográfico”. Acesso em: abril 2007).

Tradicionalmente, as professoras alfabetizadoras tinham um método (qualquer que fosse este!) e o seguiam categoricamente utilizando um manual do professor (alfabetizador) o qual explicitava, exatamente, como utilizar o material a ser empregado: tipicamente cartilhas. Sem discutir o mérito ou não das cartilhas, podemos dizer que as professoras tinham, pelo menos, um método a seguir.

Com o impacto do *construtivismo* na alfabetização, observamos no país a discussão de aspectos teóricos da alfabetização, mas há falta de um método. O *construtivismo* sugere a primazia da interação com o objeto de conhecimento (leitura-escrita) na construção do código escrito pela criança (FERREIRO, 1995). Em linhas gerais, a criança constrói o seu conhecimento a partir da formulação progressiva de hipóteses a respeito da escrita. Neste processo construtivo, a criança então aprende a ler e a escrever. O letramento seria muito contemplado no *construtivismo*, mas aspectos de forma ligados à alfabetização (como a grafia correta das palavras) apresentariam deficiências. Como consequência desse processo, as crianças alfabetizadas numa abordagem *construtivista* conseguiriam ler um texto, interagir com ele, compreendê-lo, atribuindo-lhe significado, contudo, apresentariam problemas, como, por exemplo, muitos desvios ortográficos na escrita. Claro que há méritos na abordagem construtivista, mas certamente falta um método que ofereça às professoras um impacto de ação na sala de aula a partir de pressupostos teóricos quaisquer.

Recentemente, tem sido fomentado no país o debate sobre o *Construtivismo* e o *Método Fônico* (SOARES, 2003, 2004). As críticas ao *método fônico* dizem respeito ao fato de este se limitar a alfabetização à (de)codificação: transformar letra em som (leitura) e som em letra (escrita). Tal método se centraria mais na forma, na consciência fonológica (que estaria relacionada ao conhecimento do aspecto sonoro da língua) e nem tanto no

significado do texto (sendo deficitário no processo de letramento). Assume-se que as crianças alfabetizadas pelo *método fônico* teriam a capacidade de ler um texto, mas de modo limitado, sem interagir com ele, sem entendê-lo, sem lhe atribuir significado de uma forma adequada e observa-se menor grau de desvios ortográficos na escrita.

Sem julgar a pertinência e adequação de qualquer uma das abordagens mencionadas nesta seção, podemos concluir que há algo que deve ser revisto no processo de alfabetização e letramento. O projeto e-Labore pretende ser uma contribuição em direção à compreensão dos desvios encontrados na escrita das crianças, principalmente no que tange ao aspecto fonológico. A partir da avaliação das dificuldades dos alunos o professor poderá tentar procurar soluções que sejam apropriadas ao tipo de problema enfrentado e à etapa na qual se encontra a criança. Na próxima seção, apresentamos alguns resultados do projeto e-Labore e sugerimos estratégias potenciais de documentação com o objetivo de promover impacto na melhoria do ensino do país.

4 Resultados parciais do projeto e-Labore

Na primeira fase, a equipe do projeto e-Labore coletou 1952 redações em 67 turmas de 1^a a 6^a série, bem como digitalizou e digitou essas produções textos. Na análise de corpora, podem-se destacar duas importantes medidas. O número total de palavras obtidas, o qual denominamos *freqüência de ocorrência*, e o número de palavras que foram atestadas individualmente, o qual denominamos *freqüência de tipo* (BYBEE, 2001). A Tabela 1 apresenta o número de turmas, de redações, de palavras distintas (freqüência de

tipo) e o total de palavras obtidas (frequência de ocorrência). Estes dados refletem o corpus ao final da primeira coleta.

Tabela 1. Situação do corpus do e-Labore após o encerramento da primeira coleta.

Série	Turmas	Redações	Palavras Distintas	Palavras
1a Série	12	324	2.285	22.476
2a Série	11	296	3.095	27.002
3a Série	12	379	4.137	42.582
4a Série	12	353	4.845	42.285
5a Série	10	296	4.319	37.659
6a Série	10	304	4.136	35.455
Total	67	1.952	11.415	207.459

O corpus descrito acima é formado por 11.415 palavras distintas que se repetem em um total de 207.459 palavras. Espera-se que, ao final das três coletas, o corpus do e-Labore atinja cerca de 1,5 milhões de palavras. A seguir, apresentamos a Tabela 2 que tenta indicar algumas características do corpus do projeto e-Labore.

Tabela 2. Lista de palavras mais frequentes no corpus do e-Labore.

	Palavras			Substantivos			Verbos		
	Palav.	Freq.	%	Subst.	Freq.	%	Verb.	Freq.	%
1	e	9.622	4,64	natal	1.137	0,55	é	2.995	1,44
2	que	5.997	2,89	dia	1.108	0,53	foi	1.218	0,59
3	o	5.975	2,88	peessoas	993	0,48	tem	1.006	0,48
4	a	5.920	2,86	casa	747	0,36	era	833	0,40
5	de	5.528	2,66	mãe	589	0,28	estava	694	0,34
6	um	3.330	1,60	escola	510	0,25	ser	584	0,28
7	eu	3.307	1,60	ano	500	0,24	são	492	0,24
8	para	3.206	1,54	gente	496	0,24	tinha	491	0,24
9	não	3.058	1,47	mundo	446	0,22	está	481	0,23
10	é	2.993	1,44	projeto	384	0,19	vai	459	0,22
11	uma	2.268	1,09	crianças	381	0,18	fazer	429	0,21
12	com	1.969	0,95	família	327	0,16	ter	408	0,20
13	os	1.802	0,87	pai	323	0,16	vou	391	0,19
14	no	1.739	0,84	anos	323	0,16	pode	350	0,17
15	na	1.667	0,80	violência	315	0,15	acho	327	0,16

A Tabela 2 apresenta as 15 palavras (ou seja, itens de qualquer natureza gramatical), os 15 substantivos e os 15 verbos mais frequentes do corpus do e-Labore. Para cada um desses três casos, além da frequência de ocorrência da palavra, apresenta-se também a porcentagem de ocorrência em relação a todo o corpus.

A respeito do primeiro caso, as 15 palavras mais freqüentes no geral, se pode observar que, excetuando-se (9) *não* e (10) *é*, todas as outras são palavras gramaticais e não palavras de conteúdo. Outra observação a ser feita é que as palavras desse primeiro caso são, em sua maioria, monossilábicas. Somente (8) *para* e (11) *uma* são dissílabos, contudo, tais palavras podem sofrer redução no discurso, passando a monossílabos: *para* → [pá] ou [prá] e *uma* → [ua].

Sobre o segundo caso, os 15 substantivos mais freqüentes, deve ser ressaltado que eles, de modo geral, refletem os temas das redações propostas pelas professoras dos alunos participantes. Ainda, por exemplo, o substantivo (1) *Natal* reflete a época em que as redações que compõem o corpus do e-Labore foram coletadas: final de ano. Muitas professoras sugeriram o Natal como tema das redações e, por esse motivo, tal substantivo aparece como o mais freqüente de todo o corpus. O caráter temporal da coleta em discussão, tendo como tema o Natal, poderá ser reavaliado a partir dos dados da segunda coleta que se dará no período pré-férias.

Quanto ao terceiro caso apresentado na tabela acima, os 15 verbos mais freqüentes, pode-se notar que, excetuando-se o verbo achar em (15) *acho*, todos os outros são verbos irregulares (“ser”, “estar”, “ir”, “ter”, “fazer” e “poder”). Uma outra observação a ser feita quanto aos verbos mais freqüentes é que há 4 verbos no tempo passado: (2) *foi*, (4) *era*, (5) *estava* e (8) *tinha*. É provável que esses verbos no passado estejam refletindo o tipo textual utilizado na maioria das redações: a narração. Na narração, é muito comum o uso do tempo passado.

Os resultados acima são parciais, mas servem para ilustrar o papel da estatística na aprendizagem da escrita (e certamente da leitura). Uma hipótese a ser investigada é a de que algumas palavras freqüentes podem servir como “âncora” no aprendizado de outras

palavras (como ocorre na aquisição da fala, Cf. VIHMAN, 1996). A partir da produção escrita das crianças, poderão ser observados quais erros são reflexo da linguagem oral e quais erros decorrem de hipóteses relacionadas à ortografia de palavras já aprendidas pela criança (erros que são consequência de generalização de padrões de palavras conhecidas). Sendo assim, o professor poderá propor soluções específicas para problemas específicos.

5 Conclusão e perspectivas

Este trabalho descreveu as principais características do projeto e-Labore buscando indicar a relevância da documentação proposta. Adicionalmente, foram apresentados aspectos metodológicos de desenvolvimento do projeto. Nesta conclusão, ressaltamos contribuições possíveis de decorrerem do projeto e-Labore e indicamos perspectivas futuras.

Uma primeira contribuição do e-Labore seria o fato de esse corpus permitir ao professor conhecer as palavras que são mais freqüentes/ familiares às crianças, já que o corpus refletiria isto: o léxico específico das crianças, com informações sobre freqüência de tipo e de ocorrência. Com base nesse corpus, o professor alfabetizador pode trabalhar com essas palavras mais freqüentes – que fazem parte do uso real das crianças –, as quais poderiam funcionar como palavras-chaves para a criança aprender a ler e a escrever e até para que haja avanço na aprendizagem de novas palavras. Seguindo a teoria de uso que rege este projeto, podemos afirmar que a freqüência de ocorrência de uma palavra tem impacto importante na organização gramatical (BYBEE, 2006).

Uma outra contribuição do e-Labore seria quanto à questão da consciência fonológica. O e-Labore faz um mapeamento dos desvios ortográficos das crianças contendo informações sobre a frequência de tais desvios, os quais, muitas vezes, têm relação com a oralidade. Ao tomar conhecimento de quais são os desvios ortográficos que se relacionam com a oralidade, o professor pode trabalhar com a questão da consciência fonológica de modo mais específico e produtivo. O professor pode, por meio do corpus do e-Labore, saber quais são os tipos de desvios ortográficos mais frequentes que as crianças cometem e utilizar tal informação em sua prática de sala de aula. Isso pode ser feito com um filtro relativo à série escolar da criança, já que o e-Labore trabalha com crianças de 1ª a 6ª série.

Supondo que, ao consultar o corpus, o professor descubra que, na 2ª série, desvios ortográficos relacionados a alguns pares de consoantes homorgânicas desvozeadas/vozeadas (f/v; t/d...) são mais frequentes do que desvios ortográficos relacionados à variabilidade de grafia para o fonema /l/ em final de sílaba, ele então, poderá trabalhar primeiro com a questão da consciência fonológica no tipo de desvio ortográfico que é mais frequente na língua. Ou seja, o professor poderá conscientizar o aluno de que existe diferença entre sons vozeados e desvozeados e que essa diferença se reflete na escrita de tais sons, antes de trabalhar com a variabilidade de escrita para o fonema /l/.

Encontra-se em andamento a digitação da segunda coleta e o planejamento da terceira coleta. Nesta terceira coleta a ser realizada, gravaremos a oralidade de algumas das crianças e professoras participantes. Pretendemos também organizar o mapeamento dos desvios ortográficos que foram atestados e disponibilizá-los para a comunidade científica. Esperamos que os resultados do projeto e-Labore contribuam para melhorar o processo de apropriação da escrita dos estudantes brasileiros, bem como ofereçam elucidações importantes para teorias lingüísticas baseadas no uso.

Referências bibliográficas

BATISTA, A. A. G. *Organização da Alfabetização no Ensino Fundamental de 9 anos*. 1. ed. Belo Horizonte: Ceale/FaE/UFMG. 2005, v. 1. 59 p.

BYBEE, Joan. From usage to grammar: the mind's response to repetition. *Language* 82. p. 529-551, 2006.

CAPOVILLA, A.; CAPOVILLA, F. *Alfabetização e método fônico*. São Paulo: Mnemom, 2001.

CHAMBERS, J. *Sociolinguistic Theory*. Cambridge, Oxford: Blackwell, 1995.

CRISTÓFARO SILVA, T. *et al.* Aquisição da escrita infantil: a construção de um corpus do português brasileiro. *4th Workshop in Information and Human Language Technology*. Ribeirão Preto. Proceedings of the International Joint Conference IBERAMIA/SBIA/SBRN, 2006.

ECKERT, P. Age as a sociolinguistic variable. In.: *Handbook of Sociolinguistics*. Oxford: Blackwell. 1996. p. 151-168.

ELLIS, N. C. Frequency effects in language processing: a review with implications for theories of implicit and explicit language acquisition. *SSLA*, Cambridge, 24, p. 143-188, 2002.

FERREIRO, Emilia, Teberosky. *Psicogênese da língua escrita*. Tradução de D. M. Lichtenstein, L. Di Marco, M. Corso. Porto Alegre: Artes Médicas, 1985.

GIERUTE, J. A.; STORKEL, H. L. Markedness and the grammar in lexical diffusion of fricatives. *Clinical Linguistics and Phonetics*, v. 16, n. 2, p. 115-134, 2002.

KEMMER, S. E.; BARLOW, M. (Ed.). *Usage-based models of language*. Stanford California: CSLI Publications, 2000.

KERSWILL, P. 1996. Children, adolescents and language change. *Language Variation and Change* 8: p. 177-202, 1996.

LABOV, W. Stages in the acquisition of standard English. In: SHUY, Roger (Ed.). *Social Dialects and Language Learning*. Ed. Champaign, Illinois, 1964, p. 77-103.

MINISTÉRIO DA EDUCAÇÃO. Parâmetros Curriculares Nacionais. 2006. Disponível em: <<http://www.mec.gov.br/sef/sef/pcn.shtm>>. Acesso em: abril 2007.

MOLLICA, C. *Influência da Fala na Alfabetização*. Rio de Janeiro: Tempo Brasileiro, 2000.

MORRISETTE, M. L. Recursive change in phonological acquisition. In: SLIFKA, J.; MANUEL, S.; MATTHIES, M. (Ed.) *Proceedings of From Sound to Sense: 50+ Years of Discoveries in Speech Communication*, Boston, p.C67-C72. 2004.

OLIVEIRA, M. A.; SOARES, M. B.; NICOLAU, E.; NASCIMENTO, M.; ALVARENGA, D. Da Forma Sonora da Fala à Forma Gráfica da Escrita. *Cadernos de Estudos Lingüísticos*. Campinas. v. 16, n. 2, p. 5-30, 1989.

PIERREHUMBERT, J. Phonetic diversity, statistical learning, and acquisition of phonology. *Language and speech*, v. 46, p. 115-154, 2003.

SOARES, Magda. 2003. A reinvenção da alfabetização. *Revista Presença Pedagógica*: 52. Disponível em: <http://www.editoradimensao.com.br/revistas/revista52_trecho.htm>. Acesso em: abril 2007.

SOARES, Magda. Letramento e alfabetização: as muitas facetas. *Revista Brasileira de Educação*. Número 25, 2004.

VIHMAN, M. M. *Phonological development: the origins of language in the child*. Cambridge: Blackwell publishers, 1996. 312 p.

YONEYAMA, K.; BECKMAN, M. E.; EDWARDS, J. Phoneme frequencies and acquisition of lingual stops in Japanese. *Language and Speech*, p. 1-56, in press.