

AVALIAR LEITURA E ESCRITA: CONTROLE DE FAMILIARIDADE ORTOGRÁFICA DAS PALAVRAS VIA LOGARITMO DA FREQUÊNCIA NO GOOGLE ADWORDS. *Fernando Cesar Capovilla, Luiz Eduardo Graton-Santos (**), Iago Lago Hamann (*), Caroline Ferreira da Silva (*), Kelly Casado (**), Miriam Damazio (**)* (Instituto de Psicologia, Universidade de São Paulo, São Paulo, SP)

Segundo a teoria de Duplo Processo de aquisição de leitura e escrita, a criança passa por três estágios: No estágio LogoGráfico ou Pré-AlfaBético, criança lê por reconhecimento grosseiro, sem ainda decifrar GraFEmas, e escreve como quem desenha, sem ainda cifrar FonEmas. Seus erros de leitura (ParaLexias) e escrita (ParaGrafias) consistem em trocas por semelhança visual grosseira. No estágio AlfaBético, criança usa rota PeriLexical-InfraVocabular para ler por decifragem GrafoFonÊMica e escrever por cifragem FonoGraFÊMica. Se Grau Médio de Decifrabilidade (GMD) GrafoFonÊMica das palavras escritas for baixo, cometem ParaLexias GrafoFonÊmicas na leitura em voz alta. E se Grau Médio de Cifrabilidade (GMC) FonoGraFÊMica das palavras ouvidas for baixo, cometem ParaGrafias FonoGraFÊmicas na escrita sob ditado. Assim, rota PeriLexical-InfraVocabular só funciona bem para leitura em voz alta se GMD da palavra escrita for alto; e para escrita sob ditado, só se GMC da palavra ouvida for alto. Em contraste, no estágio OrtoGráfico ou Pós-Alfabético, criança já usa rota Lexical-Vocabular e consegue ler palavras, mesmo com baixo GMD, e escrever, mesmo com baixo GMD, mas só se palavras tiverem elevada familiaridade ortográfica. Familiaridade ortográfica pode ser avaliada em bases de frequência de ocorrência de palavras na Internet, como Google AdWords. Um problema dessas bases é a enorme amplitude de variação de frequência mensal das palavras, que varia de 1 a 1 milhão. Este estudo testa a hipótese de que a conversão dos valores de frequência numa escala logarítmica de base 10 poderia produzir uma função apropriada para medir o efeito da familiaridade ortográfica sobre a frequência de palavras escritas corretamente sob ditado ouvido. Palavras com incidência 1 têm $\log_{10}=0$; 10: $\log_{10}=1$; 100: $\log_{10}=2$; 1.000: $\log_{10}=3$; 10.000: $\log_{10}=4$; 100.000: $\log_{10}=5$; 1.000.000: $\log_{10}=6$. Este estudo testou se, em universitários escrevendo sob ditado ouvido 563 palavras relativamente raras, a frequência de palavras escritas corretamente pode ser função positiva do $\log_{10}$ da frequência de ocorrência das palavras no Google AdWords. Participaram voluntariamente 62 estudantes de universidade pública paulista. Cada estudante recebia caderno de dez folhas de papel com pautas numeradas de 1 a 600. Estudantes eram avaliados coletivamente em sala de aula em 4 sessões de 30 min cada uma. O pesquisador dizia um número crescente 1-60 e, em seguida, articulava de modo claro e audível, por 3 vezes, a palavra que deveria ser escrita na linha contendo aquele número. Ao cabo de cada sessão, os cadernos eram recolhidos e os dados, analisados. Cada palavra escrita corretamente contava 1 ponto; palavras com qualquer erro de grafia, 0. Análise de covariância de familiaridade sobre acertos, tendo GMC como covariante, revelou que frequência de palavras acertadas cresceu significativamente com aumento na familiaridade das palavras pelo $\log_{10}$ da frequência. Média de acertos de palavras com frequência 1 foi 13,8; 10: 17; 100: 19,6; 1.000, 23,7; 10.000, 36,2; 100.000: 38,8; 1.000.000: 44,2. Conclusão: O $\log_{10}$ da frequência de ocorrência das palavras no Google AdWords é medida válida do grau de familiaridade ortográfica das palavras: prevê número de palavras escritas corretamente por universitários numa lista de 563 palavras para escrita sob ditado ouvido.

Palavras-chave: Escrita, Familiaridade, Google.

Capes, Inep, Observatório da Educação, CNPq

Nível: Pesquisador - P

Área: MET - Metodologia de Pesquisa e Instrumentação