

Sessão Coordenada 25 - NOVOS RECURSOS PARA IMPLEMENTAR O PARADIGMA MATRICIAL DE LINGUAGEM EM AVALIAÇÃO DE LEITURA-ESCRITA E SINALIZAÇÃO, EM PESQUISA, EDUCAÇÃO E CLÍNICA

DANDO ADEUS ÀS IDIOSSINCRASIAS NAS LISTAS DE PALAVRAS PARA ESCRITA SOB DITADO: COMO CALCULAR A CIFRABILIDADE DE QUALQUER PALAVRA FALADA DO PORTUGUÊS. *Fernando Cesar Capovilla, Kelly Casado (**), Miriam Damazio (**), Caroline Ferreira da Silva (*), Iago Lago Hamann (*), Luiz Eduardo Graton-Santos (**)* (Instituto de Psicologia, Universidade de São Paulo, São Paulo, SP), *Rodrigo Coelho (*)* (Processamento de Dados, Faculdade de Tecnologia de São Paulo, São Paulo, SP),

Psicologia Cognitiva da Leitura-Escrita e Alfabetização usa listas de palavras de tipo "regular", "regrada", "irregular" (em termos de relação letra-som-letra) para caracterizar estágio de desenvolvimento de leitura-escrita. No estágio AlfaBético, usando rota PeriLexical-InfraVocabular, crianças decifram e cifram palavras regulares, mas regularizam as irregulares e, assim, não as entendem (porque cometem ParaLexias GrafoFonêmicas), nem se fazem entender por elas (porque cometem ParaGrafias FonoGrafêmicas). Mas no estágio OrtoGráfico, já usam rota Lexical-Vocabular e conseguem fazer produção ortográfica e reconhecimento visual direto de palavras irregulares. Contudo, listas agrupam arbitrariamente sob mesmo tipo itens contendo diferentes relações em diferentes números: palavras classificadas como "regradas" contêm relações regulares; palavras classificadas como "irregulares" contêm relações regulares e regradas. Violações espúrias em graus variados nos critérios de classificação produzem listas com baixa precisão, validade, confiabilidade, generalidade. Como cada equipe de pesquisa tem sua própria lista arbitrária, não há articulação entre dados de equipes distintas para edificar ciência sólida com progresso sistêmico. Que variáveis contínuas subjazem às variações espúrias? Como controlá-las? Como medir grau de (ir)regularidade de qualquer palavra? Única saída é substituir noção de "palavra do tipo irregular para escrita" pelo conceito de grau médio de cifrabilidade FonoGrafêmica da palavra falada (GMC). Teoricamente, qualquer palavra falada teria determinado GMC que, se descoberto, deveria predizer dificuldade de escrita de palavras de baixa frequência ortográfica. GMC consiste na média aritmética dos índices ponderados de cifrabilidade (IPC) dos FonEmas componentes dessa palavra. Para calcular esses IPC criamos aplicativo A Voz Brasileira na Nova Ortografia. Escrevemos 60.803 palavras, e convertemos cada GrafEma no correspondente FonEma via caracteres IPA para representar variações de pronúncia regional. Obtendo 353.250 formas de pronunciar 60.803 palavras escritas, computamos incidências das relações bidirecionais entre seus 154 GrafEmas e 97 FonEmas, obtendo 236 relações GrafoFonêmicas na leitura em voz alta do Português escrito, e 238 relações FonoGrafêmicas na escrita sob ditado do Português falado. Mapeando modos de cifrar (grafar) cada FonEma, computamos incidência proporcional de cada um. Assim, descobrimos o IPC de cada relação FonoGrafêmica entre cada FonEma e os GrafEmas que o cifram. Exemplo: Cifra-se o FonEma [a] com 10 GrafEmas: [a]-[s] em 63,084% dos casos (ex: "sala"); [a]-[c]: 20,747% ("alface"); [a]-[ç]: 7,619% ("praça"); [a]-[ss]: 5,057% ("assar"); [a]-[x]: 1,566% ("expelir"); [a]-[sc]: 1,450% ("crescer"); [a]-[z]: 0,366% ("paz"); [a]-[xc]: 0,102% ("exceto"); [a]-[xs]: 0,005% ("exsudato"); [a]-[sç]: 0,004% ("cresça"). Fazendo isso com todos os FonEmas, obtivemos Tabelas de índices ponderados de cifrabilidade FonoGrafêmica (IPC). Descobrimos que, na cifragem de palavras incomuns, probabilidade de ParaGrafias FonoGrafêmicas é inversamente proporcional ao IPC das relações



FonoGrafÊmicas. ParaGrafias costumam acometer relações com IPC baixo (ex: [sç], [xs], [xc], [z], [sc], [x]). Consistem na substituição destas por relações com IPC alto (ex: [s], [c], [ç], [ss]) na grafia de palavras incomuns (conforme Google AdWords, ex: “acresção”, “imissão”, “exsudato”, “excisão”, “aboiz”, “aletriz”, “arriaz”, “excarcerar”, “sêxtuor), em especial por crianças com baixo conhecimento OrtoGráfico-MorfÊMico. Estudo com 162 escolares de EF6-EF9, que escreveram sob ditado 566 palavras incomuns contendo 177 relações FonoGrafÊmicas, corrobora modelo IPC-GMC para probabilidade diferencial de ParaGrafias FonoGrafÊmicas. Idiossincrasias nessa classificação de (ir)regularidade, nunca mais.

Escrita, Paragrafia, Software.

Capes, Inep, Observatório da Educação, CNPq

Pesquisador - P

MET - Metodologia de Pesquisa e Instrumentação

Matriz de para interpretação do teste de leitura de palavras e pseudopalavras. *Luiz Eduardo Graton-Santos***, *Fernando Cesar Capovilla* (Instituto de Psicologia, Universidade de São Paulo, São Paulo, SP),

A alfabetização é um importante marco do início da escolarização fundamental para a continuidade dos estudos a níveis mais avançados. Avaliar a leitura é uma estratégia que permite planejar seu ensino, assim como para corrigir eventuais atrasos e dificuldades. Normatizado para ouvintes a partir de 4 anos de idade e para surdos a partir de 7 anos, sensível à escolaridade (i.e., capaz de diferenciar sucessivamente os anos de escolares iniciais) e validado por comparação com diversos instrumentos que avaliam habilidades linguísticas e não linguísticas, o Teste de Competência de Leitura de Palavras e Pseudopalavras (TCLPP) permite identificar o estágio de desenvolvimento de leitura do examinando e o seu grau de proficiência em cada um dos três tipos de processamento de leitura: 1) ideovisual ou logográfico; 2) perilexical ou fonológico por decodificação grafofonêmica, e 3) lexical ou por reconhecimento visual direto assistido por decodificação grafofonêmica eventual. O TCLPP é ao mesmo tempo um instrumento psicométrico e neuropsicológico, pois possui tabelas de normatização para avaliar o grau de desvio entre o padrão de leitura de um avaliando e de seu grupo de referência, enquanto também permite interpretar os dados do padrão de leitura apresentado de acordo com o modelo de desenvolvimento de leitura e escrita e fornece uma visão integrada acerca do grau de desenvolvimento e preservação dos diferentes mecanismos, rotas e estratégias envolvidos na leitura competente, o que permite lançar luz sobre a natureza da dificuldade específica de um determinado examinando. Composto por sete subtestes que exigem diferentes estratégias de leitura, o padrão de distribuição dos tipos de erros no TCLPP é capaz de revelar a natureza específica do processamento cognitivo do examinando, além de indicar as estratégias de leitura que ele consegue usar e aquelas com que tem dificuldade. O TCLPP apresenta quais estratégias de leitura (i.e., logográfica, alfabética e lexical) podem ser usadas para resolver cada subteste e qual a natureza da dificuldade que o erro em cada subteste pode revelar. Contudo, o instrumento deixa a critério do examinador a classificação do estágio de desenvolvimento da leitura obtido através da avaliação. O presente trabalho apresenta uma matriz para interpretação das pontuações nos subtestes do TCLPP, que indica limites de pontuação para cada estágio de desenvolvimento de leitura para orientar a classificação do avaliando conforme seu desempenho e a consequente recomendação de intervenção para melhoria de eventuais dificuldades encontradas. O trabalho indica como interpretar as pontuações a partir da implementação da matriz com resultados obtidos em trabalhos de normatização do TCLPP.

Leitura, avaliação, alfabetização.

Capes, CNPq, Inep, Observatório da Educação

Doutorado - D

MET - Metodologia de Pesquisa e Instrumentação

AVALIANDO LEITURA EM VOZ ALTA EM ALFABETIZAÇÃO E DISLEXIA: COMO CALCULAR DECIFRABILIDADE DE QUALQUER PALAVRA ESCRITA PARA DAR ADEUS ÀS ARBITRARIEDADES NAS LISTAS DE PALAVRAS. *Caroline Ferreira da Silva (*), Iago Lago Hamann (*), Kelly Casado (**), Miriam Damazio (**), Luiz Eduardo Graton-Santos (**), Fernando Cesar Capovilla (Instituto de Psicologia, Universidade de São Paulo, São Paulo, SP), Rodrigo Coelho (*) (Processamento de Dados, Faculdade de Tecnologia de São Paulo, São Paulo, SP)*

Para caracterizar o estágio de desenvolvimento de leitura de determinada criança (Pré-Alfabético, AlfaBético, Pós-Alfabético), a Psicologia Cognitiva de Processamento de Informação emprega listas de palavras de diferentes “tipos” para leitura em voz alta. Palavras classificadas como: “regulares” contêm apenas relações regulares: um FonEma por GrafEma. Palavras “regradadas por posição” incluem relações GrafoFonêmicas em que a posição do GrafEma determina o FonEma (ex: pronuncia-se “s” intervocálico como [s̺] em “casa”). Palavras “irregulares” incluem relações GrafoFonêmicas em que FonEma varia independente da posição do GrafEma, como na pronúncia de palavras com “x”. No estágio AlfaBético, usando rota PeriLexical-InfraVocabular, crianças conseguem decifrar palavras regulares, mas regularizam irregulares, cometendo ParaLexias GrafoFonêmicas que tornam sua pronúncia incompreensível para elas e os demais. Mas no estágio Pós-Alfabético, já usam rota Lexical-Vocabular e conseguem ler irregulares, desde que familiares, por reconhecimento visual direto. Contudo, listas têm baixa precisão, validade, confiabilidade, generalidade, porque classificam arbitrariamente, sob mesmo tipo, itens contendo diferentes relações em diferentes números: palavras classificadas como “regulares” contêm relações regulares; palavras classificadas como “irregulares” contêm relações regulares e regradadas. Cada equipe de pesquisa tem sua própria lista idiossincrática com violações espúrias em graus variados nos critérios de classificação. Inexiste articulação entre dados de equipes distintas para edificar ciência sólida com progresso sistêmico. Que variáveis contínuas subjazem às variações espúrias? Como controlá-las? Como medir grau de (ir)regularidade de qualquer palavra escrita? Solução: substituir noção de “palavra do tipo irregular para leitura” pelo conceito de Grau Médio de Decifrabilidade GrafoFonêmica da palavra escrita (GMD), como média aritmética dos Índices Ponderados de Decifrabilidade (IPD) dos GrafEmas componentes. Teoricamente, qualquer palavra escrita teria determinado GMD que, se descoberto, deveria prever dificuldade de leitura de palavras de baixa frequência ortográfica, conforme Google AdWords. Para obter tabelas de IPD, Capovilla e Casado criaram aplicativo A Voz Brasileira na Nova Ortografia. Escrevendo 60.803 palavras, e convertendo cada GrafEma no correspondente FonEma via caractere de International Phonetic Alphabet, obtiveram 353.250 variações de pronúncia regional daquelas palavras. Mapearam relações entre 154 GrafEmas e 97 FonEmas, obtendo 236 relações GrafoFonêmicas na leitura em voz alta do Português escrito. Computando incidência proporcional de cada modo de decifrar (pronunciar) cada GrafEma, descobriram o IPD de cada relação GrafoFonêmica. Exemplo: Decifra-se GrafEma “x” com seis FonEmas: “x”-[x̺]: 40,19% (“bacumixá”); “x”-[x̠]: em 25,03% dos casos (ex: “expelir”, “exceção”); “x”-[x̠; x̺]: 21,66% (“actinotoxemia”); “x”-[x̠; x̠]: 12,65% (“exalviçado”); “x”-[x̠; x̺; x̠]: 0,41% (“hexágono”, “oxácido”); “x”-[x̠] (áfono): 0,05% (“bordeaux”, “vitreaux”). Fazendo isso com todos os GrafEmas, obtiveram Tabelas de índices ponderados de decifrabilidade GrafoFonêmica (IPD). Segundo esse modelo IPD de ParaLexias GrafoFonêmicas, na leitura em voz alta de palavras incomuns, a probabilidade de ParaLexias GrafoFonêmicas é inversamente proporcional ao IPD das relações GrafoFonêmicas. ParaLexias acometem relações com IPD baixo (ex: “x”-[x̠], “x”-[x̠; x̠], “x”-



[x], “x”-[x];, “x”-[x]). Consistem na substituição destas por relações com IPD alto (ex: “x”-[x]) na pronúncia de palavras incomuns, especialmente por crianças com baixo conhecimento OrtoGráfico-MorfÊmico. Este modelo IPD-GMD para probabilidade diferencial de ParaLexias GrafoFonÊmicas elimina arbitrariedades na classificação de (ir)regularidade em listas de palavras de tipos “regular”x“regrada”x“irregular” para leitura em voz alta.

Leitura, Paralexia, Software

Capes, Inep, Observatório da Educação, CNPq

Iniciação Científica - IC (trabalho de graduação)

MET - Metodologia de Pesquisa e Instrumentação

ANÁLISE COMPUTADORIZADA DA ESTRUTURA SUBLEXICAL DE 10.400 SINAIS DE LIBRAS: CARACTERIZAÇÃO DAS COMBINAÇÕES CANÔNICAS ENTRE FORMA DE MÃO, ORIENTAÇÕES DE MÃO E PALMA, MOVIMENTO, E EXPRESSÃO FACIAL. Iago Lago Hamann (*), Caroline Ferreira da Silva (*), Karina Nonato Pingituro Domingues (**), Wanessa Oliveira Garcia (**), Antonielle Cantarelli Martins (**), Fernando César Capovilla (Instituto de Psicologia, Universidade de São Paulo, São Paulo, SP), Marcelo Duduchi (Processamento de Dados, Faculdade de Tecnologia, São Paulo, SP)

Assim como o Português tem sua FonoLogia e OrtoGrafia, a Língua de Sinais Brasileira (Libras) tem sua SematosEmia. Este trabalho analisou a estrutura SematosÊMica-Signumicular da Libras nos cruzamentos entre parâmetros: Articulação de Mão(s); Orientação de Palma(s); Orientação de Mão(s); Relação entre Mãos; Movimento de Mão(s); Expressão Facial Positiva e Negativa. Objetivando caracterizar Libras em busca das combinações mais características entre SematosEmas de Libras, o software BuscaSigno-2 foi usado para analisar o corpus de 10.400 sinais da 2a. edição revista e ampliada do Novo Dicionário enciclopédico ilustrado trilingue de Libras. BuscaSigno-2 computou as incidências absoluta e relativa de sinais contidos nos 13 seguintes cruzamentos entre estes SematosEmas: 1: articulação da mão direita e articulação da mão esquerda (AMD x AME); 2: articulação da mão direita e movimento da mão direita (AMD x MMD); 3: articulação da mão direita e orientação da palma direita (AMD x OPD); 4: articulação da mão esquerda e movimento da mão esquerda (AME x MME); 5: articulação da mão esquerda e orientação da palma esquerda (AME x OPE); 6: articulação da mão direita, orientação da mão direita, orientação da palma direita (AMD x OMD x OPD); 7: articulação da mão esquerda, orientação da mão esquerda, orientação da palma esquerda (AME x OME x OPE); 8: articulação da mão direita, orientação da palma direita, e movimento da mão direita (AMD x OPD x MMD); 9: articulação da mão direita, articulação da mão esquerda, relação entre mãos (AMD x AME x RM); 10: movimento da mão direita, sentimentos sensações positivas (MMD x SSP); 11: movimento da mão direita, sentimentos sensações negativas (MMD x SSN); 12: articulação da mão direita, movimento da mão direita, sentimentos sensações positivas (AMD x MMD x SSP); 13: articulação da mão direita, movimento da mão direita, sentimentos sensações negativas (AMD x MMD x SSN). Para caracterizar as combinações mais típicas entre os SematosEmas de Libras, foram analisados 122.163 cruzamentos em busca dos sinais em cada cruzamento. Foram computados os sinais em cada um desses 122.163 cruzamentos. Dos 122.163 cruzamentos analisados, foram identificados 21.328 cruzamentos válidos (i.e., com incidência superior a zero). Destes, foram identificados 369 cruzamentos canônicos (i.e., com incidência maior que 50% do total naquele cruzamento). Esses 369 cruzamentos canônicos foram, então, ordenados por incidência decrescente, de modo a revelar quais combinações entre SematosEmas que são mais típicas em Libras. Essa caracterização do "DNA" de Libras permite: 1: Comparar Libras com outras línguas de sinais em estudos de Linguística Comparada; 2: Fazer balanceamento SematosÊMico relevante à elaboração de qualquer teste de competência linguística e metalinguística em surdos; 3: Estudar processamento cognitivo de informação em sinais por surdos descobrindo como armazenam, processam e recuperam informação de SematosEmas. Por exemplo, usando BuscaSigno-1, Duduchi e Capovilla, descobriram que o grau de dificuldade de recuperação lexical de um sinal é inversamente proporcional ao logaritmo da média aritmética da popularidade dos SematosEmas que o compõem. Assim, este mapeamento é de grande relevância à Psicologia Cognitiva de Processamento de Informação em Sinais em Surdos brasileiros.

Libras, Língua de Sinal, Software



CNPq, Capes, Inep, Observatório da Educação, Fapesp.
Iniciação Científica - IC (trabalho de graduação)
MET - Metodologia de Pesquisa e Instrumentação

ANÁLISE DA ESTRUTURA SEMATOSÊMICA-SIGNUMICULAR DO CORPUS DE 10.400 SINAIS DA 2ª ED. DO NOVO DEIT LIBRAS VIA BUSCASIGNO 2. *Wanessa Garcia Santos Oliveira (Pós-Graduação em Psicologia Clínica e Institucional, Centro Universitário Adventista de São Paulo, São Paulo, SP), Fernando César Capovilla (Instituto de Psicologia, Universidade de São Paulo, São Paulo, SP)*

Assim como palavras ouvidas são compostas de FonEmas-VocÍculos e palavras escritas de GrafEmas-ScriptumÍculos, os sinais das Línguas de Sinais são compostos de SematosEmas-SignumÍculos, como unidades mínimas. Há SematosEmas para formada(s) mão(s): QuiriFormEmas-ManusModusÍculos; para local da(s) mão(s): QuiriToposEmas-ManusLocusÍculos; para movimento da(s) mãos: QuiriCinesEmas-ManusMotusÍculos; e para expressão facial: MascarEmas-PersonalÍculos. O software de recuperação lexical de sinais BuscaSigno 2 indexa os sinais a partir de seus SematosEmas componentes. E então organiza esses SematosEmas em menus gráficos selecionáveis na tela para permitir a busca e recuperação lexical dos sinais que os contêm. BuscaSigno 2 contém banco de 10.400 sinais, que corresponde ao corpus da 2ª edição do Novo Deit Libras: Novo Dicionário enciclopédico ilustrado trilingue da Língua de Sinais Brasileira. - Libras, de Capovilla e Raphael. Este estudo empregou BuscaSigno 2 para mapear a estrutura sublexical (SematosÊMica-SignumIcular) da Língua de Sinais Brasileira (Libras) a partir do corpus de 10.400 sinais da 2ª edição do Novo Deit Libras. Assim, BuscaSigno2 permite ao usuário obter acesso lexical direto a qualquer um de 10.400 sinais de Libras e resgatar esses sinais sem precisar saber os verbetes em Português que correspondem aos sinais que procura. Para fazer essa busca SematosÊMica e resgate lexical de sinais, basta que o usuário selecione com o mouse os menus de representações gráficas de SematosEmas. BuscaSigno 2 oferece ao consultante do Dicionário de Libras uma interface de busca direta dos sinais de sua língua de sinais, que respeita a estrutura linguística dos sinais de Libras, em vez de depender dos verbetes do Português ordenados alfabeticamente. Ou seja, BuscaSigno 2 permite aos surdos recuperar diretamente os sinais de sua língua sem ter de depender de uma outra língua (Português escrito) como metalinguagem para acessar sua língua materna. O mapeamento da incidência de 501 SematosEmas no corpus de 10.400 sinais revelou 5 parâmetros: 1: 163 SematosEmas de Mão (QuiriFormEmas-ManusÍculos), 2: 34 SematosEmas de Dedo (DactilEmas-DigitumÍculos), 3: 55 SematosEmas de Local da Articulação (ArtrotoToposEmas-ArticulatiLocusÍculos), 4: 173 SematosEmas de Movimento (CinesEmas-MotusÍculos), 5: 76 SematosEmas de Expressão Facial (MascarEmas-PersonalÍculos). Esses 5 parâmetros se subdividem em dezenas de subníveis que se desdobram em 501 SematosEmas. Computando a frequência de cada um desses SematosEmas, BuscaSigno 2 identificou os canônicos (acima da média), que consistem no DNA da Libras. Tal mapeamento computadorizado da incidência absoluta e proporcional de cada SematosEma componente do léxico de Libras é crucial à Linguística Comparativa das Línguas Sinalizadas, e a uma análise experimental psicolinguística capaz de elucidar como surdos armazenam, processam, e recuperam informação a partir das unidades mínimas de que se compõem os sinais. Ilustrando a importância destes recursos para a Psicolinguística dos Surdos, Duduchi e Capovilla descobriram que o grau de dificuldade de recuperação lexical de sinais é inversamente proporcional ao logaritmo da média aritmética dos coeficientes de popularidade dos SematosEmas que os compõem. Quanto mais raros os SematosEmas componentes dos sinais, maior é a eficácia desses SematosEmas como recurso para recuperar esses sinais. Tais dados coadunam-se com as expectativas derivadas da Teoria do Pandemônio, na Psicologia Cognitiva de Processamento de Informação.

Língua de Sinais, Libras, Acesso lexical informatizado



Capes, Inep, Observatório da Educação, CNPq.
Mestrado - M
MET - Metodologia de Pesquisa e Instrumentação